Myrto Barrdahl
Dr. sc. hum.

Gene-environment interactions in relation to breast cancer risk and the role of gene variants in breast cancer survival.

Fach/Einrichtung: DKFZ
Doktorvater: Prof. Dr. Rudolf Kaaks

Breast cancer is the most common cancer and the most common cause of cancer related death among women worldwide. There are several established epidemiological risk factors, as well as a growing number of common, low-penetrance, single nucleotide polymorphism (SNP) alleles related to the disease. However, it is not clear if, and to what extent, the effect of these genetic variants on breast cancer risk varies according to established epidemiological risk factors. Identification of gene-environment interactions could improve risk models and provide insight into the biological mechanisms underlying breast carcinogenesis.

The primary aim of the doctoral work reported in this thesis was to investigate the multiplicative statistical interaction between a set of confirmed genetic breast cancer risk variants and several epidemiological risk factors in relation to breast cancer risk. Identified gene-environment interactions can be useful for the improvement of risk prediction models as well as for understanding the biology of breast cancer. A further aim was to investigate whether these confirmed genetic variants also influence the survival of breast cancer patients. Identification of alleles that affect both risk and prognosis could aid the development of targeted treatments and further the understanding of the biological mechanisms behind disease progression. Further aims consisted of exploring the association between the genetic variants and the risk of invasive breast cancer with certain properties determined by tumor receptor status or classification of the tumor, as well as the risk of breast cancer in situ (BCIS). It is not known whether BCIS is an obligate precursor of invasive breast cancer, thus identification of BCIS specific alleles could shed light on the genetic architecture of non-invasive as well as of invasive breast cancer and further early detection and prevention strategies.

The analyses for this doctoral work were conducted using data from the National Cancer Institute's Breast and Prostate Cancer Cohort Consortium (BPC3). The BPC3 consists of

eight case-control studies nested in prospective cohorts from Europe and the United States, with large scale genetic data as well as extensive questionnaire information at its disposal.

First, main effects of 39 breast cancer risk alleles as well as interactions between these and 12 established epidemiological risk factors collected at baseline, were investigated in relation to breast cancer risk, using 16,285 cases and 19,376 controls. Previously identified associations with breast cancer risk were replicated for the vast majority of the investigated SNP alleles.

There were no significant gene-environment interactions found in the analyses of the BPC3 data. However, in a meta-analysis of results from the BPC3 and the Breast Cancer Association Consortium (BCAC) on up to 79,000 subjects, a suggestive interaction between *SLC4A7*-rs4973768 and smoking was identified, which has biological plausibility. The minor allele of rs4973768 may have a functional impact on the *SLC4A7* gene which regulates the influx of lead into erythrocytes. Smokers have higher levels of lead in the blood than non-smokers through their exposure to cigarette tar, and lead exposure has been implied in breast cancer. Hence, it is possible that a polymorphic variant that influences SLC4A7 functionality might exert its effect in subjects with an increased exposure to lead, such as smokers.

Third, survival analyses of breast cancer patients were conducted on 35 confirmed breast cancer risk alleles, using 10,255 cases of which 1,379 were fatal, including 754 disease specific deaths. The most novel result was the association between the C allele of *LSP1*-rs3817198 and improved overall survival, which could be explained by the potential association of the allele with an up-regulating effect on the tumor suppressor gene *CDKN1C*.

Finally, association analyses of 39 SNP alleles and risk of BCIS and invasive breast cancer in subgroups determined by tumor prognostic factors were carried out, using 1,317 BCIS cases, 16,285 invasive breast cancer cases, and 19,376 controls. Suggestive evidence on a preferential association of *CDKN2BAS*-rs1011970 with BCIS, rather than with invasive disease was found, which could be explained by an association of the allele with the *CDKN2B* gene, which is involved in early breast carcinogenesis. Previously identified associations with either estrogen receptor positive or estrogen receptor negative breast cancer risk were also replicated for all SNPs.

In summary, the results from the doctoral work presented in this thesis suggest that moderate gene-environment interactions may be present. However, functional studies are needed to clarify whether the statistical interactions have biological plausibility. Furthermore, the

present findings are not in line with the notion of any significant role of the investigated breast cancer risk alleles in the survival of breast cancer patients, or in non-invasive breast carcinogenesis. Still, the results presented in this thesis contribute to the global body of knowledge on gene-environment interactions by generating hypotheses on where these could harbor, thereby providing guidance for future functional studies and large scale replication studies.