

CLAIMS

1. A method for prediction of breast cancer prognosis, comprising:
 - allowing a computer device to use RNA sequencing data of a subject's tissue, to measure expression levels of target genes;
 - wherein the computer device has, as input values, expression levels of target genes for a plurality of samples and has output values by inputting the expression levels of the target genes to an artificial neural network depending on Oncotype DX recurrence scores for the plurality of samples,
 - wherein the target genes include all of the following 149 genes:

ACTB, APOBEC3B, ASF1B, ASPM, AURKA, AURKB, BAG1, BCL2, BIRC5, BLM, BUB1, BUB1B, C14orf45, C16orf61, C7orf63, CACNA1D, CCNA2, CCNB1, CCNB2, CCNE1, CCNE2, CCT5, CD68, CDC20, CDC25A, CDC45, CDC6, CDCA3, CDCA8, CDK1, CDKN3, CENPA, CENPE, CENPF, CENPM, CENPN, CEP55, CHEK1, CIRBP, CKS2, CRIM1, CTSL2, CX3CR1, CYBRD1, DBF4, DDX39, DLGAP5, DNMT3B, DONSON, DTL, E2F1, E2F8, ECHDC2, ERBB2, ERCC6L, ESPL1, ESR1, EXO1, EZH2, FAM64A, FANCI, FBXO5, FEN1, FOXM1, GAPDH, GINS1, GRB7, GSTM1, GTSE1, GUSB, HJURP, HMMR, HN1, IFT46, KIF11, KIF13B, KIF14, KIF15, KIF18A, KIF18B, KIF20A, KIF23, KIF2C, KIF4A, KIFC1, KPNA2, LMNB1, LMNB2, LRIG1, LRRC48, LRRC59, MAD2L1, MARCH8, MCM10, MCM2, MCM6, MELK, MKI67, MLF1IP, MMP11, MYBL2, NCAPG, NCAPG2, NCAPH, NDC80, NEK2, NUP93, NUSAP1, OIP5, PBK, PDSS1, PGR, PKMYT1, PLK1, PLK4, PRC1, PTTG1, RACGAP1, RAD51, RAD51AP1, RAI2, RFC4, RPLP0, RRM2, SCUBE2, SETBP, SF3B3, SHCBP1, SHMT2, SLC25A1, SLC7A5, SPAG5, SPC25, SQLE, STARD13, STIL, STMN1, SYNC, TACC3, TFRC, TK1, TOP2A, TPX2, TRIP13, TROAP, TTK, UBE2C, UBE2S, and ZWINT,
 - wherein the subject is estrogen receptor positive.
2. The method according to claim 1,
 - wherein the computer device measures the expression levels of the target genes with a next-generation sequencing (NGS) technique.

3. The method according to claim 1,
wherein the target genes include 16 genes used for calculating the Oncotype DX recurrence scores and additional genes of which a correlation with respect to the Oncotype DX recurrence scores is equal to or higher than a reference value, and
the genes are selected from the genes with a correlation coefficient equal to or higher than the reference value by performing correlation analysis with the Oncotype DX recurrence scores for each of the sample genes, which have no lymph node metastasis.
4. The method according to claim 1,
wherein the RNA sequencing data includes a conserved exon region that is commonly expressed in the plurality of sample data.
5. The method according to claim 1,
wherein the computer device normalizes the expression levels of the target genes using Trimmed Mean of M-value (TMM) technique.
6. The method according to claim 1,
wherein the computer device estimates the prognosis that the subject is at high risk in a case where the output value is equal to or higher than a reference value, and estimates the prognosis that the subject is at low risk in a case where the output value is lower than the reference value.
7. A computer-readable recording medium having recorded therein a program for executing, on a computer, the method for prediction of breast cancer prognosis according to any one of claims 1 to 6.
8. A system for prediction of breast cancer prognosis, comprising:
a client device for storing RNA gene data of a subject's tissue; and
an analysis server in which RNA sequencing data, which is obtained by performing next-generation sequencing (NGS)-based sequencing on the gene data, is used to measure expression levels of target genes, and breast cancer prognosis for the

subject is estimated; and

wherein the analysis server has, as input values, expression levels of target genes for a plurality of samples and has output values by inputting the expression levels of the target genes to an artificial neural network depending on Oncotype DX recurrence scores for the plurality of samples,

wherein the target genes include all of the following 149 genes:

ACTB, APOBEC3B, ASF1B, ASPM, AURKA, AURKB, BAG1, BCL2, BIRC5, BLM, BUB1, BUB1B, C14orf45, C16orf61, C7orf63, CACNA1D, CCNA2, CCNB1, CCNB2, CCNE1, CCNE2, CCT5, CD68, CDC20, CDC25A, CDC45, CDC6, CDCA3, CDCA8, CDK1, CDKN3, CENPA, CENPE, CENPF, CENPM, CENPN, CEP55, CHEK1, CIRBP, CKS2, CRIM1, CTSL2, CX3CR1, CYBRD1, DBF4, DDX39, DLGAP5, DNMT3B, DONSON, DTL, E2F1, E2F8, ECHDC2, ERBB2, ERCC6L, ESPL1, ESR1, EXO1, EZH2, FAM64A, FANCI, FBXO5, FEN1, FOXM1, GAPDH, GINS1, GRB7, GSTM1, GTSE1, GUSB, HJURP, HMMR, HN1, IFT46, KIF11, KIF13B, KIF14, KIF15, KIF18A, KIF18B, KIF20A, KIF23, KIF2C, KIF4A, KIFC1, KPNA2, LMNB1, LMNB2, LRIG1, LRRC48, LRRC59, MAD2L1, MARCH8, MCM10, MCM2, MCM6, MELK, MKI67, MLF1IP, MMP11, MYBL2, NCAPG, NCAPG2, NCAPH, NDC80, NEK2, NUP93, NUSAP1, OIP5, PBK, PDSS1, PGR, PKMYT1, PLK1, PLK4, PRC1, PTTG1, RACGAP1, RAD51, RAD51AP1, RAI2, RFC4, RPLP0, RRM2, SCUBE2, SETBP, SF3B3, SHCBP1, SHMT2, SLC25A1, SLC7A5, SPAG5, SPC25, SQLE, STARD13, STIL, STMN1, SYNC, TACC3, TFRC, TK1, TOP2A, TPX2, TRIP13, TROAP, TTK, UBE2C, UBE2S, and ZWINT,

wherein the subject is estrogen receptor positive.

9. The system according to claim 8,

wherein the target genes include genes that have a correlation coefficient equal to or higher than a reference value in a case where correlation analysis with the Oncotype DX recurrence scores is performed for each of the sample genes which have no lymph node metastasis.

10. The system according to claim 8,

wherein the RNA sequencing data includes a conserved exon region that is

commonly expressed in the plurality of sample data.

11. The system according to claim 8,
wherein the analysis server normalizes the expression levels of the target genes
using Trimmed Mean of M-value (TMM) technique.