Breast Cancer Risk after Diagnostic Gene Sequencing (BRIDGES)

EU – H2020 project grant agreement no 634935

Website: bridges-research.eu

Deliverable Number	3.2
Deliverable Title	Combined analysis of exome and gene panel sequencing
Short Title	ExomeResults1
Lead beneficiary	UCAM
Del. Date (Annex 1)	01/09/2018
Achieved Date	31/08/2018
Nature	Report
Dissemination Level	PU
Document Filename	D32- ExomeResults1-PU-v1.0

Date	Authors/Reviewers	Remarks	Version
01/08/2018	Thomassen	Provided format	0.1
30/08/2018	Douglas Easton		0.2
31/08/2018	Thomassen	Send delivery to EC	1.0

Contents

Overview

The goal of D3.2 was to use existing sequence data in breast cancer cases and controls to identify candidate genes for replication in panel 2. Specifically, we used data generated by the PERSPECTIVE project (Partner 15), COMPLEXO (Partners 11, 14) and the TCGA project.

The PERSPECTIVE project performed whole-exome analysis of 1,500 familial breast cancer cases and 3,700 controls. Association tests were carried out based on burden analysis of protein truncating variants and rare, likely deleterious missense variants. As part of these analyses, we also included data from the TCGA (The Cancer Genome Atlas) which has exome sequence data on ~1,200 breast cancer cases.

Based on these association tests, 216 genes showing the most promising evidence of association were selected for replication in 6,500 cases and 6,100 controls from Germany, the Netherlands, UK and Canada. Targeted sequencing was performed using the Agilent SureSelect system. In burden analyses, none of the "novel" genes replicated, after adjustment for the number of genes tested.

The COMPLEXO consortium compiled exome sequence data on ~1,000 breast cancer cases from multiple case families. Based on analyses of these data, 18 putative susceptibility genes not on BRIDGES panel 1 were identified. Targeted sequencing of these genes was performed in ~14,000 cases and ~7,000 controls from the SEARCH study in the UK. This sequencing was performed using the Fluidigm Juno system, as for the sequencing in BRIDGES panel 1. In burden analyses, none of the genes showed evidence for association, either for protein truncating variants, rare missense variants (minor allele frequency <0.001) or likely deleterious rare missense variants (defined by CADD score >20).

Discussion

Analysis of the existing exome sequence data, with additional targeted replication, did not yield compelling evidence for additional breast cancer susceptibility genes, beyond those included on BRIDGES Panel 1. This is likely to be for two main reasons. The available exome sequencing studies were small (~1,500 cases for PERSPECTIVE, ~1,000 for COMPLEXO and TCGA). The power to detect plausible associations in these datasets is low, given that the frequency of susceptibility variants is likely to be low (for example, the combined frequency of truncating variants in the known susceptibility genes *BRCA1*, *BRCA2*, *ATM*, *CHEK2* and *PALB2* is ~0.1% - 0.3%). The frequencies and effect sizes of novel genes are likely to be similar to or lower than the known genes.

In addition, the existing exome sequence data were generated on breast cancer cases and compared with sequence data from other control studies (e.g. ExAC and GnomAD). These external control series, while large, were sequenced using different technologies and used different bioinformatics pipelines to perform quality control and call variants. While every effort was made to ensure that the calling of the cases and controls was as similar as possible, there will inevitably be some differences leading to spurious associations. Additionally, the ExAC and GnomAD datasets are drawn from a large number of studies from different populations and hence the variant frequencies may not represent accurately the frequencies in the breast cancer cases. The lack of comparability between cases and controls could have further reduced

the power of the analyses, and hence the likelihood that susceptibility genes were taken through to replication.

Conclusion

Analysis of existing exome sequence data has not identified susceptibility genes (beyond those already being evaluated on BRIDGES panel 1) with sufficiently compelling evidence to select for replication in panel 2. A much larger and better designed exome sequencing will be required to identify additional genes. As a result, the next phase of BRIDGES 2 will be redesigned, with panel 2 being replaced by exome sequencing in breast cancer cases and controls from the BCAC.