Bettina Fiege

Dr. sc. hum.

Genetic variants in susceptibility to childhood acute lymphoblastic leukemia

Fach/Einrichtung: DKFZ (Deutsches Krebsforschungszentrum)

Doktorvater: Professor Dr. Rajiv Kumar

The goal of the present study was the identification of novel germline genetic variants in susceptibility to childhood ALL and the investigation of a functional contribution of associated variants to the disease. This was achieved by a validation study which was carried out in continuation of a previously conducted genome-wide association study in childhood B-cell precursor ALL cases. The validation study was followed by fine-mapping of disease associated genes and determination of a risk model for childhood B-cell precursor ALL. Finally, the functional aspects of the new and previously through GWAS identified variants were investigated by using genome-wide data generated by the ENCODE project. In conclusion, the findings of the present study significantly improve our understanding of the germline genetics of childhood ALL.

The validation study comprised genotyping of 21 polymorphisms in 1,501 German B-cell precursor ALL cases of the Berlin-Frankfurt-Münster trial and 1,516 German controls. Through this approach the association of two novel variants with risk of childhood ALL was identified. The first variant identified was the polymorphism rs10828317 on chromosome 10p12.2, which localizes to exon 7 of *PIP4K2A* and introduces a benign amino acid change from asparagine to serine in the enzyme phosphatidylinositol-5-phosphate 4-kinase type II α (PI5P4K2α). This enzyme is involved in the regulation of phosphatidylinositol-5-phosphate levels. The second variant identified was the polymorphism rs3824662 on chromosome 10p14, which localizes to intron 3 of *GATA3*. GATA3 is a member of the GATA family of transcription factors. Fine-mapping of these two genes underscored their role in genetic susceptibility to childhood ALL as additional variants in the genes *PIP4K2A* and *GATA3* were associated with risk of the disease. In a multivariate statistical analysis only one variant in *PIP4K2A* gene and two variants in *GATA3* gene remained statistically significant associated with disease risk. The risk model of childhood B-cell precursor ALL was based on these three polymorphisms and showed increased disease risk

for carriers of at least four out of six risk alleles of the three disease associated variants in *PIP4K2A* and *GATA3* gene.

The subsequent functional annotation of the new and previously identified disease associated variants further supported the assumed contribution of sequence variants to gene regulation and ultimately to the development of childhood ALL. The data of the ENCODE project emerged as the essential tool to gain insights into the role of sequence variants in the non-coding part of the human genome.