Advancing transcriptome analysis in models of disease and ageing

de Jong, Tristan Vincent

DOI:
10.33612/diss.99203371

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2019

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
CHAPTER 1

OUTLINE OF THIS THESIS

Tristan V. de Jong
RNA-sequencing technologies allow us to peak behind the curtain of biomolecular mechanisms which drive all known life on earth. The process of RNA synthesis is influenced by many factors, both intrinsic and extrinsic, on both micro and macro scales. In most RNA-sequencing experiments a component of a genetic pathway is manipulated after which changes on RNA abundances are investigated. In this thesis we exemplify multiple RNA-sequencing experiments and reveal the dynamic range of information which can be taken from different experimental conditions using a rich toolbox of analytical methods.

In chapter 2 of this thesis we review the quantitative nature of RNA-sequencing data and the underlying factors which are known to influence the variation in gene expression. Due to the incomplete understanding that we currently have of transcriptome dynamics we emphasize the power of RNA-sequencing over other technologies, which include the discovery of new genes, the potential of combining these data with other omics layers such as proteomics, and the utilization of the variation in RNA read counts to unlock a new dimension of RNA-sequencing experiments.

Many transcriptome profiling studies limit themselves to protein-coding genes, as their deregulation can more easily be interpreted in the context of their molecular function. However, beyond the expression of protein coding genes, the presence of non-coding RNA can contain valuable information on mechanisms underlying biomolecular function. In Chapter 3 we explored RNA-sequencing data taken from C. elegans, with and without the expression of a transgene containing an aggregation-prone stretch of 40 glutamine residues (Q40), in combination with mutations to MOAG-2/LIR-3. Investigation of different types of RNA—snoRNAs, snRNAs, ncRNAs, and tRNAs— revealed MOAG-2/LIR-3 functions as a positive regulator of Pol III-mediated transcription of small non-coding RNAs. The result exemplify that “answers” to transcriptome alteration can be found outside of “protein coding domains”.

It is important to have a good strategy to follow up the analysis of transcriptome data. Thus, identifying systemic changes from a list of differentially expressed genes can be
instrumental for deciding on the direction of further research. Chapter 4 explores a conventional differential expression analysis to find genes which are linearly changed in expression among mutants mimicking constitutively acetylated, wildtype, and constitutively not-acetylated C/EBPα. Differential expression and over-representation analysis revealed that the hypoacetylated mutant of C/EBPα induces the transcription of genes located on mitochondrial DNA, thus resulting in an increased mitochondrial respiration. This insight helped to formulate the next experiments to prove that this gene impacts mitochondrial function.

As we get more affordable transcriptome profiling methods, we also extend our studies beyond simple two-group comparisons. This results in complex study designs that include multiple factors, such as intervention type, time after intervention, and cell type. Chapter 5 delves into tackling more complex experimental designs. Here we analyze RNA-seq data taken from proliferating cells and cells taken 2, 4, 10, and 20 days after ionizing radiation from three different cell types: fibroblasts, keratinocytes, and melanocytes. Due to the large variation in transcriptional response dynamics after ionizing radiation among these three cell types, temporal patterns were used for the identification of genes that respond to ionizing radiation.

To investigate potential causes that drive deterioration of transcriptional programs, we explored genome changes associated with advanced age. In chapter 6 we analyzed different hallmarks of ageing using whole-genome sequencing data. Telomere length, mitochondrial DNA content, somatic aneuploidy for sex chromosomes, relative T-cell content, and insertions of retro-transposable elements all are factors which change with time, and simultaneously, our biological age increases with these changes. We created an overview of these changes as well propose a method to explore hallmarks of ageing based on genome sequencing markers.

The composition of nucleotide content around transcription start site can have a huge effect on a gene’s mode of expression. In chapter 7 we analyzed the mechanisms underlying variation in gene expression and show that the intrinsic variation lies with the underlying genetic sequence. The variation is then further modified by epigenetic
changes, dietary influences, and increases with age. Predictive models of the variation in gene expression allowed for the identification of genes which are generally robust, which genes are more variable in their expression, and which genes show the strongest response in variability upon exposure to external factors like changes in diet and age. Knowing determinants of robustness of gene expression can help in studies of disease- and ageing.

To further explore the effect of ageing on robustness of gene expression, we used transcriptome profiling in wildtype mice and compared it to one from a model that shows increase in lifespan. In **chapter 8** we apply the calculation of variability in gene expression for samples of young and aged mice with a genetic modification which delays the development of age-associated phenotypes in mice. The reduced C/EBPβ-LIP expression due to the genetic ablation of the uORF not only delays the development of age-associated phenotypes, but also decreases the overall inter-individual variability in gene expression.

In **chapter 9** we discuss the merits of the findings made in this thesis and re-iterate the value of approaching RNA-sequencing experiments from multiple angles to gain more knowledge into transcriptional (dis)-regulation characteristic for ageing and disease.