RIF1, ZBTB24 and repeat silencing

Wu, H.

Appendix

Summary

Nederlandse samenvatting

Curriculum vitae

List of publications

Acknowledgements

Appendix

Summary

More than 50% of the human genome is composed of repetitive DNA. Repeat elements, which were for long time regarded as "Junk DNA", are getting increasing attention from scientists from different fields. Their important roles in multiple biological processes such as maintenance of genome stability, regulation of gene expression, and tumorigenesis have been reported. In mammals, repetitive regions are generally transcriptionally silent in somatic tissue and associated with repressive histone marks such as H3K9me2, H3K9me3, and DNA methylation. Improper re-activation of repetitive regions can cause genomic instability and aberrant transcription, and this has been related to disease. To date, around 20 diseases have been associated with repetitive regions, and this number is still increasing. In this thesis, we focus on two proteins RIF1 and ZBTB24, both of which have been linked to regulation of repetitive sequences in eukaryotic genomes.

RIF1 is a multifunctional protein that is involved in many biological processes including transcriptional regulation. In **chapter 2**, we analyzed the function of RIF1 in regulating repetitive DNA in both mouse and human model systems. We found that RIF1 deficiency leads to de-repression of X-linked ampliconic gene clusters. Rif1 was reported to stabilize the H3K9 methyltransferase complexes, and we observed decreased H3K9me2 enrichment at the affected X-linked ampliconic regions in the absence of RIF1, suggesting the involvement of H3K9me pathway. Consistent with this, knockdown of EHMT1 results in re-activation of X-linked gene clusters. Furthermore, both in mouse and human models, RIF1 disruption lead to de-repression of a set of germline genes, and this regulation is partially DNA methylation dependent. In sum, our results indicate that RIF1 participates in transcriptional regulation of X-linked gene clusters and a set of germline genes, and that the underlying mechanism is conserved between mouse and human.

Immunodeficiency, Centromeric instability, and Facial anomalies (ICF) syndrome is a rare, genetically heterogeneous, autosomal recessive disorder. DNA hypomethylation at pericentromeric regions presented in all ICF subgroups is the molecular hallmark of this syndrome. Four genes have been found so far that, when mutated, can cause the disease, namely *DNMT3B*, *ZBTB24*, *CDCA7*, and *HELLS*. How the 4 ICF genes functionally converge in the context of the ICF syndrome remains unclear. In addition, the molecular functions of ZBTB24 and CDCA7 are not fully understood. **Chapters 3**, **4** and **5** focus on the ICF2 gene, *ZBTB24*. In **chapter 3**, we generated *Zbtb24*$^{\Delta BTB/\Delta BTB}$ mutant mESC lines, and this mutation strategy resulted in a loss of Zbtb24 protein. Using RNA-sequencing we showed that Zbtb24 positively regulates a set of genes, and *Cdca7*, which is the ICF3 gene, is among them. Furthermore, we discovered that ZBTB24 functions as a transcription factor, and activates CDCA7 expression via directly binding to and activating its promoter. We validated our finding in human cell models and found that this regulation is highly conserved between mouse and human.

The utilization of the sequencing technologies has provided clinical researchers with a vast number of genetic variants from both healthy individuals and patients. Due to the limited knowledge about the pathogenicity of the detected variants, the biggest challenge now is to assess potential phenotypic consequences of the variants of unknown significance (VUSs). In **chapter 4**, we established a robust method to functionally distinguish pathogenic from neutral *ZBTB24* variants. We used ChIP-qPCR and a luciferase reporter assay to assess how *ZBTB24* variants influence binding and activating ability of ZBTB24 protein at the *CDCA7* promoter. We measured *ZBTB24* variants identified in ICF2 patients, patients of two other diseases, likely "healthy" *ZBTB24* SNPs and some rare variants detected in cancer patients. Our data shows that we can separate ICF2 missense variants from *ZBTB24* SNPs, and suggests that the two missense variants from other diseases are likely benign. Furthermore, our results imply that the rare cancer variants tested, which are located in key residues in the C2H2-ZF domain are likely pathogenic. We believe that our method can aid in the clinical diagnosis of *ZBTB24*-associated ICF syndrome.

In **chapter 5**, we investigated how loss of Zbtb24 affects DNA methylation genome-wide and compared methylome changes between Zbtb24 and Cdca7 knockout mESCs. We found that both Zbtb24 and Cdca7 depletion caused a global loss of DNA methylation, which was more pronounced in Cdca7 knockout mESCs. Remarkably, Zbtb24 depleted mESCs also showed increased 5mC levels at around half of the detected differentially methylated regions (DMRs), and they included promoters of actively transcribed genes. ChIP-seq performed in both mESCs and U2OS showed that ZBTB24 preferentially binds to open chromatin regions in a sequence-dependent manner. We further reported an integrative genome-wide analyses of DNA methylation, histone modification, Zbtb24 binding and gene expression profiles in mESCs. We found that Zbtb24 occupancy protects nearby CpGs from being methylated. Loss of Zbtb24 binding leads to a gain of DNA methylation at a set of Zbtb24 targets, accompanied by reduced H3K4me3 levels and decreased expression of the downstream genes. This suggests a direct link between Zbtb24 occupancy, H3K4me3 and aberrant DNA methylation gain at some Zbtb24 binding sites. We also identified a ZBTB24-mediated promoter hypermethylation footprint in both mESCs and ICF2-patient derived material. We believe this ZBTB24 dependent hypermethylation signature provides a potential DNA methylation biomarker for ICF2. Altogether, our research shed light on a role of Zbtb24, a transcription factor, in regulating DNA methylation genome-wide.