INTRODUCTION

Breast cancer is the most frequent malignancy and the leading cause of cancer-associated mortality in women worldwide.¹,² In clinical practice, individual specific targeted therapy has attracted more and more attention. Thus, exploring potential prognostic biomarkers and promising targets is regarded to be an essential step to achieve this goal.

Autophagy is a highly conserved process to maintain cellular homeostasis by lysosomal degradation system.³ Autophagy plays
a crucial role in many physiological processes and various pathological events, including stress and starvation adaptation, metabolism, inflammation, neurodegenerative disorders and cancer.\textsuperscript{4–7} Over the past few years, an increasing number of studies have indicated that autophagy participates in the development and progression of breast cancer.\textsuperscript{8,9} Therefore, identifying key regulators of autophagy is of great importance for both theoretical basis and clinical practice.

Long non-coding RNAs (lncRNAs) are a series of transcript RNAs longer than 200 nucleotides without the capacity of protein-coding.\textsuperscript{10} LncRNAs are considered as one of the most sensitive and specific cancer biomarkers, which participate in the development and progression of various cancers at different levels including epigenetic, transcriptional and post-transcriptional regulation.\textsuperscript{11–14} Moreover, accumulating studies have suggested that lncRNAs promote cancer progression and predict worse prognosis in numerous cancers via regulating autophagy.\textsuperscript{15–17} Therefore, it is valuable to identify key lncRNAs closely related to autophagy and prognosis in breast cancer.

In the present study, we analysed a data set of lncRNA expression in breast cancers from The Cancer Genome Atlas (TCGA) and screened out autophagy-related lncRNAs with prognostic value. We identified an eleven autophagy-related lncRNA signature with the potential to predict the survival prognosis of breast cancer patients.

2 METHODS AND MATERIALS

2.1 Patient data sets

Breast cancer patients with clinical information and pathology records were obtained from the TCGA (https://cancergenome.nih.gov/). Normalize gene expression was performed by the edgeR package. In this study, a total of 1053 TCGA female breast cancer patients with lncRNA expression profiles were used. Among them, 986 patients with complete follow-up information and survival time ≥ 30 days and 539 patients with complete clinicopathological data were selected into subsequent analyses. The clinical features are detailed in Table 1.

2.2 Identification of autophagy-related lncRNAs in breast cancer

A total of 395 autophagy-related encoding genes (mRNAs) were extracted from the Molecular Signatures Database of Gene Set Enrichment Analysis (GSEA; M27935, M6328 and M10281). Finally, 912 autophagy-related lncRNAs were identified by constructing autophagy-related mRNA-lncRNA co-expression network according to the criteria of |Correlation Coefficient| > 0.4 and \( P < .001 \) by Pearson correlation analysis using the Limma R package.\textsuperscript{18}

2.3 Identification of autophagy-related lncRNA prognostic signatures for breast cancer

To identify autophagy-related lncRNAs associated with survival, we performed univariate Cox proportional hazards analysis according to the criteria of \( P < .01 \). Subsequently, multivariate Cox analysis was conducted to construct the optimal prognostic risk model based on the Akaike information criterion (AIC = 1444.62), using the survival R package. Based on the following formula, the risk score for each patient was calculated.

\[
\text{Risk score} = \text{coef (lncRNA1)} \times \text{expr (lncRNA1)} + \text{coef (lncRNA2)} \times \text{expr (lncRNA2)} + \ldots + \text{coef (lncRNA11)} \times \text{expr (lncRNA11)}.
\]

\( \text{coef (lncRNA11)} \) was defined as the coefficient of lncRNAs correlated with survival.

\( \text{expr (lncRNA11)} \) was defined as the expression of lncRNAs.

Based on the median risk score, breast cancer patients in the TCGA were divided into a high-risk group and a low-risk group. Kaplan-Meier survival analysis was performed to estimate the survival difference between the two groups by using the survival and survminer R packages.

2.4 Independent prognostic analysis and ROC curve plotting

To assess the relationship of survival prognosis with clinicopathological factors and risk score, we performed univariate and multivariate Cox regression analyses using the Survival R package. Time-dependent receiver operating characteristic (ROC) curves were drew to estimate the predictive accuracy for survival time by different clinical pathological factors and risk score using the survivalROC R package.

2.5 Statistical analysis

All statistical analyses were performed using R software (version 3.6.2). A co-expression network of the 11 autophagy-related lncRNAs-mRNAs with prognostic value was established and visualized using Cytoscape and Sankey diagram. The correlation between 11 autophagy-related lncRNA expressions and clinicopathological factors was analysed by ggpubr R package. Principal component analysis (PCA) was performed for effective dimension reduction, pattern recognition and exploratory visualization of high-dimensional data of the whole-genome, 395 autophagy-related encoding genes and the risk model of the 11 autophagy-related lncRNA expression profiles, respectively.\textsuperscript{19,20} Gene Set Enrichment Analysis (GSEA) was used for functional annotation. GSEA (https://www.gsea-msigdb.org/gsea/index.jsp) is a powerful analytical approach for interpreting genome-wide expression profiles.\textsuperscript{21} GSEA focuses on gene
sets rather than just high scoring genes, which can detect biological processes such as several cancer-related pathways, metabolic pathways, transcriptional programmes and stress responses. GSEA tends to be more reproducible and more interpretable to analyse molecular profiling data. Two-tailed $P < .05$ was considered statistically significant.

### RESULTS

#### 3.1 Identification of autophagy-related lncRNAs with significant prognostic value in breast cancer

A total of 912 autophagy-related lncRNAs were obtained by constructing the co-expression networks with 395 autophagy-related encoding genes (mRNAs). Among them, 24 autophagy-related lncRNAs were significantly associated with the survival of breast cancer patients from the TCGA ($P < .01$) by Cox proportional hazards analysis, including 19 lncRNAs with low risk (hazard ration (HR)<1) and 5 lncRNAs with high risk (hazard ration (HR)>1) (Figure 1A). Subsequently, multivariate Cox analysis further screened 11 lncRNAs from the above 24 autophagy-related lncRNAs with prognostic significance, namely, U62317.4, LINC01016, LINC02166, C6orf99, LINC00992, BAIAP2-DT, AC245297.3, AC090912.1, Z68871.1, LINC00578 and LINC01871 (Table 2). These 11 lncRNAs were constituted into the optimal prognostic risk model of autophagy-related lncRNAs. As shown in Figure 1B-C, A visualization co-expression network of the 11 autophagy-related lncRNAs-mRNAs with prognostic value was established. On the basis of the risk score formula and the calculated median risk score, breast cancer patients were divided into a high-risk group and a low-risk group. Kaplan-Meier survival analysis showed that the high-risk group displayed worse overall survival (OS) than the low-risk group ($P = 3.048 \times 10^{-10}$) (Figure 2A), indicating that the risk score has prognostic value. The risk curve and scatterplot were applied to illustrate the risk score and the relevant survival statuses of breast cancer patients. The results indicated that the mortality occurrence depended on the risk score (Figure 2B-C). The heatmap of these 11 autophagy-related lncRNA expressions in breast cancer samples showed that C6orf99, LINC00992, Z68871.1 and LINC00578 were highly expressed in the high-risk group, while U62317.4, LINC01016, LINC02166, BAIAP2-DT, AC245297.3, AC090912.1 and LINC01871 were up-regulated in the low-risk group (Figure 2D). Therefore, these above studies identified 11 autophagy-related lncRNAs with prognostic significance for breast cancer.

#### 3.2 Evaluation of the risk model of 11 autophagy-related lncRNAs as an independent prognostic factor for breast cancer patients

To evaluate whether the risk model of the above 11 autophagy-related lncRNAs is an independent prognostic factor for breast cancer, univariate and multivariate Cox regression analyses were conducted. The hazard ratio (HR) of the risk score and 95% CI were 1.507 and 1.308-1.736 ($P < .001$) in univariate Cox regression analysis (Figure 3A), and 1.489 and 1.280-1.732 ($P < .001$) in multivariate Cox regression analysis (Figure 3B), respectively, suggesting that the risk model of the 11 autophagy-related lncRNAs is the most significant prognostic factor for breast cancer, independent of clinicopathological parameters such as oestrogen receptor (ER) expression; progesterone receptor (PR)

---

### TABLE 1 Clinical pathological parameters of patients with breast cancer

| Feature                      | N (539) | %  |
|------------------------------|---------|----|
| Age (years)                  |         |    |
| >60                          | 227     | 42.1|
| ≤60                          | 312     | 57.9|
| T classification             |         |    |
| T1 (<2 cm)                   | 147     | 27.3|
| T2 (2-5 cm)                  | 323     | 59.9|
| T3 (≥5 cm)                   | 55      | 10.2|
| T4 (chest wall and/or skin invasion) | 14 | 2.6 |
| N classification (pN)        |         |    |
| N0 (no metastasis)           | 259     | 48.1|
| N1 (1-3 metastasis)          | 178     | 33  |
| N2 (4-9 metastasis)          | 64      | 11.9|
| N3 (≥10 metastasis)          | 38      | 7   |
| M classification             |         |    |
| M0 (no distant metastasis)   | 528     | 98  |
| M1 (distant metastasis)      | 11      | 2   |
| TNM stage                    |         |    |
| I                            | 96      | 17.8|
| II                           | 318     | 59  |
| III                          | 114     | 21.2|
| IV                           | 11      | 2   |
| ER                           |         |    |
| Negative                     | 127     | 23.6|
| Positive                     | 412     | 76.4|
| PR                           |         |    |
| Negative                     | 175     | 32.5|
| Positive                     | 364     | 67.5|
| HER2                         |         |    |
| Negative                     | 440     | 81.6|
| Positive                     | 99      | 18.4|
| Molecular subtypes           |         |    |
| HER2 amplification           | 92      | 17.1|
| Luminal A/B                  | 419     | 77.7|
| TNBC                         | 28      | 5.2 |

Note: T, tumour size; N, lymph node; M, distant metastasis; TNM stage, according to AJCC 8th classification; TNBC, triple-negative breast cancer.
expression; human epidermal growth factor receptor (HER2) expression; tumour, node, metastasis (TNM) stage; tumour size; lymph node metastasis; and distant metastasis. To assess the predictive sensitivity and specificity of the risk score on the prognosis of breast cancer patients, the area under the ROC curve (AUC) of the risk score was calculated. The AUC of the risk score was 0.774, similar to the AUC of tumour size, followed by the AUC of age and more than the AUCs of other clinicopathological factors (Figure 3C), indicating that the prognostic risk model of the 11 autophagy-related lncRNAs for breast cancer is considerably reliable. Taken together, these all indicated that

| LncRNA    | Coef  | HR   | HR.95L | HR.95H | P-value | Risk |
|-----------|-------|------|--------|--------|---------|------|
| U62317.4  | −0.308| 0.735| 0.488  | 1.106  | .140    | low  |
| LINCO1016 | −0.173| 0.842| 0.678  | 1.045  | .119    | low  |
| LINCO2166 | −0.596| 0.551| 0.292  | 1.041  | .066    | low  |
| C6orf99   | 0.284 | 1.329| 1.017  | 1.738  | .038    | high |
| LINCO0992 | 0.188 | 1.207| 0.988  | 1.473  | .065    | high |
| BAIAP2-DT | −0.385| 0.680| 0.525  | 0.882  | .004    | low  |
| AC245297.3| −0.381| 0.683| 0.520  | 0.897  | .006    | low  |
| ACO90912.1| −0.475| 0.622| 0.342  | 1.131  | .119    | low  |
| Z68871.1  | 0.573 | 1.774| 1.133  | 2.778  | .012    | high |
| LINCO0578 | 0.146 | 1.158| 0.979  | 1.369  | .087    | high |
| LINCO1871 | −0.417| 0.659| 0.495  | 0.878  | .004    | low  |

Note: coef: the coefficient of lncRNAs correlated with survival; HR: hazard ratio; HR.95L: low 95% CI of HR; HR.95H: high 95% CI of HR.

FIGURE 2 The prognostic value of the risk model of the 11 autophagy-related lncRNAs in the TCGA cohort. A, Kaplan-Meier survival analysis of the high-risk and low-risk groups based on the risk model and median risk score. B, The risk curve based on the risk score of each sample. C, The scatterplot based on the survival status of each sample. The green and red dots represent survival and death, respectively. D, The heatmap displayed the expression levels of autophagy-related lncRNAs in the high-risk and low-risk groups.
3.3 Correlation of the expression of the 11 autophagy-related lncRNAs with clinicopathological factors

To further assess whether the 11 autophagy-related lncRNAs participated in the development of breast cancer, we investigated the association of the expression of the 11 autophagy-related lncRNAs with clinicopathological factors. There were remarkably correlations between a majority of the 11 autophagy-related lncRNAs and ER expression, PR expression, HER2 expression and molecular subtypes, as shown in Figure 4.

3.4 Different autophagy statuses in the low-risk and high-risk groups

PCA was performed to compare the difference between low-risk and high-risk groups based on the risk model of the 11 autophagy-related lncRNAs, 395 autophagy-related encoding genes and the whole-genome expression profiles from the TCGA, respectively (Figure 5). The results showed that the low-risk and high-risk groups were distributed in distinct directions, prior to the other two patterns, suggesting that the risk model could divide breast cancer patients into two parts and that the autophagy status of breast cancer patients in the high-risk group differed from those in the low-risk group. Functional annotation was further conducted using GSEA, and the results displayed that the differentially expressed genes between the high-risk and low-risk groups based on the risk model.
of the 11 autophagy-related genes were enrichment in autophagy processes and oncogenic signatures (Figure 6).

Above these, it all indicated that the low-risk and high-risk groups showed different autophagy and oncogenic statuses.

4 | DISCUSSION

In the field of clinical treatment, although the overall survival of breast cancer patients has gained great improvements, metastasis and recurrence of breast cancer have constantly grown, which were the major causes of breast cancer mortality. Considerable evidence has revealed that the role of autophagy in the development of cancer is a double-edged sword. That is, autophagy may serve as either a pro-survival or pro-death mechanism under different circumstances. Consistent with the theory, autophagy functioned dual roles in breast cancer. Autophagy induced the recurrence of metastatic breast cancer by elevating the survival of dormant breast cancer cells. Cytostatic autophagy suppressed the proliferation and metastasis of triple-negative breast cancer cells. Furthermore, numerous researches indicated the crucial role of IncRNAs in autophagy-inducing progression or inhibition of various cancers, such as hepatocellular carcinoma, lung cancer and breast cancer. Based on the above, it led us to find potential specific IncRNAs associated with autophagy and the survival prognosis. In this study, we identified the risk model of the 11 autophagy-related IncRNAs as an independent prognostic factor for breast cancer. So far, among these 11 autophagy-related IncRNAs, only LINC01016 and LINC00578 have been studied in breast cancer or other cancers. LINC01016-miR-302a-3p/miR-3130-3p/NFYA/SATB1 axis plays a crucial role in the occurrence of endometrial cancer. In addition, LINC01016 is a direct target of ERα, associated with survival prognosis of breast cancer. It has been reported that LINC00578 is associated with OS in pancreatic cancer and lung adenocarcinoma. In accordance with our results, LINC01016 was low-risk autophagy-related IncRNA and LINC00578 was high-risk autophagy-related IncRNA, both with prognostic value in breast cancer patients. Breast cancer is a highly heterogeneous tumour and has been divided into distinct molecular subtypes as follows: luminal A/B (ER and/or PR positive), HER2 enriched (HER2 positive) and triple-negative breast cancer (ER, PR and HER2 negative). Distinct ER, PR and HER2 statuses and molecular subtypes suggested different biological processes of breast cancer and survival outcomes. The significant correlations of most of the 11 autophagy-related IncRNAs with ER expression, PR expression, HER2 expression and molecular subtypes indicated that the 11 autophagy-related IncRNAs might be involved in regulating ER, PR and HER2 gene expression, further affecting molecular subtype formation.
More interestingly, most of the 11 autophagy-related lncRNAs had no significant association with tumour size, lymph node status and TNM stage, implying that there is no close relationship between the timing of breast cancer diagnosis and the risk model of the 11 autophagy-related lncRNAs. All above results indicated that the risk model strongly links to the intrinsic biological characteristics of distinct subtypes of breast cancer. Our results also showed that the risk model of the 11 autophagy-related lncRNAs had superior prognostic value to other clinicopathological factors. The results of PCA and GSEA indicated that the significant difference in OS between the high-risk and low-risk groups might result from different autophagy and oncogenic statuses induced by the risk model. These results also suggested that the risk model of the 11 autophagy-related lncRNAs had superior prognostic value to other clinicopathological factors. The results of PCA and GSEA indicated that the significant difference in OS between the high-risk and low-risk groups might result from different autophagy and oncogenic statuses induced by the risk model. These results also suggested that the risk model of the 11 autophagy-related lncRNAs represented more significant autophagy characteristics than autophagy-related genes. Taken together, these results suggested that the prognostic risk model of the 11 autophagy-related lncRNAs might be a feasible independent prognostic factor for breast cancer in clinical practice.

To date, the focus of precision genomic medicine is to find out accurate specific predictive factors for survival prognosis from large medical data sets with clinical outcomes. Thus, there have been some researches aiming to explore autophagy-related prognostic factors using bioinformatics analysis in recent years. Over the past year, three distinct prognostic risk models of autophagy-related encoding genes in breast cancer have been established based on TCGA database by using different screening criteria and statistics methods. Meanwhile, due to the crucial function of lncRNAs in autophagy, autophagy-related lncRNAs in cancer also arouse more attention. Recently, the prognostic risk models of autophagy-related lncRNAs were constructed in several cancers including bladder urothelial carcinoma and glioma. However, the role of autophagy-related lncRNAs in prognosis of breast cancer remains unclear. We therefore conducted this research and found out a novel eleven autophagy-related lncRNA prognostic risk model, which may assist clinicians in making individual effective therapeutic decisions.

However, our study has some limitations. First, we applied traditional statistical analyses to build and evaluate the prognostic risk model of 11 autophagy-related lncRNAs. Although the methods have been utilized and validated in many researches, it is critical to improve our further studies with more advanced methodologies and technologies in the future. To further verify our bioinformatics prediction results, in-depth studies on the 11 autophagy-related lncRNAs, including functional experiments and molecular mechanisms, are needed.

5 | CONCLUSION

In conclusion, we identified a novel autophagy-related prognostic risk model consisting of 11 lncRNAs (U62317.4, LINC01016, LINC02166, C6orf99, LINC00992, BAIAP2-DT, AC245297.3, FIGURE 6 Functional enrichment analysis based on the risk model of the 11 autophagy-related IncRNAs by GSEA. Significantly enriched autophagy of Gene Ontology (GO), KEGG pathways and oncogenic signatures in the low-risk group.
AC090912.1, Z68871.1, LINCC00578 and LINCC1871) in breast cancer. In the future, with prospective validation, the 11 autophagy-related lncRNAs signature may improve predictive accuracy and guide individualized therapy for breast cancer patients.

ACKNOWLEDGEMENTS
The data of this study were downloaded from The Cancer Genome Atlas (TCGA), and we gratefully acknowledge the patients and operations. This study was supported by grants from the National Natural Science Foundation of China (No. 81773163, 81602564).

CONFLICTS OF INTEREST
The authors declare no conflicts of interest.

AUTHOR CONTRIBUTION
Xiaoying Li: Conceptualization (equal); Data curation (equal); Formal analysis (lead); Writing-original draft (lead); Writing-review & editing (equal). Feng Jin: Conceptualization (equal); Funding acquisition (equal); Supervision (lead); Writing-review & editing (lead). Yang Li: Conceptualization (equal); Funding acquisition (lead); Supervision (lead); Writing-review & editing (lead).

DATA AVAILABILITY STATEMENT
All data utilized in this study are included in this article, and all data supporting the findings of this study are available on reasonable request from the corresponding author.

ORCID
Xiaoying Li https://orcid.org/0000-0002-4569-5649

REFERENCES
1. DeSantis CE, Ma J, Gaudet MM, et al. Breast cancer statistics, 2019. CA Cancer J Clin. 2019;69(6):438-451.
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.
3. Jiang P, Mizushima N. Autophagy and human diseases. Cell Res. 2014;24(1):69-79.
4. Santana-Codina N, Mancias JD, Kimmelman AC. The Role of Autophagy in Cancer. Annu Rev Cancer Biol. 2017;1:19-39.
5. Janku F, McConkey DJ, Hong DS, Kurzrock R. Autophagy as a target for anticancer therapy. Nat Rev Clin Oncol. 2011;8(9):528-539.
6. Nah J, Yuan J, Jung YK. Autophagy in neurodegenerative diseases: from mechanism to therapeutic approach. Mol Cells. 2015;38(5):381-389.
7. Roos WP, Thomas AD, Kaina B. DNA damage and the balance between survival and death in cancer biology. Nat Rev Cancer. 2016;16(1):20-33.
8. Vera-Ramirez L, Vodanaka SK, Nini R, Hunter KW, Green JE. Autophagy promotes the survival of dormant breast cancer cells and metastatic tumour recurrence. Nat Commun. 2018;9(1):1944.
9. Chang CH, Bijian K, Wernic D, et al. A novel orally available sele-no-purine molecule suppresses triple-negative breast cancer cell proliferation and progression to metastasis by inducing cytostatic autophagy. Autophagy. 2019;15(8):1376-1390.
10. Fazal FM, Chang HY. lncRNA structure: Message to the heart. Mol Cell. 2016;64(1):1-2.
international expert consensus on the primary therapy of early breast cancer 2013. Ann Oncol. 2013;24(9):2206-2223.

34. Coates AS, Winer EP, Goldhirsch A, et al. Tailoring therapies–improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. Ann Oncol. 2015;26(8):1533-1546.

35. Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. Comput Struct Biotechnol J. 2015;13:8-17.

36. Lai J, Chen BO, Mok H, Zhang G, Ren C, Liao N. Comprehensive analysis of autophagy-related prognostic genes in breast cancer. J Cell Mol Med. 2020;24(16):9145-9153.

37. Jun-Xian DU, Chen C, Luo Y-H, et al. Establishment and validation of a novel autophagy-related gene signature for patients with breast cancer. Gene. 2020;762:144974.

38. Lin Q-G, Liu W, Mo Y-Z, et al. Development of prognostic index based on autophagy-related genes analysis in breast cancer. Aging (Albany NY). 2020;12(2):1366-1376.

39. Sun Z, Jing C, Xiao C, Li T. An autophagy-related long non-coding RNA prognostic signature accurately predicts survival outcomes in bladder urothelial carcinoma patients. Aging (Albany NY). 2020;12(15):15624-15637.

40. Luan F, Chen W, Chen M, et al. An autophagy-related long non-coding RNA signature for glioma. FEBS Open Bio. 2019;9(4):653-667.

How to cite this article: Li X, Jin F, Li Y. A novel autophagy-related lncRNA prognostic risk model for breast cancer. J Cell Mol Med. 2021;25:4–14. https://doi.org/10.1111/jcmm.15980