# **Supplementary Online Content**

Karn T, Jiang T, Hatzis C, et al. Association between genomic metrics and immune infiltration in triple-negative breast cancer. *JAMA Oncol.* Published online July 27, 2017. doi:10.1001/jamaoncol.2017.2140

eMethods. Supplementary Methods

eFigure 1. Strategy of RNA-Seq and Whole-Exome-Seq Analyses for TNBC Classification

eFigure 2. ER, PR, and HER2 Expression Assessed by RNA-Seq and Agilent Arrays

eFigure 3. Dependency of Platform Correlation on Gene Expression Level

eFigure 4. Correlation Between RNA-Seq and Affymetrix for Metagene Clusters

eFigure 5. Classification of TNBC (n = 208) Based on RNA-Seq Data

**eFigure 6.** Correlation of MHC2 Metagene Expression and Histological Quantification of TILs in TCGA Samples

eFigure 7. Classification Algorithm of the Prognostic Immune Signature

**eFigure 8.** Validation of Improved Prognosis of TNBC Patients With "Good Prognosis" Signature in RNA-Seq Data

eFigure 9. Mutational Count Distribution in 186 TNBC

eFigure 10. Relationship Between SCNA Levels and MATH in TNBC From TCGA

**eFigure 11.** Validation of Inverse Relationships Between Measures of Genomic Complexity and Immune Cell Infiltration in TNBC Using Different Immune Metagenes

eFigure 12. Prognostic Value of Histologically Quantified TILs in the TCGA TNBC Data Set

eFigure 13. Validation of Inverse Relationship of Genomic Heterogeneity and Immune Cell

Infiltration Using Histologically Quantified TILs in the TCGA TNBC Data Set

**eFigure 14.** Differences in Mutation Count by Immune Cell Infiltration Metagenes and IL8/VEGF Metagene Expression Categories

eFigure 15. Correlation of the Number of Predicted Neoantigens and Mutational Load

eFigure 16. Independence of MATH Score and Total Mutation Counts

eFigure 17. Validation Analyses in METABRIC Data Set

**eFigure 18.** Differences in Mutation Count, Neoantigen Count, and CYT by Molecular Subtype in Breast Cancer

eFigure 19. Confounding of Molecular Breast Cancer Subtypes on Predicted Neoantigen Count and CYT

eFigure 20. High Intercorrelation of TIL Metagenes

eFigure 21. Association Between Clonal Heterogeneity and Immune Metagene Expression

eFigure 22. Association of MATH and SCNA With Prognostic Groups in TNBC

**eTable 1.** Annotated Cancer Genes Mutated in  $\geq$ 3 Samples

eTable 2. TCGA Samples Included in the Study

eTable 3. "Cancer Genes" Curated by Vogelstein and colleagues

**eTable 4.** Individual Genes Constituting TNBC Metagenes and Their Correlation With Affymetrix Microarray

This supplementary material has been provided by the authors to give readers additional information about their work.

# eMETHODS

# Contents

eMETHODS 1
Data sources
Transfer of immune and stromal prognostic metagenes from the Affymetrix to RNA-seq data 3
Selection of a gene expression based TNBC cohort from TCGA
Gene filtering in RNA-Seq data
Metagene construction
Transfer of prognostic signature
Analysis of somatic mutations
Overall mutation counts
Differences in mutational load according to prognostic signature 4
Mutation frequencies of known cancer genes in prognostic groups5
Analysis of tumor genomic heterogeneity
MATH scores and SCNA levels in TNBC subgroups by prognostic immune signature
Deconvolution of the effects of immune cell infiltration and tumor cellularity on genomic heterogeneity of TNBC
Validation by histological quantification of immune infiltration
Validation in independent METABRIC dataset7
Relationship to other recent studies7
References
eFIGURES11
eTABLES
R-MarkDown document

All analyses were performed according to the "*REporting recommendations for tumour MARKer prognostic studies*" (REMARK) <sup>16,17</sup>. A diagram of the complete analytical strategy and the flow of patients through the study, including the number of patients included in each stage of the analysis, is given in **eFigure 1**. The R software environment (<u>http://www.r-project.org/</u>) using RStudio (<u>www.rstudio.com</u>) and IBM SPSS version 22.0 (<u>http://www.ibm.com/</u>) were used for all analyses. Chi square test was applied to assess associations between categorical parameters. All reported P values are two sided and P < 0.05 was considered significant. Detailed R-code and results of analyses are given in the accompanying R-MarkDown document (*R-MarkDown-document.pdf*). Code and data are also available at <u>https://github.com/tkarn/TNBC-TIL</u>.

#### **Data sources**

- Level 3 RNA-Seq V2 data for 20530 genes from 1215 treatment naïve BRCA samples processed 2015-01-28 were downloaded from TCGA (https://tcgaon data.nci.nih.gov/tcgafiles/ftp\_auth/distro\_ftpusers/anonymous/tumor/brca/cgcc/unc.edu/illuminahi seq\_maseqv2/rnaseqv2/) and  $\log_2(x+1)$  transformed (filename <  $TCGA\_BRCA\_exp\_HiSeqV2$ -2015-02-24.tgz >). A current version (2016-08-16) of these data is available from UCSC Xena browser (http://xena.ucsc.edu/) TCGA hub (https://xenabrowser.net/datapages/?dataset=TCGA.BRCA.sampleMap/HiSeqV2&host=https://tcg and a.xenahubs.net) can be dowloaded at (https://tcga.xenahubs.net/download/TCGA.BRCA.sampleMap/HiSeqV2; filename  $\langle HiSeqV2 \rangle$ ) <sup>18,19</sup>. **eTable 2** catalog the TCGA samples included in the study.
- Agilent 244K custom gene expression G4502A\_07\_3 microarrays by the University of North Carolina TCGA genomic characterization center for 597 samples were obtained from UCSC cancer genome browser ((<u>https://genome-cancer.ucsc.edu/proj/site/hgHeatmap</u>; filename < *TCGA\_BRCA\_G4502A\_07\_3-2015-02-24.tgz* >). A current copy is available at UCSC Xena (<u>https://xenabrowser.net/datapages/?dataset=TCGA.BRCA.sampleMap/AgilentG4502A\_07\_3&ho st=https://tcga.xenahubs.net</u>).
- Gene-level mutation data (nonsilent somatic mutation; wustl curated) for 982 samples were obtained from UCSC cancer genome browser processed on 2015-01-27 ((<u>https://genome-cancer.ucsc.edu/proj/site/hgHeatmap</u>; filename < *TCGA\_BRCA\_mutation\_curated\_wustl\_gene-2015-02-24.tgz* >) and can be downloaded from <u>https://github.com/tkarn/TNBC-TIL</u>. A current copy is available at UCSC Xena (<u>https://xenabrowser.net/datapages/?dataset=TCGA.BRCA.sampleMap/mutation\_curated\_wustl\_gene&host=https://tcga.xenahubs.net</u>).
- Updated clinical data and sample information for TCGA samples were obtained from cBIO portal (www.cbioportal.org) using the R library *cgdsr*<sup>20</sup>.
- Mutant variant allele frequencies (vaf) of all genes for MATH score calculation were obtained from file < genome.wustl.edu\_BRCA.IlluminaGA\_DNASeq.Level\_2.1.1.0.curated.somatic.maf.txt > of 2014-Feb-08 from the TCGA Portal Data (https://tcgadata.nci.nih.gov/tcgafiles/ftp\_auth/distro\_ftpusers/anonymous/). A copy of the file is available from GDC (https://portal.gdc.cancer.gov/legacy-archive/files/50d6fb1d-5bb1-4a30-9e91-6d45bd9b1c3f). The vaf data are also available at https://github.com/tkarn/TNBC-TIL.
- The numbers of predicted neo-epitopes based on tumor-specific HLA typing for each patient <sup>21</sup> were obtained for 760 BRCA samples from Supplementary Table S4 from Rooney et al. <sup>5</sup> available at <u>http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4856474/bin/NIHMS717941-</u> <u>supplement-Table S4.xlsx</u>
- Somatic copy number alteration (SCNA) levels were obtained for 941 BRCA samples from Supplementary Table S7 from Davoli et al.<sup>8</sup> available at <u>http://science.sciencemag.org/highwire/filestream/689461/field\_highwire\_adjunct\_files/7/aaf8399</u> -Davoli-SM-table-S7.xlsx
- Histological quantification of mononuclear cells from sections of 180 BRCA samples from TCGA were obtained from Supplementary Table S1 from Lehmann et al. <sup>22</sup> available at

http://journals.plos.org/plosone/article/file?type=supplementary&id=info:doi/10.1371/journal.pon e.0157368.s006

- Spearman correlation values between RNA-Seq and Affymetrix microarray for 16,097 Jetset probes for 57 paired frozen breast cancer samples was obtained from Supplementary Table S2 of Fumagalli et al.<sup>23</sup> and are available at <a href="https://github.com/tkarn/TNBC-TIL">https://github.com/tkarn/TNBC-TIL</a>
- Mutational profiles and patterns of immune infiltration of the METABRIC dataset from Pereira et al.<sup>24</sup> and Ali et al.<sup>25</sup> were obtained from <a href="http://github.com/cclab-brca">http://github.com/cclab-brca</a>
- R-code and respective data files are available at <a href="https://github.com/tkarn/TNBC-TIL">https://github.com/tkarn/TNBC-TIL</a>

# Transfer of immune and stromal prognostic metagenes from the Affymetrix to RNA-seq data

We transferred triple negative breast cancer (TNBC) prognostic and immune gene expression classification based on DNA microarray data to RNA-Seq data from the TCGA using the following steps (eFigure 1):

#### Selection of a gene expression based TNBC cohort from TCGA

We first compared RNA-Seq and Agilent array expression results for estrogen (ESR1) and progesterone (PR) receptors and HER2 mRNA to ensure that we can use RNA-Seq to identify TNBC from TCGA (eFigure 2 A-C). We identified n=208 TNBC from the TCGA based on bimodal distribution of RNA-Seq for ESR1 and HER2 (eFigure 2 D-F), full details in Section-1 in *R*-*MarkDown-document.pdf*.

#### Gene filtering in RNA-Seq data

To reduce noise from the RNA-Seq data we applied a gene filtering step. A recent report <sup>23</sup> provides Spearman correlation coefficients for expression of 16097 genes measured both on Affymetrix microarray and RNA-Seq platform in a series of 57 breast cancers. We analyzed these correlation scores based on median expression of the genes among our sample set of 208 TNBC from TCGA and found a poor correlation for genes with median expression in the lower quartile (<2 in log<sub>2</sub> expression space) (**eFigure 3**), (see Section-2 in *R-MarkDown-document.pdf*). Based on that criterium we selected 15876 of 20530 genes from RNA-Seq data with log<sub>2</sub> expression values above 2 in at least 50% of the 208 samples.

#### Metagene construction

We performed unsupervised hierarchical clustering of the 15876 genes based on RNA-Seq data from the 208 TNBC samples (with single linkage and Pearson correlation as distance metric) using GenePattern <sup>26</sup> (http://genepattern.broadinstitute.org) and confirmed co-expression clusters of 15 metagenes including a total of 304 distinct genes (eTable 4). Median correlation between RNA-Seq and Affymetrix for the 304 genes was 0.88 (eFigure 4) (Section-3.1 in *R-MarkDown-document.pdf*). These co-expression clusters correspond to the previously reported Affymetrix expression data based TNBC phenotypes and immune and stromal metagenes, including Basal-like, Molecular-Apocrine, Claudin, Proliferation, B-cell, T-cell, MHC1, MHC2, IFN, Collagen/Stroma, Endothel, Histone, HOXA metagenes <sup>9</sup>. The expression of inflammation (IL8, CXCL1) and angiogenesis markers (VEGFA, adrenomeddulin, ANGPTL4 a.o.), were highly correlated in this dataset and therefore were combined into a single metagene IL8-VEGF. Metagene values were calculated as the average expression of all member genes (Section-3.2 in *R-MarkDown-document.pdf*). The metagenes allowed to reproduce the previously described classification of TNBC into several subtypes (eFigure 5) <sup>9</sup>. As an alternative method to calculate immune cell activity based on gene expression, we also used the

CYT gene expression score score (geometric mean of GZMA and PRF1 mRNA expression) as described by Rooney and colleagues<sup>5</sup>.

#### **Transfer of prognostic signature**

Different immune metagenes (e.g. T-Cell, B-Cell, MHC2, CYT) are highly correlated in breast cancer gene expression data<sup>3,5,14,27</sup>, and because of that may all be similarily applicable as surrogates for a  $T_{H1}$ type immune response<sup>4</sup>. In contrast, inflammation markers <sup>4,9,10</sup> and myeloid cell markers<sup>28</sup> may capture the signal of tumor promoting inflammation and immunosuppressive factors. We have previously shown that extensive immune infiltration reflected by high expression of either Bcell, or MHC2, or T-cell metagenes together with low expression of inflammatory metagenes (IL8, and VEGF) define a group of good prognosis TNBC 9,10,29. We tested if these metagenes retained prognostic function in the RNA-Seq data from TCGA. Since distributions of the immune metagenes did not show bimodality which would have allowed clear cutoff selection<sup>30</sup> we stayed with previously used cutoffs values to avoid overfitting of the data (Section-5 in R-MarkDown-document.pdf). We used two approaches to transfer the signatures from microarray to RNA-Seq. First, we considered the highest quartile of MHC2 metagene expression to represent high immune infiltration and high MHC2 expression<sup>27</sup> in combination with low (below the median) IL8/VEGF metagene expression' to define the good prognosis category (eFigure 7 for signature definition), full details in Section-5.1 in R-MarkDown-document.pdf. The upper quartile cutoff for MHC2 metagene has been used in one of our previous papers based on earlier histological TIL data<sup>27</sup>. A robust median split of the IL8/VEGF metagene has also been applied before<sup>9</sup>. This method assigned 25 cases to good prognosis and these patients had significantly better survival compared to the poor prognosis group (P=0.019, eFigure 8B). As a second strategy, we used B-cell metagene expression above the lowest quartile and IL8-VEGF metagene expression below the median to categorize a case as good prognosis (in this case we applied the original cutoffs from our previous description of this signature in the microarray dataset<sup>9</sup>) (eFigure 8C) (Section-5.2 in *R-MarkDown-document.pdf*). We only observed a trend for increased survival in the respective group of 76 samples in the TCGA in Kaplan-Meier analysis (P=0.22 log rank test, eFigure 8C). Many studies have demonstrated that different immune metagenes are highly correlated as shown for the MHC2 and T-cell metagenes in eFigure 20 contributing redundant information ( $R^2=0.807$ ). Since the prognostic value of these immune metagenes has been repeatedly and independently shown in many datasets, the modest effect of our second strategy involving the B-cell metagene in the TCGA data may be related to the limited power of the TCGA data set due to short median follow up of 24 months and only 29 events in 193 cases <sup>31</sup>. In our previous microarray study 139 events occurred in 402 TNBC with a median follow up of 60 months. To assure robustness in our analysis, we used both prognostic classifications methods when we performed correlation with genomic metrics and report results for both.

#### Analysis of somatic mutations

#### **Overall mutation counts**

Matching whole exome sequencing (WES) data was available for 186 of the 208 TNBC samples. For these we obtained curated somatic mutation data from UCSC cancer browser (see above section *Data Sources* for details). The median number of non-silent mutations was 53 per tumor (range 1-1138) (**eFigure 9**). The most frequently mutated gene was TP53 (74.7%). Six samples displayed a hypermutated phenotype with more than 300 mutated genes. Inclusion or exclusion of these six cases did not change the results, indicating that our results are not driven by these rare outliers with high mutation rate.

#### Differences in mutational load according to prognostic signature

When we compared total mutation counts at gene level (in 39741 curated genes/pseudogenes from WES), good prognosis TNBC had significantly lower mutation count for the two signatures using MHC2 (P=0.021, **Figure 2D**) and B-cell (P=0.014), respectively. This difference was primarily driven

by the lower mutation count in samples with high immune infiltration (i.e high MHC2 (P=0.003) or Bcell (P=0.018) metagene expression (eFigure 14). There was no significant difference in mutation counts by IL8-VEGF metagene expression levels or by TNBC molecular subtype (i.e. apocrine, basallike or claudin-low, P=0.7, not shown). We performed the same analyses using the predicted neoantigen load obtained from <sup>5</sup> and obtained similar, significant inverse association between prognostic category and neo-antigen load (Figure 2D). We noted that mutation count and neoantigen load are highly correlated (eFigure 15). However, since the number of predicted neoantigens is nearly a magnitude smaller than mutation counts, the inverse association associations with immune cell metagens were no longer significant for neo-antigen load. This is likely due to the reduced power. eFigure 18 presents mutation count, neoantigen count, and CYT metagene expression in different breast cancer molecular subtypes. The results illustrate that the previously reported association of higher neo-antigen load and higher CYT metagene expression in a combined analysis of all breast cancers is mainly driven by the differential distribution of neo-antigen load and immune cell infiltration between the different breast cancer molecular subtypes (P<0.001, Mann-Whitney U-Test, for all three parameters). When TNBC is analyzed separately, neo-antigen load and CYT metagene expression is significantly inversely correlated.

#### Mutation frequencies of known cancer genes in prognostic groups

We examined the frequency of mutations in 119 annotated "cancer genes" (eTable 3) curated by Vogelstein and colleagues <sup>12,13</sup> in the 186 TNBC. 45 genes were mutated in at least three samples (eTable 1). We analyzed whether we observed differences in mutation frequencies among these 45 genes between TNBC classified as "Good" or "Poor" prognosis by the MHC2/IL8-VEGF gene signature. The only differentially altered gene with nominal significance was CASP8 with 2 of 25 samples mutated in the "Good Prognosis" group and only 1 of 161 samples mutated in the "Poor Prognosis" group (P=0.007, Chi<sup>2</sup> test without adjustment for multiple testing). When we used the B-cell/IL8-VEGF signature to assign prognosis, all 3 cases with CASP8 mutation were in the "Good Prognosis" group (n=76) (P=0.037, Chi<sup>2</sup> test). In addition, another 3 cases with mutation in the TSC2 gene were also found in this subgroup (P=0.037, Chi<sup>2</sup> test). We next studied enrichment of mutations in 12 different "cancer pathways" to which Vogelstein and colleagues have assigned the 119 cancer genes. There was no significantly differentially mutated pathway between good and poor prognosis TNBC.

#### Analysis of tumor genomic heterogeneity

Intratumor clonal heterogeneity can be inferred from genomic sequence data by several different methods including PyClone, SciClone, or EXPANDS <sup>32–35</sup>. Unfortunately these methods are suboptimal for breast cancer data because they detect only one, or very few, subclonal populations in the majority of breast cancers <sup>7</sup>. We therefore adopted another method, MATH (mutant allele tumor heterogeneity), which uses the broadness of the distribution of mutant allele frequencies as a measure of mixed cell population of the tumor <sup>11,36</sup>. The MATH score is calculated as the median absolute deviation of each somatic mutation's allelic fraction from the median allelic fraction for all mutations in the tumor, divided by the median variant allelic fraction (vaf). The use of a ratio corrects for the confounding effect of normal tissue in the sample. A detailed description of the method can be found in the article <sup>11</sup>. MATH uses the overall variance of the VAF distribution to approximate clonal heterogeneity, however this metric is influenced by the combined effect of clonality and CNAs. We calculated MATH values for TCGA samples from mutant variant allele frequencies (vaf) of all genes (see above section *Data Sources* for details). Median vaf and MAD of all mutated genes were calculated for each tumor sample and MATH calculated as 100\*MAD/median (Section-4 in *R-MarkDown-document.pdf*).

We also analyzed the level of somatic copy number alterations (SCNA) in samples stratified by the prognostic signature. TNBC harbor larger numbers of SCNAs than other breast cancer subtypes <sup>37–39</sup> and their ability to generate immunogenic epitopes has been suggested <sup>40</sup>. SCNA levels for TCGA

samples were obtained from a recent publication<sup>8</sup>. Both single nucleotide variants (SNV) and SCNA can effect MATH scores. However, MATH scores and mutational load did not correlate (**eFigure 16**), and the correlation between MATH and SCNA levels was only limited (**eFigure 10**). This suggests that all three parameters can provide independent information on tumor and clonal heterogeneity.

#### MATH scores and SCNA levels in TNBC subgroups by prognostic immune signature

MATH scores were significantly lower in good prognosis TNBC compared to poor prognosis (eFigure 22B, P=0.001, Mann-Whitney U-Test). This difference was driven by a strong inverse relationship between the MATH score and immune metagene expression (Figure 2B). There was no correlation between MATH and the IL8-VEGF metagene expression, but there was a significantly lower MATH score in MHC2 or B-cell or T-cell high TNBC (eFigure 21). Within particularly the good prognosis group, defined by high MHC2 and low IL8/VEGFR, there was a strong negative correlation between T cell infiltration measured by the T-cell metagene and clonal heterogeneity ( $R^2$ =0.479, Figure 3B, P<0.001). When MATH scores were plotted against the T-cell metagene in the poor prognosis group, we continued to observe a weak negative correlation (Figure 3A). Analogous results were obtained for SCNA levels (eFigure 22C, Figure 2C, 3C, 3D, respectively). The inverse relationship between immune infiltration and both MATH score and SCNA levels, respectively, was validated using different surrogate measures for immune cell infiltration as MHC2 or B-cell metagenes<sup>9</sup> and the CYT metric<sup>5</sup> (eFigure 11). In contrast no significant correlation with expression of the IL8-VEGF metagene was detected.

# Deconvolution of the effects of immune cell infiltration and tumor cellularity on genomic heterogeneity of TNBC

Since tumor purity can affect the power of mutation calling it could be a confounding factor in our analysis. While MATH calculation inherently controls for the amount of normal cells, tumor cellularity may effect mutational and neoantigen count data. Even if a requirement for TCGA samples was tumor cellularity of  $\geq$ 50% and most normal cell contamination may be non-TIL stroma we tried to control for read depth and for tumor purity in multivariate analysis of genomic metrics. We calculated both the median values of variant reads for all mutated genes for each tumor and obtained tumor purity estimates from ASCAT<sup>37</sup>. In multivariate regression of MATH only T-cell metagene expression was significant (P=0.003) as well as the interaction term (P=0.019), but not tumor purity (P=0.357). In a model including T-cell metagene and median read depth both parameters were significant (P<0.001) but the contribution of T-cell metagene expression to Chi square was four fold higher than for read depth. Similar to the results for MATH, we also obtained for SCNA levels independent significance of the T-cell metagene (P<0.001) in multivariate models with purity estimates from ASCAT or median read depth. For lower mutation and neoantigen load we were not able to fully deconvolute whether high immune cell infiltration or the consequently resulting lower tumor cellularity may be the major driving force, since ASCAT derived surrogate for tumor cellularity is also negatively correlated with mutation and neoantigen counts.

#### Validation by histological quantification of immune infiltration

TIL quantification by immune gene signatures from bulk tumor biopsies may be expected to be more reliable than histological quantification from single sections. Nevertheless, studies have demonstrated reproducibility of histological quantification<sup>41</sup> and good correlation with molecular methods<sup>42</sup>. Thus, we also analyzed data on histologically quantification of TILs in TCGA (see above section *Data Sources* for details). As shown in **eFigure 6** we observed a strong correlation of expression of the MHC2 metagene with histologically quantified TILs (R<sup>2</sup>=0.367) and a trend for a better prognosis in samples with  $\geq$ 50% TILs (P=0.127, **eFigure 12**). Moreover, the negative relationship between immune cell infiltration and tumor heterogeneity was also highly significant (**eFigure 13** for both MATH and SCNA levels, respectively).

#### Validation in independent METABRIC dataset

We also performed a validation of our results from the TCGA data in the independent METABRIC dataset<sup>24,25</sup> (see above section *Data Sources* for details). We calculated MATH scores based on only a small panel of 173 sequenced genes from the METABRIC dataset. Thus we expect a lower precision of the corresponding MATH score. Still we observed a highly significant (P=1e-6) negative correlation between MATH and the CYT metric for immune cell infiltration from gene expression (Spearman's rho= -0.286, **eFigure 17A**). Since no SCNA levels were available for this dataset we compared the fraction of the genome affected by CNAs<sup>24</sup> to the CYT metric and observed a trend for a negative correlation (rho=-0.104, P=0.138; **eFigure 17B**). Interestingly, we also detected significantly lower MATH scores (P<0.001) in samples from the "integrative cluster" IntClust4- (which contains most TNBC with TILs) as compared to IntClust10 (TNBC with no TILs, **eFigure 17C**).

#### **Relationship to other recent studies**

Two recent pan-cancer genome studies reported results in line with our hypothesis: One large pancancer study on lymphocyte infiltration<sup>5</sup> observed a lower than expected number of predicted neoantigens in some cancer types suggesting immune-mediated elimination. In our dataset this ratio between observed vs. predicted neoantigens was somewhat lower in the good prognosis TNBC group (0.84, SD 0.41, n=19) than in the poor prognosis group (1.03, SD 0.57, n=111), but this difference was not significant (P=0.15, T-test). Despite only such large studies allow sound statistical proof of associations in the long tailed distributions of mutations (e.g. CASP8 mutation), inclusion of different tumor types can lead to confounding effects. We show this in **eFigures 18 and 19** for comparisons across different breast cancer subtypes. Therefore we based our study only on a single subtype of breast cancer (TNBC). A second pan-cancer study observed enrichment of immune infiltration in tumors types that are characterized by lower intratumor heterogeneity (defined as number of subclonal populations from PyClone) and suggested that this may, in part, reflect results of immunoediting<sup>7</sup>.

A recent Science paper<sup>43</sup> presents data that are all in perfect agreement with our model. However, the authors of this article follow a distinct hypothesis: They suggest that only clonal neoantigens elicit an immune response and put the "cause" of the association on the cancer cells, while our hypothesis suggests that the "source" of the observed association may be the immune systems' effect on the tumor. Similarily, a very recent pan-cancer study in Nature<sup>8</sup> observed less SCNA in tumors with immune infiltration measured by gene signatures. The authors of that study conclude that aneuploidy adversely affects immune cell action. We have also used data from that study in our analyses leading to similar correlations (**Figure 2 and 3**). Several hypothetical models may explain how aneuploidy might create an immune suppressing microenvironment. However, according to our model (**Figure 1**) the situation would be vice versa with the stage of immunoediting (equilibrium/escape) "responsible" for the clonal heterogeneity and the observed aneuploidy of the tumor. We think that it is important to view the immunological and clonal stage of a tumor not as an endpoint but as a snapshot in time of evolution of the tumor-immune interaction. This view would be highly important when interpreting such biomarkers and their relationship with response to new immunological therapies.

## References

- 1. Pusztai L, Karn T, Safonov A, Abu-Khalaf MM, Bianchini G. New Strategies in Breast Cancer: Immunotherapy. *Clin. Cancer Res.* 2016. doi:10.1158/1078-0432.CCR-15-1315.
- 2. Schreiber RD, Old LJ, Smyth MJ. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. *Science*. 2011;331(6024):1565-1570. doi:10.1126/science.1203486.
- 3. Bianchini G, Qi Y, Alvarez RH, et al. Molecular Anatomy of Breast Cancer Stroma and Its Prognostic Value in Estrogen Receptor-Positive and -Negative Cancers. *Journal of Clinical Oncology*. 2010;28(28):4316-4323. doi:10.1200/JCO.2009.27.2419.
- 4. Karn T, Pusztai L, Rody A, Holtrich U, Becker S. The Influence of Host Factors on the Prognosis of Breast Cancer: Stroma and Immune Cell Components as Cancer Biomarkers. *Current Cancer Drug Targets*. 2015;15(8):652-664.
- 5. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell*. 2015;160(1-2):48-61. doi:10.1016/j.cell.2014.12.033.
- 6. Brown SD, Warren RL, Gibb EA, et al. Neo-antigens predicted by tumor genome meta-analysis correlate with increased patient survival. *Genome Res.* 2014;24(5):743-750. doi:10.1101/gr.165985.113.
- Morris LGT, Riaz N, Desrichard A, et al. Pan-cancer analysis of intratumor heterogeneity as a prognostic determinant of survival. *Oncotarget*. 2016;7(9):10051-10063. doi:10.18632/oncotarget.7067.
- 8. Davoli T, Uno H, Wooten EC, Elledge SJ. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science*. 2017;355(6322). doi:10.1126/science.aaf8399.
- 9. Rody A, Karn T, Liedtke C, et al. A clinically relevant gene signature in triple negative and basallike breast cancer. *Breast Cancer Res.* 2011;13(5):R97. doi:10.1186/bcr3035.
- Karn T, Pusztai L, Holtrich U, et al. Homogeneous Datasets of Triple Negative Breast Cancers Enable the Identification of Novel Prognostic and Predictive Signatures. *PLoS ONE*. 2011;6(12):e28403. doi:10.1371/journal.pone.0028403.
- 11. Mroz EA, Rocco JW. MATH, a novel measure of intratumor genetic heterogeneity, is high in poor-outcome classes of head and neck squamous cell carcinoma. *Oral Oncol.* 2013;49(3):211-215. doi:10.1016/j.oraloncology.2012.09.007.
- 12. Kandoth C, McLellan MD, Vandin F, et al. Mutational landscape and significance across 12 major cancer types. *Nature*. 2013;502(7471):333-339. doi:10.1038/nature12634.
- 13. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer genome landscapes. *Science*. 2013;339(6127):1546-1558. doi:10.1126/science.1235122.
- 14. Safonov A, Jiang T, Bianchini G, et al. Immune gene expression is associated with genomic aberrations in breast cancer. *Cancer Res.* 2017. doi:10.1158/0008-5472.CAN-16-3478.
- Jiang T, Shi W, Wali VB, et al. Predictors of Chemosensitivity in Triple Negative Breast Cancer: An Integrated Genomic Analysis. *PLoS Med.* 2016;13(12):e1002193. doi:10.1371/journal.pmed.1002193.
- 16. McShane LM, Altman DG, Sauerbrei W, et al. Reporting recommendations for tumor marker prognostic studies. *J Clin Oncol*. 2005;23:9067-9072.
- 17. Simon RM, Paik S, Hayes DF. Use of archived specimens in evaluation of prognostic and predictive biomarkers. J. Natl. Cancer Inst. 2009;101(21):1446-1452. doi:10.1093/jnci/djp335.
- 18. Zhu J, Sanborn JZ, Benz S, et al. The UCSC Cancer Genomics Browser. *Nat Methods*. 2009;6(4):239-240. doi:10.1038/nmeth0409-239.

- 19. Cline MS, Craft B, Swatloski T, et al. Exploring TCGA Pan-Cancer data at the UCSC Cancer Genomics Browser. *Sci Rep.* 2013;3:2652. doi:10.1038/srep02652.
- 20. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* 2012;2(5):401-404. doi:10.1158/2159-8290.CD-12-0095.
- 21. Rajasagi M, Shukla SA, Fritsch EF, et al. Systematic identification of personal tumor-specific neoantigens in chronic lymphocytic leukemia. *Blood*. 2014;124(3):453-462. doi:10.1182/blood-2014-04-567933.
- Lehmann BD, Jovanovic B, Chen X, et al. Refinement of Triple-Negative Breast Cancer Molecular Subtypes: Implications for Neoadjuvant Chemotherapy Selection. *PLoS ONE*. 2016;11(6):e0157368. doi:10.1371/journal.pone.0157368.
- 23. Fumagalli D, Blanchet-Cohen A, Brown D, et al. Transfer of clinically relevant gene expression signatures in breast cancer: from Affymetrix microarray to Illumina RNA-Sequencing technology. *BMC Genomics*. 2014;15:1008. doi:10.1186/1471-2164-15-1008.
- 24. Pereira B, Chin S-F, Rueda OM, et al. The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. *Nat Commun.* 2016;7:11479. doi:10.1038/ncomms11479.
- 25. Ali HR, Chlon L, Pharoah PDP, Markowetz F, Caldas C. Patterns of Immune Infiltration in Breast Cancer and Their Clinical Implications: A Gene-Expression-Based Retrospective Study. *PLoS Med.* 2016;13(12):e1002194. doi:10.1371/journal.pmed.1002194.
- 26. Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP. GenePattern 2.0. Nat Genet. 2006;38(5):500-501. doi:10.1038/ng0506-500.
- 27. Rody A, Holtrich U, Pusztai L, et al. T-cell metagene predicts a favorable prognosis in estrogen receptor-negative and HER2-positive breast cancers. *Breast Cancer Res.* 2009;11(2):R15. doi:10.1186/bcr2234.
- Ruffell B, Au A, Rugo HS, Esserman LJ, Hwang ES, Coussens LM. Leukocyte composition of human breast cancer. *Proc. Natl. Acad. Sci. U.S.A.* 2012;109(8):2796-2801. doi:10.1073/pnas.1104303108.
- 29. Hanker LC, Rody A, Holtrich U, et al. Prognostic evaluation of the B cell/IL-8 metagene in different intrinsic breast cancer subtypes. *Breast Cancer Res. Treat.* 2013;137(2):407-416. doi:10.1007/s10549-012-2356-2.
- 30. Karn T, Metzler D, Ruckhäberle E, et al. Data driven derivation of cutoffs from a pool of 3,030 Affymetrix arrays to stratify distinct clinical types of breast cancer. *Breast Cancer Res Treat*. 2010;120(3):567-579. doi:10.1007/s10549-009-0416-z.
- 31. Iglesia MD, Vincent BG, Parker JS, et al. Prognostic B-cell Signatures Using mRNA-Seq in Patients with Subtype-Specific Breast and Ovarian Cancer. *Clin. Cancer Res.* 2014. doi:10.1158/1078-0432.CCR-13-3368.
- 32. Roth A, Khattra J, Yap D, et al. PyClone: statistical inference of clonal population structure in cancer. *Nat Methods*. 2014;11(4):396-398. doi:10.1038/nmeth.2883.
- 33. Shah SP, Roth A, Goya R, et al. The clonal and mutational evolution spectrum of primary triplenegative breast cancers. *Nature*. 2012. doi:10.1038/nature10933.
- 34. Miller CA, White BS, Dees ND, et al. SciClone: inferring clonal architecture and tracking the spatial and temporal patterns of tumor evolution. *PLoS Comput Biol*. 2014;10(8):e1003665. doi:10.1371/journal.pcbi.1003665.
- 35. Andor N, Graham TA, Jansen M, et al. Pan-cancer analysis of the extent and consequences of intratumor heterogeneity. *Nat. Med.* 2016;22(1):105-113. doi:10.1038/nm.3984.

- 36. Mroz EA, Tward AD, Tward AM, Hammon RJ, Ren Y, Rocco JW. Intra-tumor genetic heterogeneity and mortality in head and neck cancer: analysis of data from the Cancer Genome Atlas. *PLoS Med.* 2015;12(2):e1001786. doi:10.1371/journal.pmed.1001786.
- 37. van Loo P, Nordgard SH, Lingjærde OC, et al. Allele-specific copy number analysis of tumors. *Proceedings of the National Academy of Sciences of the United States of America*. 2010;107(39):16910-16915. doi:10.1073/pnas.1009843107.
- 38. Curtis C, Shah SP, Chin S-F, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*. 2012;486(7403):346-352. doi:10.1038/nature10983.
- Ciriello G, Miller ML, Aksoy BA, Senbabaoglu Y, Schultz N, Sander C. Emerging landscape of oncogenic signatures across human cancers. *Nat. Genet.* 2013;45(10):1127-1133. doi:10.1038/ng.2762.
- 40. Schumacher TN, Schreiber RD. Neoantigens in cancer immunotherapy. *Science*. 2015;348(6230):69-74. doi:10.1126/science.aaa4971.
- 41. Salgado R, Denkert C, Demaria S, et al. Harmonization of the evaluation of tumor infiltrating lymphocytes (TILs) in breast cancer: recommendations by an international TILs-working group 2014. *Ann. Oncol.* 2014. doi:10.1093/annonc/mdu450.
- 42. Denkert C, Minckwitz G von, Brase JC, et al. Tumor-Infiltrating Lymphocytes and Response to Neoadjuvant Chemotherapy With or Without Carboplatin in Human Epidermal Growth Factor Receptor 2-Positive and Triple-Negative Primary Breast Cancers. J. Clin. Oncol. 2014. doi:10.1200/JCO.2014.58.1967.
- 43. McGranahan N, Furness AJS, Rosenthal R, et al. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science*. 2016;351(6280):1463-1469. doi:10.1126/science.aaf1490.

# eFigures



eFigure 1: Strategy of RNA-Seq and Whole-Exome-Seq analyses for TNBC classification. The analytical strategy and the flow of patients through the study is presented as recommended by the REMARK criteria ("REporting recommendations for tumour MARKer prognostic studies"

McShane et al. J Clin Oncol. 2005;23:9067).



#### eFigure 2: ER, PR and HER2 expression assessed by RNA-Seq and Agilent arrays. Correlation of Agilent-microarray (y-axis) and RNA-Seq data (x-axis) on mRNA expression of ESR1(A), PgR(B), and HER2(C) in 529 breast cancers. Expression distribution of RNA-Seq results for the same three genes (D, E, F).





#### eFigure 3: Dependency of platform correlation on gene expression level.

Spearman correlation coefficients of 16097 genes measured both on Affymetrix microarray and RNA-Seq platform in a series of 57 breast cancers were compared to the median RNA-Seq expression of the genes in the 208 TNBC samples. Poor correlation is frequent for genes with median log2 RNA-Seq expression < 2.



#### Gene correlations RNA-Seq vs Affy

# e Figure 4: Correlation between RNA-Seq and Affymetrix for Metagene clusters. Boxplots of Spearman correlation coefficients between Affymetrix microarray and RNA-Seq platform for 304 genes from Metagene clusters.



#### eFigure 5: Classification of TNBC (n=208) based on RNA-Seq data.

Heatmap of RNA-Seq data for 208 TNBC (in colums) and 304 genes (in rows) representing the metagenes given on the right. Samples were classified and sorted according the previously described TNBC subtypes (given above heatmap), a median split according to IL8-VEGF metagene, and increasing order according to immune cell infiltration as measured by T-Cell metagene expression (Rody et al. 2011).



eFigure 6: Correlation of MHC2 metagene expression and histological quantification of TILs in TCGA samples. Expression of the MHC2 metagene in TNBC samples from TCGA was compared to the quantification of mononuclear cells from Lehmann et al. 2016 from tissue slides of TCGA.



13

#### eFigure 7: Classification algorithm of the prognostic immune signature.

TNBC with high lymphocyte infiltration measured by MHC2 metagene expression (upper quartile) in combination with low IL8/VEGF-metagene expression (below median) were classified as "good prognosis" while all other samples were classified as "poor prognosis". Metagene values represent the mean of log<sub>2</sub>-transformed RNA-Seq data for the given genes.



eFigure 8: Validation of improved prognosis of TNBC patients with "Good prognosis" signature in RNA-Seq data. Survival of 402 previously published TNBC patients from Affymetrix microarray dataset (A,C) (Rody et al. 2011) and 193 TNBC new patients from TCGA RNA-Seq data (B,D) according to prognostic signatures. In Panels A and B the patients classified as "Good Prognosis Group" are defined by highest quartile of the MHC2 metagene as immune infiltration marker AND low IL8-VEGF metagene (below median of cohort). In Panels C and D samples were stratified using high B-Cell metagene expression (above lowest quartile) IL8-metagene expression (below median).



#### eFigure 9: Mutational count distribution in 186 TNBC.

The number of curated mutations for each of the 186 TNBC samples is given (sorted in decreasing order).



#### eFigure 10: Relationship between SCNA levels and MATH in TNBC from TCGA.

SCNA (somatic copy number alteration) levels are compared to the MATH score of clonal intratumor heterogeneity for 186 TNBC samples from TCGA. Despite some correlation of the two parameters ( $R^2$ =0.214) both measures mainly provide independent information.

14



# eFigure 11: Validation of inverse relationships between measures of genomic complexity and immune cell infiltration in TNBC using different immune metagenes.

The inverse relationship between immune infiltration and both MATH score (A-C) and SCNA levels (D-E), respectively, is validated using different surrogate measures for immune cell infiltration in TNBC from TCGA. The following metagene expression values are plotted on the x-axes: MHC2 metagene (A, E), B-cell metagene (B, E), and the CYT score from Rooney et al. 2015 (C, F). Spearman's rho correlation coefficient and corresponding P-values are given.



#### eFigure 12: Prognostic value of histologically quantified TILs in the TCGA TNBC dataset.

Kaplan-Meier analysis of disease free survival in TNBC from TCGA according to lymphocyte infiltration. Lymphycte predominant breast cancer (LPBC) was defined as  $\geq$ 50% mononuclear cells according to the quantification by Lehmann et al. 2016 from tissue slides of TNBC samples from the TCGA.



eFigure 13: Validation of inverse relationship of genomic heterogeneity and immune cell infiltration using histologically quantified TILs in the TCGA TNBC dataset.

MATH scores and SCNA levels are compared to lymphocyte infiltration as percent mononuclear cells according to the quantification by Lehmann et al. 2016 from tissue slides of TNBC samples from the TCGA (Spearman's rho = -0.266 and -0.396, respectively).



eFigure 14: Differences in mutation count by immune cell infiltration metagenes and IL8/VEGF metagene expression categories.

Box plots show the numbers of mutated genes in 186 TNBC stratified by B-cell metagene (q1 vs. q2-q4, n=47 vs. 139, P=0.003), MHC2 metagene (q1-3 vs. q4, n=141 vs. 45, P=0.018), and IL8/VEGF metagene expression (below versus above the median, n=95 vs. 91, n.s.). Note, that the y-axis has been croped at 170 mutated genes per sample for this presentation, which will exclude six individual hypermutated samples with 300-1200 (see **eFigure 9**). P-values are from Mann-Whitney U-Tests.



#### eFigure 15: Correlation of the number of predicted neo-antigens and mutational load.

Scatter plot showing the high correlation of total mutation count and the numbers of predicted neoantigens in 760 breast cancers ( $R^2$ =0.68).



#### eFigure 16: Independence of MATH score and total mutation counts.

Scatter plot showing lack of correlation between MATH score and the total number of mutated genes per sample ( $R^2$ <0.001).



#### eFigure 17: Validation analyses in METABRIC dataset.

# A) Negative relationship of clonal heterogeneity (MATH score) and immune cell infiltration (CYT metric from gene expression).

We calculated MATH scores based on only a small panel of 173 sequenced genes available from the METABRIC dataset. Thus we expected only a very low precision of the corresponding MATH score. Still we observed a highly significant (P=1e-6) negative correlation between this MATH score and the CYT metric for immune cell infiltration from gene expression data (Spearman's rho = -0.286).

#### B) Relationship of chromosomal instability and immune cell infiltration.

Chromosomal instability (defined as the fraction of the genome affected by CNAs by Pereira et al 2016) is compared to immune cell infiltration (measured by the CYT score according to Rooney et al. 2015) in 204 ER-/HER2- breast cancers from the METABRIC dataset (Spearman's rho= -0.104).

#### C) Difference of clonal heterogeneity between Integrative Clusters.

Boxplots of MATH scores from 222 TNBC samples from IntClust10 (characterized by low amounts of TILs, n=163) and IntClust4- (showing high TILs, n=59).





18: Differences in mutation count, neoantigen count, and CYT by molecular subtype in breast cancer.

- A) Increased total mutation counts (Rooney et al. 2015) in TNBC compared to other breast cancer subtypes (Luminal A, n=239; Luminal B, n=345; HER2, n=35; TNBC, n=141).
- B) Increased predicted neoantigen load (Rooney et al. 2015) compared to other breast cancer subtypes.
- C) Increased immune cytolytic activity (CYT) according to Rooney et al. 2015 in TNBC subtype.



eFigure 19: Confounding of molecular breast cancer subtypes on predicted neoantigen count and CYT. Relationship of CYT (y-axis) and predicted neoantigen load (x-axis) over all breast cancer subtypes (panel A), and in subgroups of Luminal A (panel B), Luminal B (panel C), all luminal (panel D), HER2 (panel E), and TNBC (panel F). Lines represent lowess fit, rho-values from Spearman rank correlation and corresponding Pvalues are given. Significance was only observed for a small positive correlation in Luminal B (rho=0.122, panel C) and a negative correlation in TNBC (rho= -0.187, panel F).



eFigure 20: High inter-correlation of TIL-metagenes

Scatter plot showing the high correlation of MHC2- and T-cell metagene expression in 208 TNBC from TCGA ( $R^2$ =0.807). The sets of individual genes represented by the two metagenes are also listed left and right from the scatter plot.



#### eFigure 21: Association between clonal heterogeneity and immune metagene expression.

MATH score difference between individual components of the prognostic signatures: Box plots of TNBC cohorts with low or high expression of either B-cell (q1 vs. q2-q4, n=47 vs. 139), MHC2 (q1-3 vs. q4, n=141 vs. 45), T-Cell (q1-3 vs. q4, n=140 vs. 46), or IL8-VEGF (below versus above the median, n=95 vs. 91) metagenes (P-values from Mann-Whitney U-Tests).



#### eFigure 22: Association of MATH and SCNA With Prognostic Groups in TNBC.

- A) Kaplan-Meier analysis of disease-free survival of 193 TNBC samples with follow-up data from TCGA. Patients classified as "Good Prognosis Group" are defined by highest quartile of the MHC2 metagene as immune infiltration marker AND low IL8-VEGF metagene (below median of cohort) as in Figure 2 (P=0.019, log-rank test).
- B) Box plot showing differences in MATH value between prognostic groups in A (P=0.001, Mann Whitney test).
- C) Box plot showing differences in SCNA levels between prognostic groups in A (P=0.001, Mann Whitney test).

19 © 2017 American Medical Association. All rights reserved.

# eTables

ĺ	Gene	TNBC with	Proportion	Cancer pathway
		mutations (186 total)	of samples	
	TP53	139	74.7%	Cell Cycle/Apoptosis; DNA Damage Control
	PIK3CA	16	8.6%	РІЗК
	MLL3	10	5.4%	Chromatin Modification
	PTEN	9	4.8%	РІЗК
	MLL2	9	4.8%	Chromatin Modification
	FANCD2	7	3.8%	DNA Damage Control
	BRCA1	7	3.8%	DNA Damage Control
	NF1	7	3.8%	RAS
	NCOR1	7	3.8%	Chromatin Modification
	FBXW7	7	3.8%	NOTCH
	ARID1B	7	3.8%	Chromatin Modification
	CREBBP	6	3.2%	Chromatin Modification; Transcriptional Regulation
	APC	6	3.2%	APC
	NOTCH2	6	3.2%	NOTCH
	RB1	6	3.2%	Cell Cycle/Apoptosis
	BRCA2	5	2.7%	DNA Damage Control
	ATM	4	2.2%	DNA Damage Control
	PALB2	4	2.2%	DNA Damage Control
	PMS1	4	2.2%	DNA Damage Control
	BAP1	4	2.2%	DNA Damage Control
	KDM6A	4	2.2%	Chromatin Modification
	SMARCA4	4	2.2%	Chromatin Modification
	MED12	4	2.2%	Cell Cycle/Apoptosis; TGF-b
	BCOR	4	2.2%	Transcriptional Regulation
	ARID2	4	2.2%	Chromatin Modification
	CARD11	4	2.2%	Cell Cycle/Apoptosis
	NOTCH1	4	2.2%	NOTCH
	ASXL1	4	2.2%	Chromatin Modification
	MSH6	3	1.6%	DNA Damage Control
	STAG2	3	1.6%	DNA Damage Control
	PIK3R1	3	1.6%	РІЗК
	EP300	3	1.6%	Chromatin Modification; APC; TGF-b; NOTCH
	TET2	3	1.6%	Chromatin Modification
	ACVR1B	3	1.6%	TGF-b
	JAK2	3	1.6%	STAT
	ALK	3	1.6%	PI3K; RAS
	ARID1A	3	1.6%	Chromatin Modification
	CASP8	3	1.6%	Cell Cycle/Apoptosis
	TRAF7	3	1.6%	Apoptosis
	FAM123B	3	1.6%	APC
	MAP2K1	3	1.6%	RAS
	TSC2	3	1.6%	PI3K
	SETD2	3	1.6%	Chromatin Modification
	CDH1	3	1.6%	APC
	EXT1	3	1.6%	Hedgehog

#### eTable 1: Annotated cancer-genes mutated in $\ge$ 3 samples.

#### eTable 2: TCGA samples included in the study.

TCGA-Sample Patient Barcode	TCGA Sample Patient Barcode	TCGA Sample Patient Barcode	TCGA-Sample Patient Barcode	TCGA Sample Patient Barcode	TCGA-Sample Patient Barcode	TCGA-E2-A159-01 TCGA-E2-A159	TCGA Sample Patient Barcode
TCGA-3C-AAAU-01 TCGA-3C-AAAU	TCGA-A7-A3IY-01 TCGA-A7-A3IY	TCGA-AC-A8OP-01 TCGA-AC-A8OP	TCGA-AR-A255-01 TCGA-AR-A255	TCGA-BH-A0DX-01 TCGA-BH-A0DX	TCGA-C8-A12P-01 TCGA-C8-A12P		TCGA-EW-A1J1-01 TCGA-EW-A1J1
TCGA-3C-AALI-01 TCGA-3C-AALI	TCGA-A7-A3IZ-01 TCGA-A7-A3IZ	TCGA-AC-A80Q-01 TCGA-AC-A80Q	TCGA-AR-A256-01 TCGA-AR-A256	TCGA-BH-A0DZ-01 TCGA-BH-A0DZ	TCGA-C8-A12Q-01 TCGA-C8-A12Q	TCGA-E2-A15A-01 TCGA-E2-A15A	TCGA-EW-A1J2-01 TCGA-EW-A1J2
TCGA-3C-AALI-01 TCGA-3C-AALJ	TCGA-A7-A3J0-01 TCGA-A7-A3J0	TCGA-AC-A80R-01 TCGA-AC-A80R	TCGA-AR-A2LE-01 TCGA-AR-A2LE	TCGA-BH-A0DZ-11 TCGA-BH-A0DZ	TCGA-C8-A12T-01 TCGA-C8-A12T	TCGA-E2-A15A-06 TCGA-E2-A15A	TCGA-EW-A1J3-01 TCGA-EW-A1J3
TCGA-3C-AALK-01 TCGA-3C-AALK	TCGA-A7-A3I1-01 TCGA-A7-A3I1	TCGA-AC-A80S-01 TCGA-AC-A80S	TCGA-AR-A2LH-01 TCGA-AR-A2LH	TCGA-BH-A0E0-01 TCGA-BH-A0E0	TCGA-C8-A12U-01 TCGA-C8-A12U	TCGA-E2-A15C-01 TCGA-E2-A15C	TCGA-EW-A1J5-01 TCGA-EW-A1J5
TCGA-4H-AAAK-01 TCGA-4H-AAAK	TCGA-A7-A3RF-01 TCGA-A7-A3RF	TCGA-AN-A03X-01 TCGA-AN-A03X	TCGA-AR-A2LJ-01 TCGA-AR-A2LJ	TCGA-BH-A0E0-11 TCGA-BH-A0E0	TCGA-C8-A12V-01 TCGA-C8-A12V	TCGA-E2-A15D-01 TCGA-E2-A15D	TCGA-EW-A1J6-01 TCGA-EW-A1J6
TCGA-5L-AAT0-01 TCGA-5L-AAT0	TCGA-A7-A425-01 TCGA-A7-A425	TCGA-AN-A03Y-01 TCGA-AN-A03Y	TCGA-AR-A2LK-01 TCGA-AR-A2LK	TCGA-BH-A0E1-01 TCGA-BH-A0E1	TCGA-C8-A12W-01 TCGA-C8-A12W	TCGA-E2-A15E-01 TCGA-E2-A15E	TCGA-EW-A10V-01 TCGA-EW-A10V
TCGA-5L-AAT1-01 TCGA-5L-AAT1	TCGA-A7-A426-01 TCGA-A7-A426	TCGA-AN-A041-01 TCGA-AN-A041	TCGA-AR-A2LL-01 TCGA-AR-A2LL	TCGA-BH-A0E1-11 TCGA-BH-A0E1	TCGA-C8-A12X-01 TCGA-C8-A12X	TCGA-E2-A15E-06 TCGA-E2-A15E	TCGA-EW-A10W-01 TCGA-EW-A10W
TCGA-51-A9QA-01 TCGA-51-A9QA	TCGA-A7-A4SA-01 TCGA-A7-A4SA	TCGA-AN-A046-01 TCGA-AN-A046	TCGA-AR-A2LM-01 TCGA-AR-A2LM	TCGA-BH-A0E2-01 TCGA-BH-A0E2	TCGA-C8-A12Y-01 TCGA-C8-A12Y	TCGA-E2-A15F-01 TCGA-E2-A15F	TCGA-EW-A10X-01 TCGA-EW-A10X
TCGA-A1-A0SB-01 TCGA-A1-A0SB	TCGA-A7-A4SB-01 TCGA-A7-A4SB	TCGA-AN-A049-01 TCGA-AN-A049	TCGA-AR-A2LN-01 TCGA-AR-A2LN	TCGA-BH-A0E6-01 TCGA-BH-A0E6	TCGA-C8-A12Z-01 TCGA-C8-A12Z	TCGA-E2-A15G-01 TCGA-E2-A15G	TCGA-EW-A10Y-01 TCGA-EW-A10Y
TCGA-A1-A0SD-01 TCGA-A1-A0SD	TCGA-A7-A4SC-01 TCGA-A7-A4SC	TCGA-AN-A040-01 TCGA-AN-A040	TCGA-AR-A2LO-01 TCGA-AR-A2LO	TCGA-BH-A0E7-01 TCGA-BH-A0E7	TCGA-C8-A130-01 TCGA-C8-A130	TCGA-E2-A15H-01 TCGA-E2-A15H	TCGA-EW-A107-01 TCGA-EW-A107
TCGA-A1-A0SE-01 TCGA-A1-A0SE	TCGA-A7-A4SD-01 TCGA-A7-A4SD	TCGA-AN-A04C-01 TCGA-AN-A04C	TCGA-AR-A2LQ-01 TCGA-AR-A2LQ	TCGA-BH-A0E9-01 TCGA-BH-A0E9	TCGA-C8-A131-01 TCGA-C8-A131	TCGA-E2-A15I-01 TCGA-E2-A15I	TCGA-EW-A1P0-01 TCGA-EW-A1P0
TCGA-A1-A0SF-01 TCGA-A1-A0SF	TCGA-A7-A4SE-01 TCGA-A7-A4SE	TCGA-AN-A04D-01 TCGA-AN-A04D	TCGA-AR-A2LR-01 TCGA-AR-A2LR	TCGA-BH-A0EA-01 TCGA-BH-A0EA	TCGA-C8-A132-01 TCGA-C8-A132	TCGA-E2-A15I-11 TCGA-E2-A15I	TCGA-EW-A1P1-01 TCGA-EW-A1P1
TCGA-A1-A0SG-01 TCGA-A1-A0SG	TCGA-A7-A4SF-01 TCGA-A7-A4SF	TCGA-AN-ADAJ-01 TCGA-AN-ADAJ	TCGA-AR-A5QM-01 TCGA-AR-A5QM	TCGA-BH-A0EB-01 TCGA-BH-A0EB	TCGA-C8-A133-01 TCGA-C8-A133	TCGA-E2-A15J-01 TCGA-E2-A15J	TCGA-EW-A1P3-01 TCGA-EW-A1P3
TCGA-A1-A0SH-01 TCGA-A1-A0SH	TCGA-A7-A56D-01 TCGA-A7-A56D	TCGA-AN-ADAK-01 TCGA-AN-ADAK	TCGA-AR-A5QN-01 TCGA-AR-A5QN	TCGA-BH-A0EE-01 TCGA-BH-A0EE	TCGA-C8-A134-01 TCGA-C8-A134	TCGA-E2-A15K-01 TCGA-E2-A15K	TCGA-EW-A1P4-01 TCGA-EW-A1P4
TCGA-A1-A0SI-01 TCGA-A1-A0SI	TCGA-A7-A5ZV-01 TCGA-A7-A5ZV	TCGA-AN-A0AL-01 TCGA-AN-A0AL	TCGA-AR-A5QP-01 TCGA-AR-A5QP	TCGA-BH-A0EI-01 TCGA-BH-A0EI	TCGA-C8-A135-01 TCGA-C8-A135	TCGA-E2-A15K-06 TCGA-E2-A15K	TCGA-EW-A1P5-01 TCGA-EW-A1P5
TCGA-A1-A0SJ-01 TCGA-A1-A0SJ	TCGA-A7-A5ZW-01 TCGA-A7-A5ZW	TCGA-AN-A0AM-01 TCGA-AN-A0AM	TCGA-AR-A5QQ-01 TCGA-AR-A5QQ	TCGA-BH-A0GY-01 TCGA-BH-A0GY	TCGA-C8-A137-01 TCGA-C8-A137	TCGA-E2-A15K-11 TCGA-E2-A15K	TCGA-EW-A1P6-01 TCGA-EW-A1P6
TCGA-A1-A0SK-01 TCGA-A1-A0SK	TCGA-A7-A52X-01 TCGA-A7-A52X	TCGA-AN-ADAR-01 TCGA-AN-ADAR	TCGA-86-A0I1-01 TCGA-86-A0I1	TCGA-BH-A0GZ-01 TCGA-BH-A0GZ	TCGA-C8-A138-01 TCGA-C8-A138	TCGA-E2-A15L-01 TCGA-E2-A15L	TCGA-EW-A1P7-01 TCGA-EW-A1P7
TCGA-A1-A0SM-01 TCGA-A1-A0SM	TCGA-A7-A6VV-01 TCGA-A7-A6VV	TCGA-AN-ADAS-01 TCGA-AN-ADAS	TCGA-86-A0I2-01 TCGA-86-A0I2	TCGA-BH-A0H0-01 TCGA-BH-A0H0	TCGA-C8-A1HE-01 TCGA-C8-A1HE	TCGA-E2-A15M-01 TCGA-E2-A15M	TCGA-EW-A1P8-01 TCGA-EW-A1P8
TCGA-A1-AUSN-U1 TCGA-A1-AUSN	TCGA-A7-A6VW-01 TCGA-A7-A6VW	TCGA-AN-ADAT-D1 TCGA-AN-ADAT	TCGA-86-A0I5-01 TCGA-86-A0I5	TCGA-BH-ADH3-01 TCGA-BH-ADH3	TCGA-C8-A1HF-01 TCGA-C8-A1HF	TCGA-E2-A15M-11 TCGA-E2-A15M	TCGA-EW-AIPA-01 TCGA-EW-AIPA
TCGA-A1-A0SO-01 TCGA-A1-A0SO	TCGA-A7-A6VX-01 TCGA-A7-A6VX	TCGA-AN-AOFD-D1 TCGA-AN-AOFD	TCGA-86-A0I6-01 TCGA-86-A0I6	TCGA-BH-ADH5-01 TCGA-BH-ADH5	TCGA-C8-A1HG-01 TCGA-C8-A1HG	TCGA-E2-A150-01 TCGA-E2-A150	TCGA-EW-AIPB-01 TCGA-EW-AIPB
TCGA-A1-A0SD-01 TCGA-A1-A0SD	TCGA-A7-A6VX-01 TCGA-A7-A6VX	TCGA-AN-AOFE-D1 TCGA-AN-AOFE	TCGA-86-A0I8-01 TCGA-86-A0I8	TCGA-BH-ADH5-11 TCGA-BH-ADH5	TCGA-C8-A1HL01 TCGA-C8-A1HI	TCGA-E2-A158-01 TCGA-E2-A158	TCGA-EW-AIPC-01 TCGA-EW-AIPC
TCGA-A1-A0SQ-01 TCGA-A1-A0SQ TCGA-A1-A0SQ-01 TCGA-A1-A0SQ TCGA-A2-A04N-01 TCGA-A2-A04N	TCGA-A8-A06N-01 TCGA-A8-A06N TCGA-A8-A060-01 TCGA-A8-A060	TCGA-AN-ADFJ-01 TCGA-AN-ADFJ TCGA-AN-ADFJ-01 TCGA-AN-ADFJ TCGA-AN-ADFK-01 TCGA-AN-ADFK	TCGA-86-A0I9-01 TCGA-86-A0I9 TCGA-86-A0I9-01 TCGA-86-A0I9 TCGA-86-A0IA-01 TCGA-86-A0IA	TCGA-BH-A0H5-11 TCGA-BH-A0H5 TCGA-BH-A0H6-01 TCGA-BH-A0H6 TCGA-BH-A0H7-01 TCGA-BH-A0H7	TCGA-C8-A1H/-01 TCGA-C8-A1H/ TCGA-C8-A1H/-01 TCGA-C8-A1H/ TCGA-C8-A1HK-01 TCGA-C8-A1HK	TCGA-E2-A15P-01 TCGA-E2-A15P TCGA-E2-A15R-01 TCGA-E2-A15R TCGA-E2-A15S-01 TCGA-E2-A15S	TCGA-EW-AIPC-01 TCGA-EW-AIPC TCGA-EW-AIPD-01 TCGA-EW-AIPD TCGA-EW-AIPE-01 TCGA-EW-AIPE
TCGA-A2-A04P-01 TCGA-A2-A04P	TCGA-A8-A06P-01 TCGA-A8-A06P	TCGA-AN-A0FL-01 TCGA-AN-A0FL	TCGA-B6-A0IB-01 TCGA-B6-A0IB	TCGA-BH-A0H7-11 TCGA-BH-A0H7	TCGA-C8-A1HL-01 TCGA-C8-A1HL	TCGA-E2-A15T-01 TCGA-E2-A15T	TCGA-EW-A1PF-01 TCGA-EW-A1PF
TCGA-A2-A04Q-01 TCGA-A2-A04Q	TCGA-A8-A06Q-01 TCGA-A8-A06Q	TCGA-AN-A0FN-01 TCGA-AN-A0FN	TCGA-B6-A0IC-01 TCGA-B6-A0IC	TCGA-BH-A0H9-01 TCGA-BH-A0H9	TCGA-C8-A1HM-01 TCGA-C8-A1HM	TCGA-E2-A1A2-01 TCGA-E2-A1AZ	TCGA-EW-A1PG-01 TCGA-EW-A1PG
TCGA-A2-A04R-01 TCGA-A2-A04R	TCGA-A8-A06R-01 TCGA-A8-A06R	TCGA-AN-A0FS-01 TCGA-AN-A0FS	TCGA-B6-A0IE-01 TCGA-B6-A0IE	TCGA-BH-A0HA-01 TCGA-BH-A0HA	TCGA-C8-A1HN-01 TCGA-C8-A1HN	TCGA-E2-A1B0-01 TCGA-E2-A1B0	TCGA-EW-A1PH-01 TCGA-EW-A1PH
TCGA-A2-A04T-01 TCGA-A2-A04T	TCGA-A8-A06T-01 TCGA-A8-A06T	TCGA-AN-A0FT-01 TCGA-AN-A0FT	TCGA-B6-A0IG-01 TCGA-B6-A0IG	TCGA-BH-A0HA-11 TCGA-BH-A0HA	TCGA-C8-A1HO-01 TCGA-C8-A1HO	TCGA-E2-A1B1-01 TCGA-E2-A1B1	TCGA-EW-A2FR-01 TCGA-EW-A2FR
TCGA-A2-A04U-01 TCGA-A2-A04U	TCGA-A8-A06U-01 TCGA-A8-A06U	TCGA-AN-A0FV-01 TCGA-AN-A0FV	TCGA-86-A0IH-01 TCGA-86-A0IH	TCGA-BH-A0HB-01 TCGA-BH-A0HB	TCGA-C8-A26V-01 TCGA-C8-A26V	TCGA-E2-A1B4-01 TCGA-E2-A1B4	TCGA-EW-A2FS-01 TCGA-EW-A2FS
TCGA-A2-A04V-01 TCGA-A2-A04V	TCGA-A8-A06X-01 TCGA-A8-A06X	TCGA-AN-A0FW-01 TCGA-AN-A0FW	TCGA-86-A0IJ-01 TCGA-86-A0IJ	TCGA-BH-A0HF-01 TCGA-BH-A0HF	TCGA-C8-A26W-01 TCGA-C8-A26W	TCGA-E2-A1B5-01 TCGA-E2-A1B5	TCGA-EW-A2FV-01 TCGA-EW-A2FV
TCGA-A2-A04W-01 TCGA-A2-A04W	TCGA-A8-A067-01 TCGA-A8-A067	TCGA-AN-ADFX-01 TCGA-AN-ADFX	TCGA-86-A0IN-01 TCGA-86-A0IM	TCGA-BH-ADHI-D1 TCGA-BH-ADHI	TCGA-C8-A26X-01 TCGA-C8-A26X	TCGA-E2-A180-01 TCGA-E2-A180	TCGA-EW-A2FW-01 TCGA-EW-A2FW
TCGA-A2-A04X-01 TCGA-A2-A04X	TCGA-A8-A062-01 TCGA-A8-A062	TCGA-AN-ADFY-01 TCGA-AN-ADFY	TCGA-86-A0IM-01 TCGA-86-A0IM	TCGA-BH-ADHK-01 TCGA-BH-ADHK	TCGA-C8-A26Y-01 TCGA-C8-A26Y	TCGA-E2-A18C-01 TCGA-E2-A18C	TCGA-EW-A3E8-01 TCGA-EW-A3E8
TCGA-A2-A04X-01 TCGA-A2-A04X	TCGA-A8-A075-01 TCGA-A8-A075	TCGA-AN-ADFZ-01 TCGA-AN-ADFZ	TCGA-86-A0IN-01 TCGA-86-A0IN	TCGA-BH-ADHK-11 TCGA-BH-ADHK	TCGA-C8-A267-01 TCGA-C8-A267	TCGA-F2-A18C-11 TCGA-F2-A18C	TCGA-FW-A3U0-01 TCGA-FW-A3U0
TCGA-A2-A0CK-01 TCGA-A2-A0CK	TCGA-A8-A076-01 TCGA-A8-A076	TCGA-AN-A0G0-01 TCGA-AN-A0G0	TCGA-86-A0IO-01 TCGA-86-A0IO	TCGA-BH-A0HL-01 TCGA-BH-A0HL	TCGA-C8-A273-01 TCGA-C8-A273	TCGA-E2-A1BD-01 TCGA-E2-A1BD	TCGA-EW-A423-01 TCGA-EW-A423
TCGA-A2-A0CL-01 TCGA-A2-A0CL	TCGA-A8-A079-01 TCGA-A8-A079	TCGA-AN-A0XL-01 TCGA-AN-A0XL	TCGA-86-A0IP-01 TCGA-86-A0IP	TCGA-BH-A0HN-01 TCGA-BH-A0HN	TCGA-C8-A274-01 TCGA-C8-A274	TCGA-E2-A1IE-01 TCGA-E2-A1IE	TCGA-EW-A424-01 TCGA-EW-A424
TCGA-A2-A0CM-01 TCGA-A2-A0CM	TCGA-A8-A07B-01 TCGA-A8-A07B	TCGA-AN-A0XN-01 TCGA-AN-A0XN	TCGA-B6-A0IQ-01 TCGA-B6-A0IQ	TCGA-BH-A0HO-01 TCGA-BH-A0HO	TCGA-C8-A275-01 TCGA-C8-A275	TCGA-E2-A1IF-01 TCGA-E2-A1IF	TCGA-EW-A6S9-01 TCGA-EW-A6S9
TCGA-A2-A0CO-01 TCGA-A2-A0CO	TCGA-A8-A07C-01 TCGA-A8-A07C	TCGA-AN-A0XO-01 TCGA-AN-A0XO	TCGA-B6-A0RE-01 TCGA-B6-A0RE	TCGA-BH-A0HP-01 TCGA-BH-A0HP	TCGA-C8-A278-01 TCGA-C8-A278	TCGA-E2-A1IG-01 TCGA-E2-A1IG	TCGA-EW-A6SA-01 TCGA-EW-A6SA
TCGA-A2-A0CP-01 TCGA-A2-A0CP TCGA-A2-A0CQ-01 TCGA-A2-A0CQ TCGA-A2-A0CQ-01 TCGA-A2-A0CQ	TCGA-A8-A07E-01 TCGA-A8-A07E TCGA-A8-A07F-01 TCGA-A8-A07F	TCGA-AN-A0XP-01 TCGA-AN-A0XP TCGA-AN-A0XR-01 TCGA-AN-A0XR TCGA AN A0X5 01 TCGA AN A0X5	TCGA-86-A0RG-01 TCGA-86-A0RG TCGA-86-A0RH-01 TCGA-86-A0RH	TCGA-BH-A0HQ-01 TCGA-BH-A0HQ TCGA-BH-A0HU-01 TCGA-BH-A0HU	TCGA-C8-A27A-01 TCGA-C8-A27A TCGA-C8-A27B-01 TCGA-C8-A27B TCGA-C8-A27B-01 TCGA-C8-A27A	TCGA-E2-A1IG-11 TCGA-E2-A1IG TCGA-E2-A1IH-01 TCGA-E2-A1IH TCGA-E2-A1IH-01 TCGA-E2-A1IH	TCGA-EW-A6SB-01 TCGA-EW-A6SB TCGA-EW-A6SC-01 TCGA-EW-A6SC TCGA EW A6SD 01 TCGA EW A6SD
TCGA-A2-A0CK-01 TCGA-A2-A0CK TCGA-A2-A0CS-01 TCGA-A2-A0CS TCGA-A2-A0CT-01 TCGA-A2-A0CT	TCGA-A8-A07I-01 TCGA-A8-A07I TCGA-A8-A07I-01 TCGA-A8-A07I TCGA-A8-A07J-01 TCGA-A8-A07J	TCGA-AN-ADX3-01 TCGA-AN-ADX3 TCGA-AN-ADXT-01 TCGA-AN-ADXT TCGA-AN-ADXU-01 TCGA-AN-ADXU	TCGA-B6-A0RL-01 TCGA-B6-A0RL TCGA-B6-A0RL-01 TCGA-B6-A0RL	TCGA-BH-ADHX-01 TCGA-BH-ADHX TCGA-BH-ADHX-01 TCGA-BH-ADHX TCGA-BH-ADHY-01 TCGA-BH-ADHY	TCGA-C8-A3M8-01 TCGA-C8-A3M8 TCGA-C8-A3M8-01 TCGA-C8-A3M8 TCGA-C8-A8HP-01 TCGA-C8-A8HP	TCGA-E2-A1II-01 TCGA-E2-A1II TCGA-E2-A1IJ-01 TCGA-E2-A1IJ TCGA-E2-A1IK-01 TCGA-E2-A1IK	TCGA-GI-A2C8-01 TCGA-GI-A2C8 TCGA-GI-A2C8-11 TCGA-GI-A2C8
TCGA-A2-A0CU-01 TCGA-A2-A0CU	TCGA-A8-A07L-01 TCGA-A8-A07L	TCGA-AN-A0XV-01 TCGA-AN-A0XV	TCGA-86-A0RN-01 TCGA-86-A0RN	TCGA-BH-A0RX-01 TCGA-BH-A0RX	TCGA-C8-A8HQ-01 TCGA-C8-A8HQ	TCGA-E2-A1IL-01 TCGA-E2-A1IL	TCGA-GI-A2C9-01 TCGA-GI-A2C9
TCGA-A2-A0CV-01 TCGA-A2-A0CV	TCGA-A8-A07O-01 TCGA-A8-A07O	TCGA-AN-A0XW-01 TCGA-AN-A0XW	TCGA-86-A0RO-01 TCGA-86-A0RO	TCGA-BH-A0W3-01 TCGA-BH-A0W3	TCGA-C8-A8HR-01 TCGA-C8-A8HR	TCGA-E2-A1IN-01 TCGA-E2-A1IN	TCGA-GI-A2C9-11 TCGA-GI-A2C9
TCGA-A2-A0CW-01 TCGA-A2-A0CW	TCGA-A8-A07P-01 TCGA-A8-A07P	TCGA-AO-A03L-01 TCGA-AO-A03L	TCGA-B6-A0RP-01 TCGA-B6-A0RP	TCGA-BH-A0W4-01 TCGA-BH-A0W4	TCGA-D8-A13Y-01 TCGA-D8-A13Y	TCGA-E2-A1IO-01 TCGA-E2-A1IO	TCGA-GM-A2D9-01 TCGA-GM-A2D9
TCGA-A2-A0CX-01 TCGA-A2-A0CX	TCGA-A8-A07R-01 TCGA-A8-A07R	TCGA-AO-A03M-01 TCGA-AO-A03M	TCGA-B6-A0RQ-01 TCGA-B6-A0RQ	TCGA-BH-A0W5-01 TCGA-BH-A0W5	TCGA-D8-A13Z-01 TCGA-D8-A13Z	TCGA-E2-A1IP-01 TCGA-E2-A1IP	TCGA-GM-A2DA-01 TCGA-GM-A2DA
TCGA-A2-A0CY-01 TCGA-A2-A0CY TCGA-A2-A0CZ-01 TCGA-A2-A0CZ TCGA-A2-A0D0-01 TCGA-A2-A0D0	TCGA-A8-A07U-01 TCGA-A8-A07U TCGA-A8-A07U-01 TCGA-A8-A07U	TCGA-AD-AD3N-01 TCGA-AD-AD3N TCGA-AD-AD3O-01 TCGA-AD-AD3O TCGA-AD-AD3R-01 TCGA-AD-AD3R	TCGA-86-A0RS-01 TCGA-86-A0RS TCGA-86-A0RT-01 TCGA-86-A0RT TCGA-86-A0RU-01 TCGA-86-A0RU	TCGA-BH-A0W7-01 TCGA-BH-A0W7 TCGA-BH-A0WA-01 TCGA-BH-A0WA TCGA-BH-A18E-01 TCGA-BH-A18E	TCGA-D8-A140-01 TCGA-D8-A140 TCGA-D8-A141-01 TCGA-D8-A141 TCGA-D8-A142-01 TCGA-D8-A142	TCGA-E2-A1IU-01 TCGA-E2-A1IU TCGA-E2-A1L6-01 TCGA-E2-A1L6 TCGA-E2-A117-01 TCGA-E2-A117	TCGA-GM-A2DB-01 TCGA-GM-A2DB TCGA-GM-A2DC-01 TCGA-GM-A2DD TCGA-GM-A2DD-01 TCGA-GM-A2DD
TCGA-A2-A0D0-01 TCGA-A2-A0D0 TCGA-A2-A0D1-01 TCGA-A2-A0D1 TCGA-A2-A0D2-01 TCGA-A2-A0D2	TCGA-A8-A07Z-01 TCGA-A8-A07Z TCGA-A8-A081-01 TCGA-A8-A081	TCGA-AO-A03P-01 TCGA-AO-A03P TCGA-AO-A03R-01 TCGA-AO-A03R TCGA-AO-A03T-01 TCGA-AO-A03T	TCGA-B6-A0RV-01 TCGA-B6-A0RV TCGA-B6-A0RV-01 TCGA-B6-A0RV	TCGA-BH-A18F-01 TCGA-BH-A18F TCGA-BH-A18G-01 TCGA-BH-A18G TCGA-BH-A18H-01 TCGA-BH-A18H	TCGA-D8-A142-01 TCGA-D8-A142 TCGA-D8-A143-01 TCGA-D8-A143 TCGA-D8-A145-01 TCGA-D8-A145	TCGA-E2-A1L7-01 TCGA-E2-A1L7 TCGA-E2-A1L7-11 TCGA-E2-A1L7 TCGA-E2-A1L8-01 TCGA-E2-A1L8	TCGA-GM-A2DD-01 TCGA-GM-A2DD TCGA-GM-A2DF-01 TCGA-GM-A2DF TCGA-GM-A2DH-01 TCGA-GM-A2DH
TCGA-A2-A0D3-01 TCGA-A2-A0D3	TCGA-A8-A082-01 TCGA-A8-A082	TCGA-AO-A03U-01 TCGA-AO-A03U	TCGA-B6-A0WT-01 TCGA-B6-A0WT	TCGA-BH-A18I-01 TCGA-BH-A18I	TCGA-D8-A146-01 TCGA-D8-A146	TCGA-E2-A1L9-01 TCGA-E2-A1L9	TCGA-GM-A2DI-01 TCGA-GM-A2DI
TCGA-A2-A0D4-01 TCGA-A2-A0D4	TCGA-A8-A083-01 TCGA-A8-A083	TCGA-AO-A03V-01 TCGA-AO-A03V	TCGA-B6-A0WV-01 TCGA-B6-A0WV	TCGA-BH-A18J-01 TCGA-BH-A18J	TCGA-D8-A147-01 TCGA-D8-A147	TCGA-E2-A1LA-01 TCGA-E2-A1LA	TCGA-GM-A2DK-01 TCGA-GM-A2DK
TCGA-A2-A0EM-01 TCGA-A2-A0EM	TCGA-A8-A084-01 TCGA-A8-A084	TCGA-AO-A0J2-01 TCGA-AO-A0J2	TCGA-B6-A0WW-01 TCGA-B6-A0WW	TCGA-BH-A18J-11 TCGA-BH-A18J	TCGA-D8-A1J8-01 TCGA-D8-A1J8	TCGA-E2-A1LB-01 TCGA-E2-A1LB	TCGA-GM-A2DL-01 TCGA-GM-A2DL
TCGA-A2-A0EN-01 TCGA-A2-A0EN	TCGA-A8-A085-01 TCGA-A8-A085	TCGA-AO-A0J3-01 TCGA-AO-A0J3	TCGA-B6-A0WX-01 TCGA-B6-A0WX	TCGA-BH-A18K-01 TCGA-BH-A18K	TCGA-D8-A1J9-01 TCGA-D8-A1J9	TCGA-E2-A1LB-11 TCGA-E2-A1LB	TCGA-GM-A2DM-01 TCGA-GM-A2DM
TCGA-A2-A0EO-01 TCGA-A2-A0EO	TCGA-A8-A086-01 TCGA-A8-A086	TCGA-AO-ADJ4-01 TCGA-AO-ADJ4	TCGA-86-A0W2-01 TCGA-86-A0W2	TCGA-BH-A18k-11 TCGA-BH-A18k	TCGA-D8-A1JA-D1 TCGA-D8-A1JA	TCGA-E2-AILE-01 TCGA-E2-AILE	TCGA-GM-A2D0-01 TCGA-GM-A2D0
TCGA-A2-A0EP-01 TCGA-A2-A0EP	TCGA-A8-A08A-01 TCGA-A8-A08A	TCGA-AO-ADJ5-01 TCGA-AO-ADJ5	TCGA-86-A0W2-01 TCGA-86-A0W2	TCGA-BH-A18L-01 TCGA-BH-A18L	TCGA-D8-A1JB-D1 TCGA-D8-A1JB	TCGA-E2-AILG-01 TCGA-E2-AILG	TCGA-GM-A2D0-01 TCGA-GM-A2D0
TCGA-A2-A0EO-01 TCGA-A2-A0EO	TCGA-A8-A08B-01 TCGA-A8-A088	TCGA-AO-ADJ6-01 TCGA-AO-ADJ6	TCGA-86-A0X0-01 TCGA-86-A0X0	TCGA-BH-A18L-11 TCGA-BH-A18L	TCGA-D8-A1JC-D1 TCGA-D8-A1JC	TCGA-E2-AILH-01 TCGA-E2-AILH	TCGA-GM-A3NW-01 TCGA-GM-A3NW
TCGA-A2-A0ER-01 TCGA-A2-A0ER	TCGA-A8-A08C-01 TCGA-A8-A08C	TCGA-AO-A0J7-01 TCGA-AO-A0J7	TCGA-86-A0X1-01 TCGA-86-A0X1	TCGA-BH-A18M-01 TCGA-BH-A18M	TCGA-D8-A1JD-01 TCGA-D8-A1JD	TCGA-E2-A1LH-11 TCGA-E2-A1LH	TCGA-GM-A3NY-01 TCGA-GM-A3NY
TCGA-A2-A0ES-01 TCGA-A2-A0ES	TCGA-A8-A08F-01 TCGA-A8-A08F	TCGA-AO-A0J8-01 TCGA-AO-A0J8	TCGA-86-A0X4-01 TCGA-86-A0X4	TCGA-BH-A18M-11 TCGA-BH-A18M	TCGA-D8-A1JE-01 TCGA-D8-A1JE	TCGA-E2-A1LI-01 TCGA-E2-A1LI	TCGA-GM-A3XG-01 TCGA-GM-A3XG
TCGA-A2-A0ET-01 TCGA-A2-A0ET	TCGA-A8-A08G-01 TCGA-A8-A08G	TCGA-AO-A0J9-01 TCGA-AO-A0J9	TCGA-86-A0X5-01 TCGA-86-A0X5	TCGA-BH-A18N-01 TCGA-BH-A18N	TCGA-D8-A1JF-01 TCGA-D8-A1JF	TCGA-E2-A1LK-01 TCGA-E2-A1LK	TCGA-GM-A3XL-01 TCGA-GM-A3XL
TCGA-A2-A0EU-01 TCGA-A2-A0EU	TCGA-A8-A08H-01 TCGA-A8-A08H	TCGA-AO-A0JA-01 TCGA-AO-A0JA	TCGA-86-A0X7-01 TCGA-86-A0X7	TCGA-BH-A18N-11 TCGA-BH-A18N	TCGA-D8-A1JG-01 TCGA-D8-A1JG	TCGA-E2-A1LL-01 TCGA-E2-A1LL	TCGA-GM-A3XN-01 TCGA-GM-A3XN
TCGA-A2-A0EV-01 TCGA-A2-A0EV	TCGA-A8-A08I-01 TCGA-A8-A08I	TCGA-AO-ADJB-01 TCGA-AO-ADJB	TCGA-86-AIKC-01 TCGA-86-AIKC	TCGA-BH-A18P-01 TCGA-BH-A18P	TCGA-D8-A1JH-01 TCGA-D8-A1JH	TCGA-E2-AILS-01 TCGA-E2-AILS	TCGA-GM-A4EU-DI TCGA-GM-A4EU
TCGA-A2-A0EW-01 TCGA-A2-A0EW	TCGA-A8-A08J-01 TCGA-A8-A08J	TCGA-AO-ADJC-01 TCGA-AO-ADJC	TCGA-86-AIKF-01 TCGA-86-AIKF	TCGA-BH-A18P-11 TCGA-BH-A18P	TCGA-D8-A1JI-01 TCGA-D8-A1JI	TCGA-E2-AILS-11 TCGA-E2-AILS	TCGA-GM-A5PV-01 TCGA-GM-A5PV
TCGA-A2-A0EX-01 TCGA-A2-A0EX	TCGA-A8-A08L-01 TCGA-A8-A08L	TCGA-AO-ADJD-01 TCGA-AO-ADJD	TCGA-86-AIKI-01 TCGA-86-AIKI	TCGA-BH-A18Q-01 TCGA-BH-A18Q	TCGA-D8-A1JJ-01 TCGA-D8-A1JJ	TCGA-E2-A2P5-01 TCGA-E2-A2P5	TCGA-GM-A5PX-01 TCGA-GM-A5PX
TCGA-A2-A0EY-01 TCGA-A2-A0EY	TCGA-A8-A080-01 TCGA-A8-A080	TCGA-AO-ADJE-01 TCGA-AO-ADJE	TCGA-B6-A1KN-01 TCGA-B6-A1KN	TCGA-BH-A18Q-11 TCGA-BH-A18Q	TCGA-D8-A1JK-01 TCGA-D8-A1JK	TCGA-E2-A2P6-01 TCGA-E2-A2P6	TCGA-HN-A2NL-01 TCGA-HN-A2NL
TCGA-A2-A0ST-01 TCGA-A2-A0ST	TCGA-A8-A08P-01 TCGA-A8-A08P	TCGA-AO-ADJF-01 TCGA-AO-ADJF	TCGA-B6-A2IU-01 TCGA-B6-A2IU	TCGA-BH-A18R-01 TCGA-BH-A18R	TCGA-D8-A1JL-01 TCGA-D8-A1JL	TCGA-E2-A3DX-01 TCGA-E2-A3DX	TCGA-HN-A2OB-01 TCGA-HN-A2OB
TCGA-A2-A0SU-01 TCGA-A2-A0SU	TCGA-A8-A08R-01 TCGA-A8-A08R	TCGA-AO-ADJG-01 TCGA-AO-ADJG	TCGA-86-A32X-01 TCGA-86-A32X	TCGA-BH-A18R-11 TCGA-BH-A18R	TCGA-D8-A1JM-01 TCGA-D8-A1JM	TCGA-E2-A562-01 TCGA-E2-A562	TCGA-JL-A3YW-01 TCGA-JL-A3YW
TCGA-A2-A0SV-01 TCGA-A2-A0SV	TCGA-A8-A08S-01 TCGA-A8-A08S	TCGA-AO-ADJI-01 TCGA-AO-ADJI	TCGA-86-A400-01 TCGA-86-A400	TCGA-BH-A18S-01 TCGA-BH-A18S	TCGA-D8-A1JN-01 TCGA-D8-A1JN	TCGA-E2-A570-01 TCGA-E2-A570	TCGA-JL-A3YX-01 TCGA-JL-A3YX
TCGA-A2-A0SW-01 TCGA-A2-A0SW TCGA-A2-A0SX-01 TCGA-A2-A0SX TCGA-A2-A0SY-01 TCGA-A2-A0SY	TCGA-A8-A081-01 TCGA-A8-A081 TCGA-A8-A08X-01 TCGA-A8-A08X TCGA-A8-A08Z-01 TCGA-A8-A08Z	TCGA-AO-ADJI-01 TCGA-AO-ADJI TCGA-AO-ADJI-01 TCGA-AO-ADJI TCGA-AO-ADJM-01 TCGA-AO-ADJM	TCGA-86-A402-01 TCGA-86-A402 TCGA-86-A402-01 TCGA-86-A402 TCGA-86-A408-01 TCGA-86-A408	TCGA-BH-A183-11 TCGA-BH-A183 TCGA-BH-A18T-01 TCGA-BH-A18T TCGA-BH-A18U-01 TCGA-BH-A18U	TCGA-D8-A1JF-01 TCGA-D8-A1JF TCGA-D8-A1JS-01 TCGA-D8-A1JS TCGA-D8-A1JT-01 TCGA-D8-A1JT	TCGA-E2-A572-01 TCGA-E2-A572 TCGA-E2-A573-01 TCGA-E2-A573 TCGA-E2-A574-01 TCGA-E2-A574	TCGA-LD-A74U-01 TCGA-LD-A74U TCGA-LD-A7W5-01 TCGA-LD-A7W5
TCGA-A2-A0T0-01 TCGA-A2-A0T0	TCGA-A8-A090-01 TCGA-A8-A090	TCGA-AO-A124-01 TCGA-AO-A124	TCGA-86-A409-01 TCGA-86-A409	TCGA-BH-A18U-11 TCGA-BH-A18U	TCGA-D8-A1JU-01 TCGA-D8-A1JU	TCGA-E2-A576-01 TCGA-E2-A576	TCGA-LD-A7W6-01 TCGA-LD-A7W6
TCGA-A2-A0T1-01 TCGA-A2-A0T1	TCGA-A8-A091-01 TCGA-A8-A091	TCGA-AO-A125-01 TCGA-AO-A125	TCGA-86-A408-01 TCGA-86-A408	TCGA-BH-A18V-01 TCGA-BH-A18V	TCGA-D8-A1X5-01 TCGA-D8-A1X5	TCGA-E2-A9RU-01 TCGA-E2-A9RU	TCGA-LD-A9QF-01 TCGA-LD-A9QF
TCGA-A2-A0T2-01 TCGA-A2-A0T2 TCGA-A2-A0T3-01 TCGA-A2-A0T3 TCGA-A2-A0T4-01 TCGA-A2-A0T4	TCGA-A8-A092-01 TCGA-A8-A092 TCGA-A8-A093-01 TCGA-A8-A093	TCGA-AO-A126-01 TCGA-AO-A126 TCGA-AO-A128-01 TCGA-AO-A128 TCGA-AO-A128-01 TCGA-AO-A128	TCGA-B6-A40C-01 TCGA-B6-A40C TCGA-BH-A0AU-01 TCGA-BH-A0AU	TCGA-BH-A18V-06 TCGA-BH-A18V TCGA-BH-A18V-11 TCGA-BH-A18V TCGA-BH-A18N-01 TCGA-BH-A15N	TCGA-D8-A1X6-01 TCGA-D8-A1X6 TCGA-D8-A1X7-01 TCGA-D8-A1X7 TCGA-D8-A1X8-01 TCGA-D8-A1X8	TCGA-E9-A1N3-01 TCGA-E9-A1N3 TCGA-E9-A1N4-01 TCGA-E9-A1N4 TCGA-E9-A1N4-11 TCGA-E9-A1N4	TCGA-LL-A440-01 TCGA-LL-A440 TCGA-LL-A441-01 TCGA-LL-A441 TCGA-LL-A441-01 TCGA-LL-A441
TCGA-A2-A0T5-01 TCGA-A2-A0T5	TCGA-A8-A095-01 TCGA-A8-A095	TCGA-AO-A12A-01 TCGA-AO-A12A	TCGA-BH-A0AV-01 TCGA-BH-A0AV	TCGA-BH-A1EN-11 TCGA-BH-A1EN	TCGA-D8-A1X9-01 TCGA-D8-A1X9	TCGA-E9-A1N5-01 TCGA-E9-A1N5	TCGA-LL-A50Y-01 TCGA-LL-A50Y
TCGA-A2-A0T6-01 TCGA-A2-A0T6	TCGA-A8-A096-01 TCGA-A8-A096	TCGA-AO-A12B-01 TCGA-AO-A12B	TCGA-BH-A0AW-01 TCGA-BH-A0AW	TCGA-BH-A1EO-01 TCGA-BH-A1EO	TCGA-D8-A1X9-01 TCGA-D8-A1XA	TCGA-E9-A1N5-11 TCGA-E9-A1N5	TCGA-LL-A5YL-01 TCGA-LL-A5YL
TCGA-A2-A0T7-01 TCGA-A2-A0T7	TCGA-A8-A097-01 TCGA-A8-A097	TCGA-AO-A12C-01 TCGA-AO-A12C	TCGA-BH-A0AY-01 TCGA-BH-A0AY	TCGA-BH-A1EO-11 TCGA-BH-A1EO	TCGA-D8-A1XB-01 TCGA-D8-A1XB	TCGA-E9-A1N6-01 TCGA-E9-A1N6	TCGA-LL-A5YM-01 TCGA-LL-A5YM
TCGA-A2-A0YC-01 TCGA-A2-A0YC	TCGA-A8-A099-01 TCGA-A8-A099	TCGA-AO-A12D-01 TCGA-AO-A12D	TCGA-BH-A0AY-11 TCGA-BH-A0AY	TCGA-BH-A1ES-01 TCGA-BH-A1ES	TCGA-D8-A1XC-01 TCGA-D8-A1XC	TCGA-E9-A1N6-11 TCGA-E9-A1N6	TCGA-LL-A5YN-01 TCGA-LL-A5YN
TCGA-A2-A0YD-01 TCGA-A2-A0YD TCGA-A2-A0YE-01 TCGA-A2-A0YE TCGA-A2-A0YE-01 TCGA-A2-A0YE	TCGA-A8-A09A-01 TCGA-A8-A09A TCGA-A8-A09B-01 TCGA-A8-A09B	TCGA-AO-A12E-01 TCGA-AO-A12E TCGA-AO-A12F-01 TCGA-AO-A12F TCGA-AO-A12F-01 TCGA-AO-A12F	TCGA-BH-A0AZ-01 TCGA-BH-A0AZ TCGA-BH-A0AZ-11 TCGA-BH-A0AZ	TCGA-BH-A1ES-06 TCGA-BH-A1ES TCGA-BH-A1ET-01 TCGA-BH-A1ET TCGA-BH-A1ET-11 TCGA-BH-A1ET	TCGA-D8-A1XD-01 TCGA-D8-A1XD TCGA-D8-A1XF-01 TCGA-D8-A1XF TCGA-D8-A1XC-01 TCGA-D8-A1XC	TCGA-E9-A1N8-01 TCGA-E9-A1N8 TCGA-E9-A1N9-01 TCGA-E9-A1N9 TCGA-E9-A1N0-11 TCGA-E9-A1N9	TCGA-LL-ASYO-01 TCGA-LL-ASYO TCGA-LL-ASYP-01 TCGA-LL-ASYP
TCGA-A2-A0YF-01 TCGA-A2-A0YF	TCGA-A8-A09C-01 TCGA-A8-A09D	TCGA-AO-A12H-01 TCGA-AO-A12H	TCGA-BH-A080-01 TCGA-BH-A080	TCGA-BH-A1EU-11 TCGA-BH-A1EU	TCGA-D8-A1XI-01 TCGA-D8-A1XG	TCGA-E9-AINO-11 TCGA-E9-AINO	TCGA-LL-A6FQ-01 TCGA-LL-A6FQ
TCGA-A2-A0YG-01 TCGA-A2-A0YG	TCGA-A8-A09D-01 TCGA-A8-A09D	TCGA-AO-A12H-01 TCGA-AO-A12H	TCGA-BH-A081-01 TCGA-BH-A081	TCGA-BH-A1EU-11 TCGA-BH-A1EU	TCGA-D8-A1XJ-01 TCGA-D8-A1XJ	TCGA-E9-AINA-01 TCGA-E9-AINA	TCGA-LL-A6FQ-01 TCGA-LL-A6FQ
TCGA-A2-A0YH-01 TCGA-A2-A0YH	TCGA-A8-A09E-01 TCGA-A8-A09E	TCGA-AO-A1KO-01 TCGA-AO-A1KO	TCGA-BH-A082-01 TCGA-BH-A082	TCGA-BH-A1EU-11 TCGA-BH-A1EU	TCGA-D8-A1XK-01 TCGA-D8-A1XK	TCGA-E9-AINA-11 TCGA-E9-AINA	TCGA-LL-A6FR-01 TCGA-LL-A6FR
TCGA-A2-A0YI-01 TCGA-A2-A0YI	TCGA-A8-A09G-01 TCGA-A8-A09G	TCGA-AO-A1KP-01 TCGA-AO-A1KP	TCGA-BH-A0B2-11 TCGA-BH-A0B2	TCGA-BH-A1EV-01 TCGA-BH-A1EV	TCGA-D8-A1XL-01 TCGA-D8-A1XL	TCGA-E9-A1NC-01 TCGA-E9-A1NC	TCGA-LL-A73Y-01 TCGA-LL-A73Y
TCGA-A2-A0YJ-01 TCGA-A2-A0YJ	TCGA-A8-A09I-01 TCGA-A8-A09I	TCGA-AO-A1KQ-01 TCGA-AO-A1KQ	TCGA-BH-A0B3-01 TCGA-BH-A0B3	TCGA-BH-A1EV-11 TCGA-BH-A1EV	TCGA-D8-A1XM-01 TCGA-D8-A1XM	TCGA-E9-A1ND-01 TCGA-E9-A1ND	TCGA-LL-A73Z-01 TCGA-LL-A73Z
TCGA-A2-A0YK-01 TCGA-A2-A0YK	TCGA-A8-A09K-01 TCGA-A8-A09K	TCGA-AO-A1KR-01 TCGA-AO-A1KR	TCGA-BH-A0B3-11 TCGA-BH-A0B3	TCGA-BH-A1EW-01 TCGA-BH-A1EW	TCGA-D8-A1XO-01 TCGA-D8-A1XO	TCGA-E9-A1ND-11 TCGA-E9-A1ND	TCGA-LL-A740-01 TCGA-LL-A740
TCGA-A2-A0YL-01 TCGA-A2-A0YL	TCGA-A8-A09M-01 TCGA-A8-A09M	TCGA-AO-A1KS-01 TCGA-AO-A1KS	TCGA-BH-A0B4-01 TCGA-BH-A0B4	TCGA-BH-A1EW-11 TCGA-BH-A1EW	TCGA-D8-A1XQ-01 TCGA-D8-A1XQ	TCGA-E9-A1NE-01 TCGA-E9-A1NE	TCGA-LL-A75Z-01 TCGA-LL-A75Z
TCGA-A2-A0YM-01 TCGA-A2-A0YM	TCGA-A8-A09N-01 TCGA-A8-A09N	TCGA-AO-A1KT-01 TCGA-AO-A1KT	TCGA-BH-A0B5-01 TCGA-BH-A0B5	TCGA-BH-A1EX-01 TCGA-BH-A1EX	TCGA-D8-A1XR-01 TCGA-D8-A1XR	TCGA-E9-A1NE-01 TCGA-E9-A1NE	TCGA-LL-A7T0-01 TCGA-LL-A7T0
TCGA-A2-A0YT-01 TCGA-A2-A0YT	TCGA-A8-A09Q-01 TCGA-A8-A09Q	TCGA-AQ-AD4H-01 TCGA-AQ-A04H	TCGA-BH-A0B5-11 TCGA-BH-A0B5	TCGA-BH-A1EY-01 TCGA-BH-A1EY	TCGA-D8-A1XS-01 TCGA-D8-A1XS	TCGA-E9-A1NF-11 TCGA-E9-A1NF	TCGA-LL-A8F5-01 TCGA-LL-A8F5
TCGA-A2-A1FV-01 TCGA-A2-A1FV	TCGA-A8-A09R-01 TCGA-A8-A09R	TCGA-AQ-AD4J-01 TCGA-AQ-A04J	TCGA-BH-A0B6-01 TCGA-BH-A0B6	TCGA-BH-A1F0-01 TCGA-BH-A1F0	TCGA-D8-A1XT-01 TCGA-D8-A1XT	TCGA-E9-A1NG-01 TCGA-E9-A1NG	TCGA-LL-A9Q3-01 TCGA-LL-A9Q3
TCGA-A2-A1FW-01 TCGA-A2-A1FW	TCGA-A8-A09T-01 TCGA-A8-A09T	TCGA-AQ-A04L-01 TCGA-AQ-A04L	TCGA-BH-A0B7-01 TCGA-BH-A0B7	TCGA-BH-A1F0-11 TCGA-BH-A1F0	TCGA-D8-A1XU-01 TCGA-D8-A1XU	TCGA-E9-A1NG-11 TCGA-E9-A1NG	TCGA-LQ-A4E4-01 TCGA-LQ-A4E4
TCGA-A2-A1FX-01 TCGA-A2-A1FX	TCGA-A8-A09V-01 TCGA-A8-A09V	TCGA-AQ-A0Y5-01 TCGA-AQ-A0Y5	TCGA-BH-A0B7-11 TCGA-BH-A0B7	TCGA-BH-A1F2-01 TCGA-BH-A1F2	TCGA-D8-A1XV-01 TCGA-D8-A1XV	TCGA-E9-A1NH-01 TCGA-E9-A1NH	TCGA-MS-A51U-01 TCGA-MS-A51U
TCGA-A2-A1F2-01 TCGA-A2-A1F2	TCGA-A8-A09W-01 TCGA-A8-A09W	TCGA-AQ-A1H2-01 TCGA-AQ-A1H2	TCGA-BH-A088-01 TCGA-BH-A088	TCGA-BH-A1F2-11 TCGA-BH-A1F2	TCGA-D8-A1XW-01 TCGA-D8-A1XW	TCGA-E9-A1NI-01 TCGA-E9-A1NI	TCGA-0K-ASQ2-01 TCGA-0K-ASQ2
TCGA-A2-A1G0-01 TCGA-A2-A1G0	TCGA-A8-A09X-01 TCGA-A8-A09X	TCGA-AQ-A1H3-01 TCGA-AQ-A1H3	TCGA-BH-A088-11 TCGA-BH-A088	TCGA-BH-A1F5-01 TCGA-BH-A1F5	TCGA-D8-A1XY-01 TCGA-D8-A1XY	TCGA-E9-A1QZ-01 TCGA-E9-A1QZ	TCGA-0L-ASD6-01 TCGA-0L-ASD6
TCGA-A2-A1G1-01 TCGA-A2-A1G1	TCGA-A8-A097-01 TCGA-A8-A097	TCGA-AQ-A54N-01 TCGA-AQ-A54N	TCGA-BH-A089-01 TCGA-BH-A089	TCGA-BH-A1F6-01 TCGA-BH-A1F6	TCGA-D8-A1X7-01 TCGA-D8-A1X7	TCGA-E9-A1R0-01 TCGA-E9-A1R0	TCGA-0L-ASD7-01 TCGA-0L-ASD7
TCGA-A2-A1G4-01 TCGA-A2-A1G4	TCGA-A8-A0A1-01 TCGA-A8-A0A1	TCGA-AQ-A540-01 TCGA-AQ-A540	TCGA-BH-A0BA-01 TCGA-BH-A0BA	TCGA-BH-A1F6-11 TCGA-BH-A1F6	TCGA-D8-A1Y0-01 TCGA-D8-A1Y0	TCGA-E9-A1R2-01 TCGA-E9-A1R2	TCGA-OL-ASD8-01 TCGA-OL-ASD8
TCGA-A2-A1G6-01 TCGA-A2-A1G6	TCGA-A8-A0A2-01 TCGA-A8-A0A2	TCGA-AQ-A7U7-01 TCGA-AQ-A7U7	TCGA-BH-A0BA-11 TCGA-BH-A0BA	TCGA-BH-A1F8-01 TCGA-BH-A1F8	TCGA-D8-A1Y1-01 TCGA-D8-A1Y1	TCGA-E9-A1R3-01 TCGA-E9-A1R3	TCGA-OL-ASDA-01 TCGA-OL-ASDA
TCGA-A2-A259-01 TCGA-A2-A259 TCGA-A2-A25A-01 TCGA-A2-A25A TCGA-A2-A25B-01 TCGA-A2-A25B	TCGA-A8-A0A4-01 TCGA-A8-A0A4 TCGA-A8-A0A6-01 TCGA-A8-A0A6	TCGA-AR-A0TP-01 TCGA-AR-A0TP TCGA-AR-A0TQ-01 TCGA-AR-A0TQ TCGA-AR-A0TQ-01 TCGA-AR-A0TQ	TCGA-BH-A0BC-01 TCGA-BH-A0BC TCGA-BH-A0BC-11 TCGA-BH-A0BC	TCGA-BH-A1F8-11 TCGA-BH-A1F8 TCGA-BH-A1FB-01 TCGA-BH-A1FB TCGA-BH-A1FB-11 TCGA-BH-A1FB	TCGA-D8-A1Y2-01 TCGA-D8-A1Y2 TCGA-D8-A1Y3-01 TCGA-D8-A1Y3 TCGA-D8-A375-01 TCGA-D8-A375	TCGA-E9-A1R4-01 TCGA-E9-A1R4 TCGA-E9-A1R5-01 TCGA-E9-A1R5 TCGA-E9-A1R5-01 TCGA-E9-A1R5	TCGA-OL-ASRU-01 TCGA-OL-ASRU TCGA-OL-ASRV-01 TCGA-OL-ASRV
TCGA-A2-A25B-01 TCGA-A2-A25B TCGA-A2-A25C-01 TCGA-A2-A25C TCGA-A2-A25D-01 TCGA-A2-A25D	TCGA-A8-A0A9-01 TCGA-A8-A0A9 TCGA-A8-A0A9-01 TCGA-A8-A0A9 TCGA-A8-A0AB-01 TCGA-A8-A0AB	TCGA-AR-A0TS-01 TCGA-AR-A0TS TCGA-AR-A0TS-01 TCGA-AR-A0TS	TCGA-BH-A0BD-01 TCGA-BH-A0BD TCGA-BH-A0BF-01 TCGA-BH-A0BF TCGA-BH-A0BG-01 TCGA-BH-A0BG	TCGA-BH-A1FC-11 TCGA-BH-A1FC TCGA-BH-A1FC-11 TCGA-BH-A1FC TCGA-BH-A1FC-11 TCGA-BH-A1FC	TCGA-08-A27E-01 TCGA-08-A27E TCGA-08-A27F-01 TCGA-08-A27F TCGA-08-A27G-01 TCGA-08-A27G	TCGA-E9-A1R0-01 TCGA-E9-A1R0 TCGA-E9-A1R7-01 TCGA-E9-A1R7 TCGA-E9-A1R7-11 TCGA-E9-A1R7	TCGA-OL-ASRX-01 TCGA-OL-ASRX TCGA-OL-ASRX-01 TCGA-OL-ASRX TCGA-OL-ASRY-01 TCGA-OL-ASRY
TCGA-A2-A25E-01 TCGA-A2-A25E	TCGA-A8-A0AD-01 TCGA-A8-A0AD	TCGA-AR-A0TU-01 TCGA-AR-A0TU	TCGA-BH-A0BJ-01 TCGA-BH-A0BJ	TCGA-BH-A1FD-01 TCGA-BH-A1FD	TCGA-D8-A27H-01 TCGA-D8-A27H	TCGA-E9-A1RA-01 TCGA-E9-A1RA	TCGA-OL-ASRZ-01 TCGA-OL-ASRZ
TCGA-A2-A25F-01 TCGA-A2-A25F	TCGA-AC-A23C-01 TCGA-AC-A23C	TCGA-AR-A0TV-01 TCGA-AR-A0TV	TCGA-BH-A0BJ-11 TCGA-BH-A0BJ	TCGA-BH-A1FD-11 TCGA-BH-A1FD	TCGA-D8-A27I-01 TCGA-D8-A27I	TCGA-E9-A1RB-01 TCGA-E9-A1RB	TCGA-OL-ASS0-01 TCGA-OL-ASS0
TCGA-A2-A3KC-01 TCGA-A2-A3KC TCGA-A2-A3KD-01 TCGA-A2-A3KD	TCGA-AC-A23E-01 TCGA-AC-A23E TCGA-AC-A23G-01 TCGA-AC-A23G	TCGA-AR-A0TW-01 TCGA-AR-A0TW TCGA-AR-A0TX-01 TCGA-AR-A0TX TCGA-AR-A0TX-01 TCGA-AR-A0TX	TCGA-BH-A0BL-01 TCGA-BH-A0BL TCGA-BH-A0BM-01 TCGA-BH-A0BM	TCGA-BH-A1FE-01 TCGA-BH-A1FE TCGA-BH-A1FE-06 TCGA-BH-A1FE	TCGA-D8-A27K-01 TCGA-D8-A27K TCGA-D8-A27L-01 TCGA-D8-A27L TCGA-D8-A27L-01 TCGA-D8-A27A	TCGA-E9-A1RB-11 TCGA-E9-A1RB TCGA-E9-A1RC-01 TCGA-E9-A1RC TCGA-E9-A1RC-11 TCGA-E9-A1RC	TCGA-0L-A66H-01 TCGA-0L-A66H TCGA-0L-A66I-01 TCGA-0L-A66I
TCGA-A2-A3XT-01 TCGA-A2-A3XT	TCGA-AC-A23H-11 TCGA-AC-A23H	TCGA-AR-A0TZ-01 TCGA-AR-A0TZ	TCGA-BH-A0BO-01 TCGA-BH-A0BO	TCGA-BH-A1FG-01 TCGA-BH-A1FG	TCGA-D8-A27N-01 TCGA-D8-A27N	TCGA-E9-A1RD-01 TCGA-E9-A1RD	TCGA-OL-A66K-01 TCGA-OL-A66K
TCGA-A2-A3XU-01 TCGA-A2-A3XU	TCGA-AC-A288-01 TCGA-AC-A288	TCGA-AR-A0U0-01 TCGA-AR-A0U0	TCGA-BH-A0BP-01 TCGA-BH-A0BP	TCGA-BH-A1FG-11 TCGA-BH-A1FG	TCGA-D8-A27P-01 TCGA-D8-A27P	TCGA-E9-A1RD-11 TCGA-E9-A1RD	TCGA-OL-A66L-01 TCGA-OL-A66L
TCGA-A2-A3XV-01 TCGA-A2-A3XV	TCGA-AC-A2BK-01 TCGA-AC-A2BK	TCGA-AR-A0U1-01 TCGA-AR-A0U1	TCGA-BH-A0BQ-01 TCGA-BH-A0BQ	TCGA-BH-A1FH-01 TCGA-BH-A1FH	TCGA-D8-A27R-01 TCGA-D8-A27R	TCGA-E9-A1RE-01 TCGA-E9-A1RE	TCGA-OL-A66N-01 TCGA-OL-A66N
TCGA-A2-A3XW-01 TCGA-A2-A3XW	TCGA-AC-A2BM-01 TCGA-AC-A2BM	TCGA-AR-A0U2-01 TCGA-AR-A0U2	TCGA-BH-A0BQ-11 TCGA-BH-A0BQ	TCGA-BH-A1FH-11 TCGA-BH-A1FH	TCGA-D8-A27T-01 TCGA-D8-A27T	TCGA-E9-A1RF-01 TCGA-E9-A1RF	TCGA-OL-A66O-01 TCGA-OL-A66O
TCGA-A2-A3XX-01 TCGA-A2-A3XX	TCGA-AC-A2FB-01 TCGA-AC-A2FB	TCGA-AR-A0U3-01 TCGA-AR-A0U3	TCGA-BH-A0BR-01 TCGA-BH-A0BR	TCGA-BH-A1FJ-01 TCGA-BH-A1FJ	TCGA-D8-A27V-01 TCGA-D8-A27V	TCGA-E9-A1RF-11 TCGA-E9-A1RF	TCGA-OL-A669-01 TCGA-OL-A669
TCGA-A2-A3XY-01 TCGA-A2-A3XY	TCGA-AC-A2FB-11 TCGA-AC-A2FB	TCGA-AR-A0U4-01 TCGA-AR-A0U4	TCGA-BH-A0BS-01 TCGA-BH-A0BS	TCGA-BH-A1FJ-11 TCGA-BH-A1FJ	TCGA-D8-A27W-01 TCGA-D8-A27W	TCGA-E9-A1RG-01 TCGA-E9-A1RG	TCGA-OL-A690-01 TCGA-OL-A690
TCGA-A2-A3X7-01 TCGA-A2-A3X7	TCGA-AC-A2FF-01 TCGA-AC-A2FF	TCGA-AR-A1AH-01 TCGA-AR-A1AH	TCGA-BH-A0BS-11 TCGA-BH-A0BS	TCGA-BH-A1FI-01 TCGA-BH-A1FI	TCGA-D8-A375-01 TCGA-D8-A375	TCGA-E9-A1RH-01 TCGA-E9-A1RH	TCGA-OL-A698-01 TCGA-OL-A698
TCGA-A2-A3Y0-01 TCGA-A2-A3Y0	TCGA-AC-A2FF-01 TCGA-AC-A2FF	TCGA-AR-A1AI-01 TCGA-AR-A1AI	TCGA-BH-A0BT-01 TCGA-BH-A0BT	TCGA-BH-A1FM-01 TCGA-BH-A1FM	TCGA-D8-A3Z6-01 TCGA-D8-A3Z6	TCGA-E9-A1RH-11 TCGA-E9-A1RH	TCGA-OL-A97C-01 TCGA-OL-A97C
TCGA-A2-A4RW-01 TCGA-A2-A4RW	TCGA-AC-A2FF-11 TCGA-AC-A2FF	TCGA-AR-A1AI-01 TCGA-AR-A1AJ	TCGA-BH-A0BT-11 TCGA-BH-A0BT	TCGA-BH-A1FM-11 TCGA-BH-A1FM	TCGA-D8-A4Z1-01 TCGA-D8-A4Z1	TCGA-E9-A1RI-01 TCGA-E9-A1RI	TCGA-PE-A5DC-01 TCGA-PE-A5DC
TCGA-A2-A4RX-01 TCGA-A2-A4RX	TCGA-AC-A2FG-01 TCGA-AC-A2FG	TCGA-AR-A1AK-01 TCGA-AR-A1AK	TCGA-BH-A0BV-01 TCGA-BH-A0BV	TCGA-BH-A1FN-01 TCGA-BH-A1FN	TCGA-D8-A73U-01 TCGA-D8-A73U	TCGA-E9-A1RI-11 TCGA-E9-A1RI	TCGA-PE-ASDD-01 TCGA-PE-ASDD
TCGA-A2-A4RY-01 TCGA-A2-A4RY	TCGA-AC-A2FK-01 TCGA-AC-A2FK	TCGA-AR-A1AL-01 TCGA-AR-A1AL	TCGA-BH-A0BV-11 TCGA-BH-A0BV	TCGA-BH-A1FN-11 TCGA-BH-A1FN	TCGA-D8-A73W-01 TCGA-D8-A73W	TCGA-E9-A226-01 TCGA-E9-A226	TCGA-PE-ASDE-01 TCGA-PE-ASDE
TCGA-A2-A4S0-01 TCGA-A2-A4S0	TCGA-AC-A2FM-01 TCGA-AC-A2FM	TCGA-AR-AIAM-01 TCGA-AR-AIAM	TCGA-BH-A0BW-01 TCGA-BH-A0BW	TCGA-BH-A1FR-01 TCGA-BH-A1FR	TCGA-D8-A73X-01 TCGA-D8-A73X	TCGA-E9-A227-01 TCGA-E9-A227	TCGA-PL-A8LV-01 TCGA-PL-A8LV
TCGA-A2-A4S1-01 TCGA-A2-A4S1	TCGA-AC-A2FM-11 TCGA-AC-A2FM	TCGA-AR-AIAN-01 TCGA-AR-AIAN	TCGA-BH-A0BW-11 TCGA-BH-A0BW	TCGA-BH-A1FR-11 TCGA-BH-A1FR	TCGA-E2-A105-01 TCGA-E2-A105	TCGA-E9-A228-01 TCGA-E9-A228	TCGA-PL-A8LX-01 TCGA-PL-A8LX
TCGA-A2-A4S2-01 TCGA-A2-A4S2	TCGA-AC-A2FO-01 TCGA-AC-A2FO	TCGA-AR-AIAO-01 TCGA-AR-AIAO	TCGA-BH-A0B7-01 TCGA-BH-A0B7	TCGA-BH-A1FII-01 TCGA-BH-A1FII	TCGA-F2-A106-01 TCGA-F2-A106	TCGA-F9-A229-01 TCGA-F9-A229	TCGA-PL-A8LY-01 TCGA-PL-A8LY
TCGA-A2-A4S3-01 TCGA-A2-A4S3	TCGA-AC-A2QH-01 TCGA-AC-A2QH	TCGA-AR-A1AP-01 TCGA-AR-A1AP	TCGA-BH-A0BZ-11 TCGA-BH-A0BZ	TCGA-BH-A1FU-11 TCGA-BH-A1FU	TCGA-E2-A107-01 TCGA-E2-A107	TCGA-E9-A22A-01 TCGA-E9-A22A	TCGA-PL-A8LZ-01 TCGA-PL-A8LZ
TCGA-A7-A0CD-01 TCGA-A7-A0CD	TCGA-AC-A2QI-01 TCGA-AC-A2QI	TCGA-AR-A1AQ-01 TCGA-AR-A1AQ	TCGA-BH-A0CO-01 TCGA-BH-A0CO	TCGA-BH-A201-01 TCGA-BH-A201	TCGA-E2-A108-01 TCGA-E2-A108	TCGA-E9-A22B-01 TCGA-E9-A22B	TCGA-S3-A6ZF-01 TCGA-S3-A6ZF
TCGA-A7-A0CE-01 TCGA-A7-A0CE TCGA-A7-A0CE-11 TCGA-A7-A0CE	TCGA-AC-A2QJ-01 TCGA-AC-A2QJ TCGA-AC-A3BB-01 TCGA-AC-A3BB	TCGA-AR-A1AR-01 TCGA-AR-A1AR TCGA-AR-A1AS-01 TCGA-AR-A1AS	TCGA-BH-A0C0-11 TCGA-BH-A0C0 TCGA-BH-A0C1-01 TCGA-BH-A0C1 TCGA-BH-A0C2-01 TCGA-BH-A0C1	TCGA-BH-A202-01 TCGA-BH-A202 TCGA-BH-A203-01 TCGA-BH-A203	TCGA-E2-A109-01 TCGA-E2-A109 TCGA-E2-A10A-01 TCGA-E2-A10A	TCGA-E9-A22D-01 TCGA-E9-A22D TCGA-E9-A22E-01 TCGA-E9-A22E	TCGA-S3-A62G-01 TCGA-S3-A62G TCGA-S3-A62H-01 TCGA-S3-A62H
TCGA-A7-A0CH-01 TCGA-A7-A0CH TCGA-A7-A0CH-11 TCGA-A7-A0CH	TCGA-AC-A3HN-01 TCGA-AC-A3HN TCGA-AC-A3OD-01 TCGA-AC-A3OD	TCGA-AR-A1AU-01 TCGA-AR-A1AU TCGA-AR-A1AU-01 TCGA-AR-A1AU TCGA-AR-A1AV-01 TCGA-AR-A1AV	TCGA-BH-A0C3-11 TCGA-BH-A0C3 TCGA-BH-A0C7-01 TCGA-BH-A0C7	TCGA-BH-A204-01 TCGA-BH-A204 TCGA-BH-A204-01 TCGA-BH-A204	TCGA-E2-A10C-01 TCGA-E2-A10C TCGA-E2-A10C-01 TCGA-E2-A10C TCGA-E2-A10E-01 TCGA-E2-A10E	TCGA-E9-A22H-01 TCGA-E9-A22H TCGA-E9-A243-01 TCGA-E9-A243	TCGA-S3-AA10-01 TCGA-S3-AA10 TCGA-S3-AA11-01 TCGA-S3-AA11
TCGA-A7-A0CJ-01 TCGA-A7-A0CJ	TCGA-AC-A3QP-01 TCGA-AC-A3QP	TCGA-AR-A1AW-01 TCGA-AR-A1AW	TCGA-BH-A0DD-01 TCGA-BH-A0DD	TCGA-BH-A208-01 TCGA-BH-A208	TCGA-E2-A10F-01 TCGA-E2-A10F	TCGA-E9-A244-01 TCGA-E9-A244	TCGA-S3-AA12-01 TCGA-S3-AA12
TCGA-A7-A0D9-01 TCGA-A7-A0D9	TCGA-AC-A3QQ-01 TCGA-AC-A3QQ	TCGA-AR-A1AX-01 TCGA-AR-A1AX	TCGA-BH-A0DD-11 TCGA-BH-A0DD	TCGA-BH-A208-11 TCGA-BH-A208	TCGA-E2-A14N-01 TCGA-E2-A14N	TCGA-E9-A245-01 TCGA-E9-A245	TCGA-S3-AA14-01 TCGA-S3-AA14
TCGA-A7-A0D9-11 TCGA-A7-A0D9	TCGA-AC-A31M-01 TCGA-AC-A3TM	TCGA-AR-A1AY-01 TCGA-AR-A1AY	TCGA-BH-AUDE-01 TCGA-BH-A0DE	TCGA-BH-A209-01 TCGA-BH-A209	TCGA-E2-A140-01 TCGA-E2-A140	TCGA-E9-A247-01 TCGA-E9-A247	ТСGA-S3-AA15-01 ТСGA-S3-AA15
TCGA-A7-A0DA-01 TCGA-A7-A0DA	TCGA-AC-A3TN-01 TCGA-AC-A3TN	TCGA-AR-A24H-01 TCGA-AR-A24H	TCGA-BH-A0DG-01 TCGA-BH-A0DG	TCGA-BH-A209-11 TCGA-BH-A209	TCGA-E2-A14P-01 TCGA-E2-A14P	TCGA-E9-A248-01 TCGA-E9-A248	ТСGA-S3-AA17-01 ТСGA-S3-AA17
TCGA-A7-A0D8-01 TCGA-A7-A0D8	TCGA-AC-A3W5-01 TCGA-AC-A3W5	TCGA-AR-A24K-01 TCGA-AR-A24K	TCGA-BH-A0DG-11 TCGA-BH-A0DC	TCGA-BH-A280-01 TCGA-BH-A290	TCGA-E2-A140-01 TCGA-E2-A140	TCGA-E9-A249-01 TCGA-E9-A240	ТСGA-III-AA76-01 ТССА-III-AA72
TCGA-A7-A0DB-11 TCGA-A7-A0DB	TCGA-AC-A3W6-01 TCGA-AC-A3W6	TCGA-AR-A24L-01 TCGA-AR-A24L	TCGA-BH-A0DH-01 TCGA-BH-A0DH	TCGA-BH-A28Q-01 TCGA-BH-A28Q	TCGA-E2-A14R-01 TCGA-E2-A14R	TCGA-E9-A24A-01 TCGA-E9-A24A	TCGA-UU-A93S-01 TCGA-UU-A93S
TCGA-A7-A0DC-01 TCGA-A7-A0DC	TCGA-AC-A3W7-01 TCGA-AC-A3W7	TCGA-AR-A24M-01 TCGA-AR-A24M	TCGA-BH-A0DH-11 TCGA-BH-A0DH	TCGA-BH-A218-01 TCGA-BH-A218	TCGA-E2-A14S-01 TCGA-E2-A14S	TCGA-E9-A295-01 TCGA-E9-A295	TCGA-V7-A7HQ-01 TCGA-V7-A7HQ
TCGA-A7-A0DC-11 TCGA-A7-A0DC	TCGA-AC-A3YI-01 TCGA-AC-A3YI	TCGA-AR-A24N-01 TCGA-AR-A24N	TCGA-BH-A0DI-01 TCGA-BH-A0DI	TCGA-BH-A42T-01 TCGA-BH-A42T	TCGA-E2-A14T-01 TCGA-E2-A14T	TCGA-E9-A2J5-01 TCGA-E9-A2J5	TCGA-W8-A86G-01 TCGA-W8-A86G
TCGA-A7-A13D-01 TCGA-A7-A13D	TCGA-AC-A3YI-01 TCGA-AC-A3YI	TCGA-AR-A24O-01 TCGA-AR-A24O	TCGA-BH-A0DK-01 TCGA-BH-A0DK	TCGA-BH-A42U-01 TCGA-BH-A42U	TCGA-E2-A14U-01 TCGA-E2-A14U	TCGA-E9-A2JT-01 TCGA-E9-A2JT	TCGA-WT-AB41-01 TCGA-WT-AB41
ТСGA-A7-A13E-U1 ICGA-A7-A13E	TCGA-AC-AA2E-01 TCGA-AC-A42E	ТСGA-AR-A24P-U1 ICGA-AR-A24P	TCGA-BH-AUDK-11 ICGA-BH-AODK	ТСGA-BH-A42V-01 ТСGA-BH-A42V	TCGA-E2-A14V-01 TCGA-E2-A14V	TCGA-E9-A3HO-01 TCGA-E9-A3HO	TCGA-XX-A899-01 TCGA-XX-A899
TCGA-A7-A13E-11 TCGA-A7-A13E	TCGA-AC-A5EH-01 TCGA-AC-A5EH	TCGA-AR-A24Q-01 TCGA-AR-A24Q	TCGA-BH-AODL-01 TCGA-BH-AODL	TCGA-BH-A5IZ-01 ТСGA-BH-A5IZ	TCGA-E2-A14W-01 TCGA-E2-A14W	TCGA-E9-A3Q9-01 TCGA-E9-A3Q9	TCGA-XX-A899-01 TCGA-XX-A899
TCGA-A7-A13F-01 TCGA-Δ7-Δ13F	TCGA-AC-A5XS-01 TCGA-AC-A5XS	TCGA-AR-A24R-01 TCGA-AR-A24P	TCGA-BH-AODL-11 TCGA-BH-AODI	TCGA-BH-A5IO-01 ТСGA-RH-A5IO	TCGA-E2-A14X-01 TCGA-F2-A14V	TCGA-E9-A3QA-01 TCGA-F9-A3QA	TCGA-XX-A89A-01 TCGA-XX-A89A
TCGA-A7-A13F-11 TCGA-A7-A13F	TCGA-AC-ASXU-01 TCGA-AC-ASXU	TCGA-AR-A24S-01 TCGA-AR-A24S	TCGA-BH-A0DO-01 TCGA-BH-A0DO	TCGA-BH-A6R8-01 TCGA-BH-A6R8	TCGA-E2-A14Y-01 TCGA-E2-A14Y	TCGA-E9-A3X8-01 TCGA-E9-A3X8	TCGA-27-A8R5-01 TCGA-27-A8R5
TCGA-A7-A13G-01 TCGA-A7-A13G	TCGA-AC-A62V-01 TCGA-AC-A62V	TCGA-AR-A24T-01 TCGA-AR-A24T	TCGA-BH-A0DO-11 TCGA-BH-A0DO	TCGA-BH-A6R9-01 TCGA-BH-A6R9	TCGA-E2-A14Z-01 TCGA-E2-A14Z	TCGA-E9-A54X-01 TCGA-E9-A54X	TCGA-27-A8R6-01 TCGA-27-A8R6
TCGA-A7-A13G-11 TCGA-A7-A13G TCGA-A7-A13H-01 TCGA-A7-A13H	TCGA-AC-A62X-01 TCGA-AC-A62X TCGA-AC-A62Y-01 TCGA-AC-A62Y	TCGA-AR-A24U-01 TCGA-AR-A24U TCGA-AR-A24V-01 TCGA-AR-A24V	TCGA-BH-A0DP-01 TCGA-BH-A0DP TCGA-BH-A0DP-11 TCGA-BH-A0DP	TCGA-BH-A8FY-01 TCGA-BH-A8FY TCGA-BH-A8FZ-01 TCGA-BH-A8FZ	TCGA-E2-A150-01 TCGA-E2-A150 TCGA-E2-A152-01 TCGA-E2-A152 TCGA-E2-A152-01 TCGA-E2-A152	TCGA-E9-A5FK-01 TCGA-E9-A5FK TCGA-E9-A5FL-01 TCGA-E9-A5FL	
TCGA-A7-A26F-01 TCGA-A7-A26F TCGA-A7-A26F-01 TCGA-A7-A26F TCGA-A7-A26G-01 TCGA-A7-A26G	TCGA-AC-A6IV-01 TCGA-AC-A6IV TCGA-AC-A6IV-01 TCGA-AC-A6IV TCGA-AC-A6IX-01 TCGA-AC-A6IX	TCGA-AR-A24W-U1 TCGA-AR-A24W TCGA-AR-A24X-01 TCGA-AR-A24X TCGA-AR-A24Z-01 TCGA-AR-A24Z	TCGA-BH-A0DQ-11 TCGA-BH-A0DQ TCGA-BH-A0DQ-11 TCGA-BH-A0DQ TCGA-BH-A0DS-01 TCGA-BH-A0DS	TCGA-BH-AB28-01 TCGA-BH-AB28 TCGA-C8-A12K-01 TCGA-C8-A12K	TCGA-E2-A153-01 TCGA-E2-A153 TCGA-E2-A153-11 TCGA-E2-A153 TCGA-E2-A154-01 TCGA-E2-A154	TCGA-E9-ASUD-01 TCGA-E9-ASUD TCGA-E9-ASUP-01 TCGA-E9-ASUP TCGA-E9-A6HE-01 TCGA-E9-A6HE	
TCGA-A7-A26H-01 TCGA-A7-A26H	TCGA-AC-A6IX-06 TCGA-AC-A6IX	TCGA-AR-A250-01 TCGA-AR-A250	TCGA-BH-A0DT-01 TCGA-BH-A0DT	TCGA-C8-A12L-01 TCGA-C8-A12L	TCGA-E2-A155-01 TCGA-E2-A155	TCGA-EW-A1IW-01 TCGA-EW-A1IW	
TCGA-A7-A26I-01 TCGA-A7-A26I	TCGA-AC-A6NO-01 TCGA-AC-A6NO	TCGA-AR-A251-01 TCGA-AR-A251	TCGA-BH-A0DT-11 TCGA-BH-A0DT	TCGA-C8-A12M-01 TCGA-C8-A12M	TCGA-E2-A156-01 TCGA-E2-A156	TCGA-EW-A1IX-01 TCGA-EW-A1IX	
TCGA-A7-A26J-01 TCGA-A7-A26J	TCGA-AC-A7VB-01 TCGA-AC-A7VB	ICGA-AR-A252-01 TCGA-AR-A252	TCGA-BH-A0DV-01 TCGA-BH-A0DV	ICGA-C8-A12N-01 TCGA-C8-A12N	TCGA-E2-A158-01 TCGA-E2-A158	TCGA-EW-A1IY-01 TCGA-EW-A1IY	
TCGA-A7-A2KD-01 TCGA-A7-A2KD	TCGA-AC-A7VC-01 TCGA-AC-A7VC	TCGA-AR-A254-01 TCGA-AR-A254	TCGA-BH-A0DV-11 TCGA-BH-A0DV	TCGA-C8-A12O-01 TCGA-C8-A12O	TCGA-E2-A158-11 TCGA-E2-A158	TCGA-EW-A1IZ-01 TCGA-EW-A1IZ	

				-				200						
APC	Apoptosis	Cell Cycle/Ap	Chromatin M	(DNA Damage H	Ŧ	NOTCH	PI3K	PI3K; RAS	RAS	STAT	TGF-b	Transcription	PI3K_AII	RAS_AII
APC	<b>TRAF7</b>	CASP8	ARID1B	ATM	TCH1	NOTCH2	TSC2	ERBB2	CIC	JAK3	ACVR1B	AR	TSC2	CIC
NF2		CARD11	DNMT3A	RECQL4 E	XT2	NOTCH1	PIK3R1	MET	NF1	MPL	FOXL2	IKZF1	PIK3R1	NF1
AXIN1		MYC	SMARCB1	FANCA	XT1	FBXW7	PIK3CA	PDGFRA	NRAS	JAK2		PHOX2B	PIK3CA	NRAS
RNF43		SKP2	WT1	FANCF		GATA1	FLCN	CBL	HRAS		1	RUNX1	FLCN	HRAS
CDH1		CYLD	ARID1A	XPA	-	GATA2	PTEN	SDHD	BRAF			BCOR	PTEN	BRAF
FAM123B		ABL1	SETD2	STAG2	a		TSHR	ALK	KRAS			DICER1	TSHR	KRAS
<b>HNF1A</b>		CHEK2	NCOR1	MSH6				EGFR	MAP2K1				ERBB2	MAP2K1
		FUBP1	ATRX	ERCC4				FGFR3	MAP3K1				MET	ERBB2
		MDM4	H3F3A	FANCG				FGFR2		1			PDGFRA	MET
		PPP2R1A	KDM5C	MLH1				B2M					CBL	PDGFRA
		RB1	MEN1	PMS2				GNA11					SDHD	CBL
		MDM2	PRDM1	BRCA1				KIT					ALK	SDHD
		TP53	ASXL1	FANCD2				GNAQ					EGFR	ALK
		<b>TNFAIP3</b>	MLL2	PMS1				RET					FGFR3	EGFR
		MED12	SMARCA4	PALB2					T				FGFR2	FGFR3
			ARID2	BAP1									B2M	FGFR2
			KDM6A	BLM									GNA11	B2M
			TET2	BUB1B									КІТ	GNA11
			PBRM1	ERCC5									GNAQ	КІТ
			DNMT1	BRCA2									RET	GNAQ
			EZH2											RET

eTable 3: "Cancer genes" curated by Vogelstein and colleagues (Kandoth et al. 2013, PMID 24132290; Vogelstein et al. 2013, PMID 23539594)

#### eTable 4: Individual genes constituting TNBC metagenes and their correlation with Affymetrix microarray

					Correlation RNA-Seq vs	
		-		TNBCmetagene_Affymetrix	Affymetrix	Corelation Class
TNBCmetagene_RNA-Seq	EntrezID	Gene	Affymetrix probeset	(PMID_21978456)	(PMID_25412710)	(PMID_25412710)
laudin	1305	ARUCADS	203954_x_at	Claudin-CD24	0.844374	nign.cor
Claudin	553158	PRR5-ARHGAP8				
Claudin	29841	GRHL1			0.896681	high.cor
Claudin	10256	CNKSR1	204740_at		0.594828	low.cor
Claudin	10053	AP1M2	218261_at		0.782798	high.cor
Claudin	2886	GRB7	210761_s_at		0.814012	high.cor
Claudin	54836	BSPRY	218792_s_at		0.891690	high.cor
laudin	4072	EPCAM	201839_s_at	Claudin-CD24	0.926497	high.cor
laudin	2065	ERBB3	202454_s_at		0.939461	high.cor
laudin	70077	GRHL2	219121_5_dt		0.921312	high.cor
laudin	1364	CLDNA	219308_dt 201428_at	Claudin-CD24	0.817215	high.cor
laudin	1999	ELF3	201420_at	Claudin-CD24	0.792715	high.cor
laudin	126695	C1orf172			0.736777	high.cor
laudin	92359	CRB3			0.562095	low.cor
laudin	57111	RAB25	218186_at	Claudin-CD24	0.792520	high.cor
laudin	58495	OVOL2	211778_s_at		0.894478	high.cor
laudin	114569	MAL2			0 704070	LANK STATE
laudin	149466	CIOFTZIU TNAENA12E			0.731073	high.cor
laudin	999	CDH1	201131 s at		0.891755	high cor
laudin	5652	PRSS8	202525 at		0.821947	high.cor
laudin	91862	MARVELD3			0.778001	high.cor
laudin	2041	EPHA1	205977_s_at		0.540381	low.cor
laudin	140893	C20orf151			0.011991	low.cor
laudin	57662	KIAA1543			0.588605	low.cor
laudin	11187	РКРЗ	209873_s_at		0.855263	high.cor
audin	3898	LAD1	203287_at		0.864078	high.cor
laudin	146439	PVRI 4			0.782149	high cor
laudin	3855	KRT7	209016 s at	Claudin-CD24	0.887283	high.cor
laudin	10045	SH2D3A	219513_s at		0.674034	low.cor
laudin	64787	EPS8L2	218180_s_at		0.747796	high.cor
laudin	64063	PRSS22	205847_at		0.024242	low.cor
laudin	7163	TPD52	201691_s_at		0.714221	high.cor
laudin	51361	HOOK1	219976_at		0.843013	high.cor
laudin	10207	INADL	214493_s_at		0.785196	high.cor
AolApocr	57525	KIAA1324	221874 at		0.971010	high.cor
AolApocr	10551	AGR2	209173 at	Anocrine	0.979907	high cor
10IApocr	3169	FOXA1	204667 at	Apocrine	0.921506	high.cor
folApocr	401546	C9orf152	-		0.954109	high.cor
1olApocr	25803	SPDEF	213441_x_at	Apocrine	0.895191	high.cor
1olApocr	367	AR	211110_s_at	Apocrine		
folApocr	116844	LRG1			0.932460	high.cor
1olApocr	79838	TMC5	219580_s_at	A	0.914312	high.cor
ioiApocr	79083	CARRD	218211_s_at	Apocrine Racal like	0.953980	high.cor
asal-2	58473	PLEKHB1	209504 s at	Basal-like	0.958387	high.cor
asal-2	2001	FLF5	2205504_3_at	Basal-like	0.975629	high.cor
asal-2	7368	UGT8	208358 s at		0.930127	high.cor
asal-2	2296	FOXC1	213260_at	Basal-like	0.954045	high.cor
asal-2	8645	KCNK5	219615_s_at	Basal-like	0.910552	high.cor
asal-2	56963	RGMA			0.869199	high.cor
asal-2	25984	KRT23	218963_s_at	Basal-like	0.963638	high.cor
asal-2	5/44/	NDKG2	206453_s_at	Pacal liko	0.927340	nign.cor
asal-2	54763	ROPN1	200300_5_81	Basal-like		
asal-2	152015	ROPN1B	220425 x at	Basal-like	0.885987	high.cor
asal-2	6663	SOX10	209842_at	Basal-like	0.903033	high.cor
asal-2	6271	\$100A1	205334_at		0.783899	high.cor
asal-2	399694	SHC4			0.934794	high.cor
asal-2	57348	TTYH1	219415_at		0.649728	low.cor
asal-2	30812	SOX8			0.756741	high.cor
asal-2	260429	PRSS33	202037 5 2+	Basal-like	0.457610	low.cor
asal-1	3868	KRT16	202037_5_dL	Desdinine	0.510/10	nigh.cor
asal-1	3861	KRT14	209351 at	Basal-like	0.873736	high.cor
asal-1	3872	KRT17	205157_s_at	Basal-like	0.954628	high.cor
asal-1	3857	KRT9	208188_at		0.579466	low.cor
asal-1	23650	TRIM29	202504_at	Basal-like	0.895904	high.cor
asal-1	1830	DSG3	205595_at		0.835364	high.cor
asal-1	5268	SERPINB5	204855_at	Basal-like	0.946785	high.cor
asal-1	3853	KKIDA	213680 >+	Basal-like	0 08330=	high cor
asal-1	3634	KRTGC	213000_dl 209125_at	Basal-like	0.562305	high cor
asal-1	3852	KRT5	201820_at	Basal-like	0.929155	high.cor
asal-1	6273	\$100A2	204268_at	Basal-like	0.956313	high.cor
/HC2	972	CD74	209619_at	MHC-2	0.845152	high.cor
/HC2	3108	HLA-DMA	217478_s_at	MHC-2	0.925719	high.cor
1HC2	3115	HLA-DPB1	201137_s_at	MHC-2	0.831086	high.cor
1HC2	3109	HLA-DMB	203932_at	MHC-2	0.944257	high.cor
1HC2	3113	HLA-DPA1	211990_at	MHC 2	0.835429	high.cor
-Cell	3122	MVO15	210982_s_at	wirlt-2	0.779916	high.cor
-Cell	83706	FFRMTR	dt		0.885468	high.cor
-Cell	6688	SPI1	205312 at		0.139227	low.cor
-Cell	4689	NCF4	207677_s_at		0.947693	high.cor
-Cell	7454	WAS	38964_r_at			
-Cell	7305	TYROBP	204122_at		0.895385	high.cor
-Cell	10859	LILRB1	211336_x_at		0.724916	high.cor
-Cell	920	CD4	203547_at	T 0.1	0.553020	low.cor
-Cell	7805	LAPTM5	201721_s_at	T-Cell	0.931471	high.cor
-cell -Cell	3587	IL10RA	204912_at	I-Cell	0.953526	high.cor
-cell	045 032	0153	203304_dt 203416_at	T-Cell	0.97/008	high cor
-Cell	54440	SASH3	204923 at		0.574008	nigh.cor
-Cell	1794	DOCK2	213160 at		0.871921	high.cor
-Cell	3071	NCKAP1L	209734_at		0.850532	high.cor
-Cell	57705	WDFY4			0.777223	high.cor

					Correlation RNA-Seq vs	
				TNBCmetagene_Affymetrix	Affymetrix	Corelation Class
TNBCmetagene_RNA-Seq	EntrezID	Gene	Affymetrix probeset	t (PMID_21978456)	(PMID_25412710)	(PMID_25412710)
T-Cell	2124	EVI2B	211742_s_at	T-Cell	0.927470	high.cor
T-Cell	5/88	PIPRC	212388_at	I-Cell	0.893052	high cor
T-Cell	5341	PLEK	205209_dL 203471_s_at		0.893939	high.cor
T-Cell	124460	SNX20			0.653228	low.cor
T-Cell	64092	SAMSN1	220330_s_at	T-Cell	0.913663	high.cor
T-Cell	951	CD37	204192_at		0.845865	high.cor
T-Cell	64333	ARHGAP9			0.667164	low.cor
T-Cell	64926	RASAL3	205242		0.605976	low.cor
I-Cell	9/44	ACAP1	205212_s_at		-0.041094	low.cor
T-Cell	274402	TRC1D10C	204960_at		0.859087	high.cor
T-Cell	1731	Sep 01			0.587892	low.cor
T-Cell	8698	S1PR4	206437 at		0.553150	low.cor
T-Cell	387751	GVIN1	220577 at			
T-Cell	151888	BTLA			0.892209	high.cor
T-Cell	917	CD3G	206804_at		0.906274	high.cor
T-Cell	387357	THEMIS			0.900830	high.cor
T-Cell	3702	ITK	211339_s_at	T-Cell	0.926757	high.cor
T-Cell	50852	TRAT1	217147_s_at		0.951906	high.cor
T-Cell	128611	7NE931	214015_dt		0 761926	high cor
T-Cell	962	CD48	204118 at	T-Cell	0.966425	high cor
T-Cell	3683	ITGAL	213475 s at		0.759917	high.cor
T-Cell	11262	SP140	207777_s_at		0.864143	high.cor
T-Cell	939	CD27	206150_at	T-Cell	0.928377	high.cor
T-Cell	4063	LY9	215967_s_at		0.789344	high.cor
T-Cell	6504	SLAMF1	206181_at		0.910487	high.cor
T-Cell	114836	SLAMF6			0.822855	high.cor
I-Cell	2833	CXCR3	207681_at	7.0.0	0.07/777	
I-Cell	914	CD2	205831_at	I-Cell	0.974268	high.cor
	916	CD3E	205456_at	T-Cell	0.47/508	low.cor
r-cen T-Cell	915	SIBBU	210039_dt	riceii	0.900032	high cor
T-Cell	27740	SIT1	220403_5_dt 205484_at		0.744620	high.cor
T-Cell	84174	SLA2	200-10-1 at		0.855847	high.cor
T-Cell	3003	GZMK	206666 at	T-Cell	0.964480	high.cor
T-Cell	4068	SH2D1A	210116_at			
T-Cell	10663	CXCR6	206974_at		0.909515	high.cor
T-Cell	29851	ICOS	210439_at		0.924099	high.cor
T-Cell	201633	TIGIT			0.932720	high.cor
B-Cell	643	CXCR5	206126_at		0.258556	low.cor
B-Cell B. Coll	115250	BLK ECRL1	206255_at		0.714091	high.cor
B-Cell B-Cell	031	PCRL1 MS4A1	210356 x at		0.748898	high.cor
B-Cell	79368	FCRI 2	2210330_x_at		0.852606	high.cor
B-Cell	930	CD19	206398 s at		0.829660	high.cor
B-Cell	973	CD79A	205049 s at		0.641626	low.cor
B-Cell	8755	ADAM6				
B-Cell	51237	MGC29506	221286_s_at		0.937127	high.cor
B-Cell	83416	FCRL5			0.771325	high.cor
B-Cell	608	TNFRSF17	206641_at		0.962665	high.cor
B-Cell	96610	LOC96610	217179_x_at	B-Cell		
MHC1	3107	HLA-C	216526_x_at	MHC-1	0.850207	high.cor
	3105		215515_X_dt	MHC-1		
MHC1	3106	HLA-R	209140 x at	MHC-1	0 929076	high cor
MHC1	3134	HLA-F	221875 x at	MHC-1	0.927016	high.cor
MHC1	5696	PSMB8	209040_s_at		0.961499	high.cor
MHC1	5698	PSMB9	204279_at		0.954628	high.cor
MHC1	6890	TAP1	202307_s_at		0.952295	high.cor
MHC1	6891	TAP2	208428_at		0.821040	high.cor
FN	2537	IFI6	204415_at		0.961887	high.cor
FN	4599	MX1	202086_at	IFN		
FN	10561	IFI44	214453_s_at	IFN	0.91/812	high.cor
FN	5,4730	11/144L XAF1	204435_dL 206133_at	IF IN	0.938100	high cor
FN	34739	JEIT1	200135_at	IEN	0.969795	high cor
FN	3433	IFIT2	217502 at		0.968758	high.cor
IFN	129607	CMPK2			0.974851	high.cor
FN	91543	RSAD2	213797_at	IFN	0.975888	high.cor
FN	3437	IFIT3	204747_at	IFN	0.973814	high.cor
FN	4939	OAS2	204972_at	IFN	0.816243	high.cor
IFN	4938	OAS1	202869_at	IEN	0.891172	high.cor
	8638	UASL	205660_at	IFN	0.959360	high.cor
ITIN	4940	UAS3	218400_at	IFN II-8	0.884625	nigh.cor
18-VEGE	2920	CXCL2	20774_x_dL 207850 at	10-0	0.759204	high cor
II 8-VEGE	2919	CXCL1	204470 at	II -8	0.885209	high.cor
L8-VEGF	6372	CXCL6	206336_at		0.716730	high.cor
L8-VEGF	6374	CXCL5	214974_x_at		0.657895	low.cor
L8-VEGF	3576	IL8	202859_x_at	IL-8	0.909321	high.cor
L8-VEGF	133	ADM	202912_at	VEGF	0.963832	high.cor
L8-VEGF	353322	ANKRD37			0.823114	high.cor
L8-VEGF	51129	ANGPTL4	221009_s_at	VEGF	0.825966	high.cor
L&-VEGF	7422	VEGFA	210512_s_at	VEGF	0.898626	high.cor
Indothel	2828	GPR4	206236_at		0.792326	high.cor
Indothel	51204	PCDH12	219656 at		0.746305	high cor
Endothel	54538	ROBO4	220758 s at		0.864208	high.cor
Endothel	161198	CLEC14A	220,30_3_80		0.858893	high.cor
Endothel	1003	CDH5	204677_at		0.886116	high.cor
Endothel	22899	ARHGEF15	205507_at		0.538177	low.cor
Endothel	7075	TIE1	204468_s_at		0.760695	high.cor
Endothel	947	CD34	209543_s_at		0.710721	high.cor
Endothel	79742	CXorf36	219652_s_at			
Endothel	80177	MYCT1	220471_s_at		0.914701	high.cor
Endothel	79812	MMRN2	219091_s_at		0.920210	high.cor
andothel	22918	CD93	202878_s_at		0.833809	high.cor
Indothel	221395	GPR116	212950_at		0.897394	high.cor
Endothel	64123	ELTD1	219134_at		0.815595	high.cor
Endothel	5/87	PIPRB	21/1//_s_at		0.000100	nigh.cor
Collag-Strome	1010	IEK AEDD1	200702_at	Stroma	0.032424	high cor
Collag-Stroma	1303	COI 17Δ1	201/32_dl	Juona	0.965440	high.cor

					Correlation RNA-Seq vs	
				TNBCmetagene_Affymetrix	Affymetrix	Corelation Class
TNBCmetagene_RNA-Seq	EntrezID	Gene	Affymetrix probeset	(PMID_21978456)	(PMID_25412710)	(PMID_25412710)
Collag-Stroma	1289	COLSA1	212488_at	Stroma	0.890394	high.cor
Collag-Stroma	1277	COLIAI	202311_5_at	Stroma	0.910002	high.cor
Collag-Stroma	1270	COLIAZ	202404_3_at	Stroma	0.905483	high.cor
Collag-Stroma	1201	COL5A2	221729 at	Stroma	0.899922	high.cor
Collag-Stroma	1293	COL6A3	201438_at	Stroma	0.944905	high.cor
Collag-Stroma	5159	PDGFRB	202273_at	Stroma	0.834975	high.cor
Collag-Stroma	83468	GLT8D2	221447_s_at		0.880931	high.cor
Collag-Stroma	4313	MMP2	201069_at	Stroma	0.927081	high.cor
Collag-Stroma	54796	BNC2	220272_at		0.836661	high.cor
Collag-Stroma	57616	TSHZ3			0.817475	high.cor
Collag-Stroma	1009	CDH11	207173_x_at	Stroma	0.855198	high.cor
Collag Stroma	254228	FAIVI26E	202766 c at	Stroma	0.610513	low.cor
Collag-Stroma	51220	DACT1	202700_3_at	Sciona	0.832123	high.cor
Collag-Stroma	2191	FAP	209955 s at	Stroma	0.800383	high.cor
Collag-Stroma	1634	DCN	209335_at	Stroma	0.892339	high.cor
Collag-Stroma	4060	LUM	201744 s at	Stroma	0.798613	high.cor
Collag-Stroma	283298	OLFML1	217525_at		0.848133	high.cor
Collag-Stroma	10631	POSTN	210809_s_at	Stroma	0.841716	high.cor
Collag-Stroma	22795	NID2	204114_at	Stroma	0.898367	high.cor
Collag-Stroma	6678	SPARC	212667_at	Stroma	0.863041	high.cor
Collag-Stroma	7070	THY1	213869_x_at	Stroma	0.855328	high.cor
Collag-Stroma	57125	PLXDC1	219700_at		0.747861	high.cor
Adipo	2167	FABP4	203980_at	Adipocyte	0.931747	high.cor
Adipo	729	C6	210168_at		0.650570	low.cor
Adipo	63924	LIDEL	713230 <sup>-q</sup> L		0.875292	nign.cor
Adipo	364		200613 5 2*	Adipocyte	0.056270	high cor
Adipo	125		207013_3_dL 207175_at	Adipocyte	0.530378	high.cor
Adipo	53/0	PLIN1	205913 at	Adipocyte	0.894089	high.cor
Adipo	286752	TUSCS	200010_0L	. upocyte	0.505769	low cor
Adipo	729359	PLIN4			0.647459	low.cor
Hemoglobin	3040	HBA2				
Hemoglobin	3043	HBB	211696_x_at	Hemoglobin	0.884690	high.cor
HOXA	3206	HOXA10	213150_at	HOXA	0.823892	high.cor
HOXA	3207	HOXA11	213823_at	HOXA	0.638838	low.cor
HOXA	221883	HOXA11AS				
HOXA	3205	HOXA9				
HOXA	3198	HOXA1	214639_s_at		0.754019	high.cor
HOXA	3201	HOXA4	206289_at	HOXA	0.678636	low.cor
HOXA	3199	HOXA2	214457_at		0.622245	low.cor
HOXA	3200	HOXA3	208604_s_at		0.848587	high.cor
HOXA	3202	HOXA5	213844_at	HOXA	0.863430	high.cor
HOXA	3203	HOXA6	208557_at	11074	0.638450	low.cor
Histopo	3204	HUXA/	200847_S_dL	HUXA	0.804597	nign.cor
Histone	8349	HIST2H2RE	202708 s at	Histone	0 924423	high cor
Histone	440689	HIST2H2BE	202700_3_80	Thatone	0.524425	mgn.cor
Histone	8343	HIST1H2BF				
Histone	3017	HIST1H2BD	209911 x at	Histone	0.826484	high.cor
Histone	3012	HIST1H2AE	214469_at	Histone	0.854939	high.cor
Histone	8339	HIST1H2BG	215779_s_at	Histone	0.487101	low.cor
Histone	8351	HIST1H3D	214472_at			
Histone	8334	HIST1H2AC				
Histone	8347	HIST1H2BC				
Histone	8365	HIST1H4H	208180_s_at	Histone		
Histone	8344	HIST1H2BE				
Histone	85236	HIST1H2BK	209806_at	Histone	0.847874	high.cor
Histone	3006	HIST1H1C	209398_at	Histone	0.876718	high.cor
Histone	8337		21//81 -+			
Histone	8330	HIST1H2RI	214461_dl 214502_st		0.662886	low cor
Histone	83/12	HIST1H2B0	214540 at		0.520936	low.cor
Histone	26212	OR286	216522 at		0.663404	low.cor
Prolif	51514	DTL	218585_s at	Proliferation	0.885144	high.cor
Prolif	4751	NEK2	204641_at	Proliferation	0.979258	high.cor
Prolif	25896	INTS7	218783_at		0.551854	low.cor
Prolif	79915	ATAD5	220223_at		0.730166	high.cor
Prolif	699	BUB1	209642_at	Proliferation	0.966684	high.cor
Prolif	150468	CKAP2L			0.895579	high.cor
Prolif	7153	TOP2A	201292_at	Proliferation	0.955022	high.cor
Prolif	1063	CENPF	209172_s_at	Proliferation	0.912626	high.cor
Prolif	259266	ASPM	219918_s_at	Proliferation	0.968693	high.cor
Prolif	9928	KIF14	206364_at	Proliferation	0.932136	high.cor
Prolif	6491	SIIL	2000009_00		0.904978	high.cor
Prolif	151540	SGOL 1	213300_dl		0.945035	high.cor
Prolif	73307	NCAPH	212949 at		0.956707	high.cor
Prolif	23357	FOXM1	202580 x at	Proliferation	0.971805	high.cor
Prolif	9918	NCAPD2	201774 s at		0.778714	high.cor
Prolif	10635	RAD51AP1	204146 at		0.952489	high.cor
Prolif	171017	ZNF384	212369_at		0.500065	low.cor
Prolif	78995	C17orf53	219879_s_at		0.509269	low.cor
Prolif	80174	DBF4B			0.356819	low.cor
Prolif	146909	KIF18B	222039_at	Proliferation	0.939914	high.cor
Prolif	3833	KIFC1	209680_s_at		0.543039	low.cor
Prolif	113130	CDCA5			0.968175	high.cor
Prolif	991	CDC20	202870_s_at	Proliferation	0.971092	high.cor
Prolif	55143	CDCA8	221520_s_at	Proliferation	0.955211	high.cor
Prolif	11004	KIF2C	209408_at	Proliferation	0.979129	high.cor
Prolif	8438	KAD54L	204558_at		0.913858	high.cor
Prolif	4998	URCIL	205085_at		0.834651	nign.cor
Prolif	9212	AUKKB	203404_dL		0.934004	high.cor
Prolif	10024	TROAD	221351_5_dL 204649_at		0.552400	high cor
Prolif	11065	UBF2C	202954 at	Proliferation	0.964480	high.cor
Prolif	83461	CDCA3	221436_s_at	Proliferation	0.967786	high.cor

# **R-MarkDown-document: TNBC\_TIL\_analysis**

Thomas Karn

May-18 2017

# **Table of Contents**

1
1
3
4
5
10
10
12
13
14
16
18

# SECTION-1 Selection of a gene expression based TNBC cohort from TCGA

We use the cgdsr package to access data from the cBIO Portal.

```
library("cgdsr")
cbiop <- CGDS("http://www.cbioportal.org/public-portal/")
# getCancerStudies(cbiop)$cancer_study_id
clidat = getClinicalData(cbiop,"brca_tcga_all")</pre>
```

```
1.1 Analysis of correlation of ESR1 gene expression from RNA-Seq and Agilent
microarray platform
esr1.rseq = getProfileData(cbiop,"ESR1","brca_tcga_rna_seq_v2_mrna",
"brca_tcga_all")
esr1.agi = getProfileData(cbiop,"ESR1","brca_tcga_mrna", "brca_tcga_all")
```

```
# generate matrix of cases with both data for Agilent and RNA-Seq:
esr1.comp=as.data.frame(cbind(esr1.agi$ESR1, log2(esr1.rseq$ESR1+1))
        [(!is.nan(esr1.agi$ESR1)) & (!is.nan(esr1.rseq$ESR1)), ])
colnames(esr1.comp)=c("ESR1.AGI", "ESR1.RSEQ")
```

# correlation between Agilent and RNA-Seq:
plot(esr1.comp\$ESR1.RSEQ, esr1.comp\$ESR1.AGI)



esr1.comp\$ESR1.RSEQ

cor(esr1.comp\$ESR1.RSEQ, esr1.comp\$ESR1.AGI)

## [1] 0.9821414

# bimodal distribution of RNA-Seq data
hist(log2(esr1.rseq\$ESR1+1), breaks=80)





# **1.2** Analysis of correlation of PGR gene expression from RNA-Seq and Agilent microarray platform

```
pgr.rseq = getProfileData(cbiop, "PGR", "brca_tcga_rna_seq_v2_mrna",
  "brca_tcga_all")
pgr.agi = getProfileData(cbiop, "PGR", "brca_tcga_mrna", "brca_tcga_all")
# generate matrix of cases with both data for Agilent and RNA-Seq:
pgr.comp=as.data.frame(cbind(pgr.agi$PGR, log2(pgr.rseq$PGR+1))
        [(!is.nan(pgr.agi$PGR)) & (!is.nan(pgr.rseq$PGR+1))
        [(!is.nan(pgr.agi$PGR)) & (!is.nan(pgr.rseq$PGR)), ])
colnames(pgr.comp)=c("PGR.AGI", "PGR.RSEQ")
# correlation between Agilent and RNA-Seq:
plot(pgr.comp$PGR.RSEQ, pgr.comp$PGR.AGI)
```



pgr.comp\$PGR.RSEQ

cor(pgr.comp\$PGR.RSEQ, pgr.comp\$PGR.AGI)

## [1] 0.9499931

# bimodal distribution of RNA-Seq data
hist(log2(pgr.rseq\$PGR+1), breaks=80)





**1.3** Analysis of correlation of HER2 gene expression from RNA-Seq and Agilent microarray platform

```
# correlation between Agilent and RNA-Seq:
plot(erbb2.comp$ERBB2.RSEQ, erbb2.comp$ERBB2.AGI)
```



cor(erbb2.comp\$ERBB2.RSEQ, erbb2.comp\$ERBB2.AGI)

## [1] 0.9547622

```
# bimodal distribution of RNA-Seq data
hist(log2(erbb2.rseq$ERBB2+1), breaks=80)
```

```
Histogram of log2(erbb2.rseq$ERBB2 +
```



log2(erbb2.rseq\$ERBB2 + 1)

```
1.4 Generate TNBC dataset
```

```
# Select tnbc/dnbc based on cutoffs from distribution of RNA-Seq
# define a logical selection vector
tnbc.group= !is.na(esr1.rseq) & !is.na(erbb2.rseq) &
    (log2(esr1.rseq$ESR1+1)<10) & (log2(erbb2.rseq$ERBB2+1)<14)</pre>
colnames(tnbc.group)="tnbc"
sum(na.omit(tnbc.group))
## [1] 208
# Generate tnbc dataset
tnbc.data= cbind(log2(esr1.rseq$ESR1+1)[tnbc.group],
                  log2(pgr.rseq$PGR+1)[tnbc.group],
                  log2(erbb2.rseq$ERBB2+1)[tnbc.group])
row.names(tnbc.data)= row.names(tnbc.group)[tnbc.group]
colnames(tnbc.data)=c("ESR1.RSEQ", "PGR.RSEQ", "ERBB2.RSEQ")
# Merge of Clinical data and tnbc dataset
# find subset in clidat corresponding to tnbc
clidat.sel=clidat[row.names(clidat)%in% row.names(tnbc.data),]
# merge tnbc.data and clinical data, left outer join:
```

```
tnbc.data= merge(tnbc.data, clidat.sel, by="row.names", all.x =TRUE)
```

```
# "merge" creates resorted dataframe with the row.names
# as a new first column "Row.names"
```

```
# rebuild structure (row.names):
```

```
row.names(tnbc.data)=tnbc.data$Row.names
tnbc.data=tnbc.data[,colnames(tnbc.data)!= "Row.names"]
```

# check residual receptor expression in tnbc dataset: hist(tnbc.data\$ESR1.RSEQ, xlim=c(0,20), breaks=40) # tnbc group



Histogram of tnbc.data\$ESR1.RSEQ

hist(log2(esr1.rseq\$ESR1+1),xlim=c(0,20), breaks=80) # all samples

Histogram of log2(esr1.rseq\$ESR1 + 1



hist(tnbc.data\$PGR.RSEQ, xlim=c(0,20), breaks=40) # tnbc group

# Histogram of tnbc.data\$PGR.RSEQ



hist(log2(pgr.rseq\$PGR+1), xlim=c(0,20), breaks=80) # all samples

Histogram of log2(pgr.rseq\$PGR + 1)



hist(tnbc.data\$ERBB2.RSEQ, xlim=c(5,20), breaks=40) # tnbc group





hist(log2(erbb2.rseq\$ERBB2+1),xlim=c(5,20), breaks=80) # all samples

# Histogram of log2(erbb2.rseq\$ERBB2 +



log2(erbb2.rseq\$ERBB2 + 1)

## SECTION-2 Gene filtering in RNA-Seq data

```
# Spearman correlation values between RNA-Seq and Affymetrix microarray
# for 16,097 Jetset probes for 57 paired frozen breast cancer samples
# can be obtained from:
# Suppl.Tab.S2 of Fumagalli et al. 2014, PubmedID 25412710
n208.FumagCorrel <-
read.delim("2016_05_31_median_mean_n208RNASeq_vs_FumagalliCorrel.txt")
# Plot median expression vs Spearman correlation coefficient
x=n208.FumagCorrel[,c(1,3)]
plot(x)
```













# Histogram of x\$median



# Distribution of Spearman correlation coefficients
hist(x\$cor\_Fumagalli)





rm(x)

# **SECTION-3** Metagene construction



## Gene correlations RNA-Seq vs Affy

par(mar=c(5.1, 4.1, 4.1, 2.1))
hist(metag\$Correl\_PMID\_25412710)

# Histogram of metag\$Correl\_PMID\_25412



metag\$Correl\_PMID\_25412710

boxplot(metag\$Correl\_PMID\_25412710)



#### median(metag\$Correl\_PMID\_25412710, na.rm=T)

## [1] 0.8831346

```
summary(metag$Correl_PMID_25412710, na.rm=T)
```

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
##	-0.04109	0.77800	0.88310	0.82610	0.93210	0.98810	35

#### 3.2 Metagene calculation from RNA-Seq expression

```
# Load RNAseq data of 304 genes for 208 tnbc samples
# RNAseq data of 1218 TCGA BRCA can be downloaded from UCSC Xena browser
(https://tcga.xenahubs.net/download/TCGA.BRCA.sampleMap/HiSeqV2)
```

```
n304genes <- read.table("n208tnbc_n304genes_RNAseq.csv", header=TRUE,
sep=";")
```

```
# scale transposed expression data and re-transpose
n304.expr.sca= t(scale(t(n304genes[,5:212])))
colnames(n304.expr.sca)=colnames(n304genes[,5:212])
```

tnbc.data.meta17= merge(tnbc.data, metag17, by="row.names", all.x =TRUE)
# "merge" command results in resorting of dataframe and loss of row.names
# but an additional new first column "Row.names"
# Assign new row.names from this additional column and then delete it
row.names(tnbc.data.meta17)=tnbc.data.meta17\$Row.names
tnbc.data.meta17=tnbc.data.meta17[,colnames(tnbc.data.meta17)!= "Row.names"]

# **SECTION-4 MATH** analysis of dispersion in mutant allele frequencies

```
# Copy of maf file from TCGA
genome.wustl.edu BRCA.ILLuminaGA DNASeg.Level 2.1.1.0.curated.somatic.maf.txt
(52MB) is available at https://portal.qdc.cancer.qov/legacy-
archive/files/50d6fb1d-5bb1-4a30-9e91-6d45bd9b1c3f
# The required variant allele frequencies have been extracted in the smaller
file used here: "VAF-
table_genome.wustl.edu_BRCA.IlluminaGA_DNASeg.Level_2.1.1.0.curated.somatic.m
af.txt""
maf.download <- read.delim(</pre>
    "VAF-
table_genome.wustl.edu_BRCA.IlluminaGA_DNASeq.Level_2.1.1.0.curated.somatic.m
af.txt")
all.maf = maf.download[,c("Hugo Symbol", "Tumor Sample Barcode",
"tumor_vaf")]
TCGA_Sample=substr(all.maf$Tumor_Sample_Barcode, 1, 15)
all.maf = cbind(TCGA Sample, all.maf)
# calculate for each sample the median of tumor vaf values
med=by(all.maf$tumor vaf, all.maf$TCGA Sample, median)
# convert list to dataframe and transpose
med.df = t(as.data.frame(as.list(med)))
colnames(med.df)= "med.mut.AF"
# calculate MAD (Median Absolute Deviation) for each sample
MAD=by(all.maf$tumor vaf, all.maf$TCGA Sample, mad)
```

```
# convert List to dataframe and transpose
MAD.df= t(as.data.frame(as.list(MAD)))
colnames(MAD.df)= "MAD.mut.AF"
```

```
# calculate MATH (Mutant Allele Tumor Heterogeneity) as MATH=100*MAD/median
MATH.all =100 * MAD.df / med.df
colnames(MATH.all)= "MATH"
```

hist(MATH.all)







hist(tnbc.data.meta17\$B.Cell)





tnbc.data.meta17\$B.Cell

hist(tnbc.data.meta17\$IL8.VEGF)





tnbc.data.meta17\$IL8.VEGF

# Since no clear bimodality observed in distributions, # we stay with previously established cutoffs for metagenes/signatures: # MHC2 metagene: Upper quartile (Rody 2009, PMID 19272155) # B-Cell metagene: Lower quartile (Rody 2011, PMID 21978456) # IL8.VEGF metagene: Median split (Rody 2011, PMID 21978456) 5.1 MHC2/IL8VEGF signature # Define upper quartile MHC2 metagene (based on Rody 2009, PMID 19272155) MHC2.q4=tnbc.data.meta17\$MHC2 > quantile(tnbc.data.meta17\$MHC2, probs=0.75) # Define below median IL8.VEGF metagene (cutoff from Rody 2011, PMID 21978456) IL8.VEGF.q12=tnbc.data.meta17\$IL8.VEGF < quantile(tnbc.data.meta17\$IL8.VEGF,</pre> probs=0.5) # Define prognostic signature MHC2.IL8.VEGF.sig = MHC2.q4 & IL8.VEGF.q12 ## Check MHC2.IL8.VEGF.sig in Survival analysis time=tnbc.data.meta17\$dfs.120 censor= (tnbc.data.meta17\$ev.120 =="Recurred/Progressed") strata= MHC2.IL8.VEGF.sig test=survfit(Surv(time, censor)~strata,conf.type="none") summary(test)

```
## Call: survfit(formula = Surv(time, censor) ~ strata, conf.type = "none")
##
## 14 observations deleted due to missingness
##
                   strata=FALSE
##
      time n.risk n.event survival std.err
##
      5.09
              151
                        1
                             0.993 0.00660
##
      6.80
              149
                        1
                             0.987 0.00933
##
      7.79
              145
                        1
                             0.980 0.01149
      9.89
              138
                        1
                             0.973 0.01342
##
```

##	10.02	135	1	0.966 0.01513
##	10.28	134	1	0.958 0.01665
##	12.55	128	1	0.951 0.01812
##	12.71	126	1	0.943 0.01949
##	14.98	113	1	0.935 0.02103
##	16.10	109	1	0.926 0.02252
##	18.27	103	1	0.917 0.02403
##	18.50	102	1	0.908 0.02542
##	19.32	99	1	0.899 0.02677
##	21.91	89	1	0.889 0.02831
##	22.40	88	1	0.879 0.02974
##	23.95	82	1	0.868 0.03125
##	28.22	74	1	0.857 0.03295
##	31.90	69	1	0.844 0.03474
##	32.65	67	1	0.832 0.03643
##	33.31	63	1	0.818 0.03817
##	35.22	57	1	0.804 0.04011
##	36.79	53	1	0.789 0.04212
##	37.32	52	1	0.774 0.04396
##	40.70	47	1	0.757 0.04600
##	42.81	44	1	0.740 0.04807
##	53.02	37	1	0.720 0.05076
##	53.88	36	1	0.700 0.05315
##	76.54	21	1	0.667 0.06017
##	101.05	11	1	0.606 0.07957
##				
##		S	trata=TR	UE
##	time	n.risk	n.event	survival std.err
	. /	1 11 2		

plot(test, lty=c(1,3), xlab="Time", ylab="Survival Probability")
legend(10, 0.4, c("Poor", "Good") , lty=c(1,2))



 $\ensuremath{\mathbb{C}}$  2017 American Medical Association. All rights reserved.

#### 5.2 B-Cell/IL8VEGF signature

```
# Define B-Cell metagene above lowest quartile (cutoff from Rody 2011, PMID
21978456)
B.Cell.q234=tnbc.data.meta17$B.Cell > quantile(tnbc.data.meta17$B.Cell,
probs=0.25)
# Define below median IL8.VEGF metagene (cutoff from Rody 2011, PMID
21978456)
IL8.VEGF.q12=tnbc.data.meta17$IL8.VEGF < quantile(tnbc.data.meta17$IL8.VEGF,</pre>
probs=0.5)
# Define prognostic signature
B.Cell.IL8.VEGF.sig = B.Cell.q234 & IL8.VEGF.q12
## Check B.Cell.IL8.VEGF.sig in Survival analysis
time=tnbc.data.meta17$dfs.120
censor= (tnbc.data.meta17$ev.120 =="Recurred/Progressed")
strata= B.Cell.IL8.VEGF.sig
test=survfit(Surv(time, censor)~strata,conf.type="none")
summary(test)
## Call: survfit(formula = Surv(time, censor) ~ strata, conf.type = "none")
##
## 14 observations deleted due to missingness
##
                   strata=FALSE
##
      time n.risk n.event survival std.err
##
      5.09
              108
                        1
                             0.991 0.00922
      6.80
                        1
##
              106
                             0.981 0.01303
##
      7.79
              102
                        1
                             0.972 0.01607
##
      9.89
               97
                        1
                             0.962 0.01877
               95
##
     10.02
                        1
                             0.952 0.02113
##
     10.28
               94
                        1
                             0.942 0.02320
##
     12.71
               89
                        1
                             0.931 0.02524
##
     18.27
               71
                        1
                             0.918 0.02808
##
     21.91
               62
                        1
                             0.903 0.03129
##
     23.95
               58
                        1
                             0.887 0.03440
##
     28.22
               52
                        1
                             0.870 0.03774
##
     33.31
               45
                        1
                             0.851 0.04156
##
     35.22
               41
                        1
                             0.830 0.04544
##
     36.79
               37
                        1
                             0.808 0.04944
##
     37.32
               36
                        1
                             0.785 0.05292
##
     42.81
               29
                        1
                             0.758 0.05761
##
               22
                        1
     53.02
                             0.724 0.06448
##
     53.88
               21
                        1
                             0.689 0.07002
##
     76.54
               13
                        1
                             0.636 0.08230
##
   101.05
                8
                        1
                             0.557 0.10355
##
##
                   strata=TRUE
## time n.risk n.event survival std.err
## 12.6
             61
                      1
                           0.984 0.0163
    15.0
##
             54
                      1
                            0.965
                                   0.0241
             52
                      1
## 16.1
                           0.947
                                   0.0299
```

##	18.5	49	1	0.928	0.0350
##	19.3	45	1	0.907	0.0398
##	22.4	43	1	0.886	0.0441
##	31.9	34	1	0.860	0.0499
##	32.6	32	1	0.833	0.0551
##	40.7	27	1	0.802	0.0611

plot(test, lty=c(1,3), xlab="Time", ylab="Survival Probability")
legend(10, 0.4, c("Poor", "Good") , lty=c(1,2))



Time

dir()

```
## [1] "2016_05_31_median_mean_n208RNASeq_vs_FumagalliCorrel.txt"
## [2] "2016 06 01 TNBC-metagenes gene list.txt"
## [3] "n208tnbc_n304genes_RNAseq.csv"
## [4] "TNBC_TIL_analysis_2017_05_18.Rmd"
## [5] "TNBC_TIL_analysis_2017_05_18_files"
## [6] "VAF-
table_genome.wustl.edu_BRCA.IlluminaGA_DNASeq.Level_2.1.1.0.curated.somatic.m
af.txt"
sessionInfo()
## R version 3.3.2 (2016-10-31)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 14393)
##
## locale:
## [1] LC COLLATE=German Germany.1252 LC CTYPE=German Germany.1252
## [3] LC_MONETARY=German_Germany.1252 LC_NUMERIC=C
## [5] LC_TIME=German_Germany.1252
##
## attached base packages:
## [1] stats
                 graphics grDevices utils
                                               datasets
                                                         methods
                                                                    base
##
```

```
## other attached packages:
## [1] survival_2.40-1 hexbin_1.27.1
                                       cgdsr_1.2.5
##
## loaded via a namespace (and not attached):
## [1] Rcpp_0.12.9
                          lattice_0.20-34
                                            digest_0.6.12
                          R.methodsS3_1.7.1 grid_3.3.2
## [4] rprojroot_1.2
## [7] backports_1.0.5
                          magrittr_1.5
                                            evaluate_0.10
## [10] stringi_1.1.2
                          R.oo_1.21.0
                                            Matrix_1.2-8
## [13] rmarkdown_1.3
                          splines_3.3.2
                                            tools_3.3.2
## [16] stringr_1.2.0
                          yaml_2.1.14
                                            htmltools_0.3.5
## [19] knitr_1.15.1
```