

RESEARCH ARTICLE

Open Access

Cross-species comparison of aCGH data from mouse and human *BRCA1*- and *BRCA2*-mutated breast cancers

Henne Holstege^{1,2}, Erik van Beers^{3,4}, Arno Velds⁵, Xiaoling Liu^{1,6}, Simon A Joosse^{3,7}, Sjoerd Klarenbeek¹, Eva Schut¹, Ron Kerkhoven⁵, Christiaan N Klijn¹, Lodewyk FA Wessels^{1,8}, Petra M Nederlof³, Jos Jonkers^{1*}

Abstract

Background: Genomic gains and losses are a result of genomic instability in many types of cancers. *BRCA1*- and *BRCA2*-mutated breast cancers are associated with increased amounts of chromosomal aberrations, presumably due their functions in genome repair. Some of these genomic aberrations may harbor genes whose absence or overexpression may give rise to cellular growth advantage. So far, it has not been easy to identify the driver genes underlying gains and losses. A powerful approach to identify these driver genes could be a cross-species comparison of array comparative genomic hybridization (aCGH) data from cognate mouse and human tumors. Orthologous regions of mouse and human tumors that are commonly gained or lost might represent essential genomic regions selected for gain or loss during tumor development.

Methods: To identify genomic regions that are associated with *BRCA1*- and *BRCA2*-mutated breast cancers we compared aCGH data from 130 mouse *Brca1*^{ΔΔ},*p53*^{ΔΔ}, *Brca2*^{ΔΔ},*p53*^{ΔΔ} and *p53*^{ΔΔ} mammary tumor groups with 103 human *BRCA1*-mutated, *BRCA2*-mutated and non-hereditary breast cancers.

Results: Our genome-wide cross-species analysis yielded a complete collection of loci and genes that are commonly gained or lost in mouse and human breast cancer. Principal common CNAs were the well known *MYC*-associated gain and *RB1/INTS6*-associated loss that occurred in all mouse and human tumor groups, and the *AURKA*-associated gain occurred in *BRCA2*-related tumors from both species. However, there were also important differences between tumor profiles of both species, such as the prominent gain on chromosome 10 in mouse *Brca2*^{ΔΔ},*p53*^{ΔΔ} tumors and the *PIK3CA* associated 3q gain in human *BRCA1*-mutated tumors, which occurred in tumors from one species but not in tumors from the other species. This disparity in recurrent aberrations in mouse and human tumors might be due to differences in tumor cell type or genomic organization between both species.

Conclusions: The selection of the oncogenome during mouse and human breast tumor development is markedly different, apart from the *MYC* gain and *RB1*-associated loss. These differences should be kept in mind when using mouse models for preclinical studies.

Background

Western women have a 10-12% risk of developing breast cancer, making this disease the most common malignancy in females [1]. Approximately 5-10% of breast cancer cases can be explained by a hereditary predisposition. Between 25-40% of these cases involves the

inheritance of one defective copy of either the *BRCA1* gene or the *BRCA2* gene [2-5] which predisposes women in these families to a ~50-80% lifetime risk of developing breast cancer and to a lesser extent ovarian cancer [1].

BRCA1 and *BRCA2* proteins are implicated in DNA double strand break (DSB) repair and loss of *BRCA* function results in chromosomal instability, hence to cancer predisposition [6-8]. Chromosomal instability can be evaluated by array-based comparative genomic

* Correspondence: jjonkers@nki.nl

¹Division of Molecular Biology, Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands

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

hybridization (aCGH) which measures genomic copy number alterations (CNAs) in sample (tumor) DNA relative to normal (diploid) DNA [9]. Analysis of the recurrent regions of chromosomal instability in BRCA1 or BRCA2 mutated breast tumors may point to loci and/or genes involved in development and progression of BRCA-associated hereditary breast cancer. To gain insight into the specific effects of BRCA1 or BRCA2 loss on genomic instability, we used aCGH profiles of genetically defined mouse models for BRCA1- and BRCA2-associated breast cancer as biological data filters to mine the human breast cancer genome [10,11]. In these mouse models breast cancer is driven by combined loss of p53 and BRCA1 or BRCA2. The mouse *p53^{Δ/Δ}* mammary tumors are a mix of basal carcinomas and carcinosarcomas, whereas the mouse *Brca1^{Δ/Δ};p53^{Δ/Δ}* and *Brca2^{Δ/Δ};p53^{Δ/Δ}* tumors are mainly basal-like. Human BRCA2-mutated tumors resemble sporadic tumors in that they do not skew towards any phenotype [12], whereas, 90% of human BRCA1-mutated tumors have a basal like phenotype [10,12,13]. Of note, tumors from our mouse models do not give rise to ER-positive luminal tumors, which is the cell type of 70-80% of human sporadic and BRCA2-mutated breast tumors [14,15]. It is well known that human basal-like and luminal human breast tumors have different aCGH profiles [16]; therefore, the comparison of aCGH profiles of mouse *Brca1^{Δ/Δ};p53^{Δ/Δ}* carcinomas with basal-like human BRCA1-mutated tumors may give results relevant for basal-like breast tumors as well as for BRCA1-loss.

Syntenic regions that are found recurrently gained or lost in both human and mouse have a greater probability of being relevant for development of BRCA-associated and/or sporadic breast cancer [17]. To analyze syntenic genomic regions for overlapping CNAs, we applied two statistical frameworks to BRCA1-deficient, BRCA2-deficient and BRCA1/2-proficient control mouse and human breast cancers: a total of six tumor groups. First we used KC-SMART [18], which identifies significantly recurrent gains and losses within each tumor group. Next, we identified syntenic CNAs that co-occurred in both the mouse and human BRCA1-mutated, BRCA2-mutated and control tumor groups. Our data indicate that amplification of the MYC locus and loss of the RB/INTS6 locus occurs in all mouse and human tumor groups, and that AURKA amplification occurs in mouse and human BRCA2-mutated tumors. We also applied *comparative-KC-SMART* which identifies gains and losses that occur significantly more often in one tumor group compared to another. We used this method to find CNAs that occur more often in BRCA1- or BRCA2-deficient breast tumors compared to BRCA-proficient control tumors. Next, we identified which of these CNAs occur both in mouse

BRCA1- or BRCA2-deficient mammary tumors and in human BRCA1- or BRCA2-mutated tumors. Together, our KC-SMART and *comparative-KC-SMART* analyses highlight important genotypic similarities between mouse and human cognate cancers. However, we also find several genomic aberrations to be highly recurrent in one species but not in the other.

Methods

Mouse tumors

We bred cohorts of *K14Cre;Brca1^{F/F};p53^{F/F}* female mice [10], *K14Cre;Brca2^{F/F};p53^{F/F}* female mice [11], and *K14Cre;p53^{F/F}* female mice [10]. *Cre* recombinase expression in these mice is restricted to cytokeratin-14 (K14) expressing epithelial tissues, including mammary gland epithelium. We isolated DNA from 35 *Brca1^{Δ/Δ};p53^{Δ/Δ}*, 62 *K14Cre;Brca2^{Δ/Δ};p53^{Δ/Δ}* and 33 *p53^{Δ/Δ}* mammary tumors as well as (tumor-free) spleens from the same animals. Tumor type was scored by histopathology and by expression levels of vimentin and E-cadherin: both histopathology procedures and expression array procedures were previously described in [10]. All animal experiments were approved by the local ethical review committee.

Mouse tumor aCGH profiles were obtained using a microarray containing 3080 unique mouse bacterial artificial chromosome (BAC) clones spotted in duplicate, as described previously [19]. Hybridizations were done as described in Chung et al., 2004 with minor modifications. In brief, we labeled 2 μg of tumor DNA with Cy5 and 2 μg reference (spleen) DNA with Cy3 and vice versa for the dye swap (Universal Linkage System, Kreatech Biotechnology, Cat# EA-006) according to the manufacturer's protocol. All hybridizations were performed in a hybridization station (Tecan, Cat# Hs4800). Data, including detailed hybridization protocols were added to the public repository Array Express with accession numbers E-NCMF-34 and E-NCMF-35 [20]. The intensity of the Cy5 (tumor) signal was calculated relative to the intensity of the Cy3 (reference) signal for each spot on the array and this ratio was log₂ transformed. Data was normalized by shifting these ratios by the median log₂ ratio per sub array. Dye swap data (a total of four spots per data point) were combined using the Rosetta error model [21]. Each mouse tumor DNA sample was isolated from fresh tumor tissues and hybridized against a reference DNA (spleen) obtained from the same animal to avoid picking up DNA copy number variations between mice.

To compare aCGH profiles from mouse with human BRCA1- and BRCA2-mutated tumors, we used two previously published aCGH datasets [22,23]. For the cross-species comparison: we used 27 human BRCA1-mutated and 28 BRCA2 related breast tumors for the KC-SMART

analysis, and compared them with 48 control tumors of different cell types in the *comparative*-KC-SMART analysis. Pathology, tumor type, and *BRCA1*- and *BRCA2*-mutations of human tumors are shown in Additional file 1. All experiments with human materials were approved by the local ethics review committee.

KC-SMART and comparative-KC-SMART analysis

KC-SMART

(Kernel Convolution - a Statistical Method for Aberrant Region detection) determines which CNAs are significantly recurrent *within a tumor group* [18]. KC-SMART generates a Kernel Smoothed Estimate (KSE) for gains (KSE_{gains}) and losses (KSE_{losses}) separately across the genome and across *agroup* of tumors. Briefly, at the genomic midposition of each BAC probe, KC-SMART places a Gaussian kernel with the height of the sum of either the positive log₂ values or the negative log₂ values of all tumors within a tumor group. KSE curves (one for gains and one for losses) are point estimates determined by convolution of locally weighted kernel functions such that the closer the genomic distance between two probes the more they will contribute to each others convoluted values. Significance cut-offs are calculated for gains and losses separately using the distribution of KSE peak heights from randomized data ($P < 0.05$) as explained in [18].

Comparative-KC-SMART

detects CNAs that occur significantly more often in one tumor group *compared* to another. The algorithm creates a KSE curve for each individual tumor by placing Gaussian kernel functions with the height of the log₂ value and at the genomic midposition of each probe (without separating gains and losses, as done for KC-SMART analysis). For each genomic position in the KSE curve, the KSE values from one tumor group are compared to the KSE values from another group by calculating a signal to noise ratio (SNR). We determined a cutoff that defines significant SNR values by applying a False Discovery Rate (FDR) of 0.05 on the SNR data and randomized SNR data using 6000 class-label permutations.

The width of a kernel applied to each data point determines the sensitivity of smoothing and the size of aberrations detected. To compare mouse and human tumor groups we consistently used one kernel width for all tumor profiles. We found that, when using a 20 Mb kernel width, KC-SMART best smoothed noise while detecting CNAs from aCGH profiles of both mouse and human individual tumors. To identify the genomic locations of the peaks of a KSE curve, we calculated for which position the KSE value was higher (for gains), or lower (for losses), compared to the values of its neighboring datapoints. While in many instances a peak may

be a local maximum or minimum within a larger gain or loss, these local peaks are part of the data and might harbor interesting genes that drive the larger gain or loss. An R-package of the KC-SMART algorithm (which includes the *comparative*-KC-SMART algorithm) is included in Bioconductor [24].

Combining mouse and human aCGH datasets

We used the BioMart data-mining tool in Ensembl Build 52 to cross-reference two Ensembl datasets (NCBI Build 36 and NCBI Build 37). Because NCBI Build 36 was used to map the mouse RP23-BAC clones, we mapped the genomic positions of the mouse genes using this older build. However, in order to use the most current mouse-human orthologue information, we matched the ENSMUS numbers of NCBI Build 37 with their genomic positions as listed in NCBI Build 36. We obtained a list of 19589 unique mouse-human homolog combinations. In this list we found 16679 unique human genes and 17048 unique mouse genes (one human gene may have more than one mouse homolog and vice versa). We determined which genes from this list map to the significantly gained or lost regions as determined by KC-SMART method or to the differentially gained or lost regions as determined by the *comparative*-KC-SMART method for the human and mouse tumor groups separately. Next, we queried for those genes whose homologues map to regions gained or lost in *both* the human and the mouse tumor sets. The locations of these genes in the mouse and human genome are plotted by connecting lines to their syntenic regions. Genomic locations of the overlapping syntenic regions were determined by taking the genomic position of the start of the first gene and the end of the last gene.

Results

1: aCGH analysis of mouse mammary tumors

To investigate the impact of *BRCA1* and *BRCA2* deficiency on chromosomal instability in breast epithelial cells, we performed aCGH on mammary tumors derived from our genetically engineered mouse (GEM) models for *BRCA1*- and *BRCA2*-associated breast cancer [10,11]. Mammary tumors in these mice arose from epithelial-specific loss of p53 alone ($n = 33$), or in combination with *BRCA1* ($n = 35$) or *BRCA2* ($n = 62$). Typical examples of aCGH profiles from *Brca1*^{Δ/Δ};p53^{Δ/Δ}, *Brca2*^{Δ/Δ};p53^{Δ/Δ} and p53^{Δ/Δ} tumors are shown in Figure 1a. All but one tumors in the *Brca1*^{Δ/Δ};p53^{Δ/Δ} tumor group were of basal cell type, consisting of primarily high grade invasive ductal carcinoma not otherwise specified (IDC-nos; 91%), 3% carcinosarcomas and 6% adenomyoepitheliomas (previously described in [10]). Likewise, 90% of the *Brca2*^{Δ/Δ};p53^{Δ/Δ} tumors are carcinomas and 9% are carcinosarcomas. In contrast, the p53^{Δ/Δ} tumor group consisted of mixed cell

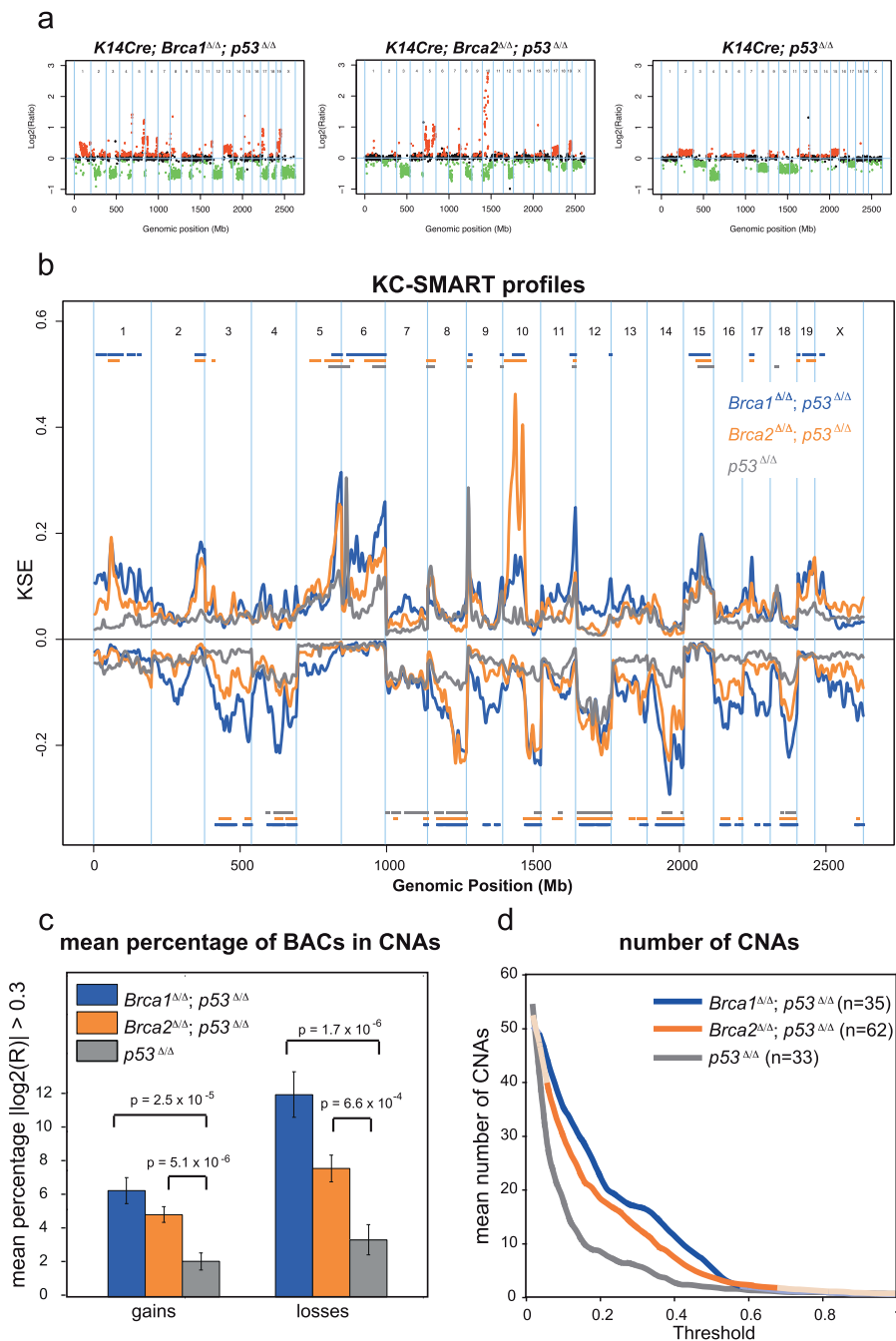


Figure 1 aCGH profiles of mouse mammary tumors. aCGH profiles typical of (a) *Brca1^{Δ/Δ}; p53^{Δ/Δ}*, *Brca2^{Δ/Δ}; p53^{Δ/Δ}* and *p53^{Δ/Δ}* mammary tumors. Each dot represents the averaged log₂ ratio (y-axis) of a BAC clone plotted at its genomic position (x-axis). Red dots (gains) and green dots (losses) represent datapoints with log₂ ratios significantly different from 0 as determined by the Rosetta error model [21]; black dots represent datapoints with log₂ ratios not significantly different from 0. Blue vertical lines represent chromosome boundaries. (b) aCGH profiles from 35 *Brca1^{Δ/Δ}; p53^{Δ/Δ}* tumors (blue), 62 *Brca2^{Δ/Δ}; p53^{Δ/Δ}* tumors (orange) and 33 *p53^{Δ/Δ}* tumors (gray) were analyzed with KC-SMART using a kernel width of 20 Mb, for gains and losses separately. Significant CNAs are depicted in the color matched bars on top (gains) and below (losses) the Kernel Smoothed Estimate curves (KSEs) for each tumor group. (c) The mean percentage of BAC clones \pm S.E.M. with absolute log₂ ratio > 0.3 is greater across the *Brca1^{Δ/Δ}; p53^{Δ/Δ}* tumors (blue) and *Brca2^{Δ/Δ}; p53^{Δ/Δ}* tumors (orange) compared with the *p53^{Δ/Δ}* tumors. P-values are determined by a two-tailed t-test with unequal variance. (d) Each individual tumor aCGH profile was smoothed using KC-SMART (kernel width: 20 Mb). For a range of thresholds, all gains exceeding a positive threshold and losses exceeding the same negative threshold were counted and averaged over each tumor group. Curves are darkened at thresholds for which the average number of CNAs per tumor group is significantly greater in the *Brca1/2^{Δ/Δ}; p53^{Δ/Δ}* tumors compared with the *p53^{Δ/Δ}* tumors, $P < 0.05$ determined by a two-tailed t-test with unequal variance.

types: 39% intermediate to high-grade IDC-nos, 50% carcinosarcomas and 11% adenomyoepitheliomas. Tumor type scoring was based on histopathology and E-cadherin/Vimentin expression, Table 1, (for tumor type data see Additional files 2 and 3).

We used the KC-SMART method [18] to identify significantly recurrent genomic aberrations in the aCGH data of the mouse mammary tumor groups. Figure 1b shows Kernel Smoothed Estimate curves (KSE curves; see Methods) for gains and losses separately for the *Brca1*^{Δ/Δ}; *p53*^{Δ/Δ}, *Brca2*^{Δ/Δ}; *p53* and *p53*^{Δ/Δ} tumor groups. Aberrations that exceeded the significance thresholds are indicated by color-matched bars above (gains) or below (losses) the KSE curves, and their genomic positions are listed in Additional file 4. Significantly recurrent gains (chromosomes 5, 6, 9-cen, 11-tel and 15) and losses (chromosomes 4, 7-tel, 8, 10-tel, 12, 14, and 18) overlapped between all tumor groups. Regional gains on chromosomes 1, 2, 10, 17, 19 and losses on chromosomes 3, 13, 16 and X were common to both *Brca1*^{Δ/Δ}; *Trp53*^{Δ/Δ} and *Brca2*^{Δ/Δ}; *Trp53*^{Δ/Δ} tumors but not to *Trp53*^{Δ/Δ} tumors. In addition, a high level amplicon on chromosome 10 occurred in >50% of the *Brca2*^{Δ/Δ}; *Trp53*^{Δ/Δ} tumors.

Compared with the *Brca1*^{Δ/Δ}; *Trp53*^{Δ/Δ} tumor group, the *Trp53*^{Δ/Δ} tumor group showed more whole chromosome gains and losses, rather than regional aberrations. The *Brca2*^{Δ/Δ}; *Trp53*^{Δ/Δ} tumor group appeared to be an intermediate version of both extremes. To quantify the level of genomic instability for each tumor group we calculated the mean percentage of BAC clones with log2 ratio >0.3 for gains and <-0.3 for losses separately and found that this percentage was significantly higher for the *Brca1*^{Δ/Δ}; *p53*^{Δ/Δ} and *Brca2*^{Δ/Δ}; *p53*^{Δ/Δ} groups compared with the *p53*^{Δ/Δ} group (Figure 1c). Also the *Brca1*^{Δ/Δ}; *p53*^{Δ/Δ} group had significantly more losses than the *Brca2*^{Δ/Δ}; *p53*^{Δ/Δ} group.

Next, to quantify the amount of CNAs in each individual tumor we smoothed their individual aCGH profiles with KC-SMART (without separating gains and losses) and determined the mean number of peaks in the KSE curves exceeding a range of KSE thresholds. We found that the average number of CNAs was significantly higher in the *Brca1*^{Δ/Δ}; *p53*^{Δ/Δ} tumor group compared with the *p53*^{Δ/Δ} tumor group between thresholds of 0.04 - 0.54, and significantly higher in the *Brca2*^{Δ/Δ}; *p53*^{Δ/Δ}

tumor group compared with the *p53*^{Δ/Δ} tumor group between thresholds of 0.06 - 0.62 (Figure 1d). To exemplify the regions counted, we depicted the regions exceeding an arbitrary threshold of 0.15 in line-plots (Figure 2). Overall, gains and losses that occurred in the *p53*^{Δ/Δ} tumor group also occurred in the *Brca1*^{Δ/Δ}; *p53*^{Δ/Δ} and *Brca2*^{Δ/Δ}; *p53*^{Δ/Δ} tumor groups but in a greater fraction of tumors, and with higher log2 ratios. Hence, *Brca1* or *Brca2* loss on top of *p53* loss appears to aggravate the tumor profile.

After stratifying the *p53*^{Δ/Δ} tumors according to tumor type, we found that the carcinomas and adenomyoepitheliomas have significantly more genomic instability compared with sarcomas (Figure 3). This could explain why the KSE curves from the *p53*^{Δ/Δ} tumor groups (which contain both carcinosarcomas and carcinomas) remain at lower level compared with the KSE curves from the *Brca1*^{Δ/Δ}; *p53*^{Δ/Δ} tumor and the *Brca1*^{Δ/Δ}; *p53*^{Δ/Δ} tumor groups which contains mainly carcinomas (Figure 1b).

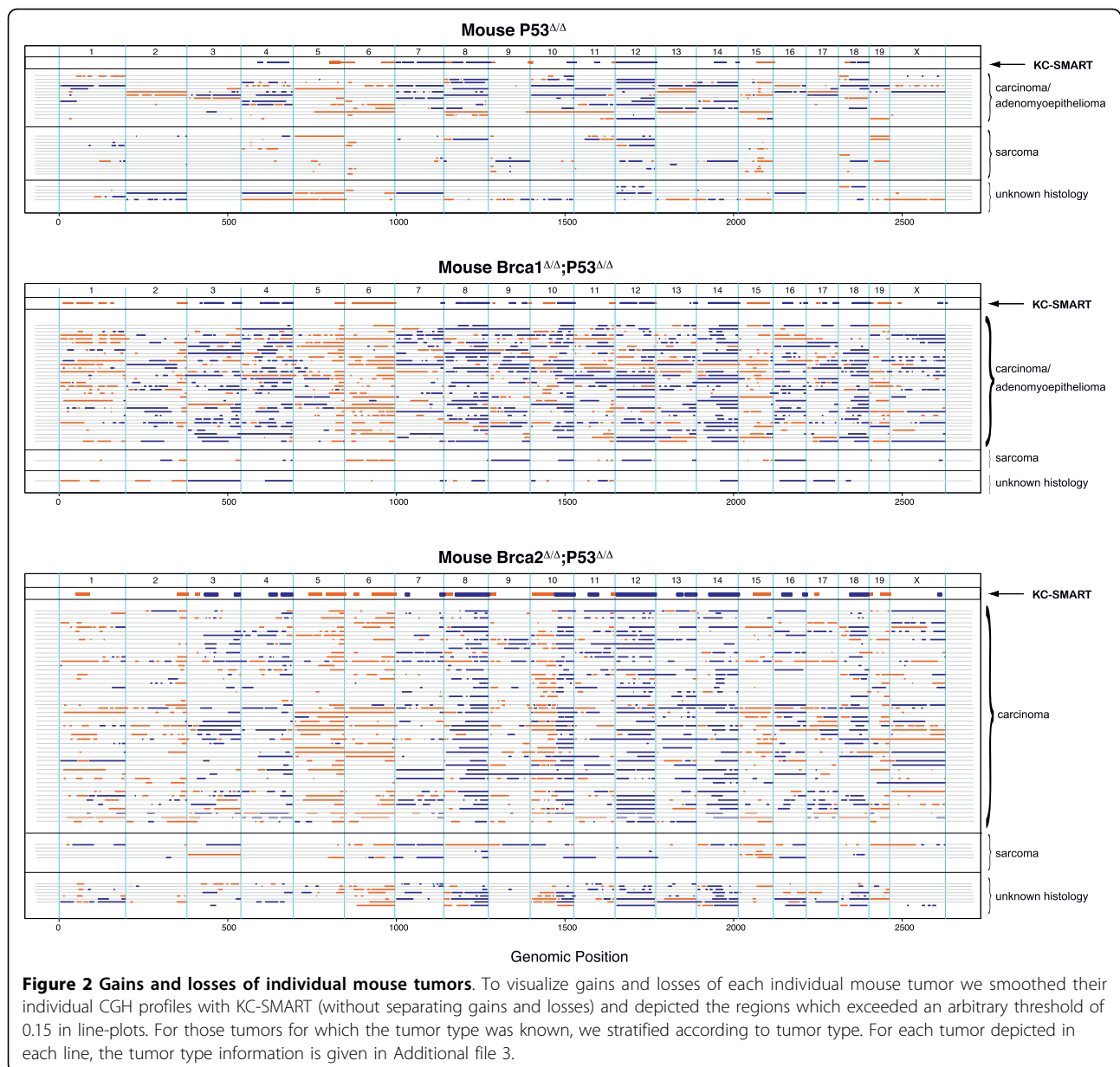
aCGH analysis of human breast tumors

We compared aCGH data from our *Brca1*^{Δ/Δ}; *p53*^{Δ/Δ}, *Brca2*^{Δ/Δ}; *p53*^{Δ/Δ} and *p53*^{Δ/Δ} tumor groups with aCGH data from previously published *BRCA1*-mutated, *BRCA2*-mutated and control human breast cancers [22,23]. First, we assessed the genomic instability of the human tumors by calculating the mean percentage of BAC clones with log2 ratio >0.2 for gains and <-0.2 for losses separately. The mean log2-ratio of the gains and the losses in the aCGH data of human tumors was lower compared to the mouse aCGH data (data not shown); hence, a lower cutoff was used to calculate the percentage of BACs reporting gains or losses in the human CNAs. This percentage was significantly higher for the gains of the *BRCA1*-mutated tumor group but not for the *BRCA2*-mutated tumor group when compared to the sporadic control tumors ($p = 2.5 \times 10^{-3}$ calculated with t-test with unequal variance) (Figure 4a). The fraction of losses in the *BRCA1* and *BRCA2* tumor groups were not significantly different from The fraction of losses found in the sporadic control group. Next, to get an indication of the amount of CNAs in each tumor group, we smoothed the individual aCGH profiles with KC-SMART and determined the mean number of gains and losses in the KSE curves exceeding a range of

Table 1 Tumor types according to histopathology

mouse tumor group	total scored	Carcinoma	Carcinosarcoma	Adenomyoepithelioma	NA
<i>Brca1</i> ^{Δ/Δ} ; <i>p53</i> ^{Δ/Δ} (n = 35)	34	91% (31)	3% (1)	6% (2)	1
<i>Brca2</i> ^{Δ/Δ} ; <i>p53</i> ^{Δ/Δ} (n = 62)	54	90% (49)	9% (5)	0% (0)	8
<i>p53</i> ^{Δ/Δ} (n = 33)	28	39% (11)	50% (14)	11% (3)	5

Histopathology of *Brca1*^{Δ/Δ}; *p53*^{Δ/Δ}, *Brca2*^{Δ/Δ}; *p53*^{Δ/Δ}, *p53*^{Δ/Δ} mammary tumors, determined by H&E and/or vimentin, SMA, as previously described in [10].

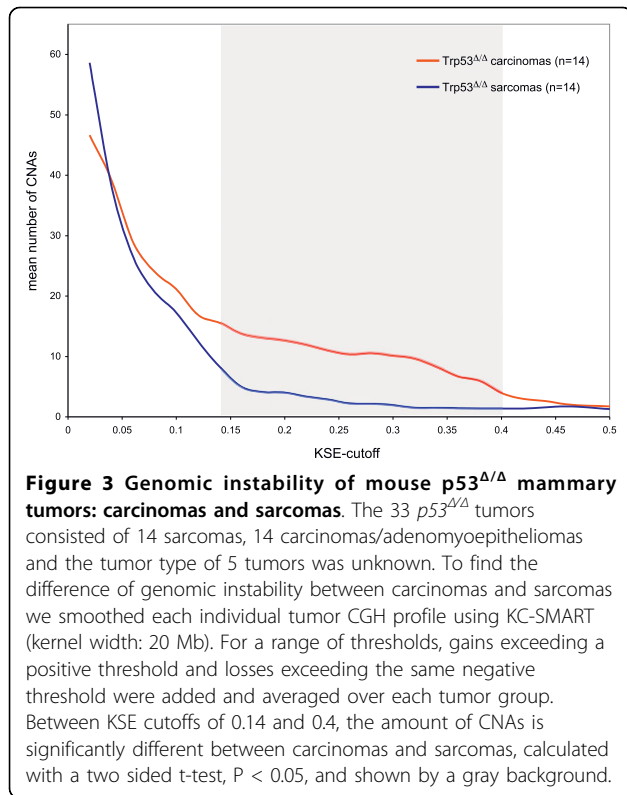


thresholds. The *BRCA1*-mutated tumor group had significantly more CNAs compared with the sporadic control tumor group between thresholds of 0.08 - 0.24. The average amount of gains and losses in the *BRCA2*-mutated tumor group was not significantly increased compared with the control group except between thresholds of 0.06 and 0.08 (Figure 4b).

Cross-species comparison of mouse and human breast tumors

Using aCGH data from genetically defined mouse mammary tumors as a data filter to mine the human breast cancer genome may be a powerful strategy to identify genomic regions with functional relevance for breast

tumorigenesis [17]. To identify and narrow down regions with highest relevance for *BRCA* related tumor formation, we compared the array-CGH profiles of human *BRCA*-mutated and sporadic breast tumors with profiles from mouse *BRCA*-deficient and -proficient mammary tumors. First, we performed a cross species analysis using KC-SMART [18], which identifies significantly recurrent CNAs within each tumor group. Second, we adapted KC-SMART to *comparative-KC-SMART* which identifies CNAs that occur with significantly different frequency in one tumor group compared to another (see Methods). Next, we performed cross species analysis using *comparative-KC-SMART*, which identified regions that occurred in *BRCA*-mutated



tumors but not in BRCA-proficient tumors (or vice versa) in both human and mouse breast tumors.

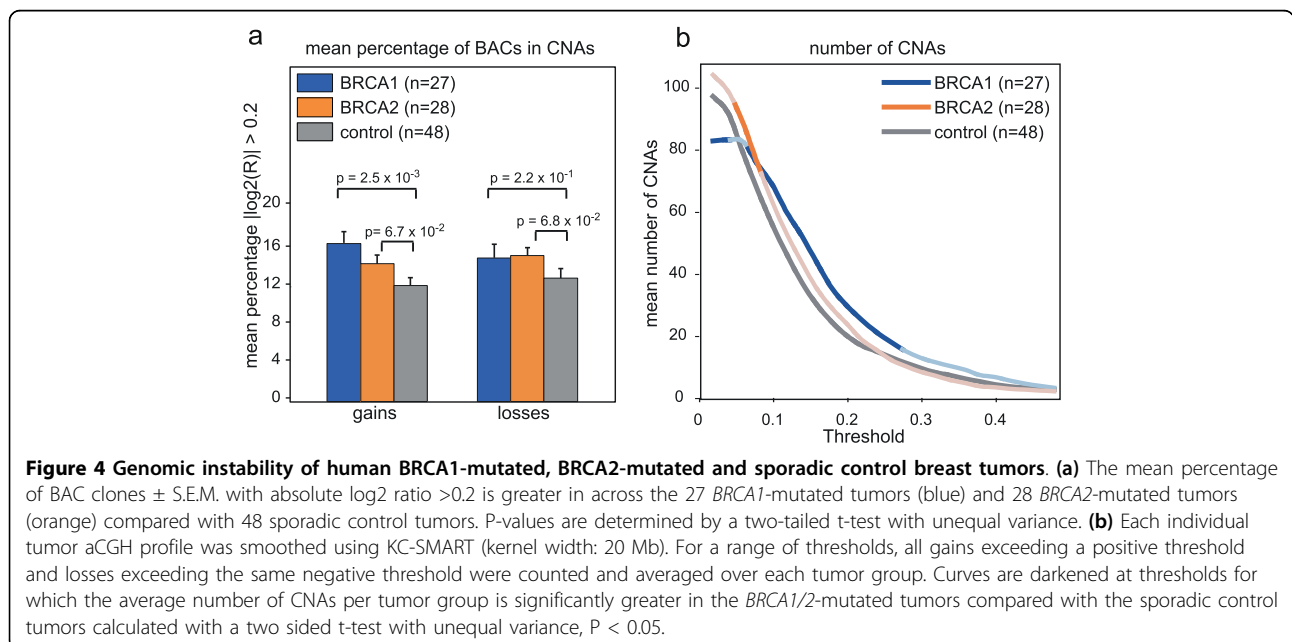
The following paragraphs describe the identification of cross-species syntenic gains and losses. However, they also show the challenges of interpreting cross-species genomic data.

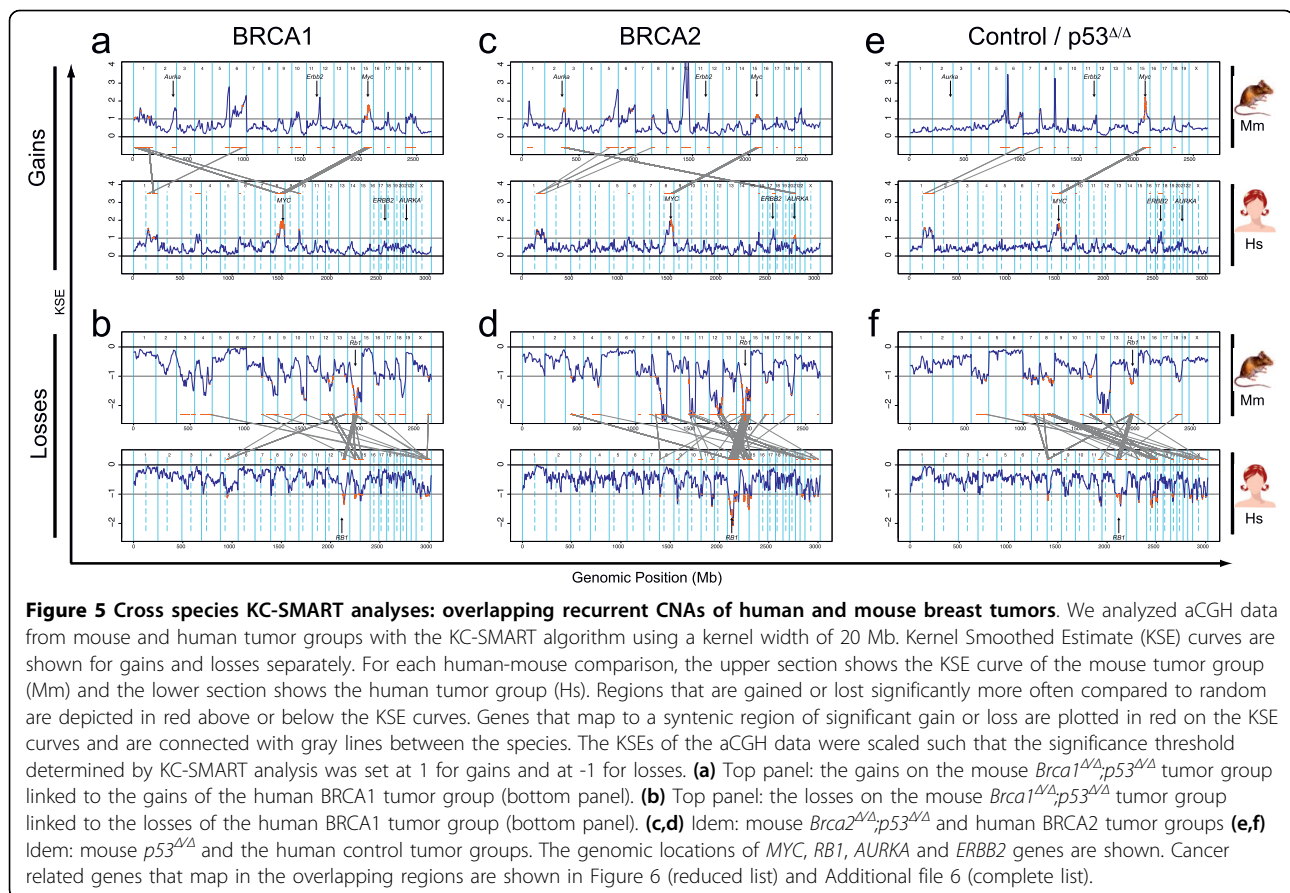
Cross-species analysis using KC-SMART

We used the KC-SMART method [18] to identify significantly recurrent CNAs in mouse and human breast tumors and compared them between the two species (Figure 5). Regions of significant gain or loss determined by KC-SMART are shown in Additional file 4 for mouse tumors and in Additional file 5 for human tumors. For comparison of mouse and human CNAs, we used a synteny list of human genes and their mouse orthologues (see Methods for details). Overlapping CNAs in our cross-species comparison, and the cancer-related genes that map within these regions are listed in Figure 6 (single overlapping genes were left out, for complete list see Additional file 6). We identified human and mouse KSE peak locations and depicted the closest cancer-related genes in Figure 6.

We identified one overlapping gain recurrent in all six tumor groups on human chromosome 8q (mouse chromosome 15). In all mouse tumor groups, the peak of this gain was located at (for $Brca2^{\Delta/\Delta};p53^{\Delta/\Delta}$ and $p53^{\Delta/\Delta}$ tumors) or near (for $Brca1^{\Delta/\Delta};p53^{\Delta/\Delta}$ tumors) the region containing the *MYC* oncogene at 128.8 Mb. The peak of the human *BRCA2*-mutated tumors mapped exactly on the *MYC* oncogene, and the peak of the human control tumor group maps slightly downstream of *MYC* (*ADCY8*, 132.21 Mb). All human tumor groups had a second peak on chromosome 8 mapping between *EXT1/TNFRSF11B* at 119.4 Mb and *ENPP2/COL14A1*, 120.87 Mb). This second peak was not shared with the mouse tumor groups.

In all tumor groups, we found a recurrent region of loss on chromosome 13q (mouse chromosome 14)





which is commonly associated with the *RB1* gene. Indeed, the overlapping regions of the mouse and human BRCA2 groups include *RB1*. The overlap of BRCA1 and control tumors mapped to the flanks of these peaks, notwithstanding that loss of *RB1* could well drive the loss of this region in tumors of both species. Interestingly, the overlap of all three tumor groups included the region containing *INTS6*, an annotated tumor related gene (also known as *DICE1*, Protein deleted in cancer 1).

A remarkable similarity between mouse and human BRCA2-mutated breast tumors was that the peak of both the human (chromosome 2) and mouse (chromosome 2) centered exactly on the Aurora kinase A (*AURKA*) oncogene. Similarly, the human control tumors showed recurrent losses on chromosomes 16q and 17p, and the peaks of syntenic losses in mouse *p53*^{Δ/Δ} tumors centered on exactly the same genes on chromosomes 8 (*MMP2*) and 11 (*MAP2K4*), respectively. This suggests that these genes (or neighboring genes) may be involved in selection of these CNAs during mouse and human tumor development. Note, however, that the *Aurka* gain also occurred in mouse *Brca1*^{Δ/Δ};*p53*^{Δ/Δ} tumors and that loss of mouse chromosomes 8 and 11 was not specific for the

Brca1^{Δ/Δ};*p53*^{Δ/Δ} tumors, suggesting that BRCA2 association of the syntenic gain and BRCA1 association of the syntenic losses are dictated by the human tumor data.

Large regions of chromosome 1 were recurrently gained in the *Brca1*^{Δ/Δ};*p53*^{Δ/Δ} tumors whereas shorter regions were gained in the *Brca2*^{Δ/Δ};*p53*^{Δ/Δ} tumors (see Figure 2). Due to this difference in the mouse tumors, there was a BRCA1-specific overlap with human chromosome 1, as all other human tumor groups showed this significantly recurrent gain. It is interesting that the region of overlap encompasses *PIK3CB* and *MDM4*.

In the BRCA1 and BRCA2 tumor groups of both species, we identified a recurrent loss on human chromosome 14 (mouse chromosome 12), peaking on *ARID4A* (57.87 Mb) and *DACT1* (58.17 Mb). All mouse tumors, including the *p53*^{Δ/Δ} group, show this loss, encompassing entire chromosome 12. Whether the region containing *ARID4A* and *DACT1* contains a driver gene for this whole chromosome loss remains to be established.

Cross species analysis using comparative-KC-SMART

To identify the CNAs that occur significantly more often in mouse BRCA1-deficient tumors compared to control tumors, we compared aCGH data from the mouse *Brca1*^{Δ/Δ};*p53*^{Δ/Δ} tumors with aCGH data from

the $p53^{\Delta/\Delta}$ tumors using *comparative-KC-SMART*. Similarly, to determine which CNAs occur significantly more often in the human *BRCA1*-mutated breast tumors compared to the control tumor group, we used *comparative-KC-SMART* on aCGH data from human tumors. Differentially occurring gains are depicted in Figure 7a and the differential losses are depicted in 7b; top panels represent mouse $Brca1^{\Delta/\Delta};p53^{\Delta/\Delta}$ tumors compared with $p53^{\Delta/\Delta}$ tumors and lower panels represent human *BRCA1*-mutated tumors compared to the control tumor group. We next determined which genomic regions are differentially recurrent in both mouse and human *BRCA1*-mutated tumors by mapping the orthologous genes of our synteny list to the differentially aberrated regions in both species. The mouse and human orthologues are connected by gray lines between the top panel and the lower panel. We applied the same analysis to the mouse and human *BRCA2*-mutated tumors (Figure 7c,d).

Regional CNAs that occur significantly more often in the $Brca1^{\Delta/\Delta};p53^{\Delta/\Delta}$ or $Brca2^{\Delta/\Delta};p53^{\Delta/\Delta}$ mouse tumor groups compared with the $p53^{\Delta/\Delta}$ control tumors are listed in Additional file 7. For the complementing human data see Additional file 8. The regions of human-mouse overlap determined by *comparative-KC-SMART* are listed in Figure 8 (regions with single overlapping genes were left out, for complete list see Additional file 9). Between the human and mouse, control tumors show no syntenic regions of overlap in any of the comparisons, therefore, the orthologous genes mapping to these regions are not connected. Regarding regions encompassing more than one gene, there are 26 regions that overlap between mouse and human *BRCA1*-mutated tumors (10 overlapping *BRCA1*-specific gains and 16 losses), compared to only 4 overlapping *BRCA2*-specific losses. This marked difference is due to the lack of differential gains and losses between the human *BRCA2* tumors and sporadic control tumors.

The largest overlapping *BRCA1*-specific gain between mouse and human *BRCA1*-mutated breast tumors maps to human chromosome 1 and to mouse chromosome 6. This region harbors the *GADD45A* gene, shown to be induced by *BRCA1* independently of *p53* [25], and implicated in regulating centrosome duplication and maintaining genome integrity [26]. We speculate that the *GADD45A* gene cannot be induced in *BRCA1*-deficient cells, and that the necessity for its activity during cellular division could select for its expression by amplification.

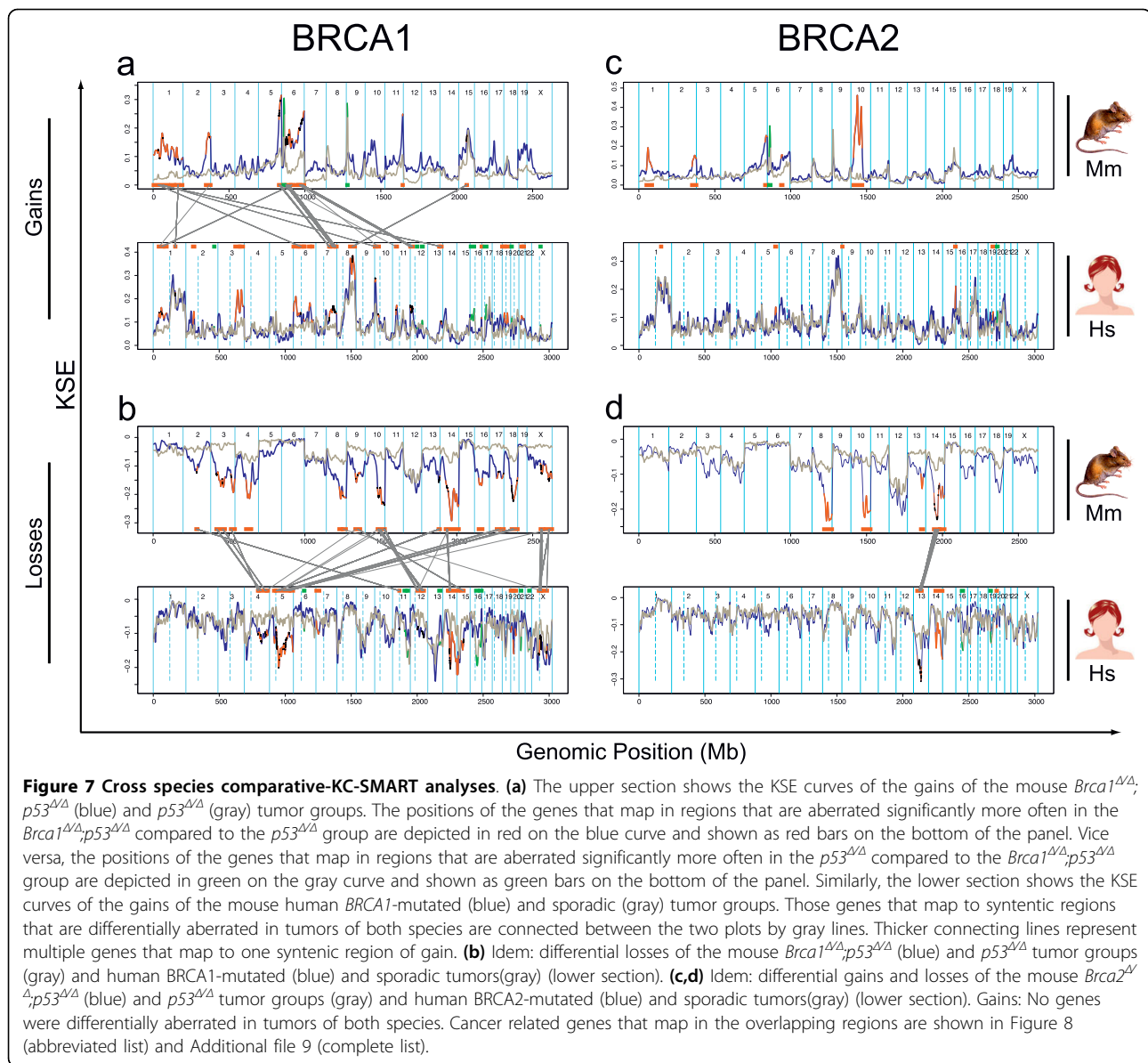
The gain on human chromosome 7 maps to mouse chromosome 6. This is one of the few gains in the human tumors that are seen in the *BRCA1*-mutated tumors but not in the control tumors. Similarly, the gain in the centromeric region of mouse chromosome 6

is also not seen in the $p53^{\Delta/\Delta}$ tumors, suggesting that in this setting, this gain is *BRCA1*-specific in tumors of both species. The mouse tumors have a peak on the *EXOC4* gene (alias *SEC8*, at 132.9 Mb), and the human peak at the *CALD1* gene (Caldesmon 1, at 134.2 Mb). *SEC8* is involved in cell-cell adhesion and is essential for the development of epithelial cell surface polarity. Caldesmon is a potential actomyosin regulatory protein found in smooth muscle and nonmuscle cells. Whether amplification of these genes selects for tumor development remains speculation. There are several interesting cancer-mutated genes in this region such as *SMO* (smoothened, at 128.63 Mb) which was previously associated with basal-like carcinomas [27] and the *BRAF* oncogene (140.17 Mb) which regulates proliferation differentiation, senescence and apoptosis through the RAS/ERK pathway. Since these two genes roughly mark the edges of the common amplicon it remains unclear whether these genes are involved in driving the development of this gain.

The overlapping gain on the telomeric end human chromosome 12p maps to the telomeric part of mouse chromosome 6. The mouse tumors have a small peak on *CLEC12A*, but the actual peak of the amplicon is at the end of chromosome 6, which encompasses the oncogenic *Kras* gene; however, this region is not part of the human-mouse overlap. The human tumors have a peak on *STRAP* (alias *MAWD*, at 15.93 Mb) which has previously been shown to be overexpressed in breast cancer [28]. The driving gene(s) of this overlapping gain might therefore differ between mouse and human breast tumors.

Loss of human chromosome 5q is strongly associated with *BRCA1*-mutated tumors [29]. This region might therefore contain tumor suppressor gene(s) that collaborate with *BRCA1*-loss in tumorigenesis. The first part of human chromosome 5q maps to the telomeric half of mouse chromosome 13, which is not lost significantly more often in the mouse $Brca1^{\Delta/\Delta};p53^{\Delta/\Delta}$ tumors compared to the $p53^{\Delta/\Delta}$ tumors, with the exception of the telomeric region. The human peak (which is only a small peak adjacent to a more intense peak) and the mouse peak localize at the region encompassing the *MAP3K1* and *PLK2* genes. There are two chromosome 5 synteny mapping to consecutive regions on mouse chromosome 18. The *SPINK7* gene (also known as *ECRG2*, esophageal cancer related gene 2) maps close to the peak of both the mouse and human loss. Again, its involvement in driving the chromosome 5 loss remains to be established.

The *BRCA1*-specific loss on human chromosome 12 maps to mouse chromosome 10 and consists of three separate losses. The first part peaks at the *GLS2* gene, whereas the mouse loss peaks at the *ITGA7* gene. This



loss contains several interesting genes such as *CD63* (melanoma-associated antigen), *CDK2*, *ERBB3*, *CDK4* and the *DDIT3*, but since these genes are part of the mitogenic signaling pathways, it is unlikely that these genes are drivers of this loss. The second loss on human chromosome 12 is a peak region encompassing five genes, of which one, *KITLG*, is listed as a putatively cancer related gene. The third loss on chromosome human 12 also maps to mouse chromosome 10. The mouse loss peaks on *APAF1*, involved in apoptosis, and the human tumors show a peak at the end of the loss, just outside the overlap.

The human *BRCA2*-mutated tumor group is very comparable to the control tumor group, and therefore shows only one overlapping region with mouse *Brca2*^{Δ/Δ},*p53*^{Δ/Δ}

tumors. The human chromosome 13 is one consecutive loss previously associated with *RB1*, which maps in different parts to mouse chromosome 14. Indeed, *RB1* (47.95 Mb) maps to this loss, close to the peaks of both mouse (*TSC22D1* at 43.9 Mb) and human tumors (*OLFM4* at 52.5 Mb).

Discussion

During tumor development, acquired genetic aberrations that confer selective advantage will become enriched in the tumor cell population. To identify the drivers underlying the gains and losses selected for during breast cancer development, we made use of the synteny between human and mouse genomes [30]. Syntenic gains and losses that occur in both the mouse and human

	Human		Mouse			Orthology				oncogenes in region of overlap bold text: annotated cancer gene / cancer gene census plain text: putative oncogenes in region of overlap
	Chr.	Start (Mb) End (Mb)	Chr.	Start (Mb) End (Mb)	Chr.	Start (Mb) End (Mb)	H	M	pair	
a. BRCA1										
Gains	1	67.40 68.07	6	66.95 67.42	5	5	5	5	-1	IL23R, IL12RB2, SERBP1, GADD45A, GNG12
		166.93 168.31	1	165.62 166.66	14	14	14	14	-1	ATP1B1, BLZF1, F5, SELP, SELL, SELE
	6	56.43 57.62	1	33.40 34.25	6	6	6	6	-1	DST, ZNF451, BAG2, RAB23, PRIM2
	7	100.52 102.09	5	136.37 143.49	25	21	28	-1	SERPINE1, AP1S1, VGF, RABL5, CUX1, ALKBH4, AC105052.3-2	
		125.87 142.43	6	27.23 41.70	107	109	109		ARF5, PAX4, SND1, LRRC4, LEP, IMPDH1, CCDC136, IRF5, TNPO3, SMO , AHCYL2, NRF1, UBE2H, ZC3HC1, MEST, COPG2, MKLN1, PODXL, EXOC4 , AKR1B1, AKR1B10, BPGM, CALD1 , NUP205, PTN, DGKI, CREB3L2 , TRIM24 , KIAA1549 , HIPK2, RAB19, ADCK2, BRAF , TAS2R38, EPHB6, TRPV6 GIMAP4, TMEM176A	
		149.90 150.19	6	48.61 48.86	7	11	11	1		
	8	117.73 120.03	15	51.62 54.11	9	9	9	1		EIF3H, RAD21, EXT1 , TNFRSF11B
	12	0.17 2.68	6	118.56 121.33	16	16	16	-1		JARID1A , B4GALNT3, RAD52, ERC1 , WNT5B, ADIPOR2
		9.80 18.78	6	129.23 140.01	66	90	103	1		CD69, CLECK12A , CLEC1B, CLEC7A, AC022075.29 , STYK1, CSDA, BCL2L14, LRP6, LOH12CR1, CREBL2, GPR19, CDKN1B , DDX47, GPRC5A, EMP1, GRIN2B, GUCY2C, ARHGDI, RERG, PTPRO, EPS8, STRAP , MGST1, LMO3, PLCZ1
	13	102.05 102.35	1	43.88 44.16	7	7	7	1		ERCC5
	Losses	4	103.77 105.64	3	134.17 135.51	7	7	7	-1	MANBA, UBE2D3, NHEDC2, BDH2, CENPE , CXXC4
			123.75 129.43	3	37.41 41.18	15	16	16	1	FGF2, NUDT6, FAT4, HSPA4L, PLK4
			159.27 163.30	3	76.16 80.03	9	9	9	-1	PPID, RAPGEF2
		5	50.00 58.18	13	111.18 118.14	31	31	31	-1	ITGA1, ITGA2, FST, HSPB3, ESM1, CCNO, DHX29, PPAP2A, DDX4, IL31RA, IL6ST , MAP3K1 , PLK2 , RAB3C
			102.91 108.55	17	58.68 63.82	4	4	4	1	EFNA5, FER
	119.83 129.55	18	51.24 59.54	30	30	30	1	LOX, CSNK1G3, ZNF608, ALDH7A1, SLC12A2		
	147.63 148.74	18	61.81 62.74	11	11	11	-1	SPINK7 , FBXO38, ADRB2		
12	53.96 61.83	10	121.85 129.14	100	113	113	-1	ITGA7, BLOC1S1, CD63, GDF11, SARNP , DNAJC14, MMP19, DGKA, SILV, CDK2, RAB5B, ERBB3, PA2G4, SMARCC2, RNF41, OBFC2B, CS, IL23A, STAT2, TIMELESS, GLS2 , PTGES3, RDH16, ZBTB39, MYO1A, STAT6, LRP1, INHBC, GLI1 , ARHGAP9, DDIT3 , PIP4K2C, GEFT, B4GALNT1, OS9, TSPAN31, CDK4 , CYP27B1, FAM119B, CTDSP2, XRCC6BP1, LRIG3, SLC16A7, PPM1H		
	86.90 87.50	10	99.45 100.05	5	5	5	-1	KITLG		
	94.86 103.06	10	82.18 92.97	39	39	39	-1	LTA4H, PCTK2, TMPO, SCYL2 , NR1H4, UTP20, ARL1, SPIC, CHPT1 AC084398.25-1 , IGF1, ASCL1 , NT5DC3, TDG, NFYB		
14	51.80 57.69	14	43.77 48.70	34	34	34	1	PTGER2, ERO1L, BMP4, CDKN3, CGRRF1, GCH1, MAPK11P1L, LGALS3, DLGAP5, FBXO34, KTN1 , PELI2, OTX2		
15	49.76 50.87	9	74.65 75.47	13	13	13	-1	TMOD2, TMOD3, BCL2L10, GNB5, ONECUT1		
X	53.98 56.61	X	146.86 148.84	16	16	16	1	FGD1, MAGED2, TRO, APEX2, MAGEH1, KLF8		
	57.63 65.78	X	90.98 93.58	13	12	13	1	ARHGFE9 , FAM123B , ZC4H2, MSN , EDA2R		
	68.64 74.06	X	96.02 100.40	41	41	41	1	EDA, ARR3, KIF4A, DLG3, FOXO4 , IL2RG, GJB1, NONO , TAF1, OGT, CXCR3 , ERCC6L, CITED1, HDAC8 , CHIC1, RNF12		
	113.72 114.37	X	142.21 142.76	3	3	3	1	IL13RA2		
b. BRCA2										
Losses	13	40.40 48.68	14	71.27 78.32	42	43	42	-1	C13orf15, AKAP11, TNFSF11, EPST11, DNAJC15, TSC22D1 , SPERT, CPB2, LCP1 , HTR2A, ITM2B, RB1 , P2RY5, RCBTB2	
		48.72 51.23	14	58.51 61.91	17	18	18	1	PHF11, RCBTB1, ARL11, KPNA3, TRIM13, KCNRG, RNASEH2B, INTS6	
		52.12 52.52	14	78.32 78.76	4	4	4	1	PCDH8, OLFM4	
		57.10 60.90	14	83.25 87.32	4	4	4	1	no cancer related genes in region	

Figure 8 Genes in overlapping recurrent CNAs of human and mouse breast tumor determined by cross species comparative-KC-SMART analyses. Overview of genes in regions found by cross species comparative-KC-SMART analysis (Figure 7). Listed are genes that map to significantly differentially gains and losses that overlap (a) between mouse *Brca1^{Δ/Δ};p53^{Δ/Δ}* mammary tumors (compared to *p53^{Δ/Δ}* tumors) and human *BRCA1*-mutated breast tumors (compared to control tumors) and (b) between mouse *Brca2^{Δ/Δ};p53^{Δ/Δ}* mammary tumors (compared to *p53^{Δ/Δ}* tumors) and *BRCA2*-mutated breast tumors (compared to control tumors). Idem Figure 6. For complete list see Additional file 9.

oncogenomes may harbor evolutionary conserved genes or genetic elements that are causal to tumor development in both species. Indeed, comparative aCGH analyses showed that many genetic aberrations are conserved between human and mouse neuroblastomas [31,32], ovarian cancers [33], hepatocellular carcinomas [34], melanomas [35] and T-cell lymphomas [36]. In this work, we analyzed and compared aCGH data from a large series of mouse and human *BRCA1*- and *BRCA2*-mutated breast cancers, in order to obtain a complete collection of genes and loci that are recurrently gained or lost in tumors of both species. We find that

amplification of the *MYC* locus and the loss of the *RB/INTS6* locus is evolutionarily conserved in mouse and human mammary tumor development, and that *AURKA* amplification is conserved in mouse and human *BRCA2*-mutated tumors. However, aside from these similarities, we also find recurrent aberrations in mouse and human oncogenomes that do not overlap.

Brca1 and Brca2 loss intensifies aCGH patterns of mouse *p53^{Δ/Δ}* mammary tumors

Because mouse models enable targeted deletion of specific gene(s) in a defined genetic background it is possible

to assess the specific function of this gene within an organism. With our mouse models, we analyzed the impact of BRCA1 or BRCA2 loss on the oncogenome, in a p53 deficient background. KC-SMART profiles of mouse *Brca1*^{Δ/Δ};p53^{Δ/Δ} and *Brca2*^{Δ/Δ};p53^{Δ/Δ} mammary tumors showed that significantly recurrent genomic aberrations largely overlap with recurrent aberrations in p53^{Δ/Δ} tumors and that these common aberrations occur more frequently in *Brca1*^{Δ/Δ};p53^{Δ/Δ} and *Brca2*^{Δ/Δ};p53^{Δ/Δ} mammary tumor compared to the p53^{Δ/Δ} tumor group. This suggests that BRCA1 or BRCA2 loss intensifies the profile of aberrations seen in the p53^{Δ/Δ} tumor group.

We find that the p53^{Δ/Δ} tumors that are BRCA proficient more often gain or lose entire chromosomes during tumor development. In contrast, *Brca1*^{Δ/Δ};p53^{Δ/Δ} and *Brca2*^{Δ/Δ};p53^{Δ/Δ} tumors have an increased amount of partial chromosomal gains and losses, presumably because they have accumulated more inter- and intra-chromosomal DNA damage as a result of error-prone DSB repair in the absence of homology directed DNA repair. The inability of these cells to properly repair DSBs might allow selection of gains and losses more specifically tailored around the relevant genes in a region. Interestingly, most gains and losses found in the p53^{Δ/Δ} tumor group are intensified by *Brca1* or *Brca2* loss but some are not, such as the gain on chromosome 15 and the losses on chromosomes 7 and 12. These regions may therefore be subject to a differential selection pressure during tumor development.

It appears that cell type also plays an important role: the carcinomas of the p53^{Δ/Δ} tumor group seems to have more genomic damage compared with the carcinosarcomas. This may be the reason that BRCA1 and BRCA2 loss gives rise to a more aggravated DNA damage profile, as >90% of *Brca1*^{Δ/Δ};p53^{Δ/Δ} tumors and *Brca2*^{Δ/Δ};p53^{Δ/Δ} mammary tumors are carcinomas, compared to half of the p53^{Δ/Δ} mammary tumors, of which the other half are sarcomatoid. This suggests that in these mouse models BRCA1 loss and BRCA2 loss blocks cellular differentiation as suggested previously [37], or alternatively, that BRCA1 and BRCA2 loss are not tolerated in more differentiated cells.

Differences between mouse *Brca1* and *Brca2* mutated tumors

Despite their common p53 deficient background, some aberrations may be more prone to arise in a BRCA1 deficient cell, whereas others are more likely to occur in a BRCA2 deficient cell. For example, the recurrent gain on chromosome 1 or the losses on chromosomes 4 and 9 of the *Brca1*^{Δ/Δ};p53^{Δ/Δ} tumors are less common in the *Brca2*^{Δ/Δ};p53^{Δ/Δ} tumors. Likewise, the prominent gain on chromosome 10 is found in half of all *Brca2*^{Δ/Δ};p53^{Δ/Δ}

tumors, and in approximately 1/8 of the *Brca1*^{Δ/Δ};p53^{Δ/Δ} tumors. In addition, it appears that in a p53 deficient background, *Brca1* loss gives rise to more genomic aberrations than does *Brca2* loss. Together, these differences in tumor profiles suggest that BRCA1 and BRCA2 have different functions in DNA repair pathways, however, we have not been able to identify the nature of this difference using our aCGH profiles.

Comparative oncogenomics identifies similarities and differences in mouse and human breast tumors

Cross species comparison of genomic aberrations in human and mouse tumors may identify conserved genomic regions likely to be important for development in both species. Tumors from genetically engineered mouse models often do not show the high levels of chromosome instability associated with human cancers [36], presumably because the engineered mutations predispose to rapid tumor formation and thereby circumvent the long incubation time required for accumulation of genomic mutations in human tumorigenesis. The increased amounts of genomic alterations in tumors from mice with engineered telomere dysfunction or defects in DNA damage checkpoints or DNA repair allows comparison of genomic alterations with human tumors that would not be possible in mouse models without chromosomal instability [36].

In our analysis, we compared aCGH profiles from human *BRCA1/2*-mutated and sporadic breast cancers with profiles from mouse p53^{Δ/Δ}, *Brca1*^{Δ/Δ};p53^{Δ/Δ} and *Brca2*^{Δ/Δ};p53^{Δ/Δ} mammary tumors. We assume that DNA damage repair deficiency in the mammary epithelium of these mice allows a comparable selective advantage in tumor formation as in human breast epithelial cells. Indeed, all human tumor groups showed a recurrent gain on chromosome 8q, likewise all mouse tumor groups showed a recurrent gain on its syntenic region, on chromosome 15, centering exactly on the *MYC* oncogene. This suggests that *MYC* gain is evolutionarily conserved in mammary tumor development from mouse to man. Likewise, the chromosome 13 loss, associated with *RB1* loss [38] occurs in all human and mouse tumor groups, and all overlapping regions encompass the *INTS6* gene. Similar to *MYC* gain, *RB1/INTS6* loss could be evolutionarily conserved in mammary tumor development.

The defined initiating lesions and similar genetic background of mouse mammary tumors permits identification of recurrent genomic aberrations linked to certain genetic predispositions. Indeed, the amplification peak on human chromosome 20 overlapped with the syntenic amplification peak on mouse chromosome 2 of the *Brca1*^{Δ/Δ};p53^{Δ/Δ} and the *Brca2*^{Δ/Δ};p53^{Δ/Δ} tumors, at the *Aurora Kinase A* gene. The *AURKA* gain has been

previously associated with *BRCA2*-mutated breast tumor development [39] but our findings suggest that the functional interaction between *AURKA* and *BRCA2* may be evolutionarily conserved. The mouse *Brca1^{Δ/Δ};p53^{Δ/Δ}* tumors also showed the gain on chromosome 2 but the *p53^{Δ/Δ}* tumors, on the other hand, did not show this gain at all, suggesting that this gain may selectively develop in a homologous repair deficient (HRD) background.

For most other overlapping gains and losses reported in the results section, it is still difficult to associate a driver gene with a gain or a loss. Also, gains known to be important in human *BRCA1*-mutated breast cancer show no (e.g. 3q gain) or only partial (e.g. the 5q loss) overlap with the mouse gains and losses. The human 3q gain is associated with *PIK3CA* overexpression; possibly, mouse tumors may find alternative ways to overexpress this gene or a functional homologue. Alternatively, *PIK3CA* overexpression might not provide a selective advantage in these mouse tumors, or mouse tumors in general. The regions susceptible for amplification or deletion in the mouse genome and the human genome may not be the same due to, for example, different nuclear chromosomal folding or genomic organization of synteny blocks, resulting in a differential development of the oncogenome during mouse and human tumor development. These findings add to previous findings of differences between mouse and human breast cancer development such as that activation of *Wnt* and/or its pathway members are commonly found in mouse but not in human spontaneous breast tumors, differences in metastasis patterns, or that few mouse models accurately represent ER positive human breast cancers [40].

Conclusions

In this work we have found some important genotypic similarities between mouse and human cognate cancers, such as the *MYC* gain and the *RB/INTS6* loss, but also important differences such as the chromosome 10 gain in mouse *Brca2^{Δ/Δ};p53^{Δ/Δ}* tumors and the 3q gain in human *BRCA1*-mutated tumors. In our mouse models, we find that *BRCA1* and *BRCA2* loss clearly have different impacts on tumor development. However, the distinct features of mouse *Brca1^{Δ/Δ};p53^{Δ/Δ}* and *Brca2^{Δ/Δ};p53^{Δ/Δ}* tumors did not translate to distinct features of human *BRCA1* and *BRCA2* mutated tumors and vice versa. A reason for this could be that tumor development in our mouse models is a fast, accelerated version of the relatively slow process of tumor development in humans. Simultaneous and homozygous deletion of the mouse *Brca1*, *Brca2* and *p53* genes leads to immediate cellular transformation, whereas human germline *BRCA1* and *BRCA2* mutations are initially heterozygous [41,42]. Human *TP53* mutations are somatic and can be

acquired in myriad different ways, possibly resulting in different phenotypes [43]. In addition, the difference in make-up of the mouse and human genome, such as the genome wide organization of syntenic regions, and/or overall differences between mouse and human mammary gland biology may result in selection of other regions and genes during the course of mouse and human breast cancer development. These differences should be taken into account when using *K14cre;Brca1^{F/F};p53^{F/F}* and *K14cre;Brca2^{F/F};p53^{F/F}* mouse models for preclinical studies.

Additional material

Additional file 1: Histopathology, BRCA1/BRCA2 mutations, tumor types of human breast tumors.

Immunohistochemistry: Presence of ER, PR, ERBB2 (HER2/neu) and TP53 was determined by immunohistochemistry staining using the antibodies: estrogen receptor AB-14 clone 1D5 + 6F11, titre 1:50 (Neomarkers); progesterone receptor clone PR-1 titre 1:400 (Immunologic), c-erbB-2 clone SP3, titre 1:25 (Neomarkers); TP53 clone D0-7, titre 1:8000 (Dako). If >70% of the tumor cells expressed ER, PR, or TP53, the tumor was scored as positive (+) for the corresponding staining, in case <10% of the cells were stained, the tumor was scored as negative (-) and between 10 and 70% the tumor was scored as ± for the corresponding staining. HER2/neu staining was scored positive when a 3+ staining was observed (*), otherwise it was scored negative (only one sporadic control case was IHC 2+, and was called negative). *) Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J Clin Oncol* 2007 Jan 1;25(1):118-45.

Additional file 2: Tumor type characterization of mouse mammary tumors by Vimentin and E-Cadherin expression analysis.

To determine tumor type, log₂ ratios of Vimentin and E-Cadherin expression were acquired by hybridization of mRNA of a subset of tumors on Mouse Operon V3 Oligo arrays and were plotted against each other. a) *Brca1^{Δ/Δ};p53^{Δ/Δ}*, *Brca2^{Δ/Δ};p53^{Δ/Δ}* and *p53^{Δ/Δ}* mouse tumors were plotted by their E-cadherin and Vimentin expression levels. b) If the difference between E-cadherin and Vimentin expression levels was less than 0.5 on a log scale these tumors were called 'ambiguous'. Tumors were scored 'Mesenchymal' if the Vimentin expression level was >0.5 higher than E-cadherin expression level. Likewise, tumors were scored 'Epithelial' when the E-cadherin expression level was >0.5 higher than the Vimentin expression.

Additional file 3: Tumor types of mouse mammary tumors determined by histopathology and gene expression analysis.

Mouse mammary tumor types were determined by histopathology and expression data. Histopathology was done as previously described in [10]. Expression data was acquired by hybridization of mRNA of a subset of tumors on Mouse Operon V3 Oligo arrays: If the difference between E-cadherin and Vimentin expression levels was less than 0.5 on a log scale we called these tumors 'ambiguous'. Tumors were scored 'Mesenchymal' if the Vimentin expression level was >0.5 higher than E-cadherin expression level. Likewise, tumors were scored 'Epithelial' when the E-cadherin expression level was >0.5 higher than the Vimentin expression level. The line numbers relate to the line-plots shown in Figure 2, line number 1 on top for each tumor group.

Additional file 4: KC-SMART analysis of mouse mammary tumors.

The significant CNAs of *Brca1^{Δ/Δ};p53^{Δ/Δ}*, *Brca2^{Δ/Δ};p53^{Δ/Δ}* and *p53^{Δ/Δ}* mouse tumor groups were obtained by running the KC-SMART algorithm over the BAC data from all tumors in each tumor group using a kernel width of 20 Mb. Significant regions are determined by the intercept of the KSE curve and the significance cutoff calculated for each tumor group, and for gains and losses separately. The upper panel lists recurrent gains and

the bottom panel lists the recurrent losses. These regions correspond to the color-matched bars on top (gains) and on the bottom (losses) of the KSE curves of the tumor groups shown in Figure 1d and the red bars of the cross species comparison in Figure 5.

Additional file 5: KC-SMART analysis of human breast tumors.

Genomic regions found recurrently gained (top) or lost (bottom) by KC-SMART analysis of human BRCA1-related, BRCA2-related, and control breast tumors. The significant CNAs were determined by running the KC-SMART algorithm over the BAC data from all tumors in each tumor group using a kernel width of 20 Mb. Significant regions are determined by the intercept of the KSE curve and the significance cutoff calculated for each tumor group, and for gains and losses separately. The upper panel lists recurrent gains and the bottom panel lists the recurrent losses.

Additional file 6: Regions identified by cross-species KC-SMART analysis.

Shown are genes within significantly recurrent regions of gain and loss that overlap between (a) human BRCA1-mutated and mouse *Brca1*^{Δ/Δ}, *p53*^{Δ/Δ} tumors, (b) human BRCA2-mutated and mouse *Brca2*^{Δ/Δ}, *p53*^{Δ/Δ} tumors, and (c) human control tumors and mouse *p53*^{Δ/Δ} tumors as shown in Figure 5. The syntenic regions in the human and the mouse genome are listed in columns 1 and 2. The "Orthology" column lists the number of mouse (M) and human (H) genes mapping within each region of overlap, and the number of unique orthologous pairs (pair). Strand inversion between the two species is indicated by 1 (no inversion) or -1 (inversion). Listed are cancer-related genes included in the Atlas of Genetics and Cytogenetics in Oncology and Haematology [44,45], and the Cancer Gene Census (CGC)[47,46]. Annotated cancer genes and CGC genes are shown in bold type, putative cancer genes are shown in normal type. Cancer-related genes that map closest to the human KSE peak are shown in blue and genes that map closest to the mouse KSE peak are shown in red. Genes that map to the mouse AND human KSE peaks are shown in green. Single genes that were not listed in the Atlas or the CGC are shown in italics.

Additional file 7: Comparative-KC-SMART analysis of mouse mammary tumors.

The significant differential gains and losses of *Brca1*^{Δ/Δ}, *p53*^{Δ/Δ} vs. *p53*^{Δ/Δ} and *Brca2*^{Δ/Δ}, *p53*^{Δ/Δ} vs. *p53*^{Δ/Δ} mouse tumor groups were obtained by running the comparative-KC-SMART algorithm over the BAC data from all tumors in each tumor group using a kernel width of 20 Mb. The upper panel shows recurrent differential gains and the bottom panel shows the differential losses. These regions correspond to the red bars plotted on top (gains) and on the bottom (losses) of the corresponding mouse KSE curves used in the cross species comparison in Figure 6.

Additional file 8: Comparative-KC-SMART analysis of human breast tumors.

Comparative KC-SMART analysis of human BRCA1-related and BRCA2 related breast tumors compared with control breast tumors. Regions of differential aberrations between BRCA1 and control breast tumors, between BRCA2 and control breast tumors. Top panel: Differential gains. Bottom panel: Differential losses.

Additional file 9: Regions identified by cross-species comparative-KC-SMART analysis.

Shown are (a) genes that map to differentially recurrent CNAs found in both mouse *Brca1*^{Δ/Δ}, *p53*^{Δ/Δ} vs. *p53*^{Δ/Δ} mammary tumors and human BRCA1-mutated vs. control breast tumors and (b) genes mapping to differentially recurrent CNAs found in both mouse *Brca2*^{Δ/Δ}, *p53*^{Δ/Δ} vs. *p53*^{Δ/Δ} mammary tumors and BRCA2-mutated vs. control breast tumors, shown in Figure 7

List of Abbreviations

aCGH: array comparative genomic hybridization; CNA: (genomic) copy number alterations; DSB: double strand break; ER: Estrogen Receptor; KC-SMART: Kernel Convolution - a Statistical Method for Aberrant Region detection; GEM: genetically engineered mouse; IDC-noS: invasive ductal carcinoma not otherwise specified; KSE: Kernel Smoothed Estimate; BAC: Bacterial Artificial Chromosome; HRD: homologous recombination deficient; SNR: signal to noise ratio; FDR: False Discovery Rate

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

HH, EvB, PN and JJ conceived and designed the study. HH, EvB, AV, XL, SAJ, SK, ES, RK and JJ performed experiments. HH, EvB, PN and JJ analyzed the data. CNK, AV and LFAW contributed analysis tools. HH, EvB, PN and JJ wrote the paper. All authors read and approved the final manuscript.

Acknowledgements

This work was supported by grants from the Dutch Cancer Society (NKI 2002-2635), the Netherlands Organization for Scientific Research (NWO Vidi 917.036.347) and the Susan G. Komen Breast Cancer Foundation (BCTR0403230). The authors thank Dr. Piet Borst and Dr. Hein te Riele for critically reading this manuscript.

Author details

¹Division of Molecular Biology, Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands. ²Division of Clinical Genetics, VU Medical Center, de Boelelaan 1118, 1081 HV Amsterdam, The Netherlands. ³Division of Experimental Therapy, Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands. ⁴Skyline-Diagnostics B.V., Dr. Molewaterplein 50, 3015 GE Rotterdam, The Netherlands. ⁵Central Microarray Facility, Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands. ⁶Biomedical Analysis Center, Tsinghua University, Beijing, 100084, China. ⁷Institute of Tumor Biology, University Medical Center Hamburg Eppendorf, Martinistrasse 52, 20246 Hamburg, Germany. ⁸Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Delft, The Netherlands.

Received: 10 February 2010 Accepted: 24 August 2010

Published: 24 August 2010

References

1. King MC, Marks JH, Mandell JB: **Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2.** *Science* 2003, **302**(5645):643-646.
2. Paakkonen K, Sauramo S, Sarantausta L, Vahteristo P, Hartikainen A, Vehmanen P, Ignatius J, Ollikainen V, Kaariainen H, Vauramo E, et al: **Involvement of BRCA1 and BRCA2 in breast cancer in a western Finnish sub-population.** *Genet Epidemiol* 2001, **20**(2):239-246.
3. Hofmann W, Schlag PM: **BRCA1 and BRCA2—breast cancer susceptibility genes.** *J Cancer Res Clin Oncol* 2000, **126**(9):487-496.
4. Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, Collins N, Gregory S, Gumbs C, Micklem G: **Identification of the breast cancer susceptibility gene BRCA2.** *Nature* 1995, **378**(6559):789-792.
5. Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, Cochran C, Bennett LM, Ding W, et al: **A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1.** *Science* 1994, **266**(5182):66-71.
6. Turner N, Tutt A, Ashworth A: **Hallmarks of 'BRCAness' in sporadic cancers.** *Nat Rev Cancer* 2004, **4**(10):814-819.
7. Loffler H, Lukas J, Bartek J, Kramer A: **Structure meets function—centrosomes, genome maintenance and the DNA damage response.** *Exp Cell Res* 2006, **312**(14):2633-2640.
8. Bertwistle D, Ashworth A: **The pathology of familial breast cancer: How do the functions of BRCA1 and BRCA2 relate to breast tumour pathology?** *Breast Cancer Res* 1999, **1**(1):41-47.
9. DeVries S, Gray JW, Pinkel D, Waldman FM, Sudar D: **Comparative genomic hybridization.** *Curr Protoc Hum Genet* 2001, **Chapter 4**(Unit4):6.
10. Liu X, Holstege H, van der Gulden H, Treur-Mulder M, Zevenhoven J, Velds A, Kerkhoven RM, van Vliet MH, Wessels LF, Peterse JL, et al: **Somatic loss of BRCA1 and p53 in mice induces mammary tumors with features of human BRCA1-mutated basal-like breast cancer.** *Proc Natl Acad Sci USA* 2007, **104**(29):12111-12116.
11. Jonkers J, Meuwissen R, van der Gulden H, Peterse H, van der Valk M, Berns A: **Synergistic tumor suppressor activity of BRCA2 and p53 in a conditional mouse model for breast cancer.** *Nat Genet* 2001, **29**(4):418-425.
12. Lakhani SR, Van De Vijver MJ, Jacquemier J, Anderson TJ, Osin PP, McGuffog L, Easton DF: **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* 2002, **20**(9):2310-2318.

13. Foulkes WD, Stefansson IM, Chappuis PO, Begin LR, Goffin JR, Wong N, Trudel M, Akslen LA: **Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer.** *J Natl Cancer Inst* 2003, **95**(19):1482-1485.
14. Althuis MD, Fergenbaum JH, Garcia-Closas M, Brinton LA, Madigan MP, Sherman ME: **Etiology of hormone receptor-defined breast cancer: a systematic review of the literature.** *Cancer Epidemiol Biomarkers Prev* 2004, **13**(10):1558-1568.
15. Bane AL, Beck JC, Bleiweiss I, Buys SS, Catalano E, Daly MB, Giles G, Godwin AK, Htibshoosh H, Hopper JL, et al: **BRCA2 mutation-associated breast cancers exhibit a distinguishing phenotype based on morphology and molecular profiles from tissue microarrays.** *Am J Surg Pathol* 2007, **31**(1):121-128.
16. Adelaide J, Finetti P, Bekhouche I, Repellini L, Geneix J, Sircoulomb F, Charafe-Jauffret E, Cervera N, Desplans J, Parzy D, et al: **Integrated profiling of basal and luminal breast cancers.** *Cancer Res* 2007, **67**(24):11565-11575.
17. Peeper D, Berns A: **Cross-species oncogenomics in cancer gene identification.** *Cell* 2006, **125**(7):1230-1233.
18. Klijn C, Holstege H, de Ridder J, Liu X, Reinders M, Jonkers J, Wessels L: **Identification of cancer genes using a statistical framework for multiexperiment analysis of nondiscretized array CGH data.** *Nucleic Acids Res* 2008, **36**(2):e13.
19. Chung YJ, Jonkers J, Kitson H, Fiegler H, Humphray S, Scott C, Hunt S, Yu Y, Nishijima I, Velds A, et al: **A whole-genome mouse BAC microarray with 1-Mb resolution for analysis of DNA copy number changes by array comparative genomic hybridization.** *Genome Res* 2004, **14**(1):188-196.
20. **Array Express Archive.** [http://www.ebi.ac.uk/microarray-as/ae/].
21. Weng L, Dai H, Zhan Y, He Y, Stepaniants SB, Bassett DE: **Rosetta error model for gene expression analysis.** *Bioinformatics* 2006, **22**(9):1111-1121.
22. Joosse SA, Brandwijk KI, Devilee P, Wesseling J, Hogervorst FB, Verhoef S, Nederlof PM: **Prediction of BRCA2-association in hereditary breast carcinomas using array-CGH.** *Breast Cancer Res Treat* 2010.
23. Joosse SA, van Beers EH, Tielen IH, Horlings H, Peterse JL, Hoogerbrugge N, Ligtenberg MJ, Wessels LF, Axwijk P, Verhoef S, et al: **Prediction of BRCA1-association in hereditary non-BRCA1/2 breast carcinomas with array-CGH.** *Breast Cancer Res Treat* 2009, **116**(3):479-489.
24. **Bioconductor.** [http://www.bioconductor.org/packages/2.4/bioc/html/KCsmart.html].
25. Harkin DP, Bean JM, Miklos D, Song YH, Truong VB, Englert C, Christians FC, Ellisen LW, Maheswaran S, Oliner JD, et al: **Induction of GADD45 and JNK/SAPK-dependent apoptosis following inducible expression of BRCA1.** *Cell* 1999, **97**(5):575-586.
26. Wang X, Wang RH, Li W, Xu X, Hollander MC, Fornace AJ Jr, Deng CX: **Genetic interactions between Brca1 and Gadd45a in centrosome duplication, genetic stability, and neural tube closure.** *J Biol Chem* 2004, **279**(28):29606-29614.
27. Xie J, Murone M, Luoh SM, Ryan A, Gu Q, Zhang C, Bonifas JM, Lam CW, Hynes M, Goddard A, et al: **Activating Smoothed mutations in sporadic basal-cell carcinoma.** *Nature* 1998, **391**(6662):90-92.
28. Matsuda S, Katsumata R, Okuda T, Yamamoto T, Miyazaki K, Senga T, Machida K, Thant AA, Nakatsugawa S, Hamaguchi M: **Molecular cloning and characterization of human MAWD, a novel protein containing WD-40 repeats frequently overexpressed in breast cancer.** *Cancer Res* 2000, **60**(1):13-17.
29. Wessels LF, van Welsem T, Hart AA, van't Veer LJ, Reinders MJ, Nederlof PM: **Molecular classification of breast carcinomas by comparative genomic hybridization: a specific somatic genetic profile for BRCA1 tumors.** *Cancer Res* 2002, **62**(23):7110-7117.
30. Waterston RH, Lindblad-Toh K, Birney E, Rogers J, Abril JF, Agarwal P, Agarwala R, Ainscough R, Alexandersson M, An P, et al: **Initial sequencing and comparative analysis of the mouse genome.** *Nature* 2002, **420**(6915):520-562.
31. Cheng AJ, Cheng NC, Ford J, Smith J, Murray JE, Flemming C, Lastowska M, Jackson MS, Hackett CS, Weiss WA, et al: **Cell lines from MYCN transgenic murine tumours reflect the molecular and biological characteristics of human neuroblastoma.** *Eur J Cancer* 2007, **43**(9):1467-1475.
32. Hackett CS, Hodgson JG, Law ME, Fridlyand J, Osoegawa K, de Jong PJ, Nowak NJ, Pinkel D, Albertson DG, Jain A, et al: **Genome-wide array CGH analysis of murine neuroblastoma reveals distinct genomic aberrations which parallel those in human tumors.** *Cancer Res* 2003, **63**(17):5266-5273.
33. Urzua U, Frankenberger C, Gangi L, Mayer S, Burkett S, Munroe DJ: **Microarray comparative genomic hybridization profile of a murine model for epithelial ovarian cancer reveals genomic imbalances resembling human ovarian carcinomas.** *Tumour Biol* 2005, **26**(5):236-244.
34. Zender L, Spector MS, Xue W, Flemming P, Cordon-Cardo C, Silke J, Fan ST, Luk JM, Wigler M, Hannon GJ, et al: **Identification and validation of oncogenes in liver cancer using an integrative oncogenic approach.** *Cell* 2006, **125**(7):1253-1267.
35. Kim M, Gans JD, Nogueira C, Wang A, Paik JH, Feng B, Brennan C, Hahn WC, Cordon-Cardo C, Wagner SN, et al: **Comparative oncogenomics identifies NEDD9 as a melanoma metastasis gene.** *Cell* 2006, **125**(7):1269-1281.
36. Maser RS, Choudhury B, Campbell PJ, Feng B, Wong KK, Protopopov A, O'Neil J, Gutierrez A, Ivanova E, Perna I, et al: **Chromosomally unstable mouse tumours have genomic alterations similar to diverse human cancers.** *Nature* 2007, **447**(7147):966-971.
37. Liu S, Ginestier C, Charafe-Jauffret E, Foco H, Kleer CG, Merajver SD, Dontu G, Wicha MS: **BRCA1 regulates human mammary stem/progenitor cell fate.** *Proc Natl Acad Sci USA* 2008, **105**(5):1680-1685.
38. Kujawski M, Rydzanicz M, Sarlomo-Rikala M, Szyfter K: **Rearrangements involving the 13q chromosome arm committed to the progression of laryngeal squamous cell carcinoma.** *Cancer Genet Cytogenet* 2002, **137**(1):54-58.
39. Bodvardsdottir SK, Hilmarsdottir H, Birgisdottir V, Steinarsdottir M, Jonasson JG, Eyfjord JE: **Aurora-A amplification associated with BRCA2 mutation in breast tumours.** *Cancer Lett* 2007, **248**(1):96-102.
40. Borowsky A: **Special considerations in mouse models of breast cancer.** *Breast Dis* 2007, **28**:29-38.
41. Smith SA, Easton DF, Evans DG, Ponder BA: **Allele losses in the region 17q12-21 in familial breast and ovarian cancer involve the wild-type chromosome.** *Nat Genet* 1992, **2**(2):128-131.
42. Collins N, McManus R, Wooster R, Mangion J, Seal S, Lakhani SR, Ormiston W, Daly PA, Ford D, Easton DF, et al: **Consistent loss of the wild type allele in breast cancers from a family linked to the BRCA2 gene on chromosome 13q12-13.** *Oncogene* 1995, **10**(8):1673-1675.
43. Petitjean A, Mathe E, Kato S, Ishioka C, Tavtigian SV, Hainaut P, Olivier M: **Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database.** *Hum Mutat* 2007, **28**(6):622-629.
44. **Atlas of Genetics and Cytogenetics in Oncology and Haematology.** [http://atlasgeneticsoncology.org/].
45. Huret JL, Dessen P, Bernheim A: **Atlas of Genetics and Cytogenetics in Oncology and Haematology, year 2003.** *Nucleic Acids Res* 2003, **31**(1):272-274.
46. Futreal PA, Coin L, Marshall M, Down T, Hubbard T, Wooster R, Rahman N, Stratton MR: **A census of human cancer genes.** *Nat Rev Cancer* 2004, **4**(3):177-183.
47. **Cancer Gene Census, Cancer Genome Project Wellcome Trust Sanger Institute.** [http://www.sanger.ac.uk/genetics/CGP/Census/].

Pre-publication history

The pre-publication history for this paper can be accessed here:
http://www.biomedcentral.com/1471-2407/10/455/prepub

doi:10.1186/1471-2407-10-455

Cite this article as: Holstege et al.: Cross-species comparison of aCGH data from mouse and human BRCA1- and BRCA2-mutated breast cancers. *BMC Cancer* 2010 **10**:455.