Defining risk factors for genomic instability in B cells: Novel insights from NGS-based technologies
Hogenbirk, M.A.

Citation for published version (APA):
SUMMARY

Cancer is caused by an accumulation of somatic DNA mutations. Knowledge on the cause and consequence of these mutations is essential for an effective prevention, diagnosis, prognosis and treatment of cancer. Our research focused on B cell non-Hodgkin lymphomas. These lymphomas are derived from germinal centers. Germinal centers are defined histological structures that develop during an immune response in secondary lymphatic tissues such as of lymph nodes, spleens or Peyer’s Patches. Within the germinal center, B cells activated by cognate antigen undergo a critical development phase during which a DNA mutator system is turned on in order to mutate antibody-encoding genes with an extremely high frequency. While this intentional mutagenesis process is of great importance for an effective immune response, it brings great risks. Disruption or aberrant targeting of this mutagenesis system may lead to genetic instability and the development of B-cell lymphomas. By investigating DNA binding profiles of activation-induced cytidine deaminase (AID), a protein central in this mutagenesis process, we aimed to answer the following questions: How specific is the mutagenesis process targeted throughout the genome? Which (onco)genes are unintentionally mutated by AID? How does AID contribute to genetic instability and especially chromosomal translocations? Chromosomal translocations are genomic mutations that occur when a DNA segment is moved from one chromosome to another (or within a chromosome). These chromosomal translocations lead to deregulation of specific genes characteristic for well-defined B cell non-Hodgkin lymphomas.

The work described in this thesis is partly focused on clarifying the targeting specificity of AID. It is well established that AID initiates mutations in immunoglobulin genes to improve immunity. Given the mutagenic potential of AID, a central question in the AID-field relates to the specificity of AID binding and activity throughout the genome. ChiP-seq is a technique which can determine the genomic binding sites of a protein of interest (here AID). As described in chapter 2, ChiP-Seq datasets were obtained and analyzed by a different laboratory. The authors concluded that AID binds to thousands of genes. As we had gathered collected results that contradicted this finding, we re-analyzed the ChiP-Seq data and reveal that their conclusions were premature. Therefore, at present there is no direct evidence supporting that AID binds and/or is active across the entire genome.

Experiments in several laboratories suggested that AID has a genome-wide function in programming cells by active DNA demethylation. Given the many transcriptional alterations at the transition from naïve to GC B cells we tested whether AID has also a role in the programming of B cells. In chapter 3 we report that AID has no role in B cell programming. Surprisingly, however, a large number of genes were found to be artificially deregulated in B cells of Aicda<sup>−/−</sup> mice, i.e. mice in which the gene encoding AID is rendered non-functional. These deregulations relate to an artificial inactivation of AID, and not AID itself. Because these mice were used in most studies on AID, these findings may have far-reaching consequences on previous conclusions made in the AID-field, including its impact on the origin of B cell non-Hodgkin lymphomas.

The concept that AID can bind throughout the genome has been to a large extent based on experiments that required ‘next-generation sequencing’ (NGS) technologies, which provide complex data sets. The analysis of these data is not trivial when compared with conventional
molecular biology techniques. In chapter 4, we therefore have a developed checklist that enables to value NGS-based results in a relative simple manner. We have applied this checklist on studies that claim that AID binds throughout the genome. We conclude that apart from the phenomena of collateral damage, none of these studies provided solid evidence for extensive genome targeting by AID. In line with this notion the number of independently confirmed AID targets is rather low.

Given that AID activity is largely restricted to immunoglobulin genes our investigation focused on the question of how translocations occur in B cells. Other laboratories reported that, in addition to DNA accessibility, transcriptional activity of a gene associates with the translocation risk. In chapter 5, we address the question whether the transcription or accessibility determines the translocation risk in active genes. Our study is the first to demonstrate that the DNA accessibility determines the translocation risk in active genes.

Studies on the origin of chromosomal translocations have long been limited by the fact that these genetic alteration occur relatively infrequent and are therefore difficult to study. As a result, most of what we know about chromosomal translocations has been obtained from experiments in which DNA breaks, the initiators of chromosomal translocations, are forced. Although informative, these studies may not reflect the physiological mechanism by which translocations actually occur. Based on analyses as presented in this thesis, and a critical literature study, we propose a new translocation model in chapter 6. This model proposes that a single-end DNA from a DSB suffices to initiate a chromosomal translocation. In B cells, this DNA end would be obtained by an AID induced DSB within an immunoglobulin locus. Our model differs from the currently predominant models proposing that at least two (AID generated) DNA breaks must pre-exist before a translocation can take place. Because this model implies that any spontaneous obtained DNA DSB can initiate a translocation, this model also applies to chromosomal translocation in other tumors.