

University of Groningen

Genetic determinants of human phenotypes

Zhang, Zhenhua

DOI:
[10.33612/diss.211254963](https://doi.org/10.33612/diss.211254963)

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:
2022

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

Citation for published version (APA):
Zhang, Z. (2022). *Genetic determinants of human phenotypes: understanding human infectious diseases by computational biology methods*. University of Groningen. <https://doi.org/10.33612/diss.211254963>

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).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

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.

-
- [83] Janek Gröhl et al. “Deep learning for biomedical photoacoustic imaging: A review”. In: *Photoacoustics* 22 (June 2021), p. 100241. DOI: 10.1016/j.pacs.2021.100241.
- [84] Haitham A. Elmarakeby et al. “Biologically informed deep neural network for prostate cancer discovery”. In: *Nature* 598.7880 (Sept. 2021), pp. 348–352. DOI: 10.1038/s41586-021-03922-4.
- [85] Riccardo Miotto et al. “Deep learning for healthcare: review, opportunities and challenges”. In: *Briefings in Bioinformatics* 19.6 (May 2017), pp. 1236–1246. DOI: 10.1093/bib/bbx044.
- [86] Sebastian Vollmer et al. “Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness”. In: *BMJ* (Mar. 2020), p. 16927. DOI: 10.1136/bmj.16927.
- [87] Yilun Zhou et al. “Do feature attribution methods correctly attribute features?” In: *arXiv preprint arXiv:2104.14403* (2021). DOI: <https://doi.org/10.48550/arXiv.2104.14403>.
- [88] Thilo Spinner et al. “explAIner: A Visual Analytics Framework for Interactive and Explainable Machine Learning”. In: *IEEE Transactions on Visualization and Computer Graphics* (2019), pp. 1–1. DOI: 10.1109/tvcg.2019.2934629.
- [89] Žiga Avsec et al. “Base-resolution models of transcription-factor binding reveal soft motif syntax”. In: *Nature Genetics* 53.3 (Feb. 2021), pp. 354–366. DOI: 10.1038/s41588-021-00782-6.

APPENDICES

Dutch Summary (Nederlandse samenvatting)

Acknowledgements

Curriculum vitae and publications

Nederlandse samenvatting

Genetische en omgevingsfactoren zijn bepalend voor menselijke eigenschappen, ook wel bekend als fenotypes. De onderliggende mechanismes van complexe eigenschappen zoals het verloop van infectieziekten zijn echter nog grotendeels onbekend. Recentelijk bieden statistische methodes zoals genome-wide association studies (GWAS) mogelijkheden om genetische varianten te identificeren die gecorreleerd zijn met fenotypische diversiteit, wat weer mogelijk wordt gemaakt door DNA-profilering met een hoge verwerkingscapaciteit zoals Next Generation Sequencing technologieën. Echter, deze associaties worden meestal niet begrepen omdat de GWAS signalen grotendeels gelokaliseerd zijn in niet-coderende regio's. Deze niet-coderende varianten worden verondersteld te functioneren als regulatorische elementen die celtype- en conditie specifiek zijn. De integratie van single-cel multi-omics en genetica stelt onderzoekers in staat om de GWAS signalen te ontleden en te valideren op single-cel resolutie. Bovendien is de associatie tussen genetica en infectieziekten vaag als gevolg van de gecompliceerde biologische informatiestroom, waardoor primitieve methodes zoals GWAS, die gebruik maken van logistische regressie, op de proef worden gesteld. Geavanceerde algoritmes zoals machine/deep learning zijn ontworpen om latente relaties tussen kenmerken en doelwitten te leren. Machine/deep learning methoden zijn al toegepast in biologisch onderzoek, maar er zijn nog meer aspecten te onderzoeken, zoals generalisatie van het model. In dit proefschrift heb ik genetische determinanten van menselijke fenotypes onderzocht, waarbij ik me specifiek heb gericht op menselijke infectieziekten en allel-specifieke expressie (ASE).

Allel-specifieke expressie en genetische varianten

Genetische en epigenetische kenmerken zijn afkomstig van twee ouders, wat resulteert in allelische verschillen tussen de twee reeksen haplotypes. De verschillen resulteren in verschillende gedragingen tussen allelen, bijvoorbeeld ASE. Epigenetische regulatie waren al bekende als mechanismes die ten grondslag liggen aan ASE, zoals genomische imprinting en X-chromosoom inactivatie. Bovendien beïnvloeden genetische loci gewoonlijk de allelische toestanden van chromatine en verstoren ze de volgorde van de bindingsplaatsen voor transcriptiefactoren (TFBSs).

In dit proefschrift veronderstellen we dat de ASE van een enkele nucleotide variant (SNV) verklaard kan worden door zijn genomische context, d.w.z., de genomische kenmerken rondom de SNV. Daarom hebben we in Hoofdstuk 2 een gradient boosting tree model gebouwd om de ASE van exonische SNVs te voorspellen met

behulp van de genomische annotaties van de SNV. Het model presteert redelijk goed (test ROC AUC 0.8), wat suggereert dat de ASE van een SNV gerelateerd is aan zijn genomische annotaties.

Ons model heeft echter twee belangrijke beperkingen. Ten eerste hebben we het model uitsluitend gebouwd op basis van coderende varianten, waarbij regulatorische effecten van niet-coderende varianten in intergenische en intronische sequenties op de ASE van een transcript werden genegeerd. Ten tweede werd het model getraind op bloed monsters, waarin slechts een subset van genen tot expressie komt.

Om de hierboven genoemde beperkingen te overwinnen, hebben we vervolgens een nieuwe methode ontwikkeld, ASDeep, in Hoofdstuk 3. ASDeep is een deep learning (DL) tool die de ASE voor het canonieke transcript van een gen voorspelt met behulp van de DNA-sequentie die mogelijk de expressie van het doelgen reguleert. Dit werk is geïnspireerd door recente studies waarin DL benaderingen krachtig werden geacht om de functies van niet-coderende DNA sequenties te ontdekken.

Het model presteert echter niet zoals verwacht door vele beperkingen. Allereerst waren er slechts 3500 monsters beschikbaar, wat resulteerde in een hoog risico op over-fitting. Vervolgens werden deze monsters gegroepeerd in drie klassen, maar het aantal monsters uit elke klasse was zeer onevenwichtig. Ten slotte leken de DNA-sequenties die voor de training van het model werden gebruikt sterk op elkaar door de beperking van imputatie algoritmes voor genotypes, wat het model belemmerde om de geïdentificeerde klassen te discrimineren.

(Epi)genetische gastheer determinanten van infectieziekten

De menselijke genomische architectuur bepaalt de verschillende immuunreacties op microbes tussen individuen waarbij dodelijke infecties veroorzaakt door pathogene virussen meestal sterke immuunreacties uitlokken tijdens het ziekteverloop dan minder gevaarlijke infecties. Daarom zal het bestuderen van de genetische determinanten van menselijke infectieziekten ons inzicht verschaffen in de onderliggende mechanismen van microbe-mens interacties en uiteindelijk richting geven aan prognose en behandelingsstrategieën. Bovendien zijn de menselijke immuunreactie op virussen ook onderhevig aan actieve epigenetische regulatie, als gevolg van microbe-mens interacties. Deze feiten hebben ons ertoe gebracht de genetische en epigenetische factoren te onderzoeken die een invloed hebben op de manifestatie van infecties door pathogene virussen, waaronder HIV-1 en het coronavirus SARS-CoV-2.

Een van de obstakels voor de genezing van mensen met HIV-1 (PLHIV) is het



latente HIV-1 reservoir, d.w.z. immuuncellen in het lichaam die geïnfecteerd zijn met HIV maar niet actief nieuw HIV produceren. De behandeling met combinatie antiretrovirale therapie (cART) verhoogt de levenskwaliteit van PLHIV patiënten aanzienlijk door de HIV-1-virale belasting in het lichaam te onderdrukken. Recente studies hebben ook virale en gastheer karakteristieken onthuld die correleren met de grootte van het HIV-1 reservoir en de lange termijn dynamiek in PLHIV. In Hoofdstuk 4 hebben we PTDSS2, IRF7, RNH1, en DEAF1 geïdentificeerd als potentiële modifierende factoren van HIV-1 reservoirs door gastheer genetische factoren te associëren met HIV-1 reservoir eigenschappen, vertegenwoordigd door cel-geassocieerd (CA) HIV-1 DNA, CA HIV-1 RNA, en hun verhouding (RNA:DNA).

Vergeleken met eerdere studies hebben wij ons gericht op het CA HIV-1 DNA en RNA geïsoleerd uit CD4+ T cellen, een keuze die de invloed van cel proportie vermijdt. Van belang is dat onze studies suggereerden dat IRF7 een potentiële rol speelt bij het reguleren van HIV-1 latentie en recente studies toonden aan dat IRF7 de HIV onderdrukking bemiddelde als gevolg van de knock-out van CNOT (het RNA deadenylase complex). Bovendien zijn er aanwijzingen dat IRF7 betrokken is bij virale latentie-initialisatie en reactivatie uit een studie waarin chronische gamma-herpesvirus infecties in IRF7-deficiënte en wild-type muizen werden vergeleken. Onze studie had echter verschillende beperkingen: 1) kleine steekproefgrootte van het ontdekkings cohort (200), 2) een beperkte hoeveelheid covariaten opgenomen in de associatie analyse hoewel vele virale en gastheer kenmerken geassocieerd zijn met de HIV-1 reservoir kenmerken, 3) ons onderzoek was beperkt tot Kaukasische mensen, en 4) onze bevindingen zouden verder gerepliceerd moeten worden door gebruik te maken van een onafhankelijk cohort, dat momenteel niet beschikbaar is.

Individen die geïnfecteerd zijn met SARS-CoV-2 vertonen een breed scala aan klinische manifestaties en deze inter-individuele variaties die te wijten zouden kunnen zijn aan de verschillen in (epi)genetische architectuur tussen populaties. Dit moedigde ons aan om de (epi)genetische regulaties in COVID-19 patiënten met verschillende mate van ziekte-ernst te evalueren.

In Hoofdstuk 6 onthulden we single-cell multi-omics integratie resultaten van transcriptomische en epigenomische profielen (bv. chromatine toegankelijkheid) met betrekking tot verschillende mate van ziekte-ernst onder COVID-19 patiënten. De scRNA-seq resultaten toonden een aantal cel proportie veranderingen door de paarsgewijze vergelijkingen tussen verschillende mate van COVID-19 ziekte-ernst, wat overeenkomt met eerdere studies. Daarnaast, gebaseerd op onze scATAC-seq resultaten, stelden we vast dat C/EBPs, JUN/FOS, en SPI1 motieven oververtegenwoordigd waren in klassieke monocyten van patiënten. Bovendien suggereerde

onze piek-naar-gen koppelingsanalyse dat het lange niet-coderende RNA, LUCAT1, gemoduleerd werd door zowel SPI1 als C/EBPs transcriptiefactoren (TFs). Deze resultaten leidden ons tot de hypothese van een ontregelde cascade, waarin verhoogde C/EBP regulatie de LUCAT1 expressie versterkte en verder de interferon reacties op SARS-CoV-2 infectie onderdrukte, wat leidde tot ernstige COVID-19. We identificeren ook ASoC SNPs en ontdekten dat ze verrijkt waren in regulatorische regio's (d.w.z. promoters, enhancers en TFBSs). Verder onthulden onze resultaten de potentiële rol van CCR2 in de ernst van COVID-19 in de context van genetica en epigenetica.

De hierboven genoemde studie heeft echter meerdere beperkingen. Ten eerste werden er slechts beperkte symptoom categorieën in de studie opgenomen, namelijk ernstige, milde en herstellende individuen, terwijl COVID-19 patiënten een breed symptomatisch spectrum hebben, variërend van asymptoom tot dodelijk ernstig. Verder hebben we alleen bloedmonsters meegenomen, die minder representatief zouden kunnen zijn in vergelijking met primaire longweefsels, zoals epitheel. Bovendien, beperkt door de steekproefgrootte en sequencing diepte, valideerde onze studie slechts één COVID-19 GWAS loci, namelijk CCR2. Ten slotte, maar daarom niet minder belangrijk, richtten onze analyses zich uitsluitend op genetische, epigenetische en transcriptomische aspecten van de ziekten in eencellige resolutie, terwijl proteomics, metabolomics en metagenomics analyses prognose- en diagnose-markers aan het licht brachten. Daarom verwachten we meer bewijs uit andere omics data om meer inzicht te krijgen in de ziektemechanismen en om onze resultaten in de toekomst te valideren.

Celtype-specifieke quantitative trait loci (QTL) van de CCR5-hoeveelheid op het celoppervlakte

De CCR5-locus is een gevestigde associatie tussen de genetische architectuur van de gastheer en HIV-1-infecties, waarbij de CCR5d32 variant voorkomt in HIV-1-controlepersonen. Deze genetische locus is consequent geassocieerd met primaire resistentie tegen HIV-1-ziekten, wat een impuls heeft gegeven aan benaderingen voor de behandeling van AIDS, zoals transplantatie van met CCR5-genen bewerkte CD4+ T-cellen waarin CCR5-genen permanent disfunctioneel worden gemaakt. Uit recente studies is gebleken dat naast CCR5 zelf, ook andere genetische loci de expressie van CCR5 beïnvloeden. Bovendien zijn naast CD4+ T-cellen ook andere immuuncellen zoals CD8+ T-cellen en dendritische cellen mogelijk betrokken bij de reactie op de HIV-infectie. Er is echter minder bekend over het feit of specifieke genetische loci de hoeveelheid CCR5 op het celoppervlak in subsets van

immuuncellen (bv. lymfocyten) van PLHIV differentieel reguleren. In Hoofdstuk 5 hebben wij een schatting gemaakt van de celtype-specifieke associaties tussen genetische varianten (SNPs) en CCR5 niveaus in lymfocyten subsets, waarbij wij de gemiddelde fluorescentie intensiteit (MFI) van cel-oppervlak CCR5 of CCR5+ cel proportie (CP) gebruikten om de CCR5 hoeveelheid weer te geven. Onze resultaten onthulden genetische loci die geassocieerd zijn met de CCR5 hoeveelheid op het celoppervlak in 11 lymfocyt subsets van PLHIV.

Daarbij vonden we QTLs die geassocieerd waren met zowel MFI als CP in CD4+ geheugen regulatoire T (mTreg) cellen, wat suggereert dat de CCR5 overvloed in mTregs wordt gereguleerd door genetische factoren. Enerzijds zijn de mTregs afkomstig van pathogeen specifieke Tregs die geactiveerd worden en zich uitbreiden bij acute virale infecties *in vivo*. In reactie op secundaire infecties en ontstekingen, verdringen mTregs naïeve Tregs (nTregs) in migratie naar weefsels en activatie richting immunosuppressie, wat suggereert dat zij de inducerende doel celpopulaties kunnen zijn voor vaccinaties. Aan de andere kant identificeerden onze resultaten onafhankelijke genetische loci geassocieerd met MFI en CP in mTregs. De leidende SNP (rs11574435) voor MFI is een proxy voor de CCR5d32, terwijl de leidende SNP (rs60939770) voor CP een proxy is voor SNP rs1015164, die een rol speelt in de bescherming van de CCR5 mRNA door invloed uit te oefenen op de expressie van *CCR5AS*. Deze resultaten suggereren verschillende genetische architecturen tussen cel-oppervlakte CCR5 intensiteiten en proporties CCR5-positieve properties in mTregs. Diepgaand toekomstig werk is echter nodig om de onderliggende regulatorische programma's in HIV-1 infecties verder te ontleden.

Een aantal beperkingen belemmerden ons echter om de genetische factoren van de gastheer te begrijpen die de CCR5 oppervlakte-expressie beïnvloeden in subpopulaties van immuuncellen van PLHIV. Ten eerste werd slechts een beperkt aantal (200) PLHIV positieve mensen geïnccludeerd in de associatie-analyse, waardoor de statistische kracht ontbreekt om loci met een kleine effectgrootte te detecteren; maar we verwachten een groter cohort te kunnen onderzoeken in de nabije toekomst. Een andere belangrijke beperking was dat er geen ander onafhankelijk PLHIV-cohort werd geïnccludeerd met het oog op replicatie, wat naar verwachting zal worden opgelost door samenwerking met andere instellingen en inspanningen om in de komende jaren meer PLHIV-deelnemers te rekruteren. Bovendien zou het gebruik van één enkel monoklonaal anti-CCR5 antilichaam een onderschatting kunnen geven van CCR5 dat niet door het antilichaam werd herkend. In een studie waarin de antivirale activiteit van verschillende CCR5-antilichamen werd geëvalueerd, bleek het monoklonale antilichaam 2D7, dat wij ook in deze studie

hebben gebruikt, het meest effectief te zijn tegen HIV in vergelijking met twee andere anti-CCR5-antilichamen (45531 en CTC5).

Acknowledgements

Time flies and steals. Though I still remember my first footprint on the ground of Groningen at the spoor 3b of Groningen Hoofdstation, many details during the past four-year Ph.D. life are becoming vague. Here, I am very lucky to meet my supervisors who trained me to be a qualified scientific researcher and help me to be a better man. In this lovely city, I also met many people, made new friends, and learned new knowledge, which was definitely a colorful and fruitful four years. However, there is always a time-point at which I have to say goodbye to the past and move forward to face new challenges and life. To give a period mark for my Ph.D. life, I would like to give my gratitude and respect to people who helped and inspired me in my research and life by the following text.

Dear **Morris**, it is definitely a good memory to work with you. Our first conversation in room 9 at the Department of Genetics is still fresh, which is the starting point of my Ph.D. life in our department. You are really supportive and inspiring, which trained me as a qualified Ph.D. and helped me to be a better man. At the beginning of my Ph.D., your Dutch working style made me a little bit anxious. Gradually, I realized it is great to work with you as you are easy-going and gave me a lot of freedom to explore the project. Many thanks for your support, encouragement, and guidance to train me as a researcher, which surely will benefit my future research career.

Dear **Yang**, I was so lucky to be supervised by and involved in your group, which brought me into the field to study genetic determinants of infectious diseases. I believe these training activities are step-stones for me to become a better scientific worker. During my four-year Ph.D., we talked a lot about the projects and your wisdom suggestions often shed a light on the problems I encountered. You really set an outstanding example to me not only for performing scientific research but also for a way to balance life and work. Many gratitudes to you for your help, support, and suggestions in research and life. Looking forward to working with you in the future.

Dear **Joeri**, I still remember the first meeting with you at GCC corner and really appreciate your help and guidance which improved my ability in many aspects. You are humorous, which makes it relaxing to work with you. You also supported and inspired me a lot, which gave me a chance to develop myself to be a qualified scientist. Hope you have a fruitful output in academia and we will keep in touch in the future.

Dear **Richard**, I have to say it is really a pity you left us before my final promotion but I am very happy to know you are satisfied with your position. I very much appreciate your suggestions for my projects from a biologist's perspective, which improved my understanding of life science. I hope you have a successful career and fruitful outputs.

I also would like to express my appreciation to Prof. **A. van der Ven**, Prof. **M.J.T. Reinders**, and Prof. **B.J.L. Eggen** for being members of my reading committee. Thanks for reading and for your positive opinions on my thesis.

My special appreciation to Prof. **A. van der Ven**, thanks for involving me in the 200HIV projects and being the reading committee member to help improve my thesis. Your knowledge on HIV infection and immunology indeed inspired me to explore further into the infectious disease field. I look forward to working with you again and hope you have a fruitful career. I very much appreciate the collaboration with Jun-Prof. **L. N. Schulte**. It is a great pleasure to work with you on COVID-19 projects, in which I learned a lot from you who is a biologist specializing in lung studies.

To my paranymphs, **Shuang** and **Esteban**, it's really a luck to meet you two and thank you very much for being my paranymphs which I appreciate it very much! Dear, **Shuang**, I still remember our first meet at the pantry of our department with you, the talk helped me to be prepared for the coming working-time in our group. In the past years, we played boardgames and enjoyed dinner/lunch with our friends and colleagues many times. I wish you enjoy your new position and have a successful career! Dear, **Esteban**, we met at the pantry and know each other because of a cup of coffee. I will miss the days we played boardgames and enjoyed lunch and dinner together. Although the pandemic situation hampered our meal and game time, I really enjoyed the days we were Wednesday's eaters. My buddy, I wish you be fruitful for your research and be successful in your career. Finally, I really look forward to getting together with you two again. Wish you guys the best luck!

It is also great to work with you **Wim** and **Marc**. We teamed up and published a great work, which I believe is a tiny but important contribution to the academia of HIV research. Also, it was a happy time to collaborate with you **Jéssica** and **Louise**. We together finalized a publishable work and I hope it can improve our understanding of HIV infections. Dear **Bowen** and **Valerie**, we together wrapped up great work on COVID-19. Both of you are knowledgeable and innovative, it is really an amazing time to work with you both closely. I look forward to working

with you all again in the future.

My serious thanks to all co-authors for their inputs, supports, and helps to finalize each project in this thesis: Mariëlle E. van Gijn, Herbert T. Kruitbosch, Sofie Rutsaert, Linos Vandekerckhove, Wouter van der Heijden, Ezio T. Fok, Nadira Vadaq, Lisa van de Wijer, Hans J.P.M. Koenen, Musa Mhlanga, Mihai G. Netea, Michelle Aillaud, Hsin-Chieh Tsay, Anke R.M. Kraft, Chai Fen Soon, Ivan Odak, Berislav Bošnjak, Anna Vlot, Uwe Ohler, Robert Geffers, Thomas Illig, Reinhold Förster, and Markus Cornberg. I look forward to collaborating with you again.

My gratitude to the Department of Genetics at UMCG for providing research facilities and environments. I am grateful to Prof. Cisca Wijmenga, Prof. Alexandra (Sasha) Zhernakova, Prof. Jingyuan Fu, and Prof. Lude Franke for your suggestions and comments on my projects at Monday Lunch meeting, Tuesday Morning meeting, and 3GI meeting. Also thank my colleagues who are or were in the department, Niek, Freerk, Martijn, Gert-Jan, Pieter, Leenart, Patrick, Alexander, Tyler, Aron, Ranko, Olivier, Harm-Jan, Dylan, Robert, Roy, Annique, Monique, and Andriaan. Especially, thanks a lot, **Kate**, for revising and polishing the language of my thesis, which makes it more fluent and logical.

Many thanks to my colleagues for their help at **CiiM** in Hannover (DE). Cheng, Manoj, Wenchao, Zhaoli, Martijn, Javi, Tina (Alevtina), Jian-Bo, Ahmad, Saumya, Liang, Philip, Xuan, He, and Beate. I appreciate Cheng's suggestions and comments on projects included in the current thesis very much, which are valuable and helped to improve the quality of the thesis. Thanks, Manoj for your revision of my thesis, which improved its quality a lot, I appreciate it very much! Bowen, Wenchao, Zhaoli, Martijn, and Javi, I still remember our hanging around the Maschsee and enjoying the beers in the summer 2021. Looking forward to working with you in the May!

No doubts I can not become a Ph.D. candidate in RUG without your help, Dr. **Wanli Zheng**, thank you so much for involving me in the DPSD program. I wish you make more success in your career!

My special thanks to all my Chinese friends who are or were in Groningen, looking forward to meeting you again: Lu Zhou, Tao Yang, Fan Yang, Jing Zhao & Yu Li, Shandong Han, Zhiwen Wang, Yanfang Wang, Lisheng Zhang, Bohuan Lin, Minpeng Liang, Haibin Wen, Zhibo Li, Jingqi Chen, Haojie (Jason) Cao, Fangfang Liu, Chunxu Song, Xiaodong Feng, Tao Zhang, Fan Liu, Jiacong Wei, Xiaojing Chu, Cancan Qi, Shixian Hu, Yanni Li, Miaozen Huang, Bingjiang Qiu, Min Wang, Pei Meng, Daoming Wang, Yichen Liu, Lianmin Chen, Chan Li, Shiqiang Sun, Mengfei Cai, Yao Lin, Chunli Kong. I hope you all have a bright future, a successful career,

and fruitful outputs! 海内存知己，天涯若比邻。

To my be loved wife, **Xinrong**. 茫茫人海，弱水三千。你我相识，相知，相爱，相携。本欲作连理，相守相依，却相隔万里。故，我愿作春风，绕你左右，让温暖驱散冬寒。又，我愿作月光，伴你夜晚，让明亮盈满黑夜。此生有你，不离不弃。

To my parents. 父母在不远游，游必有方。生我养我育我，只怨人生短而树欲静风不止。文字太过苍白，写到这里，虽然思绪万千，但话到嘴边却只能说出谢谢两字。

张振华
22ed March 2022
Groningen, NL



Curriculum vitae

Zhenhua Zhang was born on the 1st February 1990 in Ji'nan (Shandong, China). He obtained his Bachelor's degree in Agronomy at China Agricultural University (CAU, Beijing, China) in 2014. Then, he was granted his Master's degree in Plant Pathology in the field of bioinformatics at CAU in 2017. In the same year, he was granted a fellowship by China Scholarship Council (CSC) and started his Ph.D. at the University of Groningen (UG) and University Medical Center Groningen (UMCG) in Groningen (NL). Under the (co-)supervisions by Prof. Dr. Morris A. Swertz, Prof. Dr. Yang Li, and Dr. K. Joeri van der Velde, he studied genetic determinants of human infectious diseases such as HIV-1 and SARS-CoV-2 infection using computational biology methods.



Publications

1. **Zhang, Zhenhua**, Freerk van Dijk, Niek de Klein, Mariëlle E. van Gijn, Lude H. Franke, Richard J. Sinke, Morris A. Swertz, and K. Joeri van der Velde. "Feasibility of predicting allele specific expression from DNA sequencing using machine learning." *Scientific reports* 11, no. 1 (2021): 1-11.
2. **Zhang, Zhenhua***, Wim Trypsteen*, Marc Blaauw, Xiaojing Chu, Sofie Rutsaert, Linos Vandekerckhove, Wouter van der Heijden et al. "IRF7 and RNH1 are modifying factors of HIV-1 reservoirs: a genome-wide association analysis." *BMC medicine* 19, no. 1 (2021): 1-17.
3. **Zhang, Zhenhua**, Herbert T. Kruitbosch, Richard J. Sinke, Morris A. Swertz, and K. Joeri van der Velde. "*ASdeep* uses a deep-learning neural network to attribute allelic expression to non-coding genetic variants" *To be submitted* (2022)
4. dos Santos, Jéssica C.*, **Zhenhua Zhang***, Louise E. van Eekeren, Ezio T. Fok, Nadira Vadaq, Lisa van de Wijer, et al. "Host genetic variants regulates CCR5 expression on immune cells: a study in people living with HIV and healthy controls." *To be submitted* (2022)
5. Zhang, Bowen*, **Zhenhua Zhang***, Valerie A. C. M. Koeken*, Saumya Kumar, Michelle Aillaud, Hsin-Chieh Tsay, Zhaoli Liu, et al. "An altered and allele-specific open chromatin landscape reveals epigenetic and genetic regulators of innate immunity in COVID-19." *To be submitted* (2022)
6. Yi, Yanglei*, **Zhenhua Zhang***, Fan Zhao, Huan Liu, Lijun Yu, Jiwei Zha, and Gaoxue Wang. "Probiotic potential of *Bacillus velezensis* JW: antimicrobial activity against fish pathogenic bacteria and immune enhancement effects on *Carassius auratus*." *Fish & shellfish immunology* 78 (2018): 322-330.
7. Zhang, Xinrong*, **Zhenhua Zhang***, and Xiao-Lin Chen. "The redox proteome of thiol proteins in the rice blast fungus *Magnaporthe oryzae*." *Frontiers in Microbiology* 12 (2021): 278.
8. Li, Yu, Jing Zhao, Spyridon Achinas, **Zhenhua Zhang**, Janneke Krooneman, and Gert Jan Willem Euverink. "The biomethanation of cow manure in a continuous anaerobic digester can be boosted via a bioaugmentation culture containing Bathyarchaeota." *Science of the Total Environment* 745 (2020): 141042.

-
9. Li, Yu, Jing Zhao, and **Zhenhua Zhang**. “Implementing metatranscriptomics to unveil the mechanism of bioaugmentation adopted in a continuous anaerobic process treating cow manure.” *Bioresource Technology* 330 (2021): 124962.
 10. Li, Yu, and **Zhenhua Zhang**. “Recognize the benefit of continuous anaerobic co-digestion of cow manure and sheep manure from the perspective of metabolic pathways as revealed by metatranscriptomics.” *Bioresource Technology Reports* 17 (2022): 100910.
 11. Aznaourova, Marina, Nils Schmerer, Harshavardhan Janga, **Zhenhua Zhang**, Kim Pauck, Judith Hoppe, Sarah M. Volkers et al. “Single cell RNA-seq uncovers the nuclear decoy lincRNA PIRAT as a regulator of systemic monocyte immunity during COVID-19.” *Under review in PNAS* (2022).
 12. van Eekeren, Louise E., Vasiliki Matzaraki, **Zhenhua Zhang**, Lisa Van De Wijer, Marc Blaauw, Marien I De Jonge, Linos Vandekerckhove et al. “People with HIV using long-term cART have higher percentages of circulating CCR5+ CD8+ T cells and lower percentages of CCR5+ regulatory T cells compared to healthy controls - a cross-sectional study.” *Under review* (2021)
 13. Yang, Fan, Hans Meerman, **Zhenhua Zhang**, Jianrong Jiang, André Faaij. “Integral techno-economic comparison and GHG balances of different production routes of aromatics from biomass with CCS.” *Submitted* (2022)
 14. Wang, Yanfang, **Zhenhua Zhang**, Diego Javier Jiménez, and Jan Dirk van Elsas. “Metatranscriptomes of wheat straw-degrading synthetic microbial consortia reveal differential expression patterns governed by local conditions.” *To be submitted* (2022)
 15. Wang, Min, Xiaofang Li, Francis M Cavallo, Harita Yedavally, Elisa J.M. Raineri, Elias Vera Murguia, Jeroen Kuipers, Anna Salvati, **Zhenhua Zhang** et al. “Functional profiling of CHAP domain-containing peptidoglycan hydrolases of *Staphylococcus aureus* USA300 uncovers potential targets for anti-staphylococcal therapies.” *To be submitted* (2022)

* Shared first authorship.