survival time of the good prognostic group was 12.9 years, while the median survival time of the poor prognostic group was 2.4 years. Error due to unreliable survival information could be minimized by clearly dividing the poor prognostic group and the good prognostic group. The expression or not of the estrogen receptor (ER) is the most commonly used criterion when breast cancer patients are classified by subtypes. In clinical practice, the breast cancer patients are classified into ER+ and ER- through ER immunohistochemistry (IHC) readout results by pathologists. Considering that about 200 patients had no ER IHC information in the collected discovery dataset, and the determination of ER IHC was made for each of five datasets constituting the discovery dataset, independently, the ER status was determined using the mRNA expression level of ESR1 gene in the expression profile for each patient. For the patients having ER IHC information, region of convergence (ROC) analysis was conducted using the ER IHC information and the ERS1 mRNA expression level. Through a comparison between the ER IHC results and the ERS1 mRNA expression level, the cutoff for ER status was determined at the point where the accuracy was highest (accuracy=0.88). The cases showing higher expression levels than the cutoff were classified as ER+, and the cases showing lower expression levels than the cutoff were classified as ER-. In the discovery dataset, 864 patients were designated as ER+ and 240 patients as ER-.

Selection of prognostic genes

The good prognosis group was defined as ER+, and the poor prognosis group was defined as ER- in the discovery dataset. 275 patients had good prognosis, while 218 patients had poor prognosis. Through significant analysis of microarray (SAM), genes showing a difference in the expression level among the prognostic groups were investigated. By using the q-value of the SAM analysis results, overexpressed genes in the good prognosis group and the poor prognosis group were selected. As a result of combining the selected genes, a total of 302 gene

Claims

1. A method for predicting the prognosis of breast cancer, the method comprising:

(a) isolating mRNA from a sample of a tissue containing cancer cells of a breast cancer patient;

(b) measuring the mRNA expression level of TRBC1 (T cell receptor beta constant 1);

(c) normalizing the measured mRNA expression level of the TRBC1 gene; and

(d) determining the overexpression of the TRBC1 gene as indicating good prognosis if expression of TRBC1 is significantly increased in a good prognostic group wherein metastasis does not occur within five years.

2. The method of claim 1, wherein the measuring of the expression level is conducted through PCR amplification of a target gene.

3. The method of claim 1, wherein the sample is a formalin-fixed paraffin-embedded (FFPE) sample of a tissue containing cancer cells of a patient.

4. The method of claim 1, wherein the normalizing is conducted by calculating the ratio of the mean expression level compared with at least one standard gene selected from the group consisting of CTBP1 (C-terminal-binding protein 1), TBP (TATA-binding protein), HMBS (hydroxymethylbilane synthase), CUL1 (cullin 1), and UBQLN1 (ubiquilin-1).

5. A method for calculating a predictive value of the prognosis of breast cancer to provide information necessary for predicting or diagnosing the prognosis of a breast cancer patient, the method comprising:

(a) isolating mRNA from a sample of a tissue containing cancer cells of a breast cancer patient;

(b) measuring the mRNA expression level of TRBC1 (T cell receptor beta constant 1);

(c) normalizing the measured mRNA expression level of the TRBC1 gene;

(d) inserting the normalized value into a pre-determined calculation formula to calculate a predictive value; and

(e) determining the prognosis of breast cancer as being good or poor depending on the predictive value,

wherein prognosis is poor if metastasis occurs within five year and prognosis is good if metastasis does not occur within five years.

6. The method of claim 5, wherein the measuring of the expression level is conducted through PCR amplification of a target gene.

7. The method of claim 5, wherein the sample is a formalin-fixed paraffin-embedded (FFPE) sample of a tissue containing cancer cells of a patient.

8. The method of claim 5, wherein the normalizing is conducted by calculating the ratio of the mean expression level compared with at least one standard gene selected from the group consisting of CTBP1 (C-terminal-binding protein 1), TBP (TATA-binding protein), HMBS (hydroxymethylbilane synthase), CUL1 (cullin 1), and UBQLN1 (ubiquilin-1).

9. A method for predicting or diagnosing the prognosis of breast cancer, the method comprising a step of using a primer pair to measure a mRNA expression level of a selected gene from a sample of a breast cancer patient, wherein the primer pair is for TRBC1 (T cell receptor beta constant 1), and wherein the primer pair is capable of amplifying a target gene through PCR amplification.

10. Use of a primer pair as an active agent in the

manufacture of a pharmaceutical composition suitable for predicting the prognosis of breast cancer, wherein the primer pair is TRBC1 (T cell receptor beta constant 1), and wherein the primer pair is capable of amplifying a target gene through PCR amplification.

11. A kit suitable for predicting the prognosis of breast cancer, the kit comprising the composition of claim 10.

12. A method for calculating a predictive value of the prognosis of breast cancer to provide information necessary for predicting or diagnosing the prognosis of a breast cancer patient, the method comprising:

(a) isolating mRNA from a sample of a tissue containing cancer cells of a breast cancer patient;

(b) measuring the mRNA expression level of TRBC1 (T cell receptor beta constant 1), and the mRNA expression level of at least one gene selected from the p-gene group consisting of AURKA (Aurora Kinase A), CCNB2 (Cyclin B2), FOXM1 (Forkhead box protein M1), MMP11 (Matrix Metallopeptidase 11), PTTG1 (Pituitary Tumor-Transforming 1), RACGAP1 (Rac GTPase Activating Protein 1), RRM2 (Ribonucleotide Reductase M2), TOP2A (Topoisomerase II Alpha) and UBE2C (Ubiquitin-Conjugating Enzyme E2C);

(c) normalizing the measured mRNA expression levels of the TRBC1 gene and the at least one p-gene;

(d) inserting the normalized values of the TRBC1 gene and the at least one p-gene into a pre-determined calculation formula to calculate a predictive value of the prognosis of breast cancer; and

(e) determining the prognosis of breast cancer as being good or poor depending on the predictive value,

wherein prognosis is poor if metastasis occurs within five year and prognosis is good if metastasis does not occur within five years.