Molecular Heterogeneity in Primary Breast Carcinomas and Axillary Lymph Node Metastases Assessed by Genomic Fingerprinting Analysis

Rachel E. Ellsworth[1], Allyson L. Toro[2], Heather L. Blackburn[2], Alisha Decewicz[2], Brenda Deyarmin[2], Kimberly A. Mamula[2], Nicholas S. Costantino[2], Jeffrey A. Hooke[3], Craig D. Shriver[3] and Darrell L. Ellsworth[2]

[1]Clinical Breast Care Project, Murtha Cancer Center, Windber, PA, USA. [2]Clinical Breast Care Project, Windber Research Institute, Windber, PA, USA. [3]Clinical Breast Care Project, Murtha Cancer Center, Walter Reed National Military Medical Center, Bethesda, MD, USA.

ABSTRACT: Molecular heterogeneity within primary breast carcinomas and among axillary lymph node (LN) metastases may impact diagnosis and confound treatment. In this study, we used short tandem repeated sequences to assess genomic heterogeneity and to determine hereditary relationships among primary tumor areas and regional metastases from 30 breast cancer patients. We found that primary carcinomas were genetically heterogeneous and sampling multiple areas was necessary to adequately assess genomic variability. LN metastases appeared to originate at different time periods during disease progression from different sites of the primary tumor and the extent of genomic divergence among regional metastases was associated with a less favorable patient outcome ($P = 0.009$). In conclusion, metastasis is a complex process influenced by primary tumor heterogeneity and variability in the timing of dissemination. Genomic variation in primary breast tumors and regional metastases may negatively impact clinical diagnostics and contribute to therapeutic resistance.

KEYWORDS: intratumor heterogeneity, metastasis, breast cancer, allelic imbalance, axillary lymph nodes

Introduction

Primary breast carcinomas are known to exhibit intratumor differences in morphology[1] and expression of clinical biomarkers, such as the estrogen receptor (ER) and progesterone receptor (PR), and the extent of human epidermal growth factor receptor 2 (HER2) gene amplification.[2–4] Recent studies using microarray analysis and next-generation sequencing of single cells have detected additional heterogeneity in chromosomal alterations and DNA mutations.[5,6] Molecular heterogeneity within the primary carcinoma may affect diagnosis and treatment in patients where localized biopsy specimens do not accurately capture the complete genomic landscape of the primary tumor, and thus may not accurately reflect response to treatment. In addition, tumors with extensive molecular heterogeneity may be more likely to adapt to commonly used cytotoxic agents and targeted therapeutics.[7]

Previous studies have detected genomic heterogeneity and discordant patterns of chromosomal alterations in axillary lymph node (LN) metastases compared to the primary breast carcinoma.[8–11] Genomic sequencing revealed discordant mutations in invasive lobular carcinomas versus matched metastases[12] and enrichment of mutations in a breast metastasis from a patient with triple-negative breast cancer,[13] suggesting that cells with metastatic potential may arise from subpopulations of cells in the primary tumor. Expression of biomarkers used for defining treatment options also differs between primary carcinomas and metastases,[14–16] which can pose a significant challenge to cancer medicine because treatments based on the primary tumor may be ineffective on the metastases.[17]

Understanding tumor heterogeneity is critical for optimizing treatment selection and improving patient outcomes, as well as understanding mechanisms of metastatic dissemination.[18] Genetic differences between metastases and the primary carcinoma may result from early dissemination of cells with metastatic potential and ongoing molecular evolution during disease progression, or reflect the place of origin in a genetically heterogeneous primary tumor. Furthermore, evolutionary relationships between multiple axillary LN metastases remain unclear. To address these

questions, we used a genomic fingerprinting approach to examine the levels and patterns of genomic changes in primary carcinomas and axillary LN metastases from 30 node-positive female patients.

Materials and Methods

Eligibility and enrollment. Women enrolled in the Clinical Breast Care Project met the following eligibility criteria: (1) adults aged >18 years, (2) mentally competent and willing to provide informed consent, and (3) presenting with evidence of possible breast disease. Tissue and blood samples were collected with approval from the Walter Reed National Military Medical Center Human Use Committee and Institutional Review Board. All the subjects voluntarily agreed to participate and gave written informed consent. This research complied with the principles of the Declaration of Helsinki.

Tissue collection and characterization. All female patients meeting the following criteria were included in this study: invasive breast cancer ≥1.0 cm in diameter, multiple axillary LN metastases available for research, and no neoadjuvant treatment. Patients diagnosed prior to August 2004 ($n = 11$) underwent surgical excision of the primary carcinoma followed by complete axillary LN dissection. After August 2004, sentinel LN biopsy was performed, followed by axillary LN dissection ($n = 19$). Diagnosis of every specimen was conducted by a breast pathologist from hematoxylin and eosin stained slides; pathological characterization was performed as previously described.[19]

Microdissection. Formalin-fixed, paraffin-embedded (FFPE) specimens were available from 30 patients. An average of 3.8 ± 0.8 separate pieces of primary tumor embedded in paraffin blocks and 4.9 ± 1.3 paraffin blocks containing LN metastases were used per patient. Although seven cases were included in earlier studies,[8,9] all primary tumor areas and metastases were isolated and analyzed from slides created specifically for this project. The dedicated breast pathologist evaluated all the slides used for laser microdissection to identify and delineate areas within the primary tumor and metastases for dissection. Depending on cellular density, four to nine tissue sections (5 microns in thickness) were microdissected per tumor area. For each primary carcinoma, 5 to 19 areas ~2.8 mm^2 in size, which were representative of the internal areas and edges of the tumor, were isolated from multiple FFPE blocks (Fig. 1A) using an AS*LMD* laser microdissection system (Leica Microsystems).[20] Axillary LNs were assessed for metastatic deposits by examination of an additional slide and only deposits >2.0 mm in diameter were microdissected. A single area 2.8–5.3 mm^2 was obtained from each metastasis. In total, 269 primary tumor areas and 196 regional metastases were evaluated for genomic alterations.

Genomic fingerprinting. DNA was extracted from microdissected tumor cells, and referent DNA samples were obtained from blood clots from each patient using Clotspin and Puregene DNA purification kits (Qiagen Inc.). Short tandem repeats

(STRs; $n = 52$) were chosen from 26 chromosomal regions frequently altered in breast cancer (Table 1).[21] Each STR could be amplified by PCR in DNA from archival specimens after laser microdissection. STRs were amplified using either a modified PCR stepdown protocol or a hot-start protocol. Genomic alterations at each STR were detected as allelic imbalance (AI) as previously described using a threshold of ≤0.35.[22] A number of quality control/validation procedures were implemented, including (1) using a minimum acceptable signal intensity in the electropherograms of 500 relative fluorescence units to increase accuracy of the allele calls, (2) regenotyping all STRs showing evidence of AI in an independently microdissected sample, and (3) verifying the low frequency of AI events in histologically normal reductive mammoplasty specimens.[23]

Statistical analysis. Two measures of genomic heterogeneity were calculated for all primary tumors and/or all metastases of each patient. The extent of genetic difference between tumor areas or between metastases was calculated using the mean pairwise genetic divergence score (D) as

$$D = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{X_{i,j}}{Y_{i,j}}$$

where n is the number of tumor areas from a patient, $X_{i,j}$ is the number of markers that show AI in either tumor area i or j but not in both, and $Y_{i,j}$ is the number of markers that are heterozygous in normal referent DNA and for which genotyping was informative for both areas i and j. Clonal diversity (H), representing the number and abundance of genetically distinct clones within primary tumors or among LN metastases, was measured by the Shannon diversity index (H) using the formula

$$H = -\sum_{i} p_i \ln(p_i)$$

where p_i is the relative frequency of clone i in the tumor.[24]

One-way ANOVA was used to evaluate relationships between genomic heterogeneity in primary tumors and axillary LN metastases and clinical/pathological characteristics with a threshold of $P < 0.05$ defining significance. R statistical software (version 3.1.1) was used to examine the effects of sampling on our ability to detect genomic heterogeneity in the primary carcinoma. All tumor areas for each patient were selected at random, and a cumulative count of AI events was recorded as each additional tumor area was sampled. The entire sequential sampling process was repeated 10 times for each patient to determine the mean number of detectable AI events per number of tumor areas sampled, and a regression curve was derived using the best adjusted R-squared value.

Phylogenetic analysis was conducted using the Phylogeny Inference Package (PHYLIP version 3.66) to determine hereditary relationships among primary tumor areas and
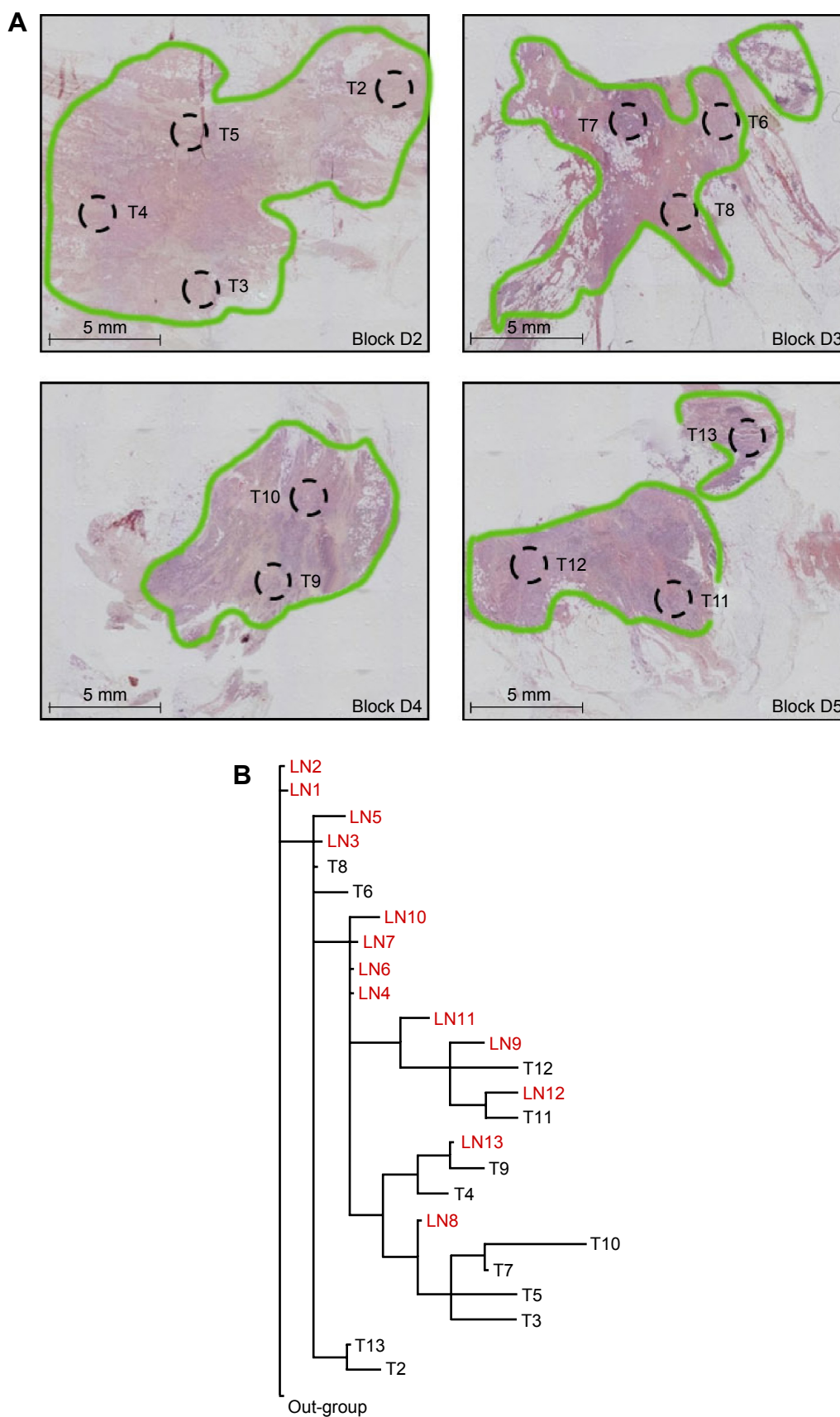
Figure 1. (**A**) Hematoxylin and eosin stained tissue sections from a representative patient showing areas of the primary breast tumor (T) microdissected to assess genomic changes. Areas of invasive carcinoma are outlined in green. (**B**) Most parsimonious phylogenetic tree depicting hereditary relationships among areas of the primary breast carcinoma (T) and axillary LN metastases from the same patient shown above in panel "**A**" based on patterns of genomic changes assessed by AI.

Table 1. Short tandem repeated sequences used to assess genomic heterogeneity in primary breast carcinomas and axillary lymph node metastases.

MARKER	MAP LOCATION	FORWARD PRIMER	REVERSE PRIMER	AI FREQUENCY IN PRIMARY TUMOR	AI FREQUENCY IN LN METASTASES
D1S468	1p36.32	AATTAACCGTTTTGGTCCT	GCGACACACACTTCCC	0.17	0.08
D1S434	1p36.22	AGCTAATTTACATTACCCAAAAAGA	GCAGGTGGCACAGTGA	0.05	0.03
D2S112	2q21.2	GAGTGGCGGTGAGAAGGTAT	AGCCATTGCTATCTTTGAGG	0.04	0.06
D2S2188	2q31.1	GGAGTCAACCTAGTGCCTATT	CATCCAGCCAGGATTC	0.03	0.02
D3S1581	3p21.31	CAGAACTGCCAAACCA	GGGTAACAGGAGCGAG	0.13	0.08
D3S1566	3p13	GCTGCTGAAGCACCAAATC	CTCGTTGGAAATCATTCTGAGG	0.16	0.14
D5S433	5q21.2	TGTAAGACATACTCTCTATCACCC	TCAGACATCCATCTGTGTG	0.19	0.10
D5S2084	5q21.3	AGCTGCTTACCACGAATGTC	TCAGACTTTGAGCCCTGCTA	0.19	0.12
D6S462	6q15	GTGACATCATAGGGAGGGCTG	ACTGGAGTTGGGGGACAT	0.06	0.02
D6S1671	6q16.2	TTTGGTCAATTTCAATCTGTAG	ATCCTCCAGGGGTGCT	0.09	0.08
D6S1698	6q21	TGCAGGTAATTTGACTACCC	ACACCCCTCATATATACTTGAGTGT	0.20	0.07
D6S287	6q22.31	ATATTAGTGCCTTATGCTTCTG	AAATTGGATATTCATGCTTG	0.08	0.07
D6S1577	6q25.2	TGACATTAGGAGGCACTGG	TTAACTTGTCTGGCTGTTTGGAT	0.12	0.16
D6S264	6q27	AGCTGACTTTATGCTGTTCCT	TTTTCCATGCCCTTCTATCA	0.09	0.08
D7S2459	7q22.3	AAGAAGTGCATTGAGACTCC	CCGCCTTAGTAAAACC	0.05	0.02
D7S530	7q32.2	TGCATTTTAGTGGAGCACAG	CAGGCATTGGGAACTTTG	0.07	0.03
D8S1827	8p22	GACAGAATCATGTGGCCTTT	TTTTGTAAAATGTAAAATTGGCTTT	0.14	0.10
D8S1734	8p21.3	GCTATCCACTTGTCCCAGA	AGCCCAGAAATAAACCCTC	0.11	0.11
D8S1720	8q24.21	GTGCCACCTGCCTGAA	CCACTACCTATTTAGAGAGGCCA	0.14	0.06
D8S256	8q24.22	GTTCAAGGGCTCAGGGTTCT	CTTCCACCTTTAGCCAAGGA	0.19	0.09
D9S171	9p21.3	AGCTAAGTGAACCTCATCTCTGTCT	ACCCTAGCACTGATGGTATAGTCT	0.09	0.04
D9S161	9p21.2	TGCTGCATAACAAATTACCAC	CATGCCTAGACTCCTGATCC	0.10	0.06
D10S1765	10q23.31	ACACTTACATAGTGCTTTCTGCG	CAGCCTCCCAAAGTTGC	0.12	0.09
D10S1709	10q24.2	GTGAGTCCAGAATCACCCC	CAGTGGAATGGCTCATTTG	0.17	0.10
D11S1760	11p15.4	GATCTCAAGTGTTTCCCCAC	AAACGATGTCTGTCCACTCA	0.09	0.04
D11S902	11p15.1	CCCGGCTGTGAATATACTTAATGC	CCCAACAGCAATGGGAAGTT	0.10	0.04
D11S987	11q13.2	ACTCCAGTCTGGGCAATAAAAGC	GGTGGCAGCATGACCTCTAAAG	0.09	0.04
D11S937	11q14.1	CTAATAAACAAATCCCTCTACCTCC	TAGTCAGTCAGGGACCCAAGT	0.15	0.08
D11S4127	11q23.3	ATGAGAAGTGCCATCCAGC	ACTATGCCCAGTGTGTGTGC	0.14	0.14
D11S4094	11q24.1	CTCAAAGAACAGCCAGTCA	GGAGTCGGGGAATTTCTAA	0.13	0.15
D13S1304	13q12.13	ACCAGCCTTTGCTTAGGA	ACATTCTAGTGCTACAGGGTACTC	0.06	0.03
D13S171	13q13.1	CCTACCATTGACACTCTCAG	TAGGGCCATCCATTCT	0.24	0.19
D13S263	13q14.11	CCTGGCCTGTTAGTTTTTATTGTTA	CCCAGTCTTGGGTATGTTTTTA	0.14	0.09
D13S153	13q14.2	AGCATTGTTTCATGTTGGTG	CAGCAGTGAAGGTCTAAGCC	0.18	0.13
D14S1037	14q31.3	TCAATCATGGAAGTTGGCTACC	TGAAGGTGGACCAAAGCATC	0.18	0.07
D14S1054	14q32.13	ACAGTCACGTGGGCCA	GGACCTGGGCATTTATTCTG	0.17	0.10

Marker	Cytoband	Primer 1	Primer 2		
D16S3136	16q12.1	ATTGCCCTCAAGAACAGC	GTGCTATGCCATCCCAG	0.09	0.17
D16S3066	16q22.3	AGTCAGGACACGATGGTTTG	CCCAGTTCCAATGGCAC	0.24	0.19
D16S3040	16q23.2	TACTCCGGCAAGGACG	GCTGCCTAGCACATGG	0.13	0.05
D16S3091	16q23.3	GGGAGATAGCCTTAAACTTTCTTAC	TGTTGCTAATAACACTAGGCCA	0.07	0.09
D17S831	17p13.3	CGCCTTTCCTCATACTCCAG	GCCAGACGGGACTTGAATTA	0.16	0.05
D17S1828	17p13.2	TGCACTCACAGATTTGCC	TTAAGCCAGTTCGGATTTG	0.19	0.08
D17S1876	17p13.2	AGCTGCTTCTGCAAAGATG	TACAAGTCCTGGGCCAC	0.22	0.15
D17S799	17p12	ATTGCCAGCCGTCAGTT	GACCAGCATATCATTATAGACAAGC	0.29	0.21
D17S927	17q12	TGTCTGACATGACACCCT	TCCAACCTGAAGGCCAGT	0.18	0.27
D17S1795	17q21.33	AGTGCCAGAGATATACCGTG	GTCTGCAAGGCAAGTTGTC	0.13	0.08
D18S1102	18q12.2	TTTCAGGATTTTGGAGCC	GGAATGACTGCGTCTGTG	0.02	0.01
D18S474	18q21.2	TGGGGTGTTTACCAGCATC	TGGCTTTCAATGTCAGAAGG	0.07	0.02
D22S1163	22q12.1	AAAAATCAAAGGTCAGCCTC	ACAATGTGTGCGTGTGC	0.16	0.12
D22S283	22q12.3	ACCAACCAGCATCATCAT	AGCTCGGGACTTTCTGAG	0.16	0.10
D22S277	22q12.3	TTCTTGTGTGGTAGTCTGGG	TACCNACTCCCCAAACTATG	0.15	0.06
D22S1170	22q13.21	ACCGTTGCCTATATCCA	AGCCCACTCCACAATTT	0.20	0.12

Abbreviations: AI, allelic imbalance; LN, lymph node.

axillary LN metastases in each patient. AI was treated as a discrete character and coded as present or absent. Analysis in PHYLIP used Wagner parsimony in the PARS program. Input options included (1) equal weighting of all chromosomal regions where AI was assessed, (2) random input order, and (3) outgroup rooting using a hypothetical (normal) tissue with no AI.

Results

Patient characteristics and outcomes. Age at diagnosis ranged from 34 to 86 years, with a mean age of 57. All patients had multiple involved LNs (range 3–31, median 10); the majority of women were diagnosed with stage III breast cancer (Table 2). Average time between diagnosis and death was 28 months (range 10–72 months). Patients with no evidence of disease were disease-free for an average of 76 months (range 28–106 months).

Genomic heterogeneity. Genomic data were available for both the primary carcinoma and the corresponding LN metastases for 25 of 30 patients; five patients had data only for the primary tumor or the axillary LN metastases. Examining a single area of the primary carcinoma captured 27% of the genomic variability (Fig. 2). Sampling at least five tumor areas was required to capture ~80% of the variability, while eight areas were needed to detect at least 90% of the genomic diversity.

The frequency of AI events was significantly higher ($P < 0.001$) in primary tumors (13.3%) than in axillary LN metastases (9.2%) using a matched pairs analysis. The most common regions showing genomic alterations—chromosomes 13q12.3 (24.1% and 19.1%), 16q11.2–22.1 (24.1% and 19.2%), and 17p13.1 (29.3% and 21.3%) in primary tumors and LN metastases—are frequently altered in primary breast carcinomas and multiple other human cancers.[25]

Within primary tumors and LN metastases, levels of genetic divergence were significantly correlated with clonal diversity (H) (Fig. 3). Clonal diversity (H) was significantly higher ($P < 0.001$) in primary tumors ($H = 2.2$, range 1.6–2.9) compared to matched LN metastases ($H = 1.4$, range 0–2.4). Mean genetic divergence within primary tumors ($D = 14.4\%$, range 4.2%–26.1%) was significantly correlated with, but not significantly different from, divergence among the matched axillary LN metastases ($D = 10.5\%$; range 0%–28.2%) (Fig. 3 and Table 3). Therefore, the process of metastatic dissemination appears similar to founder effect in population genetics where a new population (metastases) is established by a small number of migrants (cells) from a large diverse population (the primary tumor). Within a given patient, metastases tended to be less diverse because only some of the genetically distinct clones observed in the primary carcinoma were represented in the regional metastases. However, the extent of genetic divergence between clones in the metastases was representative of divergence between clones in the primary tumor.

Table 2. Clinical and pathological characteristics of patients.

PATIENT NO.	AGE AT DIAGNOSIS	ER/PR/HER2 STATUS	STAGE	GRADE	TUMOR SIZE (cm)	NO. TUMOR AREAS STUDIED	NO. POSITIVE LYMPH NODES	NO. LYMPH NODE METASTASES STUDIED	CLINICAL OUTCOME
6	45	+/+/−	IIA	2	1.8	10	3	3	NED
4	41	+/+/−	IIB	1	4.1	9	4	3	NED
1	34	+/+/+	IIIA	2	3.5	9	6	6	NED
10	49	+/+/−	IIIA	1	4.2	14	4	4	NED
14	55	+/+/−	IIIA	3	2.5	12	14	13	NED
15	55	+/−/−	IIIA	2	9.0	6	4	3	NED
18	62	−/−/−	IIIA	3	1.2	7	9	5	DOD
19	62	+/+/−	IIIA	2	2.1	6	5	5	DOD
23	66	−/−/+	IIIA	3	1.7	11	5	3	NED
25	69	+/+/−	IIIA	2	1.8	9	5	5	NED
29	84	+/+/−	IIIA	2	4.3	5	8	0	DOC
3	37	+/−/−	IIIC	3	5.6	10	9	7	NED
5	44	−/−/+	IIIC	3	1.1	12	10	6	NED
8	47	+/+/−	IIIC	2	8.0	13	10	6	NED
9	48	−/−/−	IIIC	2	2.8	0	19	11	DOD
11	50	+/+/−	IIIC	1	1.6	8	10	9	NED
12	50	+/+/−	IIIC	2	5.5	0	27	16	NED
16	55	−/−/+	IIIC	3	5.8	12	15	0	NED
22	64	+/+/−	IIIC	1	4.0	9	26	11	DOD
26	73	+/+/−	IIIC	1	7.0	16	12	10	DOC
27	75	−/−/−	IIIC	3	5.5	7	31	9	DOC
28	82	+/+/−	IIIC	1	2.0	7	15	5	NED
30	86	+/+/−	IIIC	1	na	9	10	6	NED
2	36	−/−/−	IV	3	4.0	12	18	7	DOD
7	46	−/−/−	IV	3	5.1	19	5	5	DOD
13	51	−/−/+	IV	3	5.0	10	18	14	DOD
17	61	+/−/+	IV	3	9.5	0	13	7	DOD
20	63	+/+/−	IV	3	4.5	11	8	4	DOD
21	64	+/+/−	IV	3	3.2	11	12	8	DOD
24	68	+/−/+	IV	2	4.0	5	5	5	DOD

Abbreviations: ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; DOD, deceased of disease; DOC, deceased of other causes; NED, no evidence of disease.
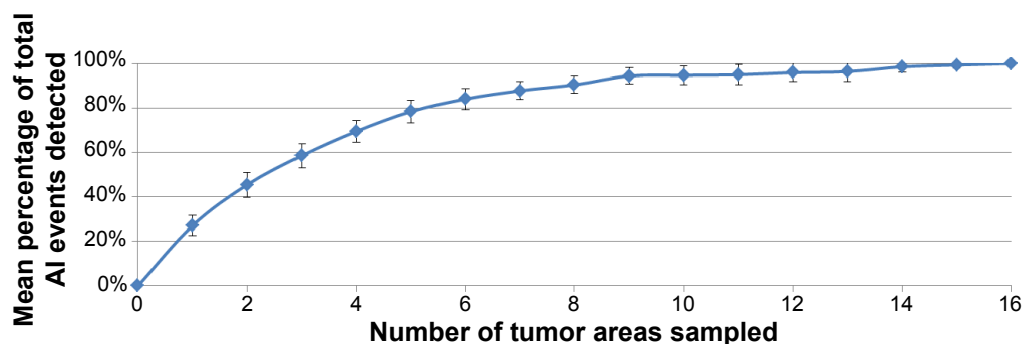
Figure 2. Mean percentage of total AI events detectable by randomly sampling areas of primary breast carcinomas. Error bars represent 95% confidence intervals for the mean percentage of AI events detected when sampling the corresponding number of tumor areas.

Clonal diversity (*H*) within primary tumors and among LN metastases did not vary significantly by menopausal status, tumor grade, stage and size, ER/HER2 status, or clinical outcome. For LN metastases only, the amount of genetic divergence between metastatic deposits (*D*) was significantly higher (*P* = 0.003) in patients diagnosed with distant metastases (stage IV) (17.5 ± 6.4%) compared to those without (stage II and III) (7.7 ± 6.1%) and was significantly higher (*P* = 0.033) in patients who died of breast disease (14.6 ± 8.1%) compared to those who remained disease free >5 years after diagnosis (7.7 ± 5.8%). Cox hazard models supported the association of higher genomic divergence among axillary LN metastases with less favorable outcomes (*P* = 0.009).

Phylogenetic analysis. Each primary tumor had a distinct pattern of genomic alterations, and thus a unique phylogenetic tree depicting genetic relationships among the tumor areas (Fig. 1B). Within each tumor, branch lengths were not equivalent for all areas, suggesting variability in rates of genomic change. Within each patient, LN metastases often shared a close common ancestry with specific areas of the primary tumor, indicating that the metastatic deposits in various LNs may have originated from different sites within the primary carcinoma (representative patients are shown in Fig. 4). Because these deposits carried genomic alterations already present in specific areas of the primary tumor, they likely originated later in disease progression. In contrast, LN metastases with low levels of
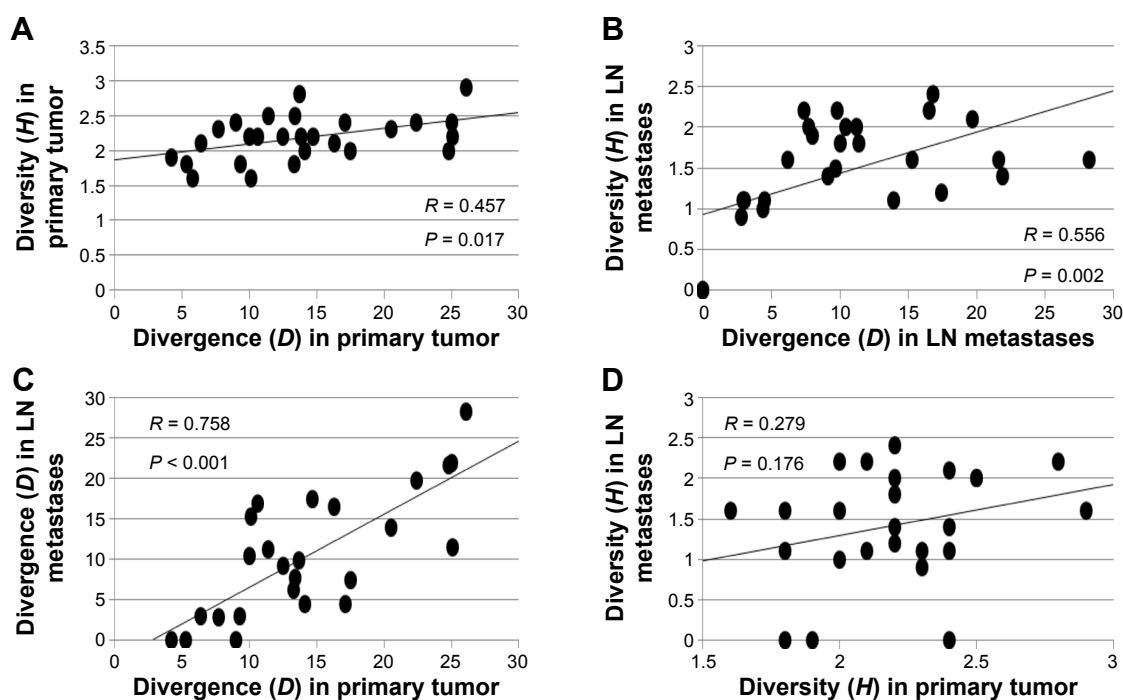

Figure 3. (Top panel) Correlation plots of mean pairwise genetic divergence (*D*) versus clonal diversity (*H*) (**A**) within primary breast tumors and (**B**) within axillary LN metastases. (Bottom panel) Relationships between (**C**) genetic divergence (*D*) and (**D**) clonal diversity (*H*) between primary breast carcinomas and axillary LN metastases.

Table 3. Number of genetically distinct clones, mean pairwise genomic divergence scores (*D*), and clonal diversity (*H*) for each patient.

PATIENT NO.	PRIMARY TUMOR			AXILLARY LN METASTASES		
	NO. CLONES	*D*	*H*	NO. CLONES	*D*	*H*
1	9	14.7	2.2	4	17.4	1.2
2	12	11.4	2.5	7	11.2	2.0
3	9	10.0	2.2	7	10.4	2.0
4	7	4.2	1.9	1	0.0	0.0
5	10	6.4	2.1	4	3.0	1.1
6	10	20.5	2.3	3	13.9	1.1
7	19	26.1	2.9	5	28.2	1.6
8	10	7.7	2.3	3	2.8	0.9
9	na	na	na	8	10.0	1.8
10	12	9.0	2.4	1	0.0	0.0
11	8	16.3	2.1	9	16.5	2.2
12	na	na	na	11	8.0	1.9
13	9	10.6	2.2	12	16.8	2.4
14	12	13.4	2.5	10	7.7	2.0
15	6	9.3	1.8	3	2.9	1.1
16	10	13.8	2.2	na	na	na
17	na	na	na	5	9.7	1.5
18	7	24.8	2.0	5	21.6	1.6
19	6	13.3	1.8	5	6.2	1.6
20	11	25.0	2.4	4	21.9	1.4
21	11	22.4	2.4	8	19.7	2.1
22	7	5.3	1.8	1	0.0	0.0
23	11	17.1	2.4	3	4.5	1.1
24	5	10.1	1.6	5	15.3	1.6
25	9	12.5	2.2	5	9.1	1.4
26	16	13.7	2.8	9	9.8	2.2
27	7	17.5	2.0	9	7.4	2.2
28	7	14.1	2.0	5	4.4	1.0
29	5	5.8	1.6	na	na	na
30	9	25.1	2.2	6	11.4	1.8

Notes: *D*, mean pairwise genetic divergence between all primary tumor areas and all regional metastases of a given patient; *H*, clonal diversity representing the number and abundance of clones within primary tumors and among regional metastases per patient.
Abbreviation: LN, lymph node.

genomic alterations did not appear to be closely related to any of the primary tumor areas sampled, suggesting that they may have been derived from cells that diverged early in disease progression before many mutations accumulated in the primary carcinoma.

Discussion

In this study, we used a genomic fingerprinting approach to evaluate genomic heterogeneity within primary breast carcinomas and among axillary LN metastases. Multiple clonal cell lineages were observed in every primary tumor and between many metastatic deposits from the same patient. Phylogenetic analyses suggest that regional metastases likely originate independently from different sites within the primary carcinoma

and may exhibit variability in the timing of metastatic dissemination.

Clonal diversification is a prominent feature of primary breast carcinomas.[26,27] Heterogeneity throughout the primary tumor in patterns of point mutations and copy-number variation develops over time as clonal lineages of cells acquire genomic aberrations during tumor growth and differentiation. Tumor areas sampled here were approximately the same diameter as 14-gauge core needle biopsy specimens. The genomic diversity and divergence between these areas support the need for routine evaluation of multiple core biopsies in patient diagnosis, as pathological characterization using a single biopsy specimen would likely underestimate underlying
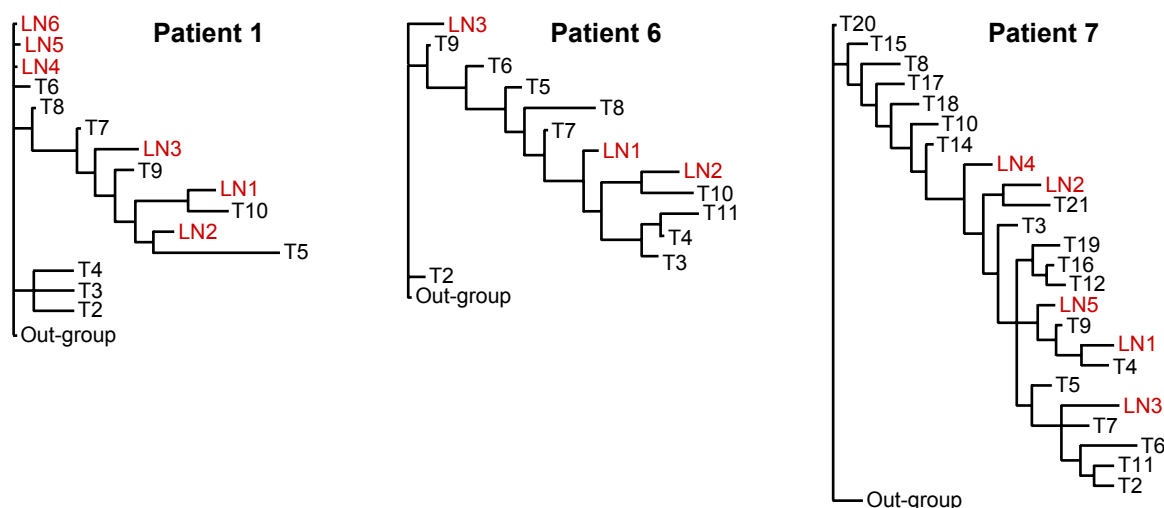
Figure 4. Most parsimonious phylogenetic trees from three representative patients depicting hereditary relationships among areas of the primary breast carcinoma (T) and axillary LN metastases based on the patterns of genomic changes assessed by AI.

molecular heterogeneity, which may have significant clinical consequences.[28] Genetic differences between cell lineages in the primary tumor may be responsible for the emergence of drug-resistant cells and eventual treatment failure. Because drug-resistant cells usually derive from minor subclones with different molecular characteristics than the bulk of the tumor, multiple biopsies of the primary carcinoma may be necessary to more thoroughly assess genetic heterogeneity. In this study, five tumor areas were needed to capture ~80% of the variability and eight areas were required to detect 90% of the genomic diversity.

Genetic differences between primary tumors and matched LN deposits may indicate that axillary node metastases arise from different clonal lineages within the primary tumor, and/or disseminate early in the disease process and evolve independently.[8–13] Largely irreversible chromosomal changes in the primary carcinoma should be present in all cells descending from that clonal lineage. Previous studies that sampled only a single region of the primary tumor may not have accurately captured the overall clonal variability, which could affect the interpretation of genetic alterations in the primary tumor and metastatic deposits. In this study, we showed that by sampling multiple areas of the primary carcinoma, genetic ancestry of some LN metastases could be mapped to specific areas of the primary tumor, while others were not closely related to any of the sampled areas, indicating spatial and temporal variability of metastatic dissemination.

Choice of therapeutic regimen is guided by pathological evaluation of the primary carcinoma; however, metastases within a given patient have been shown to differ in response to chemotherapy and hormone therapy,[29] which may be partially explained by genomic diversity. Previous research using large-scale sequencing of a lobular breast carcinoma demonstrated that the associated metastasis contained 19 mutations not present in the primary tumor.[12] Similar observations have been

reported for pancreatic cancer metastases.[30] Independent mutations may render metastases unresponsive to treatments based on the characteristics of the primary carcinoma and actionable mutations in the metastases may not be targeted. In this study, greater genetic divergence among axillary LN metastases was inversely associated with survival. This association, observed in the regional metastases but not the primary carcinoma, may be important in therapeutic resistance. Although genomic divergence among axillary metastases may be useful as a prognostic tool, less aggressive surgical approaches for managing patients with axillary disease[31] may limit accessibility to regional metastases for measuring genomic divergence. Future studies evaluating heterogeneity within sentinel LN metastases may determine whether genomic diversity is clinically useful for predicting survival.

Limitations to our study included the small number of patients with sufficient tissue from primary tumors and LN metastases available for analysis, which limited the conclusions that could be drawn from the data. Although a number of steps were taken to ensure the accuracy of AI detection,[32] DNA from FFPE samples is prone to technical artifact,[33,34] and thus may not accurately estimate the number of genetic alterations. Given the small areas of tumor sampled, we examined hypervariable STR regions representing a small portion of the genome to derive molecular fingerprints of these areas and did not use global technologies such as comparative genomic hybridization or next-generation sequencing. The possibility that AI events occurred independently in separate regions and that multiple clonal lineages were present within some of the 3 mm^2 regions examined may have led to an underestimation of genomic heterogeneity. Finally, these analyses were performed using tissues from advanced breast carcinomas; thus, these findings may not be applicable to patients with smaller primary tumors and/or fewer metastatic LNs.

Conclusions

Molecular heterogeneity represents a significant challenge for modern chemotherapeutics. We observed genomic heterogeneity within primary breast carcinomas and among regional LN metastases, which may have important implications for managing breast cancer patients and improving clinical outcomes. Next-generation treatment regimens should effectively target independent lineages of tumor cells that may escape current therapies in efforts to improve patient survival.

Author Contributions

Conceived and designed the experiments: REE and DLE. Conducted genomic assays: ALT, HLB, and AD. Performed pathology and histology: JAH and BD. Analyzed the data: KAM and NSC. Wrote the first draft of the manuscript: REE and DLE. Contributed to the writing of the manuscript: ALT, HLB, AD, BD, KAM, NSC, and JAH. Made critical revisions and approved the final version: REE, DLE, and CDS. All the authors reviewed and approved the final manuscript.

REFERENCES

1. Russnes HG, Navin N, Hicks J, Borresen-Dale AL. Insight into the heterogeneity of breast cancer through next-generation sequencing. *J Clin Invest*. 2011; 121(10):3810–3818.
2. Davis BW, Zava DT, Locher GW, Goldhirsch A, Hartmann WH. Receptor heterogeneity of human breast cancer as measured by multiple intratumoral assays of estrogen and progesterone receptor. *Eur J Cancer Clin Oncol*. 1984;20(3): 375–382.
3. Nassar A, Radhakrishnan A, Cabrero IA, Cotsonis GA, Cohen C. Intratumoral heterogeneity of immunohistochemical marker expression in breast carcinoma: a tissue microarray-based study. *Appl Immunohistochem Mol Morphol*. 2010;18(5):433–441.
4. Pertschuk LP, Axiotis CA, Feldman JG, Kim YD, Karavattayhayyil SJ, Braithwaite L. Marked intratumoral heterogeneity of the proto-oncogene Her-2/neu determined by three different detection systems. *Breast J*. 1999;5(6):369–374.
5. Navin N, Krasnitz A, Rodgers L, et al. Inferring tumor progression from genomic heterogeneity. *Genome Res*. 2010;20(1):68–80.
6. Navin N, Kendall J, Troge J, et al. Tumour evolution inferred by single-cell sequencing. *Nature*. 2011;472(7341):90–94.
7. Almendro V, Fuster G. Heterogeneity of breast cancer: etiology and clinical relevance. *Clin Transl Oncol*. 2011;13(11):767–773.
8. Becker TE, Ellsworth RE, Deyarmin B, et al. The genomic heritage of lymph node metastases: implications for clinical management of patients with breast cancer. *Ann Surg Oncol*. 2008;15(4):1056–1063.
9. Ellsworth RE, Ellsworth DL, Neatrour DM, et al. Allelic imbalance in primary breast carcinomas and metastatic tumors of the axillary lymph nodes. *Mol Cancer Res*. 2005;3(2):71–77.
10. Kuukasjärvi T, Karhu R, Tanner M, et al. Genetic heterogeneity and clonal evolution underlying development of asynchronous metastasis in human breast cancer. *Cancer Res*. 1997;57(8):1597–1604.
11. Torres L, Ribeiro FR, Pandis N, Andersen JA, Heim S, Teixeira MR. Intratumor genomic heterogeneity in breast cancer with clonal divergence between primary carcinomas and lymph node metastases. *Breast Cancer Res Treat*. 2007; 102(2):143–155.
12. Shah SP, Morin RD, Khattra J, et al. Mutational evolution in a lobular breast tumor profiled at single nucleotide resolution. *Nature*. 2009;461(7265):809–813.
13. Ding L, Ellis MJ, Li S, et al. Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature*. 2010;464(7291):999–1005.
14. Aitken E, Osman M. Factors affecting nodal status in invasive breast cancer: a retrospective analysis of 623 patients. *Breast J*. 2010;16(3):271–278.
15. Amir E, Seruga B, Niraula S, Carlsson L, Ocaña A. Toxicity of adjuvant endocrine therapy in postmenopausal breast cancer patients: a systematic review and meta-analysis. *J Natl Cancer Inst*. 2011;103(17):1299–1309.
16. Dikicioglu E, Barutca S, Meydan N, Meteoglu I. Biological characteristics of breast cancer at the primary tumour and the involved lymph nodes. *Int J Clin Pract*. 2005;59(9):1039–1044.
17. Bedard PL, Hansen AR, Ratain MJ, Siu LL. Tumour heterogeneity in the clinic. *Nature*. 2013;501(7467):355–364.
18. Klein CA. Selection and adaptation during metastatic cancer progression. *Nature*. 2013;501(7467):365–372.
19. Deyarmin B, Kane JL, Valente AL, et al. Effect of ASCO/CAP guidelines for determining ER status on molecular subtype. *Ann Surg Oncol*. 2013;20(1):87–93.
20. Ellsworth DL, Shriver CD, Ellsworth RE, Deyarmin B, Somiari RI. Laser capture microdissection of paraffin-embedded tissues. *Biotechniques*. 2003;34(1): 42–46.
21. Osborne RJ, Hamshere MG. A genome-wide map showing common regions of loss of heterozygosity/allelic imbalance in breast cancer. *Cancer Res*. 2000;60(14): 3706–3712.
22. Medintz IL, Lee CC, Wong WW, Pirkola K, Sidransky D, Mathies RA. Loss of heterozygosity assay for molecular detection of cancer using energy-transfer primers and capillary array electrophoresis. *Genome Res*. 2000;10(8): 1211–1218.
23. Ellsworth DL, Ellsworth RE, Love B, et al. Genomic patterns of allelic imbalance in disease free tissue adjacent to primary breast carcinomas. *Breast Cancer Res Treat*. 2004;88(2):131–139.
24. Maley CC, Galipeau PC, Finley JC, et al. Genetic clonal diversity predicts progression to esophageal adenocarcinoma. *Nat Genet*. 2006;38(4):468–473.
25. Beroukhim R, Mermel CH, Porter D, et al. The landscape of somatic copy-number alteration across human cancers. *Nature*. 2010;463(7283):899–905.
26. Marusyk A, Almendro V, Polyak K. Intra-tumour heterogeneity: a looking glass for cancer? *Nat Rev Cancer*. 2012;12(5):323–334.
27. Nik-Zainal S, Van Loo P, Wedge DC, et al; Breast Cancer Working Group of the International Cancer Genome Consortium. The life history of 21 breast cancers. *Cell*. 2012;149(5):994–1007.
28. Aparicio S, Caldas C. The implications of clonal genome evolution for cancer medicine. *N Engl J Med*. 2013;368(9):842–851.
29. Huyge V, Garcia C, Alexiou J, et al. Heterogeneity of metabolic response to systemic therapy in metastatic breast cancer patients. *Clin Oncol (R Coll Radiol)*. 2010; 22(10):818–827.
30. Yachida S, Jones S, Bozic I, et al. Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature*. 2010;467(7319):1114–1117.
31. Giuliano AE, McCall L, Beitsch P, et al. Locoregional recurrence after sentinel lymph node dissection with or without axillary dissection in patients with sentinel lymph node metastases: the American College of Surgeons Oncology Group Z0011 randomized trial. *Ann Surg*. 2010;252(3):426–432.
32. Holliday C, Rummel S, Hooke JA, Shriver CD, Ellsworth DL, Ellsworth RE. Genomic instability in the breast microenvironment? A critical evaluation of the evidence. *Expert Rev Mol Diagn*. 2009;9(7):667–678.
33. Farrand K, Jovanovic L, Delahunt B, et al. Loss of heterozygosity studies revisited: prior quantification of the amplifiable DNA content of archival samples improves efficiency and reliability. *J Mol Diagn*. 2002;4(3):150–158.
34. Rummel S, Valente AL, Kane JL, Shriver CD, Ellsworth RE. Genomic (in)stability of the breast tumor microenvironment. *Mol Cancer Res*. 2012;10(12): 1526–1531.