

Tangible heart, silicon brain

Citation for published version (APA):

Jin, H. (2021). *Tangible heart, silicon brain: computational modelling of cardiovascular diseases*. Ridderprint. <https://doi.org/10.26481/dis.20210921hj>

Document status and date:

Published: 01/01/2021

DOI:

[10.26481/dis.20210921hj](https://doi.org/10.26481/dis.20210921hj)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Chapter 9

Summary | Samenvatting | 总结

Summary

Cardiovascular diseases are one of the leading causes of death worldwide. The main underlying cause of cardiovascular diseases — atherosclerosis, is a chronic inflammatory disease, that implicates several complex biological processes, including lipid accumulation, inflammatory responses and fibrosis, in mutual interaction. To date, despite significant efforts that have been made towards a better understanding of atherogenesis, systematic and comprehensive modelling of atherosclerosis is still lacking. In this thesis, we deployed machine learning and computational-based data-driven methods on biological omics big-data to construct complex atherosclerotic systems and to clarify the disease mechanisms systematically. Moreover, through *in silico* analyses and modelling of cardiovascular diseases, we identified several new therapeutic targets for clinical intervention.

This thesis mostly focuses on carotid atherosclerosis, which is the main cause of ischemic stroke. In **chapter 2**, based on the transcriptional profiling of human carotid plaques, we constructed a gene co-expression network to unravel the major biological processes and pathways involved in the disease progression. The relationships between the associated pathways were subsequently described by a Bayesian inference network, highlighting the central role of T cell activity in human plaque. Furthermore, the scRNA-seq data exploration and construction of the gene regulatory network underpinned the important function of cytotoxic CD8⁺ T cells in plaque destabilization. Finally, through *in silico* screening of drugs in a publicly available drug database, we identified several epidermal growth factor receptor inhibitors (EGFRi) as top candidates to target T cell-associated gene program in haemorrhaged human plaque, which could guide the design of *in vitro* and *in vivo* experiments to test the suitability of EGFRi for the treatment of clinical atherosclerosis.

In **chapter 3**, we deployed a multi-omics machine learning prediction model integrating transcriptomics with proteomics/peptidomics to stratify high- from low-risk plaques. We compared the multi-omics with single-omics based models, and illustrated the advantages of truly integrative omics analysis in terms of accuracy, visibility and model interpretability. Our integrative multi-omics model revealed an SRF-driven gene/protein network associated with smooth muscle functions, controlling plaque instability, which could guide efforts for designing novel therapeutic strategies for plaque stabilization.

In addition, by integrating transcriptomics, proteomics, metabolomics and publicly available pathway databases, in **chapter 4** we reconstructed a genome-scale metabolic network for human plaque. Based on our knowledge, this is so far the most comprehensive metabolic atlas of human atherosclerotic plaque, providing

great details of plaque metabolism. Analysing this metabolic network for network-based topological centrality, we were able to pinpoint the pivotal role of macrophage glutamate/glutamine metabolism in plaque inflammation and vulnerability. This pathway could be a promising immunomodulation target for the prevention and treatment of atherosclerosis.

In **chapter 5**, we established a high-throughput platform to evaluate the functional changes of macrophages in the context of cardiovascular inflammation. This “functionomics” compendium of human macrophage in response to a variety of stimuli, was subsequently integrated with macrophage transcriptomics, showing macrophage plasticity at both functional and transcriptional levels. This compendium can serve as a reference for the future study of macrophage heterogeneity.

Using this MacroScreen platform, in **chapter 6** we evaluated the macrophage functional reprogramming by the systemic environment after acute myocardial infarction. In this chapter, machine learning models were deployed to stratify/distinguish the impact of AMI patients’ serum from control on macrophage function. We then integrated functionomics with network-organised transcriptomics of AMI serum exposed cells, and identified gene networks correlating with poor clinical prognosis. Finally, through *in silico* drug screening, we identified PGE2 signalling as a potential target for the prevention of adverse post-infarct remodelling.

In **chapter 7**, we showed that sex-specific differences in plaque phenotype contribute considerably to the recently reported transcriptional sex dimorphism of human atherosclerotic plaque. This information is key for future studies of the cardiovascular community to design sex-specific therapeutics for precision medicine.

Finally, in **chapter 8** we discussed the main findings and the potential perspectives of this thesis, as well as several conceptual questions for the application of *in silico* analyses in biomedical research.

In conclusion, by applying computational-based analyses, including statistical inference, machine learning, omics integration and network modelling on multiple omics data of human atherosclerosis and blood, we were able to model and analyse cardiovascular diseases at a molecular level, in a systematic manner. These computational and data-driven analyses provided novel insights into the disease mechanisms as well as into therapeutic targets and drugs for the treatment of atherosclerosis.

Samenvatting

Hart- en vaatziekten betreffen wereldwijd een van de belangrijkste doodsoorzaken. De belangrijkste onderliggende oorzaak van hart- en vaatziekten, aderverkalking of te wel atherosclerose, is een chronische ontstekingsziekte van middelgrote en grote slagaders, waarin verschillende complexe biologische processen, zoals vetopstapeling, bloedvatvorming, ontstekingsreacties en fibrose, betrokken zijn. Ondanks een sterk toegenomen begrip van dit ziektebeeld, ontbreekt het tot op heden aan een systematische analyse van de moleculaire processen en hun onderlinge samenhang die hieraan ten grondslag liggen. In dit proefschrift hebben we kunstmatige intelligentie (“machine-learning”) en computationele data-gedreven methodieken toegepast op expressieprofielen van de verkalkte vaatwand maar ook in cultuur gebrachte ontstekingscellen toegepast om het ziekteproces te modelleren. Ook hebben wij, door middel van *in silico* analyses en modellering van hart- en vaatziekten, verschillende nieuwe therapeutische markers geïdentificeerd voor klinische interventie.

Dit proefschrift richt zich vooral op aderverkalking (“plaque”) in de halsslagader, dat als gevolg van ruptuur kan leiden tot een beroerte.

In **hoofdstuk 2** hebben wij op basis van transcriptionele profilering van menselijke halsslagader endarteriëctomieën een gen co-expressie netwerk geconstrueerd om de belangrijkste biologische processen en paden van de ziekteprogressie te ontrafelen. De relatie tussen “biochemische routes” worden vervolgens beschreven via een Bayesiaans inferentienetwerk, waaruit een centrale rol van T-lymfocyt activiteit in de menselijke plaque naar voren kwam. Bovendien wezen zowel de expressedata als het afgeleide regulatoire netwerk op een belangrijke rol van cytotoxische CD8⁺ lymfocyten in de destabilisatie van de plaque. Tenslotte hebben we door middel van computergestuurde screening van geneesmiddelen uit een publiek beschikbaar databestand, epidermale groeifactor receptor remmer (EGFRi) geïdentificeerd als top-kandidaat voor interferentie met dit lymfocyt-geassocieerde netwerk in atherosclerose. Dit is een veelbelovend startpunt voor studies in proefdiermodellen naar de effectiviteit van EGFRi voor behandeling van klinische atherosclerose.

In **hoofdstuk 3** hebben we een voorspellend model ontwikkeld dat genexpressie profielen integreert met eiwit en peptideprofielen om hoog-risico plaques te kunnen onderscheiden van plaques met een laag risico. We vergeleken daarbij een model verkregen na integratie van deze datalagen met een model afgeleid van enkel genexpressieprofielen, en hebben de voordelen van eerstgenoemde in kaart gebracht (nauwkeurigheid, zichtbaarheid en interpreteerbaarheid). Uit ons integratieve model hebben we vervolgens een serum response factor-gedreven gen/eiwit netwerk geëxtraheerd. Dit netwerk correleerde met de aanwezigheid en

activiteit van gladde spierfuncties, die belangrijk zijn voor regulatie van de plaquestabiliteit. Deze studie zou op termijn kunnen leiden tot het ontwerp van nieuwe therapeutische strategieën voor plaquestabilisatie.

Door de integratie van gen-, eiwit-, en metaboliëtoprofielen van atherosclerotische plaque en combinatie met in de publieke ruimte beschikbare databestanden, hebben we in **hoofdstuk 4** een metabool netwerk voor menselijke plaque gereconstrueerd. Dit is, voor zover ons bekend, de meest uitgebreide metabole atlas van de humane atherosclerotische plaque, en laat een gedetailleerd beeld zien van het metabolisme in de plaque en haar relatie met de klinische prognose van atherosclerose. Uit een diepere analyse van dit metabole netwerk bleek een centrale rol van het glutamaat/glutamine metabolisme van macrofagen in plaque ontsteking en stabiliteit. Deze metabole route zou daarmee een veelbelovend doel kunnen zijn voor preventie en behandeling van atherosclerose.

In **hoofdstuk 5** hebben we een testplatform opgezet om de functionele veranderingen van macrofagen onder condities van hart- en vaatziekten te analyseren (de MacroScreen). "Functionomica" analyse middels dit platform resulteerde in een functioneel compendium van de reactie van humane ontstekingscellen (macrofagen) op een verscheidenheid aan stimulantia. Dit compendium werd vervolgens geïntegreerd met de corresponderende genexpressiedata van deze macrofagen blootgesteld aan dezelfde set van stimulantia. Hieruit bleek eens te meer de plasticiteit op zowel functioneel als expressie niveau van dit celtype. Dit compendium kan dienen als referentie voor de toekomstige studie van macrofaag heterogeniteit en voor screeningsdoeleinden

Met behulp van het MacroScreen platform hebben we vervolgens, in **hoofdstuk 6**, de functionele herprogrammering van macrofagen na blootstelling aan bloed van een patiënt met een acuut myocardinfarct bestudeerd. Ook hier werden machine learning modellen opgebouwd en toegepast om de impact van serum van patiënten met een hartinfarct op macrofaagfunctie te bepalen. Daaropvolgende integratie van de functionele en netwerk-georganiseerde genexpressiesignatuur van AMI serum blootgestelde cellen, leidde tot de identificatie van een genetisch netwerk dat correleerde met een ongunstige klinische prognose van deze patiënt. Tenslotte hebben we, door middel van *computergestuurde geneesmiddel* screening, een geneesmiddel geïdentificeerd dat dit netwerk mogelijk kan dempen, en dat dus mogelijk ingezet zou kunnen worden voor de preventie van ongunstige prognose na een hartinfarct.

In **hoofdstuk 7** hebben we laten zien dat sekse-specifieke verschillen in plaque morfologie in belangrijke mate bijdragen aan de recent gerapporteerde sekse gedreven genetische dimorfisme van humane atherosclerotische plaque. Deze informatie is essentieel voor toekomstige studies van de cardiovasculaire

gemeenschap om geslacht specifieke therapieën te ontwerpen voor precisie geneeskunde.

Tenslotte bespraken we in **hoofdstuk 8** de belangrijkste bevindingen en de mogelijke perspectieven van dit proefschrift, evenals een aantal conceptuele vragen voor de toepassing van *in silico* analyses in biomedisch onderzoek.

Samenvattend laten we in dit proefschrift zien dat toepassing van computergestuurde analyses, waaronder statistische inferentie, kunstmatige intelligentie, data integratie en netwerktheorie op uitgebreide en complex expressedatasets van de menselijke plaque en ontstekingscellen, ons in staat stelt om hart- en vaatziekten, op een systematische wijze, op moleculair niveau te modelleren en analyseren. Deze computationele en data-gedreven analyses leverden niet alleen nieuwe inzichten op in de ziektemechanismen maar leidde tevens tot nieuwe therapeutische doelwitten en geneesmiddelen voor behandeling van atherosclerose.

总结

心血管疾病是世界范围内人类的头号死因。引起心血管疾病主要原因——动脉粥样硬化，是一种慢性炎症性疾病。它的发生和发展牵涉到几个复杂的生物过程，包括脂质积累、炎症反应和纤维化等。这些生物过程之间相互影响和作用，共同推动着疾病的发展。尽管研究人员为更好地了解动脉粥样硬化疾病的机理做出了不懈努力，然而，迄今为止，我们仍然缺乏一个系统和全面的分子层面上的动脉粥样硬化模型来描绘该疾病的发展机制。在这篇论文中，我们将机器学习和基于计算和数据驱动的研究方法应用于生物组学大数据，进而构建复杂的动脉粥样硬化系统模型用于系统地阐明疾病机理。此外，通过对心血管疾病的计算分析和建模，我们还发现了几个新的临床干预治疗靶点。

这篇论文主要关注颈动脉粥样硬化，它的发生发展是导致缺血性中风的主要原因。在**第二章**中，基于人类颈动脉斑块的转录组分析，我们构建了一个基因共表达网络，以揭示参与疾病进展的主要生物过程和途径。随后，通过贝叶斯推理网络分析，我们刻画了相关生物过程之间的关系，并凸显了T细胞活动在人类动脉粥样硬化斑块中的核心作用。此外，基于单细胞转录组的数据分析和基因调控网络的构建揭示了细胞毒性 CD8⁺ T 细胞在斑块不稳定中的重要功能。最后，通过对公开的药物数据库中的药物进行计算筛选，我们确定了几种表皮生长因子受体抑制剂（EGFRi）作为针对出血性人类斑块中T细胞相关基因进行干预治疗的候选靶向药物。这项发现可以指导后续体外和体内实验的设计，以测试 EGFRi 是否适用于动脉硬化的临床治疗。

在**第三章**中，我们整合了转录组、蛋白质组和肽组数据并构建了一个多组学机器学习预测模型，从而实现对高风险和低风险斑块表型进行预测。我们将多组学模型与基于单一组学的机器学习模型进行了综合比较，并详细论述了多组学分析在预测准确性、结果可视化和模型可解释性方面的优势。通过解构该多组学模型，我们构建了一个与血管平滑肌功能相关的由血清应答因子（SRF）调控的基因/蛋白质网络，其在影响斑块不稳定性方面发挥着重要的作用。该发现可用于构建动脉粥样硬化的全新治疗方案，进而达到稳定动脉粥样硬化斑并预防斑块破裂的目的。

此外，通过整合转录组学、蛋白质组学、代谢组学和公开的生物通路数据库，我们在**第四章**中重建了人类斑块的基因层面的代谢网络。据我们所知，这是迄今为止最全面的人类动脉粥样硬化斑块的代谢图谱，提供着斑块代谢的丰富细节。通过分析这个代谢网络内部结点的拓扑中心性，我们能够锁定巨噬细胞中谷氨酸/谷氨酰胺代谢通路在斑块炎症和脆弱性中扮演的关键角色。这一新发现使该代谢通路有望成为预防和治疗动脉粥样硬化的一个免疫调节靶点。

在**第五章**中，我们建立了一个高通量平台（MacroScreen）并以此采集数据来评估心血管炎症背景下巨噬细胞的功能变化。通过将采集到的描述人类巨噬细胞应对各种刺激的功能性改变的数据集与巨噬细胞转录组数据进行联合分析，我们呈现出巨噬细胞在功能层面和转录层面应对不同环境和刺激下的可塑性。该数据以及平台可以作为未来研究巨噬细胞异质性的参考标准。

利用这个平台，在**第六章**中我们评估了急性心肌梗死患者发病后全身循环系统对血液中巨噬细胞功能的重塑。在这一章中，我们应用了机器学习模型来区分急性心肌梗塞患者的血清和对照组血清对巨噬细胞功能的影响。随后，我们将暴露于急性心肌梗塞患者血清的巨噬细胞的功能组学数据与网络建模后的转录组数据进行整合分析，进而发现了与不良临床预后相关联的基因调控网络。最后，通过网络建模驱动的计算药物筛选，我们发现前列腺素 E2 信号通路是干预心肌梗死不良预后发生的潜在目标。

在**第七章**中，我们证实动脉粥样硬化斑块表征的性别特异性差异对最新发现的人类动脉粥样硬化斑块的转录组下的性别二态性有着显著的贡献。这一信息对于未来心血管疾病的研究以及设计针对性别差异的精准医疗方案提供了非常关键的参考信息。

最后，在**第八章**中，我们探讨了本论文的主要发现和潜在的应用前景，以及在生物医学研究中应用计算机分析的几个概念性问题。

综上所述，通过应用计算驱动分析，包括统计推断、机器学习、多组学联合分析和网络建模等，于人类动脉粥样硬化组织和血液的多组学数据，我们以系统和全面的角度在分子水平上对心血管疾病进行了深度建模和分析。这些计算和数据驱动的分析方法为了解疾病机制，发现新的治疗靶点和药物以用于治疗动脉硬化提供了全新的见解和启示。