

Original article

Clinicopathological and genomic features of breast mucinous carcinoma



Kebin Lu, Xiao'e Wang, Wenyuan Zhang, Hui Ye, Lingling Lao, Xiaodong Zhou, Sijie Yao, Feng Lv*

Department of General Surgery, Yuyao People's Hospital of Zhejiang Province, Yuyao, Zhejiang, 315400, PR China

ARTICLE INFO

Article history:

Received 1 July 2020

Received in revised form

26 July 2020

Accepted 28 July 2020

Available online 6 August 2020

Keywords:

Breast cancer

Mucinous carcinoma

Clinicopathological characteristics

Genomic features

ABSTRACT

Introduction: Mucinous carcinoma (MC) of the breast is a special histological type of breast cancer. Clinicopathological characteristics and genomic features of MC is not fully understood.

Materials and methods: 186,497 primary breast cancer patients from SEER database diagnosed with invasive ductal carcinoma (IDC) or MC were included. 801 primary IDC or MC patients from TCGA cohort were included for transcriptomic and genomic analysis.

Results: MC patients were older, had lower tumor grade and T and N stage, higher hormone receptor positive proportions and lower HER2 positive proportions than IDC patients. Kaplan-Meier plots showed that the breast cancer-specific survival (BCSS) of MC patients was significantly better than IDC patients ($P < 0.001$). However, after adjusting for clinicopathological factors, survival advantage of MC disappeared. In terms of genomic features of MC, representative upregulated genes of MC in transcriptomic level were *MUC2*, *TFF1* and *CARTPT*. Upregulated pathways of MC included neurotransmitter-related pathways. Moreover, MC was featured by the amplification of 6p25.2, 6q12 and 11q12.3.

Conclusion: MC is a distinct histological subtype compared with IDC in terms of clinicopathological characteristics and genomic features. Further investigation need to be conducted to explore the formation of this specific histological subtype.

© 2020 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Breast cancer is one of the most common cancers, with greater than 279,100 cases and 42,690 deaths in 2020 in the United States [1]. Pathologically, breast cancer is a heterogeneous entity with over 20 histological types [2–4]. Of these distinct histological types, invasive ductal carcinoma (IDC) is the most common one, accounting for approximately 80% of all breast cancer cases [5–7]. Other histological types included invasive lobular carcinoma, adenoid cystic carcinoma, medullary carcinoma, metaplastic carcinoma and so on [5,6]. Histological type has been proved to play an important role in breast cancer prognosis [8,9]. For example, patients with the medullary carcinoma have been shown to have better prognosis whereas metaplastic carcinoma is associated with a worse prognosis [10,11]. Therefore, the histological classification of breast cancer is of great prognostic value and worth a further

research to explore the genomic background of each histological type.

Mucinous adenocarcinoma (MC) of the breast is a rare histological type of breast cancer characterized by large amount of extracellular mucin [12]. In this type of cancer, the tumor is made up of abnormal cells that “float” in pools of mucin, a key ingredient in the slimy, slippery substance known as mucus [13]. It accounts for 1–6% of all invasive breast cancers [14]. MC has been classified into two subgroups, the “pure” type consists exclusively of tumor tissue with extracellular mucin production, while the “mixed” type is defined as a tumor where 50–90% of the area is mucinous and also admixing with infiltrating ductal epithelial component [15]. These two subgroups have several fundamental differences in clinicopathological features and outcomes [16]. Recently, a new subgroup of the “mixed” type of MC, the MC with micropapillary features, was reported to be morphologically, clinically and genetically distinct from pure MC of breast [17]. In previous studies with relatively small patient numbers, MC has been exhibited to have less lymph nodes invasion, lower stage, more expression of estrogen and progesterone receptors (ER and PR) and less human

* Corresponding author. Department of General Surgery, PR China.
E-mail address: 13957465305@163.com (F. Lv).

epidermal receptor 2 (HER2) overexpression compared to IDC [12,18–21]. Genetically, MC has been illustrated to have a specific molecular identity different from invasive ductal carcinoma, such as the lower genomic instability [22,23]. There also existed genetic heterogeneity within the subgroups of MC. For example, MC with the micropapillary features was illustrated to have subtype specific *GATA3*, *TP53* and *SF3B1* mutation and 17q and 20q gains as well as 17p losses [17,24]. However, because of relative rarity of MC patients, the clinicopathological characteristics, prognostic value and genomic features of MC are not well-established.

By using the clinicopathological data from the Surveillance, Epidemiology and End Results (SEER) database and genomic data from The Cancer Genome Atlas (TCGA) database, our study aimed to investigate the clinicopathological characteristics and genomic features of MC by comparing it with IDC.

2. Materials and methods

2.1. Database

Data for this study were obtained from the recent SEER 18 registry research database (November 2015 Submission). The SEER 18 database contains data from the SEER 13 registries (Atlanta, Connecticut, Detroit, Hawaii, Iowa, New Mexico, San Francisco-Oakland, Seattle-Puget Sound, Utah, Los Angeles, San Jose-Monterey, rural Georgia, and the Alaska Native Tumor Registry) and the registries of greater California, Kentucky, Louisiana, New Jersey, and greater Georgia. The SEER database of the National Cancer Institute (NCI) is the largest population-based cancer registry in the United States and covers approximately 28% of the population (<http://seer.cancer.gov/about/>). All the data included in the present study were obtained with the approval of the independent ethics committee/institutional review board at Ethical Committee of Yuyao People's Hospital of Zhejiang Province.

2.2. Study population

Our data were obtained from the SEER database released in April 2019, which includes data from 18 population-based registries (1973–2015). The inclusion criteria were as follows: female patients, diagnosis year from 2010 to 2015, histological grades I-IV (Grade IV is the undifferentiated [UD] type), American Joint Committee on Cancer (AJCC) stages I-III, pathologic confirmation of infiltrating ductal carcinoma-not otherwise specified (IDC-NOS, ICD-O-3 8500/3) and infiltrating mucinous carcinoma (ICD-O-3 8480/3), unilateral breast cancer, breast cancer as the first and only cancer diagnosis, diagnosis not obtained from a death certificate or autopsy, only one primary site, and known ER and PR status. In all, 186,497 patients were selected, including 4578 MC patients and 181,919 IDC patients.

An analysis of the demographic and clinical characteristics of the IDC and MC subtypes included age at diagnosis, race, grade, ER, PR and HER2 status, T and N stage, radiation, chemotherapy and surgery type. The age at diagnosis was also considered as a continuous variable. Grade 3 and undifferentiated grade were merged into a single group. In addition, the types of radiation were summarized as yes, no or unknown; the types of surgery were classified as no surgery, lumpectomy, mastectomy or unknown. Detailed classification information is described in Table 1. All unknown data were excluded from the Cox analysis and subgroup analysis but were included in the generation of the Kaplan-Meier curves.

2.3. Gene set enrichment analysis (GSEA)

GSEA was performed with the GSEA software (v3.0) and the Molecular Signature Database (v6.1) (<http://www.broad.mit.edu/gsea/>). One thousand total permutations were used, IDC versus MC was used as phenotype labels. The permutation type was set to “phenotype”.

2.4. Comparison of somatic mutations and copy number alterations

Genes with mutation frequencies greater than 0% in IDC or MC were included in our comparison. P value less than 0.05 was considered significant. As most gene was low-frequency mutated genes, we conducted permutation test (“chisq_test” function in R program) to conduct the comparison. The log2 ratio of the copy numbers of each gene was compared between IDC and MC as well. P value less than 0.0 was considered significant. Chromosome fragments with more than three continuous significant genes were considered as significant peaks.

2.5. Statistical analysis

The demographic and clinical characteristics of the included cases were compared across groups by the Pearson Chi-square test or Fisher's exact test for categorical nominal data and by the Cochran-Mantel Haenszel (CMH) Chi-square test for categorical ordinal data. The BCSS and OS were considered as the primary and secondary outcomes of our study, respectively. BCSS was defined as the time from the date of diagnosis to the date of death caused by breast cancer. The OS was defined as the time from the date of diagnosis to the date of death from any cause. Kaplan-Meier curves and log-rank tests were generated with the function “Surv” and “survfit” (R package: survival and rms). Additionally, univariate and multivariate Cox proportional hazard models with HRs and 95% CIs were applied to estimate the factors associated with the BCSS and OS. Subgroup analysis and forest plots were generated with the function “forestplot” (R package: forestplot). All statistical analyses were performed using R version 3.2.2 (R Foundation for Statistical Computing, Vienna, Austria) and SPSS version 17.0 (IBM SPSS Statistics, Chicago, IL, USA). A two-sided P value < 0.001 was considered statistically significant in the comparison of clinicopathological characteristics. Two-sided P values < 0.05 were considered statistically significant in other tests.

3. Results

3.1. Demographic and clinical characteristics

We summarized the demographic and clinical characteristics of all 186,497 selected patients in Table 1, including 4578 MC patients and 181,919 IDC patients. MC patients were significantly older than IDC patients (median age at diagnosis [years]: 65 vs 59, $P < 0.001$). Additionally, MC patients had strikingly lower grade tumors (Grade 1: 56.0% vs 20.7%, $P < 0.001$), lower N stage (N0: 90.9% vs 69.3%, $P < 0.001$), higher proportion of cases with a positive ER and PR status (97.3% vs 79.7%, $P < 0.001$; % 90.0 vs 69.7%, $P < 0.001$, respectively) and lower proportion of cases with a positive HER2 status (5.1% vs 16.2%, $P < 0.001$) compared with IDC patients. Moreover, in terms of treatment, MC patients had lower proportion of receiving radiation, chemotherapy and mastectomy (51.6% vs 55.3%, $P < 0.001$; 12.6% vs 43.6%, $P < 0.001$; 31.1% vs 39.4%, $P < 0.001$, respectively). These data suggested that MC has distinct baseline characteristics from IDC.

Table 1
Baseline clinicopathological characteristics of IDC and MC patients.

Variable	IDC		MC		P
	n	%	n	%	
Age at diagnosis (years)					<0.001
Mean	59		65		
Marital status ^a					0.615
Married	145,709	80.1%	3640	79.5%	
Not married	27,269	15.0%	708	15.5%	
Unknown	8941	4.9%	230	5.0%	
Race ^b					<0.001
White	141,935	78.0%	3455	75.5%	
Black	20,307	11.2%	539	11.8%	
Asian	17,265	9.5%	529	11.6%	
Other	2412	1.3%	55	1.2%	
Grade ^c					<0.001
1	37,684	20.7%	2564	56.0%	
2	73,430	40.4%	1572	34.3%	
3	65,450	36.0%	155	3.4%	
Unknown	5355	2.9%	287	6.3%	
T stage					<0.001
T1	115,129	63.3%	3050	66.6%	
T2	54,393	29.9%	1263	27.6%	
T3	8213	4.5%	213	4.7%	
T4	3976	2.2%	50	1.1%	
Unknown	208	0.1%	2	0.0%	
N stage					<0.001
N0	126,134	69.3%	4161	90.9%	
N1	41,744	22.9%	343	7.5%	
N2	9197	5.1%	54	1.2%	
N3	4812	2.6%	20	0.4%	
Unknown	32	0.0%	0	0.0%	
ER status					<0.001
Negative	34,168	18.8%	48	1.0%	
Positive	144,995	79.7%	4456	97.3%	
Unknown	2756	1.5%	74	1.6%	
PR status					<0.001
Negative	51,801	28.5%	369	8.1%	
Positive	126,779	69.7%	4120	90.0%	
Unknown	3337	1.8%	89	2.0%	
HER2 status					<0.001
Negative	142,585	78.4%	4108	89.7%	
Positive	29,548	16.2%	233	5.1%	
Unknown	9786	5.3%	237	5.2%	
Radiation					<0.001
No	74,543	41.0%	2116	46.2%	
Yes	100,597	55.3%	2360	51.6%	
Unknown	6779	3.8%	102	2.3%	
Chemotherapy					<0.001
No	102,657	56.4%	3999	87.4%	
Yes	79,262	43.6%	579	12.6%	
Surgery					<0.001
Lumpectomy	109,848	60.4%	3141	68.6%	
Mastectomy	71,632	39.4%	1425	31.1%	
Unknown	439	0.2%	12	0.2%	

Note.

Abbreviations: IDC: invasive ductal carcinoma; MC: mucinous adenocarcinoma; ER: estrogen receptor; PR: progesterone receptor; HER2: human epidermal receptor 2.

^a Including divorced, separated, single (never married), and widowed.

^b Including American Indian/Alaskan native, Asian/Pacific Islander, and others-unspecified.

^c Including grade 3 and undifferentiated.

3.2. Survival analysis

We compared the prognosis of these histological types within the 30.2-month median follow-up period. Kaplan-Meier plots were used to evaluate the BCSS and overall survival (OS) for IDC and MC (Fig. 1). In comparison with IDC patients, MC patients had a significantly better BCSS (log-rank test $P < 0.001$). However, IDC and MC patients had insignificant difference in OS comparison (log-rank test $P = 0.4$).

The Cox proportional hazards model was used to further investigate the effect of the baseline characteristics of the disease

on BCSS and OS (Table 2). It was suggested that patients with MC exhibited a better BCSS than patients with IDC (hazard ratio [HR] = 0.22, 95% confidence interval [CI]: 0.15–0.32, $P < 0.001$), while no significant difference was observed in the OS between patients with MC and those with IDC (HR = 0.96, 95% CI: 0.83–1.10, $P = 0.526$). These significant variables were then included in the multivariate analysis to confirm their prognostic effect (Table 2). Most of the variables remained significant prognostic predictors in the multivariate analysis. However, after adjusting for other prognostic predictors, MC was no longer an independent prognostic predictor of BCSS compared with IDC (HR = 0.75, 95% CI: 0.51–1.09,

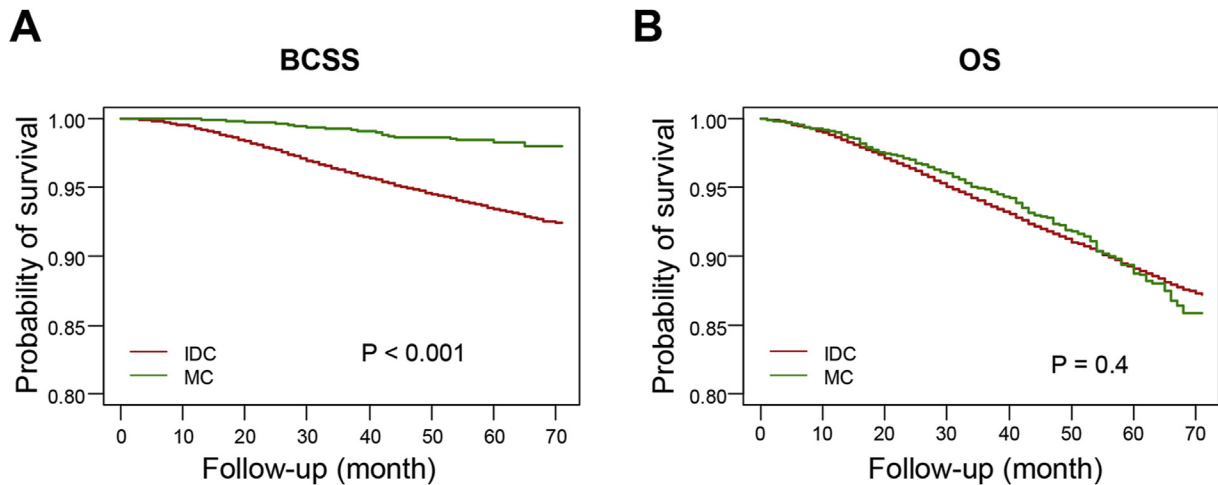


Fig. 1. Survival analysis of mucinous carcinoma. Kaplan-Meier curves of breast cancer-specific survival (BCSS) (A) and of overall survival (OS) (B) according to histological type in all patients. Log-rank tests were compared between MC and IDC. Abbreviations: MC: mucinous carcinoma; IDC: invasive ductal carcinoma; BCSS: breast cancer-specific survival; OS: overall survival.

$P = 0.125$). In contrary, MC independently predicted worse OS compared with IDC (HR = 1.36, 95% CI: 1.16–1.57, $P < 0.001$). We concluded that the prognostic value of MC was confounded by some other clinicopathological variables.

3.3. Subgroup analysis

In order to investigate whether MC predicted homogeneous prognosis when stratified by different clinical parameters, we conducted a subgroup analysis to compare the BCSS and OS between MC and IDC in each subgroup. Forest plots of HRs in the univariate Cox analysis summarized the exploratory subgroup analysis of the BCSS and OS are shown in Fig. 2. Compared with IDC, MC had lower HRs for BCSS in almost all subgroups. In terms of subgroup analysis of OS, MC had higher HRs for OS in grade 1 (HR = 2.07, 95% CI: 1.72–2.50, $P < 0.001$), T1 (HR = 1.24, 95% CI: 1.02–1.52, $P < 0.05$), N0 (HR = 1.34, 95% CI: 1.16–1.56, $P < 0.001$), ER positive (HR = 1.24, 95% CI: 1.07–1.42, $P < 0.01$), PR positive (HR = 1.33, 95% CI: 1.15–1.54, $P < 0.001$) and lumpectomy (HR = 1.23, 95% CI: 1.03–1.46, $P < 0.05$) subgroups. Meanwhile, MC had lower HRs for OS in age ≤ 60 (HR = 0.46, 95% CI: 0.31–0.70, $P < 0.001$), T3 (HR = 0.51, 95% CI: 0.31–0.85, $P < 0.01$), underwent radiation (HR = 0.69, 95% CI: 0.53–0.90, $P < 0.01$), underwent chemotherapy (HR = 0.44, 95% CI: 0.25–0.76, $P < 0.01$) and mastectomy (HR = 0.78, 95% CI: 0.62–0.98, $P < 0.05$) subgroups. These data suggested that some clinicopathological markers, such as tumor grade, T and N stage, ER, PR and HER2 status, were important confounders in determining the prognosis of MC patients.

3.4. Transcriptomic and genomic features of breast mucinous adenocarcinoma

We also utilized the TCGA dataset to explore the transcriptomic and genomic features of MC. In comparison with IDC samples, MC samples were featured by the upregulation of *MUC2*, *CARTPT* and *TFF1* (Fig. 3A). We also conducted gene set enrichment analysis (GSEA) to investigate the MC-specific pathways. As illustrated in Fig. 3B, neurotransmitter release-related pathways, including the gamma aminobutyric acid (GABA) synthesis, release and degradation pathway, was upregulated in MC. In addition, we compared the somatic mutational landscape between IDC and MC (Fig. 3C). MC

had significantly lower somatic mutation frequency than IDC in *TP53*, *PIK3CA*, *XIRP2*, *LRBA* and *EYS*. Furthermore, in terms of somatic copy number alterations (SCNA), we observed that 6p25.2, 6q12 and 11q12.3 were significantly amplified in MC (Fig. 3D). Genes located in these region including *FADS1*, *FADS2*, *FADS3*, *MYRF*, *FEN1* and *FTH1*.

4. Discussion

In this study, we retrospectively investigated the clinicopathological characteristics, prognostic significance and genomic features of MC through a comparison with IDC. The results suggested that MC has baseline characteristics that are distinct from those of IDC. A survival analysis indicated that MC was associated with a significantly better BCSS than IDC and similar OS when compared with IDC. However, the difference in BCSS between MC and IDC disappeared after adjusting for confounding factors. Furthermore, multivariate Cox analysis revealed that MC was an independent factor predicting higher HR for OS. In addition, transcriptomic features of MC were correlated with its abundant mucin secretion features. Some MC-specific genomic alterations might contribute to the formation of its pathological features.

As the largest analysis of MC to date, our research took advantage of the high number of SEER datasets to further investigate the clinicopathological characteristics and prognostic significance of MC. We illustrated that MC patients were older, had earlier stage and were more likely to be ER and PR positive and HER2 negative. Therefore, MC patients were less likely to experience radiation, chemotherapy and mastectomy. Most of our results were consistent with those of previous studies [12,20,21,25,26]. For example, Elimimian et al. compared the clinicopathological features and prognosis of rare carcinomas of the breast with the data from the National Cancer Database, which enrolled 70,341 breast cancer patients with the rare histological types, and demonstrated that MC had higher proportion of ER and PR positivity, lower HER2 positivity and an overall good prognosis [26]. However, a retrospective study with 268 MC patients demonstrated that MC patients were younger than IDC patients [12]. We considered this phenomenon as the patient bias as the enrolled patient number of this study is relatively small and the result of our study was in consistent with most of the other previous studies [20,21,25]. We also analyzed the

Table 2
Multivariate analysis of breast cancer-specific survival (BCSS) and overall survival (OS) predictors using a cox proportional hazards model.

Variables	BCSS		OS	
	HR (95% CI)	P	HR (95% CI)	P
Histological type				
MC	0.75 (0.51–1.09)	0.125	1.35 (1.16–1.57)	<0.001
IDC	Reference			
Age at diagnosis				
≤ 60 y	0.66 (0.62–0.69)	<0.001	0.44 (0.42–0.46)	<0.001
> 60 y	Reference			
Race ^a				
Black	1.26 (1.17–1.35)	<0.001	1.22 (1.16–1.30)	<0.001
Asian	0.69 (0.61–0.77)	<0.001	0.62 (0.56–0.68)	<0.001
White	Reference			
Grade ^b				
1	0.21 (0.18–0.25)	<0.001	0.50 (0.46–0.54)	<0.001
2	0.50 (0.46–0.54)	<0.001	0.64 (0.60–0.67)	<0.001
3	Reference			
T stage				
T1	0.18 (0.16–0.20)	<0.001	0.20 (0.18–0.22)	<0.001
T2	0.41 (0.37–0.45)	<0.001	0.41 (0.37–0.44)	<0.001
T3	0.73 (0.66–0.81)	<0.001	0.70 (0.63–0.77)	<0.001
T4	Reference			
N stage				
N0	0.15 (0.14–0.17)	<0.001	0.22 (0.21–0.24)	<0.001
N1	0.35 (0.32–0.38)	<0.001	0.38 (0.35–0.41)	<0.001
N2	0.67 (0.61–0.74)	<0.001	0.67 (0.61–0.73)	<0.001
N3	Reference			
ER status				
Positive	0.64 (0.59–0.69)	<0.001	0.66 (0.61–0.70)	<0.001
Negative	Reference			
PR status				
Positive	0.51 (0.47–0.55)	<0.001	0.64 (0.60–0.68)	<0.001
Negative	Reference			
HER2 status				
Positive	0.57 (0.53–0.62)	<0.001	0.67 (0.63–0.71)	<0.001
Negative	Reference			
Radiation				
No	1.60 (1.50–1.70)	<0.001	1.95 (1.86–2.05)	<0.001
Yes	Reference			
Chemotherapy				
No	1.46 (1.36–1.57)	<0.001	1.89 (1.79–1.99)	<0.001
Yes	Reference			
Surgery type				
Lumpectomy	0.95 (0.89–1.02)	0.165	1.09 (1.04–1.15)	0.001
Mastectomy	Reference			

Note.

Abbreviations: BCSS: breast cancer-specific survival; OS: overall survival; HR: hazard ratio; CI: confidence interval; IDC: invasive ductal carcinoma; MC: mucinous adenocarcinoma; ER: estrogen receptor; PR: progesterone receptor; HER2: human epidermal receptor 2.

^a Including American Indian/Alaskan native, Asian/Pacific Islander, and others-unspecified.

^b Including grade 3 and undifferentiated.

prognostic value of MC in breast cancer. MC patients showed better BCSS and similar OS when compared with IDC patients. The older median age at diagnosis of MC than IDC could explain the significant difference in BCSS but non-significant difference in OS between the two histological types. However, MC was not an independent prognostic factor for BCSS. After the adjustment for other factors, the prognostic value of MC for BCSS disappeared. This phenomenon suggested the crucial role of confounding factors in determining the BCSS of MC patients, such as the lower grade, higher proportion of ER and PR positivity and lower proportion of HER2 positivity of MC patients, which might contribute to the better BCSS of MC patients. Similarly, previous studies also demonstrated that MC patients had better disease-free survival (DFS) than IDC patients, but the significant prognostic value disappeared after the multivariate Cox analysis [12,21,27,28]. The similar OS between MC and IDC was also reported in previous studies [12,29]. We further demonstrated that after adjusting for confounding factors, MC independently predicted worse prognosis

of OS than IDC. This phenomenon suggested that after adjusting for confounding factors, MC patients were more likely to be dead from the reasons except for breast cancer. The reasons for this phenomenon might be related to the genetic background of MC patients, further explorations are needed to explore the relationship between the MC histological type with the risk of diseases other than breast cancer, such as cardiovascular and cerebrovascular diseases.

The other advantage of our study was the comprehensive transcriptomic and genomic analysis of MC. Because of the limited number of samples and the difficulties of high-throughput sequencing in earlier years, previous studies only revealed a few genetic features of MC. MC exhibited a specific molecular identity different from IDC. MC was illustrated to have the lowest levels of gene copy number changes and a lower genetic instability [30,31]. Previous study demonstrated that *GATA3*, *KMT2C* and *MAP3K1* were the most frequently mutated genes in pure MCBs [23]. Two recurrent but not pathognomonic fusion genes, *OAZ1-CSNK1G2* and

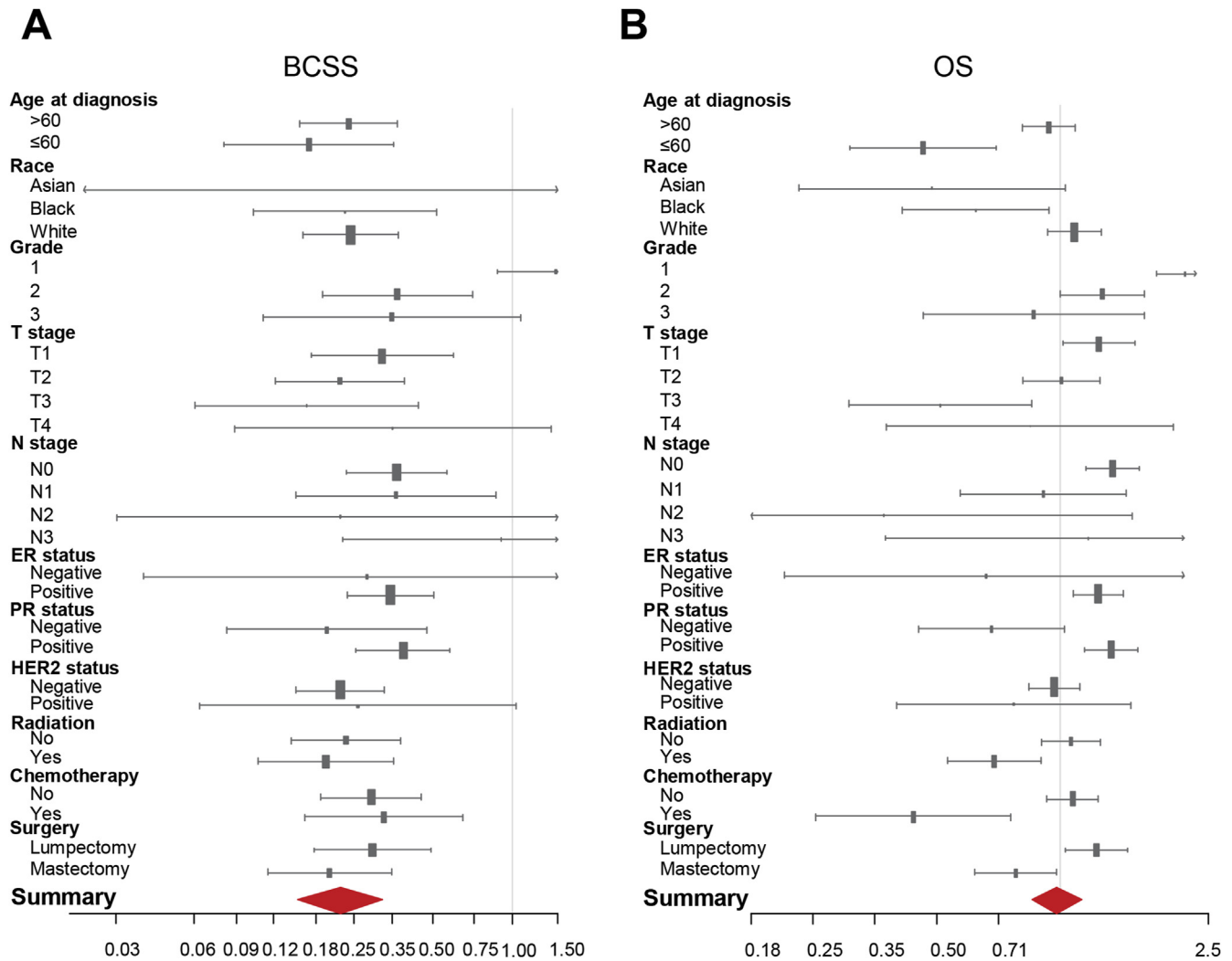


Fig. 2. Subgroup analysis of mucinous carcinoma. Forest plots of hazard ratios (HRs) of MC versus IDC for BCSS (A) and OS (B) according to the subgroup analysis. The X-axis shows the HRs and 95% confidence intervals (CIs) of each subgroup. The size of the boxes represents the relative number of patients in each subgroup. Abbreviations: HRs: hazard ratios; BCSS: breast cancer-specific survival; OS: overall survival; MC: mucinous carcinoma; IDC: invasive ductal carcinoma; ER: estrogen receptor; PR: progesterone receptor; HER2: human epidermal receptor 2.

RFC4-LPP, were detected in pure MC [23]. In addition, MC had fewer concurrent 1q gains and 16q losses [23]. Previous studies have also demonstrated that MC with the micropapillary features was illustrated to have subtype specific EMA/MUC1 “inside-out” pattern, *GATA3*, *TP53* and *SF3B1* mutation and 17q and 20q gains as well as 17p losses [17,24]. In our study, we first analyzed the transcriptomic difference between MC and IDC and illustrated that the upregulated molecules and pathways were related to neurotransmitter secretion, which might contribute to the mucin formation of MC. In addition, we revealed that MC patients had lower somatic mutation frequency of *TP53*, *PIK3CA* and *LRBA*. In addition, MC-specific somatic copy number alterations included the amplification of 6p25.2, 6q12 and 11q12.3 compared with IDC. These specific genomic alterations might be related to the histological features of MC. Some of the results in our study were similar to previous studies. For example, both of our study and the study of Pareja et al. revealed that MC had lower mutation frequency of *TP53* and *PIK3CA*, and *GATA3* was the most frequently mutated gene of MC. However, as the sample size of MC in the two studies were both small, some other top mutated genes of MC were different in the two studies. The mutational landscape of MC need further

exploration with larger sample size. In terms of SCNA, Pareja et al. revealed that MC had fewer concurrent 1q gains and 16q losses If we set the cut-off P value for SCNA comparison between MC and IDC into $P < 0.01$, we could also conclude that MC had fewer concurrent 1q gains and 16q losses in our study.

Our research also has several limitations. First, it is a retrospective study and may have some potential selection bias. In addition, as the HER2 information in the SEER datasets was not available until 2010, we selected patients from 2010 in SEER database. Therefore, the follow-up period of our study was not very long. Furthermore, as the understanding of MC in our study focused on the bioinformatics analysis of the clinicopathological characteristics and genomic features, further functional validation studies are needed to investigate the biological nature of MC.

In conclusion, MC is a histological subtype that is distinct from IDC with respect to clinicopathological characteristics, prognostic significance and genomic features. The lower grade, higher proportion of ER and PR positivity and lower proportion of HER2 positivity of MC contributed a lot to its better BCSS. In addition, some MC-specific transcriptomic and genomic alterations may contribute to the formation of its histological features.

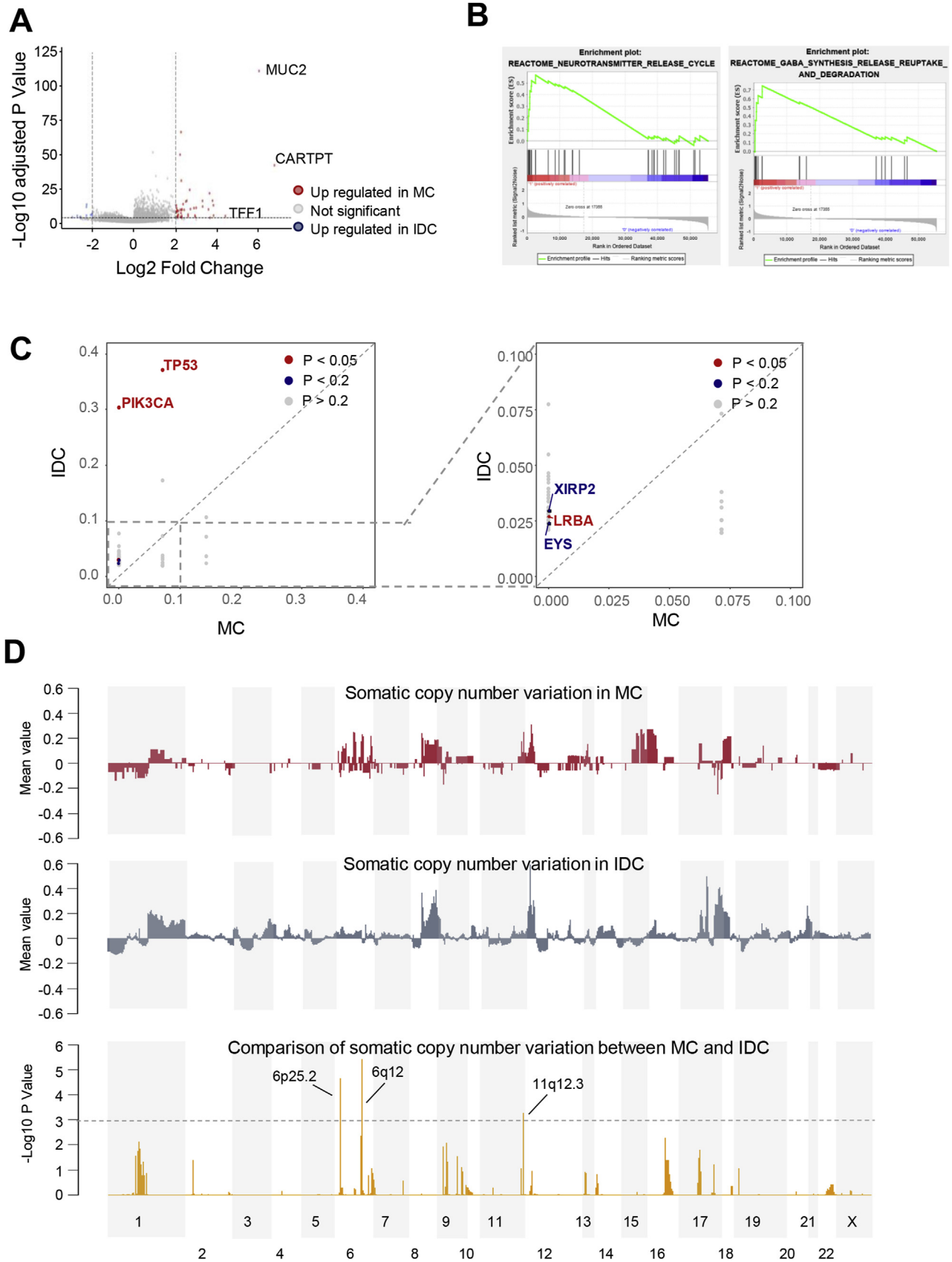


Fig. 3. Transcriptomic and genomic features of mucinous carcinoma. **(A)** Volcano plot of differentially expressed genes between MC and IDC. **(B)** GSEA analysis of representative MC-specific upregulated pathways. **(C)** Comparison of somatic mutation frequency between MC and IDC. **(D)** Comparison of somatic copy number alterations between MC and IDC. Abbreviations: MC: mucinous carcinoma; IDC: invasive ductal carcinoma; GSEA: gene set enrichment analysis.

Author contributions statement

SY and FL designed the experiments. KL, XW, WZ and XZ collected the data, KL, XW, WZ, HY, LL, XZ and SY conducted statistical analysis and data Interpretation. HY and FL wrote the original manuscript and revised the manuscript.

Declaration of competing interest

The authors declare no competing financial interests.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- [1] Siegel RL, Miller KD, Jemal A. Cancer statistics CA A Cancer J Clin 2020;70:7–30. <https://doi.org/10.3322/caac.21590>. 2020.
- [2] Woolston C. Breast cancer. Nature 2015;527:S101. <https://doi.org/10.1038/527S101a>.
- [3] Cancer Genome Atlas N. Comprehensive molecular portraits of human breast tumours. Nature 2012;490:61–70. <https://doi.org/10.1038/nature11412>.
- [4] Banerji S, et al. Sequence analysis of mutations and translocations across breast cancer subtypes. Nature 2012;486:405–9. <https://doi.org/10.1038/nature11154>.
- [5] Colleoni M, et al. Outcome of special types of luminal breast cancer. Ann Oncol 2012;23:1428–36. <https://doi.org/10.1093/annonc/mdr461>.
- [6] Li Ci, Uribe DJ, Daling JR. Clinical characteristics of different histologic types of breast cancer. Br J Canc 2005;93:1046–52. <https://doi.org/10.1038/sj.bjc.6602787>.
- [7] Li Ci, Anderson BO, Daling JR, Moe RE. Trends in incidence rates of invasive lobular and ductal breast carcinoma. Jama 2003;289:1421–4.
- [8] Zhao S, Ma D, Xiao Y, Jiang YZ, Shao ZM. Clinicopathologic features and prognoses of different histologic types of triple-negative breast cancer: a large population-based analysis. Eur J Surg Oncol 2018;44:420–8. <https://doi.org/10.1016/j.ejso.2017.11.027>.
- [9] Xiao Y, et al. Mixed invasive ductal and lobular carcinoma has distinct clinical features and predicts worse prognosis when stratified by estrogen receptor status. Sci Rep 2017;7:10380. <https://doi.org/10.1038/s41598-017-10789-x>.
- [10] Huober J, et al. Prognosis of medullary breast cancer: analysis of 13 international breast cancer study group (IBCSG) trials. Ann Oncol 2012;23:2843–51. <https://doi.org/10.1093/annonc/mds105>.
- [11] Kulkarni N, et al. Rare breast cancer: 933 adenoid cystic carcinomas from the National Cancer Data Base. Ann Surg Oncol 2013;20:2236–41. <https://doi.org/10.1245/s10434-013-2911-z>.
- [12] Bae SY, et al. Mucinous carcinoma of the breast in comparison with invasive ductal carcinoma: clinicopathologic characteristics and prognosis. J Breast Cancer 2011;14:308–13. <https://doi.org/10.4048/jbc.2011.14.4.308>.
- [13] Tan PH, Tse GM, Bay BH. Mucinous breast lesions: diagnostic challenges. J Clin Pathol 2008;61:11–9. <https://doi.org/10.1136/jcp.2006.046227>.
- [14] Anderson WF, Chu KC, Chang S, Sherman ME. Comparison of age-specific incidence rate patterns for different histopathologic types of breast carcinoma. Canc Epidemiol Biomarkers Prev 2004;13:1128–35.
- [15] Frank GA, Danilova NV, Andreeva I, Nefedova NA. [WHO classification of tumors of the breast, 2012]. Arkh Patol 2013;75:53–63.
- [16] Marrazzo E, et al. Mucinous breast cancer: a narrative review of the literature and a retrospective tertiary single-centre analysis. Breast 2020;49:87–92. <https://doi.org/10.1016/j.breast.2019.11.002>.
- [17] Sun P, et al. Mucinous carcinoma with micropapillary features is morphologically, clinically and genetically distinct from pure mucinous carcinoma of breast. Mod Pathol 2020. <https://doi.org/10.1038/s41379-020-0554-8>.
- [18] Diab SG, et al. Tumor characteristics and clinical outcome of tubular and mucinous breast carcinomas. J Clin Oncol 1999;17:1442–8. <https://doi.org/10.1200/JCO.1999.17.5.1442>.
- [19] Barkley CR, et al. Mucinous breast carcinoma: a large contemporary series. Am J Surg 2008;196:549–51. <https://doi.org/10.1016/j.amjsurg.2008.06.013>.
- [20] Di Saverio S, Gutierrez J, Avisar E. A retrospective review with long term follow up of 11,400 cases of pure mucinous breast carcinoma. Breast Canc Res Treat 2008;111:541–7. <https://doi.org/10.1007/s10549-007-9809-z>.
- [21] Lei L, Yu X, Chen B, Chen Z, Wang X. Clinicopathological characteristics of mucinous breast cancer: a retrospective analysis of a 10-year study. PLoS One 2016;11:e0155132. <https://doi.org/10.1371/journal.pone.0155132>.
- [22] Lacroix-Triki M, et al. Mucinous carcinoma of the breast is genomically distinct from invasive ductal carcinomas of no special type. J Pathol 2010;222:282–98. <https://doi.org/10.1002/path.2763>.
- [23] Pareja F, et al. The genomic landscape of mucinous breast cancer. J Natl Cancer Inst 2019;111:737–41. <https://doi.org/10.1093/jnci/djy216>.
- [24] Troxell ML. Reversed MUC1/EMA polarity in both mucinous and micropapillary breast carcinoma. Hum Pathol 2014;45:432–4. <https://doi.org/10.1016/j.humpath.2013.08.026>.
- [25] Hanagiri T, et al. Clinicopathologic characteristics of mucinous carcinoma of the breast. Int Surg 2010;95:126–9.
- [26] Elizabeth Blessing Elimimian TAS, Liang Hong, Bilani Nadeem, Elson Leah, Zeina A, Nahleh. Clinicopathological features and prognosis of rare carcinomas of the breast: a comparative analysis from the NCDB. J Clin Oncol 2020;38:e19256.
- [27] Wei YN, et al. Clinicopathologic characteristics of HER2-positive pure mucinous breast carcinoma: a systematic investigation into an unusual tumor. Int J Clin Exp Pathol 2019;12:1666–77.
- [28] Komaki K, Sakamoto G, Sugano H, Morimoto T, Monden Y. Mucinous carcinoma of the breast in Japan. A prognostic analysis based on morphologic features. Cancer 1988;61:989–96. [https://doi.org/10.1002/1097-0142\(19880301\)61:5<989::aid-cnrc2820610522>3.0.co;2-e](https://doi.org/10.1002/1097-0142(19880301)61:5<989::aid-cnrc2820610522>3.0.co;2-e).
- [29] Vo T, et al. Long-term outcomes in patients with mucinous, medullary, tubular, and invasive ductal carcinomas after lumpectomy. Am J Surg 2007;194:527–31. <https://doi.org/10.1016/j.amjsurg.2007.06.012>.
- [30] Toikkanen S, Eerola E, Eklfors TO. Pure and mixed mucinous breast carcinomas: DNA stemline and prognosis. J Clin Pathol 1988;41:300–3. <https://doi.org/10.1136/jcp.41.3.300>.
- [31] Horlings HM, et al. Genomic profiling of histological special types of breast cancer. Breast Canc Res Treat 2013;142:257–69. <https://doi.org/10.1007/s10549-013-2740-6>.