Abstract

Breast cancer, a genetically heterogeneous disease, continues to be the leading cause of death among women. Although the mortality rate is gradually declining, the incidence is once again on the rise. Tertiary and quaternary prevention focus on studying genetic differences among patients and the effects of chemotherapy. The primary hypothesis for this research is based on the evolutionary, population, and functional differences in genetic variants within the human genome. This hypothesis is already employed in numerous bioinformatics tools. The main objective of this dissertation work was to study available databases and bioinformatics tools to facilitate the processing of extensive pharmacogenomic data. Through massive parallel sequencing, we evaluated frequent germline coding and non-coding variants in a panel of 509 genes related to drug metabolism and elimination, cell death functions, or signaling pathways in oncology among Czech breast cancer patients using novel in silico tools and established pharmacogenomic databases. Prioritized variants were technically and clinically validated in a cohort of 805 patients. The association of rs2227291 in *ATP7A*, rs2293194 in *KCNAB1* (in early-stage patients), and rs4376673 in *DFFB* with response to neoadjuvant cytotoxic therapy provides new targets for subsequent functional studies, potentially contributing to tertiary prevention. In patients with luminal B or triple-negative tumors, the frequently studied variant rs1801160 in *DPYD* was significantly associated with disease-free survival, representing an additional predictive target with prognostic potential.

For rare coding variants in these patients, we employed a newly proposed strategy to address some limitations of machine learning methods. This approach revealed a significant association of variants in the *ABCC1, ABCC4, ABCB6, ATP7B, CYP2D6, CYP4F3*, and *CFTR* genes with treatment response, menopausal status, histological type, molecular subtype, and differentiation grade of breast cancer. Patients with pathogenic variants in *CFTR* had significantly reduced disease-free survival (Log Rank, $p = 0.002$) and overall survival (Log Rank, $p = 0.006$). Our study provides further evidence that *CFTR* may be involved in processes affecting the efficacy of cancer therapy or the progression of malignant disease, as supported by recent publications. Tertiary prevention in the context of treatment response could theoretically be achieved through the monitoring of *CFTR* functionality or the administration of therapeutic agents, such as currently available CFTR modulators, which partially address defects in functional variants.

By targeting all variants in the *KIF14* gene, functional non-coding variants (rs17448931 and rs3806362) identified through in silico analysis have been clinically validated, representing refined targets within the *KIF14* gene with potential applications in tertiary prevention.