


RESEARCH ARTICLE

Open Access



Chromoanasythesis is a common mechanism that leads to *ERBB2* amplifications in a cohort of early stage HER2⁺ breast cancer samples

George Vasmatzis^{1,2,3*} , Xue Wang⁴, James B. Smadbeck^{1,2}, Stephen J. Murphy^{1,2}, Katherine B. Geiersbach⁵, Sarah H. Johnson^{1,2}, Athanasios G. Gaitatzes^{1,2}, Yan W. Asmann⁴, Farhad Kosari^{1,2}, Mitesh J. Borad⁶, Daniel J. Serie^{1,2}, Sarah A. McLaughlin⁷, Jennifer M. Kachergus⁸, Brian M. Necela⁸ and E. Aubrey Thompson^{9*}

Abstract

Background: HER2 positive (HER2+) breast cancers involve chromosomal structural alterations that act as oncogenic driver events.

Methods: We interrogated the genomic structure of 18 clinically-defined HER2+ breast tumors through integrated analysis of whole genome and transcriptome sequencing, coupled with clinical information.

Results: *ERBB2* overexpression in 15 of these tumors was associated with *ERBB2* amplification due to chromoanasythesis with six of them containing single events and the other nine exhibiting multiple events. Two of the more complex cases had adverse clinical outcomes. Chromosomes 8 was commonly involved in the same chromoanasythesis with 17. In ten cases where chromosome 8 was involved we observed *NRG1* fusions (two cases), *NRG1* amplification (one case), *FGFR1* amplification and *ADAM32* or *ADAM5* fusions. *ERBB3* over-expression was associated with *NRG1* fusions and *EGFR* and *ERBB3* expressions were anti-correlated. Of the remaining three cases, one had a small duplication fully encompassing *ERBB2* and was accompanied with a pathogenic mutation.

Conclusion: Chromoanasythesis involving chromosome 17 can lead to *ERBB2* amplifications in HER2+ breast cancer. However, additional large genomic alterations contribute to a high level of genomic complexity, generating the hypothesis that worse outcome could be associated with multiple chromoanasythetic events.

Keywords: Amplification, Replication, Chromothripsis, Chromoplexy, Chromoanagenesis, Chromoanasythesis, Chromodesmy, Neochromosome

Background

Large genomic rearrangements have emerged as a potential source of oncogenic driver mutations in breast cancer [1]. The classical example of the role of large genomic rearrangements as oncogenic drivers is HER2+ breast cancer in which the *ERBB2* gene (encoding the HER2 receptor subunit) is amplified along with several other genes in the vicinity of chromosome 17q12. The

ERBB2 gene is an important oncogenic driver in at least 15% of invasive breast cancers. Amplification of *ERBB2* at the DNA level leads to over expression of HER2 receptor tyrosine kinase protein on the cell surface [2], which is believed to drive malignant transformation due to hyper-activation of downstream signaling pathways that impinge upon proliferation and survival. Breast cancers with *ERBB2* gene amplification were associated with a poor prognosis prior to the availability of HER2-targeted therapy [3, 4].

The clinical relevance of the HER2 receptor increased with the development of Herceptin (trastuzumab), the first HER2-tageted monoclonal antibody therapy for

* Correspondence: vasm@mayo.edu; thompson.aubrey@mayo.edu

¹Department of Molecular Medicine, Mayo Clinic, 200 First St., SE, Rochester, MN 55905, USA

⁹Cancer Biology, Mayo Clinic, Griffin Building 214, Jacksonville, Florida, USA

Full list of author information is available at the end of the article



treating patients with breast cancer [5, 6]. Patients treated with Herceptin showed improved survival in early clinical trials, paving the way for the clinical use of HER2-targeted therapy [7–9]. Current therapeutic options for HER2+ breast tumors target the HER2 receptor using either monoclonal antibodies (e.g. trastuzumab, pertuzumab) or small molecule receptor tyrosine kinase inhibitors (e.g. lapatinib, neratinib). Early stage HER2+ tumors are quite responsive to such therapy, with long term recurrence free survival achieved in 75–80% of patients [7–10].

The definition of what constitutes clinically significant *ERBB2* amplification or HER2 overexpression has evolved over the years since the discovery of this biomarker. For in situ hybridization methods, Slamon et al. initially used the *ERBB2* copy number (>5 copies as amplified) [2, 3], and later the HER2/chr17 centromere ratio became the preferred determinant of HER2 amplification status, with a ratio of 2.0 or greater defining amplification [11]. Immunohistochemistry was used to identify overexpression, with evolving definitions for the minimum level of staining intensity and the staining pattern on the cell membrane. Guidelines for HER2 testing in breast cancer have been published by an expert panel with members of ASCO (American Society of Clinical Oncology) and CAP (College of American Pathologists) to standardize clinical testing and to refine the criteria for a positive HER2 result [12, 13]. These observations raise two important questions. First, from the standpoint of tumor biology, what features generally define the genomic architecture of HER2+ tumors, with particular regard for large genomic rearrangements that may extend beyond the *ERBB2* amplicon? Second, from a clinical-translational perspective, to what extent do these large chromosomal rearrangements contribute to genomic complexity that might account, at least in part, for the 20–25% recurrence rates after HER2-targeted therapy?

Gene amplifications commonly arise from replications of small regions of the genome. Possible mechanisms include the generation of either multiple tandem duplications or acentric extrachromosomal DNA circles [14]. However, more recently through the analysis of whole genome sequence data, these focal amplifications were frequently observed associated with complex chromosomal shuffling events [14]. These complex-shuffling events involved either one chromosome or two or even more chromosomes and they have been referred to as chromothripsis [15, 16] or chromoplexy [17]. Structurally, this process resembles single or multiple chromosomal knotting that eventually results in generation of one or more new chromosomes that contain genomic information from the parent chromosomes. The result of this process is often referred to as chromoanagenesis [18] or chromoanasythesis [19, 20]. The newly formed

chromosome is now susceptible to focal replication that is likely driven by selective advantage conveyed by the generation of oncogenic drivers. To reconcile the definitions of this phenomenon for the purpose of this paper we decided to use a new term “chromodesmy” to encapsulate all the different generating mechanisms and the term “chromoanasythesis” to encapsulate the mechanism that leads to the resulting new chromosomes, which below will be referred to as “neochromosomes”.

The extent to which these processes are linked to *ERBB2* amplification in HER2+ breast cancer is largely unknown. We analyzed 18 HER2+ breast tumors using a combination of mate-pair genomic sequencing (MPseq) [21–23], RNA sequence analysis (RNAseq), and NanoString 3D Biology™ to assess the extent to which large chromosomal alterations both within and outside of chr 17q12 are observed and are associated with *ERBB2* amplification and overexpression. The data reveal that chromodesmic processes involving chromosome 8 and chromosome 17 are commonly, but not invariably, associated with *ERBB2* amplification and overexpression.

Methods

The aim of this study was to interrogate the genomic structure of 18 HER2+ breast tumors through integrated analysis of whole genome and transcriptome sequencing, coupled with clinical information. All tumors specimens were obtained from the Mayo Clinic biobank and the study was performed under full Mayo Clinic Institutional Review Board (IRB) approval with written consent obtained from all patients. Biospecimen handling information as outlined by the BRISQ criteria [24], is included as a supplemental table (Additional file 1: Table S1).

All tumors were clinically defined as HER2+ by ASCO/CAP guidelines: HER2 IHC 3+ and/or FISH > 2.0 in > 10% of tumor cells. HER2 IHC and *ERBB2* FISH were performed by Mayo Medical Laboratories as part of the patient's routine breast cancer diagnosis. Fresh frozen tumor specimens were cryo-sectioned and RNA and DNA extracted using routine protocols [25]. RNA integrity was evaluated by Agilent Bioanalyzer and RINs > 8.0 were observed for all samples subjected to RNAseq analysis. An alpha version of the NanoString 3D Biology™ platform was used to assess HER2 protein abundance.

MPseq was used to detect structural variants at gene level resolution through its specialized whole genome tiling with larger 2–5 kb fragment derived DNA libraries [26–35]. MPseq and RNAseq transcriptomic analysis were performed on 24 HER2+ breast cancer samples. Four samples had insufficient tumor (< 10%) to define rearrangements and were excluded from further analysis. Two samples were excluded due to ambiguities in the tumor registry. Indexed libraries for MPseq (1µg DNA) and RNAseq (150 ng total RNA) were generated using the Nextera

Mate-Pair Kit (Illumina, CA, FC-132-1001) or the TruSeq RNA Library Prep Kit v2 (Illumina, CA, RS-122-2001), following the manufacturer's instructions. Libraries were sequenced on the Illumina HiSeq2000 platform at a depth of three or six libraries per lane, respectively.

Detection of structural variants

BIMA, developed by Biomarker Discovery Lab at Mayo Clinic, mapped all MPseq fragments to GRCh38. BIMA is a binary indexing algorithm for simultaneous mapping of both reads in a fragment, specially designed for MPseq [26]. Structural variants were detected using SVAtools, version 0.34.16,, a suite of algorithms also developed by the Biomarker Discovery Lab at Mayo Clinic [22, 23]. Four main components of SVAtools includes: 1) junction detection with a customized rapid clustering algorithm to detect discordant fragments supporting a common junction, 2) a system of masks and filters to remove false-positive junctions. The mask primarily eliminates normal structural variants not present in the reference genome and eliminates mapping artifacts due to repeat or un-sequenced genomic regions. The filters use BIMA mapping scores to identify NGS library prep artifacts and to eliminate poorly qualified breakpoints. 3) CNV detection, using the read count of concordant fragments within non-overlapping bins. The detected junctions provide enhanced edge detection resolution and sensitivity. 4) Visualization of all structural variants via genome plots [21]. Putative junctions as well as any two genomic regions of interest can be visualized and further inspected via junction plots and region plots, illustrations of all fragments mapping within and between two genomic regions.

Following junction detection SVAtools' CNV detection was used to determine the location and depth of copy number variation across the genome. This algorithm uses both a sliding window statistical method to determine likely copy number edges from read depth, as well as using breakpoints locations determined in the junction detection stage to more accurately place these edges. Once the genome was segmented into likely copy number regions, the normalized read depth (NRD) was calculated as the read depth within the region was divided by the expected read depth for normal diploid level for the sample. Regions with NRD scores that deviated significantly from the expected diploid level (NRD = 2.0) were called a copy number variant. This NRD score was used to estimate the level of amplification in a region according to the following equation:

$$X_i = \frac{|NRD_i - NRD'|}{\tau}$$

X_i is the change in copy number for a region, NRD_i is the NRD value calculated for the region, NRD' is the

expected normal diploid NRD level, and τ is the fraction of tumor cells in the sample. Tumor fraction was determined by calculating the cumulative NRD score for all regions called a loss in a sample. The difference between this cumulative NRD score and expected diploid level (NRD = 2.0) was the tumor fraction. If the deleted regions were not enough to calculate the tumor fraction the gained CNV regions were used. If neither was available then the tumor fraction was considered indeterminate.

A tumor was denoted as chromoanasythetic if 5 or more junctions were found between any two chromosomes.

RNAseq

RNA-sequencing libraries for 18 samples were prepared according to the Illumina truseq protocol and run on the HiSeq2000 platform. The RNA-Seq Paired end sequence data were reran using MAP-RSeq version 3.0.0 [36], an integrated RNA-Seq bioinformatics pipeline developed at the Mayo Clinic for comprehensive analysis of raw RNA sequencing paired-end reads. MAP-RSeq employs the very fast, accurate and splice-aware aligner, STAR [37], to align reads to the reference human genome, build hg38. The aligned reads are then processed through a variety of modules in a parallel fashion. Gene and exon expression quantification is performed using the Subread [38] package to obtain both raw and normalized (RPKM – Reads Per Kilobase per Million mapped) reads. Finally, the “.count” files from the previous step were used by the edgeR (version 3.16.5, R 3.3.1) program to generate a normalized expression matrix for all samples. The *ERBB2* mRNA abundance was extracted from the normalized RNAseq data.

NanoString 3D biology™

An alpha version of the 3D Biology platform was used to assess mRNA abundance (Pan Cancer Pathways), protein expression (including *ERBB2*), and single nucleotide variants. Data were analyzed using the nSolver Advanced Applications software.

Results

MPseq is a very efficient method to examine the rearrangement landscape of tumor cells. We focused in chromosomal junctions (connections between distant breakpoint of the genome), to investigate how *ERBB2* amplifications arise in HER2+ breast cancer. All the examined tumors exhibited aberrant junctions. The total junction counts of all discordantly mapping genomic breakpoints detected in the 18 clinically determined HER2+ cases successfully analyzed by MPseq varied from 31 to 400 (Table 1). Out of these samples, 16 exhibited extensive chromoanasythesis (Table 1).

We used a whole genome visualization layout featuring the Gnome U plot [21] to inspect the rearrangement

Table 1 Summary of molecular data in all cases analyzed. The Junction-count field displays the number of junctions with 4 or more associates. The number of events field displays the number of independent chomoanasythetic events. The chromosome number when the chromoanasythesis is confined in that chromosome is color-coded red

Sample Name	Junction-count	Number of events	Chromosomes Involved in chromoanasythesis	HER2 IHC	Estimated extra copies of <i>ERBB2</i>	Log2 <i>ERBB2</i> mRNA	HER2 protein (NanoString)	Clinical outcome to 09/2017
MCJBCR-068	299	3	8,2-15-20, 10-12-17-19		47	18.8	10855	Deceased
BRB-255	47	2	17, 1-8 (NRG1 fusion)	3+	46	18.6	23654	
BRB-144	31	1	8-11-17-18	3+	40	18.3	5335	
MCJBCR-085	400	6	14,15,16,19,20,2-4-6-8-11-17	3+	29	18.3	10942	
BRB-227	211	2	1-17-7, 11-13	3+	32	18.2	4890	
BRL-451	49	1	4-17		20	18.0	1257	
BRB-085	90	3	1,17, 8-12	3+	30	17.7	6144	Relapsed
BRB-041	185	1	7-8-17 (NRG1 fusion)	3+	25	17.7	4890	
BRB-277	122	2	6-9-14-17-19, 8-10	3+	20	17.5	3767	
BRB-147	122	1	8-17	3+	24	17.4	3365	
BRB-114	320	2	1-9, 6-7-14-17-21	3+	20	17.2	5476	
MCJBCR-107	35	0		3+	1	16.9	1121	
BRB-009	94	4	5,16, 17, 8-19	2+	8	16.1	1945	
MCJBCR-064	110	1	3-14-17-18	3+	11	16.0	1945	
BRB-180	225	3	5-6, 7-8-17-19, 10-14	3+	21	15.8	1156	
BRB-239	14	1	12	2+	6	15.7	2111	
BRB-165	13			2+	1	15.1	1962	
BRB-235	20		8-11 (NRG1 amplification)	2+	1	13.8	1136	

architecture of all samples. Fig. 1 illustrates the genome plot of a representative HER2+ breast tumor (BRB-041). This tumor harbored a single 3-way chromodesmy event involving primarily chromosomes 7, 8, and 17. After this 3-way event, the tumor genome would resolve to newly synthesized neochromosomes, quite distinct from the patient's normal diploid genome. The newly synthesized neochromosomes were deprived of large portions of 7q, 8p and 17p but showed gains and focal amplifications on the remaining portions. As a consequence of this chromodesmy, *ERBB2* was greatly amplified resulting in at least 25 additional copies. *ERBB2* was the ninth most abundant transcript by RNAseq (17.7 at the log2 scale of RPKM) in this case. The dependence of this tumor on the HER2 signaling pathway was also supported by the observed high protein expression from IHC and the Nanostring assay (Table 1).

We also examined all the other genes that could be influenced by this chromodesmy event. This event also yielded a *WIPF2-NRG1* fusion on chromosome 8, corroborated by RNAseq that lead to marked up-regulation

of *NRG1* (highest *NRG1* expression among all cases). Similarly, *BRAF* was also amplified on chromosome 7 and lead to one of the highest *BRAF* expressions among all cases. Multiple other genes were influenced on chromosomes 7, 8 and 17 including an *RAI1-UNC5D* fusion. The rest of the tumor genome was relatively diploid and unaffected by the chromodesmy event with the exception of 1q gain, and deletions on the latter parts of the q arms of chromosomes 5, 11, and 14. The main questions that arise after examining the architecture of the case above are, first; how often chromodesmy events exist in HER2+ breast cancer, second; how consistent are these events, and third; what happens when there is no evidence of focal *ERBB2* amplification.

Cases with *ERBB2* amplifications

We then investigated the commonality of chromodesmy events in all the available samples (Additional file 1: Figure S1-S17). The most common phenomenon observed among all cases involved chromodesmy of two to six chromosomal bundles. The data are summarized in

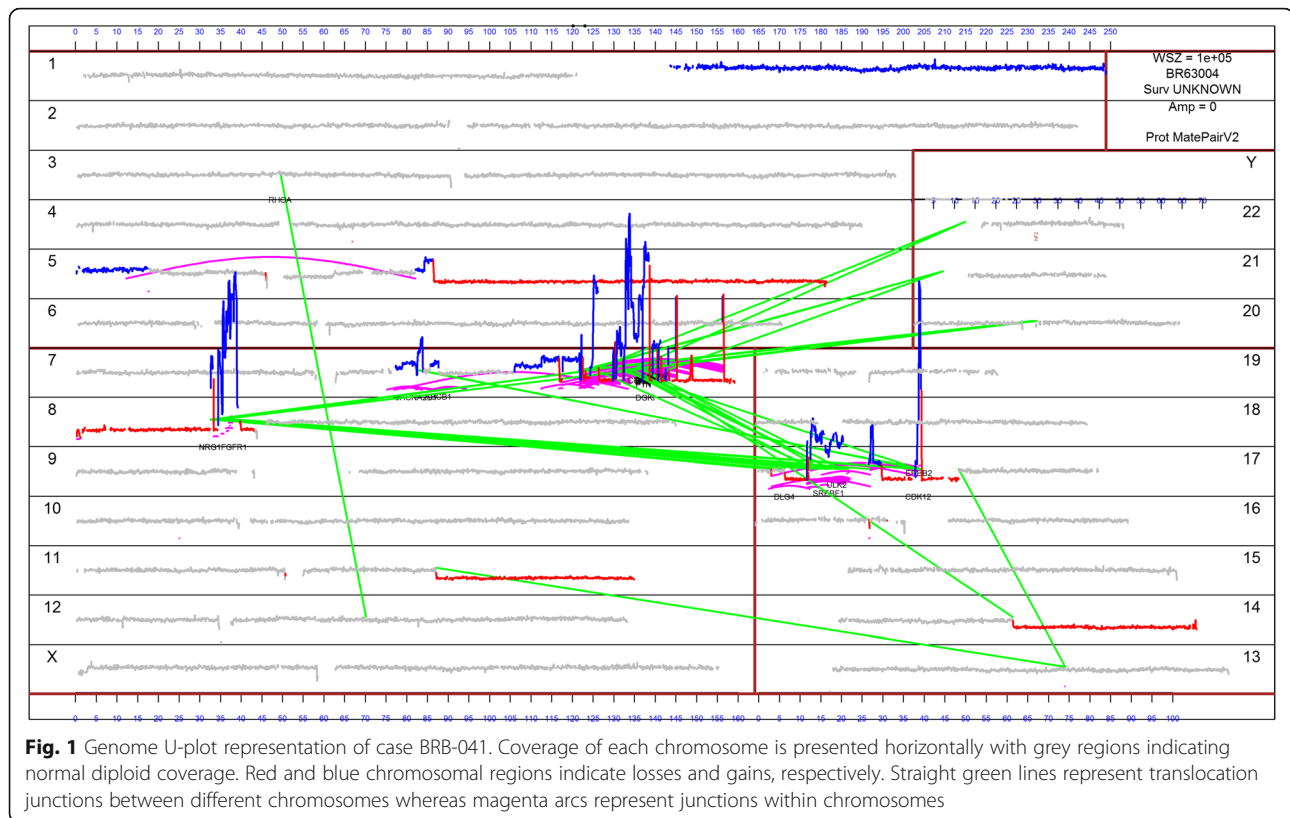


Table 1. A graphic representation of the frequency of the involved chromosomes is shown in Fig. 2. Twelve out of the 18 cases exhibit chromodesmy of multiple chromosomes that resulted in *ERBB2* amplifications. The subsequent chromoanasytosis not only resulted in focal amplifications of *ERBB2* but also involved amplifications of other areas of the genome in all cases primarily with chromosomes 8 and 7 (Fig. 2), but less frequently with other chromosomes (see below). Two cases, BRB-255 and BRB-085, exhibit *ERBB2* amplification through intra-17 chromodesmy without an apparent involvement of other chromosomes (Additional file 1: Figure S1-S17). Finally, BRB-239 was the only case that had a classic focal *ERBB2* amplification without evidence of chromodesmy. The total number of genomically-confirmed *ERBB2*-amplified HER2+ tumors was 15. As expected, the *ERBB2*-amplified cases had the highest *ERBB2* expression (Table 1). The HER2 copy number correlated very well with the mRNA abundance (Spearman's rho = 0.849, $p < 0.0001$) and protein expression (Spearman's rho = 0.777, $p = 0.0002$). The only case that did not have *ERBB2* amplification but ranked 12th on the expression order was MCJBCR-107 (see below).

Additional observational analysis between cases was performed with the intention of finding commonalities and differences. The 15 *ERBB2*-amplified cases can be roughly divided into two categories; simple and complex,

according to the number of coordinated chromodesmic events. The simple chromodesmy cases (underlined sample names in Table 1), like BRB-041, contained a single coordinated event that leads to amplification of *ERBB2*. The complex cases (bolded sample names in Table 1), like MCJBCR-068, corresponded to cases with multiple, potentially independent, chromodesmic events. MCJBCR-068 was one of the most complex cases with the highest *ERBB2* expression, containing three chromodesmic events and originated from a patient who relapsed and died after HER2-targeted therapy. A second patient, BRB-085, with complex events also relapsed.

Chromosome 8 rearrangements were the most frequently observed events (Table 1). Ten of the 18 cases involved events in a 30–40 Mb region of chromosome 8, associated with a number of genes that could be related to HER2 signaling, including *NRG1*, *FGFR1*, *UNC5D* and *ADAM5*. Two of these cases (BRB-041 and BRB-255) had *NRG1* fusions supported by both MPseq and RNAseq. *ERBB3* expression was also high in these two cases compared to other cases, whereas *EGFR* was low indicating HER3 as a likely partner of HER2 in these two cases (Additional file 1: Figure S18). Overall, *EGFR* and *ERBB3* expression were anti-correlated in the *ERBB2* amplified cases (Additional file 1: Figure S18) by a correlation coefficient -0.56 (p -value = 0.025, 95% confidence interval: -0.860 -0.085).

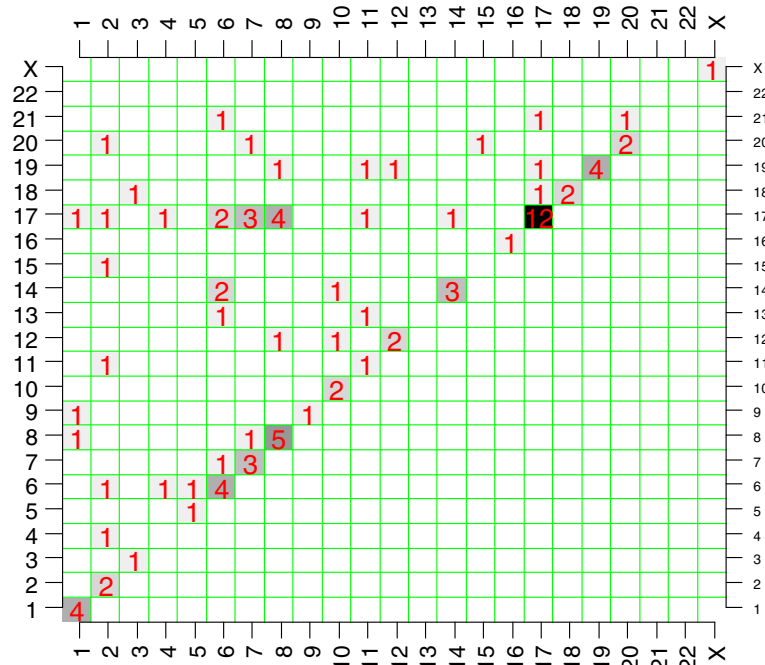


Fig. 2 A visual synthesis of all complex junctions from all cases. The chromoanasythetic events within and between chromosomes detected by MPseq in the 18 Breast cancer specimens are summarized above, with the numbers showing how many events involved the respective chromosomes

Cases that did not exhibit *ERBB2* amplifications

Although, all the cases examined were clinically rendered HER2+ by IHC, they did not all show a clear amplification signal by MPseq. The remaining three HER2+ cases did not display *ERBB2* amplification at the chromosomal level. The tumor percentage in these tumor tissues ranged from 23 to 56%, which would be adequate to demonstrate amplifications. With the exception of MCJBCR-107, which ranked 12th overall in *ERBB2* mRNA abundance, the *ERBB2* expression of the remaining cases was lower than the *ERBB2*-amplified cases described above (Table 1).

A more thorough analysis of these cases revealed additional chromosomal abnormalities involving *ERBB2*. MCJBCR-107 interestingly, had a small duplication that included the *ERBB2* gene and exhibited a pathogenic variant at position chr17:39724008 (https://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs=121913468, <http://cancerdiscovery.aacrjournals.org/content/3/2/224.long>). There are 740 reads supporting the wild type G in that case, with 2103 reads supporting the mutated T. So 75% mutated vs. 25% wild type. The mutation G->T results in a change from Asp (D) -> Tyr (Y). The HER2 IHC clinical staining score was 3 but the FISH ratio for this tumor was 1.04. The high *ERBB2* RNA expression corroborated with strong IHC HER2 staining, the small duplication around *ERBB2*, and the pathogenic mutation points to an additional mechanism and possible biomarker for HER2+ breast cancer.

Case BRB-235 contained a gain in a large region of chromosome 17 that included *ERBB2*. It also displayed a complex event between chromosomes 8 and 11 that lead to an amplification of *NRG1*. BRB-165 gained 17q and exhibits extensive aneuploidy but no chromoanasythesis (Additional file 1: Figure S16).

Discussion

Early experiments using karyotyping and metaphase FISH in *ERBB2*-amplified HER2+ breast cancers showed that the classical pattern was high copy gain with the extra *ERBB2* copies residing on one or more chromosomes, typically not the chromosome of origin (chromosome 17) but another chromosome, in “homogeneously staining regions” that contained co-amplified sequences from several other chromosomes [39]. By interphase FISH analysis, the *ERBB2* signals were present as one or more clusters, rather than diffusely scattered throughout the cell, consistent with the intrachromosomal location of the amplicons [40]. Technical limitations of earlier technologies such as comparative genomic hybridization, karyotyping, and Southern blot made it impossible to determine the composition and genomic architecture of amplicons at the sequence level. More recent large scale genomic studies have further defined the spectrum of genomic abnormalities in *ERBB2*-amplified HER2+ breast cancers, but the methods used for these studies precluded a detailed analysis of the large-scale genomic architecture of *ERBB2*-containing amplicons [41, 42]. In

the current study, for the first time, we have performed a detailed analysis of *ERBB2* amplification using mate pair whole genome sequencing.

Chromosomal rearrangements leading to fusions and amplifications are often the main structural drivers that lead to cancer progression and point to the targeted treatments that could benefit the patient. Amplifications such as *ERBB2* in HER2+ breast cancer also arise from structural rearrangements and can be examined using comprehensive sequencing technologies such as MPseq and RNAseq. To this end, MPseq provides a high resolution picture of the DNA structure and RNAseq can provide functional insight of how genes are expressed in the context of rearrangements.

An obvious question arises with respect to the clinically HER2+ tumors that do not appear to evidence *ERBB2* amplification or HER2 overexpression at the level of the bulk tumor. One obvious possibility might be a relatively low percentage of tumor cells in the sample. However, all of these samples exhibit gross chromosomal rearrangements, indicative of a high level of tumor cell enrichment. A more likely possibility is that these tumors are heterogeneous and comprised of a small percentage of HER2+ cells within a larger population of HER2-negative tumor cells. Recall that ASCO/CAP guidelines require that only 10% of tumor cells must be IHC 3+ for HER2 in order for the tumor to be called HER2+. However, in a bulk analysis of the sort described in this report, the relatively less abundant genomic contribution of the small percentage of HER2+ tumor cells might be obscured by the contribution from the more abundant HER2-negative tumor cells. This possibility obviously raises an interesting question about the efficacy of HER2-targeted therapy in such heterogeneous tumors. However, that question is beyond the scope of this study. Unfortunately, we do not have sufficient samples or long term follow up data to rigorously assess therapeutic outcome as a function of any of the genomic features that we have defined.

Conclusion

Integrated DNA and RNA genomic analysis of HER2+ breast cancers, for the majority of cases, reveals that *ERBB2* amplifications are presented in the context of chromoanasythesis involving either chromosome 17 alone, or with other chromosomes. We carefully inspected 18 clinically determined HER2+ cases with whole genome mate pair sequencing and found that 15 exhibit clear focal *ERBB2* amplification. Of those, only three cases showed amplification on 17 that did not involve any of the other chromosomes. Of the remaining 12 cases, three exhibit 2-way chromodesmy, two 3-way, four 4-way, two 5-way and one 6-way. It is unclear if all the junctions contained in these highly complex events

are a result of single or multiple progressive chromosomal catastrophe events. It is much more plausible that an initial much simpler chromosomal event renders the area sensitive for subsequent junction generating events that result in gene amplifications that give advantage to cell survival and proliferation. Three other HER2+ cases by IHC displayed single copy gains on the region that included *ERBB2*, either by small duplication (in one case) or larger regions. The one sample that had a small duplication also had a pathogenic mutation.

We also observed a preferential choice of the other chromosomes involved; specifically chromosome 8 is often one of the other partners. A possible explanation is that there exist genes in areas of chromosome 8 that collaborate with *ERBB2* in the evolutionary advantage of the cells. One of these candidates is *NRG1*, a ligand of HER3, which appears to be involved in fusions associated with high expression of *NRG1* and the HER3 gene *ERBB3*. *EGFR* and *ERBB3* expressions were anti-correlated pointing towards specificity of HER2 heterodimerization with either HER1 or HER3.

Additional file

Additional file 1: Figure S1-S17. Genome U plots of all the additional cases. Figure S18 Heat map of the *EGFR* and *ERBB3* expression log₂ expression by RNAseq. Dark red indicates low expression where yellow indicates high expression. Table S1 BRISQ summary of tumor specimens. (PPTX 3536 kb)

Abbreviations

ASCO : American Society of Clinical Oncology; CAP : College of American Pathologists; IRB : Institutional Review Board; MPseq : Mate-Pair genomic sequencing; NRD : Normalized Read Depth; RNAseq : RNA sequence analysis; RPKM : Reads Per Kilobase per Million mapped

Funding

This research was supported by the Center of Individualized Medicine and from the Breast Cancer Research Foundation. The funding sources had no role in design of the study, in collection, analysis, and interpretation of the data, or in writing the manuscript.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

GV, EAT coordinated research, Data analysis, wrote the paper. SJM, FK, MJB, Experimental Methods, wrote the paper. SHJ, AGG, Data analysis and visualization, wrote the paper. YWA, JMK, BMN, XW, JBS, KBG, Data analysis, Experimental Methods, wrote the paper. SAM, coordination the clinical sample collection, wrote the paper. All authors read and approved the final manuscript.

Ethics approval and consent to participate

All tumors specimens were obtained from the Mayo Clinic biobank and the study was performed under full Mayo Clinic Institutional Review Board approval (IRB# 07-005237) with written consent obtained from all patients.

Consent for publication

Not applicable.

Competing interests

Sarah A. McLaughlin is a member of the editorial board (Associate Editor).

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Molecular Medicine, Mayo Clinic, 200 First St., SE, Rochester, MN 55905, USA. ²Center for Individualized Medicine, Mayo Clinic, 200 First St., SE, Rochester, MN 55905, USA. ³<https://www.wholegenome.io>. ⁴Health Sciences Research, Mayo Clinic, Jacksonville, Florida, USA. ⁵Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA. ⁶Hematology/Oncology, Mayo Clinic, Phoenix, Arizona, USA. ⁷General Surgery, Mayo Clinic, Jacksonville, Florida, USA. ⁸Cancer Research, Mayo Clinic, Jacksonville, Florida, USA. ⁹Cancer Biology, Mayo Clinic, Griffin Building 214, Jacksonville, Florida, USA.

Received: 15 February 2018 Accepted: 14 June 2018

Published online: 13 July 2018

References

- Hoadley KA, Yau C, Wolf DM, Cherniack AD, Tamborero D, Ng S, Leiserson MDM, Niu B, McLellan MD, Uzunangelov V, et al. Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. *Cell*. 2014;158:929–44.
- Slamon DJ, Godolphin W, Jones LA, Holt JA, Wong SG, Keith DE, Levin WJ, Stuart SG, Udove J, Ullrich A, et al. Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. *Science*. 1989;244:707–12.
- Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science*. 1987;235:177–82.
- Pritchard KI, Shepherd LE, O'Malley FP, Andrulis IL, Tu D, Bramwell VH, Levine MN. HER2 and responsiveness of breast cancer to adjuvant chemotherapy. *N Engl J Med*. 2006;354:2103–11.
- Cobleigh MA, Vogel CL, Tripathy D, Robert NJ, Scholl S, Fehrenbacher L, Wolter JM, Paton V, Shak S, Lieberman G, Slamon DJ. Multinational study of the efficacy and safety of humanized anti-HER2 monoclonal antibody in women who have HER2-overexpressing metastatic breast cancer that has progressed after chemotherapy for metastatic disease. *J Clin Oncol*. 1999;17:2639–48.
- Carter P, Fendly BM, Lewis GD, Sliwkowski MX. Development of herceptin. *Breast Dis*. 2000;11:103–11.
- Piccant-Gebhart MJ, Procter M, Leyland-Jones B, Goldhirsch A, Untch M, Smith I, Gianni L, Baselga J, Bell R, Jackisch C, et al. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. *N Engl J Med*. 2005;353:1659–72.
- Romond EH, Perez EA, Bryant J, Suman VJ, Geyer CE Jr, Davidson NE, Tan-Chiu E, Martino S, Paik S, Kaufman PA, et al. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. *N Engl J Med*. 2005;353:1673–84.
- Joensuu H, Kellokumpu-Lehtinen PL, Bono P, Alanko T, Kataja V, Asola R, Utriainen T, Kokko R, Hemminki A, Tarkkanen M, et al. Adjuvant docetaxel or vinorelbine with or without trastuzumab for breast cancer. *N Engl J Med*. 2006;354:809–20.
- Perez EA, Romond EH, Suman VJ, Jeong JH, Davidson NE, Geyer CE Jr, Martino S, Mamounas EP, Kaufman PA, Wolmark N. Four-year follow-up of trastuzumab plus adjuvant chemotherapy for operable human epidermal growth factor receptor 2-positive breast cancer: joint analysis of data from NCCTG N9831 and NSABP B-31. *J Clin Oncol Off J Am Soc Clin Oncol*. 2011;29:3366–73.
- Pauletti G, Godolphin W, Press MF, Slamon DJ. Detection and quantitation of HER-2/neu gene amplification in human breast cancer archival material using fluorescence in situ hybridization. *Oncogene*. 1996;13:63–72.
- Wolff AC, Hammond ME, Hicks DG, Dowsett M, McShane LM, Allison KH, Allred DC, Bartlett JM, Bilous M, Fitzgibbons P, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. *J Clin Oncol Off J Am Soc Clin Oncol*. 2013;31:3997–4013.
- Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, Dowsett M, Fitzgibbons PL, Hanna WM, Langer A, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J Clin Oncol Off J Am Soc Clin Oncol*. 2007;25:118–45.
- Matsui A, Ihara T, Suda H, Mikami H, Semba K. Gene amplification: mechanisms and involvement in cancer. *Biomolecular Concepts*. 2013;4:567–82.
- Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, Mudie LJ, Pleasance ED, Lau KW, Beare D, Stebbings LA, et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell*. 2011;144:27–40.
- Forment JV, Kaidi A, Jackson SP. Chromothripsis and cancer: causes and consequences of chromosome shattering. *Nat Rev Cancer*. 2012;12:663–70.
- Shen MM. Chromoplexy: a new category of complex rearrangements in the cancer genome. *Cancer Cell*. 2013;23:567–9.
- Holland AJ, Cleveland DW. Chromoanagenesis and cancer: mechanisms and consequences of localized, complex chromosomal rearrangements. *Nat Med*. 2012;18:1630–8.
- Righolt C, Mai S. Shattered and stitched chromosomes-chromothripsis and chromoanaysynthesis-manifestations of a new chromosome crisis? *Genes Chromosomes Cancer*. 2012;51:975–81.
- Plaisancie J, Kleinfinger P, Cances C, Bazin A, Julia S, Trost D, Lohmann L, Vigouroux A. Constitutional chromoanaysynthesis: description of a rare chromosomal event in a patient. *Eur J Med Genet*. 2014;57:567–70.
- Gaitatzes AJ, Smadbeck JB, Vasmatzis G. Genome U-plot: a whole genome visualization. *Bioinformatics*. 2018;34:5.
- Johnson S, Smadbeck J, Smoley S, Gaitatzes A, Murphy S, Harris F, Drucker T, Zenka R, Pitel B, Rowsey R, et al. SVAtools for junction detection of genome-wide chromosomal rearrangements by mate-pair sequencing (MPseq). *Cancer Genet*. 2018;221:1–18.
- Smadbeck J, Johnson S, Smoley S, Gaitatzes A, Drucker T, Zenka R, RKF, Murphy SNH, Aypar U, et al. Copy number variant analysis using genome-wide mate-pair sequencing. *Genes Chromosomes Cancer*. 2018.
- Moore HM KA, Jewell SD, et al. Biospecimen reporting for improved study quality (BRISQ). *Cancer Cytopathol*. 119:92–102.
- Norton N, Sun Z, Asmann YW, Serie DJ, Necela BM, Bhagwate A, Jen J, Eckloff BW, Kalari KR, Thompson KJ, et al. Gene expression, single nucleotide variant and fusion transcript discovery in archival material from breast tumors. *PLoS One*. 2013;8:e81925.
- Drucker TM, Johnson SH, Murphy SJ, Cradic KW, Therneau TM, Vasmatzis G. BIMA V3: an aligner customized for mate pair library sequencing. *Bioinformatics*. 2014;30:1627–9.
- Feldman AL, Dogan A, Smith DI, Law ME, Ansell SM, Johnson SH, Porcher JC, Ozsan N, Wieben ED, Eckloff BW, Vasmatzis G. Discovery of recurrent t(6;7)(p25.3;q32.3) translocations in ALK-negative anaplastic large cell lymphomas by massively parallel genomic sequencing. *Blood*. 2011;117:915–9.
- Nielsen CM, Hansen K, Andersen HM, Gerstoft J, Vestergaard BF. An enzyme labelled nuclear antigen immunoassay for detection of cytomegalovirus IgM antibodies in human serum: specific and non-specific reactions. *J Med Virol*. 1987;22:67–76.
- Vasmatzis G, Johnson SH, Knudson RA, Ketterling RP, Braggio E, Fonseca R, Viswanatha DS, Law ME, Kip NS, Ozsan N, et al. Genome-wide analysis reveals recurrent structural abnormalities of TP63 and other p53-related genes in peripheral T-cell lymphomas. *Blood*. 2012;120:2280–9.
- Murphy SJ, Chevillie JC, Zarei S, Johnson SH, Sikkink RA, Kosari F, Feldman AL, Eckloff BW, Karnes RJ, Vasmatzis G. Mate pair sequencing of whole-genome-amplified DNA following laser capture microdissection of prostate cancer. *DNA research : an international journal for rapid publication of reports on genes and genomes*. 2012;19:395–406.
- Kovtun IV, Chevillie JC, Murphy SJ, Johnson SH, Zarei S, Kosari F, Sukov WR, Karnes RJ, Vasmatzis G. Lineage relationship of Gleason patterns in Gleason score 7 prostate cancer. *Cancer research*. 2013;73:3275–84.
- Lasho T, Johnson SH, Smith DI, Crispino JD, Pardanani A, Vasmatzis G, Tefferi A. Identification of submicroscopic genetic changes and precise breakpoint mapping in myelofibrosis using high resolution mate-pair sequencing. *American journal of hematology*. 2013;88:741–6.
- Feldman AL, Vasmatzis G, Asmann YW, Davila J, Middha S, Eckloff BW, Johnson SH, Porcher JC, Ansell SM, Caride A. Novel TRAF1-ALK fusion identified by deep RNA sequencing of anaplastic large cell lymphoma. *Genes, chromosomes & cancer*. 2013;52:1097–1102.
- Gao G, Johnson SH, Kasperbauer JL, Eckloff BW, Tombers NM, Vasmatzis G, Smith DI. Mate pair sequencing of oropharyngeal squamous cell carcinomas reveals that HPV integration occurs much less frequently than in cervical cancer. *Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology*. 2014;59:195–200.

35. Cradic KW, Murphy SJ, Drucker TM, Sikkink RA, Eberhardt NL, Neuhauser C, Vasmatzis G, Grebe SK. A simple method for gene phasing using mate pair sequencing. *BMC medical genetics*. 2014;15:19.
36. Kalari KR NA, Bhavsar JD, O'Brien DR, Davila JJ, Bockol MA, Nie J, Tang X, Baheti S, Doughty JB, et al. MAP-RSeq: Mayo Analysis Pipeline for RNA sequencing. *BMC bioinformatics*. 15:224.
37. Dobin A DC, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013; 29:15–21.
38. Liao Y, Smyth G, Shi W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic acids research*. 2013;41.
39. Muleris M, Almeida A, Gerbault-Seureau M, Malfroy B, Dutrillaux B. Identification of amplified DNA sequences in breast cancer and their organization within homogeneously staining regions. *Genes, chromosomes & cancer*. 1995;14:155–163.
40. Kallioniemi OP, Kallioniemi A, Kurisu W, Thor A, Chen LC, Smith HS, Waldman FM, Pinkel D, Gray JW. ERBB2 amplification in breast cancer analyzed by fluorescence in situ hybridization. *Proceedings of the National Academy of Sciences of the United States of America*. 1992;89:5321–5.
41. Cancer Genome Atlas N. Comprehensive molecular portraits of human breast tumours. *Nature*. 2012;490:61–70.
42. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, Speed D, Lynch AG, Samarajiwa S, Yuan Y, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*. 2012; 486:346–52.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

