

Computational strategies in cardiometabolic diseases

Citation for published version (APA):

Lu, C. (2022). *Computational strategies in cardiometabolic diseases: a portal to deeper mechanistic understanding*. [Doctoral Thesis, Maastricht University]. Maastricht University. <https://doi.org/10.26481/dis.20221124cl>

Document status and date:

Published: 01/01/2022

DOI:

[10.26481/dis.20221124cl](https://doi.org/10.26481/dis.20221124cl)

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 7

Summary

Samenvatting

总结

Worldwide, cardiometabolic diseases (e.g., cardiovascular disease (CVD), diabetes, and non-alcoholic fatty liver disease (NAFLD)) have acquired almost epidemic proportions in the past few decades due to the widespread adoption of a western lifestyle. They compromise heart and liver functions and underlie the two main causes of death worldwide, ischemic heart disease and stroke. The development of these diseases is characterized by lipid accumulation, inflammatory responses, and metabolic dysfunction in the arterial wall (a process referred to as atherosclerosis) and liver (“fatty liver disease” or NAFLD). For both, macrophages and their precursors, monocytes, are important contributors.

The rapid development in high-throughput and imaging technologies have enabled researchers to detect profiles of genes, proteins, and metabolites within individuals and cells. These techniques are increasingly applied to study cardiometabolic diseases. However, their potential to explore the pathogenesis has not been fully exploited. In this thesis, we aim to analyse high-dimensional omics and imaging data of plaque and liver through a combination of computational strategies including statistical inference, machine learning and image processing, to dissect the pathogenesis of atherosclerosis and NAFLD at the molecular and spatial level.

Macrophages exhibit extreme plasticity, adopting a broad range of micro-environment-driven phenotypes. Exploring their heterogeneity throughout the course of atherosclerosis will be crucial to the discovery of new diagnostic and therapeutic measures. However, most of the commonly used techniques for quantifying cell markers cannot simultaneously measure sufficient features to properly describe the heterogeneity or do not provide spatial information at single-cell resolution. In **chapter 2**, we proposed a computational pipeline to integrate multiple imaging data layers from the same sample but derived from different platforms (i.e., multispectral imaging, MALDI mass spectrometry imaging (MALDI-MSI), and histologic imaging). Our approach enabled to identify cell phenotypes, pinpoint their locations in the tissue, reveal the cell community landscape, and characterize the metabolic environment of identified cell subsets. Our method demonstrated its effectiveness in analysing macrophage heterogeneity in atherosclerotic plaques but will also be applicable to other types of tissues, diseases and cells.

The metabolism of hepatocytes has been shown to be severely disrupted with the development of NAFLD, as metabolomic studies have shown. However, most metabolomics efforts that mapped the dynamics of these changes lacked spatial information. In **chapter 3**, we fused both supervised and unsupervised MSI image segmentation algorithms to deploy a liver compartment recognition model, resulting in the identification of three distinct liver compartments in the early stage of murine NAFLD, which were assigned “parenchyma”, “sinusoids” and “vessels”, based on their localisation and morphology. Metabolites were then clustered

according to their trajectories over the disease progression for the individual compartments. Our approach identified liver compartments solely based on the metabolite features from MALDI-MSI, indicating the presence of compartment-specific metabolic pathways, while avoiding the potential errors in aligning histological images with MALDI-MSI images, as well as extensive manual annotation on histological images. In addition, spatio-temporal analysis of mouse liver helps to monitor the metabolite dynamics during the disease progression within the separate compartments, leading to a deeper understanding of the processes that occur in the early stage of NAFLD.

Several risk factors have been identified, that increase the chance of developing cardiometabolic diseases, of which some are linked to an unhealthy lifestyle. However, it is unclear if these risk factors impact circulating monocytes, precursors of macrophages and thereby important players in many diseases. Therefore, in **chapters 4 and 5**, we aimed to identify monocyte key gene programs and pathways associated with the risk factors of CVD (which is mainly caused by atherosclerosis) by analysing the monocyte expression profiles of a CVD cohort collected by the Center for Translational Molecular Medicine (CTMM). We first studied disease-relevant sex differences in the transcriptional makeup of monocytic cells (**chapter 4**). For this purpose, we compared monocyte expression profiles of male and female CVD patient in the CTMM cohort, and generated sex-biased gene signatures. Then we adopted multiple enrichment analysis approaches based on several publicly available pathway, transcription factor (TF)-target interaction, and cytokine data resources, to infer pathway, TF, and cytokine signalling activities of male vs female monocytes. This led us to pinpoint female- and male-skewed cytokine signalling cascades in CVD. Validation of our findings in an independent healthy subject cohort showed that the observed sex differences were CVD-specific. In addition, we constructed a cytokine-TF network by inferring TF activities for the expression profiles of the identified sex-specific cytokine signalling pathways, which revealed Jun-B as central activator of most female-specific pathways. In **chapter 5**, we investigated the correlation between other CVD risk factors and the response capacity of CVD patient-derived monocytes to an inflammatory stimulus (lipopolysaccharide). As we show, CVD patients with high (diastolic) blood pressure tended to have a weaker monocyte inflammatory response. We then constructed a gene co-regulation and corresponding regulatory network and identified a sub-network with strong correlation with blood pressure. Moreover, this network pointed to a dysfunction in the cells' energy production (mitochondrial respiration). Finally, we pursued to a network-guided drug repurposing approach to identify iloprost as potential candidate to target this network. This drug, which is currently used to treat pulmonary hypertension, could potentially target the network that connects a weakened inflammatory response in monocytes (and potentially other cells) to diastolic hypertension.

In **chapter 6**, I summarised the main findings of my thesis, further explained the computational approaches used from **chapters 2 to 5** and explored potential application scenarios thereof in the study of cardiometabolic diseases, including their advantages and remaining limitations, and provided an outlook on future related areas.

Wereldwijd hebben cardiometabole aandoeningen, zoals aderverkalking, diabetes, suikerziekte en leververvetting en -ontsteking, de laatste decennia bijna epidemische proporties aangenomen, deels als gevolg van een westerse levensstijl. De verminderde hart- en leverfunctie waarmee zij gepaard gaan liggen ten grondslag aan de twee belangrijkste doodsoorzaken wereldwijd, ischemische hartziekten en beroerten. De ontwikkeling van deze ziekten wordt gekenmerkt door vetophoping, ontstekingsreacties en metabolische disfunctie in de slagaderwand (een proces dat atherosclerose wordt genoemd) en de lever (leververvetting of NAFLD). Macrofagen en hun voorlopercel, de monocyt, dragen voor beide aandoeningen belangrijk bij aan het ziekteproces .

Snelle ontwikkelingen op het gebied van high-throughput en beeldvormingstechnologieën hebben onderzoekers in staat gesteld om gen-, eiwit- en metabolietprofielen van weefsel en cellen snel en volledig in kaart te brengen. Deze technieken worden in toenemende mate toegepast om cardiometabole ziekten te bestuderen. Hun onderzoekspotentieel voor deze ziekten is echter nog niet volledig benut. In dit proefschrift willen we hoog-dimensionale omics en beeldvormingsdata van plaque en lever analyseren door gebruik te maken van een combinatie van computationele strategieën, waaronder statistische inferentie, machinaal leren en beeldverwerking, met als uiteindelijke doel de ontwikkeling van atherosclerose en NAFLD op moleculair en ruimtelijk niveau te ontrafelen.

Macrofagen vertonen een extreme plasticiteit en kunnen een breed scala aan functies aannemen afhankelijk van de locatie en de fase van het ziekteproces. Het onderzoeken van hun heterogeniteit zal van cruciaal belang zijn voor de ontdekking van nieuwe diagnostische en therapeutische mogelijkheden. Echter, de meest gebruikte technieken voor het kwantificeren van cel-markers kunnen niet voldoende kenmerken tegelijk meten om de cellulaire heterogeniteit voldoende in kaart te brengen, of geven geen ruimtelijke informatie over de cellen. In **hoofdstuk 2**, stelden we een integratie-pijplijn voor om meerdere lagen van beeldvorming, uitgevoerd op eenzelfde monster via verschillende microscopische en chemische analysetechnieken (multispectrale beeldvorming, MALDI massaspectrometrie beeldvorming (MALDI-MSI) en histologische beeldvorming), te integreren tot een ruimtelijk model van het zieke weefsel. Onze aanpak maakte het mogelijk om de verschijningsvorm, locatie en directe moleculaire (en cellulaire) omgeving van elke cel in het zieke weefsel in kaart te brengen. Onze methode is al succesvol toegepast bij het analyseren van macrofaag heterogeniteit in aderverkalking maar is zeker ook toepasbaar op andere types van weefsels, ziekten en cellen.

Het metabolisme van levercellen blijkt ernstig verstoord te zijn tijdens de ontwikkeling van NAFLD, zoals metabolomics studies al eerder hebben aangetoond. Echter, de meeste metabolomics studies waarin de dynamiek van deze veranderingen geanalyseerd is misten ruimtelijke informatie, terwijl MALDI-MSI

analyses van leverweefsel geen informatie verstrekten over de veranderingen doorheen het ziekteverloop. In **hoofdstuk 3** hebben we zowel gestuurde als ongestuurde MSI beeldbewerkingsalgoritmen toegepast om de diverse levercompartimenten te kunnen onderscheiden; dit resulteerde in de identificatie van drie verschillende compartimenten in muizenlever: "parenchym", "sinusoïde" en "bloedvat". Voor elk individueel compartiment werden vervolgens de metabolieten geclusterd op basis van hun expressieprofiel in de loop van de ziekteprogressie. Onze aanpak bleek in staat de belangrijke structurele compartimenten in muizenlever te identificeren, enkel op basis van hun metabolietprofiel en ongeacht het ziektestadium. Dit wijst niet alleen op de aanwezigheid van compartiment-specifieke metabolieten, maar maakt het ook mogelijk potentiële fouten in het uitlijnen van histologische beelden bij het integreren van verschillende datalagen te omzeilen. Bovendien stelde dit ons in staat om het ziekteproces voor elk compartiment afzonderlijk te volgen in de tijd. Dit laatste leidde tot een dieper inzicht in de metabole processen in het beginstadium van NAFLD.

De risicofactoren die bijdragen aan de ontwikkeling van cardiometabole ziekten zijn inmiddels grotendeels bekend en diverse factoren blijken daarbij verband te houden met een ongezonde levensstijl. Het is echter onduidelijk of deze risicofactoren ook invloed hebben op een ontstekingsceltype in het bloed, de monocyt, die betrokken is bij de ontwikkeling van de bovengenoemde cardiometabole aandoeningen. Daarom hebben wij in hoofdstuk 4 en 5 getracht de belangrijkste genprogramma's en signaaltransductieroutes te identificeren die in dit celtype correleren met risicofactoren voor hart- en vaatziekten. Hiervoor hebben we de beschikbare expressieprofielen geanalyseerd van monocytten uit een CVD-cohort dat verzameld is in het kader van het Center for Translational Molecular Medicine programma (CTMM). We bestudeerden daarbij voor het ziekteproces relevante man-vrouw verschillen in het expressieprofiel van monocytten (**hoofdstuk 4**) en waren in staat geslachts-specifieke genetische signaturen voor hart- en vaatziekten aan te tonen. Vervolgens hebben we meerdere analyses uitgevoerd op basis van openbaar beschikbare gegevensbronnen om de belangrijkste signaaltransductieroutes in dit celtype in kaart te brengen die kenmerkend waren voor mannelijke en vrouwelijke patiënten met hart- en vaatziekten. Validatie van onze bevindingen in een onafhankelijk cohort van gezonde proefpersonen toonde aan dat de waargenomen sekseverschillen inderdaad enkel golden voor hartpatiënten. Daarnaast hebben we een genregulatiernetwerk gebouwd door transcriptiefactor-activiteiten af te leiden uit de expressieprofielen van de geïdentificeerde sekse-specifieke cytokine-signaleringspaden. Jun-B kwam daarbij naar voren als centrale activator van de het gros van de vrouw-specifieke paden.

In **hoofdstuk 5** hebben we de correlatie onderzocht tussen andere risicofactoren en de reactie van monocytten in patiënten met hart- en vaatziekten op een

ontstekingsprikkel (lipopolysaccharide). We hebben aangetoond dat CVD-patiënten met een hoge (diastolische) bloeddruk een zwakkere ontstekingsreactie vertoonden. Vervolgens hebben we een gen (co)regulatiernetwerk geconstrueerd en daarbij een regulatiernetwerk geïdentificeerd met een sterk correlatie met hoge bloeddruk. Dit netwerk wees op een verminderd functioneren van de cellulaire energieproductie (mitochondriale ademhaling). Tenslotte hebben we een aantal geneesmiddelen kunnen selecteren, via een computeralgoritme, waarvan het aannemelijk is dat ze dit netwerk (en daarmee de bloeddruk) kunnen corrigeren. Een van deze geneesmiddelen, iloprost, wordt momenteel al gebruikt om pulmonale hypertensie te behandelen, maar zou dus deels via correctie van de verzwakte afweerreactie van monocytten kunnen werken.

In **hoofdstuk 6** heb ik de belangrijkste bevindingen van mijn proefschrift samengevat, de in de hoofdstukken 2 tot en met 5 gebruikte rekenkundige technieken en mogelijke toepassingsscenario's daarvan in het onderzoek naar cardiometabole ziekten nader toegelicht, de voordelen en tekortkomingen ervan besproken en ideeën voor verdere verfijning en verbetering van deze technologieën gepresenteerd.

由于不健康的生活方式的广泛流行，在过去的几十年里代谢性心血管病 (cardiometabolic diseases)，包括心血管疾病 (CVD)、糖尿病和非酒精性脂肪肝 (NAFLD)，已成为影响全球经济社会发展的重大公共卫生问题。这些疾病严重损害心脏和肝脏的功能，已成为全世界两个主要的死亡原因--缺血性心脏病和中风的首要致病因素。这些疾病的发展特点是脂质积累、炎症反应和动脉壁（被称为动脉粥样硬化的过程）和肝脏 (NAFLD) 的代谢功能紊乱。对于这两种情况，巨噬细胞和它们的前体单核细胞，都是重要的贡献者。

高通量和成像组学技术的快速发展使研究人员能够检测个人和细胞内的基因、蛋白质和代谢物的概况。然而，目前的技术在探索心脏代谢性疾病的发病机制方面的潜力还没有得到充分的开发。在这篇论文中，我们旨在通过统计推理、机器学习和图像处理等计算策略的组合来分析高维的全息影像数据，在分子水平上研究两种最常见的代谢性心血管病--心血管疾病(CVD)和非酒精性脂肪肝(NAFLD)的发病机理。

巨噬细胞表现出极强的可塑性，在动脉粥样硬化的过程中根据环境的不同而演化出一系列表型。探索它们的异质性对于发现新的诊断和治疗措施至关重要。然而，大多数量化细胞标志物的技术不能测量足够的特征来正确描述异质性，或者不能提供单细胞分辨率下细胞的位置信息。在第二章中，我们提出了一个综合计算管道，以整合来自同一样品但不同平台的多种成像数据（即多光谱成像、MALDI 质谱成像 (MALDI-MSI) 和组织学成像）。我们的方法能够识别细胞表型，准确定位它们在组织上的位置，提供一个细胞类型相互作用的视图，并描述它们的代谢环境。我们的方法不仅证明了其在分析动脉粥样硬化斑块中巨噬细胞异质性的有效性，而且也适用于其他类型的组织、疾病和细胞。

已有研究表明，随着非酒精性脂肪肝的发展，肝细胞的代谢会受到严重破坏。然而，许多关于肝脏组织 MALDI-MSI 的研究仍然缺乏对代谢物随疾病进展的动态变化的追踪以及对空间异质性的剖析。在第三章中，我融合了有监督和无监督的 MSI 图像分割算法，部署了一个肝脏分区识别模型，从而识别了非酒精性脂肪肝早期阶段的小鼠的三个不同的肝脏分区：实质、肝窦和血管。然后，根据各个区间的疾病进展轨迹对代谢物进行聚类。该方法仅根据 MALDI-MSI 的代谢物特征来确定肝脏分区，表明分区特定代谢物的存在，从而避免了将组织学图像与 MALDI-MSI 图像对齐的潜在错误，以及对组织学图像的大量人工注释。此外，小鼠肝脏的时空分析有助于识别疾病发展过程中不同区间内代谢物的动态变化，从而更深入地了解非酒精性脂肪肝的代谢紊乱。

7

尽管一些生物学特征和不健康的生活方式与心血管疾病风险的增加有关,但目前还不清楚这些因素如何影响心血管疾病中最重要的细胞类型之一:循环单核细胞的活动。因此,在第四章和第五章中,我们旨在通过分析转化分子医学中心(CTMM)收集的 CVD 队列的单核细胞表达谱,确定与 CVD 风险因素相关的单核细胞关键基因程序和途径。我们首先在第四章中研究了明显的性别差异在多大程度上反映了单核细胞中转录构成的差异。我们通过比较 CTMM 队列中的男性和女性 CVD 患者,产生了性别偏见的基因特征,然后根据一些公开的路径、转录因子-靶基因相互作用和细胞因子资源,采用多种富集分析方法推断路径、转录因子和细胞因子信号活动。这使我们确定了心血管疾病中女性和男性激活的途径和细胞因子。在一个独立的健康队列中的验证结果也显示,这些基于单核细胞转录谱推断的性别差异是心血管疾病特有的。此外,我们通过推断细胞因子诱导的转录谱的转录因子活性,构建了一个细胞因子-转录因子网络,该网络显示 *jun-B* 是大多数女性特定途径的激活剂。在第五章中,我们研究了心血管疾病风险因素与心血管疾病受试者单核细胞先天免疫反应能力之间的相关性,这使我们观察到高血压的 CVD 患者的单核细胞对 LPS 的反应较弱。然后,我们根据 CVD 患者的 LPS 反应谱构建了一个基因共调和相应的调控网络,发现与血压水平呈现最强相关性的子网络与氧化磷酸化途径和呼吸电子链运输有关。药物再利用锁定了伊洛前列腺素,一种用于治疗肺动脉高压的药物,能够增强 LPS 反应,这在一定程度上证实了我们对高血压和 LPS 反应之间关联的推断。

第六章总结了本论文的主要发现,进一步解释了第二章至第五章所使用的计算方法及其在心脏代谢性疾病研究中的应用场景,探讨了它们的优势和局限性,并对相关领域的未来进行了展望。