

Clinical Chemistry

Manuscript Title: Next generation sequencing of circulating cell-free DNA evaluating mutations and gene amplification in metastatic breast cancer

Manuscript No: CLINCHEM/2016/261834 [R3]

Manuscript Type: Article

Date Submitted by the Author:

Complete List of Authors: Karen Page, David Guttery, Daniel Fernandez Garcia, Allison Hills, Robert Hastings, Jinli Luo, Kate Goddard, Vedia Shahin, Laura Woodley-Barker, Brenda Rosales, Raoul Charles Coombes, Justin Stebbing, and Jacqui A Shaw

Keywords: Cell-free DNA; Gene amplification; Liquid biopsy; Targeted next-generation sequencing

Dear Reviewer,

Thank you for reviewing this manuscript, please remember the following:

1. The attached manuscript is confidential and should not be circulated or shared with anyone.
2. Authors of this manuscript should never be contacted until after its publication.
3. If you have a conflict of interest regarding this work, contact the editorial office immediately.

Next generation sequencing of circulating cell-free DNA evaluating mutations and gene amplification in metastatic breast cancer

Karen Page^{1#}, David S. Guttery^{1#}, Daniel Fernandez-Garcia¹, Allison Hills², Robert K. Hastings¹, Jinli Luo¹, Kate Goddard³, Vedia Shahin³, Laura Woodley-Barker³, Brenda M. Rosales², R. Charles Coombes², Justin Stebbing², Jacqueline A. Shaw^{1*}

¹Department of Cancer Studies and Cancer Research UK Leicester Centre, University of Leicester, Leicester, LE2 7LX, UK.

²Department of Surgery and Cancer, Division of Cancer, 1st Floor ICTEM Building, Imperial College London, Du Cane Road, London, W12 0NN, UK.

³Department of Medical Oncology, Imperial College London, 1st Floor East Wing, Charing Cross Hospital, Fulham Palace Road, London, W6 8RF, UK.

#These authors contributed equally to the work

Corresponding authors: Professor Jacqui Shaw, Department of Cancer Studies, University of Leicester, Robert Kilpatrick Clinical Sciences Building, Leicester Royal Infirmary, Leicester, LE2 7LX, UK. Tel: +44 (0)1162523148; Fax: +44 (0)1162523274; E-Mail: js39@le.ac.uk

Professor Justin Stebbing, Department of Surgery and Cancer, Division of Cancer, 1st Floor ICTEM Building, Imperial College London, Du Cane Road, London, W12 0NN, UK. Tel: +44 (0)2075942793; E-Mail: j.stebbing@imperial.ac.uk

Running title: NGS of cfDNA for somatic mutations and gene amplification

Key words: Liquid biopsy, circulating free DNA, targeted next-generation sequencing.

Competing interests: The authors declare that they have no competing interests.

Acknowledgments of research support for the study: Supported by a Cancer Research UK Clinical and Translational Research Committee programme grant (C14315/A13462X), and a Breast Cancer Campaign pilot grant (2013NovSP218), the Imperial and Leicester ECMCs and the Cancer Research UK Leicester Centre.

Abstract

Background: Breast cancer tissues are heterogeneous and show diverse somatic mutations and somatic copy number alterations (CNAs). We used a novel targeted next generation sequencing (NGS) panel to examine cell-free DNA (cfDNA) to detect somatic mutations and gene amplification in women with metastatic breast cancer (MBC).

Methods: cfDNA from pre-treated patients (n=42) and 9 healthy controls were compared with matched lymphocyte DNA by NGS, using a custom 158 amplicon panel covering hot-spot mutations and CNAs in 16 genes, with further validation of results by droplet digital polymerase chain reaction.

Results: No mutations were identified in cfDNA of healthy controls, whereas exactly half the patients with MBC had at least one mutation or amplification in cfDNA (mean 2, range 1-6) across a total of 13 genes. Longitudinal follow up showed dynamic changes to mutations and gene amplification in cfDNA indicating clonal and sub-clonal response to treatment that was more dynamic than cancer antigen 15-3 (CA15-3). Interestingly, 7 patients with *ERBB2* gene amplification in their cfDNA were progressing at the time of blood sampling; of these 3 were HER2 negative at diagnosis, suggesting clonal evolution to a more aggressive phenotype. Lastly, 6 patients harbored *ESR1* mutations in cfDNA, suggesting resistance to endocrine therapy. Overall 9 of 42 patients (21%) had alterations in cfDNA that could herald a change in treatment.

Conclusion: Targeted NGS of cfDNA has potential for monitoring response to targeted therapies through both mutations and gene amplification, for analysis of dynamic tumor heterogeneity, and stratification to targeted therapy.

Introduction

Somatic mutation profiling of breast tumor tissues has identified a number of distinct breast cancer molecular subtypes (1-3) characterised by diverse somatic mutations, including single nucleotide variants (SNVs) and copy number alterations (CNAs). The two most frequently mutated genes are *TP53* and *PIK3CA*; however, a large number of other genes are less commonly mutated (4, 5). Genes that show amplification include *ERBB2*, which can be treated with anti-HER2 agents such as trastuzumab, *CCND1*, *FGFR1* and *MYC* (6).

Targeted next generation sequencing (NGS) enables detection of low-frequency somatic mutations (i.e. SNVs detected at < 5%) in heterogeneous tumor populations and in circulating cell free DNA (cfDNA), when high coverage of >5000x (7, 8) is achieved, with the potential for guidance of treatment. However, current targeted NGS approaches that focus on key driver genes (e.g. *TP53* and *PIK3CA*) do not fully capture genomic heterogeneity of breast cancer. Somatic CNA analysis has been carried out at the whole genome level, for example by single nucleotide polymorphism (SNP) 6.0 array (9) and low coverage sequencing (10), and *ERBB2* gene amplification has been investigated by real-time quantitative PCR and droplet digital PCR (ddPCR) (11). As proof of concept, in this study we evaluated targeted NGS of cfDNA for analysis of mutations and amplification in 16 genes in 42 patients with metastatic breast cancer (MBC).

Materials and Methods

Patient samples, blood processing and DNA extraction

We recruited 42 patients with radiologically-confirmed MBC (study approved by the Riverside Research Ethics Committee ref 07/Q0401/20) (Supplemental Table 1) and 9 women attending for breast screening mammography as age matched controls (study

approved by NRES: 12/LO/2019). Blood sample collection was conducted in accordance with the Declaration of Helsinki. All patients gave written informed consent prior to participation. 20 ml venous blood was collected into EDTA-containing tubes (BD Biosciences) and 3 ml of the obtained plasma processed using the Circulating Nucleic Acids kit (Qiagen) as described previously (12). DNA was isolated from 200 µl buffy coat (for germ line DNA) and breast cancer cell lines (MDA-MB-231, MCF-7, SKBR3 and ZR-75-1) as described previously (8). Cancer antigen 15-3 (CA 15-3) results were obtained from patient notes.

Targeted next generation sequencing

We designed a custom 158 amplicon panel (size range 125-175 bp) across 16 genes (Supplemental Table 2) based on previous studies (6, 9) and publically available databases (including cBioPortal (<https://www.cbioportal.org>) and ArrayExpress (<https://www.ebi.ac.uk/arrayexpress>). A minimum of 3 ng (mean 4.8 ng) cfDNA, 5 ng lymphocyte DNA and 1, 5 and 10 ng cell line DNA was used to generate libraries. Library preparation and PGM™ sequencing (Life Technologies) was performed using the Ion Ampliseq library preparation kit v2.0 as described previously (8). A maximum of 6 libraries were pooled to achieve a coverage of at least 500x per amplicon.

Detection of SNVs and CNAs

Sequencing data was accessed through the Torrent Suite v4.2.0, reads aligned against hg19 using Alignment v4.0-r77189, and variants called using the coverageAnalysis v4.0-r77897 and variantCaller v4.0-r76860, respectively. VariantCaller was configured to call on high stringency somatic variants with down sampling set to 2000 and the hotspot_min_allele_frequency set to 0.01 to detect very low frequency (< 2%) variants.

COSMIC IDs were obtained using COSMIC v72 (13). Each cfDNA sample was compared with paired lymphocyte DNA. ANNOVAR (14) was used to annotate all variants with refGene ID, functional consequence (e.g., non-synonymous), and functional predictions (using SIFT (15), Polyphen-2 (16) and MutationTaster (17)). All high confidence variant calls were reviewed manually using the Integrated Genomics Viewer (IGV) package (v2.3.25) (18).

OncoCNV (19) (v6.4) was downloaded from <https://oncocnv.curie.fr> (accessed February 2016) and installed with Samtools (0.1.19) (20) and Bedtools (2.17.0) (21) according to the authors' instructions. BAM files from 9 healthy controls were used to generate a cfDNA baseline control, which was compared to cfDNA BAM files from MBC with the OncoCNV default cghseg segmentation algorithm. Human genomic DNA (Roche) was used as the baseline for cell line DNA. GC content per target region was calculated using hg19.fasta (GRCh37) (<http://hgdownload.cse.ucsc.edu/goldenPath/hg19/bigZips/> - accessed February 2016). A panel wide profile plot was generated for each sample in addition to individual chromosome plots and tabular outputs R script processSamples.R, was modified to allow custom scale on the plots (available on request). A copy number (CN) of ≥ 3 was selected as a confident threshold for gene amplification, which was above the CN ≥ 2.5 reported previously using ddPCR (11).

Droplet digital PCR

Digital droplet PCR (ddPCR) was used to validate *TP53*, *PIK3CA* and *ESR1* mutations as described previously (8). Primers and FAM-MGB probes were designed using Primer3 (22) for *ERBB2*, *MYC*, *NOMO2*, *CCND1* and *FGFR1* (Supplemental Table 3). A mean of 5 ng cfDNA (range 1.4 ng – 8.5 ng) was pre-amplified with 12.5 μ l of TaqMan pre-amplification

mastermix (Life Technologies) and 2.5 µl of a primer pool mix at 95°C for 10 min, followed by 10 cycles of 95°C for 15 s and 60°C for 4 min, and a final incubation of 99°C for 10 min. Duplex assays were run with a *RPPH1* (labelled with VIC-MGB) reference with 5 µl 1:10 dilution of pre-amplified cfDNA, 11 µl of 2x ddPCR Supermix for probes (Bio-Rad), 1.1 µl of both target and reference assays in a final volume of 20 µl as described previously (8).

Statistical analysis

Survival analysis was performed using multiple Cox-regression as described previously (23) with each biomarker as a continuous time-dependent variable. The end of study date was selected as 01/12/2015. Comparison of total cfDNA concentrations (copies/ml) in patients with mutation or amplification compared to those without was by t-test.

Results

We detected the expected COSMIC mutations (*PIK3CA* p.E545K in MCF-7 and ZR75-1, and *TP53* p.R175H in SKBR3) and gene amplifications (*MYC* in MCF-7, ZR-75-1 and SKBR3, and *ERBB2* in SKBR3) in cell line DNA with a mean read depth of >2000x per sample (range 2177x – 2659x) (Supplemental Tables 4 and 5). Dilution of the SKBR3 cell line into human genomic DNA (Roche) showed a limit of detection of 1% for *TP53* p.R175H mutation and 10% for *MYC* and *ERBB2* gene amplification, respectively (Supplemental Table 6). There was good agreement between results by NGS by ddPCR (Supplemental Figure 1).

Detection of mutations and gene amplification in cfDNA

We sequenced paired cfDNA and lymphocytes in 42 patients with metastatic breast cancer (40 ERα positive, 2 ERα negative, median age 55 y; range 25 – 85 y; Supplemental Table 1) and 9 healthy female controls (median age 55 y; range, 37 – 66 y). Mean coverage was over

1400x (range 172x – 2790x), with greater than 360K mean mapped reads (range 48,132 – 626,148) per sample. On average, 89% of the 158 targeted regions were covered at >500x. Only 3 amplicons, which all targeted somatic mutations, had a mean coverage of <100x and were excluded from analysis. Only 2 of 69 cfDNA samples had a mean coverage below 500x, both of which had CNA detected. No mutations or CNAs were detected in cfDNA of the 9 healthy controls (Supplemental Figure 2). One cfDNA leukocyte pair had a low level gain in *NOMO2* (CN = 2.5) within an interval that can show duplication (24).

All variants reported were detected by variant caller, with one exception, a low-frequency *ESR1* mutation in a serial sample from Patient 4. Exactly half the patients had 1 or more gene specific mutations and amplifications detected in cfDNA (mean = 2, range 1-6, Supplemental Table 7, Figure 1A). The top 3 mutated genes in cfDNA were *ESR1*, *TP53* and *PIK3CA* (Figure 1B). Nine patients had 2 or more mutations detected and 10 patients had amplification in one or more genes (Supplemental Table 7). As validation, we compared NGS results with a smaller amplicon panel covering hotspot mutations in *PIK3CA*, *TP53*, *ESR1*, *FGFR1* and *FGFR2* (8) in 25 of the 42 patients. The same 9 mutations were identified across 6 patient samples and the VAFs were highly correlated ($\rho = 0.9715$; $P < 0.0001$; Figure 1C). Mutation in the *ESR1* gene was significantly associated with poorer overall survival (Figure 1D; hazard ratio (HR) 25.61; 95% CI 4.58 – 143.18; $P < 0.0001$, log-rank test), supporting a previous study (25).

Nine patients had a HER2-positive primary tumor at diagnosis. Of these, 6 were progressing on anti-HER2 agents at the time of blood sampling and 4 had *ERBB2* gene amplification in cfDNA, whereas 3 patients who were responding to their treatment had no evidence of *ERBB2* gene amplification in cfDNA (Table 1). One patient had HER2-positive metastatic

biopsy and was stable on paclitaxel and herceptin at the time of blood sampling and was also negative for *ERBB2* gene amplification in cfDNA. Lastly, 3 patients with a HER2 negative primary tumor had acquired *ERBB2* gene amplification in cfDNA and all 3 were progressing at the time of the blood sample (Table 1).

Dynamic changes in cfDNA on longitudinal follow-up.

We performed serial monitoring of alterations in cfDNA in 9 patients (Table 2, Figure 2, Figure 3 and Supplemental Figure 3).

Patient 1 (estrogen receptor (ER) pos, progesterone receptor (PR) pos, HER2 pos) was diagnosed in 2007 at age 47 y and developed metastases in November 2011. She was being treated with capecitabine and lapatinib in November 2012 at the time of the first blood sample, which had a *PIK3CA* mutation and gene amplification of *ERBB2* and *NOMO2* in cfDNA. She initially responded to therapy but relapsed in April 2013 just prior to the second blood sample. The *PIK3CA* mutation and *ERBB2* gene amplification persisted, however *NOMO2* amplification was lost, suggesting a clonal shift. Total cfDNA concentrations were rising but CA15-3 concentrations were falling with progression (Figure 2A).

Patient 2 (ER pos, PR pos, HER2 neg) was diagnosed with breast cancer in December 2006 at age 38 y and developed metastases in April 2012. She was progressing on goserelin and capecitabine at the time of first blood sample in May 2013, which had amplified *ERBB2*, *TP53* mutation (VAF = 6.5%) and a sub-clonal *PIK3CA* mutation (VAF = 2.5%) in cfDNA, suggesting resistant disease. Her disease worsened and she was switched to paclitaxel in April 2014. At the time of second blood sample in February 2015 she was responding to paclitaxel and the mutations resolved in cfDNA, however, total cfDNA and CA15-3 concentrations were rising and did not reflect response to treatment (Figure 2B). Her disease progressed again in Nov 2015.

Patient 3 (ER pos, PR pos, HER2 neg) was diagnosed at age 48 y in 1986 and developed recurrence in her regional nodes in March 2007. She was treated with anastrozole and in April 2008 was switched to monthly fulvestrant. CT scan showed partial response by RECIST at the time of the first blood sample in September 2012. A *PIK3CA* mutation was detected in cfDNA (VAF 5%) but was undetected in the second sample 14 months later suggesting a clonal response (Figure 2C), although total cfDNA concentrations were rising. Since then her disease has remained stable. CA15-3 concentrations were normal at both time points.

Patient 4 (ER pos, PR pos, HER2 neg) was diagnosed in 2004 at age 70 y and developed metastases in December 2010. At the time of the first blood sample in November 2014, she was stable on everolimus, exemestane, goserelin and zoledronic acid. Mutations were detected in cfDNA in *PIK3CA* (VAF 11.7%) and *ESR1* (VAF 10.9%). The *PIK3CA* mutation decreased to 1% two months later in January 2015. The *ESR1* mutation was undetected in variant caller, but was found at a frequency of 0.5% when manually inspected in IGV. Moreover, this was confirmed by ddPCR (data not shown). The co-reduction in these 2 mutations suggests a clonal response (Figure 2D) although CA15-3 concentrations were high at both time points. A solitary new liver metastasis was detected one month later in February 2015 and she was switched to letrozole. Her disease is currently stable.

Patient 5 (ER pos, PR pos, HER2 pos) was diagnosed in April 2012 at age 46 y and developed metastases in October 2012. Blood was collected in September and December 2014, when she was on herceptin, tamoxifen and zoledronic acid, and CT scan showed partial response by RECIST. No alterations were detected in cfDNA at either time point and CA15-3 concentrations were normal (Figure 2E), her disease remains stable.

Patient 6 (ER pos, PR pos, HER2 neg) was diagnosed with metastatic breast cancer at age 23 y in early 2009 and relapsed in 2011 on combined endocrine therapy (fulvestrant, goserelin,

letrozole and zoledronic acid). Four blood samples were collected over a 2-year period between February 2013 and March 2015. She was being treated for brain metastases at the first sample time in February 2013, but since then and up till present is stable on combined endocrine therapy. No mutations were detected in cfDNA from the first 2 blood samples and CA15-3 concentrations were normal. A low frequency *PIK3CA* mutation was detected in the third blood sample in July 2014, which increased and was accompanied by a second sub-clonal mutation in *TP53* and rising CA15-3 in the fourth sample in March 2015 (Figure 2F). Although clinically stable the presence of mutations, rising concentrations of circulating tumor DNA and rising CA15-3 suggest emergence of endocrine resistant disease.

Patient 7 (ER pos, HER2 neg) was diagnosed with breast cancer in 2002 at age 61 y. She developed metastases in 2009 and was treated with sequential tamoxifen, anastrozole and exemestane and fulvestrant. CT scan showed complete response by RECIST. At the time of the first blood sample in December 2013, she had progressed and was being treated with exemestane and everolimus. No alterations were detected in cfDNA but CA15-3 was high (314 U/ml). One month later her disease was worsening, CA15-3 rising and a single *LMTK3* mutation (VAF = 3.5%) was detected in cfDNA reflecting disease progression (Figure 2G) and she died shortly after. Based on this information an inhibitor of LMTK3 could have been used had one been available.

Patient 8 (ER pos, PR pos, HER2 neg) was diagnosed in 2004 at age 42 y and developed metastases in Sept 2012. At the time of the first sample in March 2014, she was stable on tamoxifen, but then progressively worsened and started denosumab then anastrozole. Despite analysing >3500 COSMIC mutations no alterations were detected or acquired in 3 sequential cfDNA samples during 9 months monitoring on therapy, although total cfDNA concentrations and CA15-3 were both rising. Amplification of *MYC* and *ERBB2* genes was detected in the fourth cfDNA sample at the time of pulmonary progression (Figure 3A),

accompanied by a marked rise in cfDNA concentrations and CA15-3. She had been on capecitabine for one month and CT scan showed partial response by RECIST at the time fifth blood sample, *MYC* and *ERBB2* gene amplification were undetected and CA15-3 levels were falling. As of June 2016 her disease remains stable.

Patient 9 (ER pos, PR pos, HER2 neg) was diagnosed in November 2011 at age 40 y and developed metastases soon after in January 2012. She was initially treated with tamoxifen, which was switched to zoledronic acid, letrozole and denosumab in November 2012. She responded well and was in remission at time of the first two samples in May and August 2014, which showed no alterations in cfDNA but CA15-3 concentrations were rising and remained elevated (Figure 3B). Her disease worsened over the next 6 months and paclitaxel was started. Her disease was worsening and the third blood sample in December 2014 had an *ESR1* (VAF 28.6%) and sub-clonal *PIK3CA* (VAF 17.8%) mutation and *CCND1* amplification in cfDNA. She started epirubicin in April 2015 was responding well at the time of sample 4 and sample 5 in June 2015 when CT scan showed stable disease by RECIST and the mutations resolved in cfDNA. She subsequently developed resistance and died in November 2015.

Discussion

To our knowledge this is the first targeted NGS study of cfDNA to evaluate both somatic mutations and gene specific amplification in patients with MBC. Exactly half of the 42 patients with MBC had at least one mutation or amplification in cfDNA, whereas no alterations were detected in 9 healthy female controls. The variant allele frequency (VAF) ranged from 1.1 - 35.1% for point mutations in cfDNA, whereas the limit of detection for CNA was 10% cell line DNA, similar to a previous low read depth whole genome sequencing study (10), whereas a recent study focusing on SNPs detected somatic CNA with mean allelic

imbalances as low as 0.5% (26). With respect to the analytical sensitivity of the assay, supporting that this panel can identify somatic variants with a VAF of $\geq 1\%$ for SNVs and 5% for CNA assuming sufficient cfDNA is available for analysis. A number of other methods including CAPP-Seq (27), Safe-Seq (28) and BEAMing (29) have reported greater analytical sensitivity for mutation detection with variant allele fractions of $<0.1\%$. While these approaches hold great promise, a major assumption is that ctDNA is either rare or absent in the healthy, cancer-free population. Demonstrating that a ctDNA marker has a diagnostic specificity close to 100% would also be important (30). For clinical translation, detection of low-frequency variants must first be validated for CLIA/CAP/ISO-based precision testing. Of importance, each assay would need to be quality controlled using well-controlled reference standards (for example Horizon Discovery Quantitative Multiplex Reference Standards) to establish lower detection limits and reproducibility to detect variants at the lower limit.

Nine of the patients studied had more than one mutation detected in cfDNA, and in 7 patients these mutations were at different frequencies, indicating clonal differences in the origin of circulating tumor derived DNA. Importantly, 6 patients had an *ESR1* mutation detected in cfDNA while on AI/endocrine therapy; this information could be used clinically to herald a change to chemotherapy. Ten patients had amplification in cfDNA in one or more genes, 3 patients had amplification only and 7 patients had both amplification and mutations. These data support heterogeneity of somatic alterations in breast cancer, with some characterised by mutations some by CNA and some by both. Ten patients had either a HER2-positive primary tumor or metastatic biopsy. Of these patients, 6 individuals had progressive disease at the time of blood sampling and 4 individuals had *ERBB2* gene amplification in cfDNA, whereas *ERBB2* gene amplification was undetected in 4 patients who were responding to an anti-HER2 agent. Importantly, 3 patients with a HER2 negative primary tumor had acquired

ERBB2 gene amplification in cfDNA suggesting clonal evolution to a more aggressive phenotype (11, 31). This information could be used clinically to change to an anti-HER2 therapy.

We performed serial monitoring in nine patients. Results demonstrate that cfDNA profiling of mutations and amplifications could provide useful data in terms of tumor heterogeneity, clonal evolution and response to treatment. Concentrations of circulating tumor DNA generally tracked with patient disease status, whereas when CA15-3 concentrations were high and generally remained increased. This supports the previous study by Dawson et al. (23), who suggested that ctDNA to be a highly diagnostically sensitive biomarker of MBC. Interestingly, in 3 patients we saw rising total cfDNA concentrations at a time when mutations in *PIK3CA* and/or *TP53* and *ESR1* either resolved or did not increase in circulating tumour DNA. A possible explanation for this is that another clone was shedding DNA into the blood that was not characterised by any of the alterations detectable by our NGS panel. One approach to interrogate this would be to perform whole exome analysis of plasma cfDNA. Although we sequenced ~3500 COSMIC mutations and surveyed for amplification in 16 genes, the majority of cfDNA samples had less than 5 alterations detected, as has been shown in other studies using NGS mutation hot spot panels (32). This is likely due to genomic heterogeneity as there are a large number of genes infrequently mutated in breast cancers (33, 34). However, the targeted NGS approach used here has potential clinical utility, where for example, emergence of *ERBB2* amplification in plasma cfDNA could signal a switch to an anti-HER2 therapy, and emergence of *ESR1* mutations could indicate a switch away from endocrine therapy to standard chemotherapy. Overall, 9 patients (21%) could have been offered an alternative therapy if blood based monitoring was routine in the clinic. Our

data support further investigation of the NGS approach focusing on actionable mutations and gene amplification for monitoring treatment response and clonal dynamics in MBC.

References

1. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. *Nature* 2013;500:415-21.
2. Bignell GR, Greenman CD, Davies H, Butler AP, Edkins S, Andrews JM, et al. Signatures of mutation and selection in the cancer genome. *Nature* 2010;463:893-8.
3. Leary RJ, Lin JC, Cummins J, Boca S, Wood LD, Parsons DW, et al. Integrated analysis of homozygous deletions, focal amplifications, and sequence alterations in breast and colorectal cancers. *Proc Natl Acad Sci U S A* 2008;105:16224-9.
4. Murtaza M, Dawson SJ, Tsui DW, Gale D, Forsheew T, Piskorz AM, et al. Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. *Nature* 2013;497:108-12.
5. 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.
6. Stephens PJ, Tarpey PS, Davies H, Van Loo P, Greenman C, Wedge DC, et al. The landscape of cancer genes and mutational processes in breast cancer. *Nature* 2012;486:400-4.
7. Rehm HL. Disease-targeted sequencing: A cornerstone in the clinic. *Nat Rev Genet* 2013;14:295-300.
8. Guttery DS, Page K, Hills A, Woodley L, Marchese SD, Rghebi B, et al. Noninvasive detection of activating estrogen receptor 1 (esr1) mutations in estrogen receptor-positive metastatic breast cancer. *Clin Chem* 2015;61:974-82.

9. Shaw JA, Page K, Blighe K, Hava N, Guttery D, Ward B, et al. Genomic analysis of circulating cell-free DNA infers breast cancer dormancy. *Genome Res* 2012;22:220-31.
10. Belic J, Koch M, Ulz P, Auer M, Gerhalter T, Mohan S, et al. Rapid identification of plasma DNA samples with increased ctDNA levels by a modified fast-seq approach. *Clin Chem* 2015;61:838-49.
11. Gevensleben H, Garcia-Murillas I, Graeser MK, Schiavon G, Osin P, Parton M, et al. Noninvasive detection of her2 amplification with plasma DNA digital pcr. *Clin Cancer Res* 2013;19:3276-84.
12. Page K, Guttery DS, Zahra N, Primrose L, Elshaw SR, Pringle JH, et al. Influence of plasma processing on recovery and analysis of circulating nucleic acids. *PLoS One* 2013;8:e77963.
13. Hancox RA, Allen MD, Holliday DL, Edwards DR, Pennington CJ, Guttery DS, et al. Tumour-associated tenascin-c isoforms promote breast cancer cell invasion and growth by matrix metalloproteinase-dependent and independent mechanisms. *Breast Cancer Res* 2009;11:R24.
14. Wang K, Li M, Hakonarson H. Annovar: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010;38:e164.
15. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the sift algorithm. *Nat Protoc* 2009;4:1073-81.
16. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248-9.
17. Schwarz JM, Cooper DN, Schuelke M, Seelow D. Mutationtaster2: Mutation prediction for the deep-sequencing age. *Nat Methods* 2014;11:361-2.

18. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. *Nat Biotechnol* 2011;29:24-6.
19. Boeva V, Popova T, Lienard M, Toffoli S, Kamal M, Le Tourneau C, et al. Multi-factor data normalization enables the detection of copy number aberrations in amplicon sequencing data. *Bioinformatics* 2014;30:3443-50.
20. Li H. A statistical framework for snp calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 2011;27:2987-93.
21. Quinlan AR, Hall IM. Bedtools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010;26:841-2.
22. Primer3 - <http://primer3.Ut.Ee/>. <http://primer3.ut.ee> (Accessed January 2014).
23. Dawson SJ, Tsui DW, 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.
24. Park RW, Kim TM, Kasif S, Park PJ. Identification of rare germline copy number variations over-represented in five human cancer types. *Mol Cancer* 2015;14:25.
25. Schiavon G, Hrebien S, Garcia-Murillas I, Cutts RJ, Pearson A, Tarazona N, et al. Analysis of *esr1* mutation in circulating tumor DNA demonstrates evolution during therapy for metastatic breast cancer. *Sci Transl Med* 2015;7:313ra182.
26. Kirkizlar E, Zimmermann B, Constantin T, Swenerton R, Hoang B, Wayham N, et al. Detection of clonal and subclonal copy-number variants in cell-free DNA from patients with breast cancer using a massively multiplexed pcr methodology. *Transl Oncol* 2015;8:407-16.

27. Newman AM, Bratman SV, To J, Wynne JF, Eclov NC, Modlin LA, et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat Med* 2014;20:548-54.
28. Kinde I, Wu J, Papadopoulos N, Kinzler KW, Vogelstein B. Detection and quantification of rare mutations with massively parallel sequencing. *Proc Natl Acad Sci U S A* 2011;108:9530-5.
29. Diehl F, Li M, He Y, Kinzler KW, Vogelstein B, Dressman D. Beaming: Single-molecule pcr on microparticles in water-in-oil emulsions. *Nat Methods* 2006;3:551-9.
30. Wentzensen N, Wacholder S. From differences in means between cases and controls to risk stratification: A business plan for biomarker development. *Cancer Discov* 2013;3:148-57.
31. Page K, Hava N, Ward B, Brown J, Guttery DS, Ruangpratheep C, et al. Detection of her2 amplification in circulating free DNA in patients with breast cancer. *Br J Cancer* 2011;104:1342-8.
32. Frenel JS, Carreira S, Goodall J, Roda D, Perez-Lopez R, Tunariu N, et al. Serial next-generation sequencing of circulating cell-free DNA evaluating tumor clone response to molecularly targeted drug administration. *Clin Cancer Res* 2015;21:4586-96.
33. Ciriello G, Gatza ML, Beck AH, Wilkerson MD, Rhie SK, Pastore A, et al. Comprehensive molecular portraits of invasive lobular breast cancer. *Cell* 2015;163:506-19.
34. Park NI, Rogan PK, Tarnowski HE, Knoll JH. Structural and genic characterization of stable genomic regions in breast cancer: Relevance to chemotherapy. *Mol Oncol* 2012;6:347-59.

Acknowledgements: We thank the Leicester and Imperial Experimental Cancer Medicine Centres (ECMC) and the Imperial College Tissue Bank for supporting patient recruitment and sample collection, and the Cancer Research UK Leicester Centre for bioinformatics support. **Author contributions:** J.A.S., J.S. and R.C.C. developed the study. K.P., D.F.G., J.S., R.C.C., J.A.S and D.S.G. designed and/or performed all laboratory experiments; L.W., A.H., B.R., K.G., and V.S. consented and collected patient blood samples; K.P., D.F.G., R.K.H., J. L., R.C.C., J.A.S. and D.S.G. analysed the data; J.A.S., D.S.G., K.P., D.F.G., R.C.C., and J.S., wrote the paper.

Figure Legends

Figure 1: Mutations and gene amplification in cfDNA of 42 patients with MBC. (A) Number of patient samples with SNVs, CNAs or both SNVs and CNAs detected in cfDNA by NGS. (B) Top 3 mutated genes and somatic mutations. (C) Correlation between mutation frequencies detected by 2 NGS panels. (D) *ESR1* mutation was significantly associated with poorer overall survival ((HR) 25.61; 95% CI 4.58 – 143.18; $P < 0.0001$, log-rank test).

Figure 2: Serial monitoring of alterations in cfDNA from 7 patients with MBC. Graphs show total cfDNA concentrations (copies/ml), ctDNA (copies/ml) and CA15-3 (U/ml). (A) and (B) also show OncoCNV analysis of CNAs detected in cfDNA presented as log2 normalized read count relative to healthy controls. (A – G) Patients 1 – 7 (see Table 2 for details). Treatments details are given above each graph.

Figure 3: Monitoring changes in gene amplification in cfDNA with treatment response.

Top graphs show total cfDNA concentrations (copies/ml), ctDNA (copies/ml) and CA15-3 (U/ml). Lower graphs show gene-specific copy number detected by OncoCNV. Both patients had gene-specific amplification detected in cfDNA when disease was progressing, and CA15-3 and cfDNA concentrations were rising. Both patients showed a response with loss of gene amplification and mutations in cfDNA on switching from AI therapy to chemotherapy. (A) Patient 8. (B) Patient 9 (see Table 2 for details).

Table 1. Patients with HER2 positive primary tumor and/or *ERBB2* gene amplification in cfDNA

| Patient ID | Time between diagnosis and sample (months) | HER2 status of primary tumor | Gene alteration in cfDNA | | Treatment at time of blood sample | Disease status at time of blood sample |
|------------|--|------------------------------|--------------------------------|--|--|--|
| | | | Amplification | Mutation | | |
| 10 | 16 | Positive | <i>ERBB2</i> | - | lapatinib and capecitabine | Progressing |
| 40 | 56 | Positive | <i>ERBB2</i> | - | vinorelbine and herceptin | Progressing |
| 13 | 17 | Positive | <i>ERBB2</i> and <i>NOMO2</i> | <i>TP53</i> (p.L251S) | herceptin | Progressing |
| 1 | 65 | Positive | <i>ERBB2</i> and <i>NOMO2</i> | <i>PIK3CA</i> (p.H1047R) | lapatinib and capecitabine | Progressing |
| 20 | 49 | Positive | - | - | lapatinib and capecitabine | Progressing |
| 2 | 77 | Negative | <i>ERBB2</i> | <i>PIK3CA</i> (p.H1047R); <i>TP53</i> (p.R248Q) | goserelin and capecitabine | Progressing |
| 32 | 17 | Negative | <i>ERBB2</i> and <i>NOMO2</i> | <i>PIK3CA</i> (p.H1047R); <i>TP53</i> (p. R248Q) | anastrozole | Progressing |
| 8 | 36 | Negative | <i>ERBB2</i> and <i>MYC</i> ** | - | tamoxifen, denosumab, anastrozole | Progressing |
| 15 | 56 | Positive | - | - | exemestane | Progressing |
| 36 | 42 | Positive | - | - | herceptin | Stable |
| 5 | 27 | Positive | - | - | herceptin, tamoxifen and zoledronic acid | Stable |
| 21 | 49 | Positive | - | <i>PIK3CA</i> (p.H1047R) | eribulin and herceptin | Stable |
| 26 | 79 | Negative* | - | <i>NOMO2</i> (p.R418H) | paclitaxel and herceptin | Stable |

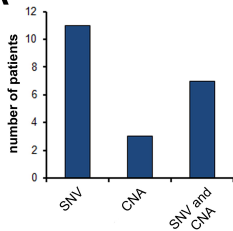
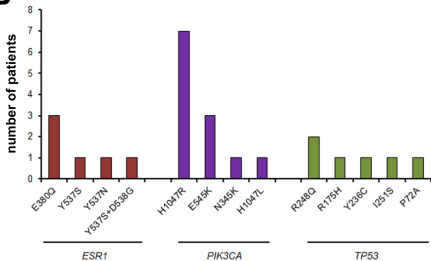
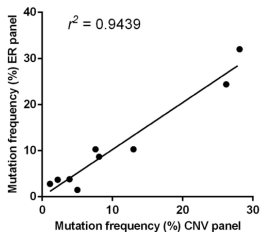
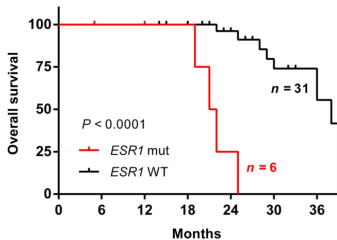
*HER2 negative primary tumor, positive metastatic biopsy, **detected in serial sample

Table 2. SNV/CNA tracking in nine metastatic breast cancer patients by NGS.

| Patient ID | Study Sample (months) | Age | ER | PR | Her2 | CA15-3 (U/ml) | cfDNA genomic copies/ml | SNV (AA mutation) and VAF | | | | COSMIC ID | Gene amplification |
|------------|-----------------------|-----|-----|-----|------|---------------|-------------------------|---------------------------|----------|---------|----------|-------------|-------------------------------|
| | | | | | | | | Gene | Mutation | VAF (%) | Coverage | | |
| 1 | Baseline | 52 | pos | pos | pos | 602 | 2,076 | <i>PIK3CA</i> | p.H1047R | 7.8 | 1988 | COSM775 | <i>NOMO2</i> and <i>ERBB2</i> |
| | 5 | | | | | 376 | 2,727 | <i>PIK3CA</i> | p.H1047R | 4.9 | 1100 | COSM775 | <i>ERBB2</i> |
| 2 | Baseline | 45 | pos | pos | neg | 25 | 4,697 | <i>PIK3CA</i> | p.H1047R | 2.5 | 1260 | COSM775 | <i>ERBB2</i> |
| | | | | | | | | <i>TP53</i> | p.R248Q | 6.5 | 876 | COSM10662 | |
| | 20 | | | | | 131 | 5,515 | | | | | | - |
| 3 | Baseline | 74 | pos | pos | neg | 29 | 3,545 | <i>PIK3CA</i> | p.H1047R | 5 | 2145 | COSM775 | - |
| | 14 | | | | | 24 | 5,742 | | - | | | | - |
| 4 | Baseline | 80 | pos | pos | neg | 150 | 13,041 | <i>PIK3CA</i> | p.N345K | 11.7 | 1357 | COSM36200 | - |
| | | | | | | | | <i>ESR1</i> | p.Y537N | 10.9 | 457 | COSM1074635 | |
| | 2 | | | | | 194 | 3,655 | <i>PIK3CA</i> | p.N345K | 1 | 1660 | COSM754 | - |
| | | | | | | | | <i>ESR1</i> | p.Y537N | 0.5 | 528 | COSM1074635 | |
| 5 | Baseline | 49 | pos | pos | pos | 23 | 12,621 | | | | | | - |
| | 2 | | | | | 23 | 13,970 | | | | | | - |
| 6 | Baseline | 27 | pos | pos | neg | 23 | 4,455 | | | | | | - |
| | 11 | | | | | 23 | 9,212 | | | | | | - |
| | 17 | | | | | 25 | 2,115 | <i>PIK3CA</i> | p.H1047R | 1 | 1181 | COSM775 | - |
| | 24 | | | | | 56 | 7,318 | <i>PIK3CA</i> | p.H1047R | 10.1 | 1738 | COSM775 | - |
| | | | | | | | | <i>TP53</i> | p.E271G | 6.2 | 64 | COSM43879 | |
| 7 | Baseline | 72 | pos | NA | neg | 314 | 6,409 | | - | | | | - |
| | 1 | | | | | 395 | 31,364 | <i>LMTK3</i> | p.E1251V | 3.5 | 1995 | | - |
| 8 | Baseline | 51 | pos | pos | neg | 20 | 1,515 | | | | | | - |
| | 7 | | | | | 46 | 2,621 | | | | | | - |
| | 9 | | | | | 221 | 4,864 | | | | | | - |

| | | | | | | | | | | | | |
|---|----------|----|-----|-----|-----|---------|-------------|-----------------|-----------|----------|-------------------|---------------|
| | 10 | | | | 567 | 11,727 | | | | | | ERBB2 and MYC |
| | 13 | | | | 252 | 9,818 | | | | | | - |
| 9 | Baseline | 43 | pos | pos | neg | 74 | 2,197 | - | | | | - |
| | 3 | | | | 110 | 4,803 | | | | | | - |
| | 6 | | | | 168 | 504,545 | PIK3CA ESR1 | p.E545K p.D538G | 17.8 28.6 | 1159 353 | COSM763 COSM94250 | CCND1 |
| | 12 | | | | 147 | 96,970 | | | | | | - |
| | 13 | | | | 138 | 86,667 | | | | | | - |

Age is the age at the time of the first blood sample.

A**B****C****D**

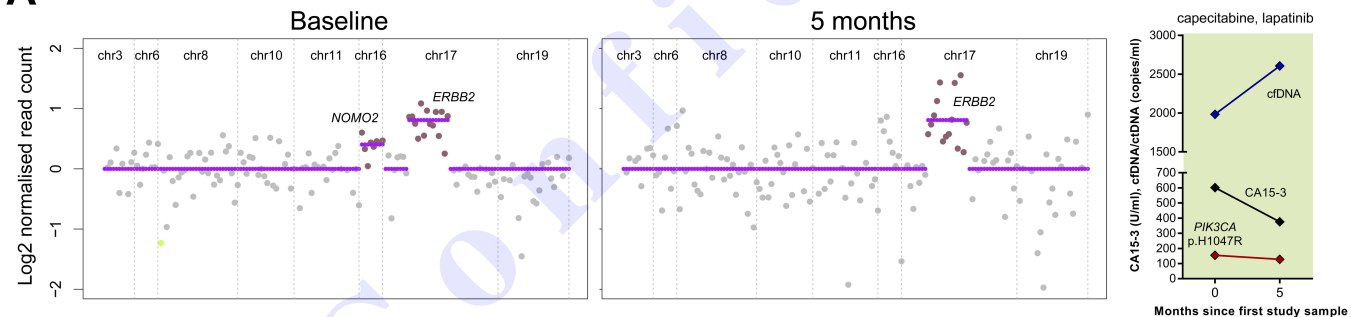
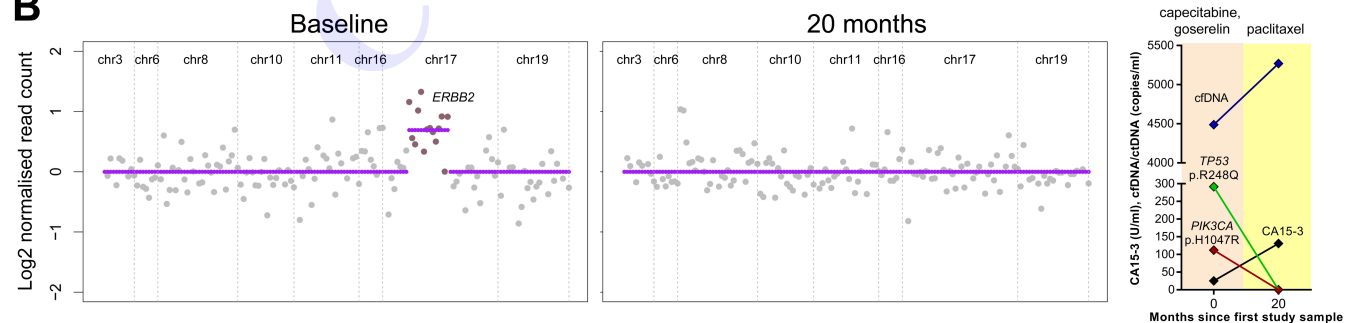
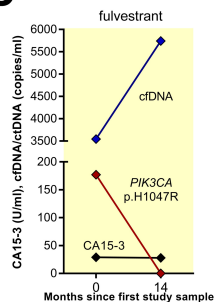
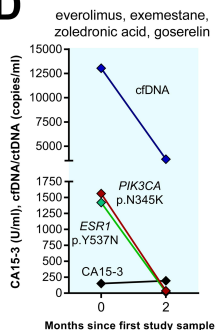
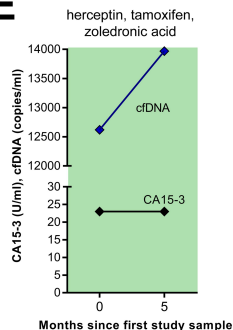
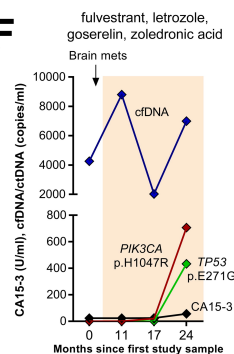
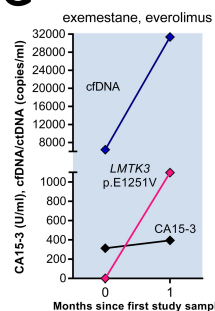
A**B****C****D****E****F****G**

Figure 3

