MOLECULAR ONCOLOGY 9 (2015) 877–888

available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/molonc

Proper genomic profiling of (*BRCA1*-mutated) basal-like breast carcinomas requires prior removal of tumor infiltrating lymphocytes

Maarten P.G. Massink[a], *Irsan E. Kooi*[a], *Saskia E. van Mil*[a], *Ekaterina S. Jordanova*[b], *Najim Ameziane*[a], *Josephine C. Dorsman*[a], *Daphne M. van Beek*[a], *J. Patrick van der Voorn*[c], *Daoud Sie*[c], *Bauke Ylstra*[c], *Carolien H.M. van Deurzen*[d], *John W. Martens*[e], *Marcel Smid*[e], *Anieta M. Sieuwerts*[e], *Vanja de Weerd*[e], *John A. Foekens*[e], *Ans M.W. van den Ouweland*[f], *Ewald van Dyk*[g], *Petra M. Nederlof*[h], *Quinten Waisfisz*[a,*], *Hanne Meijers-Heijboer*[a]

[a]*Department of Clinical Genetics, VU University Medical Center, Amsterdam, The Netherlands*
[b]*Department of Obstetrics and Gynaecology, Center for Gynaecologic Oncology, VU University Medical Center, Amsterdam, The Netherlands*
[c]*Department of Pathology, VU University Medical Center, Amsterdam, The Netherlands*
[d]*Department of Pathology, Erasmus MC Cancer Institute, Erasmus University Medical Center, Rotterdam, The Netherlands*
[e]*Department of Medical Oncology, Erasmus MC Cancer Institute, Erasmus University Medical Center, Rotterdam, The Netherlands*
[f]*Department of Clinical Genetics, Erasmus MC Cancer Institute, Erasmus University Medical Center, Rotterdam, The Netherlands*
[g]*Department of Pathology, Netherlands Cancer Institute, Amsterdam, The Netherlands*
[h]*Department of Molecular Carcinogenesis, Netherlands Cancer Institute, Amsterdam, The Netherlands*

ARTICLE INFO

Article history:
Received 5 September 2014
Received in revised form

ABSTRACT

Introduction: BRCA1-mutated breast carcinomas may have distinct biological features, suggesting the involvement of specific oncogenic pathways in tumor development. The identification of genomic aberrations characteristic for BRCA1-mutated breast carcinomas

Abbreviations: aCGH, micro-array based competitive genomic hybridization; BAF, B-allele frequency; BLC, basal-like carcinomas; BRCAX, non-BRCA1/2 mutated familial breast carcinomas; CIN, chromosomal instability; CK, cytokeratin; CNA, copy number aberration; ER, estrogen receptor; FE, Fischer exact; FFPE, formalin-fixed, paraffin-embedded; HER2, human epidermal growth factor receptor 2; HR, homologous recombination; H&E, hematoxylin and eosin; LOH, loss of heterozygosity; mBAF, mirrored B-allele frequency; MSP, methylation specific PCR; PR, progesterone receptor; SNP, single nucleotide polymorphism; TIL, tumor infiltrating lymphocyte; WGS, whole genome sequencing.

* *Corresponding author*. Tel.: +31 204449925; fax: +31 204448285.
E-mail addresses: m.massink@vumc.nl (M.P.G. Massink), ei.kooi@vumc.nl (I.E. Kooi), se.vanmil@vumc.nl (S.E. van Mil), e.jordanova@vumc.nl (E.S. Jordanova), n.ameziane@vumc.nl (N. Ameziane), jc.dorsman@vumc.nl (J.C. Dorsman), d.vanbeek@vumc.nl (D.M. van Beek), jp.vandervoorn@vumc.nl (J.P. van der Voorn), d.sie@vumc.nl (D. Sie), b.ylstra@vumc.nl (B. Ylstra), c.h.m.vandeurzen@erasmusmc.nl (C.-H.M. van Deurzen), j.martens@erasmusmc.nl (J.W. Martens), m.smid@erasmusmc.nl (M. Smid), a.sieuwerts@erasmusmc.nl (A.M. -Sieuwerts), v.deweerd@erasmusmc.nl (V. de Weerd), j.foekens@erasmusmc.nl (J.A. Foekens), a.vandenouweland@erasmusmc.nl (A.M.W. van den Ouweland), e.v.dijk@nki.nl (E. van Dyk), p.nederlof@nki.nl (P.M. Nederlof), q.waisfisz@vumc.nl (Q. Waisfisz), h.meijers-heijboer@vumc.nl (H. Meijers-Heijboer).

http://dx.doi.org/10.1016/j.molonc.2014.12.012

20 December 2014
Accepted 27 December 2014
Available online 13 January 2015

Keywords:
BRCA1
Genomic profiling
Copy number
Tumor infiltrating lymphocytes
Gene expression
FACS analysis

could lead to a better understanding of *BRCA1*-associated oncogenic events and could prove valuable in clinical testing for *BRCA1*-involvement in patients.

Methods: For this purpose, genomic and gene expression profiles of basal-like *BRCA1*-mutated breast tumors (n = 27) were compared with basal-like familial BRCAX (non-*BRCA1/2/CHEK2**1100delC) tumors (n = 14) in a familial cohort of 120 breast carcinomas.

Results: Genome wide copy number profiles of the *BRCA1*-mutated breast carcinomas in our data appeared heterogeneous. Gene expression analyses identified varying amounts of tumor infiltrating lymphocytes (TILs) as a major cause for this heterogeneity. Indeed, selecting tumors with relative low amounts of TILs, resulted in the identification of three known but also five previously unrecognized *BRCA1*-associated copy number aberrations. Moreover, these aberrations occurred with high frequencies in the *BRCA1*-mutated tumor samples. Using these regions it was possible to discriminate *BRCA1*-mutated from BRCAX breast carcinomas, and they were validated in two independent cohorts. To further substantiate our findings, we used flow cytometry to isolate cancer cells from formalin-fixed, paraffin-embedded, *BRCA1*-mutated triple negative breast carcinomas with estimated TIL percentages of 40% and higher. Genomic profiles of sorted and unsorted fractions were compared by shallow whole genome sequencing and confirm our findings.

Conclusion: This study shows that genomic profiling of in particular basal-like, and thus *BRCA1*-mutated, breast carcinomas is severely affected by the presence of high numbers of TILs. Previous reports on genomic profiling of *BRCA1*-mutated breast carcinomas have largely neglected this. Therefore, our findings have direct consequences on the interpretation of published genomic data. Also, these findings could prove valuable in light of currently used genomic tools for assessing *BRCA1*-involvement in breast cancer patients and pathogenicity assessment of *BRCA1* variants of unknown significance. The *BRCA1*-associated genomic aberrations identified in this study provide possible leads to a better understanding of *BRCA1*-associated oncogenesis.

1. Introduction

Breast tumors of *BRCA1* germ line mutation carriers have been shown to be primarily of the basal-like subtype (75–90%) (Sorlie et al., 2003; Jonsson et al., 2010). These basal-like breast carcinomas (BLCs) represent about 10–20% of all breast carcinomas (Foulkes et al., 2003; Sorlie et al., 2003) and are generally high grade, triple negative ((ER-), (PR-), and (HER2/ERBB2-) and express high molecular weight cytokeratins (CK5/6, 14 and 17) (Lakhani et al., 2002; Foulkes et al., 2003; Honrado et al., 2005; Rubinstein, 2008).

High levels of genomic instability are frequently observed in *BRCA1*-mutated carcinomas and BLCs in general, both sporadic and hereditary (Bergamaschi et al., 2006; Jonsson et al., 2010; Smid et al., 2010). Although frequent, these genomic aberrations mostly represent low copy number gains and losses of large DNA segments rather than (focal) high level amplifications (Bergamaschi et al., 2006; Stefansson et al., 2009). This observation fits well with the described role of *BRCA1* in DNA damage response and DNA repair through homologous recombination (HR) (Roy et al., 2012). Hyper-methylation of the *BRCA1* promoter region and reduced *BRCA1* mRNA expression levels have been observed in sporadic BLCs and could act as an alternative mechanism of *BRCA1* inactivation (Wei et al., 2005; Birgisdottir et al., 2006; Turner and Reis-Filho, 2006). However, the involvement of a dysfunctional *BRCA1* gene in sporadic BLCs has also been questioned due to lack of somatic mutations in the *BRCA1* gene (Futreal et al., 1994; Gonzalez-Angulo et al., 2011). Nevertheless, the observed similarities between *BRCA1*-mutated and non-mutated BLCs have led to the suggestion that targeting a dysfunctional *BRCA1*-pathway, exploiting in particular a defective HR, might be effective in all BLCs (Turner et al., 2004). However, platinum based chemotherapy and treatments with PARP inhibitors have shown inconsistent results in BLCs (Silver et al., 2010; Lips et al., 2011a). For this reason, the role of a dysfunctional *BRCA1* gene or pathway as a general characteristic of BLCs remains uncertain.

To allow discrimination between *BRCA1*-mutated and unselected sporadic/hereditary breast tumors, micro-array based comparative genomic hybridization studies (aCGH) have been used to identify characteristic genomic aberrations in *BRCA1*-mutated breast tumors (Tirkkonen et al., 1997; Wessels et al., 2002; van Beers et al., 2005; Jonsson et al., 2005; Joosse et al., 2009). Recent studies comparing BLCs with and without *BRCA1* mutations have reported few or even no differential regions of genomic aberrations (Stefansson et al., 2009; Jonsson et al., 2010; Waddell et al., 2010). A well known difficulty in DNA-copy-number determination by aCGH and SNP array analyses is variable tumor purity due to normal DNA contamination. In literature, a number of approaches have been proposed to try to mitigate this effect (Yau et al., 2010). However, such methods have not been applied in aforementioned genomic profiling efforts of *BRCA1*-mutated breast carcinomas.

The use of *BRCA1* specific classifiers would facilitate the identification of women and their family members with

unknown germ line mutation status or undetected germ line mutations. They might also prove valuable in assessing the pathogenicity of variants of unknown clinical significance (Joosse et al., 2009). Furthermore, the identification of driver genes in regions characteristic for *BRCA1*-mutated tumors could lead to a better understanding of the underlying process of oncogenesis and may provide novel clues for targeted therapies.

The aim of this study was to identify characteristic copy number aberrations (CNAs) for *BRCA1*-mutated breast. For this reason single nucleotide polymorphism (SNP) array data and gene expression data for 120 hereditary breast tumors were analyzed. As the majority of the *BRCA1*-mutated breast carcinomas are found to be of the basal-like subtype, our primary analysis was restricted to BLCs. Initially, comparison of 27 *BRCA1*-mutated versus 14 BRCAX BLCs did not show clear CNAs characteristic for the *BRCA1*-mutated BLCs. However, optimal sample selection of tumors with low amounts of TILs allowed for the identification of three known and five novel *BRCA1*-associated CNAs capable of discriminating between *BRCA1*-mutated and BRCAX BLCs.

2. Materials and methods

2.1. Ethics statement

This study has been approved by the medical ethical committee at Erasmus MC, and was performed according the Code of Conduct of the Federation of Medical Scientific Societies in The Netherlands.

2.2. Sample collection

Fresh-frozen specimens of primary breast tumors from female familial breast cancer cases were selected from the tissue bank of the Erasmus Medical Center Rotterdam. In this study, 120 tumor samples for which both SNP array and gene expression data is available were used for further analyses. The entire cohort of 154 samples has been described previously (Nagel et al., 2012). The gene expression and SNP microarray data have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number 54219.

For validation purposes three publicly available datasets were used (Supplemental Figure S1). These are: 1.) 118 BLCs selected from the discovery dataset of this study on sporadic breast carcinomas (Curtis et al., 2012) to study the effect of high numbers of TILs on copy number calling, 2.) 359 primary breast tumors, including 18 *BRCA1*-mutated BLCs, for which aCGH and gene expression data is available (Jonsson et al., 2010), used for validation of *BRCA1*-associated CNAs and the TIL gene expression profile, and 3.) 186 samples including 40 *BRCA1*-mutated breast tumors for which aCGH data and hematoxylin and eosin stained (H&E) sections of paraffin-embedded tissue is available, obtained from the Netherlands Cancer Institute (Schouten et al., 2013) used for further validation of *BRCA1*-associated CNAs. For FACS sorting, archival, formalin-fixed, paraffin-embedded, (FFPE) germ line *BRCA1*-mutated triple negative breast tumors were obtained from the FFPE tissue bank of the Department of Pathology, VUMC, Amsterdam, The Netherlands. Samples were handled according to the medical ethical guidelines described in the Code for Proper Secondary Use of Human Tissue established by the Netherlands Federation of Medical Sciences.

2.3. Mutation screening

BRCA1, *BRCA2* and *CHEK2**1100delC mutation screening has been described (Nagel et al., 2012).

2.4. Gene expression microarrays

Affymetrix hgU133_Plus_2.0 GeneChips based expression data have been previously described (Nagel et al., 2012). The gene expression data have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number 54219. The data was analyzed in Partek Genomics Suite (v6.6, Partek Inc.).

2.5. Classification of intrinsic molecular subtypes

The intrinsic gene list was used to appoint the samples to molecular subtypes as described (Smid et al., 2008). In short, the intrinsic gene list (Perou et al., 2000) was mapped to the corresponding probe-sets on the hgU_133_plus_2.0 array using Unigene Cluster Id's. The most variable probe-sets were used to cluster 120 familial samples using average linkage hierarchical clustering with correlation as a distance metric.

2.6. Copy number analyses by SNP arrays

The array intensity CEL files for 120 breast tumor samples, including 35 BRCA1, five BRCA2, 17 CHEK2*1100delC mutated and 63 BRCAX tumor samples were processed by Partek Genomics Suite using default settings for background correction and summarization, results were corrected for GC-content and fragment length. Unpaired copy number analysis was performed in Partek Genomics Suite, comparing signal Log2 ratios to a custom created reference baseline of 90 female HapMap CEPH samples with European ancestry (CEU). The genomic segmentation algorithm was used to detect breakpoint regions and estimate copy number levels with stringent parameters ($P < 0.0001$, >20 markers, signal/noise: 0.45). With an expected normal range of 2 ± 0.25 copies. Differences between the tumor groups (mutation class) for frequency of copy number aberrations (gained, lost, or unchanged) were calculated by employing a 3×2 Fisher's exact test (FE). Resulting p-values were not directly corrected for multiple testing. SNP array, gene and cytogenetic band locations are based on the hg18 Genome build.

For unsupervised hierarchical clustering of copy number data the called copy number states (gain, loss or neutral) of the segmentation data were used as distance metric. Agglomerative clustering was performed on these data by Euclidean distance and Wards method.

Identical settings for copy number analysis were applied in the validation cohorts. However, for validation dataset 3, the overall signal intensities of the samples (run on aCGH platform as opposed to SNP arrays in our primary study) was lower compared to the samples of the primary study. Therefore, copy number segmentation was also performed with

less stringent settings (signal/noise: 0.3, expected normal range of 2 ± 0.20) resulting in more samples passing the CNA calling criteria and being suited for *BRCA1* classification. Gene expression profiling was not performed on these samples. Further analyses were performed on samples for which hormone receptor status information was available. Details for correlating copy number and gene expression data can be found in Supplemental File S2.

2.7. LOH calling by detection of allelic imbalance

A segmentation based approach of allelic imbalances was used to identify regions of loss of heterozygosity (LOH). The B-allele frequencies of 120 breast tumor samples were generated in Partek Genomics Suite. Mirrored BAF profiles (mBAF) were used as previously described (Staaf et al., 2008). The resulting mBAF profiles were segmented in Partek Genomics Suite and LOH calling of segmented regions was done by applying a fixed allelic imbalance threshold of 0.76 and p-value <0.01. The same parameters used in segmentation of the copy number data were applied, except that a window size of 100 SNPs instead of 20 SNPs was used as a minimum number of genomic markers.

2.8. Analysis of BRCA1 promoter methylation

BRCA1 promoter hyper-methylation was determined by methylation-specific PCR (MSP) after bisulfite treatment of DNA. Bisulfite modification of DNA to convert unmethylated cytosine residues to uracil was carried out using the EZ DNA methylation kit (Zymo Research, Orange, CA) following the manufacturer's protocol. A genomic DNA control sample was used as a negative control, and the same sample treated *in vitro* with *SssI* bacterial methylase was used as a positive control. Primer sequences were adapted from Wei et al. (Wei et al., 2005). The PCR reaction was composed of an initial denaturation at 95 °C for 5 min, followed by 37 cycles consisting of 30 s at 95 °C, 30 s at 61 °C, 45 s at 72 °C and ending with a final 5 min extension at 72 °C. The PCR products were analyzed by electrophoresis and UV transillumination.

2.9. Scoring of TILs in tumor sections

H&E stained frozen and FFPE sections were scored for the amount of infiltrating lymphocytes in the tumor by experienced pathologists from the respective institutes. This was performed by estimating the number of lymphocytes, and reported as a percentage of all nuclei present within that area.

2.10. Scoring BRCA1-associated genomic aberrations

To discriminate between *BRCA1*-mutated and BRCAX BLCs, a class prediction algorithm using a nearest shrunken centroid methodology with equal priors, imbedded in Partek Genomics Suite software, was applied. As input, mean copy number values from the segmentation results were used. Genomic regions used by this algorithm were found to overlap with the most significant regions as reported in the Fisher's exact test results. These regions in combination with the most frequently observed genomic aberrations present in all BLCs were used to define 12 regions of *BRCA1*-BLC associated copy number aberrations. These specific regions of copy number loss or gain were analyzed in cross platform validation datasets. Samples were scored for the amount of copy number aberrations (gains/losses) in these regions, when 9 or more regions harbor similar CNAs as determined for the *BRCA1*-mutated BLCs, samples are called as *BRCA1*-like. This threshold of 9 or more was chosen based on the primary dataset whereby all cases with low immune-infiltrate levels of *BRCA1*-mutated origin were scored as *BRCA1*-like. For the Netherlands Cancer Institute validation set, samples were coded and results were disclosed after CNA scoring.

2.11. Tumor dissociation and DNA flow cytometry

H&E-stained paraffin sections taken from 10 selected triple negative *BRCA1*-mutated breast tumor samples were scored by an experienced pathologist to determine the extent of TILs. Five samples with the highest amount of TILs (>40%) were selected for further analysis. Sample processing and FACS sorting was performed as described previously with minor modifications (Corver et al., 2005; Corver and ter Haar, 2011) for details see (Supplemental File S2). After sorting, DNA was isolated using the Promega, Wizard Genomic DNA Purification kit.

2.12. Copy number profiling by shallow whole genome sequencing

DNA derived from three cellular fractions isolated from three germ line *BRCA1*-mutated breast carcinomas was sequenced at low coverage depth (~0.1–0.3 fold, for detailed sequencing statistics see Supplemental File S3). In total, 100–200 ng of DNA (depending on availability) was used for library preparation with the Illumina TruSeq DNA v2 kit, as described previously (Smeets et al., 2011; van Thuijl et al., 2013). Bar-coded libraries were pooled and run on an Illumina HiSeq 2500 (San Diego, CA, USA) for single end 50 cycle (SR50) sequencing. Sequencing reads were aligned to the human reference genome (build GRCh37/hg19) with BWA (Li and Durbin, 2009) and PCR duplicates as well as ambiguous reads were filtered out. Next, the genome was divided into non-overlapping bins of 15 kbp, and the number of reads in each bin was counted. These counts were 1.) corrected (using GC content and mappability) and 2.) filtered using a blacklist derived from ENCODE data (Bernstein et al., 2012) and 1 KG (Abecasis et al., 2012), 3.) log2-transformed, 4.) median normalized, 5.) segmented, and called with Partek Genomics Suite similarly as described for the SNP array CNA analysis. The hg18 based genomic coordinates of the *BRCA1*-associated CNAs were mapped to the hg19 genome build using the hg-Liftover tool from UCSC (http://genome.ucsc.edu/). Copy number profiles are plotted using ggplot2 software (Wickham, 2009).

3. Results

3.1. BRCA1 status and intrinsic sub-typing

Gene expression and SNP array analyses were performed on 120 familial breast carcinomas. The cohort consists of 35

BRCA1-mutated, five *BRCA2* mutated, 17 *CHEK2**1100delC mutated and 63 BRCAX carcinomas without such mutations, referring to as BRCAX. In accordance with literature, 77% of the *BRCA1*-mutated carcinomas (n = 27) in this study are of the basal-like subtype. In accordance with the incidence of basal-like cancers amongst breast cancer in general, 14 out of 63 (22%) BRCAX breast carcinomas are of the basal-like subtype (Waddell et al., 2010). The remainder of *BRCA1*-mutated breast carcinomas is predominantly of the luminal-B subtype; see Supplemental file S4 for complete sample information. Evidence for *BRCA1* promoter hyper-methylation was detected for six samples by MSP, amongst which three BRCAX tumors of the basal-like subtype.

3.2. *Tumor infiltrating lymphocytes affect genomics analyses*

To gain insight into the genomic characteristics of *BRCA1*-mutated breast carcinomas, copy number and LOH profiling based on high-resolution SNP arrays was performed.

Copy number profiles for most of the *BRCA1*-mutated tumors, and basal-like carcinomas (BLCs) in general, showed many CNAs. Supervised comparison of the copy number profiles of the *BRCA1*-mutated BLCs versus BRCAX BLCs resulted in very few regions of copy number difference. Although few, these regions are in accordance with previous reports (Joosse et al., 2009). Regions include CNAs on chromosome 3 and copy number loss on chromosomal arm 15q (data not shown). This is in line with the unsupervised hierarchical clustering results showing clustering of the majority of BLCs in two groups irrespective of mutational status (Figure 1A). One group (Figure 1A, group 1; high CIN) is characterized by many CNAs and tight co-segregation and a second group (Figure 1A, group 2; low CIN) with markedly fewer CNAs that cluster more loosely within the bulk of the luminal samples. As shown in Figure 1B, regions with loss of heterozygosity (LOH) were frequently identified in group 1 BLCs and only very few in the group 2 BLCs. This marked difference was also seen for LOH at the *BRCA1*–gene locus on chromosome 17 in *BRCA1*-mutated BLCs. Only 9 out of the 27 *BRCA1*-mutated BLCs showed LOH at the BRCA1-locus; six with copy neutral LOH and three with copy number gains and LOH. For subsequent analyses the group 1 tumors were therefore subdivided based on LOH at the *BRCA1* locus, group 1-I with LOH and group 1-II without LOH. In comparison with the group 1-II tumors (n = 10) and the group 2 tumors (n = 8), the 9 group 1-I tumors displayed highest frequency of CNAs common to all groups as well as CNAs in general, see Figure 1C.

Analyzing gene expression of *BRCA1*-mutated BLCs by means of ANOVA with LOH at the *BRCA1* locus as single factor resulted in 618 differentially expressed probe-sets passing a step-up FDR p-value of 0.05 (Supplemental File S5). Gene ontology analysis of the 458 uniquely annotated genes using the DAVID bio-informatics tool (Huang et al., 2009a, 2009b) identified the term immune response (FDR corrected p-value< 1E-5) to be the most enriched biological process. Tumor infiltrating lymphocytes (TILs) could explain the above described results. Indeed, there appeared to be a strong correlation between the relative amount of the immune signature and genome-wide measured CNAs, $r_s = -0.66$, p-value< 0.001 (Supplemental File S6). Furthermore, assessing the proportion of lymphocytic nuclei within the tumor samples using H&E-stained frozen sections showed a significant correlation with the mRNA based immune signature ($r_s = 0.74$, p-value< 0.01) (Supplemental File S7).

To confirm the correlation between the TIL mRNA signature and CNAs in BLCs, we tested this in the discovery set of a publicly available dataset (Validation dataset 1 in Supplemental Figure 1; Curtis et al., 2012). The discovery set contains high resolution SNP array and gene expression data of 998 primary sporadic breast tumors, including 118 BLCs which were selected for further analyses. A significant correlation between immune signature and CNA levels was seen ($r_s = -0.52$, p-value< 0.001). Samples with the highest immune signature levels were found to be amongst those samples with the least reported CNAs, see Supplemental File S8.

3.3. *Optimal sample selection*

To select for samples with low number of TILs in our cohort, hierarchical clustering of expression data of all BLCs based on the immune signature probe-sets was used (Figure 2A). Following this approach, eight *BRCA1* and five BRCAX were selected for further analysis. These eight *BRCA1*-mutated samples all originate from the *BRCA1* group with high CIN with LOH at the *BRCA1* locus (Figure 1, group 1-I; n = 9). The ninth sample from this group was found to have a relative high TIL immune signature and was therefore not used in follow-up analyses. Copy number analyses on this selection of samples resulted in the identification of many CNAs with frequencies reaching levels of 80–100% for both the *BRCA1*-mutated and BRCAX BLCs, something which was not observed in previous studies by others (Tirkkonen et al., 1997; Wessels et al., 2002; van Beers et al., 2005; Jonsson et al., 2005; Stefansson et al., 2009; Jonsson et al., 2010; Waddell et al., 2010). Importantly, multiple regions of significant difference between the *BRCA1*-mutated and BRCAX tumors could now be detected (Figure 2B and C).

3.4. BRCA1-*mutated BLC associated genomic aberrations*

Subsequently, a class prediction algorithm using a nearest shrunken centroid methodology was used to identify eight genomic regions associated with *BRCA1*-status. Most of these regions are similar to the regions identified by the Fischer's exact test to be significantly different between the *BRCA1*-mutated and BRCAX BLCs. These genomic regions include three regions previously described in literature, i.e. copy number gain of the proximal tip of chromosomal arm 3p, copy number loss of a more distal region on chromosomal arm 3p and copy number loss of chromosomal arm 15q (Wessels et al., 2002; Jonsson et al., 2005; van Beers et al., 2005; Joosse et al., 2009) and five novel regions, i.e. copy number losses at chromosomal arms 7q, 13q, and 17p and copy number gains on chromosomal arms 9q and 15q (Tables 1 and 2).

For samples with low amount of TILS these CNAs proved capable of discriminating between the *BRCA1*-mutated and BRCAX BLCs. However, cross-platform applicability of such algorithms is problematic, and since the used sample size was

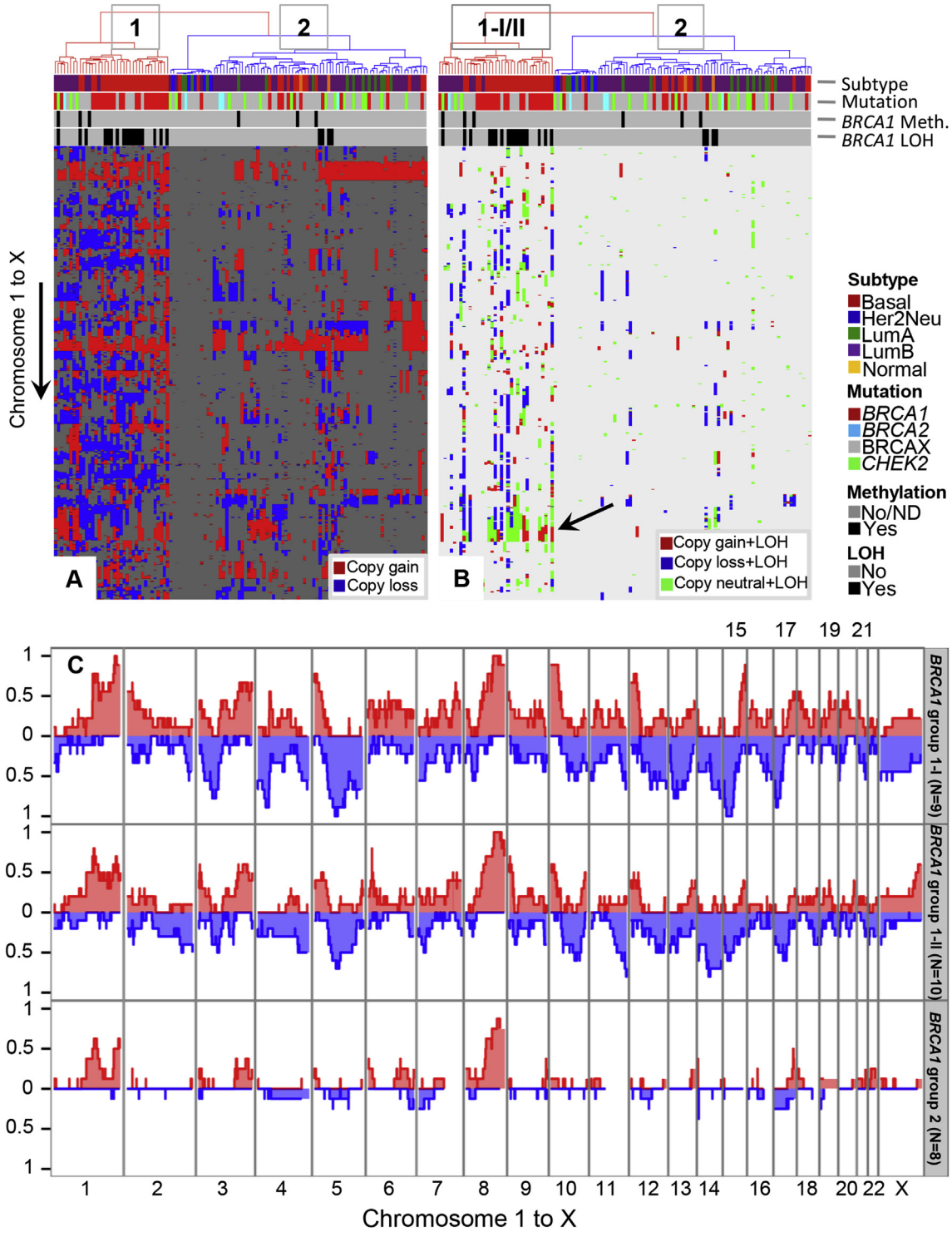

Figure 1 – Genomic profiles of 120 familial breast carcinomas. A, hierarchical clustering of CNA data. On the vertical axis, chromosomes 1 to X are displayed. Copy number gains are indicated in red and losses in blue. Subtype, mutation status and methylation of *BRCA1* are indicated for each sample in rows above the cluster plots by color. Group 1 contains the *BRCA1* BLC tumors with high CIN and group 2 contains the *BRCA1* BLC tumors with low CIN. B, integrated copy number and LOH analysis. Samples are shown as in Figure 1A. In the cluster plot green indicates copy neutral LOH, red; copy number gain with LOH and blue; copy number loss and LOH. Bottom row above the cluster plot indicates samples with LOH at the *BRCA1* locus (*BRCA1* LOH: black). The black arrow within the cluster plot indicates the *BRCA1* locus. Group 1 contains *BRCA1* BLCs with high CIN and LOH at the *BRCA1* locus (group 1-I samples) and *BRCA1* BLCs with high CIN but without LOH at the *BRCA1* locus (group 1-II samples), group 2 contains *BRCA1* BLC tumors with low CIN and no detectable LOH at the *BRCA1* locus C, the frequency (vertical axis) of gains (red) and losses (blue) is displayed for the 3 *BRCA1* groups mentioned in B.

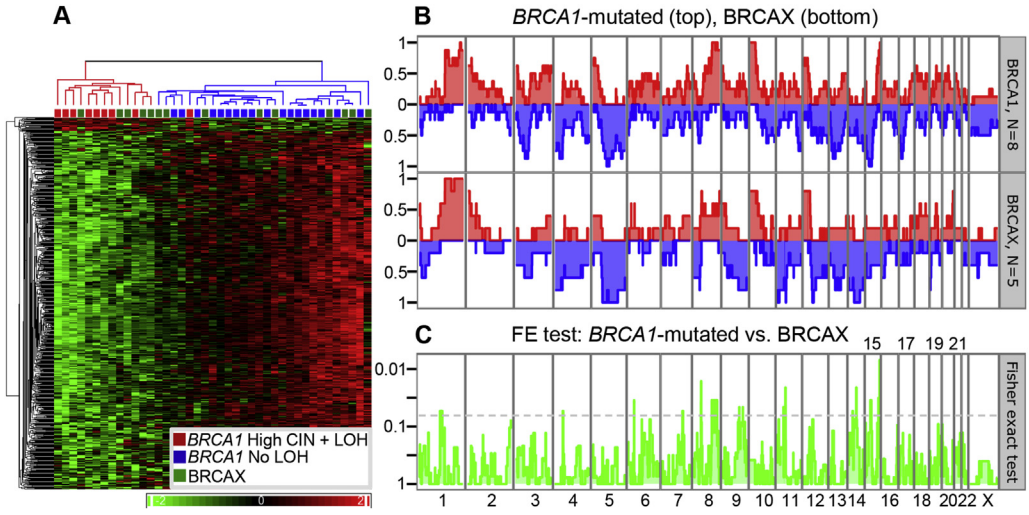

Figure 2 – Immune signature based gene expression clustering and comparison of aberration frequencies of selected *BRCA1*-mutated and BRCAX BLCs. A, hierarchical clustering of standardized expression data of all BLCs based on the immune signature. Samples from the red branch with relatively low level of immune signature are selected for further copy number analysis in panel B. Indicated in red and blue are *BRCA1*-mutated BLCs with (group 1-I, *BRCA1* BLC tumors with high CIN and LOH at the *BRCA1* locus, from Figure 1B, n = 9) and without detectable LOH (group 1-II, *BRCA1* BLC tumors with high CIN without LOH at the *BRCA1* locus, and group 2, *BRCA1* BLC tumors with low CIN without LOH at the *BRCA1* locus, from Figure 1B, n = 18) at the *BRCA1* locus respectively, green indicates the BRCAX BLCs. B, the frequency (x-axis) of gains (red) and losses (blue) are displayed along chromosomes 1 to X (y-axis) for 8 *BRCA1*-mutated (top panel) and 5 BRCAX (bottom panel) BLCs selected for a low immune signature, red branch in panel A. Note that one *BRCA1*-mutated BLC with detectable LOH was not used for further analyses due to a relative high level mRNA immune signature. C, Fisher's exact test is used to determine regions of differential copy number aberrations between the *BRCA1*-mutated and BRCAX BLCs. The dotted line represents a p-value threshold of 0.05 (not corrected for multiple testing). The regions above the threshold are considered to be significantly different between the groups. P-values are - log 10 transformed.

limited after sample selection, validation in other cohorts was a necessity. For this reason we used a simple method applicable to copy number data for all aCGH/SNP platforms, and copy number data in general. The eight regions mentioned above together with the four most frequently altered genomic regions found amongst all BLCs (two regions with copy number losses on chromosomes 4 and 5, and two with copy number gains on chromosomes 8 and 10), were used to score samples and discriminate between BRCA1-mutated BLCs and all other tumors (Tables 1 and 2). Samples were scored as described in section 2.10.

In our dataset of 120 familial breast carcinomas, the BRCA1-mutated BLCs with low TILs were found to have 9 to 12 of these specific CNAs, while all other samples had 8 or less (Figure 3). In the BRCA1-mutated BLCs with the highest amount of immune signature levels no or only very few of these CNAs are detected. High immune signature mRNA levels are also found in the three BRCAX BLCs with BRCA1 promoter hyper-methylation, making it difficult to assess a possible genomic resemblance with BRCA1-mutated BLCs for these samples.

3.5. *Validation of* BRCA1*-associated CNAs*

To validate our findings, a publicly available dataset of 359 primary breast tumors, including 18 BRCA1-mutated BLCs, for

Table 1 – *BRCA1*-associated copy number gains.

Copy number gain		
Chromosome	Start	End
3	3866308	10719396
8[a]	101981161	129288230
9	94909458	115239031
10[a]	101955	28140907
15	90224005	97940195

Hg18 based genomic regions of BRCA1-associated CNAs (gains).
a These regions are associated with tumor samples of the basal-like subtype in general.

Table 2 – *BRCA1*-associated copy number losses.

Copy number loss		
Chromosome	Start	End
3	51130455	64449381
4[a]	24424555	41258860
5[a]	75800618	86398661
7	102291383	122355559
13	33885792	49037828
15	28016572	41141219
17	6001240	12559716

Hg18 based genomic regions of BRCA1-associated CNAs (losses).
a These regions are associated with tumor samples of the basal-like subtype in general.

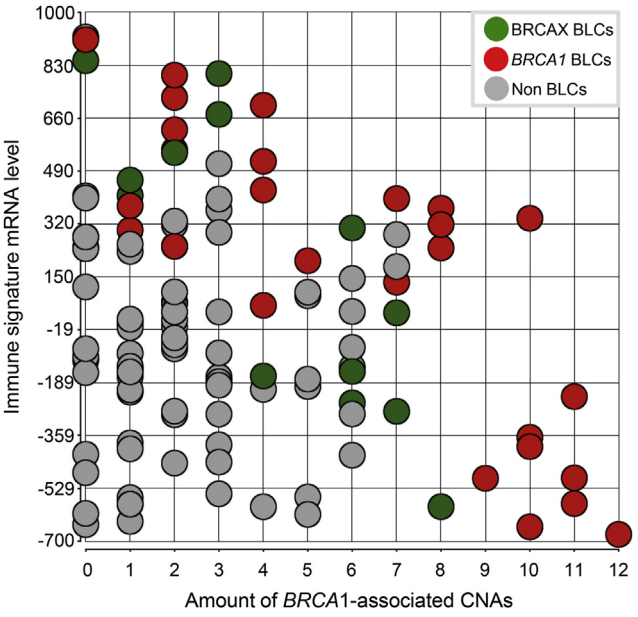

Figure 3 – Scoring of *BRCA1*-associated CNAs in relation to immune signature levels. *BRCA1*-mutated BLCs (red circles) with low immune signature levels (vertical axis) have a score of 9–12 *BRCA1*-associated CNAs (horizontal axis). BRCAX BLCs (green circles) with similar immune signature levels have less of these *BRCA1* specific CNAs. Samples of other subtypes (including *BRCA1*-mutated) are indicated by grey circles.

which aCGH and gene expression data are available, were analyzed (Validation dataset 2 in Supplemental File 1; Jonsson et al., 2005). A similar immune signature and correlation with genome wide CNAs was identified in this dataset (Supplemental File S9A).

For all 359 primary breast tumors the amount of *BRCA1*-associated CNAs was assessed, 10 samples were found to have 9 or more of these CNAs. Of these 10 samples 80% (n = 8) have a *BRCA1* defect, six samples are *BRCA1*-mutated BLCs, one sample a *BRCA1*-methylated BRCAX BLC, and one sample a *BRCA1*-methylated sporadic BLC. The other two samples are a *BRCA2*-mutated BLC and a BRCAX BLC of which no further information is available. The frequency plot of gains and losses for the six *BRCA1*-mutated BLCs shows identical regions and frequencies of CNAs as compared to the low TIL *BRCA1*-mutated BLCs from the primary study (Supplemental file S9B and Figure 2B).

A second dataset consisting of 186 micro-dissected samples including 39 *BRCA1*-mutated triple negative and 49 sporadic triple negative breast tumors for which aCGH data are available, was obtained from the Netherlands Cancer Institute (Validation dataset 3 in Supplemental Figure 1; Schouten et al., 2013). After copy number analysis of all 186 samples, 9 samples were found to have 9 or more *BRCA1*-associated CNAs. Of these 9 samples 89% (n = 8) have a BRCA1 defect, six are *BRCA1*-mutated and two samples *BRCA1*-methylated. All 9 samples are triple negative breast carcinomas.

Because the number of *BRCA1*-mutated samples scored as such was lower than expected, copy number analysis was also performed with less stringent parameters. (for details see section 2.6) With these settings 22 samples were found to have 9 or more *BRCA1*-associated CNAs. Of these samples 77% have a *BRCA1* defect; 14 are *BRCA1*-mutated and three *BRCA1*-methylated. For the called *BRCA1*-mutated samples, the frequencies and genomic coordinates of CNAs (Supplemental File S10 panel A) are again highly similar to those identified in the low TIL *BRCA1*-mutated BLCs from the primary dataset (Figure 2B) and the *BRCA1*-mutated samples identified as such in the validation dataset 2 (Supplemental File S9B). However, some differences were also observed. The frequency by which genome wide CNAs are detected is again much higher in the *BRCA1*-called samples compared to the not called samples. This is most apparent for regions with copy number losses.

For 19 *BRCA1*-mutated samples H&E stained sections of paraffin-embedded tissue were available and scored for TILs. These samples were scored within the area outlined for micro-dissection. *BRCA1*-mutated samples with 9 or more *BRCA1*-associated CNAs were found to have a lower mean percentage of TILs of 27%, as compared to 41% TILs seen in the remaining *BRCA1*-mutated samples (not-called as *BRCA1*). Although a clear trend is seen, these results appeared not significantly different (Mann–Whitney; p = 0.101, supplemental File 10 panel B). Since large numbers of TILs are often seen at the periphery of the tumor samples where micro-dissection has taken place, it is often difficult to judge by the pathologist what the actual TIL percentages are in these samples. This likely contributes to a less stringent difference in TIL percentages.

3.6. *DNA flow cytometry and shallow whole genome sequencing*

To evaluate if *BRCA1*-mutated BLCs with high amounts of TILs harbor identical CNAs as those with low amount of TILs, five triple negative germ line *BRCA1*-mutated breast carcinomas with a TIL percentage of 40% or higher were selected for purification by flow cytometry. For all samples two cell suspensions from paraffin embedded tissue sections were prepared, one unsorted (comparable to what is analyzed in the tumors from the primary dataset), and one sorted into both a tumor cell fraction and a stromal/immune cell fraction. Sufficient DNA for all three fractions was obtained for three samples to be further analyzed by shallow whole genome sequencing. The resulting Log2 ratios for 166k data points were analyzed identically to the SNP array data analysis. Figure 4 shows copy number profiles obtained by shallow whole genome sequencing analyses for the sorted and unsorted fractions of the three *BRCA1*-mutated triple negative breast carcinomas. Supplemental File S11 shows microscopy images and vimentin versus keratin plots for the three FACS sorted samples.

As expected, the copy number profiles of the unsorted fractions show few chromosomal regions with copy number gains matching those regions with the highest mean copy numbers in the sorted tumor fraction profiles (chromosome 1q, 8q, 9p and 10p). These copy number profiles resemble the SNP array copy number profiles for the BLCs with the highest amounts of TILs. The sorted tumor fraction copy number profiles display many gains and losses of large chromosomal regions, and were highly similar to those seen by SNP array analysis for

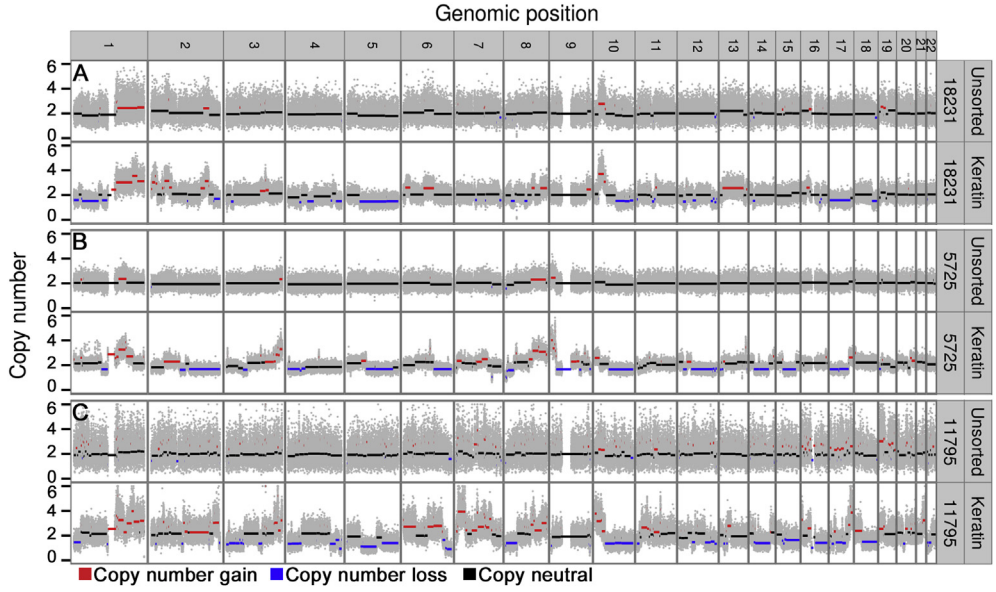

Figure 4 – Copy number profiles of Paraffin-embedded material of three *BRCA1*-mutated triple negative breast carcinomas profiled by shallow WGS. Called CNAs are shown in red (gain) and blue (losses). A, shallow WGS copy number plots of the unsorted cell fraction, and keratin positive cell fraction (tumor cells) for sample 18231. This sample was scored as 60% TIL. B, shallow WGS copy number plots of the unsorted cell fraction, and keratin positive cell fraction (tumor cells) for sample 5725. This sample was scored as 40% TIL. C, shallow WGS copy number plots of the unsorted cell fraction, and keratin positive cell fraction (tumor cells) for sample 11795. This sample was scored as 55% TIL.

the *BRCA1*-mutated BLCs with the lowest amounts of TILs. The only notable difference however is copy number loss of the entire long arm of chromosome 17 instead of copy number neutral/gain LOH at the *BRCA1* locus seen in the low TIL *BRCA1*-mutated samples by SNP array analysis.

Nine out of twelve *BRCA1*-associated CNAs were present in two of the samples. The remaining sample was found to have 6 *BRCA1*-associated CNAs. After merging the copy number profiles of the three unsorted and sorted tumor cell fractions with the SNP array copy number profiles, this sample (tumor fraction 18231) with lowest *BRCA1*-associated CNAs appeared to cluster together with high CIN luminal B carcinomas, within a predominantly BLC branch, in unsupervised hierarchical clustering of copy number data. Whereas the tumor fractions of the other two tumors cluster together with the bulk of the BLCs (Supplemental File S12).

4. Discussion

The aim of this study was to identify specific genomic characteristics of germ line *BRCA1*-mutated breast carcinomas. For this purpose, genome wide copy number, LOH, gene expression, and *BRCA1* promoter hyper-methylation data were analyzed for a familial cohort of 120 breast carcinomas. Since *BRCA1*-mutated breast carcinomas are primarily of the basal-like subtype and specific CNAs are known to be associated with this subtype (Bergamaschi et al., 2006; Jonsson et al., 2010; Smid et al., 2010), we restricted our main analyses to this intrinsic subtype. Based on mRNA analysis and subsequent assessment of TILs on H&E stained frozen sections, it appeared that approximately two thirds of the *BRCA1*-mutated BLCs have TIL percentages of 35% or higher, reaching 85% TILs. Similar observations are made for the BRCAX BLCs. These TIL percentages were initially underestimated when tumor cell percentages were established for these samples. Discrepancies between TIL percentages as scored by the pathologist and as measured by mRNA analysis is likely explained by the large heterogeneity regarding the distribution of TILs throughout the tumor tissue, which makes objective scoring very difficult. After sample selection for low levels of TILs, eight *BRCA1*-mutated and five BRCAX BLCs remained for further analysis. In spite of this small sample size, differential regions of CNAs could be readily identified between these tumor groups. Three of these regions were reported before; while five other regions were previously not considered as significantly associated with *BRCA1*-mutation status (Wessels et al., 2002; van Beers et al., 2005; Jonsson et al., 2005; Joosse et al., 2009; Prat et al., 2014). Moreover, the frequencies of recurring CNAs reach levels of 80–100%, much higher compared to the analysis of BLCs without the selection based on TIL percentages as well as literature reported observations. Exemplary for this is the *BRCA1*-mutated tumor specific loss of chromosome 15q11-21, which has been implicated in the loss of *TP53BP1* in *BRCA1* related oncogenesis (Bouwman et al., 2010). In literature, this region was found to be lost in approximately 35% of *BRCA1*-mutated tumors (Wessels et al., 2002; van Beers et al., 2005; Joosse et al., 2009). Our initial observations yielded similar results. However, after sample selection, this frequency increased to 100% in the *BRCA1*-mutated BLCs.

The eight differential CNAs proved to be capable of discriminating between the *BRCA1*-mutated and BRCAX BLCs with low amount of TILs. Together with the four most

recurring CNAs found in all BLCs, these regions proved capable of discriminating between *BRCA1*-mutated BLCs and breast carcinomas of other subtypes in our dataset as well. These results were validated in two independent cohorts, typically identifying one third of the *BRCA1*-mutated BLCs as such, i.e. without taking TILs into account. This is in line with our findings that approximately two thirds of the *BRCA1*-mutated BLCs exhibit extensive lymphocytic infiltration. Interestingly, in both validation-sets, a number of sporadic basal-like/triple negative breast carcinomas reported to have *BRCA1* promoter hyper-methylation are also scored as *BRCA1*-like. This suggests that at least a proportion of *BRCA1* promoter hyper-methylated breast carcinomas follow an oncogenic pathway leading to similar CNAs as observed in *BRCA1* germ line mutated BLCs.

The observed homogeneity of CNAs in *BRCA1*-mutated BLCs implies that *BRCA1* driven oncogenesis essentially leads to similar CNAs in most *BRCA1*-mutated BLCs. Obviously, the ability to measure these *BRCA1*-related CNAs depends on the tumor cell percentage of the sample, especially in relation to the quantity of TILs within these carcinomas. However, a biased selection could have occurred for a subgroup of *BRCA1*-mutated BLCs with low TILs characterized by specific CNAs favoring the escape of immune system responses. To address this we performed multi-parameter DNA flow cytometry analysis of three germ line *BRCA1*-mutated triple negative breast carcinomas which were selected for high amounts of TILs. The copy number profiles obtained by shallow whole genome sequencing (WGS) of the sorted tumor fractions were highly similar to those of *BRCA1*-mutated BLCs with low amounts of TILs, except for the striking observation that the sorted tumor fractions of the high TIL cases showed a loss of the complete long arm of chromosome 17, instead of copy neutral LOH or copy gain with LOH observed in the low TIL cases. This difference in acquiring complete *BRCA1* deficiency may thus be somehow linked to the immune response.

A well known difficulty in DNA-copy-number determination by aCGH and SNP array analyses is variable tumor purity due to normal DNA contamination. Other such problems arise from intra-tumor heterogeneity, aneuploidy/polyploidy and technical artifacts. For these reasons various methodologies have been proposed to deal with such issues (see e.g. Popova et al., 2009; Yau et al., 2010). However, such methods have not been applied in genomic profiling of *BRCA1*-mutated breast carcinomas. To assess the impact of such methodologies on our data we have analyzed all *BRCA1*-mutated BLCs with the Tumor Aberration Prediction Suite (TAPS) software package which is recommended for the identification of allele specific copy number aberrations in samples with low tumor percentages (Rasmussen et al., 2011). Overall, the resulting CNA profiles were found to be highly similar to our primary results (data not shown). Therefore we believe that *in silico* data 'cleaning' does not adequately resolve the effect of relative large quantities of TILs on genomic profiling, at least in our samples, and that therefore FACS methods to clean the biological signal are preferred. This is best seen by comparing the genomic profiles generated after FACS sorting and subsequent shallow whole genome sequencing analyses. The genomic profiles in the unsorted tumor fractions display, besides a few apparently low-level copy number gains, flat-line copy number profiles. Whereas the sorted tumor fractions show many copy number gains and losses (Figure 4).

In summary, we noticed that the effect of TILs on genomic profiling of tumors is often not taken into account potentially leading to misinterpretation of data, at least in breast cancer research focused on BLCs and therefore *BRCA1*-mutated breast tumors. Micro-dissection of tumor material can be used to exclude TILs located at the outer borders of the tumor cell mass. However, TILs located in between tumor cells cannot be removed in this way. Consequently, the use of genomic classifiers on tumor material to identify potential *BRCA1*-mutated patients, by for example Multiplex Ligation-dependent Probe Amplification (MLPA) (Lips et al., 2011b), or to evaluate *BRCA1* variants of unknown significance most likely suffers from the consequences of very high quantities of TILs in a sizeable fraction of patients. Care must be taken so that *BRCA1*-mutated breast carcinomas with high quantities of TILs are not misinterpreted.

5. Conclusion

In conclusion, the presence of large quantities of TILs in especially basal-like breast carcinomas (BLCs) impacts on copy number (and LOH) profiling by SNP, aCGH and shallow WGS. Based on differential CNAs between *BRCA1*-mutated and BRCAX BLCs with low amounts of TILs, known as well as novel *BRCA1* tumor specific CNAs were identified. A simple scoring method based on these CNAs was constructed and capable of identifying *BRCA1*-mutated basal-like/triple negative breast carcinomas with high specificity. A similar analysis with a larger sample set in combination with a machine learning approach could fine-tune these results. In addition, future studies may give insight into the differences between *BRCA1*-mutated BLCs with differential mechanisms for losing the wild-type *BRCA1*-allele, i.e. either by copy neutral/gain LOH or copy loss. Such studies should preferably include a selection step specifically aimed at selecting for samples with low quantities of TILs, or otherwise diminish the effect on genomic profiling of these cells for example by multi-parameter DNA flow cytometry as shown in this study and other studies performed on cervical and colorectal carcinomas (Middeldorp et al., 2008; Mehta et al., 2009). Also, this approach will most likely prove valuable in next generation sequencing studies on tumor material with high number of TILs. The CNAs specific for *BRCA1*-mutated BLCs will facilitate the identification of genes contributing to *BRCA1* driven oncogenesis, which could ultimately lead to a better understanding of the etiology of *BRCA1* tumors and possibly to novel therapeutic options.

Authors' contributions

Conception and design of the study was by HMH and JAF. MPGM, SEM, ESJ, NA, DMB, DS, BY, JPV, CHMD, JWM, AMS, AMWO, VW, MS, ED, and PMN participated in acquisition of materials and data. Analysis and interpretation of the data was performed by MPGM, IEK, SEM, ESJ, DS, BY, and QW. The manuscript was drafted by MPGM, IEK, and QW and

critically revised by ESJ, NA, JCD, BY, MS, and PMN. All authors read and approved the final manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

Funding

The public foundation that supported the study had no role in the design and conduct of the study; in the collection, management, analysis, and interpretation of the data; or in data preparation, review, or approval of the manuscript.

Acknowledgments

Funded by the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research NWO (91756341). Hanne Meijers-Heijboer is a fellow from the NWO Vidi Research Program.

We would like to thank Dr. W. Corver (Dept. of Pathology, LUMC, Leiden, The Netherlands) for technical assistance on DNA flow cytometry.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.molonc.2014.12.012.

REFERENCES

Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., et al., 2012. An integrated map of genetic variation from 1,092 human genomes. Nature 491, 56–65.

Bergamaschi, A., Kim, Y.H., Wang, P., Sorlie, T., Hernandez-Boussard, T., Lonning, P.E., Tibshirani, R., et al., 2006. Distinct patterns of DNA copy number alteration are associated with different clinicopathological features and gene-expression subtypes of breast cancer. Genes Chromosomes Cancer 45, 1033–1040.

Bernstein, B.E., Birney, E., Dunham, I., Green, E.D., Gunter, C., Snyder, M., 2012. An integrated encyclopedia of DNA elements in the human genome. Nature 489, 57–74.

Birgisdottir, V., Stefansson, O.A., Bodvarsdottir, S.K., Hilmarsdottir, H., Jonasson, J.G., Eyfjord, J.E., 2006. Epigenetic silencing and deletion of the BRCA1 gene in sporadic breast cancer. Breast Cancer Res. 8, R38.

Bouwman, P., Aly, A., Escandell, J.M., Pieterse, M., Bartkova, J., van der Gulden, H., Hiddingh, S., et al., 2010. 53BP1 loss rescues BRCA1 deficiency and is associated with triple-negative and BRCA-mutated breast cancers. Nat. Struct. Mol. Biol. 17, 688–695.

Corver, W.E., Ter Haar, N.T., Dreef, E.J., Miranda, N.F., Prins, F.A., Jordanova, E.S., Cornelisse, C.J., et al., 2005. High-resolution multi-parameter DNA flow cytometry enables detection of tumor and stromal cell subpopulations in paraffin-embedded tissues. J. Pathol. 206, 233–241.

Corver, W.E., ter Haar, N.T., 2011. High-resolution multiparameter DNA flow cytometry for the detection and sorting of tumor and stromal subpopulations from paraffin-embedded tissues. Curr. Protoc. Cytom (Chapter 7):Unit 7.37.

Curtis, C., Shah, S.P., Chin, S.F., Turashvili, G., Rueda, O.M., Dunning, M.J., Speed, D., et al., 2012. The genomic and transcriptomic architecture of 2,000 breast tumors reveals novel subgroups. Nature 486, 346–352.

Foulkes, W.D., Stefansson, I.M., Chappuis, P.O., Begin, L.R., Goffin, J.R., Wong, N., Trudel, M., et al., 2003. Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer. J. Natl. Cancer Inst. 95, 1482–1485.

Futreal, P.A., Liu, Q., Shattuck-Eidens, D., Cochran, C., Harshman, K., Tavtigian, S., Bennett, L.M., et al., 1994. BRCA1 mutations in primary breast and ovarian carcinomas. Science 266, 120–122.

Gonzalez-Angulo, A.M., Timms, K.M., Liu, S., Chen, H., Litton, J.K., Potter, J., Lanchbury, J.S., et al., 2011. Incidence and outcome of BRCA mutations in unselected patients with triple receptor-negative breast cancer. Clin. Cancer Res. 17, 1082–1089.

Honrado, E., Osorio, A., Palacios, J., Milne, R.L., Sanchez, L., Diez, O., Cazorla, A., et al., 2005. Immunohistochemical expression of DNA repair proteins in familial breast cancer differentiate BRCA2-associated tumors. J. Clin. Oncol. 23, 7503–7511.

Huang, d.W., Sherman, B.T., Lempicki, R.A., 2009a. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 37, 1–13.

Huang, D.W., Sherman, B.T., Lempicki, R.A., 2009b. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4, 44–57.

Jonsson, G., Naylor, T.L., Vallon-Christersson, J., Staaf, J., Huang, J., Ward, M.R., Greshock, J.D., et al., 2005. Distinct genomic profiles in hereditary breast tumors identified by array-based comparative genomic hybridization. Cancer Res. 65, 7612–7621.

Jonsson, G., Staaf, J., Vallon-Christersson, J., Ringner, M., Holm, K., Hegardt, C., Gunnarsson, H., et al., 2010. Genomic subtypes of breast cancer identified by array-comparative genomic hybridization display distinct molecular and clinical characteristics. Breast Cancer Res. 12, R42.

Joosse, S.A., van Beers, E.H., Tielen, I.H., Horlings, H., Peterse, J.L., Hoogerbrugge, N., Ligtenberg, M.J., et al., 2009. Prediction of BRCA1-association in hereditary non-BRCA1/2 breast carcinomas with array-CGH. Breast Cancer Res. Treat 116, 479–489.

Lakhani, S.R., van de Vijver, M.J., Acquemier, J., Anderson, T.J., Osin, P.P., McGuffog, L., Easton, D.F., 2002. The pathology of familial breast cancer: predictive value of immunohistochemical markers estrogen receptor, progesterone receptor, HER-2, and p53 in patients with mutations in BRCA1 and BRCA2. J. Clin. Oncol. 20, 2310–2318.

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

Lips, E.H., Mulder, L., Hannemann, J., Laddach, N., Vrancken Peeters, M.T., van de Vijver, M.J., Wesseling, J., et al., 2011a. Indicators of homologous recombination deficiency in breast cancer and association with response to neoadjuvant chemotherapy. Ann. Oncol. 22, 870–876.

Lips, E.H., Laddach, N., Savola, S.P., Vollebergh, M.A., Oonk, A.M., Imholz, A.L., Wessels, L.F., Wesseling, J., Nederlof, P.M., Rodenhuis, S., 2011b. Quantitative copy number analysis by Multiplex Ligation-dependent Probe Amplification (MLPA) of BRCA1-associated breast cancer regions identifies BRCAness. Breast Cancer Res. 13, R107.

Mehta, A.M., Jordanova, E.S., Corver, W.E., van Wezel, T., Uh, H.W., Kenter, G.G., Jan Fleuren, G., 2009. Single nucleotide polymorphisms in antigen processing machinery component ERAP1 significantly associate with clinical outcome in cervical carcinoma. Genes Chromosomes Cancer 48, 410–418.

Middeldorp, A., van Puijenbroek, M., Nielsen, M., Corver, W.E., Jordanova, E.S., ter Haar, N., et al., 2008. High frequency of copy-neutral LOH in MUTYH-associated polyposis carcinomas. J. Pathol. 216, 25–31.

Nagel, J.H., Peeters, J.K., Smid, M., Sieuwerts, A.M., Wasielewski, M., de Weerd, V., Trapman-Jansen, A.M., van den Ouweland, A., et al., 2012. Gene expression profiling assigns CHEK2 1100delC breast cancers to the luminal intrinsic subtypes. Breast Cancer Res. Treat 132, 439–448.

Perou, C.M., Sorlie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., Rees, C.A., Pollack, J.R., et al., 2000. Molecular portraits of human breast tumors. Nature 406, 747–752.

Prat, A., Cruz, C., Hoadley, K.A., Diez, O., Perou, C.M., Balmana, J., 2014. Molecular features of the basal-like breast cancer subtype based on BRCA1 mutation status. Breast Cancer Res. Treat 147, 185–191.

Popova, T., Manie, E., Stoppa-Lyonnet, D., Rigaill, G., Barillot, E., Stern, M.H., 2009. Genome Alteration Print (GAP): a tool to visualize and mine complex cancer genomic profiles obtained by SNP arrays. Genome Biol. 10, R128.

Rasmussen, M., Sundstrom, M., Goransson, K.H., Botling, J., Micke, P., Birgisson, H., Glimelius, B., Isaksson, A., 2011. Allele-specific copy number analysis of tumor samples with aneuploidy and tumor heterogeneity. Genome Biol. 12, R108.

Roy, R., Chun, J., Powell, S.N., 2012. BRCA1 and BRCA2: different roles in a common pathway of genome protection. Nat. Rev. Cancer 12, 68–78.

Rubinstein, W.S., 2008. Hereditary breast cancer: pathobiology, clinical translation, and potential for targeted cancer therapeutics. Fam. Cancer 7, 83–89.

Schouten, P.C., van Dyk, E., Braaf, L.M., Mulder, L., Lips, E.H., de Ronde, J.J., Holtman, L., et al., 2013. Platform comparisons for identification of breast cancers with a BRCA-like copy number profile. Breast Cancer Res. Treat 139, 317–327.

Silver, D.P., Richardson, A.L., Eklund, A.C., Wang, Z.C., Szallasi, Z., Li, Q., Juul, N., et al., 2010. Efficacy of neoadjuvant cisplatin in triple-negative breast cancer. J. Clin. Oncol. 28, 1145–1153.

Smeets, S.J., Harjes, U., van Wieringen, W.N., Sie, D., Brakenhoff, R.H., Meijer, G.A., Ylstra, B., 2011. To DNA or not to DNA? That is the question, when it comes to molecular subtyping for the clinic! Clin. Cancer Res. 17, 4959–4964.

Smid, M., Wang, Y., Zhang, Y., Sieuwerts, A.M., Yu, J., Klijn, J.G., Foekens, J.A., et al., 2008. Subtypes of breast cancer show preferential site of relapse. Cancer Res. 68, 3108–3114.

Smid, M., Hoes, M., Sieuwerts, A.M., Sleijfer, S., Zhang, Y., Wang, Y., Foekens, J.A., et al., 2010. Patterns and incidence of chromosomal instability and their prognostic relevance in breast cancer subtypes. Breast Cancer Res. Treat 128, 23–30.

Sorlie, T., Tibshirani, R., Parker, J., Hastie, T., Marron, J.S., Nobel, A., Deng, S., et al., 2003. Repeated observation of breast tumor subtypes in independent gene expression datasets. Proc. Natl. Acad. Sci. USA 100, 8418–8423.

Staaf, J., Lindgren, D., Vallon-Christersson, J., Isaksson, A., Goransson, H., Juliusson, G., Rosenquist, R., et al., 2008. Segmentation-based detection of allelic imbalance and loss-of-heterozygosity in cancer cells using whole genome SNP arrays. Genome Biol. 9, R136.

Stefansson, O.A., Jonasson, J.G., Johannsson, O.T., Olafsdottir, K., Steinarsdottir, M., Valgeirsdottir, S., Eyfjord, J.E., 2009. Genomic profiling of breast tumors in relation to BRCA abnormalities and phenotypes. Breast Cancer Res. 11, R47.

Tirkkonen, M., Johannsson, O., Agnarsson, B.A., Olsson, H., Ingvarsson, S., Karhu, R., Tanner, M., et al., 1997. Distinct somatic genetic changes associated with tumor progression in carriers of BRCA1 and BRCA2 germ-line mutations. Cancer Res. 57, 1222–1227.

Turner, N., Tutt, A., Ashworth, A., 2004. Hallmarks of 'BRCAness' in sporadic cancers. Nat. Rev. Cancer 4, 814–819.

Turner, N.C., Reis-Filho, J.S., 2006. Basal-like breast cancer and the BRCA1 phenotype. Oncogene 25, 5846–5853.

van Beers, E.H., van Welsem, T., Wessels, L.F., Li, Y., Oldenburg, R.A., Devilee, P., Cornelisse, C.J., et al., 2005. Comparative genomic hybridization profiles in human BRCA1 and BRCA2 breast tumors highlight differential sets of genomic aberrations. Cancer Res. 65, 822–827.

van Thuijl, H.F., Ylstra, B., Wurdinger, T., van Nieuwenhuizen, D., Heimans, J.J., Wesseling, P., Reijneveld, J.C., 2013. Genetics and pharmacogenomics of diffuse gliomas. Pharmacol. Ther. 137, 78–88.

Waddell, N., Arnold, J., Cocciardi, S., da Silva, L., Marsh, A., Riley, J., Johnstone, C.N., et al., 2010. Subtypes of familial breast tumors revealed by expression and copy number profiling. Breast Cancer Res. Treat 123, 661–677.

Wei, M., Grushko, T.A., Dignam, J., Hagos, F., Nanda, R., Sveen, L., Xu, J., et al., 2005. BRCA1 promoter methylation in sporadic breast cancer is associated with reduced BRCA1 copy number and chromosome 17 aneusomy. Cancer Res. 65, 10692–10699.

Wessels, L.F., van Welsem, T., Hart, A.A., van't Veer, L.J., Reinders, M.J., Nederlof, P.M., 2002. Molecular classification of breast carcinomas by comparative genomic hybridization: a specific somatic genetic profile for BRCA1 tumors. Cancer Res. 62, 7110–7117.

Wickham, H., 2009. ggplot2: Elegant Graphics for Data Analysis. Springer, New York.

Yau, C., Mouradov, D., Jorissen, R.N., Colella, S., Mirza, G., Steers, G., Harris, A., et al., 2010. A statistical approach for detecting genomic aberrations in heterogeneous tumor samples from single nucleotide polymorphism genotyping data. Genome Biol. 11, R92.