## METABRIC dataset Microarray mRNA expression levels of 20,603 genes and clinical data from 233 patients with TNBC. Correlation between RFS and OS for TNBC patients Dataset used: METABRIC Software used: JMP Pro 17 Result: There was a strong positive correlation between time to recurrence and time to death for TNBC patients. Wilcoxon test Highly significant differences (p<0.005) in RFS and OS between the high and low expression groups were evaluated for each of the 20,603 genes. Dataset used: METABRIC Software used: Python 3.8 Result: 123 genes were differentially expressed. Multivariable Cox proportional hazards analysis Hazard ratios of high versus low expression for recurrence and death, adjusted for NPI and age at diagnosis, were estimated. Dataset used: METABRIC Software used: JMP Pro 17 Result: The 123 genes were indepedent prognositic biomarkers for TNBC recurrence and progression. Low expression of 88 genes and high expression of 35 genes were associated with poor clinical outcome GSE96058 dataset for validation Spearman's correlation analysis Expression level correlations between differentially expressed genes RNA-seg mRNA expression data and clinical data from 143 patients with TNBC. and immune checkpoint molecules were examined. Dataset used: METABRIC Software used: Python 3.8 Validation analysis Results: Of the 88 genes, 71 except for 17 showed positive The 123 genes (88 and 35) were validated by multivariable cox analysis. correlations with each other and with immune checkpoint molecules. Dataset used: GSE96058 The 35 genes showed no correlation with each other or with immune Software used: IMP Pro 17 checkpoints. Results: Of the 88 and 35 genes, low expression of 16 and high expression of 2 genes, respectively, were associated with poor clinical outcome. Gene Ontrogy enrichment analysis GO analysis was performed to further investigate biological functions of genes whose expression levels correlated. Spearman's correlation analysis Signature analysis Dataset used: METABRIC Expression level correlations between the 18 genes identified in both datasets TNBC progression was compared between the high and low signature Software used: Metascape and immune checkpoint molecules were investigated. score groups using a gene set consisting of the 18 differentially expressed Results: Almost all the 71 genes were immune-related genes. Dataset used: METABRIC and GSE96058 genes identified in both datasets. Dataset used: METABRIC and GSE96058 Software used: Python 3.8 Results: Each of the 16 genes that were positively correlated in all pairs was Software used: Python 3.8 and JMP Pro 17 also positively correlated with immune checkpoint molecules. None of the 2 Results: Significant differences between the two groups were observed in Evaluation of immune cell infiltration For the 88 genes whose low expression was associated with poor genes had correlation with immune checkpoint molecules. both datasets. The gene expression signature can predict the risk of early recurrence and shorter survival outcome identified in the METABRIC, the proportions of 22 tumorinfiltrating immune cells in the high and low expression groups were estimated. Evaluation of immune cell infiltration Dataset used: METABRIC For each of the 16 genes whose low expression was associated with poor Software used: CIBERSORTx outcome identified in both datasets, and for each of the eight immune Results: Immune-related genes showed significant differences in the checkpoint molecules, the proportions of 22 tumor-infiltrating immune cells in level of immune cell infiltration in the tumor microenvironment the high and low expression groups were estimated. between the two groups. Genes not correlated with immune Dataset used: METABRIC and GSE96058 checkpoint molecules showed no significant differences in the level Software used: CIRERSORTy of immune cell infiltration. Results: TNBC patients at low and high risk of recurrence/progression differ in the degree of infiltration of immune cells in the tumor microenvironment. particularly CD8+ T cells, activated memory CD4+ T cells, activated NK cells, and M0, M1, and M2 macrophages.