Karolinska Institutet

This is the peer reviwed version of the following article *Breast Cancer Res Treat. 2015 Jun;151(2):319-24.. doi: 10.1007/s10549-015-3403-6. Epub 2015 Apr 29.* which has been published in final form at http://dx.doi.org/10.1007/s10549-015-3403-6

This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

Exome sequencing of contralateral breast cancer identifies metastatic disease. Breast Cancer Res Treat. 2015 Jun;151(2):319-24. doi: 10.1007/s10549-015-3403-6.

Klevebring, Daniel; Lindberg, Johan; Rockberg, Julia; Hilliges, Camilla; Hall, Per; Sandberg, Maria; Czene, Kamila..

Doi: [10.1007/s10549-015-3403-6](http://dx.doi.org/10.1007/s10549-015-3403-6)

Access to the published version may require subscription.
Published with permission from: **Springer Verlag**

Title

Exome sequencing of contralateral breast cancer identifies metastatic disease

Authors

Daniel Klevebring[1,2,*], Johan Lindberg[1,2], Julia Rockberg[1], Camilla Hilliges[3,4], Per Hall[1], Maria Sandberg[5], Kamila Czene[1]

Affiliations

[1]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden.

[2]Science For Life Laboratory, Stockholm, Sweden.

[3]Department of Oncology-Pathology, Karolinska Institutet, Stockholm, Sweden.

[4]Department of Clinical Pathology, Karolinska University Hospital, Stockholm, Sweden

[5]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

[*]Corresponding author

Email addresses

DK: daniel.klevebring@ki.se

JL: johan.lindberg@ki.se

JR: julia@rockberg.se

CH: camilla@hilliges.com

PH: per.hall@ki.se

MS: maria.sandberg@ki.se

KC: kamila.czene@ki.se

Abstract

Background

Women with contralateral breast cancer (CBC) have significantly worse prognosis compared to women with unilateral cancer. A possible explanation of the poor prognosis of patients with CBC is that in a subset of patients, the second cancer is

not a new primary tumor, but a metastasis of the first cancer that has potentially obtained aggressive characteristics through selection of treatment. Exome and whole-genome sequencing of solid tumors has previously been used to investigate the clonal relationship between primary tumors and metastases in several diseases.

Patients and methods

In order to assess the relationship between the first and the second cancer, we performed exome sequencing to identify somatic mutations in both first and second cancers compared paired normal tissue of 25 patients with metachronous CBC.

Results

For three patients, we identified shared somatic mutations indicating a common clonal origin thereby demonstrating that the second tumor is a metastasis of the first cancer, rather than a new primary cancer. Accordingly, these patients all developed distant metastasis within 3 years of the second diagnosis, compared with 7 out of 22 patients with non-shared somatic profiles.

Conclusions

Genomic profiling of both tumors help the clinicians distinguish between true contralateral breast cancers and subsequent metastases.

Keywords

Breast cancer, Contralateral, Bilateral, recurrence, metastasis

Introduction

Breast cancer is one of the world's most common forms of cancer, accounting for approximately 1.4 million new cases annually worldwide[1]. Women with breast cancer have, 20 years after their initial diagnosis, 10-15% cumulative incidence of second primary breast cancer in the opposite breast; contralateral breast cancer (CBC) [4,5]. We and others have previously shown that women with CBC have significantly worse prognosis compared to women with unilateral cancer[2, 3]. It is likely that for a subset of the patients, the second cancer is not a new primary tumor, but instead a metastasis of the first cancer, which could, in part, explain the poor prognosis of these patients. Several earlier attempts have been made to use molecular methods to investigate this hypothesis; Janschek *et al*. compared mutations in the p53-gene[4], Imyaitov *et al*. investigated the allelic imbalance profiles[5] and Brommesson *et al*. and Teixeira *et al.* used comparative genomic hybridization[6, 7].

Exome sequencing has proved useful to identify somatic mutations and investigate the clonal relationship between primary tumors and various forms of recurrences in a wide range of cancers, including breast cancer[8-11]. Taken together, the results from exome sequencing in the field of oncology show a large genetic diversity between tumors from different patients, with only few positions in the genome being mutated in 10% or more of cancer cases. When combined, mutations across patients show a limited set of gene categories that have oncogenic or tumor suppressive capacity. However, exome sequencing it has not previously been applied to the case of contralateral breast cancer.

The aim of this study was to use exome sequencing to determine whether the second tumor was a metastasis of the first. As the treatment and prognosis for patients with CBC differs from that of patients with metastatic breast cancer, a technique for determining the origin of the second cancer would be of great clinical importance.

Methods

Study population

From the Stockholm Breast Cancer Registry, we identified women who were diagnosed with primary breast cancer with a time between the two diagnoses of 6 months to 5 years. Based on availability of archived material, we selected a subset of 25 patients for whom we collected formalin fixed paraffin embedded tissue (FFPE) (total 50 tumors), medical record information (on adjuvant treatment for the two cancers and ER-status) and follow-up information from Stockholm Breast Cancer Register (date of distant metastasis and date of death). Clinical characteristics of the selected patients are available in Table 1.

Tissue collection and Laser Microdissection

In order to identify somatic variants from private or uncommon germline variants, a paired normal (non-cancerous) sample is required for each patient. FFPE tissue for these tumors was collected, sectioned and stained with haematoxylin and eosin (HE) and reviewed by a pathologist to identify cancer cells and distinguish them from non-cancerous cells. In order to increase the tumor cell content of the tumor samples as well as to get a normal germline control for downstream sequencing, we performed laser microdissection on HE-stained 20-μm sections. Areas of tissue containing non-cancerous tissue but a high fraction of nucleated cells were sorted to one tube, and cancerous tissue was sorted to a separate tube.

DNA extractions, library preparation and sequence capture

DNA was extracted from the microdissected tissue using QiaAmp DNA Micro (QiaGen) as instructed by the manufacturer. Quantification of the total amount of extracted DNA was performed using Qubit DS DNA HS kit (Life Technologies). Sequencing libraries were prepared by tagmentation using Nextera DNA Sample Prep Kit (Epicentre) using the HMW buffer and barcode option as instructed by the manufacturer. Libraries were quantified using Qubit (as above), pooled in equimolar amounts and subjected to exome capture using EZ Exome version 1.0 (Nimblegen)

using blocking adapters designed to match the Nextera library constructs otherwise performed as instructed by the manufacturer. Libraries were diluted for sequencing on HiSeq2000 (Illumina) with the addition of Nextera-specific sequencing primers as instructed by Epicentre.

Raw sequence data processing and somatic variant calling

Raw paired-end reads were adapter-trimmed and merged with SeqPrep [12] with a maximum merged base quality of 50 and adapter sequences CTGTCTCTTATACA and CTGTCTCTTATACA (-A and -B flags, respectively). Merged reads and unmerged pairs were aligned to the human reference genome GRCh37 using bwa v0.6.2[13] and sorted and indexed with samtools v0.1.18[14]. Amplification duplicates were removed with Picard MarkDuplicates v1.85[15] after which Base Quality Score Recalibration and Local Realignment was performed using GATK v2.4.7[16]. Somatic mutations in the target intervals +/- 200 base pairs were identified using MuTect v1.1.4[17]. Any variant present in dbSNP v137 was removed from further analysis. In order to improve the sensitivity to detect shared variants, we used the total set of variants detected in tumors 1 and 2 for each patient, and calculated the allele fraction of these variants in both tumors. This methodology help detecting variants with low allele fraction, that are present in both tumors but might otherwise be detected in only one of the tumors.

Somatic mutations that have arisen early in the development of the first tumor (both those that confer an increase in proliferation; driver mutations, and those that do not; passenger mutations) will be present in most or all tumor cells of that tumor. If a cell from the first tumor seeds a metastasis, the mutations present in that cell (both passengers and drivers) will be present in all subsequent tumor cells of the metastasis. Therefore we only used clonal mutations to infer relatedness. In practice, this was accomplished by only retaining the top 25% with highest allele fraction. If a patient's tumors shared two or more somatic mutations at this high fraction, they were considered to be of common clonal origin.

Ethics statement

The study was approved by the Regional Ethical Review Board in Stockholm with registration numbers 2005/1288 -31/1 and 2006/362-32. The study did not require written consent from the participating individuals.

Results

The patients we selected represent a diverse set of breast cancer patients in relation to age of onset, treatment and histopathological parameters (Table 1). Across the 50 tumors and paired normal samples in this study, we sequenced 2385583114 reads yielding a mean target coverage of 22.3x per sample. In total, we called 3308 somatic mutations, with a median of 52 mutations per tumor that passed the default quality filters of MuTect. No somatic mutations were shared between patients, similar to previous studies on breast cancer exome sequencing[18].

We identified three patients where the somatic mutations were shared between the first and the second cancers, indicative of metastatic spread rather than a second primary tumor. Notable, these three patients we're all diagnosed at a relatively young age (39, 42 and 51, mean age of cohort was 54) and all had ER-negative first tumors, as well as second tumors (ER status is missing for the second tumor of patient 354). In Figure 1, examples of a patient with a shared profile (patient 177, a) and a patient with a non-shared profile (patient 806, b) are shown. For patient 177, shared mutations are present at high fraction in both tumors, whereas no mutations with a high allelic fraction are found for patient 806.

Prognosis

We followed the outcome of the patients, recording metastasis and death, until Dec 31 2012. All three patients predicted to have metastatic disease were clinically diagnosed with metastatic disease within three years of their second diagnosis (Figure 2). In contrast, among the 22 patients with no shared mutations between the tumors, 7 developed distant metastasis within a three-year period, whereas 15 did not develop distant metastasis or did so more than three years from their second breast cancer diagnosis (3 of 3 vs. 7 of 22, one-tailed Fisher's exact test p=0.05).

Discussion

In this study, we show that 3 of 25 CBC-patients have a shared mutation pattern between the first and the second tumor, indicating that the second cancer is likely to be a metastasis of the first. We also find this shared pattern to correlate with an increased risk for clinical diagnosis of distant metastasis within a three-year period. Other studies have previously investigated the relationship between the first and second cancer, using various methodologies. Banelli and colleagues investigated the methylation patterns of the androgen receptor in 19 patients with contralateral breast cancer, with successful assay data for only 10 patients[19]. In none of the 10 patients, the second tumor was determined to be a recurrence of the first. Janschek *et al*. investigated TP53 mutations across 33 patients, but were due to the single-gene assay and lack of normal tissue, only able to determine the clonal status for 13 patients, of which two were determined to have a metastasis rather than a second primary. Imyanitov *et al*. determined status for all but one of 28 investigated cases, using 14 fixed markers for each tumor[5], showing a higher similarity between synchronous CBC than metachronous CBC. Two studies have used CGH to investigate the relationship between the tumors, showing that investigation of genomic imbalances can be used to detect recurrent breast tumors in both the ipsilateral and contralateral settings[6, 7]. In each study, a single patient was detected to have a metastasis instead of a second primary, out of 8 bilateral cases in the Brommesson *et al.* study, and 9 bilateral cases in the Teixeira *et al.* study.

In our study, we used exome sequencing to identify if two tumors carry shared mutations at high mutational fractions as a proxy for a common origin. The benefits of using somatic mutations is two-fold: 1) Mutations are not dynamic in nature. A mutation can be lost if a segment of a chromosome is deleted, but this will affect only a minority of the mutations. 2) The majority of mutations identified in the exome are random passenger mutations, under little or no selective pressure. The resulting mutational pattern is unique for each tumor, offering means to identify common somatic denominators to metastatic lesions. Since de novo mutations occur in every cell division, a tumor mass is a mosaic of cells carrying slightly different

somatic mutations. We used mutations with high mutational frequency since those mutations arose early in the development of the tumor and therefore have a high likelihood of being present in any cells originating from the primary tumor. The same passenger mutations are not expected to occur in an independent second primary tumor. If common mutations would be found, it would strongly indicate a common somatic origin.

Two limitations have to be acknowledged in our study. Firstly, we used FFPE tumors, the oldest of which were resected in the late 1970s. Since no paired blood samples were available, we performed laser capture microscopy to obtain pure populations of tumor and normal cells. Working with very small amounts of DNA extracted from FFPE material limits both the quality of the sequencing data, as well as the amount of data obtained, as indicated by a relatively low average coverage in our data.

Clinical sequencing of cancer holds great promise to improve the treatment of cancer. Molecular subtyping and grading in breast cancer has been shown to improve upon the standard immunohistochemical and histopathological analyses[20, 21] and with the advent of Next Generation Sequencing, molecular pathology is rapidly advancing. We show that investigation of the somatic mutation patterns aids in identifying patients with metastatic disease over patients with a new primary cancer. In a clinical setting, the limitations of our study will not be applicable. Access to a blood sample, a skin punch, or unaffected lymph nodes means that no laser microdissection needs to be performed, thereby simplifying the laboratory procedure. Thus, downstream molecular methods can then be carried out with greatly increased amounts of input DNA, increasing both the throughput and the quality of the acquired data.

Conclusions

The treatment regimens for breast cancer are very different in the primary setting and the metastatic setting. Therefore, being able to determine whether a tumor is a new primary or a recurrence of a previous cancer is of utmost importance. In addition, as our knowledge on genomic patterns associated with treatment

resistance increases, genomic profiling will be able to provide a more direct guidance on what treatment to select. As sequencing steps in to the clinical routine analysis of cancer and their recurrences, this will be an important part in directing the right treatment to the right patient, and thereby improve survival.

Authors Contributions

KC, PH and MS conceived of and coordinated the study. DK and JL performed data analysis. DK drafted the manuscript. JR and CH performed laser microdissection and tissue evaluation. All authors reads and approved the final manuscript.

Acknowledgements

The authors would like to thank the Science for Life Laboratory Stockholm Application lab for sequencing and Gabriela Prochazka, Anna Westring, Simon Sundling and Heidi Talvitie for help with library preparation of the samples.

Conflicts of interest

The authors have no conflicts of interest.

References

1. *Globocan 2008* [http://globocan.iarc.fr]

2. Brenner H, Engelsmann B, Stegmaier C, Ziegler H: **Clinical epidemiology of bilateral breast cancer.** *Cancer* 1993, **72**:3629–3635.

3. Hartman M, Czene K, Reilly M, Adolfsson J, Bergh J, Adami H-O, Dickman PW, Hall P: **Incidence and prognosis of synchronous and metachronous bilateral breast cancer.** *J Clin Oncol* 2007, **25**:4210–4216.

4. Janschek E, Kandioler-Eckersberger D, Ludwig C, Kappel S, Wolf B, Taucher S, Rudas M, Gnant M, Jakesz R: **Contralateral breast cancer: molecular differentiation between metastasis and second primary cancer.** *Breast Cancer Res Treat* 2001, **67**:1–8.

5. Imyanitov EN, Suspitsin EN, Grigoriev MY, Togo AV, Kuligina ES, Belogubova EV, Pozharisski KM, Turkevich EA, Rodriquez C, Cornelisse CJ, Hanson KP, Theillet C: **Concordance of allelic imbalance profiles in synchronous and metachronous bilateral breast carcinomas**. *Int J Cancer* 2002, **100**:557–564.

6. Brommesson S, Jönsson G, Strand C, Grabau D, Malmström P, Ringnér M, Fernö M, Hedenfalk I: **Tiling array-CGH for the assessment of genomic similarities among synchronous unilateral and bilateral invasive breast cancer tumor pairs**. *BMC Clin Pathol* 2008, **8**:6.

7. Teixeira MR, Ribeiro FR, Torres L, Pandis N, Andersen JA, Lothe RA, Heim S: **Assessment of clonal relationships in ipsilateral and bilateral multiple breast carcinomas by comparative genomic hybridisation and hierarchical clustering analysis**. *Br J Cancer* 2004.

8. Ding L, Ley TJ, Larson DE, Miller CA, Koboldt DC, Welch JS, Ritchey JK, Young MA, Lamprecht T, Mclellan MD, Mcmichael JF, Wallis JW, Lu C, Shen D, Harris CC, Dooling DJ, Fulton RS, Fulton LL, Chen K, Schmidt H, Kalicki-Veizer J, Magrini VJ, Cook L, Mcgrath SD, Vickery TL, Wendl MC, Heath S, Watson MA, Link DC, Tomasson MH, et al.: **Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing.** *Nature* 2012, **481**:506–509.

9. Ding L, Ellis MJ, Li S, Larson DE, Chen K, Wallis JW, Harris CC, Mclellan MD, Fulton RS, Fulton LL, Abbott RM, Hoog J, Dooling DJ, Koboldt DC, Schmidt H, Kalicki J, Zhang Q, Chen L, Lin L, Wendl MC, Mcmichael JF, Magrini VJ, Cook L, Mcgrath SD, Vickery TL, Appelbaum E, Deschryver K, Davies S, Guintoli T, Lin L, et al.: **Genome**

remodelling in a basal-like breast cancer metastasis and xenograft. *Nature* 2010, **464**:999–1005.

10. Van Allen EM, Foye A, Wagle N, Kim W, Carter SL, McKenna A, Simko JP, Garraway LA, Febbo PG: **Successful whole-exome sequencing from a prostate cancer bone metastasis biopsy**. *Prostate Cancer Prostatic Dis* 2013, **17**:23–27.

11. Haffner MC, Mosbruger T, Esopi DM, Fedor H, Heaphy CM, Walker DA, Adejola N, GÃ¼rel M, Hicks J, Meeker AK, Halushka MK, Simons JW, Isaacs WB, De Marzo AM, Nelson WG, Yegnasubramanian S: **Tracking the clonal origin of lethal prostate cancer**. *J Clin Invest* 2013, **123**:4918–4922.

12. *SeqPrep* [https://github.com/jstjohn/SeqPrep]

13. Li H, Durbin R: **Fast and accurate short read alignment with Burrows-Wheeler transform.** *Bioinformatics* 2009, **25**:1754–1760.

14. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup: **The Sequence Alignment/Map format and SAMtools.** *Bioinformatics* 2009, **25**:2078–2079.

15. *Picard* [http://picard.sourceforge.net]

16. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA: **The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data**. *Genome Research* 2010, **20**:1297–1303.

17. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel SB, Meyerson M, Lander ES, Getz G: **Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples.** *Nat Biotechnol* 2013, **31**:213–219.

18. Cancer Genome Atlas Network: **Comprehensive molecular portraits of human breast tumours.** *Nature* 2012, **490**:61–70.

19. Banelli B, Casciano I, Di Vinci A, Gatteschi B, Levaggi A, Carli F, Bighin C, Salvi S, Allemanni G, Ghiorzo P, Pronzato P, Venturini M, Romani M, Del Mastro L: **Pathological and molecular characteristics distinguishing contralateral metastatic from new primary breast cancer**. *Annals of Oncology* 2010, **21**:1237–1242.

20. Bastien RR, guez-Lescure LR, Ebbert MT, Prat A, rriz BM, Rowe L, Miller P, Ruiz-Borrego M, Anderson D, Lyons B, lvarez I, Dowell T, Wall D, Segu MN, Barley L, Boucher KM, Alba E, Pappas L, Davis CA, Aranda I, Fauron C, Stijleman IJ, Palacios J, n AA, Carrasco E, Caballero RA, Ellis MJ, Nielsen TO, Perou CM, Astill M, et al.: **PAM50 Breast Cancer Subtyping by RT-qPCR and Concordance with Standard Clinical Molecular Markers**. *BMC Medical Genomics* 2012, **5**:1–1.

21. Ivshina AV, George J, Senko O, Mow B, Putti TC, Smeds J, Lindahl T, Pawitan Y, Hall P, Nordgren H, Wong JEL, Liu ET, Bergh J, Kuznetsov VA, Miller LD: **Genetic reclassification of histologic grade delineates new clinical subtypes of breast cancer.** *Cancer Res* 2006, **66**:10292–10301.

Figure Legends

Figure 1

Figure 1. Allele frequencies for two example cases. (A) Several somatic mutations present at high mutational fraction in both tumors, indicating that tumor 2 was formed by a clonal expansion of cells from tumor 1. (B) No somatic mutations shared between tumor 1 and 2, indicating that the two tumors are independent events. Plots for all patients are available in the Additional File 1.

Figure 2

Figure 2. Time from second diagnosis and potential diagnosis of distant metastasis. For three patients (168, 177 and 354, blue and marked with asterisk) we predict that the second tumor is not an independent tumor, but a clonal expansion (metastasis) of the first. These three patients were all diagnosed with distant metastasis within three years after their second breast cancer diagnosis. Death or end of follow-up is indicated by solid or open squares, respectively.

Tables

Table 1

Table 1. Characteristics, treatment and estrogen receptor status of the patients included in the study. RT, radiotherapy; HT, hormone therapy; CT, chemotherapy.

Patient	Year of first cancer	Age at first cancer	Time between cancers (years)	Time from second cancer to metastasis (years)	Treatment for first cancer	Treatment for second cancer	ER-status at first cancer	ER-status at second cancer
779	1990	62	2,08		RT+HT	HT	Positive	Negative
351	2000	48	4,07		RT+HT	RT+HT	Positive	Positive
941	1997	63	1,03		CT	HT	Negative	Positive
683	2002	77	1,86		NT	HT	Negative	Positive
621	1985	62	1,79	2,15	RT	NT	Positive	Positive
354 *	1992	39	3,39	0,28	RT+CT+HT	HT	Negative	Unknown
177 *	1983	42	1,42	2,30	RT	RT+CT	Negative	Negative
409	1999	37	2,47		RT+CT+HT	RT+CT+HT	Positive	Negative
886	2003	71	1,53	1,40	RT	HT	Negative	Positive
168 *	1994	51	2,91	2,59	RT+CT	RT+CT+HT	Negative	Negative
806	2000	71	4,41		HT	HT	Positive	Positive
169	1987	51	1,28	3,13	RT+CT	HT	Negative	Positive
742	1989	62	4,50	7,99	NT	NT	Positive	Unknown
3029	2001	51	4,78		RT+CT+HT	CT	Positive	Negative
209	1991	46	4,21		HT	NT	Negative	Positive
198	1999	55	4,85		RT+CT+HT	CT	Positive	Negative
126	2000	58	4,74		RT	RT+HT	Positive	Positive
3030	1996	45	0,41	3,51	NT	RT+CT	Negative	Negative
639	1985	61	3,53		HT	NT	Negative	Negative
340	2001	50	3,13		RT+CT	RT+HT	Negative	Positive
919	1998	65	4,59		RT+CT	RT+HT	Negative	Positive
30276	1981	74	4,90	1,74	RT+HT	HT	Positive	Negative
383	1987	31	0,65		RT	RT+CT	Positive	Positive
275	1990	43	3,91		NT	CT+HT	Positive	Positive
404	1998	39	2,68	2,40	CT+HT	RT+CT+HT	Positive	Negative
Mean		54,16	3,01	2,75				