

University of Groningen

Core gene identification using gene expression

Claringbould, Annique

DOI:
[10.33612/diss.145227875](https://doi.org/10.33612/diss.145227875)

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:
2020

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
Claringbould, A. (2020). *Core gene identification using gene expression*. University of Groningen.
<https://doi.org/10.33612/diss.145227875>

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.

Core gene identification using gene expression

Annique Claringbould

Annique Claringbould
Core gene identification using gene expression

First printing, 2020
Printed by: Gildeprint

Cover design by: **Sophie Neeleman**

Printing of this thesis was financially supported by: University of Groningen, University
Medical Center Groningen

Copyright © 2020 Annique Claringbould. All rights reserved. No part of this book may be
reproduced or transmitted in any form or by any means without permission of the author.

DOI: <https://doi.org/10.33612/diss.145227875>



university of
 groningen

Core gene identification using gene expression

PhD thesis

to obtain the degree of PhD at the
 University of Groningen
 on the authority of the
 Rector Magnificus Prof C. Wijmenga
 and in accordance with
 the decision by the College of Deans.

This thesis will be defended in public on
 Wednesday 2 December 2020 at 18.00 hours

by

Annique Juliëtte Claringbould

born on 28 January 1992
 in
 's-Hertogenbosch, the Netherlands

Supervisors

Prof. L.H. Franke

Prof. C. Wijmenga

Assessment Committee

Prof. A.G. Uitterlinden

Prof. H.M. Boezen

Prof. P. Visscher

Paranymphs

Niek de Klein

Anne-Grete Märtson

Propositions

1. Genome-wide association studies have successfully uncovered the genetic architecture of numerous complex traits, but additional layers of data are required to uncover the molecular mechanism leading to disease.

2. Bulk gene expression datasets reflect their cell types or tissue of origin, and the resulting patterns need to be accounted for when identifying (causal) disease genes to avoid false positive results.

3. The process of healthy aging can be described as a change in cell populations in blood, rather than a change in gene expression within the cells.

4. Because each methodology has its flaws, integrating multiple independent lines of evidence is essential for trustworthy results.

5. Despite evolutionary constraints, local genetic regulation of gene expression can have large effects. Therefore, such cis-regulation is of limited use when understanding common complex diseases.

6. Common and rare disease genetics are traditionally viewed as independent areas of research, but they are at two ends of the same spectrum and can benefit from each other's insights.

7. While disease associations are generally small, their consequences ultimately lead to disease. Large population-based biobanks are required to detect the subtle patterns that lead to the development of disease.

8. In as far as they exist, finding core genes for common complex diseases will be the key to understand and treat these diseases.

9. Biology is infinitely complex: each cell in each organ in each (diseased or healthy) individual is unique. Every level complexity will expose more information, leading to new questions and knowledge.

The more we know, the more we know we don't know
(attributed to Aristotle)

10. Modern science is collaborative, and better for it.

“The scientific enterprise as a whole does from time to time prove useful, open up new territory, display order, and test long-accepted belief. Nevertheless, the individual [or team] engaged on a normal research problem is almost never doing any one of these things. Once engaged, [their] motivation is of a rather different sort. What then challenges [them] is the conviction that, if only [they are] skillful enough, [they] will succeed in solving a puzzle that no one before has solved or solved so well.”

Thomas Kuhn, *The Structure of Scientific Revolutions*

Square brackets indicate modified from the original to acknowledge the diversity and collaborative nature of modern science.

Contents

Chapter 1	Introduction	11
Chapter 2	The genetic architecture of molecular traits	25
Chapter 3	Evaluation of commonly used analysis strategies for epigenome- and transcriptome-wide association studies through replication of large-scale population studies	39
Chapter 4	Correction for both common and rare cell types in blood is important to identify genes that correlate with age	69
Chapter 5	Mendelian randomization while jointly modeling <i>cis</i> genetics identifies causal relationships between gene expression and lipids	89
Chapter 6	Large-scale <i>cis</i> - and <i>trans</i> -eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression	123
Chapter 7	Linking common and rare disease genetics to identify core genes using Downstreamer	171
Chapter 8	Discussion	195
Appendices	Summary	228
	Samenvatting	232
	Acknowledgments	236
	Curriculum vitae and publications	241

