ABSTRACT

Somatic mutations and copy number alterations in basal cell carcinoma- focus on gene-associated noncoding variants and miRNome

Basal cell carcinoma (BCC) of the skin is the most common human cancer, especially frequent in the Caucasian population. It is known that the development of BCC is driven by mutations in genes involved in the Sonic Hedgehog (Shh) signaling pathway, however, the pathogenesis of BCC is not fully understood. As in other cancer types, genetic analyses in BCC have been focused almost exclusively on the coding regions, constituting only ~2% of the genome.

The main aim of this study was the preliminary characteristics of somatic mutations in noncoding elements of protein-coding genes (5'UTR, fragments of 3'UTR, and introns) and in miRNA genes. DNA samples isolated from 27 pairs of BCC and corresponding healthy tissues were sequenced with the use of two next generation sequencing procedures, the routine Whole Exome Sequencing (WES) and the unique procedure of Whole miRNome Sequencing (WMS), developed at the Department of Molecular Genetics, Institute of Bioorganic Chemistry, Polish Academy of Sciences.

In total, over 80,000 somatic mutations were detected of which more than half in noncoding regions. The average tumor mutational burden (TMB) was >50 mutations/Mb, which indicates that BCC is the most frequently mutated human cancer. The most common mutational signature found in all samples was signature 7, confirming that the most important mutagenic factor in BCC is UV radiation.

The analysis of mutations in noncoding regions showed that some of the mutations were clustered in specific regions, including hotspots. Moreover, some mutations in noncoding regions strongly suggest the cancer-driving potential of the mutated genes, e.g., mutations in 3'UTR in *BAD*, mutations in the Kozak sequence in the *DHODH* 5'UTR, and mutations in 5'UTR in *CHCHD2*. All these genes are functionally involved in processes related to carcinogenesis (apoptosis, mitochondrial metabolism, and de novo pyrimidines synthesis). The analysis of external resources of cancer genomic data such as The Cancer Genome Atlas

(TCGA) demonstrated that mutations in *BAD* and *CHCHD2* occur also in melanoma, while mutations in *DHODH* are specific for BCC.

Among the noncoding mutations, 171 were detected in miRNA genes, including mutations in various functional elements of these genes, crucial for miRNA functioning and biogenesis. The most frequently mutated miRNA gene was *MIR3928* whose functional role in cancer is well documented. All the mutations in *MIR3928* were localized in the 5' flanking region of the miRNA precursor of which 3 were in a single hotspot position.

Taking advantage of the generated sequencing data allowed for the first comprehensive analysis of copy number variation in BCC. Among the identified most frequent copy number alterations are deletions of the long arm of chromosome 9 (chr9q), encompassing *PTCH1*, a key BCC tumor suppressor, and duplications of the short arm of chromosome 9 (chr9p), encompassing the oncogene *JAK2* and the genes encoding the key immune checkpoint ligands PD-L1 and PD-L2.

Among the genes identified as most frequently mutated in coding sequences, are genes such as *PTCH1*, *TP53*, *MYCN*, and *NOTCH* well known as key drivers in BCC. This high consistency of the results of the analysis of coding sequences with previous BCC genetic analyzes confirms the reliability of the performed analyzes and the credibility of the obtained results.

In summary, the study is the first analysis of noncoding mutations in BCC. The analysis enabled the identification of noncoding mutations with the potential of cancer drivers in BCC. The results provide a strong basis for further analyses of noncoding variants in BCC and cancer in general.