Genetic and immune characteristics of sentinel lymph node metastases and multiple lymph node metastases compared to their matched primary breast tumours

Bo Chena,b,1, Guochun Zhanga,b,c,d,1, Jianguo Lai1,a, Weikai Xiaoa,f, Xuerui Liang, Cheukfai Lia, Hsiaopei Mokf, Kai Lia, Yulei Wanga, Minghan Jiaa, Chongyang Renf, Lingzhu Wena, Guangnan Weia,b, Jiali Linac, Yingzi Lia,d, Yuchen Zhanga,b, Xiaoqing Chenb, Xueying Wuf, Henghui Zhangbg, Min Linh, Charles M. Balch*, Ning Liaoab,c,d,*

a Department of Breast Cancer, Guangdong Provincial People’s Hospital and Guangdong Academy of Medical Sciences, Guangzhou, Guangdong, China
b School of Medicine, South China University of Technology, 106 Zongshan Er Road, Guangzhou 510080, Guangdong, China;
c The Second School of Clinical Medicine, Southern Medical University, Guangzhou, Guangdong, China;
d Shantou University Medical College, Shantou, Guangdong, China

e Department of Breast Surgical Oncology, Foshan Maternity and Children's Healthcare Hospital Affiliated to Southern Medical University, Foshan, Guangdong, China
f Genecast Biotechnology Co., Ltd; Beijing, China;
g Institute of Infectious Diseases, Beijing Ditan Hospital, Capital Medical University, Beijing, China
h Burning Rock Biotech, Guangzhou, Guangdong, China
i Department of Surgical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

ARTICLE INFO

Article History:
Received 21 December 2020
Revised 27 July 2021
Accepted 28 July 2021
Available online xxx

Keywords:
Breast cancer
Sentinel lymph node metastases
Multiple lymph node metastases
Genome
Immune characteristics

ABSTRACT

Background: Patients with breast cancer presenting with single lymph node metastasis (from a sentinel node) experience prolonged survival compared to patients with multiple lymph node metastases (≥3). However, little information is available on the genetic and immunological characteristics of breast cancer metastases within the regional lymph nodes as they progress from the sentinel lymph node (SLN) downstream to multiple regional lymph nodes (MLNs).

Methods: Genomic profiling was performed using a next-generation sequencing panel covering 520 cancer-related genes in the primary tumour and metastatic lymph nodes of 157 female patients with breast cancer. We included primary tumours, metastatic lymph nodes and adjacent clinically normal lymph nodes (20 patients from the SLN group and 28 patients from the MLNs group) in the whole transcriptome analysis.

Findings: The downstream metastatic lymph nodes (P = 0.029) and the primary breast tumours (P = 0.011) had a higher frequency of PIK3CA mutations compared to the SLN metastasis. We identified a distinct group of 14 mutations from single sentinel node metastasis and a different group of 15 mutations from multiple nodal metastases. Only 4 distinct mutations (PIK3CA, CDK4, NFKBIA and CDKN1B) were conserved in metastases from both lymph node settings. The tumour mutational burden (TMB) was significantly lower in single nodal metastasis compared to the paired primary breast cancer (P = 0.0021), while the decline in TMB did not reach statistical significance in the MLNs group (P = 0.083). In the gene set enrichment analysis, we identified 4 upregulated signatures in both primary tumour and metastatic lymph nodes of 157 female patients with breast cancer. We included primary tumours, metastatic lymph nodes and adjacent clinically normal lymph nodes (20 patients from the SLN group and 28 patients from the MLNs group) in the whole transcriptome analysis.

Interpretation: Single lymph node metastases and multiple lymph node metastases have significant differences in their molecular profiles and immune profiles. The findings are associated with more aggressive tumour characteristics and less favourable immune characteristics in patients with multiple nodal metastases compared to those with a single metastasis in the sentinel node.

* Corresponding authors.
E-mail addresses: CMBalch@mdanderson.org (C.M. Balch), nsyling@scut.edu.cn (N. Liao).

1 The four authors contributed equally to this work.

https://doi.org/10.1016/j.ebiom.2021.103542
2352-3964/© 2021 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
1. Introduction

Regional lymph nodes (LN) are often the first sites of breast cancer metastasis and the first to encounter host immune surveillance mechanisms intended to destroy foreign invaders [1]. At a clinical level, the presence of nodal metastases has a strong predictive role in estimating the survival of breast cancer patients, depending on prognostic factors such as the number of lymph nodes with metastases and the tumour burden (microscopic versus macroscopic) [2]. At a biological level, their presence elicits an immune response to either reject the mutated breast cancer cells if the patient is to survive or the metastatic tumour engages a mechanism to evade the immune surveillance function [3-7]. Based on the existing evidence, breast cancer metastases undergo many molecular switches or alterations when comparing the molecular profiles of primary tumours with distant breast cancer metastases, but this difference has not been reported in regional lymph nodes from patients with breast cancer [8-11]. On the other hand, in melanoma studies have reported evidence of both somatic mutational changes in regional nodal metastases and immune alterations in the sentinel node that favour the metastatic cascade [12, 13].

Dr. Donald Morton, who pioneered the technology of the sentinel lymph node, proposed an "incubator hypothesis" in which metastatic melanoma clones in the sentinel node would become the source of systemic metastases because surviving metastatic cells had undergone further mutational events while incubating in the lymph nodes that enhanced their metastatic behaviour and/or triggered a state of immune tolerance to the new clones of metastatic cells [14, 15]. While this hypothesis is attractive, little molecular information is available that support this proposal in early breast cancer. At a clinical level, patients with breast cancer presenting single lymph node metastasis (from a sentinel node) experienced prolonged survival compared to patients with multiple lymph node metastases (≥3) [2]. At a biological level, little molecular information is available to determine if surviving metastatic clones in the lymph nodes have undergone additional mutational changes that would allow them to grow in downstream lymph nodes. Furthermore, information is limited on the host/tumour interaction as in the lymph nodes or on the immunological characteristics within the regional lymph nodes as breast cancer metastases progress from the sentinel lymph node to downstream regional lymph nodes.

We hypothesized that the molecular profiles of breast early nodal metastases and the characteristics of the immune response are in a different state of metastatic development compared to more advanced stages of regional lymph node metastases. Thus, we analysed the molecular characteristics of primary breast cancer and matched nodal metastases from a single sentinel nodal metastasis compared to more advanced metastatic lymph nodes (MLNs). We also analysed the immune characteristics of lymph nodes harbouring early metastases compared with lymph nodes with more advanced metastases. This strategy of comparing tumour and immune characteristics from patients with breast cancer presenting with early and advanced nodal metastases may provide insights into the host/tumour interactions that might enhance their capacity for preventing metastatic progression within the regional lymph node basin.

2. Methods

2.1. Ethics committee approval and patient selection

The present study was approved by the Ethics Committee at Guangdong Provincial People’s Hospital (Nos. GDREC2014122H and GDREC2019497H). All patients provided written informed consent to participate in this study. We performed a retrospective screen of patients with breast cancer who underwent surgery in the Department of Breast Cancer, Cancer Center, Guangdong Provincial People’s Hospital from January 1, 2017, to November 30, 2018. The inclusion criteria of this study were as follows: (1) female patients with breast...
cancer who were not treated with neoadjuvant chemotherapy and neoadjuvant radiotherapy; (2) a sentinel lymph node biopsy was performed intraoperatively, and if nodal metastases were identified, then axillary dissection was performed; (3) complete clinicopathological information were available and no loss to follow-up; and (4) pretreatment tissue specimens were collected and available. Sentinel lymph node biopsy was conducted using the blue dye mapping technique. During surgery, 1-2 mL of methylene blue was injected subcutaneously in the peri-subareolar region. Lymphatic drainage was observed for up to 10 min following the injection. The surgically isolated sentinel lymph node was processed for an intraoperative frozen section diagnosis and then for delayed histological and immunohistochemical examinations. All patients with a positive sentinel lymph node underwent axillary lymph node dissection.

We identified 157 female patients with breast cancer (SLN n=79; MLNs n=78) who met our inclusion criteria. The study flow chart is shown in Supplementary Figure S1. After quality control, 75 pairs of breast tumour and SLN tissues were profiled using a next-generation sequencing (NGS). In the MLN group, no sample failed quality control, and 78 pairs of breast tumour and metastatic lymph node tissues were profiled using NGS. The clinical characteristics are provided in Supplementary Table S1. Therefore, 61 pairs of samples from the SLN group and 72 pairs of samples from the MLNs group were subjected to a paired analysis comparing the molecular and immunological profiles of primary breast cancer and their metastases in the regional lymph nodes.

2.2. Next-generation sequencing library preparation and capture-based targeted Deoxyribonucleic acid (DNA) sequencing

DNA was extracted using a QIAamp DNA Formalin-Fixed and PARRlun-Embedded (FFPE) tissue kit (Qiagen). The DNA concentration was measured using Qubit dsDNA assay. Genomic profiling was performed using a next-generation sequencing panel covering 520 cancer-related genes (OncoScreen Plus; Supplementary Table S2) [16]. Sequencing assays were performed by investigators who were blinded to the clinicopathological parameters in the CLIA-certified Burning Rock Biotech laboratory (Guangzhou, China). We performed NGS of the primary tumour and metastatic lymph nodes from each patient. In addition, we tested nonmetastatic lymph nodes from 10 patients. None of these nonmetastatic lymph nodes contained tumour-related somatic mutations.

2.3. Mutation categories

The allelic fraction (AF) of the single-nucleotide variants (SNVs) and the copy number (CN) for the copy number variations (CNVs) were adjusted based on the tumour cell percentage. AF was defined as the absolute proportion of the mutant allele detected in the tumour sample. The tumour cell percentage was defined as the proportion of tumour cells relative to nontumour cells, including normal stromal, epithelial, and lymphoid cells. The evaluation of haematoxylin and eosin (H&E)-stained FFPE tissue sections were performed by two independent pathologists who were unaware of the patient outcomes. A minimum tumour cell percentage of 20% was required for DNA isolation and subsequent NGS. The mutations with adjusted AF and CN were then stratified into 5 categories: matched, LN AF elevated, uncertain, breast-specific, and LN-specific. Mutations that were detected in both primary breast tumour (breast) and lymph node (LN) samples (i.e.: the difference between the AF/CN between the two samples was <2) were categorized as “Matched” mutations that were detected in both samples. When the AF/CN in the LN samples was more than two-fold higher than that of the breast sample, they were categorized as “LN AF elevated”. Mutations that were detected in either breast or LN samples, but the AF/CN was below the lower LOD of the assay after adjustment and could not be conclusively determined as unique mutations were categorized as “uncertain”. Mutations that were only detected in the breast samples where the AF/CN remained higher than the LOD were categorized as “breast-specific”. Mutations that were only detected in the LN samples and the AF/CN remained higher than the LOD were categorized as “LN-specific”.

2.4. Analysis of the tumour mutational burden (TMB)

The TMB (mutations per megabase (Mb) DNA) was extrapolated using sequencing data from the gene panel containing 520 cancer-related genes and was determined by analysing the number of somatic mutations per Mb. The TMB was defined as the number of nonsynonymous alterations per megabase (Mb) of genome examined.

2.5. RNA isolation, quality control and mapping of sequenced data

RNA was extracted from formalin-fixed paraffin wax-embedded tissue by Genecast Biotechnology. Raw reads were preprocessed by removing low-quality sequences and dejection contamination (Trimmatic [17]), determining the distribution of the A/T/G/C content (RSeQC [18]), an removing rRNA (bowtie2 [19]) to obtain high-quality sequences (clean reads), and all subsequent analyses were based on clean reads to ensure consistent and reproducible data. Reference genome and gene annotation files were downloaded from the GENCODE website (https://www.gencodegenes.org/human/). Clean data were aligned to the reference genome using HISAT [20] (http://ccb.jhu.edu/software/htsat2/index.shtml). The FeatureCounts [21] algorithm was used to estimate the expression level of each gene. Gene expression was quantified using the FPKM (fragments per kilobase million mapped reads) method.

2.6. Differential gene expression analysis

The input data for differential gene expression analysis were read counts from the gene expression level analysis. The resulting P values were subjected to multiple test corrections using the Benjamini and Hochberg method to exclude false positives. Differentially expressed genes were identified when \( P < 0.05 \) and \( |\log_{2}(\text{fold change})| > 1 \). For experiments without biological replicates, we used the gene set enrichment analysis (ssGSEA) [23] for analysis, and the threshold was normally set to \( |\log_{2}(\text{fold change})| > 1 \text{ and } P < 0.05 \).

2.7. Analysis of the immune-related signature s

Single-sample gene set enrichment analysis (ssGSEA) (R library GSVA, method = “ssgsea”) [23] was used to analyse 28 immune cell subsets in the tumour immune infiltrate [24]. The innate anti-PD-1 resistance signature (IPRES) was calculated by determining the average Z score across all gene sets associated with tumour metastasis, as described in a published study [25].

2.8. Statistics

All statistical analyses were performed using R version 3.6.2 software (Institute for Statistics and Mathematics, Vienna, Austria; www.r-project.org). We used the DESeq2 [26] package in R software to screen differentially expressed genes between comparisons. DESeq2 package assumes gene counts generated from genomics data are distributed according to a negative binomial dispersion, normalizes for read depths and fits a generalized linear model. Demographic, clinical, genetic and pathologic characteristics were compared using the Chi-squared test or Fisher’s exact test (categorical variables), as applicable. The Wilcoxon rank-sum test was used to compare means
between groups, and paired Wilcoxon test was used to compare the dynamic changes of TMB from paired tissues.

2.9. Role of funders

The funders had no role on the study design, execution or interpretation of findings.

3. Results

3.1. Mutation spectrum of the matched primary breast tumour and lymph node metastases

The somatic mutations in samples from a single SLN metastasis and from MLNs were compared using an NGS panel of 520 cancer-related genes. In the primary breast tumours, the three most highly mutated genes were TP53, PIK3CA and GATA3, and the mutation frequency of PIK3CA was higher in the MLNs group (34/72, 47.22%) than in the SLN group (15/61, 24.59%) (Chi-squared test, \( P = 0.011 \)) (Fig. 1a).

In metastatic lymph nodes, TP53, PIK3CA and AKT1 were the most frequently mutated genes, and the mutation frequency of PIK3CA was also higher in the MLNs group (31/72, 43.06%) than in the SLN group (15/61, 24.59%) (Chi-squared test, \( P = 0.029 \)) (Fig. 1b). The genomic profile of paired primary breast tumours from patients with a single sentinel lymph node (SLN) metastasis was compared with patients with multiple metastatic lymph nodes (MLNs). Overall, 614 mutations spanning 228 genes were detected in the SLN metastasis group (Fig. 1c), and 657 mutations spanning 223 genes were detected in the MLNs group (Fig. 1d). The overall distribution of mutations was not significantly different between metastases in the SLN group and the MLNs group (Chi-squared test, \( P = 0.199 \); Table 1). In the SLN group, CN amplifications of FGFR3 and FGFR4 were detected in the primary breast tumours but not in the SLN metastases (FGFR3 11.5% (n=7) vs 0%; Fisher’s exact test, \( P = 0.013 \); FGFR4 9.8% (n=6) vs 0%; Fisher’s exact test, \( P = 0.027 \); Fig. 1e). In the MLNs group, no significant differences in mutations or amplifications were detected in the paired specimens from the primary tumour and nodal metastases (Fig. 1f and Supplementary Table S3).

3.2. Lymph node-specific gene mutations compared with primary breast tumours

Eighteen specific genetic abnormalities were detected in metastases from the SLN group (Supplementary Table S4) and 19 specific genetic abnormalities were identified in metastases in the MLNs group (Supplementary Table S5) that were not detected in matched breast primary tumours. When comparing the molecular profile for metastasis in the sentinel node versus the MLNs only 4 mutations (PIK3CA, CDK4, NFKBIA and CDKN1B) were shared between metastases from both the SLN and MLNs groups (Fig. 2). However, unique molecular features were observed in SLN metastasis (MEN1, MYC, DAXX, IL7R, NUP93, SUFU, NSD1, NFI, ATRX, ERCC1, FGR2, PREX2, PTCH1 and TTRAP) that were absent in the paired primary breast tumour, and a separate group of unique molecular features was identified in the MLN metastasis group (ASXL1, BRIP1, ERBB2, ERBB3, GSK3B, MAP3K1, NOTCH3, SPEN, SPN, HIST1H3B, PIK3R2, PRKDC, RARA, SPTA1 and CBL) that were absent in the paired primary breast tumour (Fig. 2). These changes were either lymph node-specific mutations or alterations with elevated abundance that were unique to lymph nodes. A heat map of the results summarized the correlation between novel lymph node-associated mutations and the genetic or clinical features of patients with either SLN (Fig. 3a) or MLNs (Fig. 3b) (Chi-squared test). Interestingly, significantly more novel lymph node-associated mutations were detected in SLN metastases from patients who also carried ERBB2 mutations (Chi-squared test, \( P = 0.045 \); Fig. 3c) or HER2 amplifications (Chi-squared test, \( P = 0.029 \); Fig. 3d). In metastases with novel lymph node-associated mutations, the tumour mutational burden (TMB) was significantly higher than that in metastases without novel mutations in both the SLN metastasis group (Chi-squared test, \( P = 0.045 \); Fig. 3e) and the MLNs group (Chi-squared test, \( P = 0.005 \); Fig. 3g). Significantly fewer novel lymph node-associated mutations were observed among AKTI mutant tumours in the MLNs group than in wild-type AKTI tumours (Chi-squared test, \( P = 0.046 \); Fig. 3f).

3.3. Genomic features of SLN- and MLN-positive breast cancer

We compared the tumour mutational burden (TMB) of metastases in the SLN group and in the MLNs group. No difference in the TMB was observed between the primary breast tumours from the SLN and MLNs group (\( P = 0.55 \); Fig. 4a). However, the TMB of metastases in the MLNs group was significantly higher than that of a single metastasis from the SLN group (Wilcoxon test, \( P = 0.0085 \); Fig. 4a). Interestingly, the TMB was significantly lower in metastases in the SLN group than in paired primary breast tumours (fold change = -1.20497 and paired Wilcoxon test \( P = 0.001 \); Fig. 4b). In contrast, the difference in the TMB did not reach statistical significance between the primary tumour and the metastases from the MLNs group (fold change = -0.42248 and paired Wilcoxon test \( P = 0.05 \); Fig. 4c).

3.4. Enhanced invasive capacity of tumour cells in the MLNs group compared with the SLN group

The tumour invasive capacity of the two groups was compared using gene set enrichment analysis (GSEA). We calculated the scores of 26 transcriptomic signatures associated with the tumour invasive capacity in both the SLN group and the MLNs group. We only selected samples from January 2018 to November 2018 for RNA extraction and detection to ensure the quality of the extracted RNA: 1) 20 from the SLN group (primary tumour n=20; metastatic sentinel lymph node n=18; and adjacent normal lymph nodes n=18) and 2) 28 from the MLNs group (primary tumour n=28; metastatic lymph nodes n=26; and adjacent normal lymph nodes n=17). A significant enrichment of invasive signatures was identified in both the primary tumour (Fig. 5a and c) and metastatic nodes (Fig. 5b and d) in the MLNs group compared to samples from the SLN group (Wilcoxon test, \( P < 0.05 \)). Four upregulated signatures were detected in both primary tumour and nodal metastases from the MLNs group, including 3 EMT signatures (VECCHI GASTRIC CANCER ADVANCED VS EARLY UP, ANASTASSIOU MULTICANCER INVASIVENESS SIGNATURE and LIEN BREAST CARCINOMA METAPLASTIC) and 1 angiogenesis signature (LU TUMOR VASCULATURE UP). Consistently, epithelial to mesenchymal transition (EMT) signaling and TGF-\( \beta \) signaling were significantly enriched in the MLNs group according to GSEA analyses (false discovery rate (FDR) = 0.049, normalized enrichment score (NES) = -1.94; FDR = 0.036, NES = -1.74, respectively) (Supplementary Fig. S2).

3.5. The antitumour immune response was different in patients with SLN-positive and MLN-positive breast cancer

The ssGSEA score was used to quantify the activity of 28 immune cell subpopulations (Fig. 6a). In primary tumours, the activities of central memory CD8 T cells, immature DCs and memory B cells were slightly increased in the MLNs group, while the activities of other immune cell subpopulations were not significantly different between the two groups. However, multiple immune cell populations, including activated and inhibited subpopulations, were enriched in metastatic and normal lymph nodes derived from the MLNs group compared with the SLN group (\( P \)-value in Supplementary Table S6). In patients with breast cancer, a high level of immune cell infiltration is associated with better clinical outcomes [27]. In primary tumours, the CD8/Treg ratio and CD8/EMT ratios were not significantly
Fig. 1. Molecular profiles of matched primary breast tumours and lymph node metastases. Oncoprint summarizing the genomic profiles of both sites of metastatic lymph nodes (SLNs and MLNs) and primary breast tumours. (a) In primary breast cancers, the mutation frequency of PIK3CA was higher in the MLNs group (34/72, 47.22%) than in the SLN group (15/61, 24.59%) (Chi-squared test, P = 0.011). (b) In the metastatic tumours, the mutational frequency of PIK3CA was also higher in the MLNs group (31/72, 43.06%) than in the SLN group (15/61, 24.59%) (Chi-squared test, P = 0.029). Oncoprint summarizing the genomic profiles of paired primary breast tumours and either (c) SLN or (d) MLN specimens. Only somatic alterations with a frequency of 4% or greater (≥4%) in the whole cohort are displayed. Stacked bar plot illustrating the distribution of mutation rates of various genes detected in either the primary breast tumour (breast) or lymph node (LN) specimens among the patients harbouring either (e) SLNs or (f) MLNs.
different between the SLN group and the MLNs group (Wilcoxon test, \( P < 0.05 \), Fig. 6b; Wilcoxon test, \( P > 0.05 \), Fig. 6c). In metastatic lymph nodes, the CD8/Treg ratio also showed no significant difference between the SLN group and MLNs group (Wilcoxon test, \( P > 0.05 \), Fig. 6d); however, the CD8/EMT ratio was higher in the MLNs group (Wilcoxon test, \( P < 0.05 \), Fig. 6f) and CD8/EMT ratio (Wilcoxon test, \( P < 0.05 \), Fig. 6g) were significantly decreased in normal lymph nodes from the MLNs group.

### Table 1

| Mutation types       | Dataset | Total mutations | Matched + Uncertain n(%) | Breast-specific n(%) | LN-specific n(%) | P-value |
|----------------------|---------|-----------------|--------------------------|----------------------|-----------------|---------|
| CN amplification     | SLN     | 272             | 148(54.4%)               | 77(28.3%)            | 47(17.3%)       | 0.116   |
| CN deletion          | MLN     | 264             | 166(62.9%)               | 57(21.6%)            | 41(15.3%)       |         |
| CN deletion          | SLN     | 1               | 1(100%)                  | 0(0%)                | 0(0%)           |         |
| CN deletion          | MLN     | 4               | 4(100%)                  | 0(0%)                | 0(0%)           | 1       |
| Frameshift           | SLN     | 36              | 23(63.9%)                | 9(25%)               | 4(11.1%)        |         |
| Frameshift           | MLN     | 41              | 28(68.3%)                | 5(12.2%)             | 8(19.5%)        | 0.265   |
| Fusion               | SLN     | 14              | 3(21.4%)                 | 5(35.7%)             | 6(42.9%)        |         |
| Fusion               | MLN     | 19              | 9(47.4%)                 | 6(31.6%)             | 4(21.1%)        | 0.247   |
| Indel                | SLN     | 8               | 6(75%)                   | 0(0%)                | 2(25%)          |         |
| Indel                | MLN     | 12              | 8(66.7%)                 | 4(33.3%)             | 0(0%)           | 0.057   |
| LGR                  | SLN     | 0               | 0(0%)                    | 0(0%)                | 0(0%)           |         |
| LGR                  | MLN     | 2               | 2(100%)                  | 0(0%)                | 0(0%)           |         |
| Missense             | SLN     | 232             | 138(59.5%)               | 48(21.1%)            | 45(19.4%)       |         |
| Missense             | MLN     | 260             | 154(59.2%)               | 50(19.2%)            | 56(21.5%)       | 0.782   |
| Nonsense             | SLN     | 21              | 14(66.7%)                | 4(19%)               | 3(14.3%)        |         |
| Nonsense             | MLN     | 34              | 20(58.8%)                | 8(23.5%)             | 6(17.6%)        | 0.844   |
| Splice-site          | SLN     | 30              | 18(60%)                  | 7(23.3%)             | 5(16.7%)        |         |
| Splice-site          | MLN     | 20              | 12(60%)                  | 4(20%)               | 4(20%)          | 0.935   |
| Stop-lost            | SLN     | 0               | 0(0%)                    | 0(0%)                | 0(0%)           |         |
| Stop-lost            | MLN     | 1               | 0(0%)                    | 1(100%)              | 0(0%)           |         |
| Total                | SLN     | 614             | 351(57.2%)               | 151(24.6%)           | 112(18.1%)      | 0.199   |
| Total                | MLN     | 657             | 403(61.3%)               | 135(20.5%)           | 119(18.1%)      |         |

![Fig. 2](image)

**Fig. 2.** Mutations unique to breast cancer lymph node metastases

Eighteen genetic mutations in the single metastasis (SLN) group and 19 genetic mutations in multiple metastases (≥ 3) in the MLNs group were not identified in the matched primary breast tumour. Only 4 distinct mutations (PIK3CA, CDK4, NFKBIA and CDKN1B) were identified in metastases from both lymph node locations (in red).

### 4. Discussion

The molecular signature of breast cancer metastases and the host’s tumour immune characteristics were different when comparing a single nodal metastasis with multiple nodal metastases. We identified a large number of molecular abnormalities unique to nodal metastases that were not expressed in paired primary breast tumours. Thus, we identified a distinct group of 14 mutations in single sentinel node metastasis and a different group of 15 mutations in multiple nodal metastases. Only 4 distinct mutations were conserved in both lymph node settings. Interestingly, the TMB was significantly lower in a single nodal metastasis than that of the TMB in the paired primary tumour, while, in contrast, the TMB in the MLN tumor was higher than that in the SLN tumor and at the same higher level as the paired primary tumor. Parameters related to the invasive properties of the tumors were also higher in both the primary tumours and their regional metastases in the MLNs group compared to the SLN group. Finally, the distribution of immune subpopulations in the tumour-infiltrating lymphocytes was different when comparing the SLN and MLNs groups. Taken together, these results are consistent with a hypothesis that the molecular profile of mutational events and TMB are distinct when comparing a single nodal metastasis from the SLN group compared to multiple metastases from the MLNs group. Furthermore, the profiles of a single metastases in the sentinel node correlate with a more favourable biology for the host. We are only able to conjecture whether the more adverse molecular profile in the MLNs group results from a less differentiated primary breast tumour or whether these nodal metastases developed additional mutations and a higher TMB while they “incubated” in the sentinel lymph node.

Previous findings indicate that PIK3CA is one of the most commonly mutated genes in breast cancer [28]. PI3K-Akt-mTOR signalling pathways play important roles in cell proliferation, signalling, and the metastatic potential [29]. In our study, metastases from the MLNs group more frequently contained mutations in PIK3CA (46%) and a higher TMB while they exhibited a less differentiated primary breast tumour, while, in contrast, the TMB in the MLN tumor was higher than that in the SLN tumor and at the same higher level as the paired primary tumor. Parameters related to the invasive properties of the tumors were also higher in both the primary tumours and their regional metastases in the MLNs group compared to the SLN group. Finally, the distribution of immune subpopulations in the tumour-infiltrating lymphocytes was different when comparing the SLN and MLNs groups. Taken together, these results are consistent with a hypothesis that the molecular profile of mutational events and TMB are distinct when comparing a single nodal metastasis from the SLN group compared to multiple metastases from the MLNs group. Furthermore, the profiles of a single metastases in the sentinel node correlate with a more favourable biology for the host. We are only able to conjecture whether the more adverse molecular profile in the MLNs group results from a less differentiated primary breast tumour or whether these nodal metastases developed additional mutations and a higher TMB while they “incubated” in the sentinel lymph node.

**4.1. PIK3CA**

PIK3CA is one of the most commonly mutated genes in breast cancer [28]. PIK3-Akt-mTOR signalling pathways play important roles in cell proliferation, signalling, and the metastatic potential [29]. In our study, metastases from the MLNs group more frequently contained mutations in PIK3CA (46%) than a metastasis from the SLN group (26%). Donovan et al. showed that tumours in which breast cancer stem cells have a genetic abnormality in PIK3/Akt signalling are significantly more likely to manifest nodal metastases [30]. In both the SLN and MLNs groups, PIK3CA mutations were identified in metastatic lymph nodes that were not found in primary breast cancer. Our results support a potential role...
Fig. 3. Distinct molecular features of paired primary breast tumours and either single nodal metastasis (SLN group) or multiple nodal metastases (MLNs group).

Heat map summarizing the correlation between novel lymph node-associated mutations and the genetic or clinical features of patients with either (a) SLN or (b) MLNs (P-values were obtained using the Chi-squared test; only significant P-values are shown in the graph). Clinical and molecular features showing statistically significant correlations with novel lymph node-associated mutations are expanded in Figs C-E for patients with SLN and F-G for patients with MLNs. (c-e) Bar plot illustrating the significantly more novel lymph node-associated mutations among patients with (c) ERBB2-mutated (Chi-squared test, P=0.045) or (d) HER2+ (Chi-squared test, P=0.029) SLN. (e and g) Box plots illustrating the significantly higher TMB in the lymph nodes of patients with (e) SLN and (g) MLN who carried novel lymph node-associated mutations than those with non-novel mutations. (f) Bar plot illustrating the lower detection of novel lymph node-associated mutations among patients with MLNs who harbour AKT1 mutations (Chi-squared test, P=0.046).
for PIK3CA in the development of breast cancer lymph node metastases. Arsenic et al. also found that PIK3CA mutations were strongly correlated with the lymph node status (N+) \[31\]. The correlation between PIK3CA mutations and lymph node metastasis suggests that activation of the PI3K/Akt pathway might increase the invasive properties of cancer cells in the lymph nodes. This hypothesis is supported by the fact that mutations of PIK3CA increase PIP3 levels and induce cellular transformation and increased mobility \[32, 33\].

In addition to PIK3CA mutations, CDK4, NFKBIA and CDKN1B were lymph node-specific gene mutations that varied in both the SLN and MLNs groups compared with the primary breast tumours. According to previous studies, CDK4 is a crucial promoter of tumour growth in HR+ breast cancer, cooperating with ER pathway activation \[34\]. CDK4/6 inhibitors prevent cell cycle progression and improve survival outcomes in patients with advanced HR+/HER2- breast cancer \[35, 36\]. CDKN1B was identified as an inhibitor of cell cycle progression because of its antiproliferative activity that inhibits cyclin-CDK complexes \[37, 38\]. Recent studies have revealed that CDKN1B mutations are driver genetic lesions in a significant percentage of cases, particularly in luminal breast cancer \[38\]. Researchers have universally acknowledged that the NF-κB protein complex is required for the initiation of inflammation and the development of immunity \[39\]. Patients with genetic defects in NFKBIA generally experience severe immunodeficiencies with impaired cellular responses to immune stimuli \[40\].

The FGF3 and FGF4 genes are located side-by-side and are also located in close proximity to the FGF19 and CCND1 genes (within 0.2 Mb of the 11q13 region) \[41, 42\]. Coamplification of both genes was observed in various human tumours at frequencies of 13–60% \[41\]. In the SLN group, among the genes detected from the paired specimens, CN amplifications in FGF3 and FGF4 were primarily detected in primary breast tumours. However, in the MLNs group, no differences were observed among the genes detected in the paired specimens. Therefore, the detection of FGF3 or FGF4 amplifications in metastatic lymph nodes might predict metastases in multiple lymph nodes. In the SLN group, significantly more novel lymph node-associated mutations were detected in patients with ERBB2-mutated tumours or HER2-positive tumours. Therefore, the interplay of HER2 amplifications might play a role in promoting the metastatic process in the early stage of lymph node metastasis.

The tumour invasive capacity of the two groups was compared using GSEA. Three EMT signatures and 1 angiogenesis signature were upregulated in both primary tumours and in metastases from the

---

**Fig. 4. Genomic features of SLN- and MLN-positive breast cancer**

(a) TMB for primary tumours and metastatic lymph nodes from the SLN group and MLNs group (Chi-squared test). The TMB was remarkably reduced in metastatic LNs from the (b) SLN group (fold change = -1.20497 and paired Wilcoxon test P = 0.001), but not in the (c) MLNs group (fold change = -0.42248 and paired Wilcoxon test P = 0.05), and was analysed with paired Wilcoxon tests. Met means metastatic lymph nodes.
MLNs group. EMT programs are believed to reflect a loss of epithelial gene expression signatures and morphologies that are associated with mesenchymal cells and their enhanced migratory and invasive behaviours [43]. Based on these results, both primary tumours and their regional metastases had more characteristics in the MLNs group than those in the SLN group. Tumour-draining lymph nodes are modified prior to cancer cell arrival by stromal remodelling and immune cell recruitment [44]. Numerous studies have shown that the normal ratio of effector T cells to regulatory T cells changes during cancer progression. For example, in ovarian cancer, a low CD8+/Treg ratio is associated with a poor prognosis [45]. Baras AS et al. [46] found that although CD8 and Treg infiltration in bladder cancer is not related to the neoadjuvant chemotherapy response, the ratio of CD8 to Treg cells is indeed related to neoadjuvant efficacy. The CD8/Treg ratio has also been shown to predict the prognosis of cervical squamous cell carcinoma [47] and colorectal cancer [48]. Mao W et al. found that the CD8/Treg ratio is closely related to the effect of anti-CTLA-4 immunotherapy on prostate cancer [49]. In addition, some recent studies have found that the tumour EMT phenotype is closely related to tumour immunosuppression by inducing PD-L1 expression. Research by Chen L et al. showed that PD-L1 expression is controlled by the EMT transcription factor zeb1 [50]. Hugo et al. also found that the transcriptomic characteristics of the upregulated EMT genes are associated with resistance to anti-PD-1 therapy [51]. Based on these findings, we are eager to understand whether differences exist in the CD8/Treg ratio and CD8/EMT ratio between the primary lesion, SLN and MLNs. In addition, we investigated whether the nearby lymph nodes (clinically normal in appearance and size) were affected by tumour cells in upstream lymph nodes. Surprisingly, both the CD8/Treg ratio and CD8/EMT ratio were significantly higher in normal lymph nodes harvested from patients in the SLN group compared to those from the MLNs group, indicating that even normal appearing lymph nodes from the MLNs group may have a favourable microenvironment for growth of multiple lymph node metastases. These data suggest that normal lymph nodes with a high antitumour immune response might help prevent multiple lymph node metastases.

Fig. 5. Tumour cells in the MLNs group had an enhanced invasive capacity compared with the SLN group

Functional enrichment was computed using a gene set enrichment analysis (GSEA). We observed an enrichment of invasive signatures in both (a and c) primary tumours and (b and d) metastatic lymph nodes derived from the MLNs group compared to the SLN group (Wilcoxon test, P < 0.05).
This study had some limitations. The abundance of somatic mutations detected in primary breast tumours and lymph nodes might have been affected by the size of the tumour cell population used for isolation. For instance, the use of samples with a relatively low tumour cell population might result in a false negative call of a CNV due to a copy number value lower than the limit of detection of the assay. Hence, the allelic fraction of the SNVs and the CN value for the CNVs were adjusted based on the tumour cell percentage. This study was conducted at a single centre and hence may lack generalizability, which may be considered one of the limitations.

In conclusion, single lymph node metastasis (from a sentinel node) and multiple lymph node metastases (i.e., > 3 metastases) contained some significant differences in their molecular profiles and anti-tumour immune characteristics. The results are associated with more aggressive tumour characteristics and less favourable immune responses in patients with multiple nodal metastases compared to those with a single nodal metastasis in the sentinel node.

Declaration of Competing Interest

Xueying Wu and Henghui Zhang are employees of Beijing Genecast Biotechnology Co., Beijing, China. Min Li and Jing Liu are employees of Burning Rock Biotech, Guangzhou, Guangdong, China. The other authors declare that they have no competing of interest.

Contributors

All authors participated in the planning and execution of this study or analysis of the study data. NL, BC and GZ designed the study. BC, GZ, JL, WX, XL, CI, HM, KL, YW, LC, MJ, CR, LW, GW, JiaL, YL, XC and YZ participated in sample collection, sample processing, collection of the clinical information and data analysis. HZ and XW verified the underlying data. BC, XW, HZ, ML, and JingL performed the statistical analyses. NL, BC, CB, XW and HZ provided critical comments and suggestions, and revised the manuscript. All authors read and approved the final version of the manuscript. Burning Rock Biotech (Guangzhou, China) conducted the NGS panel covering 520 cancer-related genes (OncoScreen Plus); Beijing Genecast Biotechnology Co. conducted the whole-exome sequencing of all included samples.

Acknowledgements

We thank Dr. Jiayan Wu from Berry Oncology Corporation for her comments on the revision of this manuscript. This work was supported by funds from the High-level Hospital Construction Project (DfJH201921, Bo Chen), the National Natural Science Foundation of China (81902828, Bo Chen; 82002928, Jianguo Lai), the Fundamental Research Funds for the Central Universities (yzy12192230, Bo Chen), and the Medical Scientific Research Foundation of Guangdong
Province (B2019039, Bo Chen). The funding agencies had no role in the design of the study; the collection, analysis, and interpretation of data; or the writing of the manuscript.

Data Sharing Statement
All data can be viewed in NODE (http://www.biosino.org/node) by pasting the accession (OEP001295 and OEP001992) into the text search box or through the URLs http://www.biosino.org/node/project/detail/OEP001295 and http://www.biosino.org/node/project/detail/OEP001992.

Genomic data could be shared after a reasonable request to the corresponding authors in accordance with Chinese law for genomic data. The data used or analysed during this study are included in this article and available from the corresponding author upon reasonable request.

Supplementary materials
Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ebiom.2021.103542.

References
[1] Kim R, Emi M, Tanaka K, Aihrio K. Immunology of the sentinel lymph node and its potential role for antitumor immunity. Lancet Oncol 2006;7(12):1006–16.
[2] Giuliano AE, Edge SB, Hortobagyi GN. Eighth Edition of the AJCC Cancer Staging Manual. Breast Cancer. Ann Oncol 2012;23(8):1783–5.
[3] Matsuura K, Yamaguchi Y, Ueno H, Osaki A, Aihrio K, Toge T. Maturation of dendritic cells and T-cell responses in sentinel lymph nodes from patients with breast carcinoma. Cancer 2009;106(6):1227–36.
[4] Shu S, Cochran AJ, Huang RR, Morton DL, Maercker HT. Immune responses in the draining lymph nodes against cancer: implications for immunotherapy. Cancer Metastasis Rev 2006;25(2):213–42.
[5] Zuckerman NS, Yu H, Simons DK, Bhattacharyya N, Carcano-Cavaosas V, Yan N, et al. Altered local and systemic immune profiles underlie lymph node metastasis in breast cancer patients. Int J Cancer 2013;132(11):2573–47.
[6] McAllister SS, Weinberg RA. The tumour-induced systemic environment as a critical regulator of cancer progression and metastasis. Nat Cell Biol 2016;18(7):717–27.
[7] Smith HA, Kang Y. The metastasis-promoting roles of tumor-associated immune cells. J Mol Med 2013;91(4):411–29.
[8] Klebe M, Freund C, Kriegunnn M, Kriegunnn K, Albrecht T, Thewes V, et al. Frequent molecular subtype switching and gene expression alterations in lung and pleural metastasis from luminal A-type breast cancer. JCO Precision Oncol 2020;4.
[9] Thompson AM, Jordan LB, Quinlan P, Anderson E, Skene A, Dewar JA, et al. Prognostic and Predictive Markers of Response to Anti-PD-1 Therapy in Metastatic Melanoma. Ann Surg Oncol 2017;24(9):2561–71.
[10] Turner NC, Slamon DJ, Boer J, Bondarenko I, Knudtson JD, Saaristie AL, et al. Elements of the PD-1/PD-L1 Pathway Are Essential for Response of Metastatic Melanoma to Pembrolizumab. Clin Cancer Res 2017;23(21):5686–97.
[11] Nowosad A, Besson A, CDK11B/p27 regulates autophagy via the control of Ragu-1 and MTOR activity in amino acid-deprived cells. Autophagy 2020.
[12] Cusan M, Mungo G, De Marco Zompit M, Segatto I, Belletti B, Baldassarre C. Land- scape of CDKN1B mutations in luminal breast cancer and other hormone-driven human tumors. Front Endocrinol (Lausanne) 2018;9:393.
[13] Nowosad A, Besson A, CDK11B/p27 regulates autophagy via the control of Regulator 1 and MTOR activity in amino acid-deprived cells. Autophagy 2020.
[14] Finns C, Crown JP, Lang I, Boer J, Bondarenko I, Kylhgk SY, et al. The cyclin-depent kinase 4/6 inhibitor panobinobilin in combination with letrozole versus letrozole alone as first-line treatment of oestrogen receptor-positive, HER2-negative, advanced breast cancer (PALOMA-1/BR3-18): a randomised phase 2 study. Lancet Oncol 2015;16(1):25–35.
[15] de Groot AF, Kuipers CJ, Kroep JR. CDK4/6 inhibition in early and metastatic breast cancer: A review. Ther Adv Breast Dis 2017;10:60–8.
[16] Nowosad A, Besson A. CDK11B/p27 regulates autophagy via the control of Regulator 1 and MTOR activity in amino acid-deprived cells. Autophagy 2020.
[17] Finns C, Crown JP, Lang I, Boer J, Bondarenko I, Kylhgk SY, et al. The cyclin-dependent kinase 4/6 inhibitor panobinobilin in combination with letrozole versus letrozole alone as first-line treatment of oestrogen receptor-positive, HER2-negative, advanced breast cancer (PALOMA-1/BR3-18): a randomised phase 2 study. Lancet Oncol 2015;16(1):25–35.
[18] Matsuura K, Yamaguchi Y, Ueno H, Osaki A, Aihrio K, Toge T. Maturation of dendritic cells and T-cell responses in sentinel lymph nodes from patients with breast carcinoma. Cancer 2009;106(6):1227–36.
[19] Shu S, Cochran AJ, Huang RR, Morton DL, Maercker HT. Immune responses in the draining lymph nodes against cancer: implications for immunotherapy. Cancer Metastasis Rev 2006;25(2):213–42.
[20] Zuckerman NS, Yu H, Simons DK, Bhattacharyya N, Carcano-Cavaosas V, Yan N, et al. Altered local and systemic immune profiles underlie lymph node metastasis in breast cancer patients. Int J Cancer 2013;132(11):2573–47.
[21] McAllister SS, Weinberg RA. The tumour-induced systemic environment as a critical regulator of cancer progression and metastasis. Nat Cell Biol 2016;18(7):717–27.
[22] Smith HA, Kang Y. The metastasis-promoting roles of tumor-associated immune cells. J Mol Med 2013;91(4):411–29. 