Author: Klerk, Eleonora de
Title: Mechanisms controlling mRNA processing and translation: decoding the regulatory layers defining gene expression through RNA sequencing
Issue Date: 2015-09-30

SUMMARY

SAMENVATTING (SUMMARY IN DUTCH)

CURRICULUM VITAE

LIST OF PUBLICATIONS

SUMMARY

In 1941, geneticists G.W. Beadle and E.L. Tatum formulated the ground-breaking hypothesis “one gene, one enzyme”, which led to their Nobel Prize in Physiology or Medicine in 1958. Today’s researchers have reformulated Beadle and Tatum’s hypothesis, as the link between DNA and proteins has been proven to be much more complex: even though the human genome contains less than 20,000 genes, these genes encode for more than 80,000 protein-coding messenger RNAs (mRNAs, intermediate molecules between DNA and proteins), which have been estimated to give rise to hundreds of thousands of proteins. One major remaining challenge in cellular biology is to understand the mechanisms regulating the diversity of mRNAs and proteins expressed from a single gene.

The work described in this thesis focuses on the mechanisms that give rise to alternative mRNAs and their alternative translation into proteins. Each of the described studies has been based on a specific set of high-throughput RNA sequencing technologies. Together these provide a comprehensive view of these alternative regulatory mechanisms. An overview of the available RNA sequencing methods, together with an introduction to different regulatory layers which define the expression of a gene, are presented in Chapter 1. This Chapter describes the processes of alternative transcription, alternative mRNA processing and alternative translation, focusing on what we have learnt from RNA sequencing studies.

Our work in Chapter 2 and Chapter 3 investigates the process of alternative polyadenylation, which is one of the steps during mRNA processing, and results in the inclusion or exclusion of sequences that affect the stability of an mRNA or the nature of the protein isoform formed.

Chapter 2 shows the role of alternative polyadenylation in the context of oculopharyngeal muscular dystrophy (OPMD), an autosomal dominant and progressive muscle disorder caused by mutation in the PABPN1 gene. In this study, we identified and quantified the usage of alternative polyadenylation sites in affected skeletal muscles using a novel high-throughput single-molecule poly(A)-site sequencing method. We demonstrated transcriptome-wide shortening of mRNAs in OPMD and propose a novel role for the PABPN1 protein in poly(A) site selection.

Chapter 3 describes genetic variants associated with alternative polyadenylation. In this study we used RNAseq and DeepSAGE to identify genetic variants affecting the usage of alternative polyadenylation sites, by disrupting or forming polyadenylation signal sequences. We confirmed the known genotype-dependent alternative polyadenylation in the gene IRF5 (explaining its genetic association with systemic lupus erythematosus), and we reported novel causative variants affecting alternative polyadenylation by changes in the polyadenylation signal, seven of which had been reported as associated with diseases.

Chapter 4 focuses on mechanisms controlling protein synthesis (translation) during skeletal muscle differentiation, highlighting changes in the use of alternative translation initiation sites. This chapter demonstrates that skeletal muscle differentiation is not only regulated at the level of mRNA transcription and processing, but that also mRNA translation is tightly controlled for specific subsets of functionally correlated genes and contributes to the diversity of proteins required for skeletal muscle function.

In Chapter 5 we investigated the interdependence between alternative regulatory events in gene expression. In this study, based on single-molecule full-length RNA sequencing, we demonstrated coordination and interdependence between alternative transcription initiation, alternative splicing, and alternative polyadenylation in nearly half of the detected genes, and suggested a coordinating role for RNA binding proteins from the muscle blind family (MBNL) in the regulation of splicing and

polyadenylation.

The alternative regulatory mechanisms described in Chapter 1 and investigated in this thesis represent only a portion of all the mechanisms affecting gene and protein expression. Additional regulatory mechanisms are shortly discussed in Chapter 6, to give a more comprehensive picture of the complexity of the process of gene expression. Finally, Chapter 6 connects fundamental research in the RNA field with clinical care, describing new diagnostic and therapeutic approaches that are based on the alternative modes regulating gene expression at transcriptional, post-transcriptional and translational level.