Discussion

In this thesis we have performed high-resolution SNP-array and gene expression profiling of 120 female familial breast tumours selected from a larger familial cohort of 155 breast tumours, including tumours from *BRCA1*, *BRCA2*, and *CHEK2* mutant tumours. We combined copy number, loss of heterozygosity (LOH) and gene expression profiling to investigate whether *BRCA1*-mutated and *CHEK2**1100delC-mutated breast tumours harbour characteristic genomic and/or transcriptional aberrations as compared to non-*BRCA1*/*2*-*CHEK2**1100delC-mutated familial breast tumours (BRCAX). In case this approach would prove successful, a similar approach for classifying BRCAX tumours might facilitate the possible discovery of yet unknown breast cancer associated risk genes and/or pathways involved in breast tumourigenesis. An important additional finding in our study is the fact that a large proportion of breast tumour samples contain varying amounts of tumour infiltrating lymphocytes (TILs). This finding proved pivotal in optimal genomic profiling of standard processed (not-sorted) tumour samples. The prognostic relevance of TILs with regard to breast cancer molecular subtypes is also discussed.

In chapter 2 we performed integrated genomic and gene expression analyses of *BRCA1*-mutated and non-*BRCA1/2*-mutated familial breast tumours (BRCAX) based on the assumption that *BRCA1*-mutated breast tumours may have distinct biological features resulting from specific oncogenic pathways involved in tumour development. As the majority of the *BRCA1*-mutated breast tumours are found to be of the Basal-like subtype [1], our analysis was restricted to Basal-like carcinomas (BLCs) with and without *BRCA1* germline mutations to avoid confounding subtype specific copy number aberrations (CNAs). Initial results indicated that there were no large differences in CNA profiles between *BRCA1*-mutated and BRCAX BLCs. However, an unanticipated inter-tumour heterogeneity regarding the copy number and LOH profiles of the *BRCA1*-mutated BLCs was observed. The frequently reported BRCAness profile (high amounts of CNAs) was only seen in roughly two thirds of the *BRCA1*-mutated BLCs [2-4]. While, the remainder of samples showed no or only very few CNAs. A similar observation was made for the BRCAX BLCs.

We questioned whether the observed heterogeneity of CNA profiles amongst the *BRCA1*-mutated BLCs could be explained by haplo-insufficiency of the *BRCA1* gene. The conservation of a functional copy of the *BRCA1* gene could be causative for the difference in CNA profiles. If tumours with few CNAs have retained a functional copy of the *BRCA1* gene

this could explain their relative stable profile. To explore this possibility, *BRCA1* mutation scanning was performed on tumour material in 6 *BRCA1*-mutated tumours with high amounts of CNAs, and 4 *BRCA1* tumours with few CNAs. In all samples the reported germline mutations could be confirmed but no additional mutations were identified. Complete loss of the *BRCA1* wild-type allele is therefore evident in the *BRCA1*-mutated tumours with many CNAs and LOH at the *BRCA1*-locus, whereas the *BRCA*1-mutated tumours with few CNAs appear to have retained a copy of the wild-type *BRCA1* allele. To further explore this unexpected observation, supervised mRNA analysis was performed.

By supervised gene expression analysis varying amounts of tumour infiltrating lymphocytes (TILs) were found to be responsible for the observed differences in CNA profiles. To confirm the presence of varying amounts of TILs amongst the BLCs, the proportion of lymphocytic nuclei within the tumour samples was assessed on H&E-stained frozen sections. The distributions of TILs throughout the sections were found to be very heterogeneous. Furthermore the amounts of TILs were seen to sometimes vary greatly amongst multiple sections taken from the same tumour, making quantification a daunting task. This could explain why very few reported genomic studies have taken TIL quantification into account in their study designs. Nevertheless, TIL percentages as scored by histopathological assessment and mRNA based immune signature values were found to be significantly correlated. An inverse correlation was observed between the amount of TILs and genome-wide levels of CNAs measured in the BLCs. Also, allelic imbalances (studied by B-allele frequencies (BAF)) were readily observed in samples with low amounts of TILs, as exemplified by frequent detection of copy neutral LOH on chromosome 17. In contrast, samples with high levels of TILs often showed aberrant or scrambled BAF-profiles even when (low-amplitude) CNAs were detectable. Therefore, we believe that the presence of large quantities of TILs hamper copy number and LOH profiling in a significant proportion of the (*BRCA1*-mutated) BLCs. Furthermore, the mutation scanning results for the *BRCA1*-mutated samples with and without complete loss of the wild type allele can be explained by the absence and presence of large quantities of TILs respectively. Interestingly, a small number of low-TIL BRCAX BLCs were found to share specific CNAs and large overlapping regions of copy neutral LOH on the long arm of chromosome 17.

After sample selection for low levels of TILs, differential regions of CNAs could be identified between the *BRCA1*-mutated and BRCAX BLCs. Some of these regions were reported before; while others were previously

not considered as significantly associated with *BRCA1*-mutation status. A combination of specific CNAs proved to be capable of discriminating between the selected *BRCA1*-mutated and *BRCAX* BLCs. These results were validated in two publically available datasets. Our classification method uses segmented data for copy number gains and losses, composed of signal intensities of 20 or more SNPs. The rationale behind this is that true oncogenic copy number changes will most likely target genomic regions containing genes which are beneficial to tumour growth, resulting in larger genomic stretches of copy number change rather than for example single SNPs or BAC-clones as used in previous reports [2,3]. Furthermore, this classification method is applicable to all copy number platforms.

Curiously, the detrimental effect that TILs have on the ability to reliably measure genomic and expression profiles in *BRCA1*-mutated samples (and tumour samples in general) is largely neglected in literature. Reports of *BRCA1*-mutated breast carcinomas with few CNAs are scarce. Perhaps this is due to the presumption that all *BRCA1*-mutated tumours are aggressive, genomically instable tumours. And samples not fulfilling these criteria are perhaps discarded. Nevertheless, genomic stable *BRCA1*-mutated BLCs are reported by Stefansson et al. [5]. These authors report a simple genomic profile subgroup resembling a *BRCA1*-related subgroup but displaying considerably less complex genomes. These tumours display gains at chromosomes 1q and 8q, typically reported as regions with the highest mean copy number gains in BLCs. Sporadic BLCs with few CNAs are reported in a study by Chin et al. [6]. Here a novel genomic stable subtype of ER negative breast cancer is reported, characterized by a 37 gene expression classifier with marginal statistical associations with immune and inflammatory responses. Also, in a recent study on chromosomal break patterns to identify *BRCA1*/*2* mutated BLCs, it is reported that 15 of the 80 SNP arrays were discarded because of low hybridization quality, low tumour content, or ambiguous profile interpretation [7]. We suspect the presence of large amounts of TILs in a number of these discarded samples.

To further evaluate the impact of TILs on copy number profiling in BLCs, the discovery set of a publically available dataset described in the introduction of this thesis was studied [8]. In this report, clustering analysis of joint copy number and gene expression data revealed 10 novel molecular subgroups, each associated with distinct CNAs and gene expression changes. The discovery set contains high resolution SNP array and gene expression data of 998 primary sporadic breast tumours including 118 BLCs, which we analyzed similar to our own samples. Again a significant correlation between immune signature levels and CNA levels was seen for

the BLCs. Interestingly, BLCs with the highest immune signature levels were found to be amongst those samples with the least reported CNAs. These samples were reported to belong to a specific subgroup of CNA devoid tumours exhibiting a strong immune and inflammation signature (molecular subgroup 4). This subgroup 4 is reported to contain both ER+ and ER- tumours from all intrinsic subtypes, with a disproportional large amount of Normal-like tumours. Importantly, this subgroup is associated with favourable outcome and a 10-year disease-specific survival of around 80%, which is in agreement with our results described in chapter 4. Our study suggests that the absence of CNAs in the BLC samples in subgroup 4 is likely due to measuring mostly genomic DNA of TILs instead of tumour DNA and therefore these BLCs are misplaced. The same probably holds true for the Normal-like samples in this subgroup.

To further substantiate and validate our findings we performed DNA flow cytometry on paraffin-embedded, formalin-fixed, material of *BRCA1*-mutated breast carcinomas selected for large numbers of TILs (> 40% of nuclei). Copy number profiles obtained by shallow whole genome sequencing analysis of sorted tumour cell derived DNA clearly showed profiles that were not detected in the matching unsorted tumour samples. To our knowledge this approach is unique in breast cancer research and the results clearly demonstrate the impact of TILs on copy number profiling. For these samples the reported pathogenic *BRCA1* germline mutations were all confirmed by Sanger sequencing in all fractions. Confirming our previous results, the mutations were seen in a heterozygous state in the unsorted and normal/immune fractions, while the mutant allele was seen predominantly in the sorted tumour fractions.

The resulting copy number profiles of the sorted tumour fractions were highly similar to those of *BRCA1*-mutated BLCs with low amounts of TILs, except for the striking observation that the sorted tumour cell fractions of the high TIL cases showed a loss of the complete long arm of chromosome 17, instead of copy neutral LOH as observed in the low TIL cases from the SNP array. This difference in acquiring complete *BRCA1* deficiency, i.e. copy neutral LOH or deletion, may thus be somehow linked to the immune response and is an appealing subject for further research.

Our findings have direct consequences on current clinical genetic tools for assessing *BRCA1* involvement in breast cancer patients and pathogenicity assessment of *BRCA1* variants of unknown significance [9,10]. This holds true for both array-technology as well as Multiplex Ligation-dependent Probe Amplification (MLPA) based tests. In cooperation with MRC-Holland (www.mlpa.com), we constructed a MLPA-assay based on the

BRCA1-associated CNAs identified in this study. We performed MLPA analysis on tumour material derived from the same samples as used in the SNP array analysis. The results show that TILs have negative effects on measuring copy number changes by means of MLPA-analysis as seen by SNP array analysis.

Finally, we offer a direct explanation for CNA devoid BLCs which have been reported by others. Also, we demonstrate methods to overcome the effects of TILs by using DNA flow cytometry, which can prove invaluable in the clinic, future research studies on (*BRCA1*-mutated) BLCs, but also for other tumour types as well as next generation sequencing efforts on formalin-fixed, paraffin-embedded tumour material. For future efforts, similar analysis with larger sample sizes could fine-tune our results and perhaps give insight into the differences of *BRCA1*-mutated BLCs with differential mechanisms for losing the wild-type *BRCA1*-allele, i.e. copy neutral LOH or large deletions.

Chapter 3 describes integrated genomic and gene expression analysis of luminal and HER2+ breast tumours to identify specific CNAs or gene expression abnormalities in *CHEK2**1100delC mutated breast tumours. Literature on genomic profiling of *CHEK2* associated breast tumours is scarce with a single report published so far. Unlike the previous report, in our study, samples were again selected for low amounts of tumour infiltrating lymphocytes by mRNA profiling. High mRNA immune signature values were found in only a small number of luminal and approximately half of the HER2+ tumours. In contrast to basal-like *BRCA1* mutated breast cancers, no apparent specific somatic copy number aberration (CNA) profile for *CHEK2**1100delC breast cancers was found. With the possible exception of copy number loss of chromosomal arm 1p present in a subset of tumours, which might be involved in *CHEK2* tumourigenesis. In addition, genome wide gene expression analysis identified only a small number of genes to be differentially expressed between *CHEK2**1100delC and BRCAX breast cancers, of which none overlapped with a previously reported *CHEK2* gene expression signature [11], possibly due to the small number of samples used in both studies.

The existence of specific CNAs in *BRCA1*-mutated tumours and the apparent lack of such CNAs for *CHEK2*-mutated tumours might be explained by the need for *BRCA1*-deficient tumour cells to acquire survival factors, by for example specific copy number aberrations, to expand. Such factors may not be needed for breast tumours with a defect in a non-essential gene such as *CHEK2* therefore resulting in copy number profiles that largely reflect the tumour subtype. However, in this study and the previously reported study

[11], loss of the wild-type *CHEK2* allele by means of genomic losses or copy neutral LOH, as observed in *BRCA1*-mutated tumours, is rarely detected in *CHEK2* tumours. Therefore, an alternative explanation for a lack of specific CNAs in the *CHEK2* tumours might be due to heterogeneity among these tumours where part of the tumours have retained a wild type copy of the *CHEK2* gene. Genomic profiling of tumours from homozygous *CHEK2**1100delC mutation carriers could therefore yield more conclusive answers to whether CNAs associated with *CHEK2* tumourigenesis do exist. Currently such an analysis is perfromed in collaboration with the Dutch cancer institute.

The results presented in chapter 2 and to a lesser extend in chapter 3 are guided by the observation of large numbers of TILs within the tumour samples. In breast cancer there have been conflicting reports regarding the prognostic value of the local inflammatory response. This can in part be explained by the use of small heterogeneous patient groups and varying methodologies for assessing tumour cell infiltrates such as core biopsies, tissue microarrays and gene expression arrays. We have defined a gene expression immune infiltrate signature representing a general tumour inflammatory cell infiltrate which can easily be applied to large genome wide expression datasets to identify relatively poor from highly immune infiltrated breast cancer samples. We used the mRNA immune infiltrate signature as a standardized assessment of general tumour inflammatory cell infiltrate to investigate its association with patient survival and molecular subtypes in large cohorts of lymph node negative primary breast cancer patients who did not receive any adjuvant therapy. We used hierarchical clustering to stratify samples into high and low-TIL groups. While this approach is valid in large tumour cohorts containing both high and low-TIL samples, such as the breast cancer cohorts used in our study, for smaller sample sizes or even single sample assessments, this method needs refinement.

In line with the results from chapters 2 and 3, samples with a relative high amount of TILs were found more than expected in the ER-/HER2- and the HER2+ subtype. From the analyses we can conclude that in lymph node negative primary breast cancer, high levels of tumour inflammatory cell infiltrates as determined by gene expression analysis, are associated with better metastasis-free survival, especially in HER2+ breast cancer patients. The contradicting results regarding the prognostic value of a general tumour inflammatory cell infiltrate in breast cancer reported in literatue is most likely due to confounding heterogeneity with regard to breast cancer subtype and treatment amongst the study samples. Where we find the presence of TILs to be associated with improved survival in all breast cancer subtypes,

previous reports find no association or even poorer association of TILs and survival.

In chapter 5, we show the results of whole-exome sequencing analysis of germline DNA for 5 BRCAX BLCs. These samples were found to share overlapping regions of copy neutral LOH on chromosome 17. For the low-TIL *BRCA1*-mutated BLCs, copy neutral LOH was seen at the *BRCA1*-locus. Therefore, the overlapping regions of copy neutral LOH in the BRCAX samples could hint towards the genomic location of one or more breast cancer susceptibility genes involved in breast cancer predisposition in these patients. Analogues, in retinoblastoma, regions of copy number neutral LOH are frequently and specifically reported on chromosome 13, where the *RB1* gene resides [12].

For one sample, within the region of LOH in matching tumour material, a novel nonsense mutation in the established cancer-predisposing *RAD51C* gene was identified. The histopathological characteristics of this sample are in line with previously reported cases of breast tumours with large deletions in the *RAD51C* gene [13]. Therefore, *RAD51C*-mutated breast tumours seem to be aggressive, genomically instable Basal-like tumours resembling *BRCA1*-mutated tumours to a certain extent. However, very few cases involving *RAD51C* pathogenic germline mutations in breast tumours have been described.

We did not find conclusive evidence for the involvement of novel breast cancer susceptibility genes in the remaining four patients. This result could partially be explained by the fact that causal mutations might have been missed by the exome sequencing procedure. The data also show that the selected group of patients, based on the tumour subtype and CNA profile, is still heterogeneous at the level of genetic predisposition. Nevertheless, the identification of a novel nonsense mutation in the *RAD51C* gene also suggests the approach to be of value and could perhaps be optimized for future endeavors. The combination of tumour profiling (gene expression and genomic profiling) and whole exome/genome sequencing in large cohorts could prove useful in unraveling the underlying genetic basis of a proportion of breast cancer cases.

References

1. 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, 95:1482-1485.

2. 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 tumours.** *Cancer Res* 2002, 62:7110-7117.

3. van Beers EH, van Welsem T, Wessels LF, Li Y, Oldenburg RA, Devilee P, Cornelisse CJ, Verhoef S, Hogervorst FB, van't Veer LJ et al.: **Comparative genomic hybridization profiles in human BRCA1 and BRCA2 breast tumours highlight differential sets of genomic aberrations.** *Cancer Res* 2005, 65:822-827.

4. Jonsson G, Naylor TL, Vallon-Christersson J, Staaf J, Huang J, Ward MR, Greshock JD, Luts L, Olsson H, Rahman N et al.: **Distinct genomic profiles in hereditary breast tumours identified by array-based comparative genomic hybridization.** *Cancer Res* 2005, 65:7612-7621.

5. Stefansson OA, Jonasson JG, Johannsson OT, Olafsdottir K, Steinarsdottir M, Valgeirsdottir S, Eyfjord JE: **Genomic profiling of breast tumours in relation to BRCA abnormalities and phenotypes.** *Breast Cancer Res* 2009, 11:R47.

6. Chin SF, Teschendorff AE, Marioni JC, Wang Y, Barbosa-Morais NL, Thorne NP, Costa JL, Pinder SE, van de Wiel MA, Green AR et al.: **High-resolution aCGH and expression profiling identifies a novel genomic subtype of ER negative breast cancer.** *Genome Biol* 2007, 8:R215.

7. Popova T, Manie E, Rieunier G, Caux-Moncoutier V, Tirapo C, Dubois T, Delattre O, Sigal-Zafrani B, Bollet M, Longy M et al.: **Ploidy and large-scale genomic instability consistently identify basal-like breast carcinomas with BRCA1/2 inactivation.** *Cancer Res* 2012, 72:5454-5462.

8. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, Speed D, Lynch AG, Samarajiwa S, Yuan Y et al.: **The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups.** *Nature* 2012, 486:346-352.

9. Schouten PC, van Dyk E, Braaf LM, Mulder L, Lips EH, de Ronde JJ, Holtman L, Wesseling J, Hauptmann M, Wessels LF et al.: **Platform comparisons for identification of breast cancers with a BRCA-like copy number profile.** *Breast Cancer Res Treat* 2013, 139:317-327.

10. Lips EH, Laddach N, Savola SP, Vollebergh MA, Oonk AM, Imholz AL, Wessels LF, Wesseling J, Nederlof PM, Rodenhuis S: **Quantitative copy number analysis by Multiplex Ligation-dependent Probe Amplification (MLPA) of BRCA1-associated breast cancer regions identifies BRCAness.** *Breast Cancer Res* 2011, 13:R107.

11. Muranen TA, Greco D, Fagerholm R, Kilpivaara O, Kampjarvi K, Aittomaki K, Blomqvist C, Heikkila P, Borg A, Nevanlinna H: **Breast tumours from CHEK2 1100delC-mutation carriers: genomic landscape and clinical implications.** *Breast Cancer Res* 2011, 13:R90.

12. Mol BM, Massink MP, van der Hout AH, Dommering CJ, Zaman JM, Bosscha MI, Kors WA, Meijers-Heijboer HE, Kaspers GJ, Riele H et al.: **High resolution SNP array profiling identifies variability in retinoblastoma genome stability.** *Genes Chromosomes Cancer* 2014, 53:1-14.

13. Schnurbein G, Hauke J, Wappenschmidt B, Weber-Lassalle N, Engert S, Hellebrand H, Garbes L, Becker A, Neidhardt G, Rhiem K et al.: **RAD51C deletion screening identifies a recurrent gross deletion in breast cancer and ovarian cancer families.** *Breast Cancer Res* 2013, 15:R120.