Robust BRCA1-like classification of copy number profiles of samples repeated across different datasets and platforms


Departments of Molecular Pathology, Netherlands Cancer Institute, Amsterdam, The Netherlands
Breakthrough Breast Cancer Research Unit, Department of Research Oncology, Guy’s Hospital, King’s College London School of Medicine, London, United Kingdom
Division of Molecular Oncology, Netherlands Cancer Institute, Amsterdam, The Netherlands
Division of Molecular Oncology, Amsterdam, The Netherlands
Cancer Research UK Cambridge Research Institute, Li Ka Shing Centre, Cambridge, UK
Department of Medical Oncology, University Medical Center Utrecht, Utrecht, The Netherlands
Netherlands Center for Personalized Cancer Treatment, Utrecht, The Netherlands
Division of Biological Stress Response, Netherlands Cancer Institute, Amsterdam, The Netherlands
The Breakthrough Breast Cancer Research Centre, The Institute of Cancer Research, London, UK
Genomics Core Facility, Netherlands Cancer Institute, Amsterdam, The Netherlands
Department of Pathology, Netherlands Cancer Institute, Amsterdam, The Netherlands
Department of Oncology, University of Cambridge, Addenbrooke’s Hospital, Cambridge, UK
Cambridge Experimental Cancer Medicine Centre and NIHR Cambridge Biomedical, Research Centre, Cambridge University Hospitals NHS, Cambridge, UK
Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Delft, The Netherlands
Department of Pathology, University Medical Center Utrecht, Utrecht, The Netherlands
Division of Medical Oncology, Netherlands Cancer Institute, Amsterdam, The Netherlands

Abbreviations: aCGH, array Comparative Genomic Hybridization; BAC, Bacterial Artificial Chromosome; BAC32K, Bacterial Artificial Chromosome aCGH, 32K platform; BAC3K, Bacterial Artificial Chromosome aCGH, 3K platform; BRCA1, Breast Cancer Early Onset 1; CN, Copy number; DNA, Deoxyribonucleic acid; dsDNA, double-stranded DNA; FFPE, Formalin Fixed Paraffin Embedded; hg 18, human reference genome version 18; hg19, human reference genome version 19; MIP, Molecular Inversion Probe; NG135, Nimblegen 135k oligonucleotide aCGH; NG720, Nimblegen 720K oligonucleotide aCGH; NGS, Low coverage next generation sequencing; SNP6, Affymetrix SNP6 array; SNR, Signal to Noise Ratio; VN, Variance of the Noise.
* Corresponding author. Departments of Medical Oncology and Molecular Pathology at Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands. Tel.: +31 20 512 2951.
E-mail address: s.linn@nki.nl (S.C. Linn).
http://dx.doi.org/10.1016/j.molonc.2015.03.002
1574-7891/© 2015 Published by Elsevier B.V. on behalf of Federation of European Biochemical Societies.

Please cite this article in press as: Schouten, P.C., et al., Robust BRCA1-like classification of copy number profiles of samples repeated across different datasets and platforms, Molecular Oncology (2015), http://dx.doi.org/10.1016/j.molonc.2015.03.002
Breast cancers with BRCA1 germline mutation have a characteristic DNA copy number (CN) pattern. We developed a test that assigns CN profiles to be 'BRCA1-like' or 'non-BRCA1-like', which refers to resembling a BRCA1-mutated tumor or resembling a tumor without a BRCA1 mutation, respectively. Approximately one third of the BRCA1-like breast cancers have a BRCA1 mutation, one third has hypermethylation of the BRCA1 promoter and one third has an unknown reason for being BRCA1-like. This classification is indicative of patients’ response to high dose alkylating and platinum containing chemotherapy regimens, which targets the inability of BRCA1 deficient cells to repair DNA double strand breaks. We investigated whether this classification can be reliably obtained with next generation sequencing and copy number platforms other than the bacterial artificial chromosome (BAC) array Comparative Genomic Hybridization (aCGH) on which it was originally developed.

We investigated samples from 230 breast cancer patients for which a CN profile had been generated on two to five platforms, comprising low coverage CN sequencing, CN extraction from targeted sequencing panels (Copywriter), Affymetrix SNP6.0, 135K/720K oligonucleotide aCGH, Affymetrix Oncoscan FFPE (MIP) technology, 3K BAC and 32K BAC aCGH. Pairwise comparison of genomic position-mapped profiles from the original aCGH platform and other platforms revealed concordance. For most cases, biological differences between samples exceeded the differences between platforms within one sample. We observed the same classification across different platforms in over 80% of the patients and kappa values of at least 0.36. Differential classification could be attributed to CN profiles that were not strongly associated to one class. In conclusion, we have shown that the genomic regions that define our BRCA1-like classifier are robustly measured by different CN profiling technologies, providing the possibility to retro- and prospectively investigate BRCA1-like classification across a wide range of CN platforms.

© 2015 Published by Elsevier B.V. on behalf of Federation of European Biochemical Societies.

1. Introduction

Breast cancer arising in patients with a germline BRCA1 mutation are thought to be genomically unstable due to the impairment of error-free homologous recombination DNA repair in which BRCA1 has a role (Venkitaraman, 2009; Vollebergh et al., 2012). DNA copy number (CN) profiles provide a snapshot of a result of genomic instability in cancer, namely the CN aberrations. The copy number profiles of patients with a BRCA1 mutation have specific gains and losses (Alvarez et al., 2005; Tirkkonen et al., 1997; Wessels et al., 2002). We previously developed a shrunken centroids classifier which uses 371 genomic regions to assign a CN profile to the BRCA1-like (sharing characteristics of BRCA1 mutated breast cancer) or non-BRCA1-like phenotype (Vollebergh et al., 2011). This classifier not only identifies germline BRCA1-mutated cases (approximately 1/3 of the BRCA1-like tumors) but also enriches for tumors with other mechanisms of BRCA1 inactivation, for example promoter hypermethylation (approximately 1/3 of the BRCA1-like tumors, mutually exclusive with BRCA1 mutation) (Joosse et al., 2011; Vollebergh et al., 2011; Lips et al., 2011) which can confer to non-familial cases a tumor phenotype that is similar to BRCA1 mutation carriers. Alternative modes of BRCA1 inactivation and similarity of these tumors to BRCA1-mutated tumors have been observed in other datasets as well (Turner et al., 2004; Esteller et al., 2000; Alvarez et al., 2005; Tung et al., 2010; Cancer Genome Atlas Network, 2012) and has been referred to as ‘BRCaness’ (Turner et al., 2004). The cases with unknown cause for being classified as BRCA1-like may thus be subject to BRCA1 dysfunction due to yet unidentified causes, or reflect a broader pathway dysfunction. Subsequently, we demonstrated that BRCA1-like patients benefit significantly more from high dose DNA double strand break-inducing chemotherapy, containing both platinum and alkylating agents, than from a conventional second generation chemotherapy regimen (Vollebergh et al., 2011). Two follow-up studies with different chemotherapy regimens demonstrated that BRCA1-like patients benefit also from tandem high dose (both including alkylating agents, one including platinum) compared to conventional, and from tandem high dose compared to dose dense chemotherapy, underlining the clinical relevance of the BRCA1-like profile (Schouten et al., 2015, 2014, 2013b). Technological advances in experimental platforms have provided many datasets to study BRCA1-like profiles next to those generated on the original BAC (BAC3K) platform and 135k oligonucleotide aCGH (NGI135), on which we reported in a previous manuscript (Schouten et al., 2013a). Given this reported reproducibility between different CN profiling platforms, we investigated whether BRCA1-like classification of CN profiles of repeated samples could be reliably obtained across multiple platforms (Baumbusch et al., 2008; Curtis et al., 2009; Hester et al., 2009; Krijgsman et al., 2012; Schouten et al., 2013a; Wicker et al., 2007). For this study we compared data from samples for which data from at least two of the following platforms were available: low coverage genome-wide sequencing,
targeted sequencing panels (extracted with the CopywriteR algorithm, Kuilman et al., 2015), Affymetrix SNP6.0 arrays (SNP6), Nimblegen 720k (NG720) oligonucleotide aCGH, Affymetrix Oncoscan molecular inversion probe (MIP) technology, 3K (BAC3K) and 32K BAC aCGH (BAC32K) We investigated whether these alternative methods can be used to obtain copy number profiles suitable for reliable and accurate BRCA1-like classification, as defined by being similar to the original BAC aCGH-based classification.

2. Methods

2.1. Samples

We investigated 5 cohorts of patients: 1) 118 FFPE DNA samples from a Dutch randomized controlled clinical trial dataset comparing high dose chemotherapy with conventional chemotherapy (termed ‘N4+’; Rodenhuis et al., 2003; Vollebergh et al., 2011); 2) 27 fresh frozen samples from the EU FP7 RATHER project (termed ‘RATHER’, http://www.rather-project.com/); 3) A cohort of 11 samples (5 HER2+ and 6 TN patients) for which both FFPE and fresh frozen tissue was available from the Breakthrough Breast Cancer Research Unit, King’s College London, UK, termed ‘KCL’; 4) A cohort of 76 FFPE DNA samples from BRCA1 and -2 mutated breast cancer samples and sporadic controls (Joosse et al., 2012, 2011, 2009) termed ‘BC’; 5) a cohort of triple negative patients treated with neo-adjuvant chemotherapy termed ‘NAC’ (Lips et al. submitted). Tissue was used according to national guidelines regarding the use of archival material and with approval of the respective medical ethical review committees.

2.2. DNA isolation

N4+ samples and BC samples: Formalin fixed Paraffin Embedded (FFPE) sections were macrodissected to contain at least 60% tumor cells and isolated with the Qiagen DNA mini kit as described previously (Vollebergh et al., 2011).

KCL samples: FFPE sections of tumor were microdissected to achieve a minimum of 70% composition of tumor cells, and DNA was extracted using the DNeasy Kit (Qiagen Ltd, Crawley, UK) according to the manufacturer’s recommendations. DNA from the fresh frozen tumor samples were extracted with the DNeasy kits (Qiagen, Hilden, Germany) using the manufacturer’s protocols.

RATHER samples: DNA was isolated from fresh frozen tumor samples containing at least 30% tumor cells and DNA was isolated using the DNeasy Blood & Tissue kit (Qiagen, Hilden, Germany). NAC samples: DNA was isolated from Fresh Frozen sections containing at least 50% tumor cells with the Qiagen DNA mini kit.

2.3. Micro-array copy number profiling and data processing

All copy number profiling and processing was performed as described in previous publications unless further specified (Buffart et al., 2008; Curtis et al., 2012; Joosse et al., 2012, 2011, 2009; Natrajan et al., 2014, 2009; Schouten et al., 2013a; Vollebergh et al., 2014, 2011; Wang et al., 2005). The unsegmented data (i.e. raw pre-processed data, according to established methods per platform) were used as input in the analysis. Table 1 refers to the respective references for the individual data and platforms. Summarizing, these steps included labeling, hybridization, scanning and converting images to background-corrected log2 ratios or copy number estimates (MIP). DNA was hybridized to Affymetrix SNP 6.0 arrays (SNP6) as described before for RATHER samples (Curtis et al., 2012). Processing of KCL samples with Affymetrix SNP6.0 arrays was outsourced to Atlas Biolabs GmbH (Berlin, Germany) and standard manufacturer protocols were followed for the amplification, hybridisation, washing, and scanning of the samples hybridization. R package “aroma.affymetrix” was used for the preprocessing of the Affymetrix SNP6.0 data (Bengtsson et al., 2008).

2.4. Low coverage copy number sequencing and data processing

The amount of double stranded DNA in genomic DNA samples was quantified using the Qubit® dsDNA HS Assay Kit (Invitrogen). Up to 250 ng of double stranded genomic DNA was fragmented by Covaris shearing to obtain fragment sizes of 160–180 bp. Samples were purified with the Agencourt AMPure XP PCR Purification beads according to manufacturer’s instructions (Beckman Coulter, cat no A63881). DNA library preparation for Illumina sequencing was done with the
TruSeq® DNA LT Sample Preparation kit (Illumina). The double-stranded DNA input amount was lower than advised by the TrueSeq protocol, so we used up to 250 ng of double-stranded DNA, such that 2.5 times less adapter concentration was used than prescribed in the TrueSeq protocol. During enrichment PCR, 10 cycles were necessary to obtain enough yield for sequencing. All DNA libraries were analyzed on a BioAnalyzer system (Agilent Technologies) using the DNA7500 chips for determining the molarity. Up to ten uniquely indexed samples were pooled equimolarly to give a final concentration of 10 nM. Pools were then sequenced using an Illumina HiSeq2000 machine to a coverage of 5.5×. This was done in one lane of a single-end 50 bp run according to manufacturer’s instructions.

Reads were aligned to the reference genome (hg19) using the BWA backtrack algorithm (Li and Durbin, 2009) and counted in tumour's instructions. In one lane of a single-end 50 bp run according to manufacturer's instructions.

2.5. Targeted sequencing and data processing

Three ug of DNA from N4+ samples was used to prepare paired-end fragment libraries using a genomic DNA library preparation kit (Illumina). The libraries were hybridized to a SureSelect custom-based bait library (Agilent) enriching for 565 genes involved in DNA repair and cancer ("DNA repairome"). After washing the captured DNA was amplified. Enriched libraries were barcoded, pooled and sequenced on an Illumina HiSeq 2000 machine using a 2×75 bp paired-end protocol. Reads were filtered for quality and aligned to the human genome (GRCh37/hg19) using Samtools.

Bar-coded sequence libraries for the NAC samples were generated based on (Vermaat et al., 2012). 300–600 ng of input DNA was used (Harakalova et al., 2011). These pools were enriched for 1977 ("Cancer mini-genome") cancer-related genes using SureSelect technology. Enriched libraries were sequenced on a SOLiD 5500x1 instrument according to the manufacturer’s protocol. Variant calling was done using a custom pipeline as described in (Lips et al., submitted).

Sequence reads were mapped on the human reference genome version 19 (GRCh37), using BWA (Li and Durbin, 2009). To obtain copy number profiles from these targeted reads we used the CopywriteR tool (Kuilman et al. 2015, https://github.com/PeeperLab/CopywriteR). In brief, this tool extracts the off-target reads obtained with targeted sequencing and uses these for copy number detection. The reads were then mapped to the BAC clone regions, and subsequently corrected for GC content and mappability and filtered for CNV regions as described above with the exception that mappability was corrected using a loess.

2.6. Mapping and BRCA1-like classification

The BRCA1-like classifier was originally trained on unsegmented BAC3K aCGH copy number profiles (Joosse et al., 2009; Vollebergh et al., 2011). The BRCA1-like classifier is a shrunk centroid classifier based on 371 (out of 3277) BAC clones (Vollebergh et al., 2011). For each platform we mapped raw copy number data-points to the3277 BAC clones.

BAC3K, BAC32K, NG135, NG720, and SNP6 data were obtained as log2 copy number ratio; NGS, and targeted sequencing data were log2 read counts; MIP data was obtained as continuous copy number estimate (i.e. no ratio or log2). The MIP data was log2 transformed and subtracted by 1 to obtain 0-centered log2 values. We subsequently mapped these log2 ratio/value profiles to the original BAC3K aCGH platform on genome version hg18. Subsequently, we averaged the log2 ratios/values that fall within the chromosomal start and end position of the BAC clones (Schouten et al., 2013a). We used custom functions using the functionality from the following R packages in the mapping process: DNAcopy (Venkatraman and Oshen, 2007), cghseg (Picard et al., 2011), Genomic Ranges (Aboyoun et al., 2013), KCsmart (De Ronde et al., 2009) and, copy number (Nilsen et al., 2012). The median BAC size was approximately 150 kb and the median number of probes averaged 3 for BAC32K (range 1–12), 6 for NG135 (range 1–30), 36 for NG720 (range 1–107), 92 for SNP6 (range 1–328), 15 for MIP (range 1–215) and the median number of 20 KB bins averaged for NGS was 8 (range 1–15). Missing BAC clones were filled by linear interpolation of the surrounding probes to obtain the ‘mapped profile’. This mapped profile was classified to be BRCA1-like or non-BRCA1-like as described previously and used for all other further analyses (Schouten et al., 2013a; Vollebergh et al., 2011).

2.7. Statistical analysis

To evaluate the quality of each sample and to exclude low quality CN profiles we employed two statistical measures; ‘variance of the noise’ (VN) and ‘signal to noise ratio’ (SNR). The VN is defined as the variance between the processed signal (segmented copy number value) and the unprocessed signal (raw copy number value). The signal to noise is defined as the variance of the biological signal (log2 ratio of the segmented value) divided by the noise, as measured by the VN.

Profiles that had less signal than noise (SNR < 1) and high noise (VN > 0.025, as obtained from the density plot, Supplementary Figure 2) were considered low quality and excluded from the analysis. The similarity of samples analyzed by two platforms was visually assessed by plotting the average profile for each platforms. Hierarchical clustering was performed with a distance measure of 1-Pearson correlation and ward linkage. Subsequently, we checked whether repeated samples from the same patient clustered together.
For those samples we had BRCA1 methylation or mutation data available we calculated the sensitivity and the proportion of mutated/methylated samples in the BRCA1-like group. We calculated the inter-rater agreement between repeated samples using two measures and their respective confidence intervals: 1) the statistical accuracy values defined as the percentage of samples on the diagonal of the cross table of BRCA1-like status on one platform vs. the other platform and 2) Cohen’s kappa value (R package epiR (Stevenson, 2012)). We used Table1Heatmap for plotting (Schouten, 2014). Cohen’s kappa can be interpreted as follows: 0–0.4: poor agreement between tests; 0.4–0.8: moderate agreement between tests; and >0.8: near-perfect agreement (Schouten et al., 2013a). We calculated strength-of-classification and its standard deviation. Strength-of-classification is the Euclidean distance to the closest class. The value of this measure increases when a sample is closer and thus more strongly assigned to the class profile. 0 indicates that a sample is equally close to both classes. For example, if the Euclidean distance between a sample and the BRCA1-like average profile is 0.6, and the Euclidean distance between that sample and the non-BRCA1-like class is 0.9, the closest class is BRCA1-like and the measure is 0.9–0.6 = 0.3. If for another sample the distance between a sample and the respective classes are 0.75 and 0.8, the measure would be 0.05, indicating less strong favor for any of the classes. Second, we used the standard deviation of the strength-of-classification. The larger the standard deviation the more likely a difference in classification.

All analyses were performed with R version 3.0.2.

3. Results

To establish the robustness of our BRCA1-like classifier on multiple CN platforms we used breast cancer samples that were analyzed by at least two genomic profiling platforms. The classifier, which was originally developed on a BAC aCGH platform, was tested on 263 tumor samples. The samples were analyzed using seven different technologies, with the overlap per technology ranging between zero (some platforms had no overlap) and 173 (NGS versus BAC3K) tumors. This resulted in 616 CN profiles, with 198 tumors overlapping between two, 43 between three, 19 between four, and two between five technologies. Forty profiles had an SNP smaller than 1 and a VN larger than 0.025 and were therefore removed, with another 31 profiles that lost a counterpart on another platform, resulting in 545 CN profiles spread over 230 patients. Table 1 describes the total number of profiles overlapping between two platforms and the number of profiles after quality control.

3.1. Mapped profiles resemble original profiles and biological signal overrules platform-specific characteristics

For every platform, we mapped the CN data to the BAC3K aCGH locations by averaging the log2 ratios of positions overlapping each BAC clone, and investigated both the genome-wide and classifier region specific similarity between two platforms. We calculated the average genome-wide profile of samples that overlap per technology. Visual inspection revealed high concordance between, segmented CN profiles (Figure 1), unsegmented CN profiles and CN profiles limited to the 371 classifier regions (Supplementary Figures 3 and 4) The distributions of the measurements were similar (Supplementary Figure 5). MIP, NGS and CopywriteR data demonstrated a larger dynamic range while Affymetrix SNP6 data displayed a compressed dynamic range. Within-sample variation (clustering by patient) is smaller than between-sample variation (clustering by technology) as is observed from hierarchical clustering analysis (Supplementary Figure 6). In conclusion, we observed high similarity between CN profiles after filtering low-quality genomic profiles and reducing dimension by averaging measurements that fall within a BAC clone. This is independent of the technology used.

3.2. BRCA1-like classification of mapped CN profiles is highly concordant with gold standard BAC3K classification

Having established the similarity between repeated samples from different datasets we investigated whether the minor differences in CN profiles would influence sample classification. We therefore performed a comparison between the tumor classification based on the original BAC3K profiles and profiles of the same tumors profiled on all other platforms. The class labels obtained from the BAC3K classifier served as gold standard. We then calculated the classification accuracy (how well does a classifier on another platform reproduce the BAC3K labels) and Cohen’s kappa (what is the concordance between the BAC3K labels and the labels from another platform (Table 2).

The Cohen’s kappa values between classification with mapped and original profiles ranged from 0.36 (BAC3K vs. CopywriteR) to 1 (BAC3K vs. BAC32K/NG720). With the exception of the MIP and CopywriteR results all Cohen’s Kappa values are close to or above 0.8 indicating almost perfect concordance. However, some datasets are small resulting in wide confidence intervals (BAC3K vs. BAC32K), and some datasets do not have any samples classified differently, suggesting a potential bias of having only good quality and/or very concordant profiles in this particular analysis (BAC3K/NG720). A less stringent measure than Cohen’s Kappa is the accuracy. For all technologies the percentage of samples that classify identically as the original profile was over 80%.

3.3. BRCA1-like classification is highly concordant with consensus classification across datasets

Although BAC3K is considered as the gold standard we should note that 1) the platform is not in operation anymore and 2) the gold standard classification may have been based on lower quality CN profiles (see below, section 3.7). Since we demonstrated high agreement for BRCA1-like classification of samples using the original BAC3K aCGH classification as a reference we investigated the overall agreement for each patient with all available data (Figure 2). In this analysis, we compared the class assigned with data from a particular platform to the class assigned based on the profiles from all other platforms.
Figure 1 — Average copy number profiles compared between two technologies. Comparison of average of all samples based on their segmented copy number profiles showing the genomic position on the x-axis and the average log2 ratio on the y-axis. Original profiles are plotted in black and mapped profiles in red. A) BAC vs. MIP segmented; B: BAC vs. NG135 segmented; C: BAC vs NG720 segmented; D: BAC vs. NGS segmented; E: BAC3K vs BAC32K segmented; F: NG135 vs. SNP6 segmented; G: CopywriteR vs BAC segmented; H: CopywriteR vs NG135 segmented.
206 out of 230 (90%, 95% CI of proportion: 85%–93%) samples had the same classification on all platforms, as far as they were profiled. As before, we calculated Cohen’s kappa and accuracy values comparing one platform to all the others. In this analysis a sample was called BRCA1-like if it was BRCA1-like in any of the other platforms (Table 3, supplementary Table 1). Supplementary Figure 8 shows cross tables and kappa values for comparisons between all platforms. Accuracies of over 80% were obtained and kappa values of over 0.70. Only classification based on CopywriteR-extracted data had a lower kappa value (0.37), however, the accuracy remained high.

3.4. BRCA1-like status identifies BRCA1-mutated or methylated cases

We investigated the performance in finding BRCA1-mutated and methylated cases for the sequencing based datasets. In this series, the BAC classifier identified 89% (33/37), NGS 93% (28/30) and CopywriteR 100% (24/24) of the BRCA1-mutated or methylated cases. BRCA1-like tumors were thus enriched for known causes of BRCA1 inactivation with respectively 33/48 (69%, BAC), 28/47 (60%, NGS) and 24/35 (69%, CopywriteR). The other cross tables are shown in Supplementary Table 2.

3.5. Sources of differential classification

Overall, classification between the tested platforms was similar. Subsequently, we performed a descriptive analysis to identify the causes underlying differential classification of samples (Supplementary Figure 9). We therefore re-analyzed including samples that were excluded due to quality control issues.

22 out of 31 patients that classified differently had CN profiles on two technologies, while eight had CN profiles on three and one on four technologies; 24 of these passed quality control with at least two profiles and were in the previous analysis. The shrunken centroid classifier will compare whether a sample is closer to the average profile of the BRCA1-like class or closer to the non-BRCA1-like class. We found that samples that have an inconsistent classification within one patient have a lower strength-of-classification and a larger standard deviation of strength-of-classification, compared to those patients that have the same classification (Figure 3). The filtering of samples with low quality partially removed this effect, indicating that samples with low signal and high noise are less strongly associated with a class. We observed that samples that failed to meet quality criteria were more likely those samples with different classification (p = 0.04, Fisher’s exact test).

Figure 1 — (Continued)
Poor quality CN data could not explain all misclassifications, indicated by the fact that misclassification still occurred after quality control. Visual inspection of the profiles (Supplementary Figure 9) with different classifications demonstrated differences in quality between the profiles. Combining these findings we found a combination of three reasons for misclassification: 1) quality differences between CN profiles of the same patient, 2) lack of strong association in one of the CN profiles with one of the classes, 3) the presence of aberrations in genomic regions that are used for classification. The third point means that misclassification cannot occur if aberrations that are affected by differences in signal to noise ratio are absent from classifier regions. However, only a minority of samples is affected by a combination of these causes. Misclassification can be observed in Figure 2: weakly assigned cases have intermediate mean BRCA1-like probabilities and a small maximum difference between the BRCA1-like probabilities. Cases that are clearly discordant have a high maximum difference between the BRCA1-like probabilities. This could be due to either an incorrect original classification because of a lower BAC3K profile quality, or an incorrect classification based on the mapped profile. An incorrect mapped classification could occur because the original classifier was trained to recognize uncertainty in the classification (which results in probabilities around 0.5 for both classes) in the BAC3K profiles. Using mapped data with differences in, for example, noise or dynamic range could then increase the association with the wrong/correct class.

4. Discussion

In this study we investigated the robustness of our previously established BRCA1-like classification of CN profiles from breast cancer samples that were profiled on two to five different experimental platforms. We found that genomic position-based mapping between platforms results in comparable CN profiles and subsequently similar BRCA1-like classification with high accuracy and agreement between platforms.

The overall comparability of CN profiling platforms for large aberrations (chromosome-arm level) (Baumbusch et al., 2008; Curtis et al., 2009; Krijgsman et al., 2012; Schouten et al., 2013a; Wicker et al., 2007) could be a beneficial characteristic for applying our BRCA1-like CN profile classifier (Joosse et al., 2009; Vollebergh et al., 2011). This classifier is based on genomic changes that arise in patients with BRCA1 mutation carriers or that have BRCA1 promoter methylation (Joosse et al., 2011, 2009; Lips et al., 2011; Vollebergh et al., 2011). We have demonstrated in three independent cohorts that this test could be used to predict benefit from high dose alkylating chemotherapy, which induces DNA double-strand breaks that cannot be repaired in BRCA1-deficient cancer cells (Schouten et al., 2015, 2014, 2013b; Vollebergh et al., 2011). Technological advances and the availability of many datasets have prompted an expansion of the methods deployed to obtain BRCA1-like classification of breast cancer samples.

Given that we did not aim to change the classifier itself, it is of importance to mimic the original BAC3K aCGH profiles as closely as possible. This limits the manipulations that can be done to improve the similarity between profiles. For example, we demonstrated previously (Schouten et al., 2013a) and in this manuscript that segmenting CN data results in repeated profiles that are more similar than unsegmented profiles. However, the segmentation negatively influences classification as the original classifier was trained on unsegmented data (Joosse et al., 2009; Schouten et al., 2013a; Vollebergh et al., 2011). As Curtis et al. described there is no ideal platform comparison because it is practically impossible to remove all differences between technologies when obtaining the actual copy number data (Curtis et al., 2009). In our case we also had experiments performed in the labs that were specialized in a certain technology. Therefore, it is not possible to derive the exact influence of the experimental platform and
experimental variations on classification. Furthermore, the applied measures of performance could be influenced by biological characteristics of the cohort. The biological characteristics of the cohort may confound estimating the performance. These characteristics alter the proportion of (non-)BRCA1-like tumors. For example, since BRCA1-like status associates with triple negative status, a cohort which contains mostly triple negative tumors (CopywriteR and MIP analysis) has few non-BRCA1-like cases. Few cases that misclassify result in a low kappa value while maintaining high accuracy (most concordantly classified to reference). An equal amount of misclassifications in a cohort with more non-BRCA1-likes does not result in such a drop in kappa value. Using these two measures we hope to thus control for optimistic accuracy estimates with the kappa value, and to control for pessimistic estimates with accuracy, acknowledging that both can’t control for confounding or a bias in selection of some of the cohorts. Keeping these limitations in obtaining copy number from different technologies and of the performance measures in mind, we observed a high concordance of classification when comparing to the original BAC aCGH-based classification. 7 out of 7 datasets reached the same classification.
classification in over 80% of the samples with MIP and CopywriteR possessing a low Cohen’s kappa value for concordance. We attribute the lower kappa value for MIP and CopywriteR based data to the number and distribution of patients in the analysis as well as to differences in the quality of the profile or the mapping process. Mapping the MIP data to BAC clones requires transformation of CN to log2 ratio and a relatively high number of mapped MIP profiles did not meet the required quality criteria.

We tried to enlarge our datasets by pooling the data to obtain ‘consensus classification’ (in contrast to the gold standard) in a larger dataset. We found high accuracies (>0.8) in these pooled analyses, while kappa value was over 0.7 for most comparisons. Only CopywriteR-based classification was low at 0.37. We attribute this to the distribution of patients (see above) and two cases that were not strongly assigned with probability scores around the cutoff.

In the direct comparison we used the BRCA1-like classification obtained by BAC aCGH as gold standard. Unfortunately, BRCA1 mutation or methylation status was not available for all samples. For those datasets with available BRCA1 status the classifier reproducibly found the BRCA1-mutated and methylated cases. This indicates that patients with a known BRCA1-inactivation are being identified. However, the classifier identifies BRCA1-like cases without BRCA1-mutation or −methylation that benefit from a chemotherapy regimen

<table>
<thead>
<tr>
<th>Non-BRCA1-like</th>
<th>BRCA1-like</th>
<th>accuracy</th>
<th>95% CI</th>
<th>Kappa</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>BAC3K</td>
<td>non-BRCA1-like</td>
<td>93</td>
<td>14</td>
<td>0.91</td>
<td>0.85–0.95</td>
</tr>
<tr>
<td>NGS</td>
<td>non-BRCA1-like</td>
<td>78</td>
<td>1</td>
<td>0.93</td>
<td>0.87–0.96</td>
</tr>
<tr>
<td>NG720</td>
<td>non-BRCA1-like</td>
<td>11</td>
<td>55</td>
<td>0.97</td>
<td>0.90–1.0</td>
</tr>
<tr>
<td>NG135</td>
<td>non-BRCA1-like</td>
<td>30</td>
<td>0</td>
<td>0.94</td>
<td>0.85–0.98</td>
</tr>
<tr>
<td>SNP6</td>
<td>non-BRCA1-like</td>
<td>9</td>
<td>3</td>
<td>0.89</td>
<td>0.75–0.96</td>
</tr>
<tr>
<td>BAC32K</td>
<td>non-BRCA1-like</td>
<td>12</td>
<td>3</td>
<td>0.87</td>
<td>0.68–0.95</td>
</tr>
<tr>
<td>MIP</td>
<td>non-BRCA1-like</td>
<td>6</td>
<td>2</td>
<td>0.89</td>
<td>0.71–0.97</td>
</tr>
<tr>
<td>CopywriteR</td>
<td>non-BRCA1-like</td>
<td>1</td>
<td>0</td>
<td>0.92</td>
<td>0.76–0.98</td>
</tr>
</tbody>
</table>

Figure 3 — Strength of classification Density plots of the strength and variation of classification within one patient for concordantly and discordantly classified patients. A) density plot indicating the strength of assignment to the BRCA1-like or non-BRCA1-like of all 243 patients. The strength of assignment is calculated as the mean absolute difference between one patient’s profiles and the BRCA1-like and non-BRCA1-like class average profile. Zero indicates that the profile is equally close to the BRCA1-like as the non-BRCA1-like average. The higher the value the stronger its association with a particular class. B) Density plot of the standard deviation of the strength of assignment, indicating the association of the profiles from a patient with a particular class. In black are the patients for which all copy number profiles classified the same class, in red the patients that have different class assignment across technologies.
targeting the BRCA1 defect (Vollebergh et al., 2011; Schouten et al., 2015, 2014), which makes BRCA1-like classification the relevant read-out for predictive biomarker analyses.

We found that differential classification occurred in approximately 10% of the cases. The misclassification is caused by samples that lack strong characteristics of one class and/or varying quality profiles within a patient’s samples. In cases with different classification it is uncertain which classification is correct, since the classifier may account for uncertainty arising from lower resolution BAC aCGH data. Conversely, it could be that lower quality BAC aCGH data obscured the true class and the use of higher resolution platform removed this uncertainty. In general, we observed high concordance and therefore applicability, at least for research purposes. The hypothetical ideal approach is always to optimize a classifier on the new platform with the same samples. However our results indicate that this may not be necessary. If optimizing is not feasible, similar classification can be obtained by mapping profiles. Furthermore, one should balance whether the benefit of optimizing the classifier on a platform outweighs (i.e. results in a much better classification) changing the test. It is advisable to obtain CN profiles for at least some tumor DNAs for which an original CN profile is available for comparison. Furthermore, a large reference set of profiles obtained with different platforms can be useful for finding outliers or potential errors in previous experiments (for example, our investigation of differential classification).

In conclusion, we demonstrated that BRCA1-like classification of mapped CN profiles is robust across multiple datasets and experimental platforms. This allows for the further investigation of the clinical benefit of treatments targeting the BRCA1 defect in existing datasets with CN profiles. Furthermore, the high concordance in CN profiles across different technologies encourages use on a range of current and established platforms.

Conflict of interest

SC Linn and FM Nederlof are named inventors on a patent application for the BRCA1 and -2-like array comparative genomics hybridization classifiers. The other authors do not disclose any conflict of interest.

Acknowledgments

This work was supported by an unrestricted research grant from Roche Life Science to translate the BRCA1-like classifier to a next generation sequencing platform, a grant from the Life Sciences Center Amsterdam (LSCA) Validation fund and the FP7 Collaborative Project “RATHER” (www.ratherproject.com).

Appendix A.
Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.molonc.2015.03.002.

REFERENCES


Schouten, P., 2014. Table1Heatmap. http://cran.us.r-project.org/web/packages/Table1Heatmap/index.html.


