Analysis of Opposing Histone Modifications H3K4me3 and H3K27me3 Reveals Candidate Diagnostic Biomarkers for TNBC and Gene Set Prediction Combination

Hyoung-Min Park¹, HuiSu Kim¹, Kang-Hoon Lee¹, and Je-Yoel Cho*¹

¹Department of Biochemistry, BK21 Plus and Research Institute for Veterinary Science, School of Veterinary Medicine, Seoul National University, Seoul, Korea.

Running Title: Opposing Histone Modifications in TNBC

*Corresponding author:
Je-Yoel Cho, DVM, PhD Professor,
Department of Biochemistry,
College of Veterinary Medicine Seoul National University,
Gwanak-ro1, Gwanak-gu, Seoul, Korea
Tel) +82-02-800-1268
Fax) +82-02-886-1268
E-mail) jeycho@snu.ac.kr
Supplement Figure 1. Candidate accuracy comparison plot. Each x-axis represents ChIP-seq peak fold change and y-axis represents RNA-seq log2 fold change of MDA-MB-436 compared to HMEC cell lines. H3K4me3 and H3K27me3 are analyzed separately and combined in candidates of up-regulated (A) and down-regulated (B). The matching correlation accuracy is presented in each plot. The red dots indicate the top 10 up-regulated candidates and blue dots indicate the top 10 down-regulated candidates.
Supplement Figure 2. Expression box plot. Box plot of 10 up-(A) and down (B)-regulated candidates using TCGA expression data.
Supplement Figure 3. Survival rate analysis. Kaplan-Meier (KM) plots representing each TNBC candidate gene. All genes’ risk of death was compared using the TCGA database. A. KM plots of up-regulating genes. B. KM plots of down-regulating genes.
| Gene   | Affymetrix ID | Gene   | Affymetrix ID |
|--------|--------------|--------|--------------|
| CDH2   | 203440_at    | DUSP6  | 208891_at    |
| DCLK2  | 227666_at    | IRX2   | 228462_at    |
| NOVA1  | 205794_s_at  | ATP2B1 | 212930_at    |
| PLCL2  | 213309_at    | VSNL1  | 203797_at    |
| SOX5   | 207336_at    | ADRB2  | 206170_at    |
| SALL1  | 229273_at    | PLXDC2 | 227995_at    |
| SYTL4  | 227703_s_at  | PLD5   | 1563933_s_at |
| DNER   | 226281_at    | TPD5L1 | 203786_s_at  |
| NAT8L  | 228880_at    | FAM84A | 231439_at    |
| MMP16  | 207012_at    | SNX19  | 202359_s_at  |

**Supplementary data Table 2. Affymetrix ID of top 10 potential up and down regulated candidates.** Each up and down regulated candidate is listed with its individual gene ID and matching affymetrix ID.
## Supplementary data Table 3. Primer list

| Target size | Gene | Forward            | Backward                         | TM |
|-------------|------|--------------------|----------------------------------|----|
| 233         | CDH2 | CCTTTCACACACGCCACGG | TGTTCGGGTTCGTTCTGGATG           | 61 |
| 235         | DCLK2| GTGAAGCCTAGAAAAGCGGTG| CCAGGACAAAGTCATCTCTGGG          | 61 |
| 274         | NOVA1| GGCAGGAACTGCAACACATGA| AGACAGCTTGATGGTGGCTC            | 61 |
| 193         | PLCL2| GCATAGGCACCAAGGAAGGT| GCATTTTCGTTGTCCTGGTTT           | 61 |
| 162         | SOX5 | TCTCCGTATGGGGAAGCAGA| CGTCAGCAGAGAAGCTGCTGG          | 61 |
| 212         | SALL1| TTTCAATCCGACCCCGAAG| GTGGGGAGGGCTGGATTTTCA           | 61 |
| 293         | SYTL4| TACCCCGATCTCTCTCGGA| CCTCTTGGCCCTCTCTTTT            | 61 |
| 163         | DNER | TGCAGGGACCTGTTGTAGG | TAAACCCGGGTGCACTAGTGG         | 61 |
| 212         | NAT8L| ACTACTACGCGCAAGGTG  | TCGGAAACGTTGAGTCACAG           | 61 |
| 236         | MMP16| ACCGACTGACCCCAAGATGT| TGATGTGCTTGCTGGCAT            | 61 |
Supplementary data Table 3. Primer list

| Target size | Gene   | Forward       | Backward                     | TM |
|-------------|--------|---------------|------------------------------|----|
| 124         | DUSP6  | CGACTGGAACGAGAATACGGG | GGAGAACTCGGGCTTGGAACCTT | 61 |
| 271         | IRX2   | GCCCTTCTACGGCAACTACA | TGCTCGGCCCCCTCTATAGGT | 61 |
| 279         | ATP2B1 | CAGGTCCACAGATGCATTACGA | TCCCCCTTCTGGAGGCTGATAAA | 61 |
| 207         | VSNL1  | CCAAGTGGAGGCCTAATCTCG | GAAGTCAATGGTGCCGTCCC | 61 |
| 245         | ADRB2  | TGATCGCAGTGGATCGCTAC | GGACACGATGGAAGGCAA | 61 |
| 161         | PLXDC2 | TTCAGTTCCGCGATGGGAAA | GTTCGTGTCTACCACCTTGA | 61 |
| 168         | PLD5   | CATTCTCTGCCTCTGAAGGCT | TCAGTGCAACCAGAATGGCA | 61 |
| 244         | TPD52L1 | TTTGCTCCCTTTCTGAGCTA | TGCAACGTTTCATCTCCAA | 61 |
| 247         | FAM84A | CGTCAGCGCCTCATACCTTTCAT | AGCCCTGTTTTTCTCCTGG | 61 |
| 172         | SNX19  | GCGATGAGTCCTCAATCAGGT | CCCGATAGCGACGATTCA | 61 |

**Supplementary data Table 3. Primers.** Forward and Backward primer sequences for every 10 up- and down-regulated genes. TM are matched to 61 degrees.