

What is claimed is,

1. A set of molecular markers consisting of *BCL2*, *PGR*, *SCUBE2*, *ESR1*, *MKi67*, *CCNB1*, *MYBL2*, *GRB7*, *ERBB2*, *MMP11*, *CD68*, *BAG1*, *MAPT* and *MS4A1*.

5

2. The set of molecular markers according to claim 1, further comprising internal reference genes *ACTB*, *GAPDH*, *GUSB*, *NUP214* and *VCAN*.

3. Use of the set of molecular markers according to claim 2 for risk prediction of recurrence and/or mortality in 3-10 years after breast cancer surgery.

10

4. The use according to claim 3, wherein the risk prediction of recurrence and/or mortality in 3-10 years after breast cancer surgery is performed by:

obtaining total RNA of a test sample, obtaining cDNA by reverse transcription, obtaining Ct values of the set of molecular markers and the internal reference genes by a fluorescent quantitative PCR, calculating average Ct value of the internal reference genes to obtain an average Ct value (Ct') of internal reference gene combination, normalizing Ct value of the set of molecular markers respectively by subtracting Ct' value of the internal reference gene combination to obtain ΔCt , and then analyzing ΔCt value, and age, pT value and LN value of the subject by a model for risk prediction of recurrence or mortality in 3-10 years after breast cancer surgery constructed by random forest algorithm to obtain a result.

15

20

5. The use according to claim 4, wherein the constructing of the model for risk prediction of recurrence or mortality in 3-10 years after breast cancer surgery is performed by:

constructing a matrix by using ΔCt value of the set of molecular markers of the test sample, and age, pT value and LN value of the subject, randomly selecting half as a training set and another half as a verification set, establishing a prediction model containing 10,000 decision trees by

25

random forest algorithm, randomly sampling for ≥ 1000 times and establishing ≥ 1000 prediction models, selecting ≥ 39 optimal models having the highest concordance rate with follow-up information from the ≥ 1000 prediction models as submodels of final model, and setting the median of the ≥ 39 submodels as a final prediction value of prognostic risk.

5

6. The use according to claim 4 or 5, wherein the test sample is a FFPE sample from an untreated early or intermediate stage ER or PR positive breast cancer newly diagnosed patient.

7. A primer set for amplifying the set of molecular markers according to claim 1, having sequences represented by SEQ ID NO: 1 to SEQ ID NO: 28.

10

8. A probe set for amplifying the set of molecular markers according to claim 1, having sequences represented by SEQ ID NO: 29 to SEQ ID NO: 42.

9. A test kit for risk prediction of recurrence and/or mortality in 3-10 years after breast cancer surgery, comprising the primer set according to claim 7 and/or the probe set according to claim 8, and reagents commonly used in a kit.

15

10. A method for constructing the model for risk prediction of recurrence or mortality in 3-10 years after breast cancer surgery, comprising:

20

constructing a matrix by using ΔC_t value of a set of molecular markers according to claim 1 of a test sample, and age, pT value and LN value of the subject, randomly selecting half as a training set and another half as a verification set, establishing a prediction model containing 10,000 decision trees by random forest algorithm, randomly sampling for ≥ 1000 times and establishing ≥ 1000 prediction models, selecting ≥ 39 optimal models having the highest concordance rate with follow-up information from the ≥ 1000 prediction models as submodels of final model, and setting the median of the ≥ 39 submodels as a final prediction value of prognostic risk.

25

11. A method for risk prediction of recurrence or mortality in 3-10 years after breast cancer surgery, comprising:

5 obtaining total RNA of a test sample, obtaining cDNA by reverse transcription, obtaining Ct values of a set of molecular markers according to claim 1 and internal reference genes according to claim 2 by a fluorescent quantitative PCR, calculating average Ct value of the internal reference genes to obtain an average Ct value (Ct') of internal reference gene combination, normalizing Ct value of the set of molecular markers respectively by subtracting Ct' value of the internal reference gene combination to obtain Δ Ct, and then analyzing Δ Ct value, and age, pT value and LN value
10 of the subject by a model for risk prediction of recurrence or mortality in 3-10 years after breast cancer surgery constructed by random forest algorithm to obtain the result.

12. The method according to claim 10 or 11, wherein the test sample is a FFPE sample from an untreated early or intermediate stage ER or PR positive breast cancer newly diagnosed patient.