

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.

Appendices

1

1

Summary

There are many factors involved in the development of human diseases and traits. In recent years the field of human genetics has been very successful in linking genetic variation to diseases and traits. By conducting large-scale studies comparing the genetic make-up of affected versus non-affected participants, we have identified thousands of variants in the human genome that are more or less commonly found in cases compared to controls. These genome-wide association studies (GWAS) have been instrumental in the identification of genes linked to a multitude of diseases and traits. Variants in the functional parts of a gene can be relatively straightforward to interpret. However, not all the variants linked to disease can be directly interpreted. By using intermediate molecular data layers, such as gene expression, DNA-methylation or protein levels, we can gain more insight into the genetic variants identified by GWAS.

However, we only get a limited picture of disease by focusing on genetic variation. Another important factor related to disease is the environment. But it is much harder to quantify environmental factors than to determine the genetic differences between two individuals. Using the intermediate molecular data, or biological omics, we can gain insights into the environment of individuals. The environment surrounding individuals can, for instance, influence the composition of their microbiome, but also their gene expression, DNA-methylation and protein levels. By studying the differences in these biological omics in relation to phenotypes and disease, we can learn more about the environmental factors that lead to disease. However, as with GWAS studies, we do not always know what the differences in these biological data layers mean.

In this thesis we have focused on two biological omics, the gut microbiome composition and the DNA-methylome. The gut microbiome is the collection of micro-organisms that live together in the human gut; DNA-methylation is the occurrence of a methyl group bound to the DNA and this mainly occurs at cytosine-guanine pairs. In the first part of the thesis we have focused on inter-individual differences influencing, or influenced by, differences in the microbiome composition, while in the second part, we have focused on changes in DNA-methylation associated to tissue differences and on the influence of genetic variation on DNA-methylation.

The microbiome

There are many factors which influence the gut microbiome. In **chapter two** we describe a study into the specific influence of a gluten-free diet (GFD) on the human microbiome. A gluten-free diet is the most commonly adopted special diet worldwide. It is the only effective “treatment” for coeliac disease, but is also often adopted by individuals with other gastrointestinal complaints. We found that inter-individual variation in the gut microbiota remained stable during the GFD intervention, although on an individual level the most striking shift in composition was seen for the bacterial family *Veillonellaceae*. *Veillonellaceae* is considered to be a pro-inflammatory family of bacteria. Because of this it is conceivable that a decrease in *Veillonellaceae* abundance might be one of the mediators of the GFD’s beneficial effect observed in patients with IBS and gluten-related disorders.

Apart from diet, we also specifically tested for the influence of commonly used drugs on the microbiome, this study is described in **chapter three**. We found that proton pump inhibitors (PPI), which are among the top ten most widely used drugs in the world, have an even more prominent effect on a population level than antibiotics. PPI use was associated with a significant decrease in microbial diversity and with changes in 20% of the bacterial taxa. Multiple oral bacteria were overrepresented in the fecal microbiome of PPI users. The differences between PPI users and non-users seen in this study were consistently associated with changes towards a less healthy gut microbiome.

1
2
3
4

Since many studies have suggested there is a relation between the gut microbiome and the development of cardiovascular disease, we studied the relation between lipid levels, body mass index (BMI) and the microbiome. In **chapter four** we report a study that identified 34 bacterial taxa associated to BMI and blood lipids. Furthermore, we built a model to explain differences in lipid and BMI levels between different individuals based on both their genetics and microbiome composition. This showed that the microbiome composition could explain up to 6% of variance in lipid levels (triglycerides), on top of the variation explained by age, gender and host genetics. Our findings support the potential of therapies altering the gut microbiome to control body mass, triglycerides and HDL.

Most microbiome studies to date have been focused on the abundances of microbial species, however, there are complex interactions between microbes living in the gut. Since looking at species only gives a limited resolution, we used a strategy to not only look at the abundance of the different microbes, but also at the number of reads per gene present in the microbiome. In **chapter five** we report on a study into the factors influencing both the microbes and microbial functions. We identified 126 exogenous and intrinsic host factors, including 31 intrinsic factors, 12 diseases, 19 drug groups, 4 smoking categories, and 60 dietary factors that all influence the microbiome in some way. These factors collectively explain 18.7% of the variance seen in the inter-individual microbial composition.

After identification of the main factors influencing the microbiome, we went on to search specifically for the relation between host genetics and the gut microbiome, as described in **chapter six**. We assessed the influence of host genetics on microbial species and function, which we investigated by looking at gene pathways and GO categories in 1,514 subjects. We identified 41 genomic loci that were significantly linked to a difference in the microbiome composition. In addition, we investigated genomic regions involved in diseases, immunity or food preferences, and found 32 loci that were associated with microbiome composition.

DNA-methylation

In the second part of this thesis we report on the relation between genomic variation and the DNA-methylome, and factors related to the DNA-methylome. In **chapter seven** we describe the differences between DNA-methylation and gene expression in fetal and adult liver and compare the local genetic relation on DNA-methylation and gene expression in adult liver to two types of fat and muscle. When comparing adult versus fetal liver we identified 1,657 differentially methylated genes; these genes were enriched for transcription factor binding sites of HNF1A, HNF4A, GATA1, STAT5A, STAT5B, and YY1. We also identified, 2,673 differentially expressed genes; these genes were enriched for metabolic and developmental pathways. When comparing the genetic control on liver we observed strong liver-specific effects from single nucleotide polymorphisms (SNPs) on both methylation levels (28,447 unique CpG sites) and gene expression levels (526 unique genes).

In **chapter eight** we report on tests to discover whether DNA-methylation profiles account for the inter-individual variation in BMI and height. We derived methylation predictors for both BMI and height by estimating probe-trait effects in discovery samples and tested them in external samples. Methylation profiles associated with BMI based on the LifeLines-DEEP cohort explained 4.9% and 3.6% of the variation in BMI seen in two other datasets (Lothian Birth Cohorts and the Brisbane System Genetic Study, respectively). Methylation profiles predicted BMI independently of genetic profiles in an additive manner: 5%, 9%, and 13% of variance of BMI in LifeLines-DEEP subjects were explained, respectively, by the methylation predictor, genetic predictor, and a model containing both. In contrast, methylation profiles accounted for almost no variation in height. The BMI results suggest that combining genetic and epigenetic information might have greater utility for predicting complex traits.

In **chapter nine** we describe a study on disease-associated genetic variants and their relation to DNA-methylation. We show that disease variants have widespread effects on distal DNA methylation sites, and these likely reflect differential occupancy of *trans*-binding sites (i.e. sites located far from the genetic risk factor of interest) by *cis*-regulated transcription factors (located near the genetic risk factor of interest). For 1,907 established trait-associated SNPs, we found that they affect distal methylation levels of 10,141 different CpG sites (false discovery rate <0.05). They included SNPs that affect both the expression of a nearby transcription factor (like *NFKB1*, *CTCF* and *NKX2-3*) and the methylation of its respective binding site across the genome.

Eight main points from this thesis

1. The gut microbiome is influenced by many exogenous and intrinsic factors.
2. A gluten-free diet has limited but significant effects on microbiome composition.
3. Genetic differences between individuals influence their microbiome composition.
4. DNA-methylation and gene-expression in liver differs between adult liver and fetal liver.
5. The genetic control on DNA-methylation varies per tissue type.
6. Genetic variants found to be associated by GWAS have effects on distal methylation sites.
7. Distal effects on methylation work fully or partly by influencing local changes in the gene expression of transcription factors.
8. Using other biological omics datasets, in addition to genomic data, helps in predicting complex phenotypes such as BMI.

Samenvatting

1 Er zijn vele factoren die betrokken zijn bij de ontwikkeling van ziekten en fenotypen. In de afgelopen jaren zijn er veel successen geboekt in het koppelen van genetische variatie aan ziektes en uiterlijke kenmerken. Door het uitvoeren van grootschalige studies waarin de genetische variatie tussen getroffen versus niet-getroffen deelnemers wordt vergeleken, hebben zijn er duizenden varianten in het menselijk genoom geïdentificeerd die minder of meer voorkomen in de cases in vergelijking met controles. Deze genomwijde associatiestudies (GWAS) hebben een grote rol gespeeld bij de identificatie van de genen die een relatie hebben met ziekten en patiënt karakteristieken. Varianten in de functionele delen van een gen kunnen relatief eenvoudig te interpreteren zijn. Echter, niet alle varianten gekoppeld aan ziekte zijn direct te interpreteren. Via moleculaire data lagen, zoals genexpressie, DNA-methylatie of eiwitniveaus kunnen we meer inzicht krijgen in de genetische varianten die door GWAS zijn gevonden.

2 Echter, als er alleen wordt gefocust op genetische variatie in relatie tot ziekten krijgen we maar een beperkt beeld van de ziekten. Een andere belangrijke factor in de ontwikkeling van ziekte is de leefomgeving van een persoon. Maar omgevingsfactoren zijn veel moeilijker te kwantificeren en te vergelijken tussen twee individuen dan de genetische verschillen. Met behulp van de intermediaire moleculaire gegevens, of biologische-omics (omics), kunnen we inzicht krijgen in de omgeving van individuen. De omgeving van individuen kan bijvoorbeeld invloed hebben op de samenstelling van zijn of haar microbiom, maar ook de genexpressie, DNA-methylatie en eiwitniveaus kunnen verschillend zijn hierdoor. Door het bestuderen van de verschillen in deze omics ten opzichte van fenotypes en ziekten, kunnen we de omgevingsfactoren die leiden tot ziekte identificeren. Echter, zoals met GWAS-studies, we weten niet altijd wat de verschillen in deze biologische data lagen betekenen.

3 In dit proefschrift hebben we ons gericht op twee biologische-omics, de samenstelling van het microbiom in de darm en het DNA-methylom. Het darm microbiom is de verzameling van micro-organismen die in de menselijke darm samenleven, DNA-methylering is het optreden van een methylgroep gebonden aan het DNA, dit komt vooral voor bij cyteïne-guanine (CpG) paren. In het eerste deel van dit proefschrift hebben we ons gericht op interindividuele verschillen in de samenstelling van het microbiom, terwijl we in het tweede deel van het proefschrift ons hebben gericht op veranderingen in de DNA-methylatie geassocieerd met weefsel verschillen en de invloed van genetische variatie op DNA-methylatie.

Het microbiom

4 Er zijn vele factoren die het darm microbiom beïnvloeden. In **hoofdstuk twee** wordt een onderzoek beschreven naar de specifieke invloed van een glutenvrij dieet (GFD) op het menselijk microbiom. Een glutenvrij dieet is het meest gebruikte speciale dieet wereldwijd. Het is de enige effectieve "behandeling" voor coeliakie, maar wordt ook vaak door mensen met andere darmklachten gebruikt. In de studie vonden wij dat interindividuele variatie in de darmflora stabiel is gebleven tijdens de GFD-interventie, de meest opvallende verschuiving die gezien werd binnen de individuen was de veranderde hoeveelheid van de bacteriële familie *Veillonellaceae*. De *Veillonellaceae* familie wordt beschouwd als een pro-inflammatoire familie van bacteriën. Hierdoor is het denkbaar dat de vermindering van de aanwezigheid van deze bacteriële familie een van de redenen is dat mensen een dieet een GFD gebruiken het als positief ervaren.

Naast voeding, hebben we ook specifiek onderzocht wat de invloed van veel gebruikte medicatie is op het microbiom, dit is beschreven in **hoofdstuk drie**. We vonden dat protonpompremmers (proton pump inhibitors, PPI), die behoren in de top tien van meest gebruikte medicijnen in de wereld, op een populatieniveau nog prominentere effect hebben dan antibiotica. PPI-gebruik werd geassocieerd met een significante daling van microbiële diversiteit en veranderingen in

20% van de bacteriële taxa. Meerdere orale bacteriën zijn oververtegenwoordigd in het fecale microbiom van de PPI-gebruikers. De verschillen tussen PPI-gebruikers en niet-gebruikers die zijn gevonden in dit onderzoek zijn consistent geassocieerd met veranderingen naar een minder gezond darm microbiom.

Aangezien veel studies hebben gesuggereerd dat er een verband is tussen het microbiom en de ontwikkeling van hart- en vaatziekten, hebben we de relatie tussen lipide niveaus, body mass index (BMI) en het microbiom onderzocht. In **hoofdstuk vier** beschrijven we een studie waar in we hebben gevonden dat 34 bacteriële taxa geassocieerd zijn met BMI en bloedlipiden. Bovendien bouwden we een model om verschillen in lipide en BMI-niveaus te voorspellen, op basis van zowel genetische informatie en microbiom samenstelling. Hieruit bleek dat met behulp van de samenstelling van het microbiom 6% van de variatie in lipide niveaus (triglyceriden) kan worden verklaard, bovenop de variatie verklaard door leeftijd, geslacht en humane genetica. Onze vindingen onderschrijven de mogelijkheid om BMI en lipiden te beïnvloeden door middel van interventie op microbiom niveau.

De meeste microbiom studies hebben zich tot op heden gericht op het kwantificeren van microbiële soorten, maar er zijn complexe interacties tussen verschillende microben in de darm. Omdat kijken naar soorten slechts een beperkt beeld geeft gebruikten wij een strategie om niet alleen te kijken naar de hoeveelheden van de bacterie maar ook naar de genen in de bacterie en de aanwezigheid van het gen in het microbiom. In **hoofdstuk vijf** beschrijven we een onderzoek naar de factoren die van invloed zijn op zowel de microben en microbiële functies. We identificeerden 126 exogene en intrinsieke factoren, waaronder 31 intrinsieke factoren, 12 ziekten, 19 medicatie groepen, 4 rook categorieën en 60 voedingsfactoren die allemaal invloed hebben op het microbiom. Deze factoren tezamen verklaren 18,7% van de variatie in de interindividuele microbiële samenstelling.

Na identificatie van de belangrijkste factoren die invloed hebben op het microbiom, gingen we specifiek op zoek naar de relatie tussen humane genetica en het darm microbiom, zoals beschreven in **hoofdstuk zes**. Wij hebben de invloed van humane genetische variatie op de microbiële soorten en functie, die we onderzochten door te kijken naar gen pathways en GO-categorieën, onderzocht in 1.514 individuen. We identificeerden 41 genomische loci die significant zijn gerelateerd aan een verschil in de microbiom compositie. Daarnaast onderzochten we genomische regio's die betrokken zijn bij ziekten, immuniteit of voedsel voorkeuren, en hebben 32 loci gevonden die suggestief zijn geassocieerd met de samenstelling van het microbiom.

DNA-methylatie

In het tweede deel van dit proefschrift beschrijven we de relatie tussen genomische variatie en het DNA-methylom, en factoren die verband houden met de DNA-methylom. In **hoofdstuk zeven** beschrijven we de verschillen tussen DNA-methylatie en genexpressie in foetale levers versus volwassen levers en vergelijken we de lokale genetische invloed op DNA-methylatie en genexpressie in een volwassen lever in vergelijking tot twee soorten vet en spieren. We identificeerden 1.657 differentieel gemethyleerde genen; deze genen zijn verrijkt voor transcriptiefactor bindingsplaatsen van HNF1A, HNF4A, GATA1, STAT5A, STAT5b en YY1. We vonden ook 2.673 differentieel tot expressie komende genen, deze genen zijn verrijkt voor genen die te maken hebben met metabolisme en ontwikkeling. We zagen sterk lever-specifieke effecten van genetische variatie op beide niveaus, methylatie (28.447 unieke sites) en genexpressie (526 unieke genen).

In **hoofdstuk acht** beschrijven we een studie waar in we DNA-methylatie-profielen gerelateerd hebben aan interindividuele variatie in BMI en lengte. We gebruikten DNA-methylatie verschillen als voorspeller voor zowel BMI en lengte door relaties te schatten in LifeLines-DEEP individuen en deze relaties te valideren in externe individuen. Met behulp van DNA-methylatie profielen geassocieerd met BMI op basis van het LifeLines-DEEP cohort kan 4,9% en 3,6% van de variatie

1
2
3
4

in BMI worden verklaard in de twee andere datasets, de Lothian birth cohorts en Brisbane System genetische studie. We vonden dat we met methylatie profielen BMI onafhankelijk van genetische profielen kan worden verklaard op een additieve wijze: 5%, 9% en 13% van de variantie van BMI in LifeLines-DEEP participanten kan worden verklaard door respectievelijk, de methylering voorspeller, genetische voorspeller, en een model dat zowel methylatie als genetische informatie gebruikt. Daarentegen verklaren methylatie profielen bijna geen variatie in lengte, wat te verwachten is aangezien lengte een eigenschap is die erg over erfbaar is. De BMI resultaten suggereren dat het combineren van genetische en epi-genetische informatie van grote waarde kunnen zijn bij het voorspellen van complexe menselijke fenotypen.

In **hoofdstuk negen** beschrijven we een studie waar we kijken naar de relatie van DNA-methylatie met ziekte-geassocieerde genetische varianten. We laten zien dat de ziekte-geassocieerde varianten grote gevolgen hebben op DNA-methylatie op andere gebieden van het genoom. Dit wordt waarschijnlijk veroorzaakt door differentiël gebruik van *trans*-bindingsplaatsen (d.w.z. locaties die zich ver van de genetische risicofactor bevinden) van door *cis*-gereguleerde transcriptiefactoren (gelegen in de buurt van de genetisch risico factor). Voor 1.907 ziekte-geassocieerde genetische varianten hebben we gevonden dat ze invloed hebben op in totaal 10.141 verschillende CpG sites. Een deel van de genetische risicofactoren beïnvloeden zowel de expressie van een nabijgelegen transcriptiefactor (zoals *NFKB1*, *CTCF* en *NKX2-3*) en de methylering van zijn respectieve downstream bindingsplaats in het genoom.

Acht hoofdpunten uit dit proefschrift

1. Het darm microbioom wordt beïnvloed door vele exogene en intrinsieke factoren.
2. Een glutenvrij dieet heeft een beperkt, maar significant effect op de samenstelling van het microbioom.
3. Genetische variatie beïnvloed de persoonlijke microbioom samenstelling.
4. DNA-methylatie en gen-expressie in de lever verschilt tussen volwassen level en foetale lever,.
5. De genetische invloed op DNA-methylatie verschilt per weefsel.
6. Genetische varianten gevonden door GWAS, hebben invloed op distale methylatie sites.
7. Distale gevolgen op methylatie zijn deels het gevolg van lokale genexpressie verschillen van transcriptiefactoren.
8. Het gebruik van andere biologische-omics, naast genomische data, helpt bij het voorspellen van complexe fenotypen zoals BMI.

List of abbreviations

16s rRNA	16s ribosomal RNA gene	TC	total cholesterol
ADME	absorption, distribution, metabolism and excretion	TCAs	tricyclic antidepressants
AUC	area under the curve	TG	triglycerides
BGI	Beijing Genomics Institute	TMAO	trimethylamine N-oxide
BMI	Body mass index	TSS	transcription start site
CGI	CpG islands	VAT	visceral adipose tissue
CVD	cardiovascular disease		
ECLIA	electro-chemiluminescence immunoassay		
EDTA	ethylenediaminetetraacetic acid		
ELISA	enzyme-linked immunosorbent assay		
eQTL	expression quantitative trait loci		
eQTM	expression quantitative trait methylation		
FDR	false discovery rate		
FISH	fluorescence in situ hybridization		
GC_MS	gas chromatography-mass spectrometry		
GFD	Gluten-free diet		
GWA	genome-wide association		
HD	Habitual diet		
HDL	high-density lipoprotein		
HPLC	high performance liquid chromatography		
IBD	Inflammatory Bowel Diseases		
IBS	Irritable Bowel Syndrome		
IBS	Irritable bowel syndrome		
KEGG	Kyoto encyclopedia of genes and genomes		
LDL	low-density lipoprotein		
meQTL	methylation quantitative trait loci		
NSAID	nonsteroidal anti-inflammatory drug		
OTU	Operational taxonomic unit		
PCoA	Principal Coordinate Analysis		
PPI	Proton Pump Inhibitor		
qPCR	quantitative real-time polymerase chain reaction		
RIA	radioimmunoassay		
SAT	subcutaneous adipose tissue		
SCFA	short chain fatty acids		
SD	standard deviation		
SNP	single nucleotide polymorphism		
SNRIs	serotonin-norepinephrine reuptake inhibitors		
SSRI	serotonin-specific reuptake inhibitors		

Acknowledgments

Dear family, friends and colleagues,

I want to thank all of you for your big or small contribution to my thesis! I really enjoyed the years that I spent in Groningen working and learning, and enjoying my time with all of you and I am proud of the work presented here. Some of you deserve a personal thank you, but these are by no means the only people who have contributed to this work. Big thanks to everybody I do not mention specifically.

Lude, Sasha and Cisca: Thank you so much for giving me the space to be able to do this work! I really enjoyed working in the lab, before and during my PhD! Thanks for pushing me, all the opportunities you have given me, and all the support and guidance in the various projects! Thanks for creating such a great atmosphere and such a nice group. Without the three of you, this work would not have been possible! I wish you all the best for your future projects.

Dear **Patrick:** I do not know where to start or how to thank you. Without you, my PhD, Master's and Bachelor's periods would have been so different. Thanks for all the great things and everything you have taught me!

Dear **Arnau and Floris:** The other two microbiome guys, thanks for all the teamwork. I really enjoyed working together in our exploration of the microbiome and, in particular, on the PPI paper. Also, I want to say sorry for not using the correct cover, but please go ahead and use the Telegraaf picture on your thesis. I would also like to thank all the other members of the Poepgroep, in particular **Jing, Alex, Wouter, Rinse, Jackie D** and **Marten**. Thanks for all your help and I enjoyed working with you!

The P? -Medicine platform team (**Freerk, Urmo, Niek, Sipko, Annique, Marion, Joeri, Lucas** and **Adriaan**), I no longer know what the amount of P's is or what they were all supposed to represent, but I would like to thank all of you for your help and the nice atmosphere in the Franke-Swertz group. **Harm-Jan, Juha, Dasha, Isis, Javier**, thanks for the warm welcome to the group, the good times and everything you taught me. **Pieter** and **Morris**, thanks for the great help and managing the cluster, lots of this work would not be possible without this!

A big thanks to all the co-authors of the many projects I have had the opportunity to play a part in. Thanks to **Yang, Cheng, Vinod**, and **Sebo** and all other people in the group, I enjoyed working with you in the Wijmenga group!

Kate & Jackie S: Thank you so much for your unbelievable help in editing my sometimes horrible texts. Without you, the messages in this thesis would have been much less clear and the text full of "*kromme zinnen*".

Mentje, Helene, Bote and **Joke:** Thanks for taking care of much of the administrative things, which have come along during my period at the genetics department.

Dennis: Thanks for your help with my thesis lay-out and making the awesome cover.

Dear **Steven, Papa** and **Mama:** Thanks for all the interest you have shown in my research, even though most of the time it was hard for you to follow what I was doing. And thanks for all the support you have given me.

Dear **Suus:** Thanks for all the patience you have had, every time when I wanted to finish one more thing and I lost track of the time again. Without you, this thesis would not have been possible! I cannot thank you enough!

Curriculum vitae

Marc Jan Bonder was born 28 march 1989 in Tolbert, The Netherlands. He finished his Bachelor degree in bioinformatics at the Hanze University of Applied Sciences in 2010 and received his Master degree in bioinformatics at the Vrije Universiteit Amsterdam in 2012. After this he moved back to Groningen to do a PhD under joint supervision of Dr. Alexandra Zhernakova and Prof. Dr. Lude Franke at the Department of Genetics, University of Groningen and University Medical Centre Groningen. He is currently starting a Post-Doc at the EMBL European Bioinformatics Institute in the lab of Dr. Oliver Stegle.

List of publications

The influence of proton pump inhibitors and other commonly used medication on the gut microbiota. Imhann et al. Gut Microbes, 00-00

The emerging landscape of dynamic DNA methylation in early childhood. CJ Xu et al. BMC genomics 18 (1), 25

Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. S Wahl et al. Nature

Disease variants alter transcription factor levels and methylation of their binding sites. MJ Bonder & R Luijk et al. Nature Genetics (2016)

The effect of host genetics on the gut microbiome. MJ Bonder & A Kurilshikov et al. Nature Genetics 48 (11), 1407-1412.

Identification of context-dependent expression quantitative trait loci in whole blood. DV Zhernakova et al. Nature Genetics

Linking the human gut microbiome to inflammatory cytokine production capacity. M Schirmer et al. Cell 167 (4), 1125-1136. e8

Genome-wide analysis identifies 12 loci influencing human reproductive behavior. N Barban et al. Nature genetics

Evidence for mitochondrial genetic control of autosomal gene expression. I Kassam et al. Human Molecular Genetics, ddw347

Age-related accrual of methylomic variability is linked to fundamental ageing mechanisms. RC Slieker et al. Genome Biology 17 (1), 191

A GWAS meta-analysis suggests roles for xenobiotic metabolism and ion channel activity in the biology of stool frequency. SA Jankipersadsing et al. Gut, gutjnl-2016-312398, 2016

Blood lipids influence DNA methylation in circulating cells. KF Dekkers et al. Genome Biology 17 (1), 13, 2016

Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. A Zhernakova et al. Science 352 (6285), 565-569, 2016

Population-level analysis of gut microbiome variation. G Falony et al. Science 352 (6285), 560-564, 2016.

Tobacco smoking is associated with DNA methylation of diabetes susceptibility genes. S Ligthart et al. Diabetologia 59 (5), 998-1006, 2016

1 **Proton pump inhibitors affect the gut microbiome.** F Imhann & MJ Bonder & AV Vila et al. Gut 65 (5), 740-748, 2016

2 **The influence of a short-term gluten-free diet on the human gut microbiome.** MJ Bonder & EF Tigchelaar, et al. Genome Medicine 8 (1), 1, 2016

3 **Improving Phenotypic Prediction by Combining Genetic and Epigenetic Associations.** S Shah & MJ Bonder et al. The American Journal of Human Genetics 97 (1), 75-85, 2015

4 **Gut microbiota composition associated with stool consistency.** EF Tigchelaar et al. Gut, gutjnl-2015-310328, 2015

Trans-ancestry genome-wide association study identifies 12 genetic loci influencing blood pressure and implicates a role for DNA methylation. N Kato et al. Nature genetics 47 (11), 1282-1293, 2015

Genotype harmonizer: automatic strand alignment and format conversion for genotype data integration. P Deelen & MJ Bonder et al. BMC Research Notes 7 (1), 901, 2015

Calling genotypes from public RNA-sequencing data enables identification of genetic variants that affect gene-expression levels. P Deelen & DV Zhernakova et al. Genome Medicine 7 (30), 13, 2015

The Gut Microbiome Contributes to a Substantial Proportion of the Variation in Blood Lipids. J Fu et al. Circulation Research 117 (9), 817-824, 2015

Genetic and epigenetic regulation of gene expression in fetal and adult human livers. MJ Bonder & S Kasela et al. BMC genomics 15 (1), 1, 2014

Breast Cancer Subtype Specific Classifiers of Response to Neoadjuvant Chemotherapy Do Not Outperform Classifiers Trained on All Subtypes. JJ de Ronde & MJ Bonder et al. PLOS ONE 9 (2), e88551, 2014

Comparing clustering and pre-processing in taxonomy analysis. MJ Bonder et al.. Bioinformatics 28 (22), 2891-2897, 2012

The Relation between Oral Candida Load and Bacterial Microbiome Profiles in Dutch Older Adults. EA Kraneveld et al. PLoS ONE 7 (8), e42770, 2012

TaxMan: a server to trim rRNA reference databases and inspect taxonomic coverage. BW Brandt, et al. Nucleic Acids Research, 2012

Trans-eQTLs reveal that independent genetic variants associated with a complex phenotype converge on intermediate genes, with a major role for the HLA. RS Fehrmann et al. PLoS Genet 7 (8), e1002197, 2011