

University of Groningen

The interplay between genetics, the microbiome, DNA-methylation & gene-expression

Bonder, Marc Jan

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:

2017

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

Citation for published version (APA):

Bonder, M. J. (2017). The interplay between genetics, the microbiome, DNA-methylation & gene-expression [Groningen]: University of Groningen

Copyright

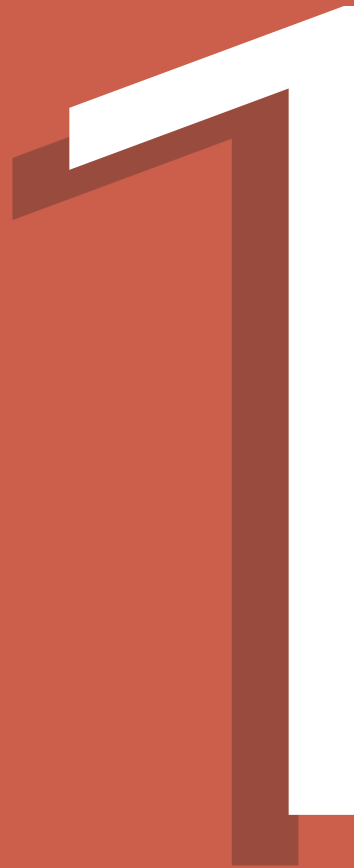
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

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.

General introduction



In recent years we have been highly successful in identifying the genetic basis of disease¹. DNA genotyping and, in particular, sequencing technologies have advanced tremendously and prices have come down substantially. It is now possible to collect genotyping information on large numbers of patients and controls, which enables researchers to identify genetic variants that are associated to (complex) diseases or traits by conducting so-called genome-wide association studies (GWAS). For instance, for inflammatory bowel disease over a hundred genetic variants have been identified by systematically comparing thousands of patients with thousands of controls^{2,3}. However, although GWAS have identified hundreds of associated loci, they do not provide mechanistic information on how these variants ultimately cause disease. This is particularly challenging, since the majority of the identified GWAS variants are not changing protein structure, but are non-coding and must thus have regulatory effects.

To gain a better functional understanding of these variants, quantitative trait loci (QTL) mapping studies are now being conducted⁴. The most common form of QTL mapping is expression QTL (eQTL) mapping, which allows us to link a genetic variant to its effects on gene expression. Two different types of eQTLs have been defined: *cis* local QTL effects and *trans* distal QTL effects. To date, most large-scale studies on eQTLs have been performed in blood, since it is easy to collect from patients and controls. However, the *cis* and *trans* eQTLs identified can be very tissue- and context-specific, so the effects observed in blood might not be representative for expression in other tissues⁵. In the largest *trans*-eQTL study to date⁶, 233 GWAS associated variants have been linked to expression variation, giving insights into the mechanism of action of these variants.

Besides expression, much effort has been put into the mapping of genetic influence on DNA-methylation. DNA-methylation is a key component of the epigenome. By studying DNA-methylation levels in a genomic region, one can gain insight into the regulatory potential of the genomic region. Using DNA-methylation QTL (meQTL) mapping, we acquire complementary data to eQTLs⁷, which also helps to provide more insight into the downstream effects of genetic variation in health and disease.

However, the great majority of complex diseases are not solely caused by genetic factors, but also by environmental factors. Unfortunately, for many diseases, these environmental risk factors still need to be identified. A major challenge is that in many diseases it is not yet clear what these environmental risk factors might be or even how they can be identified.

Paradoxically, a promising way to overcome this problem is to take advantage of the massive improvement in genotyping technologies. For instance, DNA methylation chips provide information on over 485,000 different CpG sites at once, and the variation in measured DNA methylation can be a strong proxy for phenotypic status or environmental exposures: CpG sites have now been found to be near-perfect proxies for age and many other associations are being determined through epigenome-wide association studies (EWAS)⁸. This suggests that other CpG sites might be representative proxies for more environmental factors, some of which might represent risk factors for certain diseases.

DNA sequencing improvements now also make it possible to identify and quantify micro-organisms. There are a comparable number of microbial and human cells within the human body and on its surface⁹, but there are roughly 150 times more microbial genes than human genes¹⁰. The microbiome is a collection of bacteria, archaea and viruses living together in a community, and they collectively perform important functions for the host. The largest fraction of the human microbiome is located in the digestive tract, where it has important functions in the metabolism but has also been shown to interact with the immune system. The gut microbiome is linked to multiple environmental and intrinsic factors, for instance age¹¹, gender¹² and diet¹³. In inflammatory bowel disease, differences in the composition of the microbiome have been linked to the disease. This suggests that the onset and progression of disease could be altered by changing the microbiome.

1 **The aim of this thesis was to use the new biological data based on sequencing technology, to study the role of genetic and environmental factors in disease, and to ascertain how genetic variation and environmental factors are related to variations in phenotype, gene expression, methylation and microbial composition.**

2 In the first part of this thesis, the relationship between several phenotypic factors and the microbiome were studied. Studies on the links between variation in the human microbiome and diet (chapter two), medication use (chapter three) and lipid levels (chapter four) are presented. Chapter five presents an integrative analysis in which 126 exogenous and intrinsic factors influencing the gut microbiome were identified. In chapter six, we report on how the host genome influences the microbiome composition.

3 In the second part of the thesis, the relationship between DNA-methylation, gene expression, phenotypes, and genetic variation were studied. Chapter seven describes the link between genetic variation and DNA-methylation and expression levels in several different tissues. In chapter eight, epi-genomic and genomic risk scores have been used to explain variation in individual traits like height and BMI. In chapter nine we studied the effects of genetic risk factors on downstream molecular data. To investigate the effects of the risk factors we integrated multiple omics layers, including DNA-methylation, RNA-sequencing, gene expression and genetics.

4 In the third part of the thesis, I discuss the results and show the combined interpretation of the results presented in the previous chapters. More specifically I discuss two approaches that can be used for multi-omics integration studies. These results showcase future research possibilities and highlights possibilities to gain more insight into health and disease by integrating multiple biological omics data.

Definitions

Epigenome	The set of chemical compounds attached to the DNA.
Microbiome	The set of micro-organisms, as represented by the genetic information, in a particular environment.
Locus	A site on the human genome.
GWAS	Genome-Wide Association Study, the study of relating variation in the human genome to a trait or disease.
EWAS	Epigenome-Wide Association Study, the study of relating variation in the human epigenome to a trait or disease.
QTL	Quantitative Trait Loci, a locus in the human genome having a relation with a quantitative trait, such as expression (eQTL), DNA-methylation (meQTL), or the microbiome (miQTL).
CpG site	A site on the human genome where a cytosine is followed by a guanine; this is a site where DNA-methylation can take place.

References

1. Welter, D. *et al.* The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* **42**, D1001-6 (2014).
2. Jostins, L. *et al.* Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491**, 119–124 (2012).
3. Liu, J. Z. *et al.* Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat. Genet.* **47**, 979–986 (2015).
4. Edwards, S. L., Beesley, J., French, J. D. & Dunning, A. M. Beyond GWASs: illuminating the dark road from association to function. *Am. J. Hum. Genet.* **93**, 779–797 (2013).
5. Fu, J. *et al.* Unraveling the regulatory mechanisms underlying tissue-dependent genetic variation of gene expression. *PLoS Genet.* **8**, e1002431 (2012).
6. Westra, H.-J. *et al.* Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat. Genet.* **45**, 1238–1243 (2013).
7. Banovich, N. E. *et al.* Methylation QTLs Are Associated with Coordinated Changes in Transcription Factor Binding, Histone Modifications, and Gene Expression Levels. *PLoS Genet.* **10**, 1–12 (2014).
8. Rakyan, V. K., Down, T. a, Balding, D. J. & Beck, S. Epigenome-wide association studies for common human diseases. *Nat. Rev. Genet.* **12**, 529–41 (2011).
9. Sender, R. *et al.* Are We Really Vastly Outnumbered? Revisiting the Ratio of Bacterial to Host Cells in Humans. *Cell* **164**, 337–340 (2016).
10. Qin, J. *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59–65 (2010).
11. T, Y. *et al.* Anokhin {AP} *et al.*: {Human} gut microbiome viewed across age and geography. *Nat.* 2012 **486**, 222–227 (2012).
12. Yurkovetskiy, L. *et al.* Gender bias in autoimmunity is influenced by microbiota. *Immunity* **39**, 400–412 (2013).
13. David, L. A. *et al.* Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505**, 559–563 (2014).