Published OnlineFirst June 11, 2018; DOI: 10.1158/1078-0432.CCR-17-3374

Precision Medicine and Imaging

Clinical Cancer Research

The Spatiotemporal Evolution of Lymph Node Spread in Early Breast Cancer

Peter Barry[1], Alexandra Vatsiou[2], Inmaculada Spiteri[2], Daniel Nichol[2], George D. Cresswell[2], Ahmet Acar[2], Nicholas Trahearn[2], Sarah Hrebien[3], Isaac Garcia-Murillas[3], Kate Chkhaidze[2], Luca Ermini[4], Ian Said Huntingford[5], Hannah Cottom[3], Lila Zabaglo[6], Konrad Koelble[6], Saira Khalique[6], Jennifer E. Rusby[1], Francesca Muscara[1], Mitch Dowsett[6], Carlo C. Maley[2,7], Rachael Natrajan[3], Yinyin Yuan[4], Gaia Schiavon[3], Nicholas Turner[3], and Andrea Sottoriva[2]

Abstract

Purpose: The most significant prognostic factor in early breast cancer is lymph node involvement. This stage between localized and systemic disease is key to understanding breast cancer progression; however, our knowledge of the evolution of lymph node malignant invasion remains limited, as most currently available data are derived from primary tumors.

Experimental Design: In 11 patients with treatment-naïve node-positive early breast cancer without clinical evidence of distant metastasis, we investigated lymph node evolution using spatial multiregion sequencing ($n = 78$ samples) of primary and lymph node deposits and genomic profiling of matched longitudinal circulating tumor DNA (ctDNA).

Results: Linear evolution from primary to lymph node was rare (1/11), whereas the majority of cases displayed either early divergence between primary and nodes (4/11) or no detectable divergence (6/11), where both primary and nodal cells belonged to a single recent expansion of a metastatic clone. Divergence of metastatic subclones was driven in part by APOBEC. Longitudinal ctDNA samples from 2 of 7 subjects with evaluable plasma taken perioperatively reflected the two major evolutionary patterns and demonstrate that private mutations can be detected even from early metastatic nodal deposits. Moreover, node removal resulted in disappearance of private lymph node mutations in ctDNA.

Conclusions: This study sheds new light on a crucial evolutionary step in the natural history of breast cancer, demonstrating early establishment of axillary lymph node metastasis in a substantial proportion of patients. *Clin Cancer Res; 24(19); 4763–70. ©2018 AACR.*

Introduction

Breast cancer is characterized by high genomic and transcriptomic diversity, both between (1–3) and within patients (4–8). This inherent complexity is fully consistent with a clonal evolution model of cancer (9, 10). The cancer evolution paradigm provides a biological explanation of experimental observations and may also lead to more accurate predictions of the future course of the disease, in particular prognostication and the emergence of treatment resistance (9).

Currently, clinicopathologic parameters such as age, tumor grade and stage, ER expression, and HER2 expression have been integrated into scoring systems to estimate the probability of recurrence and death from breast cancer (11, 12). Moreover, gene expression profiles provide additional prognostic and/or predictive information regarding adjuvant chemotherapy in ER-positive early breast cancer, and their clinical utility is being prospectively evaluated in large randomized clinical trials (13). Large meta-analyses have indicated that in early breast cancer, the most important prognostic factor is lymph node involvement (14–16). This clinical stage represents a potentially intermediate evolutionary step between localized disease and metastatic dissemination, and it is therefore of crucial importance to understand progression. Micrometastases can also be present at diagnosis of some early breast cancers, and ultrasensitive methods to analyze circulating tumor DNA (ctDNA) have recently helped interrogate such deposits that can subsequently result in overt metastatic recurrence (17). Hence, combined genomic analyses of primary, lymph nodes, and ctDNA are necessary to understand metastatic progression in cancer—also in light of findings in other cancer types where metastatic dissemination was found to be decoupled from lymphatic spread in a subset of cases (18, 19).

[1]Department of Surgery, Breast Unit, Royal Marsden Hospital, London, United Kingdom. [2]Evolutionary Genomics and Modelling Lab, Centre for Evolution and Cancer, The Institute of Cancer Research, London, United Kingdom. [3]Breast Cancer Now Research Centre, The Institute of Cancer Research, London, United Kingdom. [4]Centre for Evolution and Cancer, The Institute of Cancer Research, London, United Kingdom. [5]Department of Pathology, Mater Dei Hospital, Msida, Malta. [6]Ralph Lauren Breast Cancer Research Centre, Royal Marsden Hospital, London, United Kingdom. [7]Biodesign Institute, Arizona State University, Tempe, Arizona.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

Current address for K. Koelble: Neuropathology, University Hospital, Erlangen, Germany; and current address for G. Schiavon: IMED Biotech Unit, AstraZeneca, Cambridge, United Kingdom.

P. Barry, A. Vatsiou, and I. Spiteri contributed equally to this article.

Corresponding Author: Andrea Sottoriva, The Institute of Cancer Research, Brookes Lawley Building, 2E12.6, London SM2 5NG, UK. Phone: 44-20-8722-4072; E-mail: andrea.sottoriva@icr.ac.uk

doi: 10.1158/1078-0432.CCR-17-3374

Published OnlineFirst June 11, 2018; DOI: 10.1158/1078-0432.CCR-17-3374

Translational Relevance

Lymph node involvement is an important prognostic factor in breast cancer, but the patterns of cancer evolution during lymph node spread are poorly understood. Furthermore, circulating tumor DNA as a potential biomarker to guide treatment is relatively unexplored in this context. In this prospective study, we used multiregion sequencing of multiple samples of primary tumors and multiple matched lymph nodes and showed remarkable differences in the patterns of lymph node infiltration between subsets of patients. Specifically, one subgroup of cases showed early lymph node divergence. Those patterns were reflected in the circulating tumor DNA of the same patients and demonstrated (to our knowledge for the first time) disappearance of private nodal mutations after surgical resection. Divergence was also associated with APOBEC activity. Together, these results suggest that evolutionary patterns of lymph node infiltration may be important to predict the course of the disease in individual patients and, combined with circulating tumor DNA sampling, could aid patient stratification and personalized, precision medicine.

Here, we sought to study lymph node spread from an evolutionary perspective by analyzing 78 multiregion samples taken from untreated primary tumors and lymph nodes, as well as 7 longitudinal ctDNA samples of a selected cohort of 11 primary breast cancers that had biopsy-proven ipsilateral axillary lymph node spread without clinical evidence of distant metastatic disease (see Fig. 1 and Supplementary Table S1 for clinical details). We used whole-exome sequencing (WES), whole-genome sequencing (WGS), and targeted deep sequencing data, combined with phylogenomics analysis, to understand the dynamics of lymph node spread.

We found evident patterns of early divergence between primary and lymph node deposits in a subset of patients in our cohort and showed that these patterns were reflected in ctDNA. We also found that APOBEC activity contributed to such early divergence. Finally, we show proof-of-principle loss of ctDNA mutations private to nodes following surgical resection.

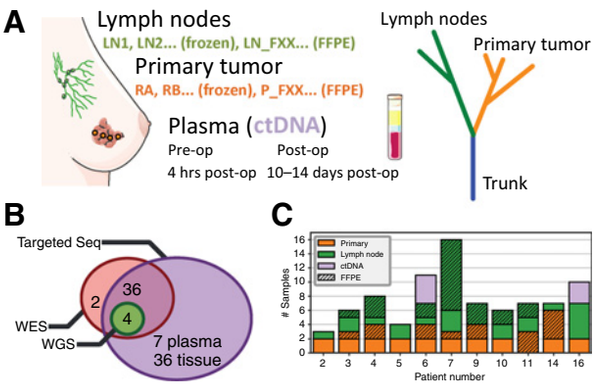

Figure 1.

Spatiotemporal genomic profiling of lymph node evolution in breast cancer. **A,** Multiregion sampling and genomic profiling of primary (regions RA, RB, . . ., 1–6 cm apart) and lymph node (LN1, LN2, . . .) samples (total samples $n = 78$) from a selected cohort of 11 patients with early breast cancer with lymph node involvement without distant metastases. Longitudinal ctDNA samples were taken before and after surgery from 2 patients ($n = 7$ samples). **B,** Sequencing analysis was performed with WES ($n = 42$), WGS ($n = 4$), and targeted deep sequencing (Targeted Seq) of a cohort-specific panel ($n = 76$) to identify mutational and copy-number profiles. **C,** The number of samples per location for each patient in the cohort [solid = fresh-frozen (FF), stripes = formalin-fixed paraffin-embedded (FFPE)].

Materials and Methods

Patient cohort and samples

Samples were collected from 11 patients with breast cancer with positive axillary nodes. Patients had not received any treatment prior to surgery. The median age of patients in this cohort was 56 years (range, 38–81). Several lymph nodes and primary tumor specimens were collected from each patient. Samples were either paraffin embedded after formalin fixation or snap frozen immediately after resection.

Whole peripheral blood was collected from each patient for germline DNA and plasma ctDNA taken at four time points: (1) intraoperatively before tumor resection, (2) intraoperatively immediately after tumor resection, (3) 4 hours postoperatively, and (4) 10 to 14 days postoperatively, at the follow-up visit. Blood at each time point (1–4) was collected into 3 × 10 mL EDTA or Streck tubes and was centrifuged at 1,600 × *g* for 20 minutes for a single spin. Plasma and buffy coat were collected and stored at −80°C. ctDNA was extracted from 5 mL plasma using the QIAamp circulating nucleic acid kit (catalog number 55114) from QIAGEN according to the manufacturer's instructions. Briefly, plasma was lysed with proteinase K and ACL for 30 minutes at 60°C, with carrier RNA in AVE added. Buffer ACB was added, and the sample was passed through a QIAamp mini spin column to bind the DNA. DNA was washed with ACW1, ACW2, and 100% ethanol before centrifugation at 14,000 RPM for 3 minutes, dried for 10 minutes, eluted into 50 μL AVE buffer, and stored at −20°C. Extracted DNA was quantified by digital droplet PCR (ddPCR) using a TaqMan copy-number reference assay for RNAse P (Life Technologies). ddPCR reactions were assembled using 1 μL of eluate and 10 μL of ddPCR supermix for probes (Bio-Rad Laboratories) for a total reaction volume of 20 μL. The reaction was partitioned into approximately 20,000 droplets on a Bio-Rad Laboratories QX200 droplet generator. PCR of the emulsified reaction was performed in 96-well plates on a G-Storm GS4 thermocycler for 40 PCR cycles with 60°C annealing temperature. The plates were read on a Bio-Rad Laboratories QX200 droplet reader, and the DNA concentration was calculated using Quantasoft software (Version 1.4.0.99). At least 2 NTC wells were included in each quantification run.

Clinical and histopathologic data from the patient cohort can be found in Supplementary Table S1. The study protocol was approved by an institutional research ethics committee (reference number 13/LO/1015). All patients gave their written informed consent to participate before enrolling in the study. The study was carried out in accordance with the principles of the International Conference on Harmonization Guideline for Good Clinical Practice and the Declaration of Helsinki. See Supplementary Material and Methods for details on sample preparation.

WES, WGS, and targeted sequencing

For each of the 11 patients in our cohort, 500 ng of DNA from 2 primary breast tumor specimens and 1 to 5 involved lymph nodes

Published OnlineFirst June 11, 2018; DOI: 10.1158/1078-0432.CCR-17-3374

were whole-exome sequenced (SureSelect human all exon V2). The set comprised 40 fresh-frozen (FF) tissue samples (20 primary tumor and 20 lymph node), 2 formalin-fixed paraffin-embedded (FFPE) specimens (both primary tumor tissues), and 11 germline samples (buffy coats). Exome sequencing data had a mean coverage of 154X. Whole-genome libraries were prepared from 30 to 100 ng of genomic DNA with the NebNext Ultra II kit following the manufacturer's instructions. Genomic DNAs were sheared in a Diagenode sonicator prior to library preparation. Whole-genome median coverage was 38. Further, a total of 807 exonic single-nucleotide variants (SNV) were selected for targeted validation. All but two samples used for WES were included in the targeted validation panel. In addition, we included DNA from 36 FFPE manually microdissected specimens and 7 ctDNA samples. A custom SureSelect XT2 panel (Agilent Technologies) was used to generate targeted capture libraries from these 83 samples (for 2 samples, there was not enough DNA for targeted sequencing; therefore, information from only the exome sequencing was used) following the manufacturer's recommendations. Mean coverage for targeted sequencing was 1,813X with 98% validation. All libraries were sequenced on an Illumina HiSeq 2500. See Supplementary Material and Methods for details regarding bioinformatics analysis.

Results

Intratumor heterogeneity in lymph node–positive breast cancers

Using WES, we profiled 40 FF and 2 FFPE samples from the 11 patients, as well as matched normal (FF buffy coat), obtaining a mean depth of 154X. For each patient, we had at least two regions from the primary tumor, taken 1 to 6 cm apart, and one lymph node (Fig. 1). Extensive intratumor heterogeneity (ITH) was evident in our cohort, with an average of 73.5% of variants considered to be subclonal (Supplementary Figs. S1 and S2). A total of 807 mutations were selected for custom targeted deep sequencing validation (mean depth of 1,813X, 98% validation rate). We also applied the same panel to 36 additional FFPE samples from the same patients (all those available with >50% tumor content, see Supplementary Table S2), confirming that the original FF samples were representative of ITH both in the primary and the lymph node deposits, and the observed patterns were not due to sampling bias (Fig. 2A). Estimated purity, ploidy, and copy-number profiles (Fig. 2B) were used to calculate CCFs for both WES and targeted sequencing profiles, as presented in Fig. 2A and Supplementary Fig. S2 (see Supplementary Tables S3–S6 for values). For those samples where only targeted sequencing was available (i.e., additional FFPE samples), purity and ploidy estimates could not be calculated, and we therefore reported presence/absence of the mutations (e.g., Fig. 2A, FFPE samples are set to CCF = 1/0). As macrodissected samples represented a small localized region of the tumor, we did not find any evident subclonal structure *within* each sample, although the limited mutational burden of breast cancer, combined with exome sequencing, precluded reliable subclonal analysis within such samples.

The mutational landscape of our cohort was consistent with previous studies (2, 3, 6), with TP53 and PIK3CA being the most commonly mutated drivers (Fig. 2A: tier 1 cancer genes, most likely drivers in black; tier 2 cancer genes, possibly drivers

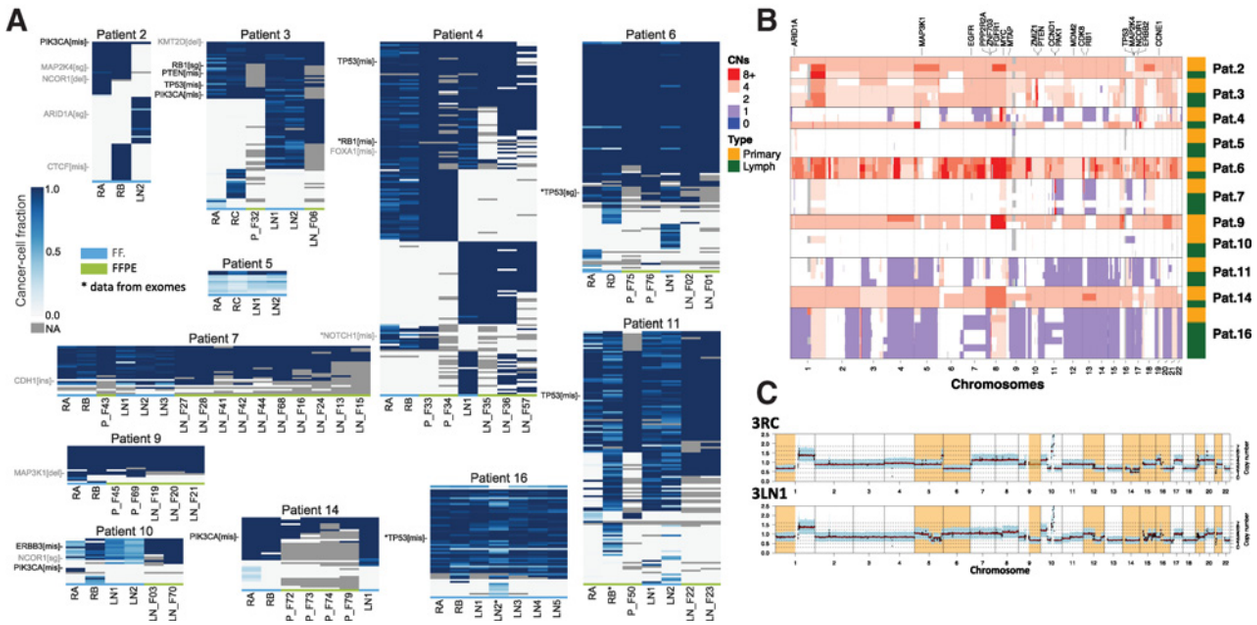

Figure 2.

Distinct modes of lymph node evolution. **A,** Deep targeted sequencing of a cohort-specific panel derived from WES. Heatmaps indicate CCF of a mutation or indel in different samples from the same patient (gray = NA, not enough coverage or variant does not overlap with a copy-number segment). For FFPE samples (marked in green at the bottom), CCFs were not available and presence/absence is reported (CCF = 1 or CCF = 0). Tier 1 cancer genes, most likely drivers, are annotated in black, and tier 2 cancer genes, possible drivers with uncertain pathogenicity, are annotated in gray. For samples/variants marked with (*), we report exome data as targeted were not available. **B,** Copy-number aberrations in all samples showing differences in copy-number status between primary and lymph nodes in the divergent subgroup. CNs, copy numbers; Pat., patient. **C,** WGS for a tumor and lymph node sample of patient 3; divergent copy-number regions shown in orange.

Published OnlineFirst June 11, 2018; DOI: 10.1158/1078-0432.CCR-17-3374

but uncertain pathogenicity in gray). Copy-number alterations (CNA) were widespread, with patterns consistent with the profile of primary tumors (1, 3), such as 1q and 8q gains and 8p loss (Fig. 2B and Supplementary Fig. S3). Copy-number profiles were confirmed by WGS (performed only for patients 3 and 4; e.g., Fig. 2C). Mutations in tumor-suppressor genes frequently co-occurred with LOH, consistent with the inactivation of the gene. Thus at the genomic level, our cohort was consistent with other cohorts of early breast cancers (1–3).

Distinct modes of lymph node evolution

The combination of point mutations, indels, and CNAs clearly identified two major evolutionary patterns.

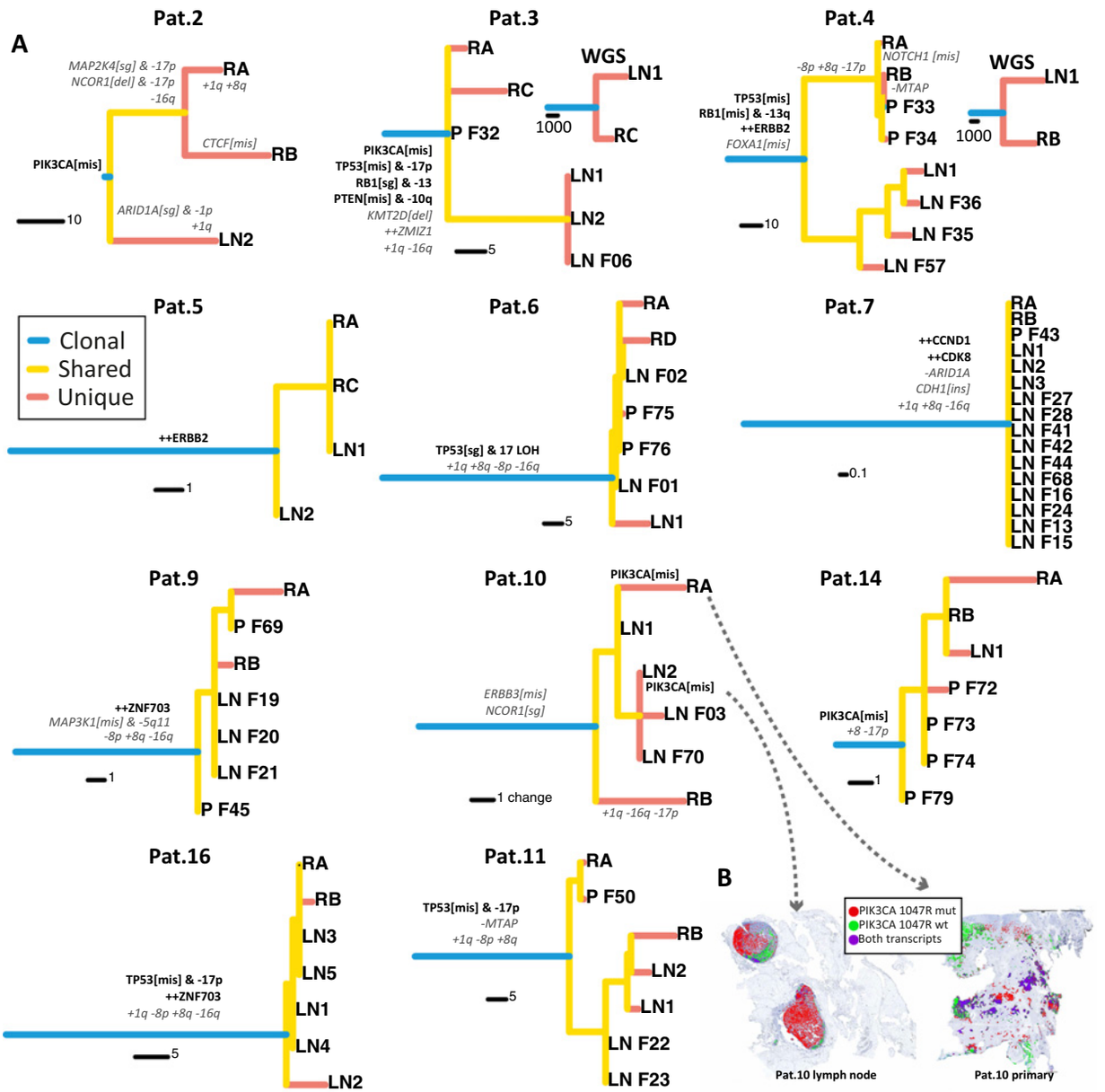

Figure 3.

Evolutionary trajectories during lymph node invasion. **A,** Phylogenetic trees reconstructed with maximum parsimony for each patient illustrate the patterns of lymph node spread. For patients 3 and 4, results were validated with WGS. When multiple lymph node samples were available for a patient, those clustered together in a single clade, indicating a recent common ancestor that led to lymph node colonization. See Supplementary Fig. S4 for tree bootstrap values. Putative driver genes and recurrent CNAs in breast cancers are annotated in the trees (tier 1, likely driver genes, in bold, black type; recurrent CNAs; and tier 2, possible drivers of uncertain pathogenicity, in italic, gray type). **B,** We assessed the spatial heterogeneity of subclonal mutation PIK3CA 1047R in patient 10 at single-cell resolution using ISH of mutant versus wild-type transcripts, revealing spatial segregation of mutant and wild-type subclones (only signals from cancer cells are represented). Pat., patient.

Published OnlineFirst June 11, 2018; DOI: 10.1158/1078-0432.CCR-17-3374

Sequential evolution where lymph node metastasis originated from a localized subclone in the primary was rare (1/11—only patient 11; Fig. 2A). This was confirmed using phylogenetic analysis. In this case, the tumor phylogenetic tree showed expansion of a metastatic cancer lineage that originated within region RB and spread to both LN1 and LN2. This is shown in the tree as RB, LN1, and LN2 having a recent common ancestor and forming a clade distinct from RA (Fig. 3A). Additional unique mutations within the clade are most likely passengers.

The two predominantly observed patterns were early divergence or a complete lack of divergence. Early divergence between primary and lymph nodes was observed in patients 2, 3, 4, and 6. In these cases, the mutational and in part the CNA landscapes were very different between the primary and the lymph nodes, with significant heterogeneity at the level of putative driver alterations. Phylogenetic analysis revealed that the lymph node deposits diverged very early during the evolutionary history of the tumor in these patients (Fig. 3A; see Supplementary Fig. S4 for bootstrap values and Supplementary Fig. S5 for WES trees). This is particularly interesting because recent breast cancer studies found similar patterns of divergence between primary and metastatic lesions (20, 21). In our cohort, multiple samples from the lymph nodes were also very similar to each other, consistent with a recent common ancestor of the lymph node lesions, which indicates a clonal bottleneck. The fact that additional samples profiled with targeted sequencing corroborated the original phylogenetic topology constructed with WES (Supplementary Fig. S5) confirms that divergence patterns were not due to sampling bias. This is important because phylogenetic divergence could appear simply due to undersampling of the lineages in the primary tumor (22). Furthermore, early divergence was confirmed by WGS in patients 3 and 4 (Figs. 2C and 3A). To test the impact of possible subclonal structure that may confound the phylogenies (23), we also reconstructed the phylogenetic trees with only clonal mutations in each sample with CCF >80% [using the most recent common ancestor (MRCA) in each sample], and the topologies were unchanged (Supplementary Fig. S6), again highlighting early divergence in a subset of cases.

The rest of the cohort (patients 5, 7, 9, 10, 14, and 16) was characterized by a palm-tree topology, with relatively short branches and no detectable divergence between primary and lymph node lesions (Fig. 3A). Putative drivers and recurrent alterations in this subgroup were almost invariably truncal (all apart from PIK3CA in patient 10, as also reported by others; ref. 6). We investigated the spatial heterogeneity of PIK3CA 1047R in primary and lymph node lesions of patient 10 further using single-cell level chromogenic ISH with BaseScope (Fig. 3B; primary: 44.17% mutant, lymph node: 80.96% mutant; signal from cancer cells only is reported), demonstrating segregation of PIK3CA mutant and wild-type subclones. Divergent patterns were quantified using Node Cophenetic Distance (24) and confirmed significant divergence measured both on targeted exome sequencing ($P = 0.0043$, Wilcoxon rank-sum test) and WES ($P = 0.0079$, Wilcoxon rank-sum test) data (Supplementary Fig. S7). Divergence was not correlated with number of samples.

Importantly, we found no evidence of genes recurrently altered in lymph nodes with respect to the primary lesions, although we cannot exclude the presence of weakly recurrent drivers that we do not have the power to detect in our cohort.

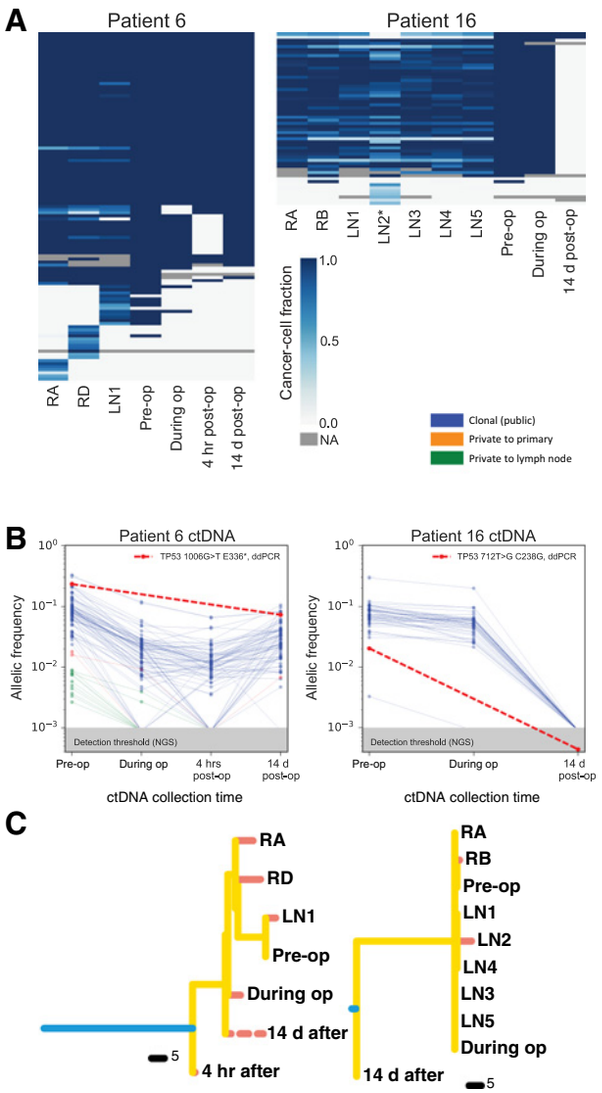

Figure 4.

Longitudinal ctDNA analysis recapitulates tissue evolution. **A,** The cohort-specific targeted panel was applied to ctDNA for 2 patients, and heatmaps show presence (blue) and absence (yellow) of variants at four time points (pre-op, immediately postresection, 4 hours post-op, and 14 days post-op) compared with the corresponding primary and lymph node samples per patient. NA, not enough coverage or variant does not overlap with a copy-number segment. **B,** Variant allele frequency (VAF) changes of all mutations at different time points. **C,** Phylogenetic trees reconstructed with both tissue and ctDNA data confirm these patterns. Post-op ctDNA appears early in the tree in patient 6, suggesting early disseminated micrometastatic disease (dashed line indicates possible additional variants not detectable with a targeted approach). d, days; hrs, hours.

Different modes of lymph node spread are recapitulated in ctDNA

In order to follow the evolutionary dynamics of node-positive early breast cancers through time, we collected cell-free DNA at multiple time points on 11 patients; however, only 7 had enough DNA (20 ng in total) to allow genomic profiling. We applied the targeted sequencing custom panel used in Fig. 2A to those 7 patients, but somatic mutations were detected in only 2. For these 2 patients, we had four time

Published OnlineFirst June 11, 2018; DOI: 10.1158/1078-0432.CCR-17-3374

points: pre-op, immediately after resection, 4 hours post-op, and 12 to 14 days post-op. Genomic profiling shows the dramatic impact of tumor resection on the ctDNA, as proof of principle indicating that the resected lesions were responsible for shedding detectable tumor DNA in the plasma (Fig. 4A). Indeed, the frequency of mutations dramatically drops after tumor (primary and nodes) resection (Fig. 4B). However, patient 6 showed mutations increasing again 14 days after the operation. Remarkably, the majority of private mutations found in the ctDNA samples before the operation were unique to the lymph nodes, corroborating the divergence patterns observed in solid samples. After lymph node resection, private mutations from the nodes disappeared from plasma, confirming the origin of the shedding. A subset of truncal mutations, however, persisted in the plasma 14 days after the operation. This was unlikely due to ctDNA remnants due to its short half-life and instead suggests the presence of residual micrometastatic disease shedding ctDNA in the blood. For patient 16, the lack of divergence reported in the tissue was observed in plasma as well. Phylogenetic reconstruction confirmed these patterns for both patients. In particular, for patient 6, the preoperative ctDNA profile clustered with the lymph node sample, whereas the postoperative ctDNA sample showed an earlier divergence event of the micrometastatic disease. Because we do not know which mutations are private to the micrometastatic deposits, the postoperative ctDNA branch appears shorter in our data than it actually is. These results indicate that the patterns of lymph node spread observed in the tissue, even in this early node-positive cohort, are recapitulated in the plasma.

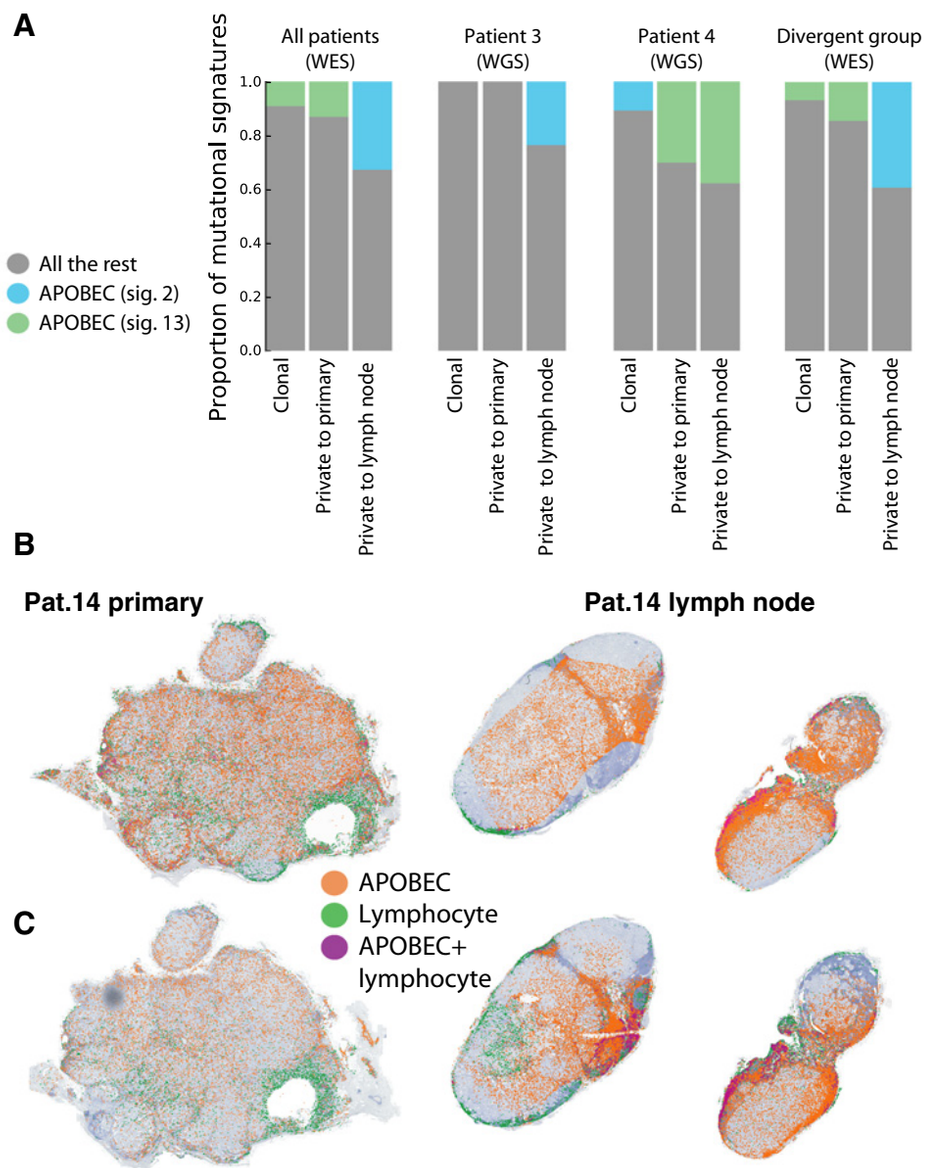

Figure 5.
APOBEC signature is increased in lymph nodes. **A,** Mutational signature analysis applied to WES mutations from the whole cohort, and WGS for patients 3 and 4 showed an increase in APOBEC signatures in mutations private to the lymph nodes. sig., signature. Spatial heterogeneity in expression of APOBEC3A (**B**) and APOBEC3B (**C**) was assessed at the single-cell level using ISH for patient (Pat.) 14, revealing higher expression of both APOBEC3A and APOBEC3B in the lymph node lesion (only signals from cancer cells are represented).

Published OnlineFirst June 11, 2018; DOI: 10.1158/1078-0432.CCR-17-3374

APOBEC activity is increased in lymph nodes

Mutational signature analysis revealed the presence of a common age-related cancer signature 1, as well as signatures specific to breast cancer (ref. 25; signatures 2, 3, and 13). Interestingly, APOBEC signatures, which were detected in 5 of 11 patients, were found to be increased in lymph nodes with respect to the primary tumor, especially in the divergent subgroup. These results were confirmed using WGS on patients 3 and 4 (Fig. 5A). RNA ISH of APOBEC3A and APOBEC3B transcripts for patient 14 using RNAScope (Fig. 5B and C) revealed spatial heterogeneity, with 3.17-fold higher expression of APOBEC3A and 3.72-fold higher expression of APOBEC3B in the lymph node with respect to the primary tumor (measured in dots/mm^2, only signal from cancer area is reported, see "Materials and Methods"). Automatic identification of lymphocytes indicated that APOBEC signal came predominantly from cancer cells, although in some areas, cell density was so high that the two signals overlapped (purple). This was consistent with the mutational signatures in Fig. 5A. This suggests that APOBEC may be involved in driving ITH during metastatic spread to lymph nodes.

Discussion

Breast cancer can spread from one organ system to another via hematogenous and lymphatic routes. Understanding lymph node spread from an evolutionary perspective is crucial to improve the understanding of progression to metastatic disease. In this study, we focused on untreated lymph node–positive patients without evidence of distant deposits and performed a spatiotemporal analysis of the evolution of lymph node invasion. We found striking patterns of early divergence in a significant proportion of patients. Remarkably, ctDNA analysis identified the divergent lymph nodes as the main contributor to ctDNA at resection, thus reflecting the evolutionary patterns identified in the tissue. This implies that ctDNA may partially inform on the biology of axillary lymph node spread. The divergent lesions were highly distinct in terms of mutations and partly in terms of copy-number changes, suggesting a clonal bottleneck during lymph node spread in these patients. Importantly, our data are consistent with a model of punctuated evolution in breast cancer, where tumorigenesis is driven by relatively rare but dramatic selection events (7). Moreover, from a therapeutic perspective, inhibiting APOBEC may prevent or slow down metastatic evolution. The question on whether evolutionary patterns such as lymph node divergence have prognostic and/or predictive value remains open and will require testing in larger cohorts.

This study has several limitations, for example, the limited number of patients and the lack of ctDNA longitudinal tracking beyond day 14 after surgery. Moreover, the follow-up is relatively short, and, to date, we have not had the opportunity to profile distant metastatic deposits to understand the representation of the clones from primary and nodes in the three subgroups. Further efforts on larger cohorts of patients are needed to validate the three subgroups, and to determine their utility in determining patients' prognosis in addition to already established prognostic factors and potentially direct more effective treatment strategies.

Data and Materials Availability

Sequence data have been deposited at the European Genome-phenome Archive (EGA), which is hosted by the EBI and the CRG, under accession number EGAS00001002947. Further information about EGA can be found on https://ega-archive.org. High-resolution images have been deposited in BioStudies with accession number S-BSST110.

Disclaimer

The findings, opinions, and recommendations expressed here are those of the authors and not necessarily those of the universities where the research was performed, the National Institutes of Health or the Arizona Department of Health Services, Arizona Biomedical Research Centre.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: P. Barry, M. Dowsett, C.C. Maley, G. Schiavon, A. Sottoriva

Development of methodology: P. Barry, I. Spiteri, D. Nichol, G.D. Cresswell, Y. Yuan, G. Schiavon

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): P. Barry, I. Spiteri, A. Acar, S. Hrebien, I. Garcia-Murillas, L. Zabaglo, K. Koelble, J.E. Rusby, F. Muscara, R. Natrajan, A. Sottoriva

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): P. Barry, A. Vatsiou, I. Spiteri, D. Nichol, G.D. Cresswell, N. Trahearn, K. Chkhaidze, L. Ermini, I.S. Huntingford, H. Cottom, K. Koelble, S. Khalique, Y. Yuan, G. Schiavon, N. Turner, A. Sottoriva

Writing, review, and/or revision of the manuscript: P. Barry, A. Vatsiou, I. Spiteri, D. Nichol, G.D. Cresswell, A. Acar, N. Trahearn, S. Hrebien, I. Garcia-Murillas, K. Koelble, J.E. Rusby, M. Dowsett, G. Schiavon, N. Turner, A. Sottoriva

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): P. Barry, I. Spiteri, A. Acar, S. Hrebien, H. Cottom, S. Khalique, R. Natrajan

Study supervision: A. Sottoriva

Other (performed histologic assessment and review of the primary tumors and lymph node tissue samples from the patient cohort): H. Cottom

Acknowledgments

A. Sottoriva is supported by the Wellcome Trust (202778/B/16/Z), Cancer Research UK (A22909), and the Chris Rokos Fellowship in Evolution and Cancer. Y. Yuan acknowledges support by Cancer Research UK (A21808) and Breast Cancer Now (2015NovPR638). This work was also supported by Wellcome Trust funding to the Centre for Evolution and Cancer (105104/Z/14/Z). The authors also acknowledge funding from Breast Cancer Now and the Royal Marsden and ICR NIHR Biomedical Research Centre. C.C. Maley was supported in part by NIH grants U54 CA217376, P01 CA91955, R01 CA170595, R01 CA185138, and R01 CA140657 as well as CDMRP Breast Cancer Research Program Award BC132057 and by the Arizona Biomedical Research Centre as made available through the Arizona Department of Health Services. The authors thank Advanced Cell Diagnostics for providing the BaseScope PIK3CA probes. Image from Fig. 1A was taken from Servier Medical Art (licensed under a Creative Commons Attribution 3.0 Unported License).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received February 6, 2018; revised April 11, 2018; accepted June 5, 2018; published first June 11, 2018.

Published OnlineFirst June 11, 2018; DOI: 10.1158/1078-0432.CCR-17-3374

References

1. Curtis C, Shah SP, Chin S-F, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 2012;486:346–52.
2. Shah SP, Roth A, Goya R, Oloumi A, Ha G, Zhao Y, et al. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. Nature 2012;486:395–9.
3. Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature 2016;534:47–54.
4. Nik-Zainal S, Van Loo P, Wedge DC, Alexandrov LB, Greenman CD, Lau KW, et al. The life history of 21 breast cancers. Cell 2012;149:994–1007.
5. Wang Y, Waters J, Leung ML, Unruh A, Roh W, Shi X, et al. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. Nature 2014;512:155–60.
6. Yates LR, Gerstung M, Knappskog S, Desmedt C, Gundem G, Van Loo P, et al. Subclonal diversification of primary breast cancer revealed by multiregion sequencing. Nat Med 2015;21:751–9.
7. Gao R, Davis A, McDonald TO, Sei E, Shi X, Wang Y, et al. Punctuated copy number evolution and clonal stasis in triple-negative breast cancer. Nature Genetics 2016;48:1119–30.
8. Eirew P, Steif A, Khattra J, Ha G, Yap D, Farahani H, et al. Dynamics of genomic clones in breast cancer patient xenografts at single-cell resolution. Nature 2015;518:422–6.
9. Greaves M, Maley CC. Clonal evolution in cancer. Nature 2012;481:306–13.
10. McGranahan N, Swanton C. Clonal heterogeneity and tumor evolution: past, present, and the future. Cell 2017;168:613–28.
11. Blamey RW, Pinder SE, Ball GR, Ellis IO, Elston CW, Mitchell MJ, et al. Reading the prognosis of the individual with breast cancer. Eur J Cancer 2007;43:1545–7.
12. Wishart GC, Bajdik CD, Azzato EM, Dicks E, Greenberg DC, Rashbass J, et al. A population-based validation of the prognostic model PREDICT for early breast cancer. Eur J Surg Oncol 2011;37:411–7.
13. Senkus E, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rutgers E, et al. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2015:v8–30.
14. Early Breast Cancer Trialists' Collaborative Group (EBCTCG), Davies C, Godwin J, Gray R, Clarke M, Cutter D, et al. Relevance of breast cancer hormone receptors and other factors to the efficacy of adjuvant tamoxifen: patient-level meta-analysis of randomised trials. Lancet 2011;378:771–84.
15. Early Breast Cancer Trialists' Collaborative Group (EBCTCG), Peto R, Davies C, Godwin J, Gray R, Pan HC, et al. Comparisons between different polychemotherapy regimens for early breast cancer: meta-analyses of long-term outcome among 100,000 women in 123 randomised trials. Lancet 2012;379:432–44.
16. EBCTCG Early Breast Cancer Trialists' Collaborative Group. Effect of radiotherapy after mastectomy and axillary surgery on 10-year recurrence and 20-year breast cancer mortality: meta-analysis of individual patient data for 8135 women in 22 randomised trials. Lancet 2014;383:2127–35.
17. Garcia-Murillas I, Schiavon G, Weigelt B, Ng C, Hrebien S, Cutts RJ, et al. Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer. Sci Transl Med 2015;7:302ra133–3.
18. Naxerova K, Reiter JG, Brachtel E, Lennerz JK, van de Wetering M, Rowan A, et al. Origins of lymphatic and distant metastases in human colorectal cancer. Science 2017;357:55–60.
19. Olmeda D, Cerezo-Wallis D, Riveiro-Falkenbach E, Pennacchi PC, Contreras-Alcalde M, Ibarz N, et al. Whole-body imaging of lymphovascular niches identifies pre-metastatic roles of midkine. Nature 2017;546:676–80.
20. Dawson SJ, Tsui DWY, Murtaza M, Biggs H, Rueda OM, Chin SF, et al. Analysis of circulating tumor DNA to monitor metastatic breast cancer. N Engl J Med 2013;368:1199–209.
21. Ng CKY, Bidard FC, Piscuoglio S, Geyer FC, Lim RS, de Bruijn I, et al. Genetic heterogeneity in therapy-naïve synchronous primary breast cancers and their metastases. Clin Cancer Res 2017;23:4402–15.
22. Hong WS, Shpak M, Townsend JP. Inferring the origin of metastases from cancer phylogenies. Cancer Res 2015;75:4021–5.
23. Alves JM, Prieto T, Posada D. Multiregional tumor trees are not phylogenies. Trends Cancer 2017;3:546–50.
24. Paradis E, Claude J, Strimmer K. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 2004;20:289–90.
25. Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, et al. Mutational processes molding the genomes of 21 breast cancers. Cell 2012;149:979–93.

Published OnlineFirst June 11, 2018; DOI: 10.1158/1078-0432.CCR-17-3374

Clinical Cancer Research

The Spatiotemporal Evolution of Lymph Node Spread in Early Breast Cancer

Peter Barry, Alexandra Vatsiou, Inmaculada Spiteri, et al.

Clin Cancer Res 2018;24:4763-4770. Published OnlineFirst June 11, 2018.