Multi-omics dissection of acquired and inherited heart failure

Jiayi Pei

Multi-omics dissection of acquired and inherited heart failure

University Medical Center Utrecht, Utrecht University, the Netherlands

PhD thesis

Cover design and layout: Jiayi Pei, quote taken from Jalaluddin Muhammad Balkhi.

Printed by: Proefshcriftmaken.nl

ISBN: 978-94-6419-027-4

Financial support by the Dutch Heart Foundation for the publication of this thesis is gratefully acknowledged.

Multi-omics dissection of acquired and inherited heart failure

Multi-omics ontleding van verworven en erfelijk hartfalen

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

donderdag 15 oktober 2020 des ochtends te 11.00 uur

door

Jiayi Pei

geboren op 21 augustus 1991 te Jiangxi, China

Promotoren:

Prof. dr. M.C. Verhaar
Prof. dr. F.W. Asselbergs

Copromotoren:

Dr. K.L. Cheng
Dr. M. Harakalová

Contents

Chapter 1

Introduction

The heterogeneity of heart failure

Heart failure (HF) is a cardiovascular disease with a high prevalence, severe morbidity, and high mortality.[1,2] Despite the enormous effort of the researchers in the cardiovascular field, it still remains challenging to tailor the optimal treatment for HF patients due to the complex nature of HF, including different clinical presentations, various underlying etiologies, and additional risk factors and modifiers. Based on the left ventricular ejection fraction (LVEF), HF can be divided into three types, namely HF with preserved ejection fraction (HFpEF, LVEF≥50%), HF with mid-range ejection fraction (HFmrEF, LVEF =40%-49%), and HF with reduced ejection fraction (HErEF, LVEF<40%).[3] Although HFpEF accounts for more than 50% of the HF population,[4,5] available treatments (i.e. β-blockers) fail to show promising effects in HFpEF patients when compared to their efficiency in treating HFrEF patients.[6] Common causes of HF are ischaemic heart disease, hypertension, valvular disease, and cardiomyopathies.[7] Within cardiomyopathies, which are defined as cardiac muscle disorders of known etiology (i.e. genetic mutations) or unknown etiology, there are several subtypes based on the ventricular morphology and physiology, such as hypertrophic cardiomyopathy (HCM), dilated cardiomyopathy (DCM), arrhythmogenic cardiomyopathy (ACM), etc.[8] This heterogeneity of cardiomyopathies further contributes to the complexity of HF. Therefore, it is critical to investigate the precise pathomechanism(s) and key regulators underlying each HF-causing etiology and/or risk factor in order to optimize the HF prevention and management.

HF induced by risk factors, such as hypertension, and comorbidities, are considered as acquired HF. One major risk factor of acquired HF is chronological aging. Age-related features, including reduced endurance capacity, weakened muscle strength, increased arterial stiffness, and declined mitochondrial function, are associated with HF.[9,10] In addition, prolonged exposure of the heart to deleterious stimuli, like hypertension and disrupted metabolism, also contributes to increased HF prevalence in the elderly.[10] At the molecular level, an increased accumulation of genetic mutations, which is accompanied by a decreased repair ability due to a decline in mitochondrial function, directly contributes to HF development upon aging.[10] Interestingly, HFpEF patients are generally older and often have hypertension when compared to HFrEF patients.[3,11] In line with this clinical observation, studies have already shown the mechanism of hypertension in triggering the hypertrophic responses in the myocardium, which subsequently leads to HF.[12] Within the elderly population, women have a higher risk of development HFpEF when compared to men, and women outnumber men by a 2:1 ratio in the HFpEF patient population, suggesting the sex-related effects on the development of HFpEF.[13,14] Studies have revealed multiple factors that contribute to the high prevalence of HFpEF in elderly women, such as the higher LVEF and the

predominant pattern of hypertrophic remodeling in women when compared to men.[15] Strikingly, estrogen deficiency shows a significant impact on these pathophysiological changes of HFpEF, possibly by affecting several mechanisms, including mitochondrial function, calcium hemostasis, renin-angiotensin-aldosterone system, and interstitial inflammation.[13] Higher estrogen level, on the other hand, shows the potential in attenuating cardiac hypertrophy.[16,17] Combined, these findings point out a putative important role of estrogen during HFpEF development, and the decline and/or loss of estrogen during and after menopause could contribute to the higher HFpEF prevalence in the elderly women.

Due to the bidirectional interaction between heart and kidney function, kidney disease is also a major risk factor of acquired HF. More than 50% of HF patients have kidney disease,[18,19], and ~20% of patients with chronic kidney diseases have the *de novo* HF.[20] Strikingly, around 80% of kidney patients who receive dialysis have HFpEF.[21] Kidney disease-induced hypertension is one key component that contributes to this tightly linked cardio-renal syndrome.[22,12] Additionally, the removal of uremic toxins is insufficient due to kidney dysfunction, resulting in the accumulation of uremic toxins circulating in the bloodstream, which has been shown in *in vitro* and animal studies to impair the (micro)vasculature and subsequently contribute to the HF development.[23,24] Uremic toxins impaired the endothelial function by triggering the coagulant response, inflammation, and oxidative stress.[25] Protein-bound uremic toxins that are poorly removed during hemodialysis, such as p-cresyl sulfate and indoxyl sulfate, exhibit significantly deleterious effects on the endothelium.[26] Therefore, uremic toxins-related microvascular diseases play a critical role in the heart and kidney interactions.[27] Although a systematic overview of the molecular and clinical perspectives have been presented in the past, reliable biomarkers and evident-based beneficial treatment for cardio-renal diseases are currently not available.[20]

Unlike the causes in acquired HF, inherited HF is a cardiac disorder associated with genetic mutations, which pass down from generation to generation. Although inherited HF is commonly divided into four subtypes, namely HCM, DCM, ACM, and restrictive cardiomyopathy,[28] it is important to note that there is no one-to-one genotype-to-phenotype association. For example, more than 1,500 pathogenic mutations have been identified in HCM, which is characterized by the thickened left ventricular wall with preserved LVEF (HFpEF).[29] The majority of these mutations occur in the cardiac β-myosin heavy chain (MYH7) and cardiac myosin binding protein-C (MYBPC3). Similarly, MYH7 and MYBPC3 mutations are presented in DCM, which is associated with impaired systolic function (HFrEF).[30,31] The PLN mutations are associated with DCM and ACM with a high risk for life-threatening ventricular arrhythmias.[32] Within the Netherlands, a rather large population of patients (>1,500) exists that carry the same founder PLN R14del mutation.[33]

PLN plays a key role in the calcium uptake and PLN R14del mutation inhibits the transportation of cytosolic calcium to the sarcoplasmic reticulum, which disturbs the cardiac contraction and relaxation and results in systolic dysfunction (HFrEF).[34] However, individuals with PLN R14del mutation may present as symptomatic and asymptomatic carriers, and the precise signaling that drives towards the symptomatic phenotype remains unclear.[33,35] Besides the complex genotype-to-phenotype association among different individuals, recent studies also demonstrate the cell-to-cell differences within the same individual, which further contribute to the variable phenotypes observed for the same mutation.[36,37] Next to genetic mutations, environmental factors, such as inflammation, toxins, and diet, may further modify the phenotypes in inherited HF.[38] Therefore, studies that aid in the understanding of the affected biological processes per mutation on the single-cell basis as well as the gene-environmental interactions are required.

Next-generation sequencing: From fundamental research to therapeutic strategies

Next-generation sequencing (NGS) techniques, such as whole-genome sequencing (WGS), chromatin immunoprecipitation and sequencing (ChIP-seq), and RNA sequencing (RNA-seq), have been widely used to study the molecular mechanisms underlying complex diseases and serve in clinical diagnostics.[39,40] Due to its broad scale and sensitive sequencing feature, NGS can provide comprehensive information on the genome-wide scale, even with limited materials.[41,42] The ChIP-seq technique can reveal the chromatin reorganization, which is a hall hallmark of the disease.[43,44] Changes in chromatin profile can, directly and indirectly, affect the transcription factor binding affinity, which subsequently alters the gene transcription and the biological signaling.[45] RNA-seq reveals the transcriptome landscape, which further illustrates the cellular responses to the disease.[46] Oncology studies using NGS have successfully identified multiple mutations and provide novel and direct targets for treatments.[47] Besides oncology, NGS has also identified autoimmune diseases-related mutations, and it further demonstrates the potential influences of environmental factors on genetic allele expression and subsequently biological signaling.[48]

Likewise, NGS is increasingly employed to study HF. A previous study using WGS identified novel pathogenic variants in HCM patients, not previously included in the genetic testing panel for the diagnosis of HCM.[49] Another study used RNA-seq and revealed a set of dysregulated long non-coding RNAs in HCM hearts, next to the known coding RNAs that are prevalent in HCM.[50] By pooling the transcriptional profiles obtained by RNA-seq from multiple DCM studies, Alimadadi and co-workers presented a set of genes that

participate in the deleterious biological processes in DCM.[51] Within the diseased hearts, Sweet and colleagues showed disease-specific gene expression patterns in DCM versus ischemic HF using RNA-seq, which subsequently elucidated the unique biological functions per disease etiology.[52] Next to these studies in the bulk myocardium, Gilsbach and colleagues showed a distinct epigenomic landscape between fetal, infant, and adult cardiomyocytes, which served as a unique dataset to understand the development and maturation of the cardiomyocytes and the heart.[53] By comparing failing and non-failing human cardiomyocytes, they further revealed a set of chromatin regions and genes that could be specific for HF. Another study used ChIP-seq coupled with different antibodies targeting either active DNA regulatory regions or poised DNA regulatory regions and showed considerable changes of chromatin activities in cardiomyocytes from acquired hypertrophic hearts.[54] Connecting the chromatin remodeling data to the transcriptome data using RNA-seq, the authors identified a set of genes and their enhancers that played key roles in the hypertrophic response.

In parallel, accumulating studies are investigating new therapeutic approaches, namely the use of epigenetic drugs, which are designed to target specific chromatin regions in treating HF.[55] Several studies have shown the potential of histone deacetylase inhibitors in attenuating the cardiac maladaptations by regulating pathways, such as inflammatory responses and autophagy.[56,57] Several oncology clinical trials have been launched to study the bromodomain and extraterminal domain (BET) inhibitors, which mediate the activation of transcription factors.[58] Recently, the clinical potential of BET inhibitors in HF starts to be recognized as well.[59,60] Besides the potential use of these drugs that target epigenetic modifications, RNA-seq has become a powerful tool to study drug repurposing in cardiac and non-cardiac disorders.[61,62,63]

Thesis outline

The primary aim of this thesis is to investigate the epigenetic and transcriptional changes in diseased hearts using multi-omics approaches (**Figure 1**), enabling us to better understand the pathological mechanisms underlying the diseases. The first part of the thesis focuses on acquired heart failure due to several common causes, including kidney disease, aging, gender, and aortic stenosis-induced hypertension. The second part of the thesis focuses on inherited heart failure due to multiple mutations. Both hypertrophic and dilated cardiomyopathies are addressed, followed by a set of promising candidates and their regulated biological functions per disease.

Since kidney dysfunction, female gender, and aging are three key risk factors of acquired heart failure, in **Chapter 2** we studied the influence of estrogen on cardiorenal syndrome. Particularly, we presented an overview of the biological signaling of estrogen on the microvasculature, which plays a central role in cardiac and kidney interactions. In **Chapter 3** we investigated the deleterious effect of accumulated uremic toxin (indoxyl sulfate) due to kidney dysfunction on the endothelium using RNA-seq. We obtained the transcriptome changes in endothelial cells exposed to the toxin and studied their affected pathways. In **Chapter 4** we revealed the histone modification in remodeled non-failing human hearts due to aortic stenosis-induced hypertension using ChIP-seq. Connecting these chromatin changes to the gene expression changes obtained by RNA-seq, we presented a list of transcription factors that could pose a dominant influence on myocardial remodeling. In **Chapter 5** we employed ChIP-seq, RNA-seq, and proteomics to investigate HCM hearts carrying MYBPC3 mutations. Using two independent analytical methods to process the multi-omics data, we obtained a set of candidates that were driving the biological processes in HCM hearts. **Chapter 6** presents the database of the altered protein profiles in HCM *versus* control hearts. Moreover, by comparing HCM hearts with and without known mutations, we identified mutation-specific changes and investigated the affected signaling in disease-mimicking mouse models.

In **Chapter 7** we examined the association between 52 QRS-related loci and altered DNA regulatory regions in HCM *versus* control hearts. We performed a comprehensive *in silico* functional annotations and identified a novel set of potentially disease-causing genes. In **Chapter 8** we first studied the epigenomic and transcriptional changes in end-stage DCM hearts carrying PLN R14del mutation. From the promising candidates obtained at the tissue level, we further investigated their mRNA and protein expression patterns in human induced pluripotent stem cell-derived cardiomyocytes with and without PLN R14del mutation. **Chapter 9** provides a general summary of the key findings in the previous chapters and discusses their potentials with future therapeutic and diagnostic directions.

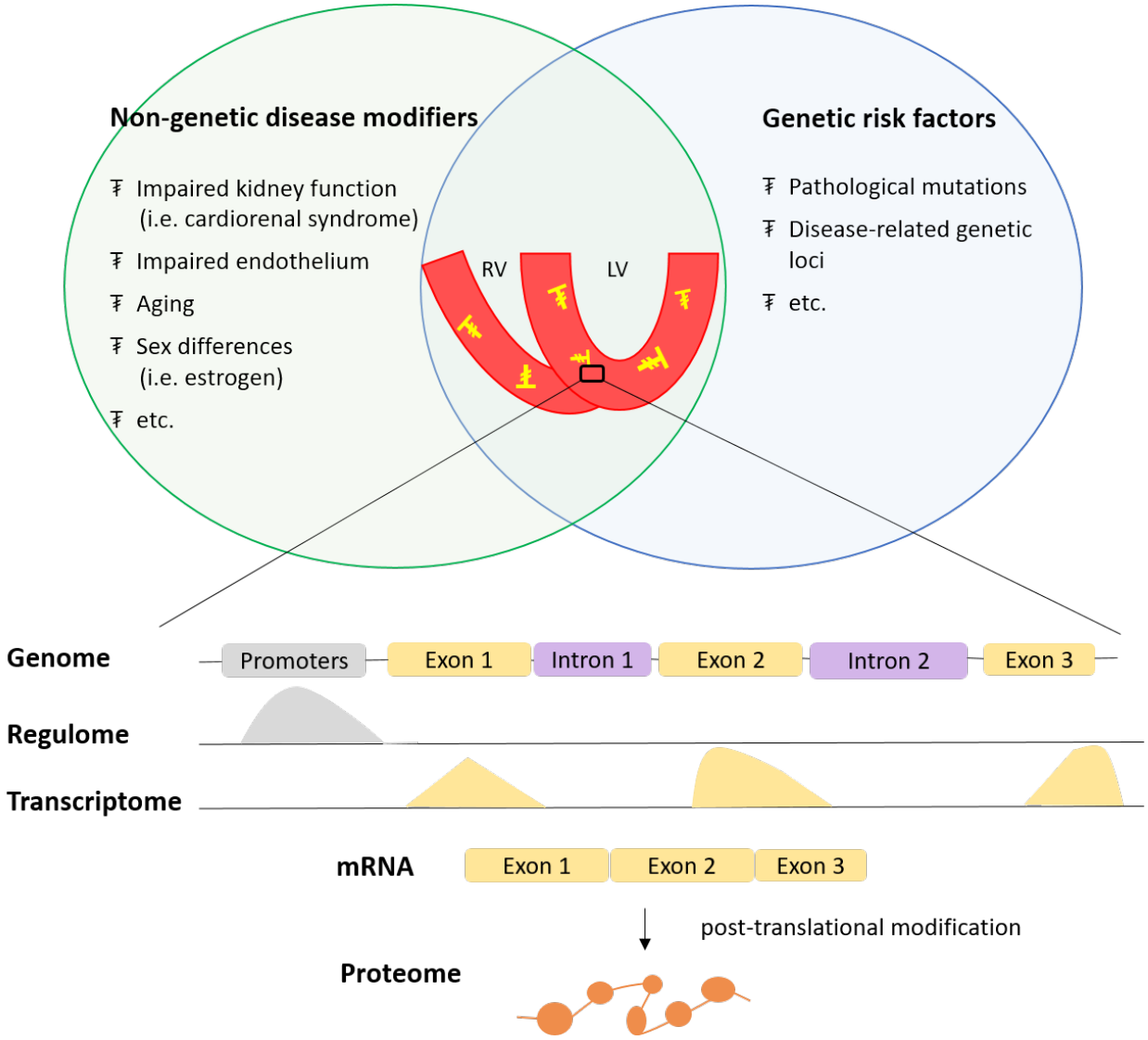


Figure 1. Genetic risk factors and/or non-genetic disease modifiers of heart failure can interfere with the cellular genome, regulome, transcriptome, and/or proteome, which subsequently affect cardiac function. LV: left ventricle; RV: right ventricle.

References

1. Heart failure: Overview. in *InformedHealth.org [Internet]* (Institute for Quality and Efficiency in Health Care (IQWiG), 2018).
2. Boback Ziaeian, G. C. F. Epidemiology and aetiology of heart failure. *Nat. Rev. Cardiol.* **13**, 368 (2016).
3. Choi, H.-M., Park, M.-S. & Youn, J.-C. Update on heart failure management and future directions. *Korean J. Intern. Med.* **34**, 11 (2019).
4. pubmeddev & Dunlay SM, E. al. Epidemiology of heart failure with preserved ejection fraction. - PubMed - NCBI. https://www.ncbi.nlm.nih.gov/pubmed/28492288.
5. Pfeffer, M. A., Shah, A. M. & Borlaug, B. A. Heart Failure With Preserved Ejection Fraction In Perspective. *Circ. Res.* **124**, (2019).
6. Yoon, S. & Eom, G. H. Heart failure with preserved ejection fraction: present status and future directions. *Exp. Mol. Med.* **51**, 1–9 (2019).
7. Arend Mosterd, A. W. H. Clinical epidemiology of heart failure. *Heart* **93**, 1137 (2007).
8. Elliott, P. *et al.* Classification of the cardiomyopathies: a position statement from the european society of cardiology working group on myocardial and pericardial diseases. *Eur. Heart J.* **29**, 270–276 (2008).
9. Wagner, J. *et al.* Functional aging in health and heart failure: the COmPLETE Study. *BMC Cardiovasc. Disord.* **19**, 1–17 (2019).
10. Li, H. *et al.* Targeting Age-Related Pathways in Heart Failure. *Circ. Res.* **126**, (2020).
11. Kazzam, E., Ghurbana, B. A., Obineche, E. N. & Nicholls, M. G. Hypertension — still an important cause of heart failure? *J. Hum. Hypertens.* **19**, 267–275 (2005).
12. Messerli, F. H., Rimoldi, S. F. & Bangalore, S. The Transition From Hypertension to Heart Failure. *J Am Coll Cardiol HF* **5**, 543–551 (2017).
13. Tadic, M. *et al.* Sex and Heart Failure with Preserved Ejection Fraction: From Pathophysiology to Clinical Studies. *J. Clin. Med. Res.* **8**, (2019).
14. Scantlebury, D. C. & Borlaug, B. A. Why Are Women More Likely Than Men to Develop Heart Failure With Preserved Ejection Fraction? *Curr. Opin. Cardiol.* **26**, (2011).
15. Bharathi Upadhya, D. W. K. Heart Failure With Preserved Ejection Fraction in Older Adult. *Heart Fail. Clin.* **13**, 485 (2017).
16. Xu, Y., Arenas, I. A., Armstrong, S. J. & Davidge, S. T. Estrogen modulation of left ventricular remodeling in the aged heart. *Cardiovasc. Res.* **57**, 388–394 (2003).
17. Iorga, A. *et al.* The protective role of estrogen and estrogen receptors in cardiovascular disease and the controversial use of estrogen therapy. *Biol. Sex Differ.* **8**, (2017).
18. Löfman, I. *et al.* Prevalence and prognostic impact of kidney disease on heart failure patients. *Open Heart* **3**, e000324 (2016).
19. Kevin Damman, J. M. T. The kidney in heart failure: an update. *Eur. Heart J.* **36**, 1437 (2015).
20. Heart failure in chronic kidney disease: conclusions from a Kidney Disease: Improving Global Outcomes (KDIGO) Controversies Conference. *Kidney Int.* **95**, 1304–1317 (2019).

21. Berl, T. & Henrich, W. Kidney-Heart Interactions: Epidemiology, Pathogenesis, and Treatment. *Clin. J. Am. Soc. Nephrol.* **1**, 8–18 (2006).

22. Middleton, J. P. & Pun, P. H. Hypertension, chronic kidney disease, and the development of cardiovascular risk: a joint primacy. *Kidney International* vol. 77 753–755 (2010).

23. Davenport, A. Role of dialysis technology in the removal of uremic toxins. *Hemodialysis International* vol. 15 S49–S53 (2011).

24. Fujii, H., Goto, S. & Fukagawa, M. Role of Uremic Toxins for Kidney, Cardiovascular, and Bone Dysfunction. *Toxins* **10**, (2018).

25. Jourde-Chiche, N. *et al.* Endothelium structure and function in kidney health and disease. *Nat. Rev. Nephrol.* **15**, 87–108 (2019).

26. Jourde-Chiche, N., Dou, L., Cerini, C., Dignat-George, F. & Brunet, P. Vascular Incompetence in Dialysis Patients--Protein-Bound Uremic Toxins and Endothelial Dysfunction. *Semin. Dial.* **24**, (2011).

27. Zhang, J., Bottiglieri, T. & McCullough, P. A. The Central Role of Endothelial Dysfunction in Cardiorenal Syndrome. *Cardiorenal Med.* **7**, 104 (2017).

28. Daniel Jacoby, W. J. M. Genetics of inherited cardiomyopathy. *Eur. Heart J.* **33**, 296 (2012).

29. Viswanathan, S. K. *et al.* Hypertrophic cardiomyopathy clinical phenotype is independent of gene mutation and mutation dosage. *PLoS One* **12**, e0187948 (2017).

30. Elizabeth M McNally, L. M. Dilated cardiomyopathy: genetic determinants and mechanisms. *Circ. Res.* **121**, 731 (2017).

31. Ferro, M. D., Severini, G. M., Gigli, M., Mestroni, L. & Sinagra, G. Genetics of Dilated Cardiomyopathy: Current Knowledge and Future Perspectives. in *Dilated Cardiomyopathy: From Genetics to Clinical Management [Internet]* (Springer, 2019).

32. Hof, I. E. *et al.* Prevalence and cardiac phenotype of patients with a phospholamban mutation. *Neth. Heart J.* **27**, 64–69 (2018).

33. van der Zwaag, P. A. *et al.* Phospholamban R14del mutation in patients diagnosed with dilated cardiomyopathy or arrhythmogenic right ventricular cardiomyopathy: evidence supporting the concept of arrhythmogenic cardiomyopathy. *Eur. J. Heart Fail.* **14**, 1199–1207 (2012).

34. van Opbergen, C. J. M., Delmar, M. & van Veen, T. A. B. Potential new mechanisms of pro-arrhythmia in arrhythmogenic cardiomyopathy: focus on calcium sensitive pathways. *Neth. Heart J.* **25**, 157–169 (2017).

35. van Rijsingen, I. A. W. *et al.* Outcome in phospholamban R14del carriers: results of a large multicentre cohort study. *Circ. Cardiovasc. Genet.* **7**, 455–465 (2014).

36. Montag, J. *et al.* Burst-Like Transcription of Mutant and Wildtype MYH7-Alleles as Possible Origin of Cell-to-Cell Contractile Imbalance in Hypertrophic Cardiomyopathy. *Front. Physiol.* **9**, (2018).

37. Parbhudayal, R. Y. *et al.* Variable Cardiac Myosin Binding protein-C Expression in the Myofilaments Due to MYBPC3 Mutations in Hypertrophic Cardiomyopathy. *J. Mol. Cell. Cardiol.* **123**, (2018).

38. Teekakirikul, P., Zhu, W., Huang, H. C. & Fung, E. Hypertrophic Cardiomyopathy: An Overview of Genetics and Management. *Biomolecules* **9**, (2019).

39. Donlin, L. T. *et al.* Insights into rheumatic diseases from next-generation sequencing. *Nat. Rev. Rheumatol.* **15**, 327–339 (2019).

40. Yska, H. A. F. *et al.* Diagnostic Yield of Next Generation Sequencing in Genetically Undiagnosed Patients with Primary Immunodeficiencies: a Systematic Review. *J. Clin. Immunol.* **39**, 577–591 (2019).

41. Koboldt, D. C., Steinberg, K. M., Larson, D. E., Wilson, R. K. & Mardis, E. The Next-Generation Sequencing Revolution and Its Impact on Genomics. *Cell* **155**, 27 (2013).

42. Serratì, S. *et al.* Next-generation sequencing: advances and applications in cancer diagnosis. *Onco. Targets. Ther.* **9**, 7355 (2016).

43. Constanze, B. & Cockerill, P. N. Chromatin Mechanisms Regulating Gene Expression In Health And Disease. in *Madame Curie Bioscience Database [Internet]* (Landes Bioscience, 2013).

44. Johnson, C. A. Chromatin modification and disease. *J. Med. Genet.* **37**, 905–915 (2000).

45. Milne, T. A., Zhao, K. & Hess, J. L. Chromatin Immunoprecipitation (ChIP) for Analysis of Histone Modifications and Chromatin-Associated Proteins. *Methods Mol. Biol.* **538**, 409 (2009).

46. Al-Dalahmah, O. *et al.* Single-nucleus RNA-seq identifies Huntington disease astrocyte states. *Acta Neuropathologica Communications* **8**, 1–21 (2020).

47. Azim, F. S., Houri, H., Ghalavand, Z. & Nikmanesh, B. Next Generation Sequencing in Clinical Oncology: Applications, Challenges and Promises: A Review Article. *Iran. J. Public Health* **47**, 1453 (2018).

48. Applications of Next-generation Sequencing in Systemic Autoimmune Diseases. *Genomics Proteomics Bioinformatics* **13**, 242–249 (2015).

49. Bagnall, R. D. *et al.* Whole Genome Sequencing Improves Outcomes of Genetic Testing in Patients With Hypertrophic Cardiomyopathy. *J. Am. Coll. Cardiol.* **72**, 419–429 (2018).

50. Liu, X. *et al.* Long non-coding and coding RNA profiling using strand-specific RNA-seq in human hypertrophic cardiomyopathy. *Scientific Data* **6**, 1–7 (2019).

51. Alimadadi, A., Munroe, P. B., Joe, B. & Cheng, X. Meta-Analysis of Dilated Cardiomyopathy Using Cardiac RNA-Seq Transcriptomic Datasets. *Genes* **11**, (2020).

52. Sweet, M. E. *et al.* Transcriptome analysis of human heart failure reveals dysregulated cell adhesion in dilated cardiomyopathy and activated immune pathways in ischemic heart failure. *BMC Genomics* **19**, 1–14 (2018).

53. Gilsbach, R. *et al.* Distinct epigenetic programs regulate cardiac myocyte development and disease in the human heart in vivo. *Nat. Commun.* **9**, 1–14 (2018).

54. Papait, R. *et al.* Genome-wide analysis of histone marks identifying an epigenetic signature of promoters and enhancers underlying cardiac hypertrophy. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 20164–20169 (2013).

55. Liu, C.-F. & Wilson Tang, W. H. Epigenetics in Cardiac Hypertrophy and Heart Failure. *J Am Coll Cardiol Basic Trans Science* **4**, 976–993 (2019).

56. McKinsey, T. A. Targeting Inflammation in Heart Failure With Histone Deacetylase Inhibitors. *Mol. Med.* **17**, (2011).

57. Cao, D. J. *et al.* Histone deacetylase (HDAC) inhibitors attenuate cardiac hypertrophy by suppressing autophagy. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 4123–4128 (2011).

58. Cochran, A. G., Conery, A. R. & Sims, R. J. Bromodomains: a new target class for drug development. *Nat. Rev. Drug Discov.* **18**, 609–628 (2019).

59. Duan, Q. *et al.* BET bromodomain inhibition suppresses innate inflammatory and profibrotic transcriptional networks in heart failure. *Sci. Transl. Med.* **9**, (2017).

60. Cully, M. BET inhibitor attenuates heart failure. *Nat. Rev. Drug Discov.* **16**, 453–453 (2017).

61. Yang, G., Ma, A. & Qin, Z. S. An Integrated System Biology Approach Yields Drug Repositioning Candidates for the Treatment of Heart Failure. *Front. Genet.* **10**, (2019).

62. Sun, J., Yang, J., Chi, J., Ding, X. & Lv, N. Identification of drug repurposing candidates based on a miRNA-mediated drug and pathway network for cardiac hypertrophy and acute myocardial infarction. *Hum. Genomics* **12**, (2018).

63. Yang, X. *et al.* High-Throughput Transcriptome Profiling in Drug and Biomarker Discovery. *Front. Genet.* **11**, (2020).

Chapter 2

Cardiorenal disease connection during post-menopause: The protective role of estrogen in uremic toxins induced microvascular dysfunction

Jiayi Pei1*, Magdalena Harakalova2,5*, Hester den Ruijter4, Gerard Pasterkamp4, Dirk J. Duncker3, Marianne C. Verhaar1, Folkert W. Asselbergs2,5,6‡, Caroline Cheng1,3‡

*Equal contribution of first authors, ‡equal contribution of last authors

[1]Department of Nephrology and Hypertension, Division of Internal Medicine and Dermatology, University Medical Center Utrecht, The Netherlands
[2]Department of Cardiology, University Medical Center Utrecht, The Netherlands
[3]Experimental Cardiology, Department of Cardiology, Thoraxcenter Erasmus University Medical Center, Rotterdam, The Netherlands
[4]Experimental Cardiology Laboratory, Department of Experimental Cardiology, University Medical Center Utrecht, The Netherlands
[5]Netherland Heart Institute, Utrecht, The Netherlands
[6]Institute of Cardiovascular Science, University College London, United Kingdom

International Journal of Cardiology

DOI: 10.1016/j.ijcard.2017.03.050

Abstract

Female gender, post-menopause, chronic kidney disease (CKD) and (CKD linked) microvascular disease are important risk factors for developing heart failure with preserved ejection fraction (HFpEF). Enhancing our understanding of the interrelation between these risk factors could greatly benefit the identification of new drug targets for future therapy. This review discusses the evidence for the protective role of estradiol (E2) in CKD-associated microvascular disease and related HFpEF. Elevated circulating levels of uremic toxins (UTs) during CKD may act in synergy with hormonal changes during postmenopause and could lead to coronary microvascular endothelial dysfunction in HFpEF. To elucidate the molecular mechanism involved, published transcriptome datasets of indoxyl sulfate (IS), high inorganic phosphate (HP) or E2 treated human derived endothelial cells from the NCBI Gene Expression Omnibus database were analyzed. In total, 36 genes overlapped in both IS- and HP-activated gene sets, 188 genes were increased by UTs (HP and/or IS) and decreased by E2, and 572 genes were decreased by UTs and increased by E2. Based on a comprehensive in silico analysis and literature studies of collected gene sets, we conclude that CKD-accumulated UTs could negatively impact renal and cardiac endothelial homeostasis by triggering extensive inflammatory responses and initiating dysregulation of angiogenesis. E2 may protect (myo)endothelium by inhibiting UTs-induced inflammation and ameliorating UTs-related uremic bleeding and thrombotic diathesis via restored coagulation capacity and hemostasis in injured vessels.

Key words: cardiorenal syndrome; menopause; microvascular dysfunction; estrogen; indoxyl sulfate; high phosphate.

1. Complex interrelationship between chronic kidney disease and heart failure with preserved ejection fraction

Heart failure (HF) is a growing major public health problem that affects ~2% of the western population [1]. It has two main subtypes, HF with reduced ejection fraction (HFrEF) and HF with preserved ejection fraction (HFpEF). Within the HF population, more than 50% suffer from HFpEF [2]. Interestingly, chronic kidney disease (CKD) occurs in 26% to 53% of the HFpEF population and the subclinical diastolic dysfunction appears to be the most common echocardiographic feature in asymptomatic CKD patients on hemodialysis, suggesting a strong link between CKD and HFpEF [3, 4]. Furthermore, clinical studies showed a linear relationship between the progression of CKD and the worsening of longitudinal function of the left ventricle in the HFpEF population [5]. The cardiac parameters in patients with CKD stage 2 and 3 already resemble early HFpEF, and the cardiac mechanics have been reported to become worse in patients with CKD stage 4 and 5. In a large cohort study on the development of heart dysfunction during 11 years of follow-up, Brouwers and colleagues demonstrated that increased urinary albumin excretion and cystatin C were more associated with the onset of HFpEF when compared to HFrEF [6]. In particular, older females with increased urinary albumin excretion or cystatin C were more vulnerable to develop HFpEF. These findings indicate a clear association between CKD and HFpEF, especially in the elderly female population.

Important findings in the field further prove that impaired renal function is a major risk for developing HFpEF [7]. Although several mechanisms underlying how CKD contribute to HF in general have been well established, including increased inflammatory responses and activated neurohomronal pathways [8], studies on the driving mechanisms on CKD-related HFpEF are limited. A recent publication by Paulus et al. proposed a disease mechanism in which renal dysfunction caused systemic changes in circulating factors that activated inflammation and led to microvascular disease (MVD), cardiomyocyte stiffening, and a hypertrophic response [7]. MVD often occur ubiquitously throughout the body in patients with cardiovascular disease [9]. Paulus and co-workers further proposed that identified metabolic syndrome linked cardiovascular comorbidities, such as diabetes and obesity, could act as inducers of systemic inflammation that trigger global and coronary endothelial dysfunction, leading to myocardial hypertrophy, impaired myocardial relaxation and increased myocardial stiffness [10]. In relation to cardiovascular disease in general, endothelial dysfunction is well known to be able to serve as a strong predictor of coronary artery disease onset [9]. CKD, a common cardiovascular comorbidity associated with metabolic risk factors, leads to hyperphosphatemia and accumulation of uremic toxins (UTs) that trigger inflammation and MVD, which could subsequently contribute to HFpEF onset and progression [11].

Early stage of kidney disease can already be detected by the presence of proteinuria, and a continuous retention of UTs results in a toxic circulatory environment due to the reduction of glomerular filtration rate during the progression of CKD [12]. Serum levels of some uremic toxin compounds, such as methionine sulfoxide and hydroxyproline, accumulate significantly as glomerular filtration rate declines and have been proposed as markers for detecting early stage of CKD [13]. Most UTs circulate in the bloodstream in albumin bound form, and they are not able to directly pass an intact endothelial barrier [14, 15]. However, alterations in endothelial barrier do occur in response to various inflammatory mediators and atherogenic metabolic particles [16]. Many UTs, such as indoxyl sulfate and p-cresyl sulfate, have been shown to compromise the endothelial barrier function [17], which could promote protein leakage, causing direct exposure of surrounding non-vascular cells like cardiomyocytes to (protein bound) UTs. However, the exact mechanisms underlying CKD-triggered MVD and corresponding HFpEF remain to be further elucidated, especially at the molecular level.

In this review we focused on the microvasculature in female CKD patients before and after menopause, which will improve our understanding on the subsequent development of HFpEF. Firstly, female gender and aging, two major risk factors of HFpEF, will be addressed. Secondly, the evidence for the role of estrogen (E_2) mediated protection mechanisms in the onset and progress of renal disease-related MVD and HFpEF will be summarized. Finally, we will further discuss the new information that we have gathered from the analysis of publicly available NCBI Gene Expression Omnibus (GEO) database sets for the transcriptome response of endothelial cells (ECs) to E_2 and CKD associated circulatory factors. Based on this, we propose putative pathways of CKD-related MVD that are susceptible to E_2 protection.

2. Postmenopausal women are at high risk of developing HFpEF

The prevalence of HFpEF increases with age and HFpEF patients are typically older than those with HFrEF [18]. In general, the average age of HFpEF patients are between 73 and 79 years old [19]. Aging has been proposed as an independent risk factor for abnormal diastolic function [20]. Age-dependent increase in left ventricular mass index has been observed in humans, and age-dependent increase in cardiomyocyte size has been observed in animals. In addition, increased interstitial fibrosis has also been noticed in aged myocardium. These changes due to aging contribute to myocardial stiffness, putatively leading to diastolic dysfunction in HFpEF. Unfortunately, clinical trial data of treatments for HF were mostly collected from the young and the middle-aged patients, leading to the lack of adequate evidence in treating the elderly, not to mention the older HFpEF patients specifically [21].

Besides the elderly population, women are consistently ~2 times more at risk than men to develop HFpEF and outnumber men by a 2:1 ratio in the HFpEF patient population [22]. Women also differ from the male HFpEF population as they show less evidence of coronary artery disease but are more vulnerable for coronary MVD, indicating a sex-based difference in the underlying pathology of HFpEF [22, 23]. Left ventricular diastolic dysfunction (LVDD) can be considered as a pre-stage of HFpEF. In the female population, LVDD onset and progression into HFpEF is strongly associated to the postmenopausal period [24]. High estrogen levels appear to protect the premenopausal heart from ventricular remodeling triggered by hypertension, although the specific mechanism remains to be further defined. Therapeutic interventions for HFpEF have failed to improve the mortality rate. At the moment early detection and treatment of LVDD appears to be the only effective strategy to prevent progression into HFpEF.

Since both aging and female gender appear to be important risk factors for LVDD and HFpEF, it has been postulated that gender specific hormones and changes in hormone levels may play an important role in the higher prevalence of HFpEF in women, particularly in the postmenopausal population [25]. Studies in the early 1990s already showed the beneficial effects of menopausal hormone therapy (MHT) on preventing coronary heart disease [26]. However, subsequent randomized clinical trials failed to demonstrate that MHT prevents secondary events in ischemic heart disease, cerebrovascular events, or progression of coronary atherosclerosis in postmenopausal patients with already established coronary disease [27, 28]. Recently, reevaluation of those trails and the initiation of new clinical trials have shed light on how to improve MHT. In particular, the effects of early versus late MHT intervention were evaluated, as comparison between previous MHT responsive and non-responsive groups have indicated that women who received intervention at the early postmenopause stage without pre-existing coronary disease were more likely to benefit from MHT than older patients with pre-existing coronary disease [29]. A recent retrospective single-center study showed that MHT was significantly associated with improved left ventricular relaxation indices, which is in line with the reported improvement in diastolic function following MHT in postmenopausal women, pointing towards the need for further investigation of the use of MHT in treatment of HFpEF [30]. Together, these clinical studies indicate the postmenopausal women are at high risk of developing HFpEF.

3. Postmenopausal estrogen depletion in female CKD patients and microvascular dysfunction

A limited number of studies have started to reveal the putative disease mechanisms of LVDD and HFpEF in women with CKD. An *in vitro* study showed that the contraction rate in uremic toxin p-cresol treated cardiomyocytes was decreased, and p-cresol impaired cardiomyocytes gap junctions by

increasing the activity of protein kinase Cα [31]. UTs have also been shown to induce cardiac remodeling response via estrogen receptor dependent mitogen-activated protein kinase and nuclear factor-κB pathways, suggesting that estrogen receptor signaling could interfere with the negative effects of UTs [32].

Brunet et al. have summarized two major mechanisms of how UTs contribute to vascular dysfunction [33]: (1) UTs promote inflammation by stimulating leukocyte activation and endothelial adhesion molecule expression. Activated inflammation and immune responses increase the migration and proliferation of vascular smooth muscle cells (VSMCs). However, UTs also inhibit the proliferation of ECs and enhance the apoptosis of endothelial progenitor cells, thus impairing vascular repair. (2) UTs stimulate the transdifferentiation of VSMCs into osteoblast-like cells and reduce digestibility of collagen and other extracellular matrix proteins by forming irreversible crosslinks, which subsequently lead to an increased vessel stiffness and vascular dysfunction. Among over 150 UTs that have been listed to date, some UTs like indole-3-acetic acid strongly accumulate in the circulation of patients who are still in an early stage of CKD as compared to normal levels observed in healthy individuals [34]. The concentrations of 11 different uremic toxins have been reported to be 2.3 to 44.7 times increased in patients with stage 3 and stage 4 CKD (moderate to severe CKD stage but before stage 5 (dialysis stage)) versus the plasma levels found healthy controls [35]. Protein-bound UTs, such as ADMA, p-cresyl sulfate and indoxyl sulfate (IS), also accumulate in patients with early CKD stages with continuous concentration build up during CKD progression [36]. These protein-bound toxins have been shown to exhibit high endothelial and vascular toxicity [36, 37]. A common pathway for these UTs is the activation of NAD(P)H oxidase, a reactive oxygen species (ROS) inducer, leading to oxidative stress in ECs and a reduction of nitric oxide (NO) bioavailability in the micro-environment [38]. NO contributes to vasorelaxation and inhibits platelet aggregation, expression of adhesion molecules and proliferation of VSMCs. Through activated oxidative stress and the subsequent activation of the p38/ mitogen-activated protein kinase pathway, UTs also affects immune cells as demonstrated by increased cell-surface expression of the immune activation marker beta2-integrin Mac-1 (CD11b/CD18) in the leukocyte and monocyte cell populations of CKD patients.

Unlike the deleterious influence of UTs on endothelium, clinical studies showed that E_2 administration increased flow-mediated vasodilation response, indicating that E_2 is an important regulator of protective endothelial function [39]. The non-genomic and genomic pathways of E_2 that act via its three receptors (ERα, ERβ, and GPER) have been shown to be able to increase endothelial NO synthase (eNOS) in various cell types [39, 40]. In E_2 treated human ECs, ERα signaling via PI3K triggers the activation of protein kinase B, extracellular-signal-regulated kinase 1/2, and phosphorylation and activation of eNOS. Increased eNOS and NO bioavailability promote vascular relaxation, ECs migration

and proliferation. A rapid increase in intracellular calcium is also observed after E_2 stimulation [40]. A recent paper demonstrated that activated GPER increased the expression of calmodulin and prolonged cytoplasmic Ca^{2+} signals via the transactivation of epidermal growth factor receptor and the activation of mitogen-activated protein kinase cascade in porcine aortic ECs [41]. Calmodulin is a transducer of Ca^{2+} signals, and an activated calmodulin/Ca2+ system is able to modulate eNOS function. Interestingly, E_2 treatment was shown to increase the number of endothelial progenitor cells in a NO-dependent way, which in return restored vascular repair activities [42]. Furthermore, Osako et al. showed that the activation of ERα suppressed the cascade of the receptor activator of nuclear factor-κB and its ligand, which subsequently promoted expression of a calcification inhibitor matrix Gla protein while inhibiting the expression of a calcification inducer bone morphogenetic protein 2 in both human aortic ECs and VSMCs, implying a beneficial effect of E_2 in preventing vascular calcification [43].

A community-based study showed that 18.2% of 445 women at the mean age of 45.2 years old had reduced glomerular filtration rate (<60 mL/min/1.73m2), indicating a high prevalence of CKD in perimenopausal women [44]. Furthermore, clinical studies showed that MHT protected renal function by increasing the glomerular filtration rate in healthy postmenopausal women. This was coincided with a lower left ventricular mass index when compared to other healthy postmenopausal women with lower glomerular filtration rate who did not receive MHT [45]. Taken together, we postulate that reduced estrogen levels during and after menopause could increase the vulnerability of these CKD patients for MVD, which may lead to LVDD and HFpEF. To further elucidate this process, we analyzed published transcriptome datasets of UTs or E_2 treated human derived ECs from NCBI GEO database. We investigated genes and pathways that may be involved in UTs-triggered MVD in CKD. We obtained a gene set (group A) that poses deleterious effects on endothelial function and could be inhibited by E_2 stimulation, and a second gene set (group B) with an endothelial protective effect that could be induced by E_2 (Supplemental Table 1). Based on these two gene sets, we propose several target genes and their associated biological pathways, which may assist future studies in identifying valuable biomarkers and drug targets for early diagnosis and treatment of the CKD postmenopausal population at risk of developing MVD and subsequent HFpEF.

3.1 UTs impair microvascular homeostasis

Following a systematic pipeline (Figure 1), we obtained two GEO datasets that provide information of gene expression changes in ECs under CKD condition. Briefly, keywords “chronic kidney disease” and “uremic toxin” were coupled with “endothelial cells” separately to gather datasets from NCBI GEO

database. Available datasets were further filtered for "Organism: Homo sapiens" and "Study type: Expression profiling by array". Two GEO datasets were found according to these criteria (Table 1). GSE34259 contains information of the transcriptome profile of human umbilical vein endothelial cells (HUVECs) in response to IS treatment. GSE60937 contains information of the transcriptome profile of HUVECs in response to high inorganic phosphate (HP) treatment. Both raw datasets were downloaded and sorted by R script provided by NCBI's GEO2R program. These sorted gene expression datasheets were analyzed as followed: (1) Based on default setting stated in GEO2R, p-value below 0.05 was used to filter genes that reached significance; (2) Selected genes with logFC value above 0 represented increased expressions, and those with logFC value below 0 represented decreased expressions. (3) Genes with increased or decreased expression were mapped between IS and HP groups to obtain overlapping genes. We identified 36 genes with significantly increased expression under both IS and HP stimulation, whereas 14 gene decreased expressions under both IS and HP stimulation (Figure 2A, Supplemental Table 2).

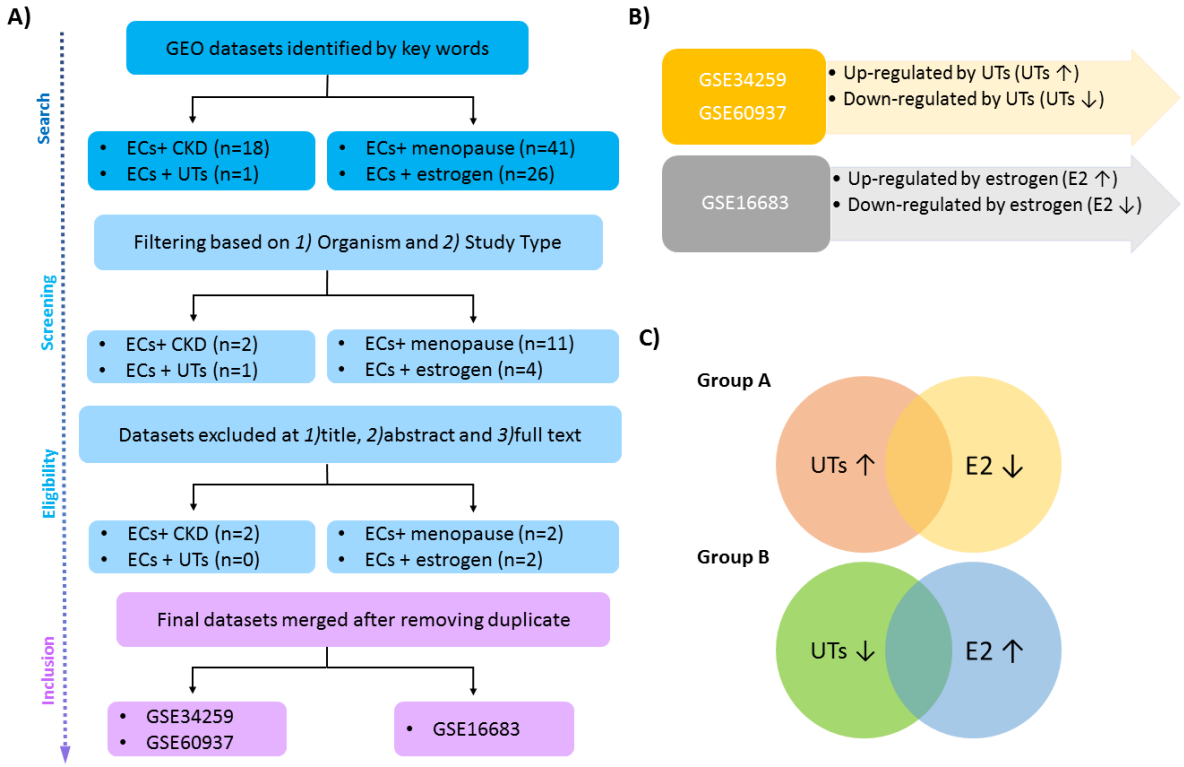


Figure 1. Overview of the workflow for identifying interesting gene groups from three included GEO datasets, representing gene expressions in HUVECs under indoxyl sulfate (IS), high level inorganic phosphates (HP) and estrogen (E_2) stimulations. A) Workflow of dataset collection by searching for key words in GEO datasets, followed by filtering "Organism: Homo sapiens" and "Study Type: Expression profiling by array" and screening for the relevance. Three datasets were finally included after removing duplicates in filtered datasets. B) Significantly up- and down-regulated genes in HUVEC under UTs (IS and HP) and E_2 were identified based on their fold change and p values. C) By mapping genes that were regulated by both UTs and E_2, we identified a

group of genes (group A) that were upregulated by UTs (IS and/or HP) and downregulated by E_2, and another gene group (group B) in which genes were inhibited by UTs and increased by E_2.

Table 1. Included GEO (Gene Expression Omnibus) datasets

GEO Series	Cell type	Stimulus	Number of up-regulated genes	Number of down-regulated genes
GSE16683	HUVECs	E_2	2816	1670
GSE34259	HUVECs	IS	286	50
GSE60937	HUVECs	HP	1633	3175

HUVECs: human umbilical vein endothelial cells; E_2: estradiol; IS: indoxyl sulfate; HP: high inorganic phosphate.

3.2 Functional annotation of UTs activated genes include positive regulation of prostaglandin synthesis and prostaglandin-related processes

Based on the 36 genes that were significantly activated by both IS and HP, we performed comprehensive functional annotations by using the ToppGene Suite tool ToppFun (Correction: FDR; P-Value cutoff: 0.05; Gene limits: 1≤ n ≤2000). The most enriched biological processes include positive regulation of prostaglandin biosynthetic process (GO:0031394) and prostaglandin-related processes (GO:2001280, GO:2001279, GO:0031392, GO:0046890, Supplemental Table 3). In addition, cardiac chamber morphogenesis (GO:0003206) was identified as the 7th most enriched biological process. By building a protein-protein interaction network with these 36 genes using STRING, we identified *PTGS2, NRG1, ICAM1, PPKCA, PLA2G4A* and *ADAMTS1* as key regulators within the constructed networks (Figure 2B, confidence score ≥0.4). Basal levels of prostaglandin production are very low but increase significantly during inflammation, indicating UTs may promote inflammation of the endothelium. Many of the identified key mediators of this gene set, including *PTGS2*, *ICAM1* and *PLA2G4*, indeed exhibit pro-inflammatory properties. During inflammation, arachidonic acid is released from the plasma membrane by *PLA2G4*-encoded phospholipase 2, followed by PTGS2 mediated conversion of arachidonic acid to prostaglandins [46]. ICAM1 is a ligand for lymphocyte function-associated antigen 1 on leukocytes, and upregulation of ICAM1 expression in ECs during endothelial activation is essential for leukocytes recruitment [47]. Interestingly, Hadad et al. showed that phospholipase 2 increased ICAM1 expression via two transcription factors nuclear factor-κB and CREB in ECs [48]. These data indicate that prostaglandin regulation may represent a common pro-inflammatory pathway underlying MVD, which is induced by both IS and hyperphosphatemia.

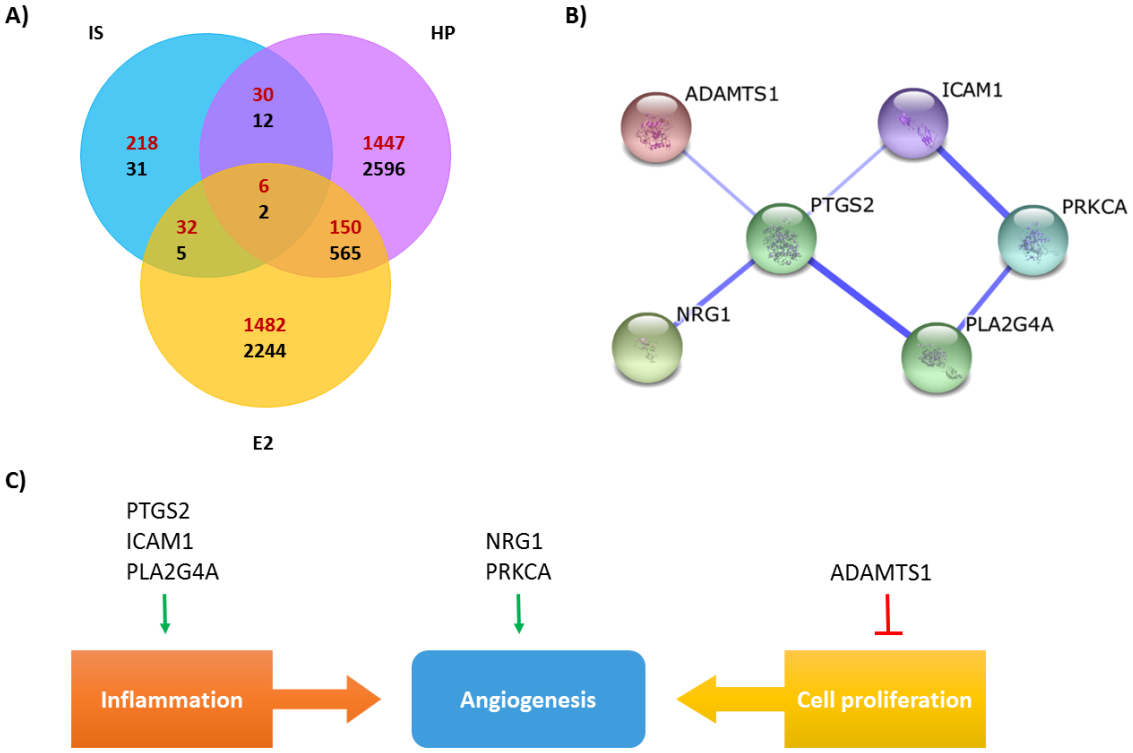


Figure 2. Identification of genes with significantly increased and decreased expressions in response to IS, HP and/or E_2 stimulation in HUVECs. A) Venn diagram of genes stimulated by UTs and suppressed by E_2 (group A, indicated in red) and genes inhibited by UTs and enhanced by E_2 (group B, indicated in black). B) Among 36 genes increased by both IS and HP, *ADAMTS1, PTGS2, NRG1, ICAM1, PRKCA* and *PLA2G4A* were identified as the key mediators by STRING (confidence score ≥0.4). Stronger associations are represented by thicker lines. C) Biological functions of identified key genes regarding endothelium homeostasis based on literature studies. Green arrow: gene promotes the process; red indicator: gene inhibits the process.

3.3 Dysregulation of angiogenesis in UTs exposed vasculature

Inflammation-induced angiogenesis aids in the replacement of lost microvasculature and restores microvascular density in CKD [49]. Other identified key mediators of this gene set such as *PRKCA* and *NRG1* are known to promote angiogenesis, whereas *ADAMTS1* blocks angiogenesis: An *in vitro* study showed that inhibition of *PRKCA* resulted in decreased expression of *VEGF*, a crucial angiogenic factor [50]. Endothelial-derived NRG1 has been reported to bind to ERBB receptors, and the downstream cascade involves the activation of angiogenesis-related tyrosine kinase receptors, like VEGF receptors and Eph receptors [51]. *ADAMTS1* has been proposed to block angiogenesis via two mechanisms, i.e. by suppressing EC proliferation by disrupting VEGF signaling or by releasing anti-angiogenic peptides from angiogenesis-related proteins thrombospondin 1 and 2 [52]. Based on our analysis, we propose that increased circulating UTs could negatively impact renal and cardiac endothelial homeostasis by

triggering inflammation and initiating vascular damage. In addition, vascular repair activity appears to be impaired by UTs-activated genes, leading to further progression of MVD.

4. E_2 actives microvascular protective mechanisms

We also searched for gene expression changes in ECs under postmenopausal condition. Keywords "menopause" and "estrogen" were coupled with "endothelial cells" separately to gather datasets from NCBI GEO database. Available datasets were further filtered for "Organism: Homo sapiens" and "Study type: Expression profiling by array" (Figure 1). One GEO dataset (GSE16683) was found according to these criteria, which showed the transcriptome response of E_2 stimulated HUVECs (Table 1). Gene transcripts that were significantly increased and decreased by E_2 were identified in the same way as previously described. In order to evaluate the susceptibility of UTs influenced genes to E_2 regulation, we mapped genes with increased and/or decreased expressions among IS, HP and E2 groups. We identified 188 overlapping genes that were increased in expression by UTs (IS and/or HP) and inhibited by E_2 in group A, and 572 overlapping genes that were decreased in expression by UTs and increased by E_2 in group B. Group A represents genes which are possibly involved in MVD-inducing mechanistic pathways that could be suppressed by E_2 protection. Group B represented genes in putative protective MVD-preventive pathways that could be induced by E_2.

4.1 E_2 protects microvasculature by suppressing UTs-induced inflammation

Based on 188 genes in group A, the most enriched biological processes (Table 2, Supplemental Table 4) that were annotated by using ToppFun include "regulations of nitrogen compound metabolic process", "potassium ion import" and "neutrophil migration", which are known to influence microvascular homeostasis directly and could lead to MVD [53, 54]. During inflammation, neutrophils extravasate the vasculature by migrating between ECs to inflamed sites. Increased ROS levels, produced by neutrophils, disrupt endothelial integrity by inhibiting endothelial occludin expression in tight junctions and by activating the phosphorylation of VE-cadherin, β-catenin and P120 catenin in adherens junctions, leading to increased diapedesis of inflammatory cells [53]. Excessive ROS also activates the JNK cascade, which is responsible for apoptosis, and could lead to tissue injury. Vasodilation regulated by "positive regulation of nitrogen compound metabolic process" and "potassium ion import", further promotes a persistent inflammatory response [55]. In the cardiac microvasculature, it has been reported that increased ROS levels lowers the activity of protein kinase G and titin hypophosphorylation, which resulted in an increased resting tension of cardiomyocytes

[56]. In addition, lower the activity of protein kinase G contributed to cardiomyocyte hypertrophy, and subsequently increased left ventricle wall stiffness. Based on these observations, we conclude that E_2 could protect the (myo)endothelium by inhibiting UTs-induced inflammation, especially via downregulation of genes involved in ROS signalling.

4.2 E_2 facilitates vascular repair activity by activating platelet coagulation and hemostasis at injured sites

Based on the 572 genes in group B, the most enriched biological processes that are annotated by using ToppFun include “responses to stress”, “wounding and growth factors” and “hemostasis” (Table 2, Supplemental Table 4). These processes have been linked in literature to mechanisms that restore microvascular homeostasis [57-60]. In addition, nerve growth factor signaling (ID106459) was one of the top pathways annotated in this gene set (Figure 3) and is known to play a regulatory role in promoting angiogenesis in vascular disease. For example, Park et al. showed that the binding of nerve growth factor to its receptor TrkA increased matrix metalloproteinase 2 expression via the activation of PI3K/Akt pathway and the AP-2 transcription factor [57]. Another significantly enriched pathway is the interleukin 6 signaling pathway (ID198864, Figure 3), which has been positively linked to coagulation [58]. Blood coagulation, a significantly enriched Biological process (GO:0007596, p-value<0.01), is in line with the observations of estrogen-related pro-coagulation capacity in clinical studies and represents an important step in hemostasis [61]. During hemostasis, tissue factor and the serine protease factor VIIa activate factor X and IX, which initiate the coagulation cascade and lead to the formation of thrombin [62]. Thrombin cleaves fibrinogen to generate insoluble fibrin, forming a fibrin mesh to strengthen and stabilize the blood clot and stop bleeding at the site of injury. By cleaving 2 protease activated receptors, PAR1 and PAR4, thrombin also activates platelets. Activated platelets express receptor GPIb-IX-V and receptor GPVI to bind von Willebrand factor and collagen in the sub-endothelial matrix, which further facilitates platelet adhesion. In addition, activated platelets secrete both pro- and anti-angiogenetic factors to repair and replace damaged blood vessels and restore homeostasis [59].

Based on these findings, we suggest that E_2 could counter the UTs-induced inhibition of the platelet coagulation and hemostasis response of the (micro)vasculature. Indeed, it has been long recognized that the coagulation system of patients with renal insufficiency and uremia is profoundly affected, demonstrating frequent symptoms of severe bleeding or thrombosis that lead to significant increase in morbidity and mortality [63]. The expression data of group B imply that UTs may contribute to the pathogenesis of uremic bleeding and thrombotic diathesis, which can be counteracted by the effects

of E_2 protection. In relation to uremic bleeding, treatment with conjugated E_2 has been proposed and shown to be effective [64]. Interestingly, uremic bleeding has been linked to high NO bioavailability. Early studies have shown that monomethyl-L-arginine mediated inhibition of NO could normalize platelet dysfunction and bleeding time of uremic rats [65]. Likewise, systemic inhibition of NO production in healthy human volunteers could significantly shorten the bleeding time [66]. E_2 has been shown to normalize vascular expression of NO-producing enzymes and NO plasma levels in uremic rats [67], whereas induction of NO diminished the protective effects of estrogen in uremic bleeding [68]. As NO has been indicated to play a central role in cardiac endothelial dysfunction, linking microvascular disease to onset of HFpEF, the uncovered putative mechanism of UTs regulation of the coagulation system may be an important step in the pathogenesis that requires further study, particularly in relation to post-menopausal estrogen levels.

Table 2. Most enriched biological processes in uremic toxins-increased and estrogen-inhibited groups (group A) and uremic toxins-decreased and estrogen-enhanced groups (group B).

GO Rank	ID	Name	FDR-corrected p-value	Number of input/annotation genes
Group A				
1	GO:0031328	Positive regulation of cellular biosynthetic process	7.410E-6	37/1872
2	GO:0097368	Establishment of Sertoli cell barrier	7.799E-6	3/5
3	GO:0051173	Positive regulation of nitrogen compound metabolic process	8.883E-6	37/1887
4	GO:0010107	Potassium ion import	1.431E-4	4/29
5	GO:0002027	Regulation of heart rate	2.602E-4	6/95
6	GO:0010657	Muscle cell apoptotic process	3.711E-4	5/66
7	GO:1902622	Regulation of neutrophil migration	4.168E-4	4/38
8	GO:0001666	Response to hypoxia	4.293E-4	10/293
Group B				
1	GO:0007010	Cytoskeleton organization	1.747E-9	72/1142
2	GO:0007049	Cell cycle	4.609E-9	97/1780
3	GO:0033554	Cellular response to stress	4.610E-3	97/1983
4	GO:0033043	Regulation of organelle organization	2.126E-6	62/1116
5	GO:0007599	Hemostasis	4.903E-6	38/574
6	GO:0016032	Viral process	6.011E-6	48/810
7	GO:0009611	Response to wounding	7.165E-6	60/1109
8	GO:0022008	Neurogenesis	8.600E-6	88/1848
9	GO:0070848	Response to growth factor	9.128E-6	53/944
10	GO:0044764	Multi-organism cellular process	9.172E-6	48/823

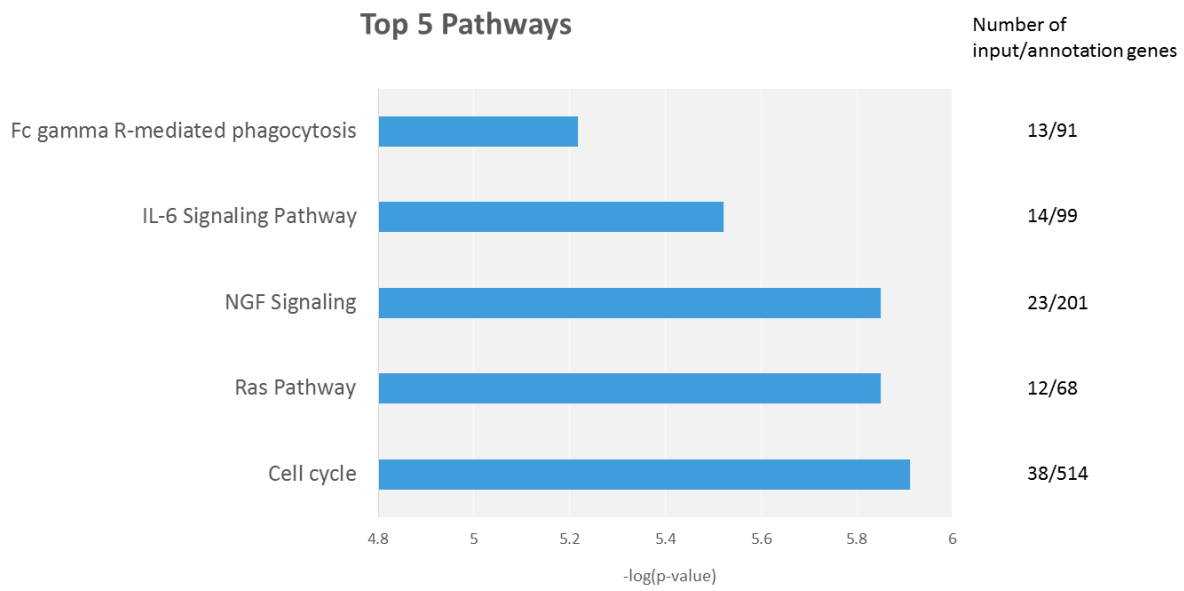


Figure 3. Top 5 statistically significant pathways based on 572 genes in group B using the ToppGene Suite tool ToppFun (Correction: FDR; P-Value cutoff: 0.05; Gene limits: 1≤ n ≤2000). The number of input genes and genes involved in each pathway are also listed.

4.3 Possible UTs and E_2 regulate gene targets in EC-cardiomyocyte paracrine crosstalk

Since ECs are located at a distance of maximum 5 µM from cardiomyocytes, UTs and/or E_2 induced signaling in ECs could further influence cardiac function in a paracrine fashion. In order to identify possible genes that are involved in EC-cardiomyocyte paracrine crosstalk, genes in Group A and B were re-analyzed using ToppFun regardless of their p-value (p-value cutoff: 1), and the obtained enrichments (including Molecular function, Biological process, Pathway, Mouse phenotype, Human phenotype and Disease) were further filtered for terms including “cardio”, “cardiac”, “myocard” and “heart”. Following this procedure, an enrichment list of 58 genes (31%) in group A, and 138 genes (24%) in group B were produced that matched the above described criteria (Figure 4C and 4D).

Gene functions annotated from these 58 cardiac related genes in group A include “leukocyte migration”, “immune system development”, “positive regulation of cell motility”, “response to lipid”, “regulation of JNK cascade”, “vascular process in circulatory system”, and “regulation of cellular response to stress” (Figure 4A). Immune responses are known to be linked to endothelial dysfunction, which in turn is related to a lower activity of protein kinase G, titin hypophosphorylation, and an increased resting tension of cardiomyocytes and a hypertrophic response [56]. Endothelial dysfunction-related impaired NO bioactivity also results in endothelial-to-mesenchymal transition, during which ECs differentiate into (myo)fibroblasts and contribute to cardiac fibrosis. The identification of inflammation and MVD related mechanisms in the gene set of group A, thus imply

that these initiating factors that contribute to LVDD and HFpEF could be mainly induced by circulating UTs during post-menopause.

Gene functions annotated from the 138 cardiac related genes in group B include “Notch signaling pathway”, “ERBB signaling pathway”, “fibroblast growth factor receptor signaling pathway” and “downregulation of immune process” (Figure 4B). During embryonic ventricular development, Notch activation induces the expression of *EPHRINB2*, *NRG1* and *BMP10*, resulting in trabecular differentiation and the proliferation of trabecular cardiomyocytes. In addition, Notch pathway, especially *NOTCH1*, is important in regulating endothelial function in the aortic valve, whereas endothelial dysfunction is associated with aortic valve calcification [69]. In response to pressure overload of the left ventricle due to aortic valve calcification, the myocardium becomes hypertrophic, eventually leading to diastolic dysfunction [70, 71]. A recent paper showed an increased level of Notch ligand DLL1 was associated with diastolic dysfunction, but no significant association between serum level of DLL1 and HF patients with LVEF>50% was found [72].

ERBB signaling is activated by endothelium-derived neuregulins. Parodi et al. indicated that during heart development, neuregulins-ERBB4 complex initiates dimerization with ERBB2, leading to a neuregulins/ERBB signaling cascade that promotes cardiomyocytes proliferation and differentiation [73]. Protein levels of ERBB2 and ERBB4 have been reported to be extremely reduced in HF patients, suggesting a key role of ERBB signaling in maintaining cardiac function [73]. In line with this notion, genetic knockdown of FRG receptor 1 and 2 in ECs of myocardial infarcted mice led to decreased vessel density, increased apoptosis and worsened cardiac function measured by echocardiography, suggesting a protective role of fibroblast growth factor signaling in vascular adaptation for cardiomyocyte function during ischemic heart injury [74].

To conclude, these findings imply that E_2 might protect the heart against UTs by suppressing inflammatory responses, protecting cardiac vessel density and facilitating cardiomyocyte proliferation in an EC-cardiomyocyte paracrine crosstalk dependent manner.

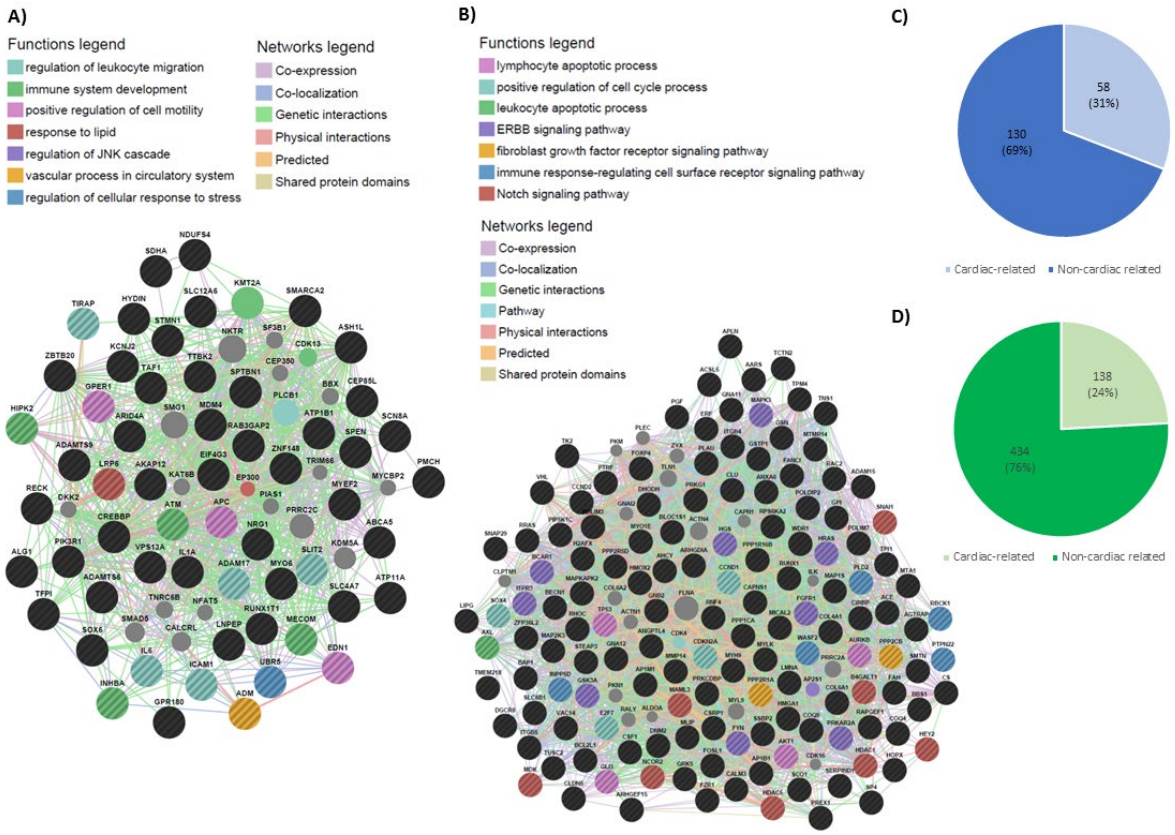


Figure 4. Gene annotations based on 38 genes and 138 genes related to cardiac function in group A and B. A, B) Gene functions including; immune response, apoptosis, response to lipid and stress signals in blood vessel cells were annotated by GeneMANIA (default parameters) in group A, whereas gene functions including; downregulation of immune-related process, ERBB signaling pathway, Notch signaling pathway and fibroblast growth factor receptor signaling were annotated in group B. Genes in black and/or with lines within the nodes are input gene and genes in grey and/or without lines within the nodes are annotation genes. C, D) Proportions of cardiac function-related genes within group A and B, respectively.

5. Conclusion

In the current review, we aimed to gain more insight into the gene expression profiles of ECs that are predisposing and preventing MVD and MVD-related HFpEF in postmenopausal women with CKD. By analyzing three published microarray datasets based on the transcriptome responses of IS-, HP- and E_2-treated HUVECs, we conclude that transcriptional changes of ECs exposed to CKD-associated circulatory risk factors IS and HP are characterized by increased expression of genes that are related to activation of inflammation and dysregulation of angiogenesis. In addition, these deleterious pathogenic expression profiles may be ameliorated by E_2 protection during the premenopausal period by a gene expression profile that promotes suppression of UTs-induced inflammation and facilitates angiogenesis by restoring coagulation and hemostasis in injured vessels. Reduced estrogen levels during and after menopause could accelerate the development of cardiorenal syndrome and CKD-associated HFpEF in CKD patients due to loss of E_2 protection of the endothelium in the deleterious

environment of UTs. Our analysis provides, for the first time to our knowledge, a comprehensive *in silico* analysis of possible genes involved in CKD associated MVD that are regulated by E_2 and may be more affected during postmenopause period. Other mechanisms, which could also contribute to CKD-related HFpEF, might interact or work in parallel with our proposed mechanism. Clinical studies of the high-risk population of postmenopausal patients with cardiorenal syndrome are required to further verify our findings and to reveal the precise disease mechanisms involved. Eventually, improved insight into the link between loss of estrogen protection, MVD onset and progression in the CKD postmenopausal population could yield new and more dedicated biomarkers and drug targets for the development of new diagnostic tools and pharmacotherapies for HFpEF for this specific patient population.

Acknowledgement

This work was supported by Netherlands Foundation for Cardiovascular Excellence [to C.C], two NWO VIDI grants [no. 91714302 to C.C., and no. 016096359 to M.V.] the ErasmusMC fellowship grant [to C.C.], the RM fellowship grant of the UMC Utrecht [to C.C], the Netherlands Cardiovascular Research Initiative: An initiative with support of the Dutch Heart Foundation [CVON2014-11 RECONNECT, to M.V., D.D., and C.C.], and the Dutch Heart Foundation [Netherlands Heart Foundation: 2013/T084, Queen of Hearts, to G.P, H.d.R., J.P. and C.C.].

Author Disclosure Statement

The authors declare that they have no competing interests.

Supplementary Files – Online

Supplementary Table 1. Overlapping genes between UTs (IS and/or HP) and E2 regulated gene sets.

Supplementary Table 2. Genes up- and down-regulated significantly by UTs (both IS and HP).

Supplementary Table 3. Input genes in top 20 biological process obtained from 36 UTs up-regulated genes.

Supplementary Table 4. Input genes in most enriched biological process obtained from both groups.

References

[1] C.S. Lam, E. Donal, E. Kraigher-Krainer, R.S. Vasan, Epidemiology and clinical course of heart failure with preserved ejection fraction, Eur J Heart Fail, 13 (2011) 18-28.
[2] F. Edelmann, R. Stahrenberg, G. Gelbrich, K. Durstewitz, C.E. Angermann, H.D. Dungen, T. Scheffold, C. Zugck, B. Maisch, V. Regitz-Zagrosek, G. Hasenfuss, B.M. Pieske, R. Wachter, Contribution of comorbidities to functional impairment is higher in heart failure with preserved than with reduced ejection fraction, Clin Res Cardiol, 100 (2011) 755-764.
[3] R. Pecoits-Filho, S. Bucharles, S.H. Barberato, Diastolic heart failure in dialysis patients: mechanisms, diagnostic approach, and treatment, Semin Dial, 25 (2012) 35-41.
[4] K. Sharma, D.A. Kass, Heart failure with preserved ejection fraction: mechanisms, clinical features, and therapies, Circ Res, 115 (2014) 79-96.
[5] E.D. Unger, R.F. Dubin, R. Deo, V. Daruwalla, J.L. Friedman, C. Medina, L. Beussink, B.H. Freed, S.J. Shah, Association of chronic kidney disease with abnormal cardiac mechanics and adverse outcomes in patients with heart failure and preserved ejection fraction, Eur J Heart Fail, 18 (2016) 103-112.
[6] F.P. Brouwers, R.A. de Boer, P. van der Harst, A.A. Voors, R.T. Gansevoort, S.J. Bakker, H.L. Hillege, D.J. van Veldhuisen, W.H. van Gilst, Incidence and epidemiology of new onset heart failure with preserved vs. reduced ejection fraction in a community-based cohort: 11-year follow-up of PREVEND, Eur Heart J, 34 (2013) 1424-1431.
[7] J.M. Ter Maaten, K. Damman, M.C. Verhaar, W.J. Paulus, D.J. Duncker, C. Cheng, L. van Heerebeek, H.L. Hillege, C.S. Lam, G. Navis, A.A. Voors, Connecting heart failure with preserved ejection fraction and renal dysfunction: the role of endothelial dysfunction and inflammation, Eur J Heart Fail, 18 (2016) 588-598.
[8] D.H. Smith, M.L. Thorp, J.H. Gurwitz, D.D. McManus, P. Robert J. Goldberg, L.A. Allen, G. Hsu, S.H. Sung, D.J. Magid, A.S. Go, Chronic Kidney Disease and Outcomes in Heart Failure With Preserved Versus Reduced Ejection Fraction, Circ Cardiovasc Qual Outcomes, (2013).
[9] R.J. Widmer, A. Lerman, Endothelial dysfunction and cardiovascular disease, Glob Cardiol Sci Pract, 2014 (2014) 291-308.
[10] W.J. Paulus, C. Tschope, A novel paradigm for heart failure with preserved ejection fraction: comorbidities drive myocardial dysfunction and remodeling through coronary microvascular endothelial inflammation, J Am Coll Cardiol, 62 (2013) 263-271.
[11] A. Clementi, G.M. Virzi, C.Y. Goh, D.N. Cruz, A. Granata, G. Vescovo, C. Ronco, Cardiorenal syndrome type 4: a review, Cardiorenal Med, 3 (2013) 63-70.
[12] A. Regeniter, H. Freidank, M. Dickenmann, W.H. Boesken, W.H. Siede, Evaluation of proteinuria and GFR to diagnose and classify kidney disease: systematic review and proof of concept, Eur J Intern Med, 20 (2009) 556-561.
[13] T. Toyohara, Y. Akiyama, T. Suzuki, Y. Takeuchi, E. Mishima, M. Tanemoto, A. Momose, N. Toki, H. Sato, M. Nakayama, A. Hozawa, I. Tsuji, S. Ito, T. Soga, T. Abe, Metabolomic profiling of uremic solutes in CKD patients, Hypertens Res, 33 (2010) 944-952.
[14] X. Tao, S. Thijssen, P. Kotanko, C.H. Ho, M. Henrie, E. Stroup, G. Handelman, Improved dialytic removal of protein-bound uraemic toxins with use of albumin binding competitors: an in vitro human whole blood study, Sci Rep, 6 (2016) 23389.
[15] R. Vanholder, U. Baurmeister, P. Brunet, G. Cohen, G. Glorieux, J. Jankowski, G. European Uremic Toxin Work, A bench to bedside view of uremic toxins, J Am Soc Nephrol, 19 (2008) 863-870.
[16] M. Aepfelbacher, M. Essler, Disturbance of endothelial barrier function by bacterial toxins and atherogenic mediators a role for RhoRho kinase, Cellular Microbiology, (2001).
[17] N. Jourde-Chiche, L. Dou, C. Cerini, F. Dignat-George, P. Brunet, Vascular incompetence in dialysis patients--protein-bound uremic toxins and endothelial dysfunction, Semin Dial, 24 (2011) 327-337.
[18] J.E. Ho, A. Lyass, D.S. Lee, R.S. Vasan, W.B. Kannel, M.G. Larson, D. Levy, Predictors of new-onset heart failure: differences in preserved versus reduced ejection fraction, Circ Heart Fail, 6 (2013) 279-286.

[19] T. Bhuiyan, M.S. Maurer, Heart Failure with Preserved Ejection Fraction: Persistent Diagnosis, Therapeutic Enigma, Curr Cardiovasc Risk Rep, 5 (2011) 440-449.
[20] F.S. Loffredo, A.P. Nikolova, J.R. Pancoast, R.T. Lee, Heart failure with preserved ejection fraction: molecular pathways of the aging myocardium, Circ Res, 115 (2014) 97-107.
[21] K. Alagiakrishnan, M. Banach, L.G. Jones, A. Ahmed, W.S. Aronow, Medication management of chronic heart failure in older adults, Drugs Aging, 30 (2013) 765-782.
[22] D.C. Scantlebury, B.A. Borlaug, Why are women more likely than men to develop heart failure with preserved ejection fraction?, Curr Opin Cardiol, 26 (2011) 562-568.
[23] A.H.E.M. Maas, Y.E.A. Appelman, Gender differences in coronary heart disease, Netherlands Heart Journal 18 (2010).
[24] A.A. Oktay, J.D. Rich, S.J. Shah, The emerging epidemic of heart failure with preserved ejection fraction, Curr Heart Fail Rep, 10 (2013) 401-410.
[25] E.Z. Gorodeski, Is early menopause a risk factor for heart failure?, Menopause, 21 (2014) 558-560.
[26] B.E. Henderson, A. Paganini-Hill, R.K. Ross, Decreased mortality in users of estrogen replacement therapy, Arch Intern Med, 151 (1991) 75-78.
[27] H.N. Hodis, W.J. Mack, S.P. Azen, R.A. Lobo, D. Shoupe, P.R. Mahrer, D.P. Faxon, L. Cashin-Hemphill, M.E. Sanmarco, W.J. French, T.L. Shook, T.D. Gaarder, A.O. Mehra, R. Rabbani, A. Sevanian, A.B. Shil, M. Torres, K.H. Vogelbach, R.H. Selzer, Hormone Therapy and the Progression of Coronary-Artery Atherosclerosis in Postmenopausal Women, The New England journal of medicine, 349 (2003).
[28] S. Hulley, D. Grady, T. Bush, C. Furberg, D. Herrington, B. Riggs, E. Vittinghoff, Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease in postmenopausal women, JAMA, 280 (1998).
[29] O.M. Reslan, R.A. Khalil, Vascular effects of estrogenic menopausal hormone therapy, Rev Recent Clin Trials, 7 (2012).
[30] M.A. Duzenli, K. Ozdemir, A. Sokmen, K. Gezginc, A. Soylu, C. Celik, B.B. Altunkeser, M. Tokac, The effects of hormone replacement therapy on myocardial performance in early postmenopausal women, Climacteric, 13 (2010) 157-170.
[31] Y.S. Peng, H.C. Ding, Y.T. Lin, J.P. Syu, Y. Chen, S.M. Wang, Uremic toxin p-cresol induces disassembly of gap junctions of cardiomyocytes, Toxicology, 302 (2012) 11-17.
[32] J.A. Tumlin, M.R. Costanzo, L.S. Chawla, C.A. Herzog, J.A. Kellum, P.A. McCullough, C. Ronco, Cardiorenal syndrome type 4: insights on clinical presentation and pathophysiology from the eleventh consensus conference of the Acute Dialysis Quality Initiative (ADQI), Contrib Nephrol, 182 (2013) 158-173.
[33] P. Brunet, B. Gondouin, A. Duval-Sabatier, L. Dou, C. Cerini, F. Dignat-George, N. Jourde-Chiche, A. Argiles, S. Burtey, Does uremia cause vascular dysfunction?, Kidney Blood Press Res, 34 (2011) 284-290.
[34] H.A. Mutsaers, L.P. van den Heuvel, L.H. Ringens, A.C. Dankers, F.G. Russel, J.F. Wetzels, J.G. Hoenderop, R. Masereeuw, Uremic toxins inhibit transport by breast cancer resistance protein and multidrug resistance protein 4 at clinically relevant concentrations, PLoS One, 6 (2011) e18438.
[35] H.A. Mutsaers, U.F. Engelke, M.J. Wilmer, J.F. Wetzels, R.A. Wevers, L.P. van den Heuvel, J.G. Hoenderop, R. Masereeuw, Optimized metabolomic approach to identify uremic solutes in plasma of stage 3-4 chronic kidney disease patients, PLoS One, 8 (2013) e71199.
[36] O. Deltombe, W. Van Biesen, G. Glorieux, Z. Massy, A. Dhondt, S. Eloot, Exploring Protein Binding of Uremic Toxins in Patients with Different Stages of Chronic Kidney Disease and during Hemodialysis, Toxins (Basel), 7 (2015) 3933-3946.
[37] P.R. Gajjala, M. Sanati, J. Jankowski, Cellular and Molecular Mechanisms of Chronic Kidney Disease with Diabetes Mellitus and Cardiovascular Diseases as Its Comorbidities, Front Immunol, 6 (2015) 340.
[38] S. Ito, M. Yoshida, Protein-bound uremic toxins: new culprits of cardiovascular events in chronic kidney disease patients, Toxins (Basel), 6 (2014) 665-678.
[39] C.W. Usselman, N.S. Stachenfeld, J.R. Bender, The molecular actions of estrogen in the regulation of vascular health, Exp Physiol, (2016).

[40] S. Menazza, E. Murphy, The Expanding Complexity of Estrogen Receptor Signaling in the Cardiovascular System, Circ Res, 118 (2016) 994-1007.
[41] Q.K. Tran, R. Firkins, J. Giles, S. Francis, V. Matnishian, P. Tran, M. VerMeer, J. Jasurda, M.A. Burgard, B. Gebert-Oberle, Estrogen Enhances Linkage in the Vascular Endothelial Calmodulin Network via a Feedforward Mechanism at the G Protein-coupled Estrogen Receptor 1, J Biol Chem, 291 (2016) 10805-10823.
[42] A. Blum, Endoethelial Progenitor Cells are affected by Meidcations and Estrogen, The Israel Medical Association Journal, 17 (2015).
[43] M.K. Osako, H. Nakagami, N. Koibuchi, H. Shimizu, F. Nakagami, H. Koriyama, M. Shimamura, T. Miyake, H. Rakugi, R. Morishita, Estrogen inhibits vascular calcification via vascular RANKL system: common mechanism of osteoporosis and vascular calcification, Circ Res, 107 (2010) 466-475.
[44] H. Salve, S. Mahajan, P. Misra, Prevalence of chronic kidney diseases and its determinants among perimenopausal women in a rural area of North India: A community-based study, Indian Journal of Nephrology, (2012).
[45] E. Vitolo, M. Comassi, M.T. Caputo, A. Solini, Hormone replacement therapy, renal function and heart ultrasonographic parameters in postmenopausal women: an observational study, Int J Clin Pract, 69 (2015) 632-637.
[46] E. Ricciotti, G.A. FitzGerald, Prostaglandins and inflammation, Arterioscler Thromb Vasc Biol, 31 (2011) 986-1000.
[47] J.M. Cook-Mills, T.L. Deem, Active participation of endothelial cells in inflammation, J Leukoc Biol, 77 (2005) 487-495.
[48] N. Hadad, L. Tuval, V. Elgazar-Carmom, R. Levy, R. Levy, Endothelial ICAM-1 protein induction is regulated by cytosolic phospholipase A2alpha via both NF-kappaB and CREB transcription factors, J Immunol, 186 (2011) 1816-1827.
[49] D.A. Long, J.T. Norman, L.G. Fine, Restoring the renal microvasculature to treat chronic kidney disease, Nat Rev Nephrol, 8 (2012) 244-250.
[50] H. Xu, P. Czerwinski, M. Hortmann, H.Y. Sohn, U. Forstermann, H. Li, Protein kinase C alpha promotes angiogenic activity of human endothelial cells via induction of vascular endothelial growth factor, Cardiovasc Res, 78 (2008) 349-355.
[51] K.S. RUSSELL, D.F. STERN, P.J. POLVERINI, J.R. BENDER, Neuregulin activation of ErbB receptors in vascular endothelium leads to angiogenesis, The American Physiological Society, (1999).
[52] S. Kumar, N. Rao, R. Ge, Emerging Roles of ADAMTSs in Angiogenesis and Cancer, Cancers (Basel), 4 (2012) 1252-1299.
[53] M. Mittal, M.R. Siddiqui, K. Tran, S.P. Reddy, A.B. Malik, Reactive oxygen species in inflammation and tissue injury, Antioxid Redox Signal, 20 (2014) 1126-1167.
[54] W.F. JACKSON, Potassium Channels in the Peripheral Microcirculation, Microcirculation, (2005).
[55] A.H. Sprague, R.A. Khalil, Inflammatory cytokines in vascular dysfunction and vascular disease, Biochem Pharmacol, 78 (2009) 539-552.
[56] C. Tschope, S. Van Linthout, New insights in (inter)cellular mechanisms by heart failure with preserved ejection fraction, Curr Heart Fail Rep, 11 (2014) 436-444.
[57] M.J. Park, H.J. Kwak, H.C. Lee, D.H. Yoo, I.C. Park, M.S. Kim, S.H. Lee, C.H. Rhee, S.I. Hong, Nerve growth factor induces endothelial cell invasion and cord formation by promoting matrix metalloproteinase-2 expression through the phosphatidylinositol 3-kinase/Akt signaling pathway and AP-2 transcription factor, J Biol Chem, 282 (2007) 30485-30496.
[58] E. de Jonge, P.W. Friederich, G.P. Vlasuk, W.E. Rote, M.B. Vroom, M. Levi, T. van der Poll, Activation of Coagulation by Administration of Recombinant Factor VIIa Elicits Interleukin 6 (IL-6) and IL-8 Release in Healthy Human Subjects, Clinical and Vaccine Immunology, 10 (2003) 495-497.
[59] T. Browder, J. Folkman, S. Steven Pirie-Shepherd, The Hemostatic System as a Regulator of Angiogenesis, BIOLOGICAL CHEMISTRY, 275 (2000).
[60] B.M. Doshi, G.A. Perdrizet, L.E. Hightower, Wound healing from a cellular stress response perspective, Cell Stress Chaperones, 13 (2008) 393-399.

[61] K.K. Koh, B.K. Yoon, Controversies regarding hormone therapy: Insights from inflammation and hemostasis, Cardiovasc Res, 70 (2006) 22-30.
[62] A.J. Gale, Current understanding of hemostasis, Toxicol Pathol, 39 (2011) 273-280.
[63] A.S. Sohal, A.S. Gangji, M.A. Crowther, D. Treleaven, Uremic bleeding: pathophysiology and clinical risk factors, Thromb Res, 118 (2006) 417-422.
[64] S.J. Hedges, S.B. Dehoney, J.S. Hooper, J. Amanzadeh, A.J. Busti, Evidence-based treatment recommendations for uremic bleeding, Nat Clin Pract Neph, 3 (2007) 138-153.
[65] S. Aiello, M. Noris, M. Todeschini, S. Zappella, C. Foglieni, A. Benigni, D. Corna, C. Zoja, D. Cavallotti, G. Remuzzi, Renal and systemic nitric oxide synthesis in rats with renal mass reduction, Kidney International, 52 (1997) 171-181.
[66] D. Simon, J. Stamler, E. Loh, J. Loscalzo, S. Francis, M. Creager, Effect of nitric oxide synthase inhibition on bleeding time in humans, Journal Cardiovasc Pharmacology, 26 (1995).
[67] M. Noris, G. Remuzzi, Uremic bleeding closing the circle after 30 years of controversies, Blood, 94 (1999).
[68] Zoja C, Noris M, Corna D, Viganò G, Perico N, de Gaetano G, R. G., L-arginine, the precursor of nitric oxide, abolishes the effect of estrogens on bleeding time in experimental uremia, Lab Invest., (1991).
[69] K. Niessen, A. Karsan, Notch signaling in cardiac development, Circ Res, 102 (2008) 1169-1181.
[70] A. Ozkan, Low gradient "severe" aortic stenosis with preserved left ventricular ejection fraction, Cardiovasc Diagn Ther, 2 (2012) 19-27.
[71] R.V. Freeman, C.M. Otto, Spectrum of calcific aortic valve disease: pathogenesis, disease progression, and treatment strategies, Circulation, 111 (2005) 3316-3326.
[72] H.M. Norum, L. Gullestad, A. Abraityte, K. Broch, S. Aakhus, P. Aukrust, T. Ueland, Increased Serum Levels of the Notch Ligand DLL1 Are Associated With Diastolic Dysfunction, Reduced Exercise Capacity, and Adverse Outcome in Chronic Heart Failure, J Card Fail, 22 (2016) 218-223.
[73] E.M. Parodi, B. Kuhn, Signalling between microvascular endothelium and cardiomyocytes through neuregulin, Cardiovasc Res, 102 (2014) 194-204.
[74] S.L. House, A.M. Castro, T.S. Lupu, C. Weinheimer, C. Smith, A. Kovacs, D.M. Ornitz, Endothelial fibroblast growth factor receptor signaling is required for vascular remodeling following cardiac ischemia-reperfusion injury, Am J Physiol Heart Circ Physiol, 310 (2016) H559-571.

Chapter 3

Indoxyl sulfate stimulates angiogenesis by regulating reactive oxygen species production via CYP1B1

Jiayi Pei[1,2,3], Rio Juni[4], Magdalena Harakalova[2,5], Dirk J Duncker[6], Folkert W Asselbergs[2,7,8], Pieter Koolwijk[4], Victor van Hinsbergh[4], Marianne C Verhaar[1,3], Michal Mokry[3,9,10], Caroline Cheng[11,12]

[1]Department of Nephrology and Hypertension, DIGD, UMC Utrecht, University of Utrecht, 3584 CX Utrecht, The Netherlands
[2]Department of Cardiology, Division Heart & Lungs, UMC Utrecht, University of Utrecht, 3584 CX Utrecht, The Netherlands
[3]Regenerative Medicine Utrecht, UMC Utrecht, University of Utrecht, 3584 CX Utrecht, The Netherlands
[4]Department of Physiology, Amsterdam UMC, VUmc location, Amsterdam Cardiovascular Science, 1081 HV Amsterdam, The Netherlands
[5]Department of Pathology, UMC Utrecht, University of Utrecht, 3584 CX Utrecht, The Netherlands.
[6]Department of Cardiology, Erasmus MC, Rotterdam, 3015 GD, The Netherlands
[7]Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College London, London, NW1 2DA, UK
[8]Health Data Research UK and Institute of Health Informatics, University College London, London, NW1 2DA, UK
9Department of Clinical Chemistry and Heamatology, University of Utrecht, 3584 CX Utrecht, The Netherlands
[10]Division of Paediatrics, UMC Utrecht, University of Utrecht, 3584 CX Utrecht, The Netherlands
[11]Department of Nephrology and Hypertension, DIGD, UMC Utrecht, University of Utrecht, 3584 CX Utrecht, The Netherlands
[12]Regenerative Medicine Utrecht, UMC Utrecht, University of Utrecht, 3584 CX Utrecht, The Netherlands

Toxins (Basel)

DOI: 10.3390/toxins11080454

Abstract

Indoxyl sulfate (IS) is an accumulative protein-bound uremic toxin found in patients with kidney disease. It is reported that IS impairs the vascular endothelium, but a comprehensive overview of all mechanisms active in IS-injury remains currently lacking. Here we performed RNA sequencing in human umbilical vein endothelial cells (HUVECs) after IS or control medium treatment and identified 1,293 genes that were affected in a IS-induced response. Gene enrichment analysis highlighted pathways involved in altered vascular formation and cell metabolism. We confirmed these transcriptome profiles at the functional level by demonstrating decreased viability and increased cell senescence in response to IS treatment. In line with the additional pathways highlighted by the transcriptome analysis, we further could demonstrate that IS exposure of HUVECs promoted tubule formation as shown by the increase in total tubule length in a 3D HUVECs/pericytes co-culture assay. Notably, the pro-angiogenic response of IS and increased ROS production were abolished when CYP1B1, one of the main target genes that was highly upregulated by IS, was silenced. This observation indicates IS-induced ROS in endothelial cells is CYP1B1-dependent. Taken together, our findings demonstrate that IS promotes angiogenesis and CYP1B1 is an important factor in IS-activated angiogenic response.

Keywords: indoxyl sulfate; chronic kidney disease; reactive oxygen species; CYP1B1; angiogenesis

1. Introduction

Indoxyl sulfate (IS) is a uremic retention solute that accumulates in the systemic circulation due to renal impairment [1]. Unlike other uremic retention solutes that are water-soluble or non-protein-bound, IS binds to albumin (66.5kDA) and cannot be cleared effectively via dialysis, which is the main method to remove uremic toxins in end-stage chronic kidney disease (CKD) patients [1]. Compared to healthy individuals, serum IS levels are nearly 50 times higher in patients with acute kidney injury (AKI) and reach the highest level in patients with end-stage CKD [2]. Circulating IS has been shown to play an important role in the progression of CKD and the development of cardiac disease, such as left ventricular hypertrophy [3]. Although treatment with the carbonaceous adsorbent AST-120 to lower serum IS level showed improvement on renal and cardiac function in both animal models and a phase II study, it failed to demonstrate promising results in the subsequent phase III trial [4, 5]. More studies based on a genome wide analysis approach could shed new light on the working mechanism of IS associated cardiorenal disease and provide new targets for the development of new therapeutic approaches.

Studies have implied that IS first impairs endothelial function, which subsequently contributes to the worsening of kidney function and the development of cardiovascular disease [6]. Flow-mediated endothelial dilation (FMD), which is a clinical parameter for endothelial function, is significantly lower in CKD patients when compared to healthy individuals and hypertensive patients [7]. Notably, by lowering IS level, FMD increases in CKD patients and correlates inversely with IS levels. *In vitro* studies have revealed that IS inhibits nitric oxide (NO) production, which is a critical regulator of vascular tone, while it promotes reactive oxygen species (ROS) release, resulting in oxidative stress [7-9]. Furthermore, IS inhibits the proliferation ability of endothelial cells (ECs) by activating aryl hydrocarbon receptor-mediated cell senescence [10]. Besides the direct deleterious effect of IS on ECs, IS also interferes with the immune system and actives inflammatory cytokines, such as IL-1β, E-selectin and TNF-α, which contribute to EC apoptosis and result in endothelial dysfunction [11, 12]. Thus IS appears to have a broad and complex effect on endothelial function. A genome-wide transcriptome study would aid in mapping the key driving signalling factor(s) underlying this important IS-induced disease mechanism.

In the present study, we conducted a genome-wide transcriptome analysis using RNA sequencing (RNA-seq) to reveal the transcriptome profile of IS-treated human umbilical vein endothelial cells (HUVECs). Based on the set of IS influenced genes, we performed gene enrichment analyses and obtained indications on altered pathways involved in cell migration, angiogenesis, apoptosis and cell metabolism. We studied these enriched biological processes at the functional level using an established 3D collagen-based *in vitro* model for angiogenesis [13, 14]. An activated angiogenic

response was observed under IS stimulation. Furthermore, cytochrome P450 1B1 (*CYP1B1*) was identified as one of the strongest up-regulated genes in IS-treated HUVECs. Silencing of *CYP1B1* decreased IS-induced ROS and attenuated angiogenic response under IS stimulation, implying a role for *CYP1B1*-dependent ROS production in the IS-induced angiogenic response.

2. Results

2.1.RNA-seq reveals differentially expressed genes in IS treated HUVECs

We studied transcriptome changes in HUVECs after 24 hr stimulation of 250 μM IS as compared to the potassium salt (KCl) control using RNA-seq. Heatmap depicts clustering of samples based on all differentially expressed genes between the two groups (Figure 1A). Volcano plot shows both fold change and p-value of all genes in log 2 scale, and differentially expressed genes were highlighted in red (Figure 1B). In total, we identified 1,293 genes that were differentially expressed between KCl control and IS groups (p-value<0.05, Table 1, Supplementary Table 1). Of these, 643 genes were up-regulated by IS as compared to the control group, and gene enrichment analysis showed that they were mostly involved in cell migration, angiogenesis and programmed cell death processes (Figure 1C); 650 genes were down-regulated by IS, and these were mainly enriched for biological processes related to cell metabolism, such as cell cycle process, chromosome segregation, and cell division (Figure 1D).

2.2. IS inhibits cell viability at high concentration

As the transcriptome profile indicated that apoptosis was enhanced by IS in HUVECs, we performed MTT assay to examine the viability of IS-treated HUVECs. IS at 250 μM decreased cell viability as compared to the control, although the decrease was not significant (p-value=0.07, Figure 2A). A significant decrease of cell viability was achieved in 500 μM and 750 μM IS treated HUVECs (p-value<0.05, Figure 2A).

2.3. IS induces cell senescence

The transcriptome data also indicated that cell cycle progression was impeded by IS exposure. For validation, we examined IS-induced senescence features. HUVECs were incubated with 250 μM IS or control buffer for 24 hr incubation before the X-gal assay for senescent cell detection. We observed more X-gal positive cells (blue) in IS group when compared to the control (Figure 2B). In line with this finding, the expression level of cell senescence marker *CDKN1A* was significantly higher in IS group

when compared to the KCl control, whereas the expression level of cell proliferation marker *KI67* was significantly lower in IS group, as shown by RT-qPCR validation (Figure 2C).

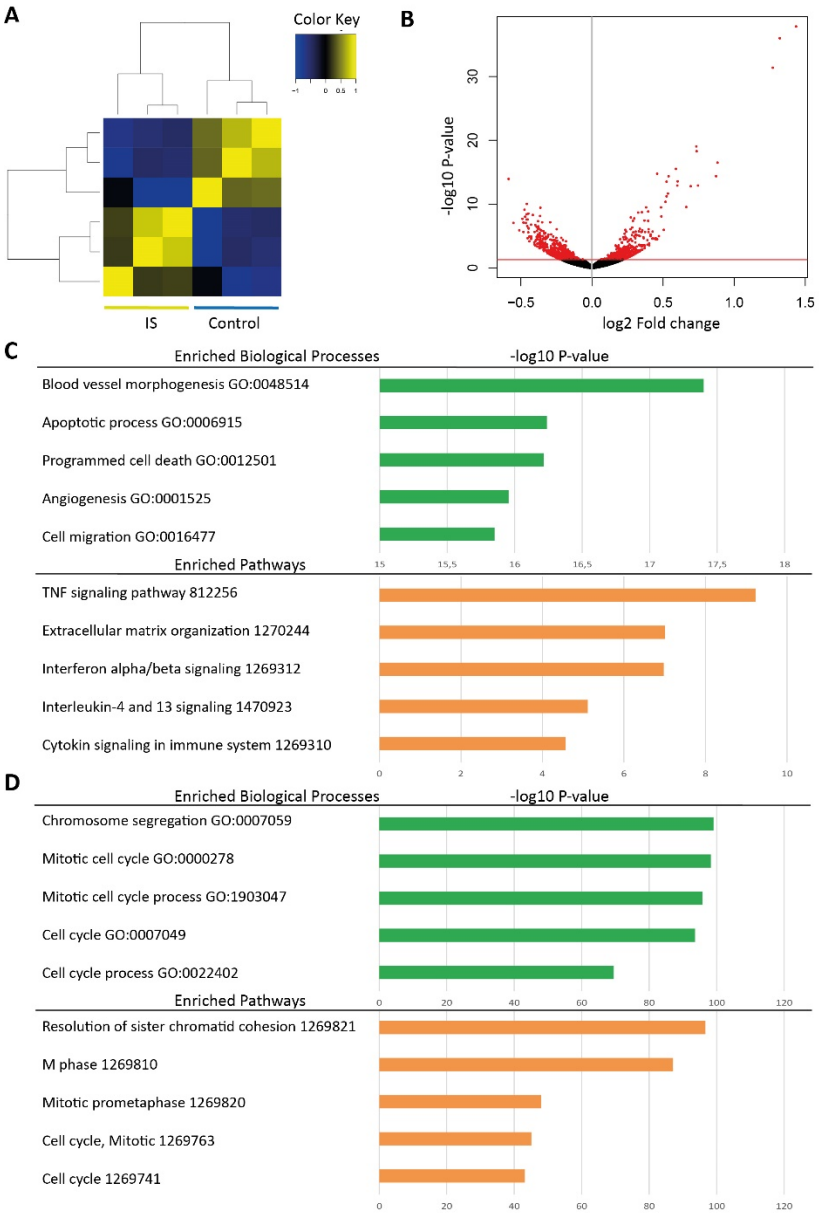


Figure 1. Comparison of the transcriptome profile of HUVECs treated with IS to KCl treated control groups. (A) Heatmap depicts clustering of samples based on all differentially expressed genes between two groups. (B) Volcano plot presents fold change (x-axis) and p-value (y-axis) of all genes in log 2 scale. Differentially expressed genes are shown in red. (C) Top 5 Enriched biological processes (green) and pathways (orange) based on IS upregulated genes. (D) Top 5 Enriched biological processes (green) and pathways (orange) based on IS downregulated genes.

Table 1. Top 10 genes that were significantly up- or down-regulated in IS-treated HUVECs when compared to the KCl treated control.

Category	Ensembl ID	Gene symbol	Gene name	Angiogenic function†	Fold change (log2)	P-value
Up-regulation	ENSG00000138061	*CYP1B1*	Cytochrome P450 Family 1 Subfamily B Member 1	Promote angiogenesis	1.434	1.390E-38
	ENSG00000114812	*VIPR1*	Vasoactive Intestinal Peptide Receptor 1	Not known	1.321	1.072E-36
	ENSG00000137809	*ITGA11*	Integrin Subunit Alpha 11	Not known	1.272	4.190E-32
	ENSG00000178695	*KCTD12*	Potassium Channel Tetramerization Domain Containing 12	Not known	0.883	3.247E-17
	ENSG00000063438	*AHRR*	Aryl-Hydrocarbon Receptor Repressor	Not known	0.872	4.144E-15
	ENSG00000007908	*SELE*	Selectin E	Not known	0.746	1.240E-13
	ENSG00000137331	*IER3*	Immediate Early Response 3	Not known	0.736	5.396E-19
	ENSG00000163659	*TIPARP*	TCDD Inducible Poly(ADP-Ribose) Polymerase	Not known	0.734	8.865E-20
	ENSG00000144476	*ACKR3*	Atypical Chemokine Receptor 3	Promote angiogenesis	0.695	1.596E-13
	ENSG00000144802	*NFKBIZ*	NFKB Inhibitor Zeta	Not known	0.663	3.022E-10
Down-regulation	ENSG00000117724	*CENPF*	Centromere Protein F	Not known	-0.586	1.114E-14
	ENSG00000143476	*DTL*	Denticleless E3 Ubiquitin Protein Ligase Homolog	Not known	-0.554	8.619E-08
	ENSG00000163808	*KIF15*	Kinesin Family Member 15	Not known	-0.510	1.195E-06
	ENSG00000138778	*CENPE*	Centromere Protein E	Not known	-0.490	1.912E-08
	ENSG00000137812	*CASC5*	Kinetochore Scaffold 1	Not known	-0.483	8.198E-08
	ENSG00000184661	*CDCA2*	Cell Division Cycle Associated 2	Not known	-0.482	2.314E-06
	ENSG00000066279	*ASPM*	Abnormal Spindle Microtubule Assembly	Not known	-0.473	7.741E-10
	ENSG00000156802	*ATAD2*	ATPase Family AAA Domain Containing 2	Inhibit angiogenesis	-0.470	9.314E-09
	ENSG00000196549	*MME*	Membrane Metalloendopeptidase	Not known	-0.465	9.572E-08
	ENSG00000132646	*PCNA*	Proliferating Cell Nuclear Antigen	Not known	-0.461	3.809E-09

†: Established biological function of each gene in relation to angiogenesis was collected from NCBI Gene (https://www.ncbi.nlm.nih.gov/gene).

2.4. IS does not influence cell migration ability

Transcriptome data implied that IS enriched transcripts of genes were involved in cell migration. We used both invasive wound healing assay and non-invasive plug assay to study the influence of IS

on the cell migration capacity. After 24 hr incubation with either IS or KCl control buffer, the migration distances of HUVECs were comparable between two groups at three difference concentrations using the wound healing assay (Fig. 2D and 2E). Likewise, no difference was detected between the two groups on the number of migrated HUVECs into the cell-free area using the plug assay (Figure 2F and 2G).

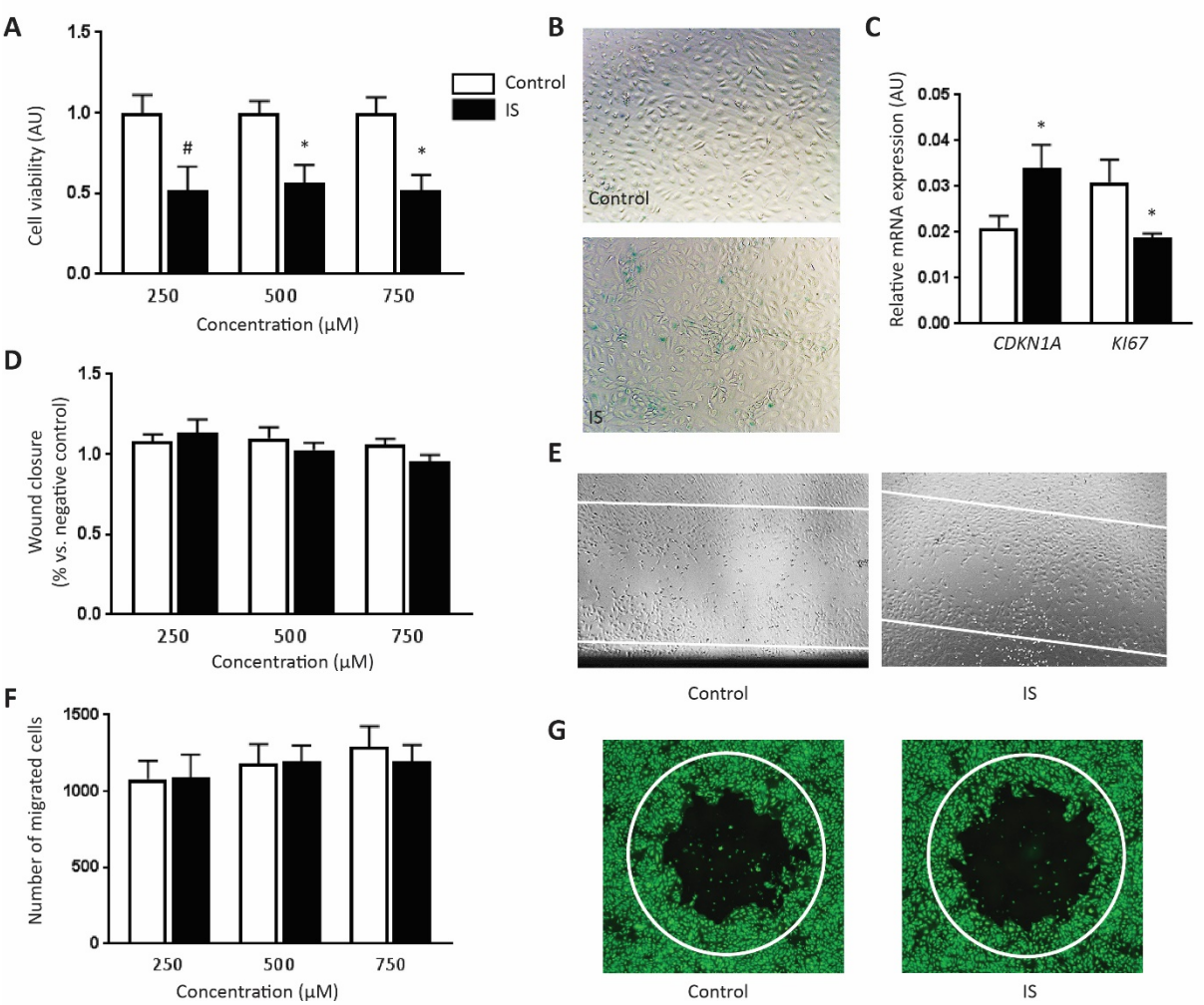


Figure 2. Functional examination of the identified IS related processes: Assessment of cell viability, senescence and migration in response to IS. (**A**) Cell viability was studied using a MTT assay. IS diminished viability of HUVECs when compared to the control at three difference concentrations (n=4). (**B**) A representative image of X-gal activity in HUVECs treated with 250 μM IS or KCl control buffer, at 10x magnification. More X-gal positive cells (blue) were observed in IS group when compared to the control. (**C**) RT-qPCR results showed a higher expression level of cell senescence marker CDKN1A and a lower expression level of cell proliferation marker KI67 in HUVECs after exposing to 250 μM IS when compared to the KCl control (n≥5). (**D**) An invasive wound healing assay was performed to study the influence of IS on cell migration ability. No difference on the migration distances of HUVECs was shown between two groups at three difference concentrations (n=3). (**E**) Examples of migrated HUVECs after exposure to 250 μM IS or KCl control after 24 hr in the wound healing assay, at 4x magnification. (**F**) A non-invasive plug assay was also performed to study the influence of IS on cell migration ability. No difference on the migrated HUVECs into the cell-free area was shown between two groups at three difference concentrations (n=6). (**G**) Examples of migrated HUVECs after exposure to 250 μM IS or KCl control after 24 hr in the plug assay. 2x magnification was used. All values are presented as mean ± SEM and they are shown in

arbitrary units (AU), #p-value<0.1, *p-value<0.05. White lines indicate migration area in wound healing and plug assay.

2.5. IS promotes angiogenic response

The main biological process enriched in upregulated genes of the IS response was blood vessel morphogenesis. To validate this, we used a 3D collagen-based co-culture model to study the influence of IS on endothelial reorganization and tubule formation. In this assay, HUVECs-GFP cells and pericytes-DsRED were cultured together in type I collagen and incubated with IS or control buffer for 3 days. In control conditions, these vascular cells will undergo EC sprouting, tubule formation and pericytes-induced stabilization of neovascular structures in 3-5 days. Compared to the standard culture medium, the KCl-adjusted control medium did not affect formation of vascular structures. No difference on the number of branches, the number of tubules and total tubule length was detected between IS and KCl control treated co-cultures at day 1 post-stimulation. At day 3 both the number of branches and tubules in IS groups were higher than the controls (p-value=0.063 and p-value=0.062 respectively). The total tubule length in IS group was significantly higher compared to control (p-value<0.05, Figure 3A and 3B). Combined, these *in vitro* data, except for the migration assays, confirm the findings from transcriptome analysis, and demonstrate the complex effects of IS on ECs homeostasis and regenerative capacity.

2.6. Depletion of CYP1B1 inhibits tubular formation

CYP1B1 showed the highest fold change among IS up-regulated genes based on the RNA-seq data. To investigate the role of *CYP1B1* in IS-treated HUVECs, we first validated its expression level using RT-qPCR and obtained a consistent result (Figure 4A). Compared to the siSham transfected HUVECs, a significant lower mRNA expression level of *CYP1B1* was observed in *CYP1B1* silenced HUVECs 3 days post transfection (Figure 4B). Next, we investigated the possible involvement of CYP1B1 in relation to IS-induced angiogenesis in the previously described co-culture assay. After exposure to 250 µM IS, the number of branches, the number of tubules and total tubule length were significantly lower in *CYP1B1* silenced HUVECs when compared to siSham transfected cells at day 3, and remained highly suppressed at day 4 after IS stimulation (Figure 4C and 4D).

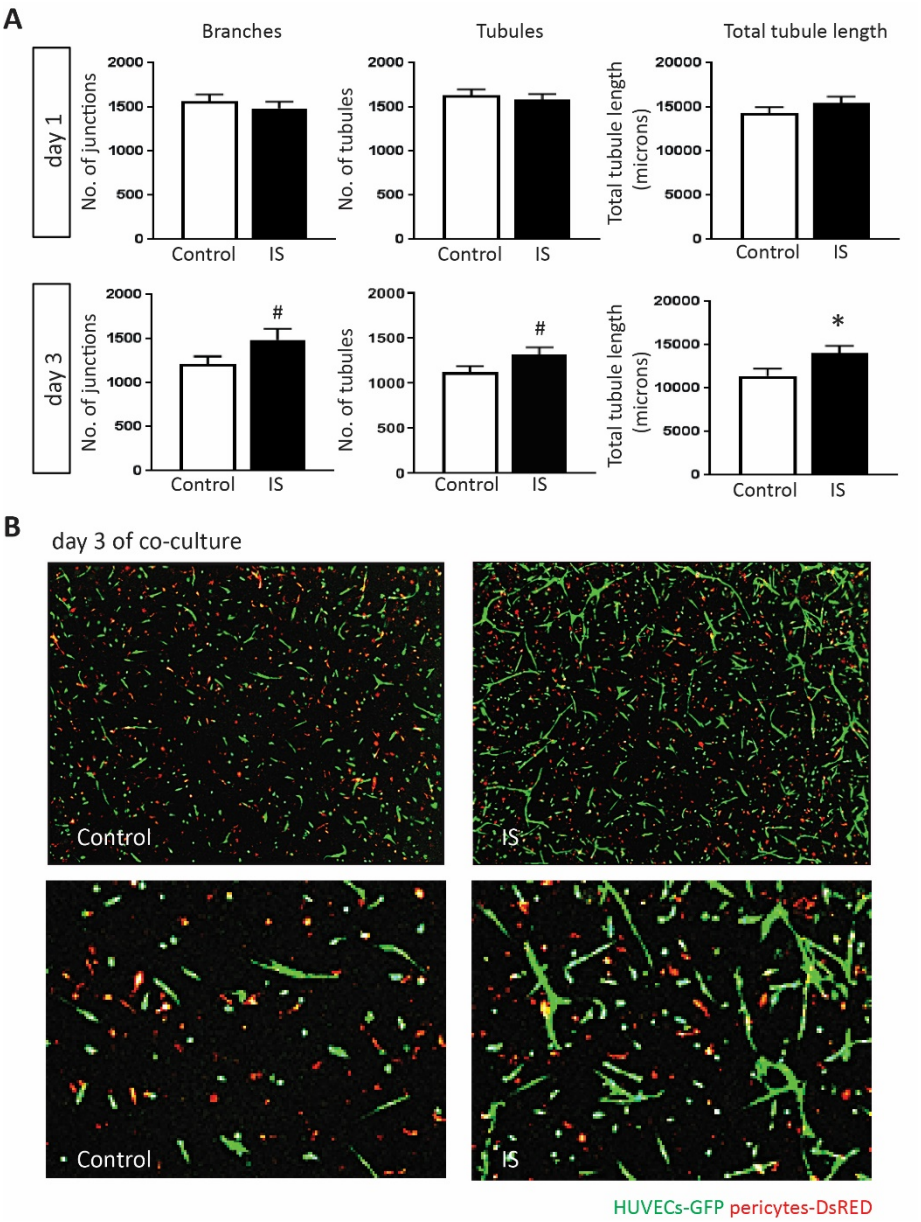


Figure 3. Examination of IS affected angiogenic responses. (**A**) A 3D co-culture model, in which HUVECs-GFP cells and pericytes-DsRED were cultured together in type I collagen, was used to study the influence of IS on angiogenesis at 250 µM. No difference on the number of branches, the number of tubules, and the total tubule length was detected between the two groups at 24 hr after incubation. After 3 days, IS showed a tendency to promote angiogenesis by increasing all three parameters when compared to the KCl control (n≥25). (**B**) Confocal images showing representative examples of vascular formation at day 3 post 250 µM IS or KCl control stimulation. Images shown in the upper row were taken at 20x magnification and zoomed-in views are shown in the lower row. In red are shown DsRED marked pericytes. In green are shown GFP marked HUVECs. All values are mean ± SEM, #p-value<0.1, *p-value<0.05.

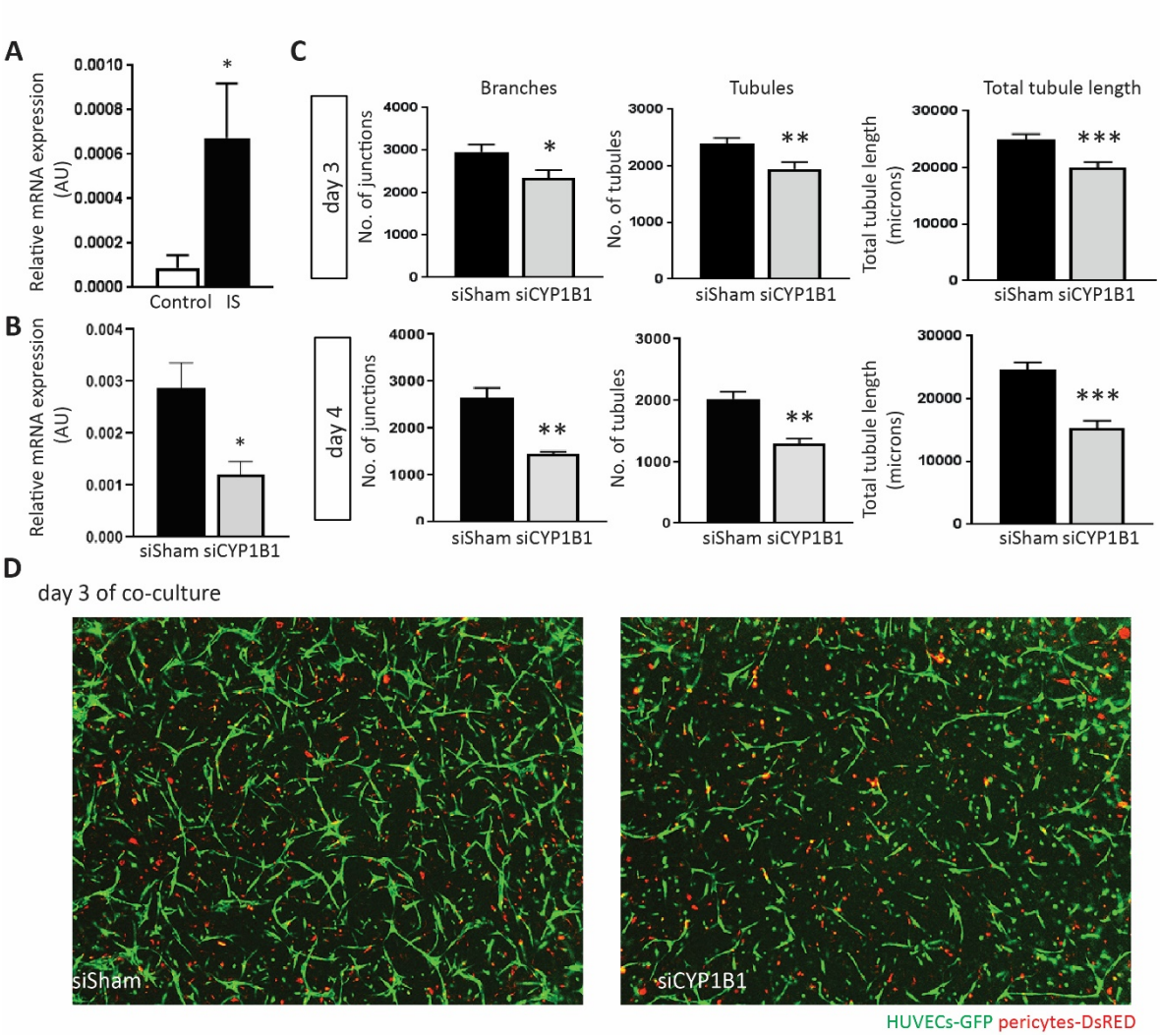


Figure 4. Investigation of increased *CYP1B1* expression in IS influenced angiogenic responses. (**A**) Bar graphs show results of RT-qPCR evaluation of CYP1B1 expression level in HUVECs treated with 250 µM IS compared to the KCl controls (n=6). (**B**) Bar graphs show results of RT-qPCR evaluation of CYP1B1 expression level in *CYP1B1* silenced HUVECs when compared to siSham transfected cells at day 3 post transfection (n=4). (**C**) Bar graphs show the number of branches, the number of tubules, and the total tubule length in *CYP1B1* targeting siRNA transfected HUVECs (siCYP1B1) compared to sham transfected HUVECs (siSham) 3 days and 4 days after 250 µM IS stimulation (n≥3). (**D**) Examples of vascular formation at day 3 post 250 µM IS in siSham and siCYP1B1 treated HUVECs. In red are shown DsRED marked pericytes. In green are shown GFP marked HUVECs. 20x magnification was used. All values are presented as mean ± SEM and they are shown in arbitrary units (AU), *p-value<0.05, **p-value<0.01, ***p-value<0.001.

2.7. CYP1B1 plays an important role in IS-increased ROS production

ROS plays an important role in the induction of endothelial dysfunction and has been shown to trigger the angiogenic response [15]. Next, we studied the possible involvement of *CYP1B1* in IS-induced ROS production. In line with previous reports, IS-treated HUVECs showed significantly higher ROS level when compared to the control at three different concentrations (p-value<0.05, Figure 5A and 5B), indicative of enhanced cellular oxidative stress. However, HUVECs transfected with siRNA targeting *CYP1B1* transcripts demonstrated a decrease in ROS production compared to control sisham

transfected groups at three difference concentrations (Figure 5C and 5D). Additionally, we further investigate this effect in cardiac microvascular endothelial cells (CMECs) to access whether it holds true in the arterial vascular bed, in particular cardiac microcirculation. Notably, we also observed a significantly increase of ROS production in IS-treated CMECs, which was attenuated after silencing *CYP1B1* (Fig. 5E). Combined, these data indicate a regulatory role of *CYP1B1* in endothelial ROS production.

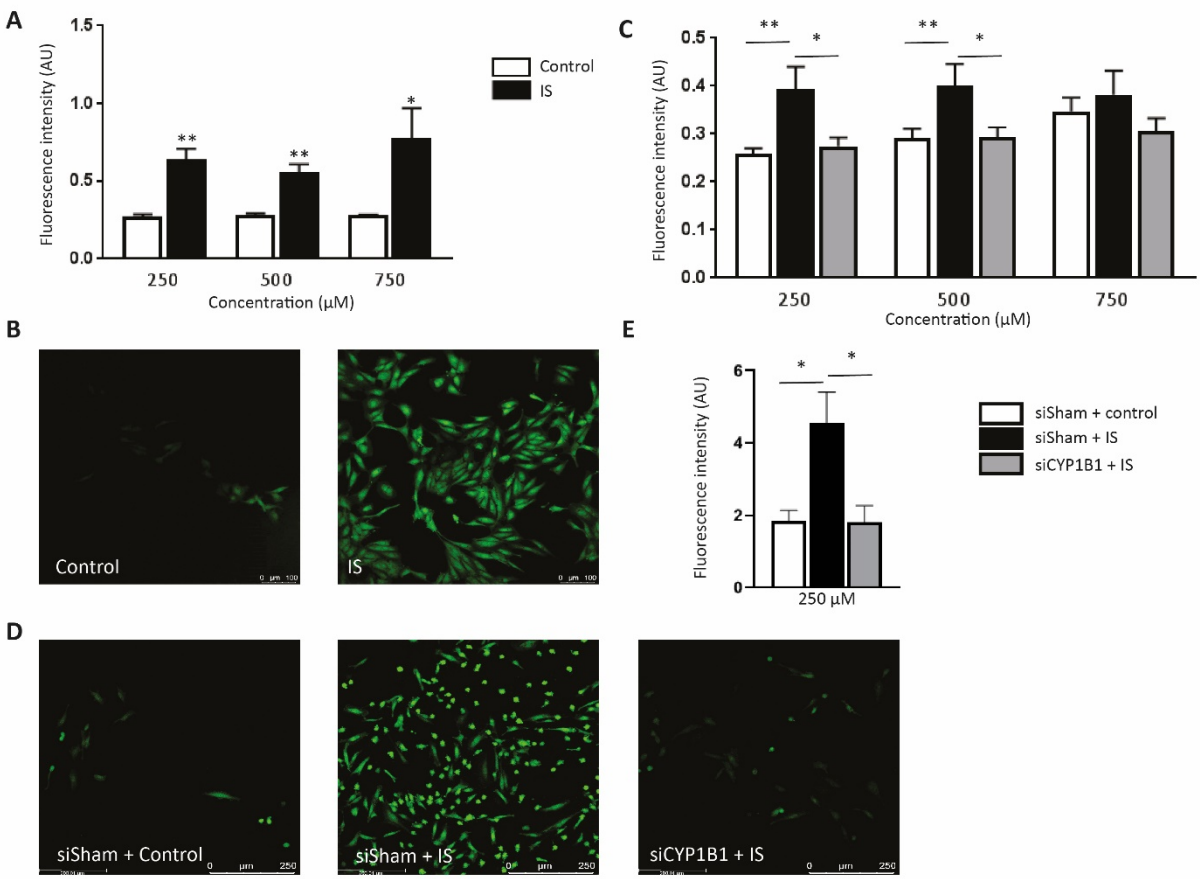


Figure 5. Detection of intracellular ROS production and the involvement of CYP1B1 in IS influenced ROS level. (**A**) IS or control treated HUVECs were loaded with CM-H2DCFDA, a fluorescent indicator for ROS. Bargraphs show detected fluorescent signal representing ROS production in IS group when compared to the KCl control at three difference concentrations (n=4). (**B**) Confocal figures showing representative examples of ROS signals in HUVECs treated with 250 μM IS compared to the KCl controls. (**C**) ROS signal measured in HUVECs treated with siRNA targeting CYP1B1 (siCYP1B1) versus non-targeting siRNA shams (siSham) with different doses of IS or equivalent doses of KCl control stimulation. (n≥9 for each group). (**D**) Confocal figures showing typical examples of ROS signals in siSham or siCYP1B1 treated HUVECs at 24hr after exposure to 250 μM IS or KCl control. 20x magnification was used. (**E**) ROS signal measured in CMECs treated with siCYP1B1 versus siSham (n=3) with or without 250 μM IS stimulation. All values are mean ± SEM and they are shown in arbitrary units (AU), *p-value<0.05, **p-value<0.01.

3. Discussion

In the present study, we demonstrated using whole genome transcriptome analysis that the IS affected genes in ECs were mostly enriched in biological functions related to vascular formation, cell

apoptosis, and cell cycle. Consistent with previous studies, we validated in *in vitro* assays that IS indeed induced an EC phenotype with reduced cell viability and increased activated cellular senescence [9, 10]. Paradoxically, we also observed enhanced angiogenic capacity of vascular cells in our 3D co-culture system in response to IS stimulation, which was in line with our transcriptome findings. We identified CYP1B1 as a new downstream target of IS and demonstrated the pro-angiogenic effect of IS was likely to be regulated via CYP1B1 modulation of endothelial ROS levels.

Dou and colleagues showed that IS decreased cell proliferation ability of HUVECs, but it did not affect cell viability at tested concentrations (from 100 µM to 1 nM) using the trypan blue exclusion test, which stains only dead cells [16]. We did not observe an effect on cell viability after 250 µM IS stimulation, however we showed that 500 µM and 750 µM impaired cell viability using the MTT assay, which more reflects metabolic activity. Dou et al. also showed that in the presence of 4% human albumin, IS decreased the wound repaired at the concentration of 125 µg/mL and 250 µg/mL. This reduction remained but was milder in IS-treated cells without the addition of albumin. Contradictory, we did not find any effect of IS on cell migration and would healing ability, which might be explained by difference in assay setups: No albumin was used in our stimulation buffer and the effect we obtained might resemble more of the non-bound IS. However, the exact effect of albumin binding of IS on endothelial cell response remains to be further evaluated.

Patients with AKI suffer from oxidative stress, during which oxygen radicals could lead to cell injury and trigger apoptosis and senescence [17, 18]. The application of antioxidants in lowering ROS level and to modulate AKI has been extensively studied and reviewed in previous studies [19]. In CKD patients, ROS level remains at a high level, especially in patients with end stage kidney disease, and has been proposed as an important mediator in CKD-associated cardiovascular diseases [20, 21]. Consistent with previous studies showing the ability of IS to induce ROS [6, 9], we also showed IS-induced ROS production. IS is a protein-bound toxin and around 90% of IS bind to plasma proteins [22, 23]. Most previous *in vitro* studies used a range from 62.5 µM to 1000 µM "free" IS [24-26] or 2 mM to 20 mM protein-bound IS [27, 28]. Additionally, 250 µM IS is comparable to the mean serum level in CKD patients [29, 30], and the maximum IS concentration in the circulation of patients is approximately 236 mg/L (939.1µM) as reported by the European Uremic Toxin Work Group [31]. Therefore, we examined IS at a broad range of concentrations from 250 µM to maximum 750 µM. Furthermore, we observed this effect in both HUVECs and CMECs, suggesting a potential role of IS-induced ROS in cardiorenal syndrome.

Excessive ROS has been shown to promote angiogenesis by inducing proangiogenic factors in ECs, such as VEGF, MMPs, ANGPT1, and VEGFRs [32]. ROS also oxidize phospholipids and the resulting oxidant products could contribute to angiogenesis via TLR signalling [15]. The transcriptome data

indicated that IS exposure significantly increased expression levels of VEGFC, MMP1, MMP24-AS1 and MMP25-AS1 in ECs. Furthermore, in our co-culture assay, we found a significant increase in total tubule length 3 days after exposing to IS, indicating IS activated a (micro)vascular angiogenic response.

A major source of ROS is cytochrome P450 activity [33]. Cytochrome P450 is a large complex of enzymes, which are actively involved in more than 70% of all drug metabolism by initiating monooxygenase or hydroxylation reaction via other substrates (i.e. oxygen and NADPH) [34]. During the reaction, P450 produces active oxygen species and subsequently contribute to excessive ROS formation [33].

CYP1B1, the biggest known human P450 protein in terms of size of mRNA and amino acids, is highly expressed in tumour cells and studies have highlighted its important role in tumour development [35]. Compared to the general population, the prevalence of cancer is higher in patients with moderate CKD and patients received dialysis or kidney transplantation [36]. Notably, McFadyen and colleagues showed a higher CYP1B1 expression in renal cell carcinoma when compared to the normal kidney [37]. Gondouin *et al.* further showed that IS increased CYP1B1 expression in HUVECs using a microarray setup [38]. Consistent with previous studies, we also demonstrated an increased in CYP1B1 expression levels in response to IS. In fact, our RNAseq based analysis showed that *CYP1B1* had the highest fold change increase among all IS-activated genes. So far, only a limited number of studies have shown the involvement of CYP1B1 in angiogenesis. Dallaglio and colleagues showed that both the RNA and protein expression levels of CYP1B1 were significantly down-regulated in HUVECs after exposing to metformin, which inhibited vascular formation *in vitro [39]*. Tang and colleagues showed that the number of retinal blood vessels was decreased in mice that lacked CYP1B1 [40]. They also confirmed that the lack of CYP1B1 impaired endothelial cell sprouting *in vitro,* which could be reversed by restoring CYP1B1 expression. Both studies used only one vascular cell type (ECs) in a 2D Matrigel based model. A later study from Palenski and colleagues examined the involvement of CYP1B1 in both retinal ECs and pericytes that were isolated from mice [41]. They also observed impaired vascular formation using the Matrigel model, in which ECs that lacked CYP1B1 were cultured with normal pericytes. Expanding on these previous findings, using our established 3D type I collagen human-derived EC and pericyte co-culture model that allows complex vascular structure formation [13, 14], we provide evidence that the enhanced angiogenic response under IS stimulation is partially mediated via CYP1B1 upregulation by IS in endothelial cells.

Interestingly, multiple studies show defective angiogenesis in CKD patients. Futrakul and colleagues showed that CKD patients had nearly 17-fold increase of circulating endothelial cells that reflected vascular injury when compared to healthy individuals. This was also linked to a decrease in VEGF/endostatin ratio that indicated a decline in angiogenic capacity [42]. A recent study included a

larger population of both CKD patients and healthy individuals, and showed a decrease in angiopoietin-1/VEGF-A ratio in CKD patients when compared to the control, indicating impaired angiogenesis and enhanced endothelial leakage [43]. In AKI, hypoxia impaired angiogenesis has also been identified, which has been proposed to contribute to the transition from AKI to CKD [44]. However, only a limited number of studies are focused on the influence of IS on angiogenesis. Hung and colleagues showed that accumulated IS in nephrectomised mice inhibited the maturation of endothelial progenitor cells (EPCs) and subsequently suppressed neovascularization [45]. By treating these mice with AST-120 that removes IS precursor indole in the intestine, they showed a decreased plasma level of IS and an improvement in the EPC-based neovascularization. Another study showed IS inhibited the chemotactic motility and the colony-forming ability of human EPCs [46]. In our study, IS activated genes were highly enriched for angiogenesis. We also showed increased branches and tubule formation at 3 days after IS stimulation *in vitro*, and the total tubule length was significantly higher in IS group. Combined, our data suggest enhanced vascular formation activity in response to IS. It is important to point out that different cell types were used in previous studies and in our study, namely EPCs and HUVECs respectively [47]. Besides, instead of vasculogenesis during which vascular formation occurs from *in situ* differentiating EPCs [48], we used a well-established model to study angiogenesis of differentiated ECs. Additionally, tubular structures seem to decline from day 3 post IS stimulation to day 4 in *CYP1B1* silenced cells, whereas it remained stable or slightly increased between day 3 and day 4 in absence of IS stimulation (Supplementary Figure 1). Further studies are required to examine the long-term effect of IS on angiogenesis. Combined, they could explain the different findings between previous studies and the present study.

IS has been shown to function as a ligand for aryl hydrocarbon receptor (AhR), and the AhR signaling is activated upon binding [49]. Notably, the expression level of AhR is relatively high in kidney [50], and the highest level of accumulated IS has also found in kidney as compared to lung, heart and liver [51]. A positive correlation between the increased activation of AhR signaling and CKD has also been shown [52], suggesting the deleterious effect of AhR signaling in kidney disease. The downstream targets of AhR signaling vary among cell types [53-56]. To investigate the regulation between IS and AhR signaling in ECs, we collected 266 established AhR targets (systematic name: M9986 and M17378) from Molecular Signature Database v6.2 and performed gene set enrichment analyses (GSEA) to examine the representation level of AhR targeted genes in differentially expressed genes from our study. We observed an over-representation of AhR downstream targets in the IS-activated genes under the default settings (FDR<25%, Supplementary Report 1 and Supplementary Figure 2), implying the activation of AhR pathway by IS in HUVECs. Taken together, these data highlight an important main regulatory pathway through which IS could negatively impact the regenerative capacity of the

renal vasculature in renal disease. More studies are required to investigate the effect of IS-mediated AhR signaling on angiogenic response in ECs.

In terms of disease, we previously reported increased capillary networks in heart and kidney of obese ZSF1 rat with cardiorenal metabolic syndrome when compared to the controls, indicating an activated angiogenic response [57]. Despite the enriched ECs foci and pericytes foci, the lack of regular vascular endothelial luminal surface and the decrease of peritubular and glomerular endothelium suggested non-functional vasculature. Furthermore, they recruited macrophages, which subsequently contributed to the fibrotic formation. It is important to note that other risk factors, such as hypertension and onset of heart failure with preserved ejection fraction, were also observed in this obese ZSF1 rat model and could interfere with the impaired vascular formation. In the light of these findings, the observed IS-activated angiogenic responses need further investigation, especially in relation to their functional activities, the paracrine signaling with the immune system and the possible influence from other metabolic risk factors.

In summary, we presented a comprehensive list of IS affected genes in ECs. Gene enrichment analyses indicated altered angiogenesis and cell metabolism. IS induced enhanced ROS production in ECs, which was CYP1B1-depedent. Furthermore, IS activated an angiogenic response in HUVECs-pericytes co-culture. CYP1B1 deficiency in ECs resulted in a suppressed angiogenic response, indicating a critical role of CYP1B1 in IS-activated angiogenesis. We hypothesize that IS induces ROS level in ECs, which initiates the activation of the observed angiogenic responses. However, the influence of a chronic status of high ROS level on the balance between pro- and anti-angiogenic factors *in vivo* is more complex and remains to be elucidated. With the transcriptome data generated in our study, we offer a detailed overview of putative functional chances in EC-behaviour in response to IS, and we identified CYP1B1 as a key regulator in the process, shedding light into the underlying mechanism of IS-regulated vascular formation and maturation.

4. Materials and Methods

4.1. Cell culture

Human umbilical vein endothelial cells (HUVECs) were cultured in EGM2 medium (Lonza) with 100 Uml^{-1} penicillin-streptomycin (PS). Pooled donor HUVEC were purchased from Lonza and used in all functional assays in this study. For the sequencing purpose, HUVECs were isolated from three newborns anonymously and obtained from the Department of Gynecology, Utrecht (The Netherlands), with the informed consent under the EPD term. Human brain vascular pericytes (ScienCell) were cultured in DMEM (Gibco) supplied with 10% FCS and 100 Uml^{-1} PS. All cells were

cultured in the gelatin-coated dishes (Greiner Bio-One) in a 5% CO_2 incubator at 37˚C. Cells between passage 3 and 8 were used in this study.

4.2. RNA-seq and data analysis

HUVECs were incubated with 250 µM IS for 24 hr. Potassium chloride (KCl) was used as a control, because IS is a potassium salt. Total RNA isolation was isolated using the RNeasy Mini Kit (Qiagen) according to the manufacturer's recommendations. Polyadenylated mRNA was further selected using Poly(A) Beads (NEXTflex™) and libraries were generated using the NEXTflex™ Rapid RNA-seq Kit (Bio Scientific). Libraries were sequenced by the Nextseq500 platform (Illumina). Sequencing data were analysed as described previously [58]. Briefly, reads were aligned to the human reference genome GRCh37 and mapped to the transcriptome. Reads per kilobase million for each refseq gene were calculated [59] and a list of differentially expressed genes between IS and control groups was obtained at p-value<0.05 [60].

4.3. Gene enrichment analysis

Differentially expressed genes were enriched for their biological functions using ToppGene Suite tool ToppFun (default setting: FDR correction, p-value cut off at 0.05 and gene limit set between and including 1 and 2000 per pathway) [61].

4.4. Cell metabolism assay

Cells were seeded to a gelatin-coated 96-well plate and incubated overnight for adhesion. Cells were washed once with PBS and incubated with either IS or control buffer for 24 hr. Afterwards, stimulation buffers were removed and cells were washed once with PBS, followed by 4 hr incubation with 100 µl MTT buffer (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide, 0.5 mg/mL, Sigma). MTT buffer was removed and 200 µl dimethyl sulfoxide (DMSO) per well was added to dissolve formed Formazan crystals. The plate was shaken gently using the IKA microplate shaker for 30 min in the dark. Absorbance was measured at 570 nm by the microplate reader (Bio-Rad). To correct for batch effects, an additional condition of cells cultured in the standard growth medium was included in each independent experiment.

4.5. Senescence-associated beta galactosidase activity

Cellular senescence was examined using the Senescence Detection Kit (Abcam). Briefly, treated HUVECs were washed once with PBS and fixed using Fixative Solution for 15 min in the incubator. Afterwards, HUVECs were washed twice with PBS and incubated with staining buffer containing 25

mg/mL X-gal overnight. TNFα (40 ng/mL) treated HUVECs were used as positive control. Images of random views were taken at 10x magnification using inverted fluorescence microscope.

4.6. Wound healing assay

Cells were seeded to a gelatin-coated 24-well plate (Greiner Bio-One) and were grown till 95% confluency in the growth medium. A scratch was made to create a cell-free area. Cells were washed once with PBS and incubated with IS or control buffer. Images were taken at 0 hr and 24 hr post stimulation. Area covered by migrated cells from the leading edge of the scratch was measured. To avoid batch effect, an additional condition of cells cultured in the standard growth medium was included in each independent experiment.

4.7. Cell migration plug assay

Cell stoppers (Oris^{TM}) were pre-inserted to a gelatin-coated 96-well plate (Greiner Bio-One) to create a cell-free area. Cells were seeded to the plate and incubated overnight for adhesion. The following day cell stoppers were removed and cells were stimulated by IS or control buffer for 24 hr. Afterwards, cells were stained by calcein-AM (BD-Bioscience) and images were taken by the fluorescence microscopy. Cells that migrated into the cell-free area were analysed.

4.8. 3D collagen co-culture assay

Lentivirus green fluorescent protein transduced HUVECs (HUVECs-GFP) and lentivirus discosoma sp. Red fluorescent protein transduced pericytes (pericytes-DsRED) were mixed at 5:1 ratio in co-culture medium, which is basal EBM medium supplied with 2% FCS, rhFGF-B, ascorbic acid and 100 Uml^{-1} PS. Cell mixture supplied with growth factors, including IL-3, SCF-1 and CXCL12 (BD Bioscience), at the volume of 300µl was added to 200 µl bovine collagen type I (Gibco). NaOH was used to adjust pH to 7.5. Cell-collagen mixture was added to 96-well plate (50µl per well). After 1 hr incubation, 100 µl EGM2 was added per well and the plate was incubated overnight. The following day IS or control buffer was added to the cells. Images were taken from day 1 till day 3 after stimulation using inverted fluorescence microscope and analysed by AngioSys 2.0. To correct for batch effects, an additional condition of cells cultured in the standard co-culture medium was included in each independent experiment.

4.9. Detection of intracellular ROS levels in HUVECs

Cells were seeded to a gelatin-coated 96-well plate and incubated overnight for adhesion. Next day, IS or control buffer was added to the cells. After 24 hr, stimulation buffer was removed and cells were exposed to 10 µM CM-H2DCFDA (Life Technologies) for 30 min in the dark. Excessive CM-

H2DCFDA was washed twice with PBS, followed by the addition of 100 µl PBS supplied with 0.2% Bovine Serum Albumin. Images were taken using the Leica SP8X confocal microscope at 20x magnification, and the fluorescence intensity was measured at the wavelength of 485 nm (excitation) and 538 nm (emission) using the fluorescent plate reader (Fluoroskan™).

4.10.Detection of intracellular ROS levels in cardiac microvascular endothelial cells (CMECs)

Besides HUVECs, we also examined the impact of IS on ROS production in cardiac microvascular endothelial cells (CMECs). The culture of CMECs (Lonza, CC-7030) and cytoplasmic reactive oxygen species measurement were performed as previously described [62]. Briefly, 6 hours after exposure to 250µm IS, CEMCs were washed once and incubated with 5µM CM-H2DCFDA (ThermoFischer, C6827) for 30min in phosphate buffer saline (Braun, 220/12257974/1110). Excessive CM-H2DCFDA was washed. Images were taken using the on a Zeiss Axiovert 200M Marianas inverted fluorescence microscope (Intelligent Imaging Innovations, Denver, CO, USA) with a 63x oil-immersion objective at 37°C and 5% CO_2 environment. All fluorescent images were corrected for background and negative controls. Quantification of all fluorescent images was performed using digital cell masking software (Slidebook 6, Intelligent Imaging Innovations).

4.11. Reverse transcription-quantitative polymerase chain reaction (RT-qPCR) analysis

RNA was isolated from HUVECs treated by 250 µM IS for 24h or control buffer. The quality and quantity of RNA was measured by a spectrophotometer (DeNoVIX). Complementary DNA was transcribed using iScript Synthesis Kit (Bio-Rad) according to the manufacturer's instructions. RT-qPCR was performed to measure mRNA level of targeted genes using SYBR-GREEN-Cycler IQ5 detection system (Bio-Rad). β-actin was used as the housekeeping gene. Primer sequences were as follows: *CYP1B1* (forward: 3'-TGATGGACGCCTTTATCCTC-5'; reverse: 5'-ACGACCTGATCCAATTCTGC-3'), *CDKN1A* (forward: 3'- GACACCACTGGAGGGTGACT-5'; reverse: 5'-ACAGGTCCACATGGTCTTCC-3'), *KI67* (forward: 3'- AAGCCCTCCAGCTCCTAGTC-5'; reverse: 5'-TCCGAAGCACCACTTCTTCT-3'), and beta-actin (forward: 3'-TCCCTGGAGAAGAGCTACGA-5'; reverse: 5'-AGCACTGTGTTGGCGTACAG-3').

4.12. Short interference RNA

HUVECs were grown to 60% confluence and transfected with 200 nM *CYP1B1*-siRNA or non-targeting Sham-siRNA (Dharmacon™) using lipofectamine according to the manufacturer's instructions. The silencing effect of *CYP1B1* expression in HUVECs was validated using RT-qPCR at 24 hr and 72 hr post transfection. Successfully transfected cells were further used for comparing the intracellular ROS production and angiogenic response between IS and control groups as explained above.

4.13. Statistical analyses

Statistical analyses were performed using GraphPad Prism 7.02. Unpaired t-test was used to evaluate the difference between treatment and control groups. One-way ANOVA was used to evaluate the difference when three groups were included. All means are reported with SEM. P-values <0.05 were considered statistically significant.

4.14. Data availability

RNA-seq data have been deposited in the National Center for Biotechnology Information Gene Expression Omnibus (GEO) and are accessible through GEO Series accession number GSE132410. Differentially expressed genes in HUVECs with or without IS stimulation is presented in the Supplementary Table 1.

Conflicts of Interest: The authors declare no conflict of interest.

Supplementary Files – Online

Supplementary Table 1. Differentially expressed genes in indoxyl sulfate stimulated HUVECs as compared to control.

Supplementary Figure 1. IS-affected angiogenic responses compared to the control group with no IS stimulation at day 3 and day 4 respectively in a 3D co-culture model, in CYP1B1 silenced HUVECs.

Supplementary Figure 2. Gene set enrichment plot for the AhR signalling.

Supplementary Report 1. GSEA Report for differentially expressed genes.

References:

1. Ellis RJ, Small DM, Vesey DA, Johnson DW, Francis R, Vitetta L, et al. Indoxyl sulphate and kidney disease: Causes, consequences and interventions. Nephrology (Carlton). 2016;21(3):170-7. doi: 10.1111/nep.12580. PubMed PMID: 26239363.
2. Wu VC, Young GH, Huang PH, Lo SC, Wang KC, Sun CY, et al. In acute kidney injury, indoxyl sulfate impairs human endothelial progenitor cells: modulation by statin. Angiogenesis. 2013;16(3):609-24. doi: 10.1007/s10456-013-9339-8. PubMed PMID: 23408148.
3. Tan X, Cao X, Zou J, Shen B, Zhang X, Liu Z, et al. Indoxyl sulfate, a valuable biomarker in chronic kidney disease and dialysis. Hemodial Int. 2017;21(2):161-7. doi: 10.1111/hdi.12483. PubMed PMID: 27616754.
4. Yamaguchi J, Tanaka T, Inagi R. Effect of AST-120 in Chronic Kidney Disease Treatment: Still a Controversy? Nephron. 2017;135(3):201-6. doi: 10.1159/000453673. PubMed PMID: 27960172.
5. Schulman G, Berl T, Beck GJ, Remuzzi G, Ritz E, Arita K, et al. Randomized Placebo-Controlled EPPIC Trials of AST-120 in CKD. J Am Soc Nephrol. 2015;26(7):1732-46. doi: 10.1681/ASN.2014010042. PubMed PMID: 25349205; PubMed Central PMCID: PMCPMC4483576.
6. DOU L, JOURDE-CHICHE N, FAURE V, CERINI C, BERLAND Y, DIGNAT-GEORGE F, et al. The uremic solute indoxyl sulfate induces oxidative stress in endothelial cells. Journal of Thrombosis and Haemostasis. 2007;5.
7. Yu M, Kim YJ, Kang DH. Indoxyl sulfate-induced endothelial dysfunction in patients with chronic kidney disease via an induction of oxidative stress. Clin J Am Soc Nephrol. 2011;6(1):30-9. doi: 10.2215/CJN.05340610. PubMed PMID: 20876676; PubMed Central PMCID: PMCPMC3022246.
8. Stinghen AE, Chillon JM, Massy ZA, Boullier A. Differential effects of indoxyl sulfate and inorganic phosphate in a murine cerebral endothelial cell line (bEnd.3). Toxins (Basel). 2014;6(6):1742-60. doi: 10.3390/toxins6061742. PubMed PMID: 24902077; PubMed Central PMCID: PMCPMC4073127.
9. Lee WC, Li LC, Chen JB, Chang HW. Indoxyl sulfate-induced oxidative stress, mitochondrial dysfunction, and impaired biogenesis are partly protected by vitamin C and N-acetylcysteine. ScientificWorldJournal. 2015;2015:620826. doi: 10.1155/2015/620826. PubMed PMID: 25839054; PubMed Central PMCID: PMCPMC4369955.
10. Koizumi M, Tatebe J, Watanabe I, Yamazaki J, Ikeda T, Morita T. Aryl Hydrocarbon Receptor Mediates Indoxyl Sulfate-Induced Cellular Senescence in Human Umbilical Vein Endothelial Cells. J Atherosclero Thromb. 2014;21.
11. Shen WC, Liang CJ, Huang TM, Liu CW, Wang SH, Young GH, et al. Indoxyl sulfate enhances IL-1beta-induced E-selectin expression in endothelial cells in acute kidney injury by the ROS/MAPKs/NFkappaB/AP-1 pathway. Arch Toxicol. 2016;90(11):2779-92. doi: 10.1007/s00204-015-1652-0. PubMed PMID: 26659566.
12. Kim HY, Yoo TH, Hwang Y, Lee GH, Kim B, Jang J, et al. Indoxyl sulfate (IS)-mediated immune dysfunction provokes endothelial damage in patients with end-stage renal disease (ESRD). Sci Rep. 2017;7(1):3057. doi: 10.1038/s41598-017-03130-z. PubMed PMID: 28596556; PubMed Central PMCID: PMCPMC5465082.
13. Chrifi I, Hermkens D, Brandt MM, van Dijk CGM, Burgisser PE, Haasdijk R, et al. Cgnl1, an endothelial junction complex protein, regulates GTPase mediated angiogenesis. Cardiovasc Res. 2017;113(14):1776-88. doi: 10.1093/cvr/cvx175. PubMed PMID: 29016873; PubMed Central PMCID: PMCPMC5852532.
14. Chrifi I, Louzao-Martinez L, Brandt MM, van Dijk CGM, Burgisser PE, Zhu C, et al. CMTM4 regulates angiogenesis by promoting cell surface recycling of VE-cadherin to endothelial adherens junctions. Angiogenesis. 2018. doi: 10.1007/s10456-018-9638-1. PubMed PMID: 30097810.
15. YKim YW, Byzova TV. Oxidative stress in angiogenesis and vascular disease. Blood. 2014;123. doi: 10.1182/blood2013-09-512749.
16. DOU L, BERTRAND E, CERINI C, FAURE V, SAMPOL J, VANHOLDER R, et al. The uremic solutes p-cresol and indoxyl sulfate inhibit endothelial proliferation and wound repair. Kidney International.

2004;65.
17. Nath KA, Norby SM. Reactive oxygen species and acute renal failure. PHYSIOLOGY IN MEDICINE. 2000;109.
18. Aksua U, Demircic C, Incea C. The Pathogenesis of Acute Kidney Injury and the Toxic Triangle of Oxygen, Reactive Oxygen Species and Nitric Oxide. Pathophysiology of Sepsis- Induced Acute Kidney Injury. 2011;174.
19. Dennis JM, Witting PK. Protective Role for Antioxidants in Acute Kidney Disease. Nutrients. 2017;9(7). doi: 10.3390/nu9070718. PubMed PMID: 28686196; PubMed Central PMCID: PMCPMC5537833.
20. Duni A, Liakopoulos V, Rapsomanikis KP, Dounousi E. Chronic Kidney Disease and Disproportionally Increased Cardiovascular Damage: Does Oxidative Stress Explain the Burden? Oxid Med Cell Longev. 2017;2017:9036450. doi: 10.1155/2017/9036450. PubMed PMID: 29333213; PubMed Central PMCID: PMCPMC5733207.
21. Cachofeiro V, Goicochea M, de Vinuesa SG, Oubina P, Lahera V, Luno J. Oxidative stress and inflammation, a link between chronic kidney disease and cardiovascular disease. Kidney Int Suppl. 2008;(111):S4-9. doi: 10.1038/ki.2008.516. PubMed PMID: 19034325.
22. Leong SC, Sirich TL. Indoxyl Sulfate-Review of Toxicity and Therapeutic Strategies. Toxins (Basel). 2016;8(12). doi: 10.3390/toxins8120358. PubMed PMID: 27916890; PubMed Central PMCID: PMCPMC5198552.
23. Maheshwari V, Thijssen S, Tao X, Fuertinger DH, Kappel F, Kotanko P. In silico comparison of protein-bound uremic toxin removal by hemodialysis, hemodiafiltration, membrane adsorption, and binding competition. Sci Rep. 2019;9(1):909. doi: 10.1038/s41598-018-37195-1. PubMed PMID: 30696874; PubMed Central PMCID: PMCPMC6351554.
24. Wang W, Liu X, Wang W, Li J, Li Y, Li L, et al. The Effects of Indoxyl Sulfate on Human Umbilical Cord-Derived Mesenchymal Stem Cells In Vitro. Cell Physiol Biochem. 2016;38(1):401-14. doi: 10.1159/000438639. PubMed PMID: 26824459.
25. Kharait S, Haddad DJ, Springer ML. Nitric oxide counters the inhibitory effects of uremic toxin indoxyl sulfate on endothelial cells by governing ERK MAP kinase and myosin light chain activation. Biochem Biophys Res Commun. 2011;409(4):758-63. doi: 10.1016/j.bbrc.2011.05.084. PubMed PMID: 21621512; PubMed Central PMCID: PMCPMC3121259.
26. Carmona A, Guerrero F, Buendia P, Obrero T, Aljama P, Carracedo J. Microvesicles Derived from Indoxyl Sulfate Treated Endothelial Cells Induce Endothelial Progenitor Cells Dysfunction. Front Physiol. 2017;8:666. doi: 10.3389/fphys.2017.00666. PubMed PMID: 28951723; PubMed Central PMCID: PMCPMC5599774.
27. Edamatsu T, Fujieda A, Itoh Y. Phenyl sulfate, indoxyl sulfate and p-cresyl sulfate decrease glutathione level to render cells vulnerable to oxidative stress in renal tubular cells. PLoS One. 2018;13(2):e0193342. doi: 10.1371/journal.pone.0193342. PubMed PMID: 29474405; PubMed Central PMCID: PMCPMC5825083.
28. Edamatsu T, Fujieda A, Ezawa A, Itoh Y. Classification of Five Uremic Solutes according to Their Effects on Renal Tubular Cells. Int J Nephrol. 2014;2014:512178. doi: 10.1155/2014/512178. PubMed PMID: 25431671; PubMed Central PMCID: PMCPMC4241681.
29. Saito S, Shimizu H, Yisireyili M, Nishijima F, Enomoto A, Niwa T. Indoxyl sulfate-induced activation of (pro)renin receptor is involved in expression of TGF-beta1 and alpha-smooth muscle actin in proximal tubular cells. Endocrinology. 2014;155(5):1899-907. doi: 10.1210/en.2013-1937. PubMed PMID: 24601883.
30. Adesso S, Popolo A, Bianco G, Sorrentino R, Pinto A, Autore G, et al. The uremic toxin indoxyl sulphate enhances macrophage response to LPS. PLoS One. 2013;8(9):e76778. doi: 10.1371/journal.pone.0076778. PubMed PMID: 24098806; PubMed Central PMCID: PMCPMC3786936.
31. Vanholder R, De Smet R, Glorieux G, Argiles A, Baurmeister U, Brunet P, et al. Review on uremic toxins: classification, concentration, and interindividual variability. Kidney Int.

2003;63(5):1934-43. doi: 10.1046/j.1523-1755.2003.00924.x. PubMed PMID: 12675874.
32. Ushio-Fukai M, Nakamura Y. Reactive oxygen species and angiogenesis: NADPH oxidase as target for cancer therapy. Cancer Lett. 2008;266(1):37-52. doi: 10.1016/j.canlet.2008.02.044. PubMed PMID: 18406051; PubMed Central PMCID: PMCPMC2673114.
33. Lewis DFV. Oxidative stress: the role of cytochromes P450 in oxygen activation. Journal of Chemical Technology & Biotechnology. 2002;77(10).
34. Munro AW, McLean KJ, Grant JL, Makris TM. Structure and function of the cytochrome P450 peroxygenase enzymes. Biochem Soc Trans. 2018;46(1):183-96. doi: 10.1042/BST20170218. PubMed PMID: 29432141; PubMed Central PMCID: PMCPMC5818669.
35. Murray G, Melvin WT, Greenlee WF, M.D. B. Regulation, Function, and Tissue-specific Expression of Cytochrome P450 CYP1B1. Annual Review of Pharmacology. 2001;41.
36. Stengel B. Chronic kidney disease and cancer: a troubling connection. J Nephrol. 2010;23.
37. McFadyen MC, Melvin WT, Murray GI. Cytochrome P450 CYP1B1 activity in renal cell carcinoma. Br J Cancer. 2004;91(5):966-71. doi: 10.1038/sj.bjc.6602053. PubMed PMID: 15280921; PubMed Central PMCID: PMCPMC2409861.
38. Gondouin B, Cerini C, Dou L, Sallee M, Duval-Sabatier A, Pletinck A, et al. Indolic uremic solutes increase tissue factor production in endothelial cells by the aryl hydrocarbon receptor pathway. Kidney Int. 2013;84(4):733-44. doi: 10.1038/ki.2013.133. PubMed PMID: 23636172.
39. Dallaglio K, Bruno A, Cantelmo AR, Esposito AI, Ruggiero L, Orecchioni S, et al. Paradoxic effects of metformin on endothelial cells and angiogenesis. Carcinogenesis. 2014;35(5):1055-66. doi: 10.1093/carcin/bgu001. PubMed PMID: 24419232; PubMed Central PMCID: PMCPMC4004203.
40. Tang Y, Scheef EA, Wang S, Sorenson CM, Marcus CB, Jefcoate CR, et al. CYP1B1 expression promotes the proangiogenic phenotype of endothelium through decreased intracellular oxidative stress and thrombospondin-2 expression. Blood. 2009;113(3):744-54. doi: 10.1182/blood-2008-03-145219. PubMed PMID: 19005183; PubMed Central PMCID: PMCPMC2628380.
41. Palenski TL, Gurel Z, Sorenson CM, Hankenson KD, Sheibani N. Cyp1B1 expression promotes angiogenesis by suppressing NF-kappaB activity. Am J Physiol Cell Physiol. 2013;305(11):C1170-84. doi: 10.1152/ajpcell.00139.2013. PubMed PMID: 24088896; PubMed Central PMCID: PMCPMC3882387.
42. Futrakul N, Butthep P, Laohareungpanya N, Chaisuriya P, Ratanabanangkoon K. A defective angiogenesis in chronic kidney disease. Ren Fail. 2008;30(2):215-7. doi: 10.1080/08860220701813335. PubMed PMID: 18300124.
43. Anderson CE, Hamm LL, Batuman G, Kumbala DR, Chen CS, Kallu SG, et al. The association of angiogenic factors and chronic kidney disease. BMC Nephrol. 2018;19(1):117. doi: 10.1186/s12882-018-0909-2. PubMed PMID: 29783932; PubMed Central PMCID: PMCPMC5963107.
44. Tanaka S, Tanaka T, Nangaku M. Hypoxia and Dysregulated Angiogenesis in Kidney Disease. Kidney Dis (Basel). 2015;1(1):80-9. doi: 10.1159/000381515. PubMed PMID: 27536668; PubMed Central PMCID: PMCPMC4934802.
45. Hung SC, Kuo KL, Huang HL, Lin CC, Tsai TH, Wang CH, et al. Indoxyl sulfate suppresses endothelial progenitor cell-mediated neovascularization. Kidney Int. 2016;89(3):574-85. doi: 10.1016/j.kint.2015.11.020. PubMed PMID: 26880454.
46. Lin CJ, Wu CJ, Wu PC, Pan CF, Wang TJ, Sun FJ, et al. Indoxyl Sulfate Impairs Endothelial Progenitor Cells and Might Contribute to Vascular Dysfunction in Patients with Chronic Kidney Disease. Kidney Blood Press Res. 2016;41(6):1025-36. doi: 10.1159/000452604. PubMed PMID: 28006782.
47. Ingram DA, Mead LE, Moore DB, Woodard W, Fenoglio A, M.C. Y. Vessel wall–derived endothelial cells rapidly proliferate because they contain a complete hierarchy of endothelial progenitor cells. HEMOSTASIS, THROMBOSIS, AND VASCULAR BIOLOGY. 2005;105. doi: 10.1182/blood-2004-083057.
48. Patan S. Vasculogenesis and angiogenesis. Cancer Treat Res. 2004;117:3-32. PubMed PMID: 15015550.

49. Schroeder JC, Dinatale BC, Murray IA, Flaveny CA, Liu Q, Laurenzana EM, et al. The uremic toxin 3-indoxyl sulfate is a potent endogenous agonist for the human aryl hydrocarbon receptor. Biochemistry. 2010;49(2):393-400. doi: 10.1021/bi901786x. PubMed PMID: 20000589; PubMed Central PMCID: PMCPMC2805781.
50. Moriguchi T, Motohashi H, Hosoya T, Nakajima O, Takahashi S, Ohsako S, et al. Distinct response to dioxin in an arylhydrocarbon receptor (AHR)-humanized mouse. Proc Natl Acad Sci U S A. 2003;100(10):5652-7. doi: 10.1073/pnas.1037886100. PubMed PMID: 12730383; PubMed Central PMCID: PMCPMC156256.
51. Deguchi T, Nakamura M, Tsutsumi Y, Suenaga A, Otagiri M. Pharmacokinetics and tissue distribution of uraemic indoxyl sulphate in rats. Biopharm Drug Dispos. 2003;24(8):345-55. doi: 10.1002/bdd.370. PubMed PMID: 14595703.
52. Sindhu RK, Vaziri ND. Upregulation of cytochrome P450 1A2 in chronic renal failure: does oxidized tryptophan play a role? Adv Exp Med Biol. 2003;527.
53. Patel RD, Murray IA, Flaveny CA, Kusnadi A, Perdew GH. Ah receptor represses acute-phase response gene expression without binding to its cognate response element. Lab Invest. 2009;89(6):695-707. doi: 10.1038/labinvest.2009.24. PubMed PMID: 19333233; PubMed Central PMCID: PMCPMC2743259.
54. Ikuta T, Kawajiri K. Zinc finger transcription factor Slug is a novel target gene of aryl hydrocarbon receptor. Exp Cell Res. 2006;312(18):3585-94. doi: 10.1016/j.yexcr.2006.08.002. PubMed PMID: 16952353.
55. Hollingshead BD, Beischlag TV, Dinatale BC, Ramadoss P, Perdew GH. Inflammatory signaling and aryl hydrocarbon receptor mediate synergistic induction of interleukin 6 in MCF-7 cells. Cancer Res. 2008;68(10):3609-17. doi: 10.1158/0008-5472.CAN-07-6168. PubMed PMID: 18483242; PubMed Central PMCID: PMCPMC2568985.
56. Patel RD, Kim DJ, Peters JM, Perdew GH. The aryl hydrocarbon receptor directly regulates expression of the potent mitogen epiregulin. Toxicol Sci. 2006;89(1):75-82. doi: 10.1093/toxsci/kfi344. PubMed PMID: 16192470.
57. van Dijk CG, Oosterhuis NR, Xu YJ, Brandt M, Paulus WJ, van Heerebeek L, et al. Distinct Endothelial Cell Responses in the Heart and Kidney Microvasculature Characterize the Progression of Heart Failure With Preserved Ejection Fraction in the Obese ZSF1 Rat With Cardiorenal Metabolic Syndrome. Circ Heart Fail. 2016;9(4):e002760. doi: 10.1161/CIRCHEARTFAILURE.115.002760. PubMed PMID: 27056881.
58. Brandt MM, Meddens CA, Louzao-Martinez L, van den Dungen NAM, Lansu NR, Nieuwenhuis EES, et al. Chromatin Conformation Links Distal Target Genes to CKD Loci. J Am Soc Nephrol. 2018;29(2):462-76. doi: 10.1681/ASN.2016080875. PubMed PMID: 29093029; PubMed Central PMCID: PMCPMC5791087.
59. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139-40. doi: 10.1093/bioinformatics/btp616. PubMed PMID: 19910308; PubMed Central PMCID: PMCPMC2796818.
60. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550. doi: 10.1186/s13059-014-0550-8. PubMed PMID: 25516281; PubMed Central PMCID: PMCPMC4302049.
61. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. Nucleic Acids Res. 2009;37(Web Server issue):W305-11. doi: 10.1093/nar/gkp427. PubMed PMID: 19465376; PubMed Central PMCID: PMCPMC2703978.
62. Juni R.P., Kuster D.W.D., Goebel M., Helmes M., van der Velden J., Musters R., et al. Loss of Cardiac Microvascular Endothelial Cell-Cardiomyocyte Interaction by Inflammatory Activation is Restored by Empagliflozin. JACC: Basic to Transl Sci. 2019;In Press.

Chapter 4

H3K27ac acetylome signatures reveal the epigenomic reorganization in remodeled non-failing human hearts

Jiayi Pei[1,2,3‡], Magdalena Harakalova[2,3,4,‡], Thomas A.Treibel[5], Tom Lumbers[5], Bastiaan J. Boukens[6], Igor R. Efimov[7], Jip T. van Dinter[2], Arantxa González[8,9], Begoña López[8,9], Hamid El Azzouzi[2], Noortje van den Dungen[10], Christian G.M. van Dijk[1], Merle M. Krebber[1], Hester M. den Ruijter[11], Gerard Pasterkamp[10], Dirk J. Duncker[12], Edward E. S. Nieuwenhuis[13], Roel de Weger[4], Manon M. Huibers[4], Aryan Vink[4], Jason H. Moore[15], James C. Moon[5], Marianne C. Verhaar[1], Georgios Kararigas[14], Michal Mokry[3,10,13]*, Folkert W. Asselbergs[2,16,17]*, Caroline Cheng[1,3,12]*

‡Equal contribution of first authors, *Corresponding authors

[1]Department of Nephrology and Hypertension, DIGD, UMC Utrecht, University of Utrecht, NL;
[2]Department of Cardiology, Division Heart & Lungs, UMC Utrecht, University of Utrecht, NL;
[3]Regenerative Medicine Utrecht (RMU), UMC Utrecht, University of Utrecht, NL;
[4]Department of Pathology, UMC Utrecht, University of Utrecht, NL;
[5]Institute of Cardiovascular Science, University College London, London, UK;
[6]Department of Medical Biology, AMC, Amsterdam, NL;
[7]Department of Biomedical Engineering, GWU, US;
[8]Program of Cardiovascular Diseases, CIMA Universidad de Navarra and IdiSNA, Pamplona, ES; [9]CIBERCV, Carlos III Institute of Health, Madrid, ES;
[10]Laboratory of Clinical Chemistry and Hematology, UMC Utrecht, NL;
[11]Department of Experimental Cardiology, UMC Utrecht, University of Utrecht, NL;
[12]Division of Experimental Cardiology, Department of Cardiology, Erasmus University Medical Center, Rotterdam, NL;
[13]Division of Paediatrics, UMC Utrecht, University of Utrecht, NL;
[14]Charité – Universitätsmedizin Berlin, and DZHK (German Centre for Cardiovascular Research), partner site Berlin, DE;
[15]Institute for Biomedical Informatics, UPENN, US;
[16]Institute of Cardiovascular Science, Faculty of Population Health Science, University College London, London, UK;
[17]Health Data Research UK and Institute of Health Informatics, University College London, London, UK.

Clinical Epigenetics

doi: 10.1186/s13148-020-00895-5

ABSTRACT

Background: H3K27ac histone acetylome changes contribute to the phenotypic response in heart diseases, particularly in end-stage heart failure. However, such epigenetic alterations have not been systematically investigated in remodeled non-failing human hearts. Therefore, valuable insight into cardiac dysfunction in early remodeling is lacking. This study aimed to reveal the acetylation changes of chromatin regions in response to myocardial remodeling and their correlations to transcriptional changes of neighboring genes.

Results: We detected chromatin regions with differential acetylation activity (DARs, $P_{adj.}<0.05$) between remodeled non-failing patient hearts and healthy donor hearts. The acetylation level of the chromatin region correlated with its RNA polymerase II occupancy level and the mRNA expression level of its adjacent gene per sample. Annotated genes from DARs were enriched in disease-related pathways, including fibrosis and cell metabolism regulation. DARs that change in the same direction have a tendency to cluster together, suggesting the well-reorganized chromatin architecture that facilitates the interactions of regulatory domains in response to myocardial remodeling. We further show the differences between the acetylation level and the mRNA expression level of cell-type-specific markers for cardiomyocytes and 11 non-myocyte cell types. Notably, we identified transcriptome factor (TF) binding motifs that were enriched in DARs and defined TFs that were predicted to bind to these motifs. We further showed 64 genes coding for these TFs that were differentially expressed in remodeled myocardium when compared with controls.

Conclusions: Our study reveals extensive novel insight on myocardial remodeling at the DNA regulatory level. Differences between the acetylation level and the transcriptional level of cell-type-specific markers suggest additional mechanism(s) between acetylome and transcriptome. By integrating these two layers of epigenetic profiles, we further provide promising TF-encoding genes that could serve as master regulators of myocardial remodeling. Combined, our findings highlight the important role of chromatin regulatory signatures in understanding disease etiology.

Keywords

Myocardial remodeling, histone acetylation, transcriptome, transcription factor

INTRODUCTION

Myocardial remodeling is defined as changes in the size, shape, structure, and function of the heart from cardiac injury due to various causes [1,2]. It is a complex process resulting from the interplay between cardiomyocytes (hypertrophy), cardiac fibroblasts (fibrosis), vascular smooth muscle cells (vascular stiffening), vascular endothelial cells (endothelial dysfunction), and leukocytes (inflammation) [2]. Pathological myocardial remodeling has a poor prognosis related to a higher risk of heart failure and sudden cardiac death [3]. Given the scarce availability of patient and control cardiac biopsies in humans, most mechanistic studies on myocardial remodeling are based on animal models or cultured cells [4–6] that do not necessarily represent the patient's situation completely [7–9]. Human myocardial biopsies have been used to investigate remodeling between health and disease on the level of the methylome, transcriptome, and proteome [10–14]. However, chromatin regulation, defined as the dynamic modification of chromatin architecture to control gene expression [15,16], has not been systematically investigated. However, a previous study mapped the epigenome in failing human hearts [17]. Although the epigenetic regulation of specific classes of genes has been suggested to contribute to the phenotypic response throughout the mild-stage to the end-stage heart failure [18], there is still a lack of insight into the early-stage of cardiac dysfunction, which could be elucidated by studying chromatin regulation changes between healthy and remodeled non-failing human hearts.

Chromatin immunoprecipitation and sequencing (ChIP-seq) is widely utilized to study chromatin regulation [19]. Key DNA regulatory regions and pathways in several (mostly inflammatory) diseases have been identified using ChIP-seq [20,21]. Histone 3 lysine 27 acetylation (H3K27ac) is found at both active enhancers and promoters, which are accessible to transcription factors (TFs), polymerases, and other members of the transcriptional complex in order to regulate transcription of genes [22]. Histone acetylome studies using the H3K27ac mark have successfully identified chromatin regulation changes under healthy and diseased conditions [23,24]. Additionally, the H3K27ac level correlates with gene expression levels [16]. Other histone marks that are indicative of transcription include promoter-regions-associated H3K4me3 and H3K9ac and gene-bodies-associated H3K36me3 and H3K79me3 [25]. A recent paper further indicates that the H3K27ac and H3K36me3 marks serve as the best predictive marks for pathological gene programming in diseased human cardiomyocytes when compared with other histone marks [17]. Additionally, studies have shown that H3K27ac exhibits the best correlation with both active promoters and enhancers when compared with other marks [26,27]. Despite the rapid increase of ChIP-

seq data over the last decade [28,29], the H3K27ac landscape that defines myocardial remodeling has been only scarcely investigated [17], mainly due to the lack of proper samples and controls.

To improve our understanding of myocardial remodeling at the DNA regulatory level, in this study we employ a ChIP-seq approach to characterize the H3K27ac binding landscape. Our findings provide new insights into the early disease etiology, revealing genome-wide histone acetylation changes between rare remodeled non-failing myocardial biopsies from patients with severe aortic stenosis (AS) [30] and control samples from non-transplanted donor hearts. Firstly, our study yields a unique list of differentially acetylated regions (DARs) between these two groups. Secondly, by studying the chromatin dynamics, we show the well-organized chromatin structure that physically facilitates gene regulation. Enrichment analysis using annotated genes from DARs indicate altered extracellular matrix (ECM) organization and cell metabolism regulation in remodeled myocardium. Thirdly, we investigate the correlation between histone acetylation changes and the transcriptome changes using data from our study and published studies. Fourthly, we present the differences between the acetylation levels and the mRNA expression level of markers specific for cardiomyocytes and 11 non-myocyte cell types in response to myocardial remodeling, suggesting the additional mechanism(s) between chromatin acetylation and gene transcription. Lastly, we have identified TF binding motifs that are enriched in DARs and genes coding for TFs predicted to bind to these enriched motifs. Notably, 64 TF-encoding genes are differentially expressed in remodeled myocardium versus controls, and they may serve as candidate master regulators of myocardial remodeling.

RESULTS

DARs between non-failing remodeled hearts and control hearts

We studied histone acetylation activity using H3K27ac ChIP-seq in myocardium from AS patient hearts when compared with control hearts, followed by multiple *in silico* analyses to illustrate chromatin structural dynamics and biological functions enriched by annotated genes in remodeled myocardium (Supplementary Figure 1). On average, we obtained 40,745±9,656 and 16,734±6,896 regions using H3K27ac ChIP-seq in control and patient samples, respectively. Among detected H3K27ac regions, 27,879 regions were supported by at least two independent samples, out of which 11,358 unique regions showed significantly different H3K27ac occupancy between two groups (adjusted p-value<0.05, Figure 1a, 1b and 1c, Supplementary Table 1). From these, we identified 5,634 regions with increased signal in patients

(hyperacetylation) and 5,724 regions with increased signal in controls (hypoacetylation). Examples of hyper- and hypoacetylated regions are shown in Figure 1d. This altogether demonstrates extensive changes in histone acetylation upon myocardial remodeling.

Adjacent DARs tend to have the same direction in their acetylation changes

A growing number of studies using chromatin conformation capture techniques (i.e. 4C-seq and Hi-C) have shown that compartments and topologically associating domains (TADs) are tissue-specific and contain preferential interactions of certain regulatory elements and genes [31]. Changes within and among TADs influence gene regulation, and such changes have been observed in cancer [31]. We found that regions with the same direction in acetylation change upon myocardial remodeling are more likely located next to each other (Figure 2a, 2b, and 2c). Using the adjacent (5kb-100kb) peak pairs we have found that one DAR is more likely followed by regions with the same direction of acetylation change when compared with randomized regions (p-value=2.4e-108, Figure 2d). Similar behavior was also observed in more distal (100kb-1Mb) peak pairs (p-value=0.00012, Figure 2d, and Supplementary Table 2). This suggests that adjacent DARs with the same direction in acetylation change might function together as larger regulatory domains or potential subdomains within TADs.

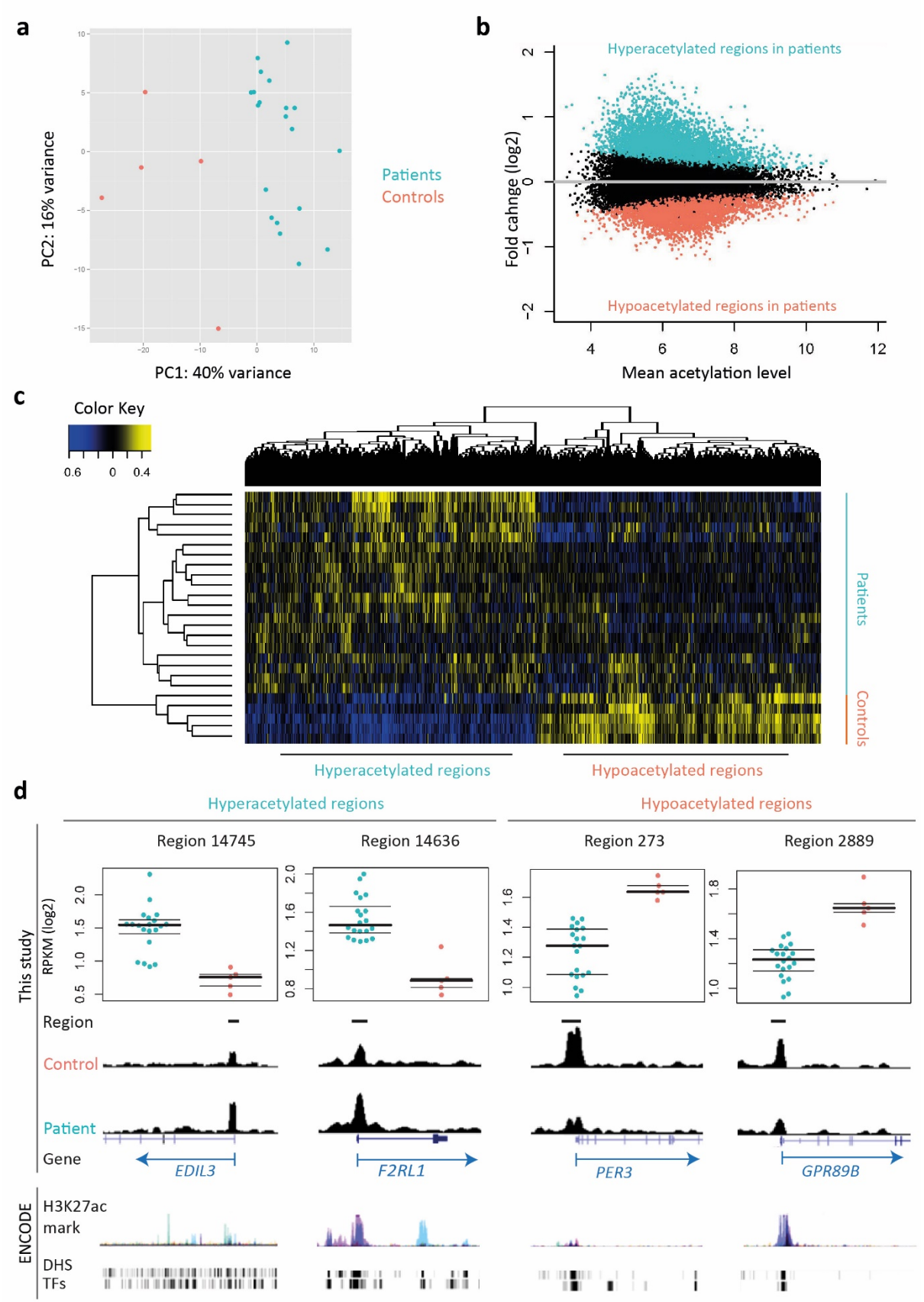

a
PC2: 16% variance
PC1: 40% variance
Patients
Controls
b
Hyperacetylated regions in patients
Hypoacetylated regions in patients
Fold cahnge (log2)
Mean acetylation level
c
Color Key
0.6
0
0.4
Patients
Controls
Hyperacetylated regions
Hypoacetylated regions
d
Hyperacetylated regions
Hypoacetylated regions
Region 14745
Region 14636
Region 273
Region 2889
This study
RPKM (log2)
Region
Control
Patient
Gene
EDIL3
F2RL1
PER3
GPR89B
ENCODE
H3K27ac mark
DHS
TFs

Figure 1. Differentially acetylated H3K27ac regions between patients and controls. **a** Principal component analysis (PCA) plot showing the clustering of patient and control samples based on H3K27ac profiles (using 500 regions with the highest variance). **b** MA plot showing the mean acetylation levels of all samples (x-axis) and the fold changes between two groups in the log2 scale (y-axis). Colored dots represent hyper (aqua)- and hypoacetylated (coral) regions in patients compared with controls respectively (adjusted p-value<0.05). **c** Heatmap depicts the clustering of samples based on all DARs acetylation levels. **d** Examples of DARs in the UCSC genome browser. Dot plots depict the acetylation level in patients (blue) and controls (orange). ENCODE =public ENCODE data default display. H3K27ac mark=ChIP-seq data of 7 cell types obtained from ENCODE. DHS =DNaseI hypersensitivity clusters in 125 cells. TFs =ChIP-seq data of 161 transcription factors.

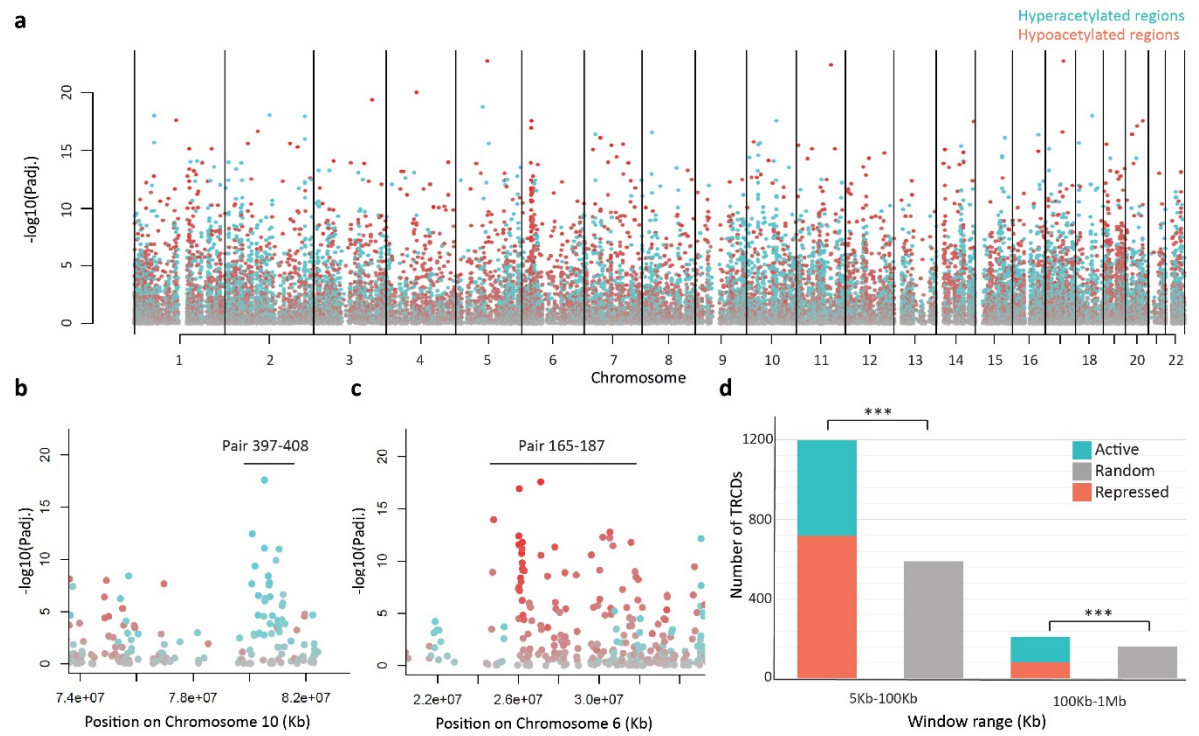


Figure 2. Distribution of tandem regulated chromatin domains (TRCDs). **a** Manhattan plot depicting the distribution of differentially H3K27 acetylated regions in patients vs. controls: non-significant regions (grey), hyperacetylated regions (aqua), and hypoacetylated regions (coral). **b** Zoomed-in view of clusters of TRCDs in the short-range (indicated with the bar with higher acetylation level in patients than in controls on chromosome 10. **c** Zoomed-in view of clusters of TRCDs in the short-range (indicated with the bar) with lower acetylation level in patients than in controls on chromosome 6. **d** Number of identified active and repressed TRCDs in short- and long-genomic distance respectively in real and randomly distributed dataset.

Annotated genes from DARs are well correlated with their transcriptome changes

We performed region-to-gene annotation from DARs. Briefly, we mapped genes located in the vicinity of DARs using a conservative window of ±5kb from the transcription start site (TSS), which is recognized to be within the promoter region range of most genes. In total, we annotated 1,294 and 5,886 genes in the

hyper- and hypoacetylated regions, respectively (Supplementary Table 3). Examples of noteworthy annotated genes in the hyper- and hypoacetylated regions are shown in Figure 1d.

First, we confirmed that the H3K27ac signal corresponded to gene expression levels as proxied by RNA polymerase II (RNAPII) occupancy and RNA sequencing (RNA-seq) per sample (Figure 3a and Supplementary Figure 3). Next, we employed gene set enrichment analysis to examine the correlation of annotated genes from DARs with their mRNA expression changes in the transcription profiles obtained in this study and in previously published studies [11,12]. Based on 4,240 differentially expressed genes between patients and controls from the RNA-seq dataset in our study (P value<0.05, Supplementary Figure 4 and Supplementary Table 4), we showed a significant correlation between annotated genes from the hypoacetylated regions and genes with lower expression levels in patients (FDR<0.001, Figure 3b). In contrast, annotated genes from the hyperacetylated regions were not enriched in the pool of genes with higher expression levels in patients when compared with controls. However, annotated genes from the hyperacetylated regions showed a statistically significant correlation with genes with higher expression levels in patients when compared with controls from two published studies (FDR=0.007 and FDR<0.001). There was no significant correlation between the annotated genes from the hypoacetylated regions in the pool of lower expressed genes in these two studies (FDR=0.84 and FDR=0.22, Figure 3b). Taken together, these data suggest that the presence of an H3K27ac signal near TSS is positively correlated with the gene expression levels at the same sample level (Figure 3a), whereas changed H3K27ac signals near TSS and differentially expressed genes between patients and controls are not always correlated (Figure 3b).

Annotated genes from DARs are enriched for remodeling-associated biological processes

Next, we studied their enriched biological functions and pathways (Supplementary Figure 5). The most enriched biological functions in genes close to the hyperacetylated regions were linked to extensive ECM regulation and cell binding (Figure 4a). STRING analysis identified key gene-encoded proteins based on all genes involved in ECM-related processes, including *TGFB1*, *FBN1, MFAP2, FBLN5, MFAP4*, and a cluster of collagen encoding genes (Figure 4c). Likewise, genes involved in enriched Pathway and Biological Process “Hemostasis” and “response to wounding” were annotated by STRING to provide information on interactions in the wound healing process (Figure 4d). The most enriched GO Biological Processes in the hypoacetylated regions were summarized in Figure 4b, and a considerable number of them were involved in (transcriptional) regulation of protein synthesis.

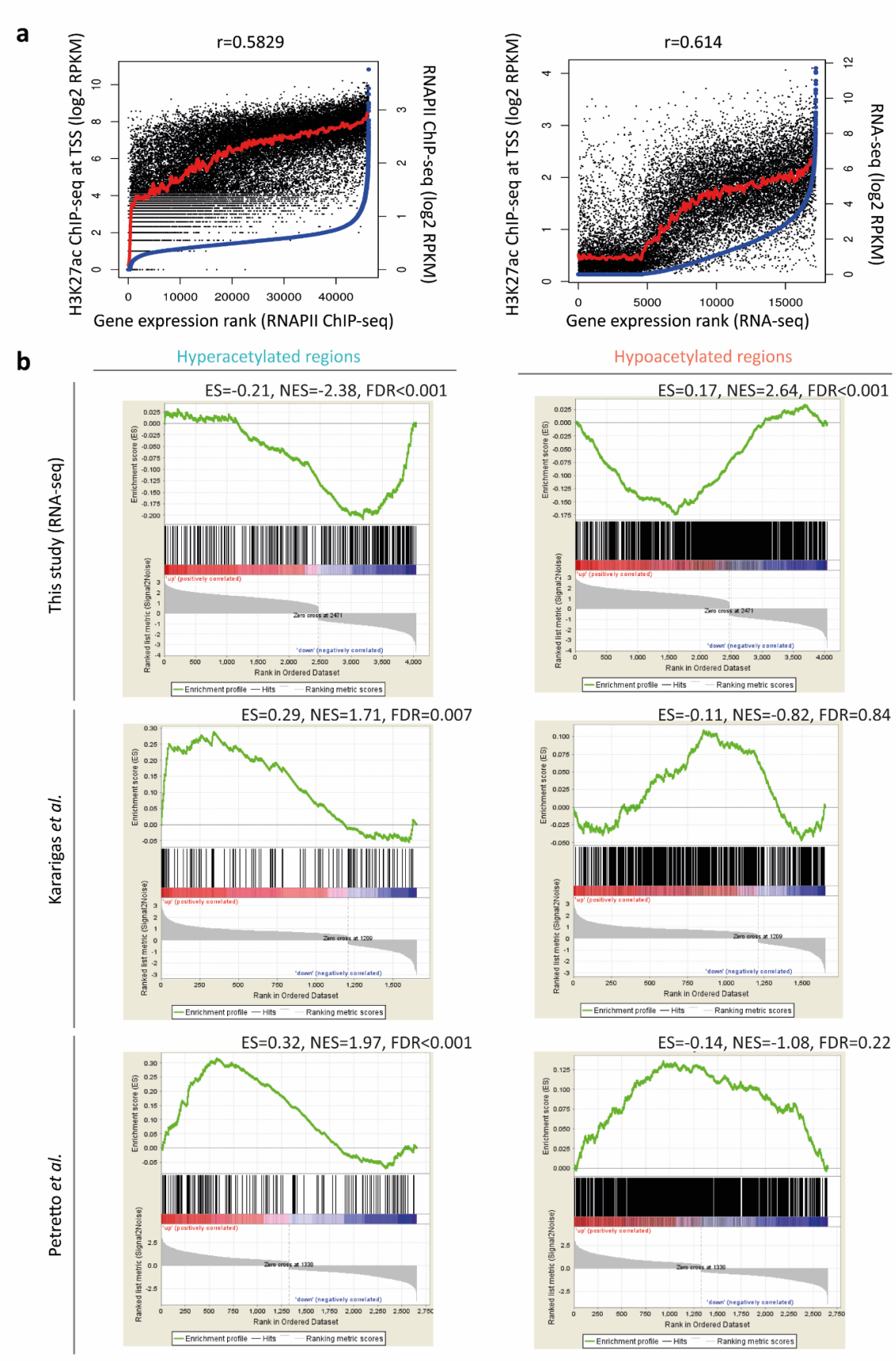

a
r=0.5829
H3K27ac ChIP-seq at TSS (log2 RPKM)
RNAPII ChIP-seq (log2 RPKM)
Gene expression rank (RNAPII ChIP-seq)
r=0.614
H3K27ac ChIP-seq at TSS (log2 RPKM)
RNA-seq (log2 RPKM)
Gene expression rank (RNA-seq)
b
Hyperacetylated regions
Hypoacetylated regions
This study (RNA-seq)
ES=-0.21, NES=-2.38, FDR<0.001
ES=0.17, NES=2.64, FDR<0.001
Karagigas et al.
ES=0.29, NES=1.71, FDR=0.007
ES=-0.11, NES=-0.82, FDR=0.84
Petretto et al.
ES=0.32, NES=1.97, FDR<0.001
ES=-0.14, NES=-1.08, FDR=0.22
Enrichment score (ES)
Ranked list metric (Signal2Noise)
'up' (positively correlated)
'down' (negatively correlated)
Rank in Ordered Dataset
Enrichment profile
Hits
Ranking metric scores

Figure 3. Correlation analysis of H3K27ac ChIP-seq data , RNAPII ChIP-seq data, and RNA-seq data. **a** Correlation between H3K27ac ChIP-seq vs. RNAPII ChIP-seq data (left plot) and RNA-seq data (right plot) in the same sample, respectively. H3K27ac ChIP-seq data is presented on the y-axis, whereas RNAPII ChIP-seq data and RNAseq data are shown on the x-axis and z-axis (log2 scale). **b** Correlation between annotated genes from differentially acetylated regions and differentially expressed genes from transcriptome profiles from our study and two published studies. Differentially expressed genes per study are ranked by their fold changes and shown on the x-axis. The running correlation throughout the gene set is shown by the curve (green) and the running enrichment score (ES) is shown in the y-axis. Enrichment score normalized for gene set size (NES) and the false discovery rate (FDR) are shown above each plot. Black bars indicate annotated genes from differentially acetylated regions that are presented among the transcriptome profiles.

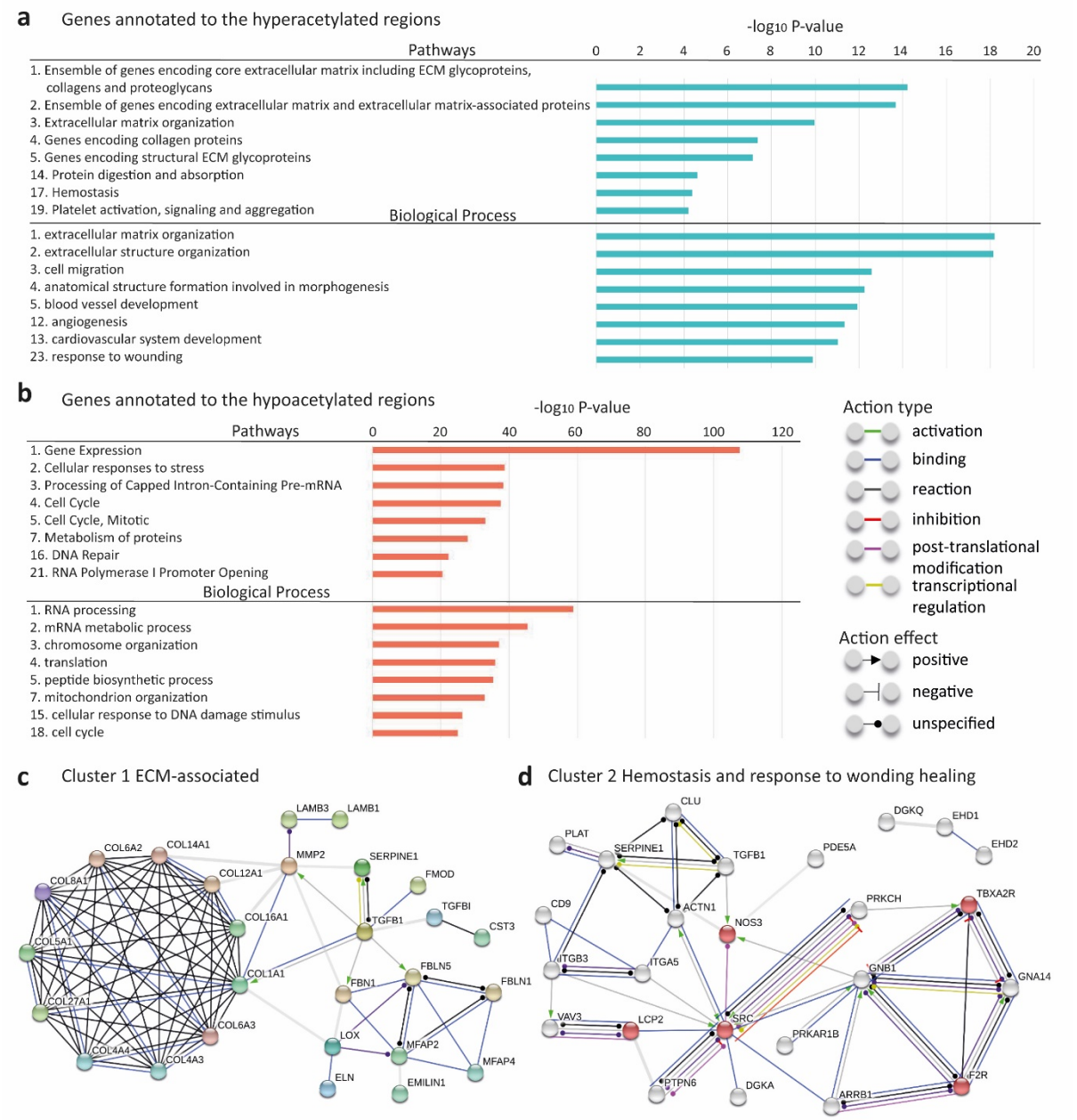


Figure 4. Functional analysis based on genes identified from DARs. **a** Top enriched Pathways and GO Biological

Process based on 1,924 annotated genes in the vicinity of hyperacetylated regions that locate within a ±5kb range from the transcription start site. **b** Top enriched Pathways and GO Biological Process based on 5885 annotated genes in the vicinity of hypoacetylated regions that locate within a ±5kb range from the TSS. **c** Interaction of extracellular matrix-related genes, including TGF**β** pathway associated genes and a cluster of collagen encoding genes detected (STRING). **d** Interaction of genes that are involved in the wound healing process, gene-encoded proteins in platelet activation are indicated in red nodes (STRING).

Acetylation levels and transcriptional levels per cell-type-specific marker

To estimate from which cardiac cell type the H3K27ac signal and RNA-seq data are mainly derived, we further examined H3K27ac occupancy and mRNA expression level of cell-type-specific markers in our bulk data set. First, we collected validated markers for cardiomyocytes and 11 types of non-myocytes as shown by two recent studies in murine hearts in which single-cell RNA sequencing was used [32]. The promoter acetylation level and the mRNA expression level of each marker per cell type in all of our samples, in patients only, and in controls only, are shown in Figure 5a and Figure 5b, respectively. Interestingly, the acetylation levels of these marker genes remained consistent among the 12 cell types, and they were comparable between patients and controls. Their transcriptional levels were also consistent among the 11 non-myocyte cell types and remained comparable between patients and controls. However, the mRNA expression levels of cardiomyocyte-specific markers in all samples were more profoundly expressed when compared with all non-myocyte markers, and this expression pattern was also observed in both patients and controls. We further highlighted the individual markers whose nearest upstream regions were significantly differentially acetylated (Figure 5c) and markers whose mRNA expression levels were significantly differentially expressed between patients and controls (Figure 5d). Out of 19 cardiomyocytes markers, the nearest chromatin regions of 7 markers were significantly differentially acetylated, while another 2 markers were significantly differentially expressed (Figure 5d). For the non-myocytes, some markers showed significant changes in both acetylation and transcriptional levels (i.e. fibroblast marker *LAMC1*). The comparable acetylation and mRNA expression levels of cell-type-specific markers between patients and controls suggest that the cell compositions in healthy and patient-derived cardiac tissues were not significantly different. Given the strikingly higher mRNA expression levels of markers for cardiomyocytes when compared with markers for 11 non-myocytes cell types, the RNA data were more likely to be representative of the transcriptional profile of the cardiomyocyte pool.

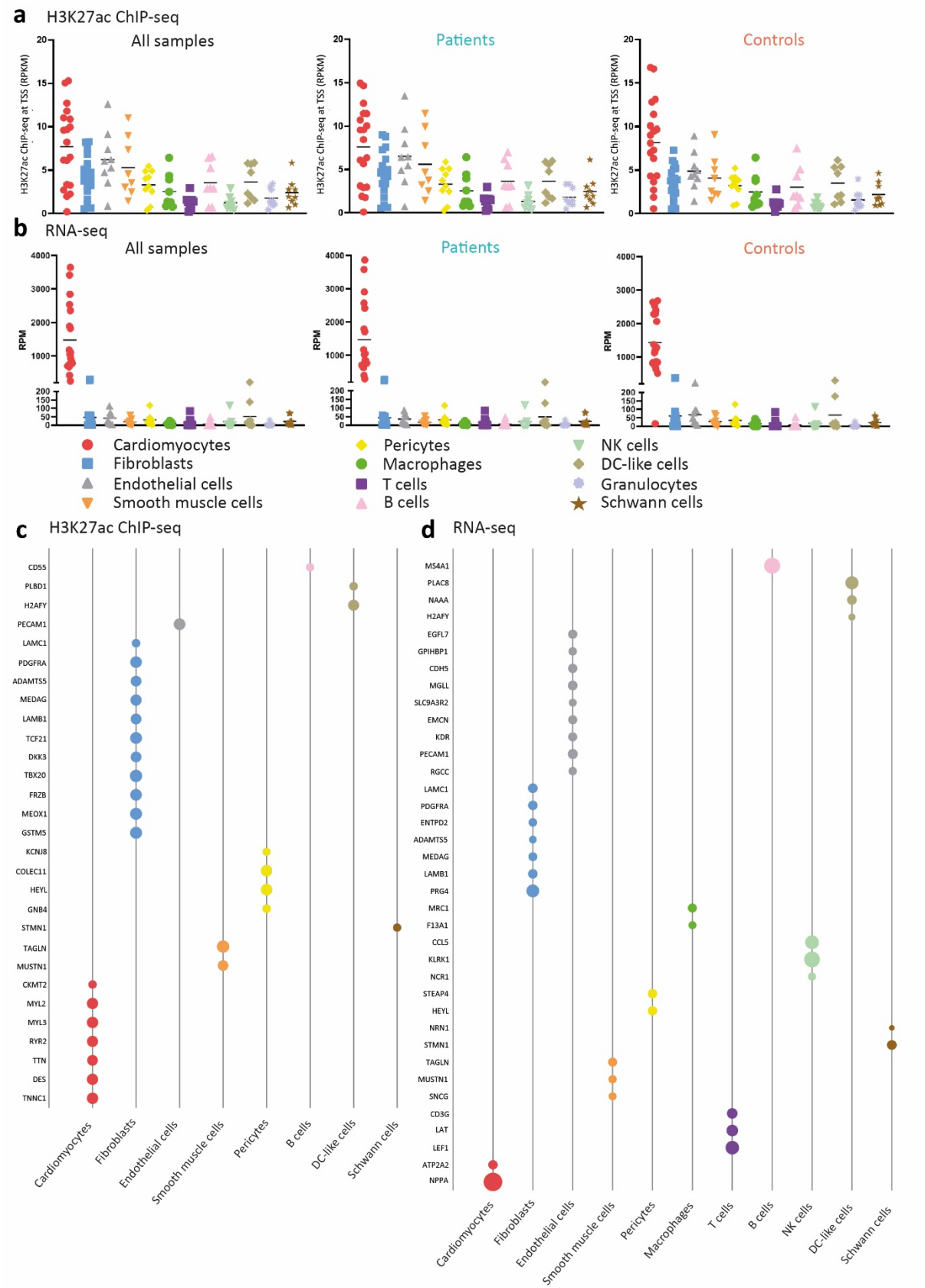


Figure 5. The promoter acetylation and mRNA expression levels of cell-type-specific markers. **a** The acetylation levels

of all makers per cell type in all samples, patient samples, or controls. **b** The mRNA expression levels of all markers per cell type in all samples, patient samples, or controls. **c** Markers, which showed significantly different acetylation levels at the 2.5kb upstream window from the transcription start sites between patients and controls, were shown. Each point represents a marker, and the fold change value was used and corresponds to the point size. **d** Markers, which showed significantly different mRNA expression level between patients and controls, were shown. Each point represents a marker, and the fold change value was used and corresponds to the point size. RPM: reads per million, which were normalized.

Discovery of transcription factor binding motifs (TFBMs) and candidate TFs in regulating myocardial remodeling

To identify the possible upstream regulators of the observed chromatin activity changes, we investigated the occurrence of TFBMs in DARs. In total, we identified 304 and 540 TFBMs in the hyper- and hypoacetylated regions, respectively (Supplementary Table 5). We could assign 48 and 288 TFs with the ability to bind to TFBMs that were exclusively overrepresented in the hyper- and hypoacetylated regions, respectively (Figure 6a). Also, we could assign 231 TFs to TFBMs that were overrepresented in both hyper- and hypoacetylated regions. Genes coding for TFs that were identified in the hypoacetylated regions significantly correlated with genes with lower mRNA expression levels in patients when compared with controls (FDR<0.001, Figure 6a). However, genes coded for the predicted TFs from the hyperacetylated regions were not enriched in the pool of upregulated genes. By overlapping the annotated TFs in the hyperacetylated regions (48+231=279 in total) with the differentially expressed genes, we identified 16 TFs (e.g. *SMAD2* and *ELF3*) that were enriched in the hyperacetylated regions and were also presented with higher mRNA levels in patients when compared with controls. In addition, the overlap analysis showed that 48 annotated TFs (e.g. *CEBPB* and *KLF6*) in the hypoacetylated regions (288+231=519 in total) also showed lower mRNA levels in patients versus controls (*Figure 6b*).

To further investigate if the observed epigenetic regulation is mainly derived from the cardiomyocyte pool, we studied the involvement of those 64 TFs, which bind to enriched TFBMs and showed altered mRNA expression levels in remodeled hearts when compared with controls, using a gene set in hypertrophic human stem cell-derived cardiomyocytes from a published study [33]. Four suppressed TFs identified in our study (*EGR1*, *EGR2*, *EPAS1,* and *ZFHX3*) also showed lower mRNA expression levels in mechanically induced hypertrophic cardiomyocytes when compared with controls. This finding suggests that some of our identified TFs could be highly relevant for cardiomyocytes.

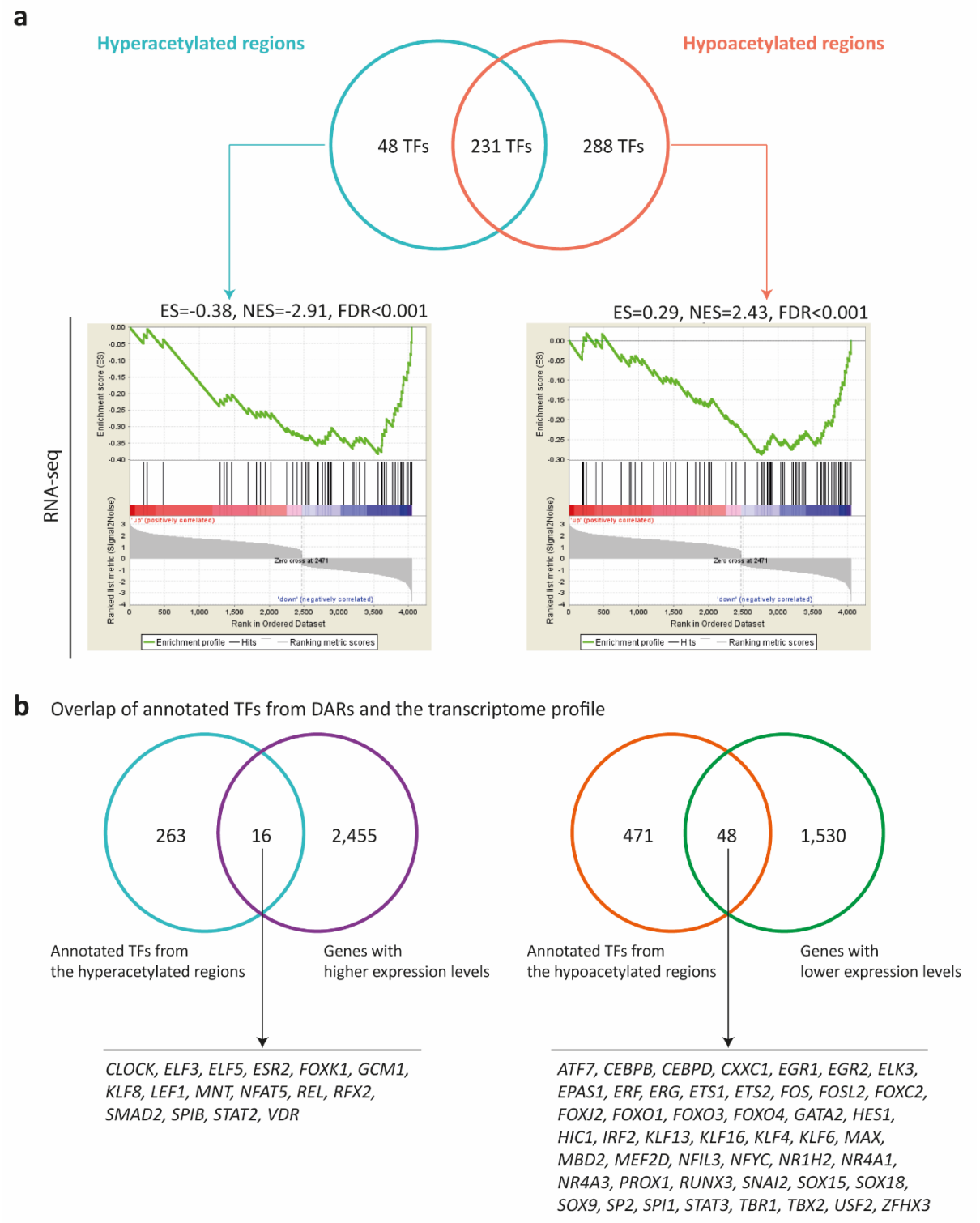


Figure 6. Transcription factors (TFs) annotated from enriched transcription factor binding motifs (TFBMs) in the differentially acetylated regions (DARs). **a** A Venn diagram shows the overlap of motif-encoded TFs that were linked to TFBMs found to be enriched in DARs. In the graph below are shown the gene set enrichment analysis of genes encoding these TFs with differentially expressed genes between patients and controls. Differentially expressed genes per study are ranked by their fold changes and shown on the x-axis. The running correlation throughout the gene set is shown by the curve (green) and the running enrichment score (ES) is shown in the y-axis. Enrichment score

normalized for gene set size (NES) and the false discovery rate (FDR) are shown above each plot. Black bars indicate annotated genes from differentially acetylated regions that are presented among the transcriptome profiles. **b** A subset of TFs that present the same direction in the acetylation activity and mRNA expression change in remodeled myocardium as compared to the controls.

DISCUSSION

In this study, we report the first histone acetylome profile based on H3K27ac ChIP-seq to reveal changed chromatin activities in remodeled non-failing myocardium from AS patients as compared to healthy donor hearts. We identified 5,634 hyperacetylated regions and 5,724 hypoacetylated regions upon myocardial remodeling. We also showed that re-organized chromatin regions, which changed in the same direction in response to myocardial remodeling, tended to cluster together. Genes annotated to these altered regulatory regions are associated with the development of fibrosis formation and the regulation of cell metabolism, and their acetylation levels are correlated with their expression levels between patients and controls, as shown in the mRNA data obtained from our study and published studies. We further demonstrated the acetylation level and mRNA level of markers specific for cardiomyocytes and 11 non-myocyte cell populations. Lastly, by integrating the acetylome and transcriptome datasets, we identified a list of TFs that presented with the same direction in the acetylation change and the mRNA expression changes in remodeled myocardium. Several of these TFs, such as *EGR1* and *EPAS1*, have already been shown to be downregulated in hypertrophic cardiomyocytes [33]. These TFs could be promising candidates in elucidating the key TFs-DNA interactions that are involved in the pathomechanisms in remodeled myocardium, and may serve as potential druggable targets for future epigenetic therapies.

Recent studies demonstrate that alterations in chromatin domains could lead to altered gene expression and disease [34]. A review from van Steensel and Furlong on the spatial organization of the genome showed that the interaction between enhancer and promoter is often found within the same TAD and subsequently regulates transcription [35]. They further indicated the possible influence of active transcription on the chromatin architecture within the subdomains of TADs (sub-TADs). However, little is known about TAD regulation in human myocardium. A previous study mapped the global chromatin structural changes in heart failure using a murine model [36]. Here, we showed that adjacent regions, which might serve as putative (sub-)TADs, were more likely to interact in the same fashion, indicating well-organized chromatin structures that could physically facilitate the interactions of regulatory domains. Yet, additional studies are required to further elucidate these identified adjacent regions as (sub-)TADs using circular chromosome conformation capture and sequencing. We are the first to demonstrate that differences of H3K27ac occupancy are not limited to quantitative and qualitative analyses of detected

regions between healthy and diseased conditions, but also provide insights into the highly-ordered chromatin structural rearrangements in response to the disease.

A recent epigenetic study showed that gene expression levels in cardiomyocytes from end-stage failing human hearts correlated with levels of active histone marks, especially the H3K27ac mark [17]. Consistent with previous studies, we also observed that the H3K27ac occupancy level correlated well with both RNA polymerase II occupancy level using RNAPII ChIP-seq and gene expression level using RNA-seq per sample. Using differentially expressed gene profiles obtained in the present study and from two published studies, we further observed that the annotated gene set in DARs was statistically significantly correlated with differentially expressed genes. Nevertheless, it is important to note that not all datasets were significantly correlated, which could be due to the different number of genes between the gene set and the expression database and the different turnover ratios between histone acetylation and mRNAs [37]. Previous epigenomic studies that aimed to identify novel targets in heart diseases also clearly demonstrated the importance of combining transcriptome data with epigenetic data to reveal different perspectives of the epigenetic regulation of the disease [38]. For example, genes, which showed the same direction between H3K27ac signal and expression level in patients and controls (i.e. chr1:209,973,800-209,977,199/*IRF6* and chr16:56,639,200-56,646,949/*MT2A*), could improve our understanding of the affected biological networks and serve as potential targets for chromatin modifiers and/or gene therapy-related strategies in remodeled hearts.

The most enriched biological processes by genes annotated to the hyperacetylated regions are related to ECM regulation. ECM organization is an important component of cardiac remodeling and extensive cardiac fibrosis is a hallmark of maladaptive hypertrophy [39,40]. TGFβ, a key mediator in ECM formation, is regulated by the activation of SMAD3 and the phosphorylation of ATF2 [41]. We also identified *SMAD3* in the hyperacetylated regions and *ATF2* in the hypoacetylated regions. An important TGF*β*1-related signaling cascade is activated in cardiac fibroblasts from patients with coronary artery disease when compared with controls, and the inhibition of IL11, a main downstream target of TGF*β*1, leads to reduced fibrosis in the heart using murine model [42]. In line with these findings, we also observed a higher mRNA expression levels of *TGFβ1* and *IL11* in patients versus controls. Furthermore, ECM homeostasis is also regulated by a tight balance between matrix metalloproteinases (MMPs), which degrade ECM proteins, and tissue inhibitors of metalloproteinases (TIMPs) [39,43]. We annotated *MMP2*, *MMP23A, TIMP3*, and *TIMP4* in hyperacetylated regions and a group of MMPs that showed higher mRNA expression levels in patients. These data imply the activation of genes involved in fibrogenic pathways at the chromatin level

and transcriptional level, which is in line with previous studies that reveal a strong fibrogenic response in left ventricular remodeling [39]. Although we performed bulk sequencing and the histological staining showed extensive fibrosis in remodeled myocardium, it is important to note that cardiomyocytes also express ECM-related genes [44]. Transcriptional processes and cell cycle were mostly enriched by genes annotated in the hypoacetylated regions, indicating the downregulation of cell metabolism in patients.

Cellular mechanotransduction has been demonstrated in the heart and mechanical forces are greatly increased in AS [45,46], suggesting the potential influence of mechanotransduction during myocardial remodeling. Among various pathways that are involved in the mechanical signaling, the Hippo pathway and its regulation of the YAP/TAZ protein complex have been highlighted to play a central role [46–48]. The core regulators of the Hippo pathways, including MST1/2, SAV1, TAO-family kinases (TAO), and MAP4K kinases, activate LATS1 and LATS2 kinases, which subsequently phosphorylate YAP/TAZ, thereby stimulating proteolysis of the YAP/TAZ complex and limiting its transcription regulatory capacity. In our study, we observed that the chromatin acetylation levels of MST1, MST2 (also known as STK3), and TAOK2 were higher in patients than controls. However, the mRNA expression levels of these genes remained comparable between both groups. In addition, the chromatin acetylation levels and the mRNA expression levels of MAP4K2-5 members and LAST1/2 were not significantly changed in patient hearts when compared with controls. Downstream targets YAP and TAZ also showed similar chromatin acetylation and mRNA expression levels in patient versus control hearts (Supplementary Table 8). As we have conducted bulk sequencing, any subtle changes in the Hippo pathway on single type cell level might remain hidden. Furthermore, changes in pathways related to mechano-activation might mainly present itself on protein level (e.g. level of phosphorylation of the YAP/TAZ complex). Future single cell (type) sequencing based studies in combination to proteomics studies are needed to further elucidate the effect of mechanical forces on different cell types in myocardial remodeling.

Sex differences in myocardial remodeling have been identified by previous studies [49–52]. For example, there is a higher expression level of genes involved in fibrosis and inflammation in male hypertrophic hearts compared with female hypertrophic hearts [53,54]. In the present study, we did not observe a sex-based clustering of hearts with concentric remodeling based on differentially acetylated regions and differentially expressed genes (Supplementary Figure 6). This was most likely due to the limited number of healthy hearts from women and men (n=2 and n=3, respectively). Additionally, unlike left ventricular hypertrophy, concentric remodeling is defined as a normal cardiac mass despite a thickened left ventricular wall [55]. Therefore, the maladaptation of myocardial remodeling is worse in left ventricular

hypertrophy than in concentric remodeling, which could also lead to the different observations on the sex dimorphism between this study and previous studies from us and others. Last but not least, the difference in location from which cardiac specimens were collected between our study and previous studies may also play a role. Combined, it is likely that the sex-specific acetylome and transcriptome were less profound in hearts with moderate myocardial remodeling than advanced myocardial remodeling. Certainly, future studies are needed to investigate sex differences with additional healthy control samples and a larger patient group size.

Notably, several TF-encoding genes that were identified in the present study have previously been shown affected on the gene expression level, either in animal models with remodeled myocardium by our group or in hypertrophic human stem cell-derived cardiomyocytes by others. Previous studies showed a list of TFs with altered mRNA expression levels in pigs with remodeled left ventricle when compared with controls, such as the upregulation of *SMAD3* and *SPIB*, and the downregulation of *FOSL2*, *MEF2D*, *STAT3* and *ATF7* [56,57]. These studies also identified TFs with significantly affected binding activities in hypertrophic conditions using protein/DNA array analysis, such as *EGR1*, *ETS1,* and *ETS2* [57]. In line with these reports, these same TFs were also found in the present study, which showed matched acetylation and/or mRNA levels changes in remodeled hearts when compared with controls. Furthermore, the mRNA expression levels of several TFs also showed consistent changing patterns in hypertrophic human stem cell-derived cardiomyocytes when compared with controls, including *EGR1*, *EGR2*, *EPAS1,* and *ZFHX3* [33]. The involvement of these TFs in cardiac remodeling has also been well established by other studies [58,59].

Taken into account that we performed tissue-level bulk sequencing and the cellular heterogeneity in the heart, our data do not represent changes only in cardiomyocytes. Recent studies have been using single-cell sequencing to further reveal the transcriptional landscape per cell type in the heart, including cardiomyocytes, fibroblasts, and endothelial cells [60]. However, these studies either sorted cells from freshly harvested hearts or they isolated nuclei from snap-frozen human hearts. Due to the difficulty of obtaining intact cells from snap-frozen hearts, no study has shown the epigenome and/or transcriptional regulation at the single-cell resolution in remodeled human myocardium. Nomura and colleagues demonstrate that the transcriptional profiles in isolated cardiomyocytes are well correlated with the bulk transcriptional profiles of murine hearts [61]. Interestingly, by examining the mRNA expression level of markers for cardiomyocytes and 11 non-myocytes in our bulk sequencing data, we observed that cardiomyocyte markers displayed around 28-225 fold higher expression levels than all other cell type markers, suggesting that the majority of signals at the transcriptional level was likely derived from

cardiomyocytes. It is important to note that cell-type-specific markers for cardiomyocytes and 11 non-myocyte cell types were based on findings derived from two single-cell sequencing studies using murine hearts. Although some studies showed conserved global gene expression between human and mouse hearts, others have shown a poor correlation between species [62–64]. Therefore, the analysis is limited in the way that using cell-type-specific markers of murine hearts may not translate one-to-one with marker expression on human cardiac cells. Additionally, these markers were not corrected to the number of cells per cell type. Yet, this observation is in line with the well correlated transcriptional profiles between cardiomyocytes and bulk cardiac tissues shown by Nomura *et al* [61]. Although the acetylation level of cardiomyocyte markers seemed to be slightly higher than other cell type markers, the profound mRNA expression pattern was not observed, suggesting additional regulatory machinery between histone acetylome and transcriptome. Additional studies on cell-type-specific enhancers are still needed to improve our understanding of the chromatin acetylation changes per cell composition. Nevertheless, top enriched TFBMs in cardiomyocytes from remodeled murine hearts after transverse aortic constriction were also obtained in our study [61]. The paracrine mechanisms between cardiomyocytes and non-cardiomyocytes play an important role in regulating cardiac function [65]. The direct cell-to-cell crosstalks (i.e. the TGF*β* signaling between cardiac fibroblasts and cardiomyocytes), and indirect cell-to-ECM crosstalks (i.e. via the integrin-related signaling), contribute throughout the (mal)adaptive process of cardiac remodeling [66]. Our data could offer valuable and extensive information on these biological crosstalks.

CONCLUSIONS

In conclusion, the identified H3K27ac chromatin regulation landscape shows significant changes in remodeled myocardium as compared to controls. Enriched genes and TFBMs in the differentially acetylated regions highlight known disease-associated processes, such as fibrosis. TFs that exhibited the same direction in their acetylation and mRNA expression changes might shed light on the upstream signaling by providing the candidates to build and validate putative key TF-DNA networks. Taken together, the dataset as presented here from our study provides valuable new information and aid in the discovery of key pathways and transcription mechanisms in the underlying disease etiology. These findings also demonstrate the value of genome-wide chromatin analysis in understanding the molecular regulation and etiology of myocardial remodeling and cardiovascular disease in general. With the development of single-cell sequencing techniques in biobanked patient material and the accumulating evidence of cell-to-cell

heterogeneity within the very same cell population in response to myocardial remodeling [29], we might gain knowledge on the chromatin profile at the single-cell resolution. This will allow us to examine gene expression, protein level, and the crosstalk between multiple cell types and to investigate the cell-to-cell variations after chromatin regulatory activity changes and structural reorganization, further contributing to the understanding of DNA mutations in regions with regulatory function and new drug target identification for future therapies.

MATERIALS AND METHODS

Study design and sample information

This study was approved by the Biobank Research Ethics Committee of University Medical Center Utrecht (protocol number 12/387), the Ethics Committee of UK National Research Ethics Service (07/H0715/101), and the Washington University School of Medicine Ethics Committee (Institutional Review Board). Written informed consent was obtained or in certain cases waived by the ethics committee when acquiring informed consent was not possible due to the death of the individual (control samples). All patients with cardiac remodeling due to AS were recruited prior to pre-operative assessment, which included a comprehensive evaluation with clinical history, resting blood pressure, 6-minute-walk test, electrocardiogram, transthoracic 2D-echocardiogram, and cardiac magnetic resonance (CMR). Inclusion criteria were patients >18 years with severe AS (2 or more of the following: aortic valve area<1 cm^2, peak pressure gradient >64 mmHg, mean pressure gradient >40 mmHg, aortic valve velocity ratio <0.25) undergoing aortic valve replacement ± coronary artery bypass grafting. Exclusion criteria were pregnancy/breastfeeding, estimated glomerular filtration rate <30 mL/min/1.73 m^2, CMR incompatible devices, inability to complete the protocol, previous valve surgery, or severe valve disease other than AS. Patient samples (n=20, 13 men and 7 women) were obtained by either intraoperative scalpel or Tru-Cut needle biopsy. Due to the difficulty in obtaining healthy donor hearts, the number of control samples was limited (n=5, 3 men and 2 women), and they were either autopsy material or non-transplanted donor hearts without signs of cardiac remodeling. Myocardial samples were collected and snap-frozen in liquid nitrogen, and stored at -80˚C. For all samples, 8 µm thick histological slices were stained for hematoxylin-eosin (**Supplementary Figure 1**). For detailed information per sample, please refer to **Supplementary Table 6A** and **6B**.

Chromatin immunoprecipitation and sequencing

To study the changes of histone acetylome between patient and control samples, we isolated chromatin from all frozen cardiac tissues. Briefly, all samples were sectioned at the thickness of 10 µm and chromatin was isolated using the MAGnify™ Chromatin Immunoprecipitation System kit (Life Technologies) according to the manufacturer's instructions. The anti-histone H3K27ac antibody (ab4729, Abcam) was used for immunoprecipitation. Captured DNA was purified using the ChIP DNA Clean & Concentrator kit (Zymo Research). Libraries were prepared using the NEXTflex™ Rapid DNA Sequencing Kit (Bioo Scientific). Samples were PCR amplified, checked for the proper size range and for the absence of adaptor dimers on a 2% agarose gel. Barcoded libraries were sequenced 75 bp single-end on Illumina NextSeq500 sequencer. Sequencing reads were mapped against the reference genome (hg19 assembly, NCBI37) using the BWA package (mem –t 7 –c 100 –M –R) [67]. Multiple reads mapping to the same location and strand were collapsed to single read and only uniquely placed reads were used for peak-calling. Peaks/regions were called using Cisgenome 2.0 (–e 150 -maxgap 200 –minlen 200) [68]. Region coordinates from all samples were stretched to at least 2000 base pairs and collapsed into a single common list. Overlapping regions were merged based on their outmost coordinates. Only regions supported by at least 2 independent datasets were further analyzed. Autosomal sequencing reads from each ChIP-seq library were overlapped back with the common region list to set the H3K27ac occupancy for every region-sample pair.

To further verify the detected H3K27ac signal corresponds to the transcription level, we incubated chromatins from 4 controls with anti-RBP1 antibody (PB-7G5), a subunit of RNA polymerase II (RNAPII), and performed the same ChIP-seq procedure [69]. To visualize the correlation between H3K27ac and RNAPII - log2 transformed Reads Per Kilobase per Million (RPKM) in the gene body were correlated to H3K27ac ChIP-seq signal on gene promoters (+/- 2.5 kb around the TSS). For more information, please refer to **Supplementary Figure 3** and **Supplementary Table 7**.

Identification of differentially acetylated regions

Next, we identified regions with different H3K27ac occupancy levels between patients and controls using DESeq2 with the standard settings (FDR Benjamini & Hochberg $P_{adj.}<0.05$) [70]. Supervised hierarchical clustering was performed with quantile normalized (limma::normalizeQantiles() function in R), $\log_2$ transformed and median centered read counts per common region. To avoid the $\log_2$ transformation of zero values, one RPKM was added to each region. Genomic regions in which hyper- or hypoacetylated

peaks/regions located in patients when compared with controls were defined based on the cut-off of adjusted P<0.05 and are referred to as 'DARs'. To visualize the clustering of the samples based on the median centered and $\log_2$ transformed levels of acetylation in DARs, we used heatmap.2() command from gplots package in R.

Identification of tandem regulated chromatin domains

Next, we studied the direction of the acetylation changes in DARs to verify the chromatin spatio-temporal dynamics that play an important role in gene regulation. Adjacent regions that had the same fold change (FC) direction and the absolute $\log_2$FC >0.4 were considered as possible (sub)compartments due to the spatial organization of the genome. Control datasets were generated by shuffling the acetylation fold change and significance values while retaining the locations and distances of adjacent peaks.

Region-to-gene annotation and Gene Ontology enrichment analysis

We also performed *in silico* region-to-gene annotation using a conservative window of ±5kb from the transcription start site (TSS). Annotated gene sets were studied for their biological functions using ToppGene Suite tool ToppFun (default setting: Probability density function, FDR correction, p-value cut off of 0.05 and gene limit set between and including 1 and 2000 per pathway) [71]. STRING was used to identify the key gene-encoded proteins and their interactions within selected gene sets (minimum required interaction score: high confidence (0.700)) [72].

Discovery of transcription factor binding motifs and their putative networks

In DARs, we also identified enriched transcription factor binding motifs (TFBMs) that could elucidate the putative networks between motif-encoded transcription factors and genes. Briefly, repeat masked sequences of the DARs were first overlapped with DNAse hypersensitivity site analysis from cardiac samples from the ENCODE project [28]. DNA sequences of those overlapping DNAse sites and control shuffled sequences were taken by MEME Suite AME tool [73] under the following settings: motif database: human (HOCOMOCO v10); background model frequencies:0.25,0.25,0.25,0.25; total pseudo-count added to a motif columns: 0.25; Wilcoxon rank-sum test with threshold P<0.05; number of multiple tests for

Bonferroni correction: #Motifs × #Partitions Tested = 641 × 1 = 641. By overlapping identified motifs and annotated genes from DARs, we obtained a list of genes encoding for TFs that were potentially master regulators of myocardial remodeling at the chromatin level.

RNA sequencing and gene expression analyses

Besides the correlation between the H3K27ac signal and RNAPII occupancy level, we also examined the H3K27ac signal in correlation with the transcriptome level as revealed by RNA sequencing (RNA-seq). Briefly, RNA was isolated from 4 controls according to the manufacturer's instructions (BioLine) with minor changes. Libraries were generated using NEXTflex™ Rapid RNA-seq Kit (Bio Scientific) and sequenced by the Nextseq500 platform (Illumina). Sequenced reads were annotated as described previously [69]. Reads Per Kilobase Million (RPKM) that presented in all samples were included, and genes with the mean RPKM>0.5 were considered expressed. RPKM on the $\log_2$ scale were correlated to call reads in H3K27ac ChIP-seq using R.

To compare the transcriptional landscapes between patient and control samples, we performed RNA-seq using the CEL-seq protocol - an adapted single-cell RNA-seq method that overcomes the challenge of the low input materials due to the limited biopsy size [74]. Briefly, 10ng RNA from each sample was used. Primer design, linear mRNA amplification, library construction, and sequencing were performed as described previously [75]. Raw read file per sample was used and differentially expressed genes between patient and control hearts were identified using the DESeq2 [70] within the Galaxy environment under the default settings. A p-value cutoff of 0.05 was used.

Gene set enrichment analysis (GSEA)

We performed gene set enrichment analyses using GSEA software v3.0 [76] and studied the enrichment of annotated genes from DARs in relation to the gene expression levels between patient and control hearts. TFs, from which their TFBMs were enriched in DARs, were also examined in the transcriptome profile. Briefly, we included genes with significantly different expression levels as revealed by the CEL-seq in the present study and two published studies using microarray [11,12]. Differentially expressed genes in each of three transcriptome profiles were ranked based on their fold change and uploaded to GSEA as the

expression datasets. The annotated gene set from hyper- or hypoacetylated regions was accessed for its positive or negative enrichment throughout each expression dataset with the standard settings.

Hypertrophic markers in cardiomyocytes during remodeling

We also investigated candidates that could play an important role in cardiomyocytes from the bulk sequencing data. Briefly, a list of genes that showed the same direction of changes at the histone acetylation level and mRNA expression level were selected from our study. They were overlapped with genes with altered mRNA expression levels in hypertrophic human stem cell-derived cardiomyocytes due to mechanical stress [33].

AVAILABILITY OF DATA AND MATERIALS

All relevant data are available within the article and the supplementary files. Because of the sensitive nature of the data collected for this study, requests to access the dataset from qualified researchers trained in human subject confidentiality protocols may be sent to the corresponding authors.

DECLARATIONS

Competing interests

None.

Funding

This work was supported by the Netherlands Foundation for Cardiovascular Excellence (to C.C.), the NWO VENI grant (no. 016.176.136 to M.H.), two NWO VIDI grants (no. 91714302 to C.C. and no. 016096359 to M.C.V), the Erasmus MC fellowship grant (to C.C.), the RM fellowship grant of the UMC Utrecht (to C.C.), Wilhelmina Children's Hospital research funding (no. OZF/14 to M.H.), the Netherlands Cardiovascular Research Initiative: An initiative with the support of the Dutch Heart Foundation (CVON2014-40 DOSIS: to J.P., M.H., F.W.A., and CVON2014-11 RECONNECT: to C.C., M.C.V., G.P., M.M., D.D.), and Dutch Heart Foundation (Queen of Heart: to C.C., H.D.R., M.C.V, G.P.), UCL Hospitals NIHR Biomedical Research Centre

(to F.W.A.), the doctoral research fellowships by the National Institute of Health Research (NIHR; DRF-2013-06-102 to T.A.T.), and National Institutes of Health (USA) grant LM010098 (to J.H.M.).

Authors' contributions

M.H., M.M., C.C., and F.W.A. designed the study; J.P. and M.H. wrote the original draft; J.P., M.H., and N.vD. performed experiments; M.M. processed the raw sequencing data; J.P., M.H., J.V.D, and M.M. analyzed and interpreted the sequencing data; G.K. analyzed microarray data; T.A.T., T.L., B.J.B., I.R.E., J.C.M., R.dW., M.M.H., and A.V. collected and provided human myocardium; A.G and B.L. provided RNA samples of the same human myocardium; HdR, G.P., D.J.D., E.E.S.N., J.H.M., M.C.V., H.E.A., C.G.M.vD, and M.M.K provided technical support and conceptual advice; all authors discussed the results and implications and provided comments to the final version of the manuscript.

Acknowledgments

We thank Joyce van Kuik and Erica Sierra-de Koning for their laboratory support. We thank Utrecht Sequencing Facility for support in processing ChIP-seq data, which is partially subsidized by Hubrecht Laboratory, Utrecht University, and UMC Utrecht.

Supplementary Files – Online

Supplementary Figure 1: An overview of the workflow in this study. †Detailed information of samples used in H3K27ac ChIP-seq, RNAPII ChIP-seq, and RNA-seq are listed in Supplementary Table 2. *: Standard RNA-seq and adjusted RNA-seq (3'-RNA-seq) were both performed, detailed information is shown in Supplementary Table 7.

Supplementary Figure 2: Histology of cardiac tissues used in this study. Overview of representative slides stained with hematoxylin-eosin from cardiac samples in control and patient groups are shown. Higher magnification showing normal myocardium in control (panel A) and patient myocardium with hypertrophy of cardiomyocytes and interstitial fibrosis in the patient sample (panel B, both 40x magnification).

Supplementary Figure 3: Correlation plots between H3K27ac ChIP-seq (red) with RNAPII ChIP-seq (blue) and RNAseq datasets (blue), respectively. Correlations in four control samples (from **a** to **d**) used in this study. Gene expression rank is based on RPKM values.

Supplementary Figure 4: The mRNA expression level of cell-type-specific markers. Genes labeled red and blue were significantly up- and down-regulated in patients versus controls ($P<0.05$).

Supplementary Figure 5: Gene Ontology (GO) analysis of genes annotated to DARs using ToppFun. a Molecular function enrichment for hyperacetylated regions, **b** Biological process for hyperacetylated regions, **c** Cellular component for hyperacetylated regions, **d** Molecular function for hypoacetylated regions, **e** Biological process for hypoacetylated regions, **f** Cellular component for hypoacetylated regions.

Supplementary Figure 6: Sex-specific H3K27ac acetylome and transcriptome profiles between male and female patients with concentric remodeling. a Principal component analysis (PCA) plot showing the clustering of man and woman cardiac samples based on H3K27ac profiles (using 500 regions with the highest variance). **b** PCA plot showing the clustering of man and woman cardiac samples based on the transcriptome profiles (using 500 genes with the highest variance).

Supplementary Table 1: Regions with different H3K27ac occupancy in patients compared with the control group (adjusted p-value <0.05).

Supplementary Table 2: a Identified active tandem regulated chromatin domains (TRCDs) in the short-range (5Kb-100Kb) in patients when compared with controls. **b** Identified active TRCDs in the long-range (100Kb-1Mb) in patients when compared with controls. **C** Identified repressed TRCDs in the short-range (5Kb-100Kb) in patients when compared with controls. **d** Identified repressed TRCDs in the long-range

(100Kb-1Mb) in patients when compared with controls. **e** Randomized TRCDs in the short-range (5Kb-100Kb) in patients when compared with controls. **f** Randomized TRCDs in the long-range (100Kb-1Mb) in patients when compared with controls.

Supplementary Table 3: a Genes in the vicinity of hyperacetylated regions using a +/-5kb window. **b** Genes in the vicinity of hypoacetylated regions using a +/-5kb window.

Supplementary Table 4: Differentially expressed genes between patients and controls.

Supplementary Table 5: Identified transcription binding motifs from differentially acetylated regions.

Supplementary Table 6: a An overview of general information of all included human cardiac samples. **b** An overview of detailed clinical parameters of the included AS patients.

Supplementary Table 7: Overview of included samples in H3K27ac ChIP-seq, RNAPII ChIP-seq, and RNA-seq experiments.

Supplementary Table 8: The chromatin acetylation changes and the mRNA expression changes of key regulators and targets in the Hippo pathway.

REFERENCES

1. Thaman R, Gimeno JR, Reith S, Esteban MTT, Limongelli G, Murphy RT, et al. Progressive left ventricular remodeling in patients with hypertrophic cardiomyopathy and severe left ventricular hypertrophy. J Am Coll Cardiol. 2004;44:398–405.

2. Burchfield JS, Xie M, Hill JA. Pathological ventricular remodeling: mechanisms: part 1 of 2. Circulation. 2013;128:388–400.

3. Kehat I, Molkentin JD. Molecular Pathways Underlying Cardiac Remodeling During Pathophysiological Stimulation. Circulation. 2010;122:2727–35.

4. Lau E, Cao Q, Lam MPY, Wang J, Ng DCM, Bleakley BJ, et al. Integrated omics dissection of proteome dynamics during cardiac remodeling. Nat Commun. 2018;9:120.

5. Sassi Y, Avramopoulos P, Ramanujam D, Grüter L, Werfel S, Giosele S, et al. Cardiac myocyte miR-29 promotes pathological remodeling of the heart by activating Wnt signaling. Nat Commun. 2017;8:1614.

6. Papait R, Cattaneo P, Kunderfranco P, Greco C, Carullo P, Guffanti A, et al. Genome-wide analysis of histone marks identifying an epigenetic signature of promoters and enhancers underlying cardiac hypertrophy. Proceedings of the National Academy of Sciences. 2013;110:20164–9.

7. Rice J. Animal models: Not close enough. Nature. Nature Publishing Group; 2012;484:S9.

8. Seok J, Warren HS, Cuenca AG, Mindrinos MN, Baker HV, Xu W, et al. Genomic responses in mouse models poorly mimic human inflammatory diseases. Proc Natl Acad Sci U S A. 2013;110:3507–12.

9. Vegter EL, Ovchinnikova ES, Silljé HHW, Meems LMG, van der Pol A, van der Velde AR, et al. Rodent heart failure models do not reflect the human circulating microRNA signature in heart failure. PLoS One. Public Library of Science; 2017;12:e0177242.

10. Meder B, Haas J, Sedaghat-Hamedani F, Kayvanpour E, Frese K, Lai A, et al. Epigenome-Wide Association Study Identifies Cardiac Gene Patterning and a Novel Class of Biomarkers for Heart Failure. Circulation. 2017;136:1528–44.

11. Kararigas G, Dworatzek E, Petrov G, Summer H, Schulze TM, Baczko I, et al. Sex-dependent regulation of fibrosis and inflammation in human left ventricular remodelling under pressure overload. Eur J Heart Fail. 2014;16:1160–7.

12. Petretto E, Sarwar R, Grieve I, Lu H, Kumaran MK, Muckett PJ, et al. Integrated genomic approaches implicate osteoglycin (Ogn) in the regulation of left ventricular mass. Nat Genet. 2008;40:546–52.

13. Heinig M, Adriaens ME, Schafer S, van Deutekom HWM, Lodder EM, Ware JS, et al. Natural genetic variation of the cardiac transcriptome in non-diseased donors and patients with dilated cardiomyopathy. Genome Biol. 2017;18:170.

14. Ren C-W, Liu J-J, Li J-H, Li J-W, Dai J, Lai Y-Q. RNA-seq profiling of mRNA associated with hypertrophic cardiomyopathy. Mol Med Rep. 2016;14:5573–86.

15. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. Cell Res. 2011;21:381–95.

16. Karlic R, -R. Chung H, Lasserre J, Vlahovicek K, Vingron M. Histone modification levels are predictive for gene expression. Proceedings of the National Academy of Sciences. 2010;107:2926–31.

17. Gilsbach R E al. Distinct epigenetic programs regulate cardiac myocyte development and disease in the human heart in vivo. - PubMed - NCBI [Internet]. [cited 2019 Mar 28]. Available from: https://www.ncbi.nlm.nih.gov/pubmed/29374152

18. pubmeddev, Berezin A. Epigenetics in heart failure phenotypes. - PubMed - NCBI [Internet]. [cited 2019 Jul 15]. Available from: https://www.ncbi.nlm.nih.gov/pubmed/27335803

19. Raha D, Hong M, Snyder M. ChIP-Seq: a method for global identification of regulatory elements in the genome. Curr Protoc Mol Biol. 2010;Chapter 21:Unit 21.19.1–14.

20. Sun W, Poschmann J, Cruz-Herrera Del Rosario R, Parikshak NN, Hajan HS, Kumar V, et al. Histone Acetylome-wide Association Study of Autism Spectrum Disorder. Cell. 2016;167:1385–97.e11.

21. Ooi WF, Xing M, Xu C, Yao X, Ramlee MK, Lim MC, et al. Epigenomic profiling of primary gastric adenocarcinoma reveals super-enhancer heterogeneity. Nat Commun. 2016;7:12983.

22. Shlyueva D, Stampfel G, Stark A. Transcriptional enhancers: from properties to genome-wide predictions. Nat Rev Genet. 2014;15:272–86.

23. Sengupta D, Kannan A, Kern M, Moreno MA, Vural E, Stack B, et al. Disruption of BRD4 at H3K27Ac-enriched enhancer region correlates with decreased c-Myc expression in Merkel cell carcinoma. Epigenetics. 2015;10:460–6.

24. Akhtar-Zaidi B, middle dot, Corradin O, Saiakhova A, Bartels CF, Balasubramanian D, et al. Epigenomic Enhancer Profiling Defines a Signature of Colon Cancer. Science. 2012;336:736–9.

25. Gates LA, Foulds CE, O'Malley BW. Histone marks in the "drivers seat": functional roles in steering the transcription cycle. Trends Biochem Sci. NIH Public Access; 2017;42:977.

26. Bogu GK, Vizán P, Stanton LW, Beato M, Di Croce L, Marti-Renom MA. Chromatin and RNA Maps Reveal Regulatory Long Noncoding RNAs in Mouse. Mol Cell Biol. American Society for Microbiology Journals; 2016;36:809–19.

27. Varshney A, Scott LJ, Welch RP, Erdos MR, Chines PS, Narisu N, et al. Genetic regulatory signatures underlying islet gene expression and type 2 diabetes. Proc Natl Acad Sci U S A. 2017;114:2301–6.

28. Rosenbloom KR, Sloan CA, Malladi VS, Dreszer TR, Learned K, Kirkup VM, et al. ENCODE data in the UCSC Genome Browser: year 5 update. Nucleic Acids Res. 2013;41:D56–63.

29. Harakalova M, Asselbergs FW. Systems analysis of dilated cardiomyopathy in the next generation sequencing era. Wiley Interdiscip Rev Syst Biol Med. 2018;10:e1419.

30. Treibel TA E al. Reverse Myocardial Remodeling Following Valve Replacement in Patients With Aortic Stenosis. - PubMed - NCBI [Internet]. [cited 2018 Jul 27]. Available from: https://www.ncbi.nlm.nih.gov/pubmed/29471937

31. Achinger-Kawecka J, Clark SJ. Disruption of the 3D cancer genome blueprint. Epigenomics. 2017;9:47–55.

32. Skelly DA E al. Single-Cell Transcriptional Profiling Reveals Cellular Diversity and Intercommunication in the Mouse Heart. - PubMed - NCBI [Internet]. [cited 2019 Apr 4]. Available from: https://www.ncbi.nlm.nih.gov/pubmed/29346760

33. Ovchinnikova E, Hoes M, Ustyantsev K, Bomer N, de Jong TV, van der Mei H, et al. Modeling Human Cardiac Hypertrophy in Stem Cell-Derived Cardiomyocytes. Stem Cell Reports. Elsevier; 2018;10:794.

34. Lupiáñez DG, Spielmann M, Mundlos S. Breaking TADs: How Alterations of Chromatin Domains Result in Disease. Trends Genet. 2016;32:225–37.

35. van Steensel B, Furlong EEM. The role of transcription in shaping the spatial organization of the genome. Nat Rev Mol Cell Biol. Nature Publishing Group; 2019;1.

36. Rosa-Garrido M, Chapski DJ, Schmitt AD, Kimball TH, Karbassi E, Monte E, et al. High-Resolution Mapping of Chromatin Conformation in Cardiac Myocytes Reveals Structural Remodeling of the Epigenome in Heart Failure. Circulation. 2017;136:1613–25.

37. Zheng Y, Thomas PM, Kelleher NL. Measurement of acetylation turnover at distinct lysines in human histones identifies long-lived acetylation sites. Nat Commun. 2013;4:2203.

38. Perrino C, Barabási A-L, Condorelli G, Davidson SM, De Windt L, Dimmeler S, et al. Epigenomic and transcriptomic approaches in the post-genomic era: path to novel targets for diagnosis and therapy of the ischaemic heart? Position Paper of the European Society of Cardiology Working Group on Cellular Biology of the Heart. Cardiovasc Res. Narnia; 2017;113:725–36.

39. Fan D, Takawale A, Lee J, Kassiri Z. Cardiac fibroblasts, fibrosis and extracellular matrix remodeling in heart disease. Fibrogenesis Tissue Repair. 2012;5:15.

40. Creemers EE, Pinto YM. Molecular mechanisms that control interstitial fibrosis in the pressure-overloaded heart. Cardiovasc Res. 2010;89:265–72.

41. Kapur NK. Transforming growth factor-β: governing the transition from inflammation to fibrosis in heart failure with preserved left ventricular function. Circ. Heart Fail. 2011. p. 5–7.

42. Schafer S, Viswanathan S, Widjaja AA, Lim W-W, Moreno-Moral A, DeLaughter DM, et al. IL-11 is a crucial determinant of cardiovascular fibrosis. Nature. Nature Publishing Group; 2017;552:110.

43. Dworatzek E, Baczko I, Kararigas G. Effects of aging on cardiac extracellular matrix in men and women. Proteomics Clin Appl. 2016;10:84–91.

44. Cardiomyocytes facing fibrotic conditions re-express extracellular matrix transcripts. Acta Biomater. Elsevier; 2019;89:180–92.

45. Esmerats JF, Heath J, Rezvan A, Jo H. Hemodynamics and Mechanobiology of Aortic Valve Calcification. Biomedical Engineering: Frontier Research and Converging Technologies. Springer, Cham; 2016. p. 237–61.

46. Gloria Garoffolo MP. Mechanotransduction in the Cardiovascular System: From Developmental Origins to Homeostasis and Pathology. Cells [Internet]. Multidisciplinary Digital Publishing Institute (MDPI); 2019 [cited 2020 May 21];8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6953076/

47. Ishani Dasgupta DM. Control of cellular responses to mechanical cues through YAP/TAZ regulation. J Biol Chem. American Society for Biochemistry and Molecular Biology; 2019;294:17693.

48. Cobbaut M, Karagil S, Bruno L, de la Loza MDCD, Mackenzie FE, Stolinski M, et al. Dysfunctional Mechanotransduction through the YAP/TAZ/Hippo Pathway as a Feature of Chronic Disease. Cells [Internet]. Multidisciplinary Digital Publishing Institute (MDPI); 2020 [cited 2020 May 21];9. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7016982/

49. pubmeddev, Patrizio M And Marano. Gender differences in cardiac hypertrophic remodeling. - PubMed - NCBI [Internet]. [cited 2020 May 19]. Available from: https://www.ncbi.nlm.nih.gov/pubmed/27364397

50. Kessler EL, Rivaud MR, Vos MA, van Veen TAB. Sex-specific influence on cardiac structural remodeling and therapy in cardiovascular disease. Biol Sex Differ. BioMed Central; 2019;10:1–11.

51. Deschepper CF, Llamas B. Hypertensive Cardiac Remodeling in Males and Females [Internet]. Hypertension. 2007. p. 401–7. Available from: http://dx.doi.org/10.1161/01.hyp.0000256279.49882.d8

52. Regitz-Zagrosek V, Kararigas G. Mechanistic Pathways of Sex Differences in Cardiovascular Disease. Physiol Rev. 2017;97:1–37.

53. pubmeddev, Kararigas G E al. Sex-dependent regulation of fibrosis and inflammation in human left ventricular remodelling under pressure overload. - PubMed - NCBI [Internet]. [cited 2020 May 19]. Available from: https://www.ncbi.nlm.nih.gov/pubmed/25287281

54. pubmeddev, Gaignebet L E al. Sex-Specific Human Cardiomyocyte Gene Regulation in Left Ventricular Pressure Overload. - PubMed - NCBI [Internet]. [cited 2020 May 19]. Available from: https://www.ncbi.nlm.nih.gov/pubmed/31954524

55. Prognostic significance of left ventricular concentric remodelling in patients with aortic stenosis. Arch Cardiovasc Dis. Elsevier Masson; 2017;110:26–34.

56. pubmeddev, Kuster DW E al. Left ventricular remodeling in swine after myocardial infarction: a transcriptional genomics approach. - PubMed - NCBI [Internet]. [cited 2020 Jan 3]. Available from: https://www.ncbi.nlm.nih.gov/pubmed/22057716

57. pubmeddev, Kuster DW E al. Gene reprogramming in exercise-induced cardiac hypertrophy in swine: A transcriptional genomics approach. - PubMed - NCBI [Internet]. [cited 2020 Jan 3]. Available from: https://www.ncbi.nlm.nih.gov/pubmed/25451387

58. Europe PMC. Update on the Pathogenic Implications and Clinical Potential of microRNAs in Cardiac Disease. - Abstract - Europe PMC [Internet]. [cited 2020 Jan 3]. Available from: http://dx.doi.org/10.1155/2015/105620

59. Endothelial PAS Domain Protein 1 (EPAS1) Induces Adrenomedullin Gene Expression in Cardiac Myocytes: Role of EPAS1 in an Inflammatory Response in Cardiac Myocytes. J Mol Cell Cardiol. Academic Press; 2002;34:739–48.

60. Jia G, Preussner J, Chen X, Guenther S, Yuan X, Yekelchyk M, et al. Single cell RNA-seq and ATAC-seq analysis of cardiac progenitor cell transition states and lineage settlement. Nat Commun. Nature Publishing Group; 2018;9:4877.

61. Nomura S, Satoh M, Fujita T, Higo T, Sumida T, Ko T, et al. Cardiomyocyte gene programs encoding morphological and functional signatures in cardiac hypertrophy and failure. Nat Commun. Nature Publishing Group; 2018;9:4435.

62. Monaco G, van Dam S, João Luis Casal, Larbi A, de Magalhães JP. A comparison of human and mouse gene co-expression networks reveals conservation and divergence at the tissue, pathway and disease levels. BMC Evol Biol [Internet]. BioMed Central; 2015 [cited 2020 May 19];15. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4654840/

63. Zheng-Bradley X, Rung J, Parkinson H, Brazma A. Large scale comparison of global gene expression patterns in human and mouse. Genome Biol. BioMed Central; 2010;11:1–11.

64. Hideki Uosaki Y-HT. Comparative Gene Expression Analysis of Mouse and Human Cardiac Maturation. Genomics Proteomics Bioinformatics. Elsevier; 2016;14:207.

65. Mayourian J, Ceholski DK, Gonzalez DM, Cashman TJ, Sahoo S, Hajjar RJ, et al. Physiologic, Pathologic, and Therapeutic Paracrine Modulation of Cardiac Excitation-Contraction Coupling. Circ Res. NIH Public Access; 2018;122:167.

66. Fountoulaki K, Dagres N, Iliodromitis EK. Cellular Communications in the Heart. Cardiac Failure Review. Radcliffe Cardiology; 2015;1:64.

67. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25:1754–60.

68. Ji H, Jiang H, Ma W, Johnson DS, Myers RM, Wong WH. An integrated software system for analyzing ChIP-chip and ChIP-seq data. Nat Biotechnol. 2008;26:1293–300.

69. Mokry M, Hatzis P, Schuijers J, Lansu N, Ruzius F-P, Clevers H, et al. Integrated genome-wide analysis of transcription factor occupancy, RNA polymerase II binding and steady-state RNA levels identify differentially regulated functional gene classes. Nucleic Acids Res. 2012;40:148–58.

70. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550.

71. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. Nucleic Acids Res. 2009;37:W305–11.

72. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 2015;43:D447–52.

73. McLeay RC, Bailey TL. Motif Enrichment Analysis: a unified framework and an evaluation on ChIP data. BMC Bioinformatics. 2010;11:165.

74. Hashimshony T, Senderovich N, Avital G, Klochendler A, de Leeuw Y, Anavy L, et al. CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. Genome Biol. BioMed Central; 2016;17:1–7.

75. Marcia A M, Rho HS, Hemerich D, Henning HHW, van Tol HTA, Hölker M, et al. An oviduct-on-a-chip provides an enhanced in vitro environment for zygote genome reprogramming. Nat Commun. Nature Publishing Group; 2018;9:4934.

76. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. National Academy of Sciences; 2005;102:15545–50.

Chapter 5

Multi-omics integration delineates the key regulators and pathomechanisms in hypertrophic cardiomyopathy due to truncating MYBPC3 mutations

J. Pei[1,2,3‡], M. Schuldt[4‡], E. Nagyova[5‡], Z. Gu[6], S. el Bouhaddani[6], M. Jansen[7], J.J.A. Calis[1,2], L.M. Dorsch[4], N.A.M. van den Dungen[5], N. Lansu[5], B.J. Boukens[8], I.R. Efimov[9], M. Michels[10], M.C. Verhaar[2,3], R. de Weger[11], A. Vink[11], F.G. van Steenbeek[1,2,12], A.F. Baas[7], H.W. Uh[6], D.W.D Kuster[4], C. Cheng[2,3,10], M. Mokry[1,2,5,13], J. van der Velden[4*], F.W. Asselbergs[1,14,15*], M. Harakalova[1,2*]

‡Equal contribution of first authors, *Equal contribution of last authors

1Department of Cardiology, Division Heart & Lungs, UMC Utrecht, University of Utrecht, NL;
2Regenerative Medicine Utrecht (RMU), UMC Utrecht, University of Utrecht, NL;
3Department of Nephrology and Hypertension, DIG-D, UMC Utrecht, University of Utrecht, NL;
4Department of Physiology, Amsterdam UMC, Vrije Universiteit Amsterdam, NL;
5Laboratory of Clinical Chemistry and Hematology, UMC Utrecht, NL;
6Department of Biostatistics and Research Support, UMC Utrecht, University of Utrecht, NL;
7Department of Genetics, Division of Laboratories, Pharmacy and Biomedical Genetics, UMC Utrecht, University of Utrecht, NL;
8Department of Medical Biology, AMC, Amsterdam, NL;
9Department of Biomedical Engineering, GWU, US;
10Department of Cardiology, Thoraxcentre, Erasmus Medical Centre, Rotterdam, the Netherlands;
11Department of Pathology, UMC Utrecht, University of Utrecht, NL;
12Department of Clinical Sciences, Faculty of Veterinary Medicine, University of Utrecht, NL;
13Division of Paediatrics, UMC Utrecht, University of Utrecht, NL;
14Health Data Research UK and Institute of Health Informatics, University College London, London, United Kingdom;
15Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College London, London, United Kingdom.

Manuscript under submission

ABSTRACT

Background: Hypertrophic cardiomyopathy (HCM) is the most common genetic disease of the cardiac muscle, frequently caused by mutations in MYBPC3. However, little is known about the upstream pathways and key regulators causing the disease. Therefore, we employed a multi-omics approach to study the pathomechanisms underlying HCM comparing patient hearts harboring MYBPC3 mutations to control hearts.

Results: Using H3K27ac ChIP-seq and RNA-seq we obtained 9,310 differentially acetylated regions (DARs) and 2,033 differentially expressed genes, respectively, between 13 HCM and 10 control hearts. We obtained 441 differentially expressed proteins between 11 HCM and 8 control hearts using proteomics. By integrating multi-omics datasets, we identified a set of DNA regions and genes that differentiate HCM from control hearts and 53 protein-coding genes as the major contributors. This comprehensive analysis consistently points towards altered extracellular matrix formation, muscle contraction, and metabolism. Therefore, we studied enriched transcription factor (TF) binding motifs and identified 9 motif-encoded TFs, including KLF15, ETV4, AR, CLOCK, ETS2, GATA5, MEIS1, RXRA, and ZFX. Furthermore, we observed an abundance of acetylation signals and transcripts derived from cardiomyocytes compared to non-myocyte populations in the heart.

Conclusions: By integrating histone acetylome, transcriptome, and proteome profiles, we identified major effector genes and protein networks that drive the pathological changes in HCM caused by mutated MYBPC3. Our work identifies 38 highly affected protein-coding genes as potential plasma HCM biomarkers and 9 TFs as potential upstream regulators of these pathomechanisms that may serve as possible therapeutic targets.

Keywords: HCM, MYBPC3, histone acetylome, transcriptome, proteome, transcription factors

INTRODUCTION

Hypertrophic cardiomyopathy (HCM), characterized by thickening of the myocardium that is not explained by abnormal loading conditions, is the most common inherited cardiac disease [1]. More than 1,500 associated mutations, primarily in genes encoding parts of the sarcomere, such as MYBPC3, MYH7, and TNNT2, have been identified. The remodeled myocardium is characterized by cardiomyocyte hypertrophy and disarray, extensive fibrosis, and reduced capillary density [2]. HCM is a heterogeneous disease as the onset, the disease phenotype, and the severity of the clinical presentations differ greatly among mutation carriers [3]. Additionally, different mutated genes exhibit distinct biological impact on cellular contractility, energy metabolism, and sarcomeric protein expression in HCM hearts [4]. However, the driving pathological mechanisms underlying the heterogenous HCM remains largely unknown.

Next-generation sequencing technologies (NGS) are instrumental in understanding disease etiology, delivering a clinical diagnosis, and discovering new treatment options. Multiple studies employed NGS to reveal the epigenetic modifications in heart failure, including DNA methylation and histone (de-)acetylation, which provided insights and identification of the driving mechanism underlying the disease [5,6]. We previously showed the influence of histone acetylation changes on QRS complex-related GWAS loci in HCM [7,8]. Studies also employed RNA sequencing to identify the affected transcription factor-mediated upstream regulatory events and the distinct gene expressions that define heart failure [9,10]. Furthermore, data-piling studies are now connecting proteomics to NGS to get comprehensive information on the disease biology for precision medicine [9].

In this study, we aim to understand the pathomechanisms driving HCM by employing a multi-omics approach, including chromatin immunoprecipitation sequencing (ChIP-seq), RNA sequencing (RNA-seq), and proteomics, using myocardial tissue obtained from clinically well-phenotyped HCM patients with truncating MYBPC3 mutations and compared these with non-failing donor hearts. We revealed altered histone acetylome, transcriptome, and proteome profiles in HCM versus control hearts and studied affected biological functions. Besides, we identified key factors that may play a critical role in regulating the pathomechanisms underlying HCM. We also evaluated the contribution of histone acetylation and transcription signals in 11 cell types in the heart. Combined, this multi-omics study gives insight into the underlying disease pathways driving HCM and identifies promising candidates for therapeutic strategies.

RESULTS

Pairwise comparison between HCM and control hearts reveals distinct histone acetylome profiles

H3K27ac ChIP-seq was used to capture acetylated DNA regions in each sample and to compare the acetylation levels between 13 HCM (n=13) and 10 control hearts (Figure 1A and Supplementary Table 1) using DESeq2. In total, we identified 4,226 regions presented higher acetylation levels (hyperacetylated regions) and 5,084 regions presented lower acetylation levels (hypoacetylated regions) in HCM versus control hearts (Padj.<0.05, Figure 1B, Supplementary Table 2A). Examples of hyper- and hypoacetylated regions are shown in Supplementary Figure 1. Region-to-gene annotation using either a 5kb or 50kb window from the transcription start site (TSS) revealed genes in the hyperacetylated regions are mostly involved in muscle contraction and extracellular matrix (ECM) development-related processes, whereas genes in the hypoacetylated regions are mostly involved in metabolic processes (Supplementary Table 2B-2E).

Pairwise comparison between HCM and control hearts reveals distinct transcriptome profiles

Following RNA-seq of the same biopsies, the transcriptional profiles between 13 HCM (n=13) and 10 control hearts were compared using DESeq2. In total, we identified 936 up-regulated genes and 1,097 down-regulated genes in HCM hearts compared to controls (Padj.<0.05, Figure 1C, Supplementary Table 3A). The top biological processes enriched by up-regulated genes are involved in the muscle system process and energy production. The top biological processes enriched by down-regulated genes are involved in lipid metabolism and cell adhesion (Supplementary Table 3B and 3C).

Pairwise comparison between HCM and control hearts reveals distinct proteome profiles

We also performed proteomics in 11 HCM samples and compared their protein expression levels to another control group (n=8) using DESeq2. In total, we identified 216 up-regulated proteins and 225 down-regulated proteins in HCM hearts compared to controls (P<0.05, Figure 1D, Supplementary Table 4A). The top enriched biological processes by up-regulated proteins are involved in muscular and ECM development. The top enriched biological processes by down-regulated proteins are involved in metabolism (Supplementary Figure 2, Supplementary Table 4B, and 4C).

Integrating histone acetylome, transcriptome, and proteome changes in HCM

We also integrated H3K27ac ChIP-seq and RNA-seq data in an unsupervised manner using O2PLS. Notably, without predefining the patient and the control groups, the first joint component of both ChIP-seq and RNA-seq plots discriminated HCM from control hearts (Figure 1E and 1F). DNA regions and genes that contributed to the separation between HCM and control hearts are shown in Supplementary Table 5. Next, we aimed to identify key genes that underlie the separation between two groups and show altered expression levels in HCM hearts versus controls. Therefore, we overlapped 2,000 genes that discriminate HCM hearts and controls from the integrative analysis (O2PLS) with the differentially expressed genes from the pairwise comparison with the top and obtained 618 up-regulated genes and 308 down-regulated genes (Figure 1G and Supplementary Table 6A). These overlapping genes are enriched for biological processes in the circulatory system and muscle contraction and pathways involved in the ECM formation and complement system (Figure 1H, Supplementary Table 6B).

Since only 11 out of 13 HCM samples were included in the proteomics experiment and a different set of control samples was used in comparison with the HCM group, we could not apply O2PLS to integrate the proteomic data with either H3K27ac ChIP-seq or RNA-seq data. We, therefore, overlapped differentially expressed genes supported by DESeq2 and O2PLS analyses with proteomic data and identified 36 up-regulated protein-coding genes and 17 down-regulated protein-coding genes in HCM versus control hearts (Table 1). Notably, the protein levels of 38 protein-coding genes are detectable in the plasma and 5 out of them were consistently changed in the same direction in HCM versus control hearts at DNA (5kb from TSS), RNA, and protein levels, including the up-regulation of ASPN, FMOD, MCAM, and NPPA and the down-regulation of AASS, highlighting them as promising candidates for biomarker discovery in HCM (Table 1).

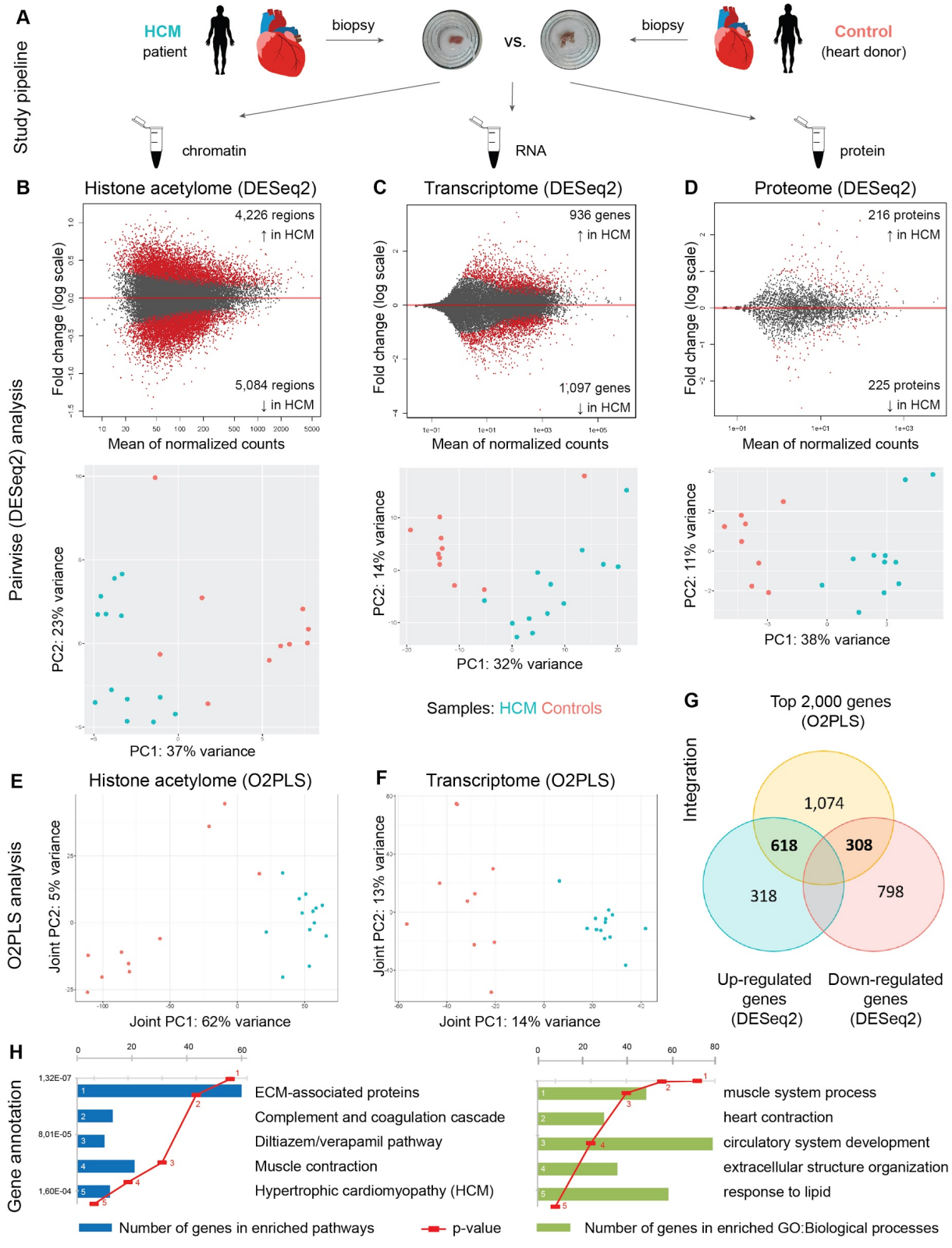


Figure 1. Pairwise analyses using DESeq2 and integrative analyses using unsupervised O2PLS in all samples. **(A)** An

overview of the study design. **(B)** Upper plot: principal component analysis (PCA) plot showing the separation between HCM and control samples based on the top differentially acetylated regions using DESeq2; Bottom plot: MA (ratio intensity) plot showing the hyper- and hypoacetylation regions in HCM samples compared to controls using DESeq2. Mean values of normalized counts in all samples are depicted on the x-axis and fold changes (log2) on the y-axis. **(C)** Upper plot: PCA plot showing the separation between HCM and control samples based on the top differentially expressed genes using DESeq2; Bottom plot: MA plot showing the up-regulated and down-regulated genes in HCM samples compared to controls using DESeq2. **(D)** Upper plot: PCA plot showing the separation between HCM and control samples based on the top differentially expressed proteins using DESeq2; Bottom plot: MA plot showing the up-regulated and down-regulated proteins in HCM samples compared to controls using DESeq2. **(E)** Score plot of the first joint component of the H3K27ac ChIP-seq data showing the separation of HCM hearts from controls. **(F)** Score plot of the first joint component of the RNA-seq data discriminating HCM hearts from controls. **(G)** Venn diagram showing the overlapping targets between the top 2,000 genes obtained using the integrative approach (O2PLS) and differentially expressed genes obtained using the pairwise comparison (DESeq2). **(H)** Enrichment analyses using the overlapping targets, which included 618 up-regulated genes and 308 down-regulated genes.

Table 1. Promising candidates showing the same changing direction in HCM versus control hearts at multi-omics levels.

Symbol	Known expression level in cardiomyocytes*	Protein level in the plasma/serum‡	Rank in the O2PLS analysis	Transcriptome change (log2FC)	Proteome change (log2FC)	Histone acetylome change in the matched direction§
AASS	Medium	Detected	1046	-1.279808444	-0.773247332	Yes
ABHD11	Medium	Detected	669	1.137180651	0.640971138	-
ACTN2	High	Detected	1597	0.570095367	0.184641682	-
ADH1B	Low	Detected	1281	-1.081956134	-0.880289941	-
ARHGAP1	Medium	Detected	1512	1.763584	0.597291165	Yes
ASPN	Not detected	Detected	454	2.204317	1.77725136	Yes
ATP2A2	High	Detected	728	-1.466783433	-0.353968239	-
BGN	Medium	Detected	1012	1.832538	1.533815477	-
C6	Not detected	Detected	130	-1.776237721	-0.695516121	-
CA3	Not detected	Detected	43	4.518344	1.686864533	-
CHCHD6	Medium	Not detected	910	2.003514	0.778880421	-
CHDH	Low	Not detected	98	-2.209806815	-1.442995937	Yes
CLGN	Not detected	Not detected	1751	-1.088152172	-0.808838009	Yes
DDAH1	Not detected	Detected	452	1.764206	1.115473264	Yes
EFHD1	Not detected	Not detected	682	1.865517	1.264459654	Yes

FGF12	No data	Not detected	62	-2.353022973	-1.079274106	Yes
FHL2	High	Not detected	440	-1.255001122	-0.411782639	-
FMOD	No data	Detected	395	2.594323	1.414783229	Yes
FSCN1	Not detected	Detected	1276	1.759951	0.971207984	Yes
GATM	Not detected	Detected	277	1.545734889	1.002695568	Yes
GPD1	Medium	Detected	1708	-1.687472289	-1.343895776	Yes
GPD1L	No data	Not detected	1342	-1.066784553	-0.526757351	-
HSPA2	Not detected	Detected	171	3.490234	2.652261545	Yes
HSPB6	High	Detected	647	2.253803	0.392977171	-
LDHA	Not detected	Detected	974	-1.219034333	-0.474108409	Yes
LTBP2	Not detected	Detected	829	2.1279	2.317830122	Yes
LUM	Not detected	Detected	217	2.135187	0.959297891	-
MAP4	Medium	Detected	717	1.851811	1.003525943	Yes
MCAM	Low	Detected	1338	1.653115	1.104716402	Yes
MFAP2	Not detected	Not detected	1303	1.84358	0.70531399	Yes
MYH6	High	Detected	40	-2.919038787	-1.009056575	-
MYL12A	Medium	Detected	64	2.419604	0.578322489	-
MYLK3	High	Not detected	1904	-1.007764366	-0.747812735	-
NES	Medium	Detected	641	1.806419	1.936943877	-
NPPA	Medium	Detected	13	1.549472808	0.747781872	Yes
NUDT4	High	Not detected	1531	-1.093022558	-0.695436878	-
PDK4	High	Not detected	1724	-1.07459529	-0.607585372	Yes
RAB24	No data	Not detected	1085	2.211788	1.041018278	-
RRAS	High	Detected	1642	1.788982	0.447281909	-
S100A6	Not detected	Detected	1486	1.859224	0.794548978	-
SAA1	No data	Detected	63	-1.173021107	-2.740052623	-
SERPINE2	No data	Detected	338	1.002427623	0.652055032	-
SGCG	High	Not detected	1805	0.707480434	0.430185184	Yes
SLC25A5	High	Detected	1634	1.811682	0.613247381	-
SNCA	Not detected	Detected	274	2.629958	1.02410125	-
SORBS2	No data	Detected	1991	1.475585	0.484980733	Yes
STMN1	Not detected	Detected	1211	0.77170211	0.807423435	-
SYNPO2L	High	Not detected	483	2.306538	1.028059491	Yes
TANGO2	Low	Detected	1086	1.981163	0.733156318	-
THBS4	Low	Detected	618	2.323482	2.106545221	-
TPM3	Medium	Detected	981	2.116673	1.224815654	-
TPPP	Not detected	Not detected	1374	-0.898226805	-0.830942802	Yes
UCHL1	Not detected	Detected	44	5.052289	0.933127805	-

*: The known expression level of each candidate in cardiomyocyte is collected from the Human Protein Atlas (https://www.proteinatlas.org/).

‡: The detectable protein level of each candidate in plasma and serum is collected from the Human Plasma Proteome Project Data Central (http://www.peptideatlas.org/hupo/hppp/).

§: 50Kb window of transcription start site.

FC=fold change, TSS=transcription start site, -: not matching.

Combined omics analyses identify ECM, muscle contraction, and metabolism as the main mechanisms altered in HCM

Genes annotated to differentially acetylated regions in the histone acetylome data, differentially expressed genes in the transcriptome data, and differentially expressed proteins in the proteome data consistently pointed to biological functions involved in the enhancement of ECM, enhancement of muscle contraction, and suppression of metabolism in HCM versus control hearts. To further evaluate these biological functions, we collected gene sets that are known to regulate ECM remodeling, muscle contraction, and metabolism, and performed GSEA to study whether they are primarily found in the up-regulated or down-regulated data sets in our study or they are randomly distributed. We observed that genes involved in ECM and cardiac muscle contraction were positively correlated with genes annotated to the hyperacetylated regions in the ChIP-seq data, the up-regulated genes in the RNA-seq data, and genes encoding the up-regulated proteins in the proteomics data (Figure 2A and 2B). Genes that are related to fatty acid metabolism were significantly correlated with genes annotated to the hypoacetylated regions in the ChIP-seq data, the down-regulated genes in the RNA-seq data, and genes encoding the down-regulated proteins in the proteomics data (Figure 2C, Supplementary Figure 3). Combined, we confirmed that pathways involved in ECM, muscle contraction, and metabolism were affected in HCM.

Consistently, protein-coding genes (36 up-regulated and 17 down-regulated ones) were enriched for comparable biological processes and pathways as shown in previous analyses, such as ECM and muscle contraction (Supplementary Table 6C). We also studied protein networks among them and observed that ECM and muscle contraction were the most enriched pathways by the up-regulated protein-coding genes and metabolism was the most enriched pathway by the down-regulated protein-coding genes (Figure 3A and 3B).

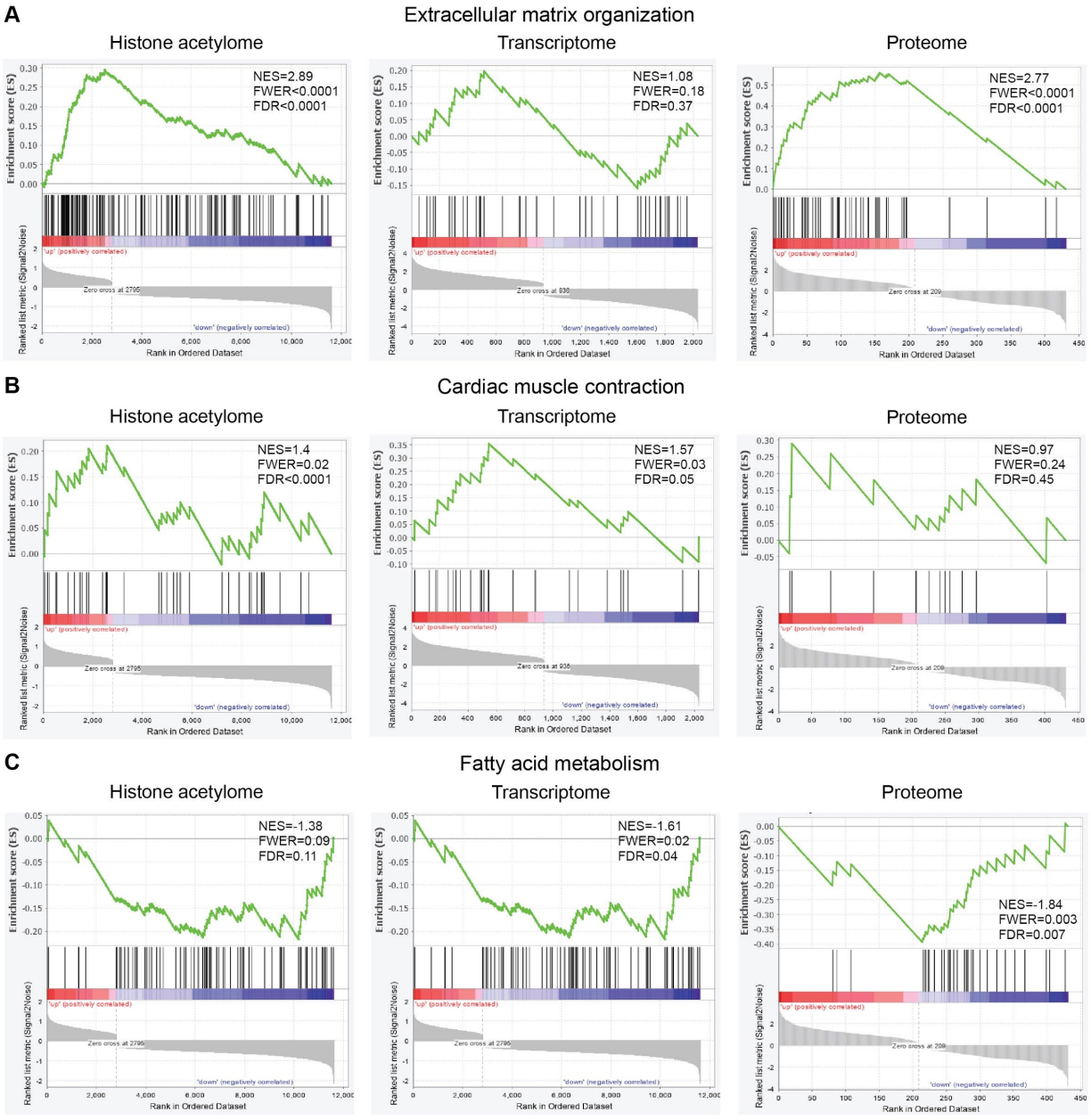


Figure 2. Gene set enrichment analysis showing the correlation of the gene set, which are established in extracellular matrix organization **(A)**, cardiac muscle contraction **(B)**, and fatty acid metabolism **(C)**, in annotated genes from differentially acetylated regions (histone acetylome), differentially expressed genes (transcriptome), and genes encoding differentially expressed proteins (proteome), respectively. Differentially expressed genes were ranked by their fold changes and shown on the x-axis. The running correlation throughout the gene set is shown by the curve (green) and the running enrichment score (ES) is shown in the y-axis. Enrichment score normalized for gene set size (NES), Familywise-error rate p-value (FWER), and the false discovery rate (FDR) are shown per enrichment plot.

Enriched transcription factor binding motifs in differentially acetylated regions

To define the actual factors that regulate altered RNA/protein expression in HCM hearts, we studied the putative transcription factors (TFs) encoded by enriched transcription factor binding motifs (TFBMs) in differentially acetylated regions by scanning through the HOCOMOCO motif database of 769 human primary and alternative binding models. We obtained 125 TFBMs in the hyperacetylated regions and 115 TFs were predicted to bind to them. In the hypoacetylated regions, 120 TFBMs were enriched and they encoded for 111 TFs. Out of those, 68 annotated TFs were enriched in both hyper- and hypo-acetylated regions (Supplementary Figure 4).

Notably, 2 TFs (KLF15 and ETV4) that were encoded by enriched TFBMs in the hyperacetylated regions showed higher mRNA levels in HCM hearts than controls, and 7 TFs (AR, CLOCK, ETS2, GATA5, MEIS1, RXRA, and ZFX) that were encoded by enriched TFBMs in the hypoacetylated regions showed lower mRNA levels in HCM hearts than controls (Figure 3C). These TFs are enriched for biological functions involved in muscle hypertrophy and lipid metabolism (Supplementary Table 7), suggesting their potential roles in regulating the upstream signaling in HCM.

A schematic overview of all analysis steps is shown in Figure 3D.

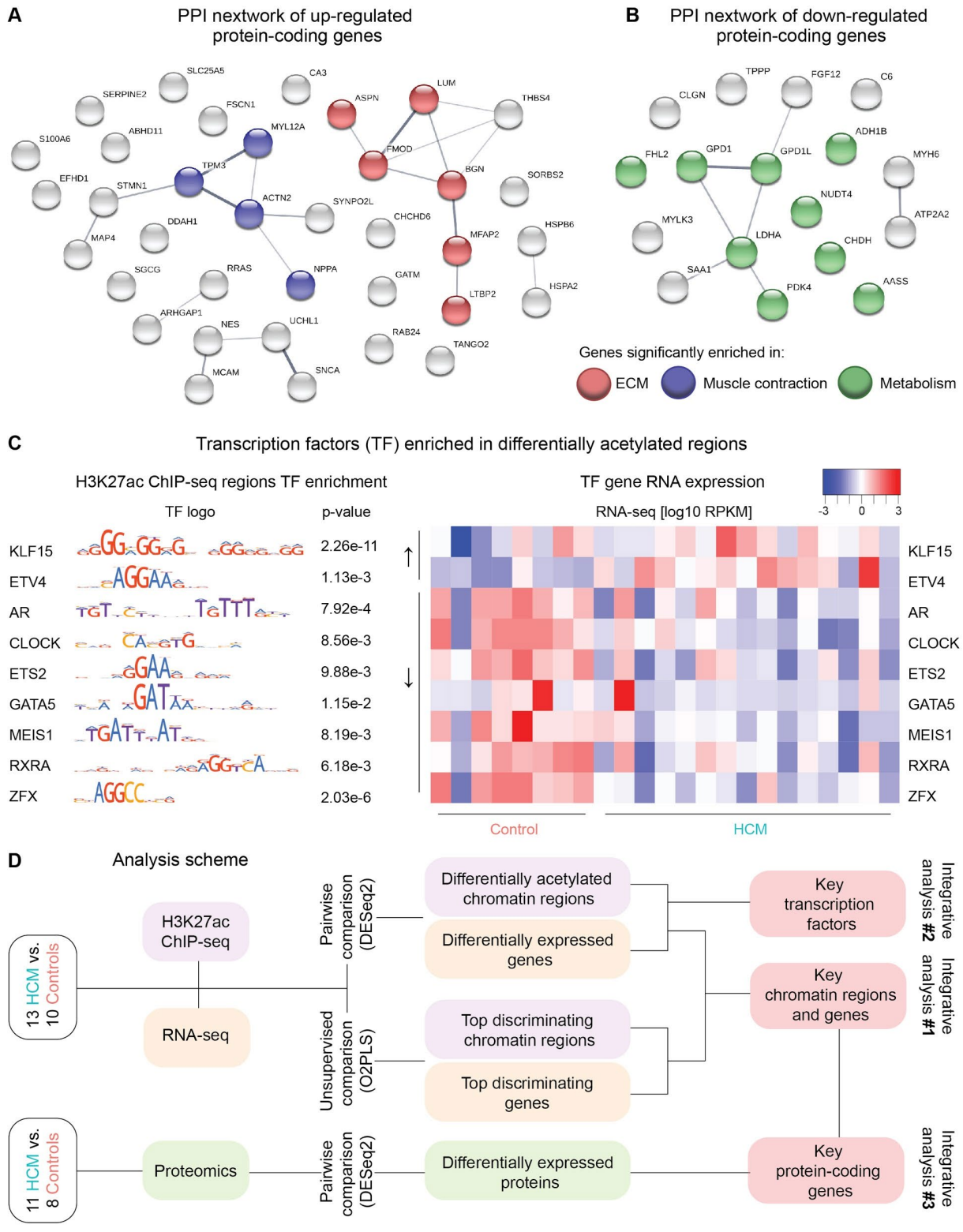


Figure 3. (A) An overview of protein-protein interactions of up-regulated proteins in HCM versus control hearts. Each node represents an individual protein. Nodes highlighted in red are involved in the top enriched pathway in

extracellular matrix organization (ECM). Nodes highlighted in blue are involved in the top enriched pathway in muscle contraction. **(B)** An overview of protein-protein interactions of down-regulated proteins in HCM versus control hearts. Each node represents an individual protein. Nodes highlighted in green are involved in the top enriched pathway in metabolism. **(C)** Transcription factors enriched in differentially acetylated chromatin regions and their corresponding p-values (Fisher test) are shown on the left, and their mRNA expression levels among all samples are shown in the heatmap on the right. RPKM: reads per kilobase million. **(D)** A schematic overview showing all analysis steps in histone acetylome, transcriptome, and proteome data used in this study.

Genes discriminating HCM from controls show consistent changes on various levels

We further selected an example of one down-regulated and one up-regulated gene and examined them in more detail to demonstrate the strength of our integrative omic analysis. ATP2A2, one protein-coding gene from the overlapping candidates, encodes the sarcoplasmic reticulum Ca2+-ATPase pump SERCA2a and plays a critical role in the regulation of calcium handling [10]. We identified ATP2A2 as one of the major candidates in discriminating HCM hearts from controls in our integrated H3K27ac ChIP-seq and RNA-seq analysis (Figure 4A). Besides, its mRNA and protein levels were significantly lower in HCM versus control hearts in the pairwise comparison, and the suppressed protein level was also validated using western blot (Figure 4B).

HSPA2, another protein-coding gene from the overlapping candidates, is involved in protein quality control [11]. We demonstrated that HSPA2 acts as another key player in discriminating HCM hearts from controls (Figure 4A). Its mRNA and protein levels were significantly higher in HCM versus control hearts in the pairwise comparison, and the enhanced protein level was also validated using western blot (Figure 4C). The upstream regulatory region of HSPA2 also showed a higher acetylation level in HCM versus control hearts (Table 1). We further showed a profound HSPA2 expression in HCM compared to the control heart (Figure 4D). Moreover, occasional HSPA2 aggregates were observed in cardiomyocytes from the HCM heart, whereas no HSPA2 aggregates were shown in the control heart (Figure 4E).

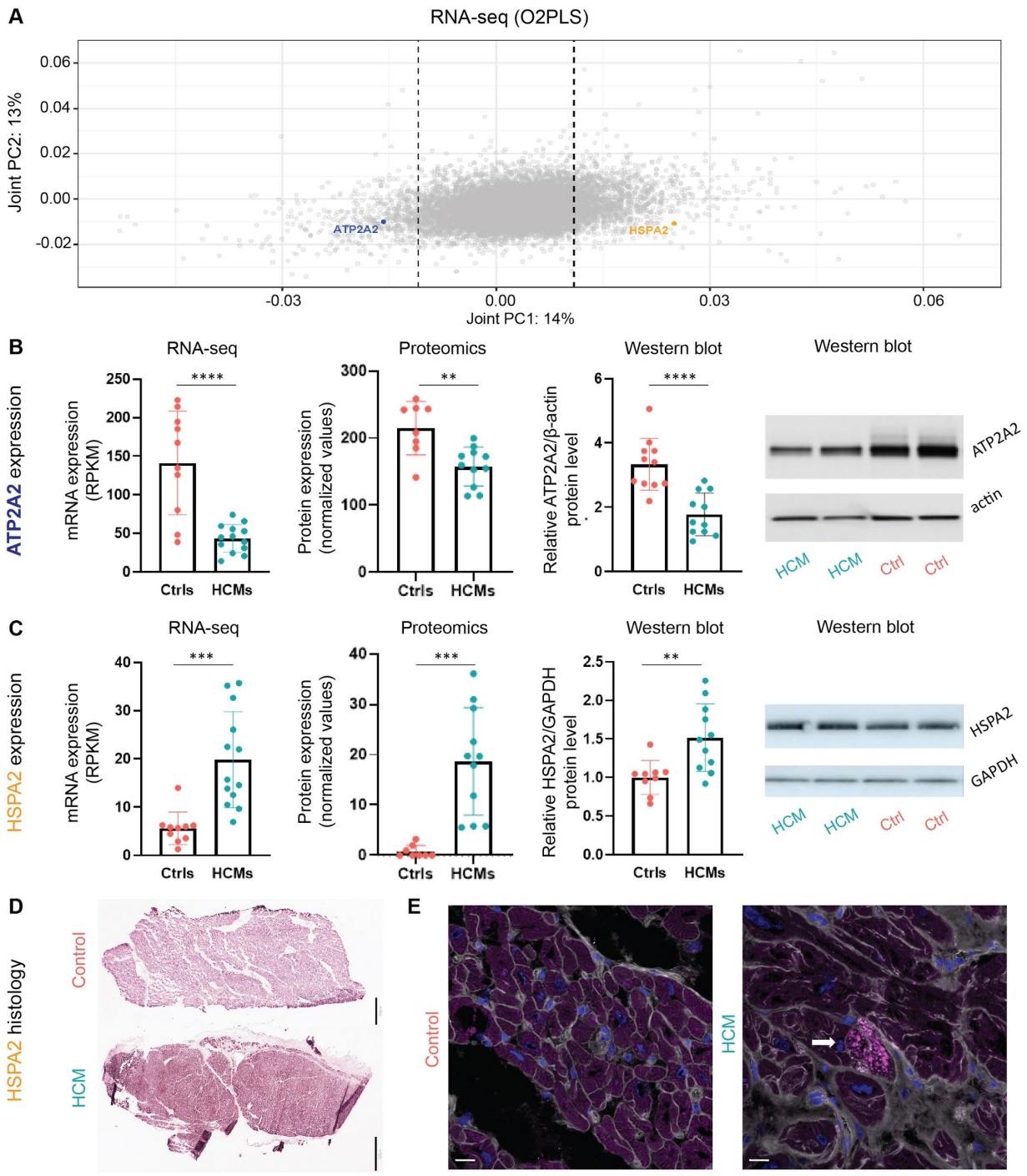


Figure 4. Changes of ATP2A/SERCA2a and HSPA2 at the mRNA and protein levels in HCM versus control hearts. **(A)** A plot of the joint component loadings of RNA-seq data showed ATP2A2 and HSPA2 were two major players in discriminating HCM hearts from controls. The dashed lines on both positive and negative sides indicate the cutoff threshold, with genes with a large contribution to the joint component falling outside of the dash lines. **(B)** ATP2A2 mRNA and protein levels in HCM and control samples at the mRNA and protein levels. **(C)** HSPA2 mRNA and protein

levels expression in HCM and control samples. **: P<0.01, ***: P<0.001, ****: P<0.0001. **(D)** Representative immunohistochemistry staining showing higher HSPA2 staining intensity in HCM heart as compared to control. Scale bar=400µm (control sample) and 800µm (HCM sample). **(E)** Representative immunofluorescence staining showing HSPA2 aggregates in an HCM heart (indicated by the arrow), whereas the control shows diffuse staining of the cytoplasm without aggregates. WGA-AF488 appears in grey to visualise cell membrane and DAPI appears in blue to visualise nuclei. Scale bar=16µm.

Allelic imbalance of MYBPC3 in HCM hearts is observed at both DNA and RNA levels

To further explore the potential of the produced data, we investigated the contribution of both MYBPC3 alleles in the sequencing datasets. In the patient cohort, three heterozygous truncating mutations were present in MYBPC3, namely c.2373dupG (n=5), c.2827C>T (n=6), and c.927-2A>G (n=2). We observed that the average acetylation ratio of MYBPC3 with c.2373dupG, c.2827C>T, and c.927-2A>G mutation to wildtype allele was 50%, 25%, and 66.6% respectively (Supplementary Figure 5A). The average mRNA expression ratio of MYBPC3 with c.2373dupG, c.2827C>T, and c.927-2A>G mutation to wildtype allele was 6.7%, 19.7%, and 43.4% respectively (Supplementary Figure 5B). The acetylation and mRNA levels of three mutations were not observed in control hearts (Supplementary Figure 5C and 5D). It has to be noted that it is not possible to effectively distinguish between wildtype and mutant alleles in the proteomics data.

Cellular fraction sub-analysis of bulk cardiac tissue transcriptome indicates cardiomyocyte enrichment in a cellular-specific response in HCM

Since there are multiple cell types present in the heart samples, we collected cell-type-specific markers for cardiomyocytes and 11 non-myocyte cell types as revealed by recent single-cell studies in the heart [12,13] and examined their expression levels in our bulk sequencing data. Cardiomyocyte-specific markers showed higher histone acetylation and mRNA levels than markers of 11 non-myocyte cell types (Figure 5A and 5B), regardless of health and disease. The mRNA expression of several cell-type-specific markers was significantly different between HCM and control hearts (Figure 5C). The upstream regulatory region of three cardiomyocyte-specific markers and one T cells-specific marker also showed significantly higher acetylation activities in HCM hearts versus controls (Figure 5D).

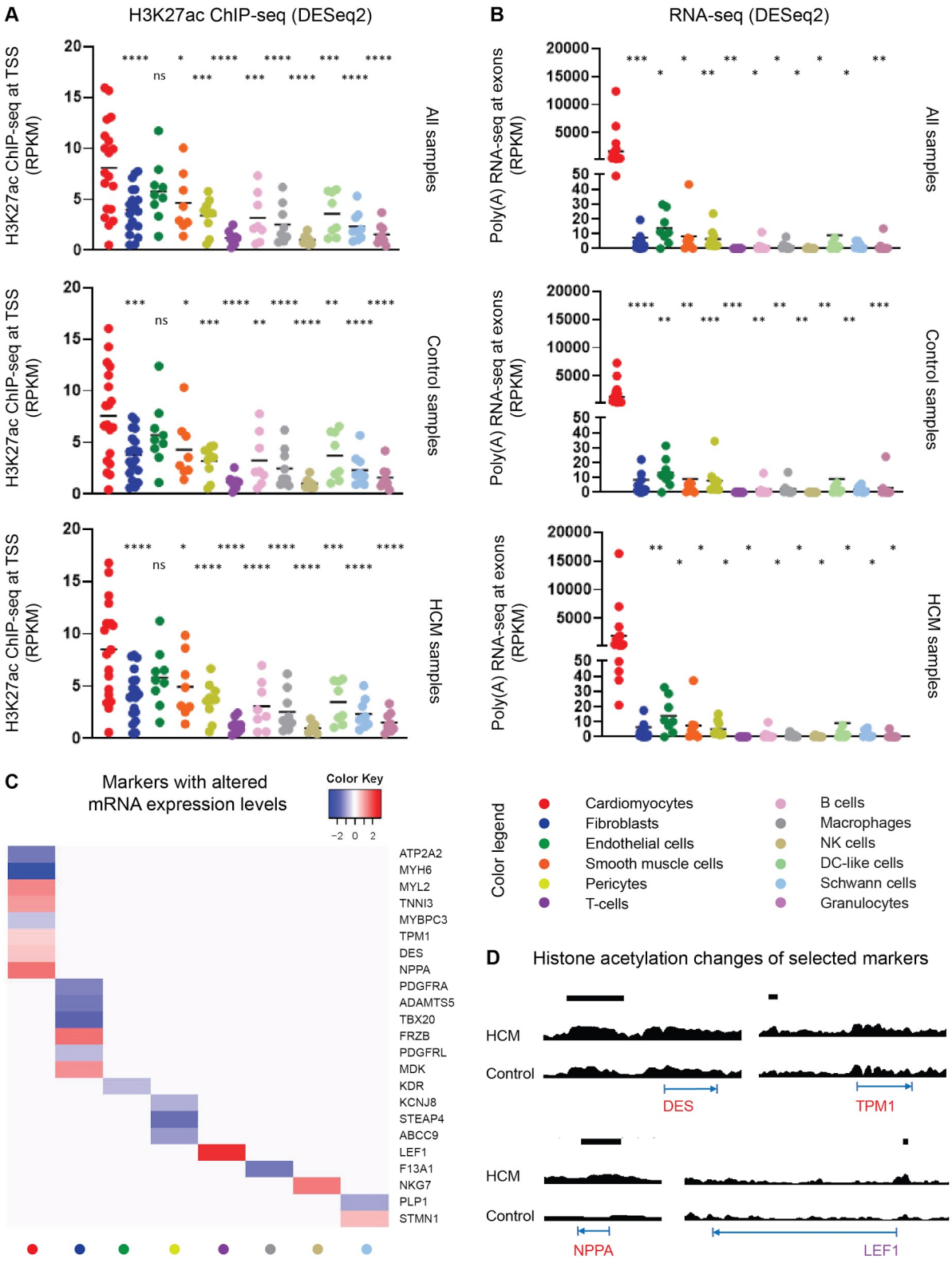


Figure 5. The expression of cell-type-specific markers in 12 cell populations. **(A)** The histone acetylation levels of 11 cell-type-specific markers. Each dot represents the average acetylation value of each marker for all samples, or only

control or HCM samples. One-way ANOVA was used to compare the mean of cardiomyocyte-specific markers with the mean of non-myocyte markers separately: ns (not significant), *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001. **(B)** The mRNA expression levels of 11 cell-type-specific markers. Each dot represents the average expression value of each marker for all samples, or only control or HCM samples. One-way ANOVA was used to compare the mean of cardiomyocyte-specific markers with the mean of non-myocyte markers separately: ns (not significant), *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001. **(C)** Heatmap showing cell-type-specific markers with significantly changed mRNA expression levels in HCM hearts compared to controls. Fold changes of these markers are depicted. Positive fold changes (red) and negative fold changes (blue) represent up-regulation and down-regulation in HCM versus control hearts respectively. **(D)** Histone acetylation levels at the upstream (50 kb) of three cardiomyocytes-specific markers (red) and one T cells-specific marker (violet) were significantly changed between HCM and control hearts. Tracks of one HCM and one control heart were scaled in the UCSC genome browser (ln(x+1): 0-10). Significantly changed upstream regions are indicated by the black bar above. The transcription start and end site of each marker are indicated by the blue arrow below.

DISCUSSION

In the present study, we studied changes in histone acetylome, transcriptome, and proteome between HCM and control hearts. Integrating these multi-omics data, we present for the first time a set of DNA regions and genes that differentiate HCM from control hearts and identified 53 protein-coding genes as the major contributors. Since these comprehensive datasets consistently indicated altered biological functions involved in ECM, muscle contraction, and metabolism, we further studied and identified the upstream TFs that could play a critical regulatory role. We also observed an abundance of acetylation signals and transcripts signals from cardiomyocyte-specific markers compared to markers of 11 non-myocyte populations in the heart.

By integrating ChIP-seq, RNA-seq, and proteomics data we identified a list of deciding elements in discriminating HCM hearts from controls. Among these, ATP2A2/SERCA2a transports cytosolic calcium into the sarcoplasmic reticulum, which subsequently regulates the contraction and relaxation of cardiomyocytes [14]. Suppressed ATP2A2/SERCA2a expression is associated with impaired relaxation in cardiomyocytes and contributes to diastolic dysfunction in HCM [10,15]. Consistent with previous findings, we also showed lower mRNA and protein levels in HCM hearts compared to controls.

Another element HSPA2 (heat shock-related 70 kDa protein 2), which protects cardiac integrity by correcting misfolded proteins upon stress [16], showed higher upstream acetylation level, mRNA level, and protein level in HCM hearts versus controls. Up-regulated HSPA2 protein level has been observed at

both the initial stage of diseased cardiomyocytes and at the end stage of diseased heart with diastolic dysfunction [17,18]. We previously also showed elevated HSPA2 protein level in HCM with mutations in different sarcomeric proteins when compared to controls and it was negatively correlated with MYBPC3 peptide counts [11]. Here we further demonstrated a cardiomyocyte-specific elevation of HSPA2 accompanied by the occasional HSPA2 aggregates in the HCM heart, which may contribute to the disease progression by hampering the protein quality control system.

Interestingly, both ATP2A2/Serca2a- and HSPA2-mediated activities are ATP-dependent. Besides the enriched metabolism using the ChIP-seq, RNA-seq, and proteomic data, we also identified several TFs as potential key upstream factors in regulating metabolism in HCM, such as KLF15, AR, and RXRA, which which are potential key upstream factors in regulating metabolism [19–21]. A review from our group summarized the metabolic changes in HCM [22]. Drugs targeting energy metabolism, such as trimetazidine, have been investigated in the clinical trial in HCM patients [23]. Combined, identified TFs could be potential therapeutic targets in restoring metabolic homeostasis in HCM patients.

Previous studies from our group and others showed around 30% ratio of mutant to wildtype MYH7- and MYBPC3-alleles in HCM patients at the mRNA and the protein levels, respectively [24,25]. In this study, we present imbalanced mRNA expression of three truncating mutations in MYBPC3. Notably, we showed for the first time the presence of such an imbalance at the DNA level by evaluating allele-specific histone acetylation signals. The mutant/wildtype expression ratio differs between truncating mutations and between DNA and mRNA levels, suggesting mutation-specific imbalance and additional machinery between DNA and mRNA. It is important to point out that H3K27ac ChIP-seq is sensitive in capturing upstream regulatory regions of a gene and is restored in covering the core gene body [26], the allelic expression based on the histone acetylation level could be limited by the nature of the technique. Nevertheless, this finding highlights another utilization of omics data in detecting allelic imbalance for genetic diseases.

As the cell composition changes in diseased hearts, we investigated cell-type-specific signals from cardiomyocytes and 11 non-myocyte cell types and observed more ChIP-seq signals and transcripts in cardiomyocyte-specific markers than all non-myocyte markers, suggesting a considerable portion of cardiomyocytes-derived signals in the bulk sequencing data. Gene set enrichment analysis showed enriched pathways in ECM, which is in line with previous reports indicating fibrosis as a hallmark of HCM [27]. Apart from fibroblasts, cardiomyocytes also express ECM-related genes, including two fibroblast-specific markers (DKK3 and ABI3BP) [28]. Because cell-type-specific markers might also be expressed by

other cell types and the limitation of capturing pure cell populations from snap-frozen tissues without damaging the DNA and RNA integrity, we can only speculate that the majority of the signals were likely obtained from cardiomyocytes and it remains a challenge to separate them in bulk sequencing. Additionally, we used cell-type-specific markers obtained from the murine models [12,13], which may not translate one-to-one across species. Therefore, future studies are required to validate the cellular-specific response in HCM by isolating single-cell populations from human cardiac tissues.

It is important to note that both H3K27ac ChIP-seq and RNA-seq data were generated from the same 13 HCM and 10 control samples. However, only 11 out of 13 HCM samples were included in the proteomics analysis in comparison with 8 different control samples. Thus, different donor samples and group sizes might obscure the integrative analysis of the multi-omics data. Besides, we used end-stage cardiac tissues, identified candidates can be further examined using early disease models, such as programmed cardiomyocytes, engineered heart tissues, and/or cardiac organoids, to study their involvements in the chronological progression of HCM, which subsequently provide valuable insights into the identification of early disease markers.

In conclusion, our study presents detailed information in HCM hearts with truncating MYBPC3 mutations. These data showed altered ECM, muscle contraction, and metabolism in HCM. Integrative analyses further identified a subset of protein-coding genes and upstream TFs that could drive these pathophysiological mechanisms and serve as promising diagnostic and/or therapeutic targets. We also showed a considerable amount of cardiomyocytes-derived signals compared to non-myocyte cell types, providing cardiomyocyte-specific insights to better understand HCM in future studies.

MATERIALS AND METHODS

Human cardiac samples

The study protocol was approved by the local medical ethics review committees, including the Biobank Research Ethics Committee of University Medical Center Utrecht (protocol number 12/387), the local ethics committee of the Erasmus MC (2010-409), the Washington University School of Medicine Ethics Committee (Institutional Review Board), and the Sydney Heart Bank (HREC Univ. Sydney 2012/030). Cardiac tissue of the interventricular septum of 13 HCM patients undergoing septal myectomy to relieve left ventricular outflow tract obstruction was collected in the Erasmus MC. Genetic analyses of all patients revealed three truncating pathogenic heterozygous mutations in MYBPC3, namely c.2373dupG in 5 HCM

patients, c.2827C>T in 6 patients, and c927-2A>G in 2 HCM patients. Control tissues of 18 non-failing donors were obtained from the Biobank of University Medical Center Utrecht, the Washington University School of Medicine, and the Sydney Heart Bank. Informed consent was obtained from each patient prior to surgery or was waived by the ethics committee when acquiring informed consent was not possible due to the death of the donor. Samples were collected and snap-frozen in liquid nitrogen and stored at -80˚C up until analysis. Detailed clinical characteristics are shown in Supplementary Table 1 and an overview of samples included in the following experiments is shown in Figure 1A.

H3K27ac chromatin immunoprecipitation and sequencing (ChIP-seq)

We performed chromatin immunoprecipitation and sequencing (ChIP-seq) using the H3K27ac mark to study the differences of the histone acetylome between patient and control samples as previously described [29]. To study the differences of the histone acetylome between patient and control samples, we performed chromatin immunoprecipitation and sequencing (ChIP-seq) using the H3K27ac mark. Briefly, all cardiac samples were sectioned at a thickness of 10µm, and chromatin was isolated using the MAGnify™ Chromatin Immunoprecipitation System kit (Life Technologies) according to the manufacturer's instructions. The anti-histone H3K27ac antibody (ab4729, Abcam) was used for immunoprecipitation. Captured DNA was purified using the ChIP DNA Clean & Concentrator kit (Zymo Research). Libraries were prepared using the NEXTflex™ Rapid DNA Sequencing Kit (Bioo Scientific). Samples were PCR amplified, checked for the proper size range and absence of adaptor dimers on a 2% agarose gel, and barcoded libraries were sequenced 75 bp single-end on an Illumina NextSeq500 sequencer. Sequencing reads were mapped against the reference genome (hg19 assembly, NCBI37) using the BWA package (mem –t 7 –c 100 –M –R)1. Multiple reads mapping to the same location and strand were collapsed to single read and only uniquely placed reads were used for peak-calling. Peaks/regions were called using Cisgenome 2.02 (–e 150 -maxgap 200 –minlen 200). Region coordinates from all samples were stretched to at least 2000 base pairs and collapsed into a single common list. Overlapping regions were merged based on their coordinates. Only regions supported by at least 2 independent datasets were further analyzed. Autosomal sequencing reads from each ChIP-seq library were overlapped back with the common region list to set the H3K27ac occupancy for every region-sample pair. Differentially acetylated regions between HCM and control hearts were identified using DESeq2 under the default setting in the Galaxy environment [30].

Annotating genes in the vicinity of differentially acetylated regions

Region-to-gene annotation was performed to study potentially affected genes in the vicinity of DNA regions with altered acetylation levels in HCM hearts when compared with controls. Differentially acetylated regions located within either +/-5 kb or +/-50 kb window from the transcription start site (TSS) of all genes were obtained, and the nearest genes of these regions were collected.

Predicting transcription factor binding motifs in differentially acetylated regions

To study the putative upstream signaling, we studied the enriched transcription factor binding motifs (TFBMs) by differentially acetylated regions and motifs-encoded transcription factors (TFs). DNAse I hypersensitivity regions in human cardiac samples, which play a key role in transcription factor footprinting [31], were collected from the ENCODE project and overlapped with differentially acetylated regions in this study [32]. Overlapping DNA sequences between differentially acetylated regions and DNAse I hypersensitivity regions were used to studying the enriched transcription factor binding motifs and motifs-encoded TFs using MEME Suite AME tool (HOCOMOCO Human v11 Full, average odds scoring method, and Fisher's exact test) [33].

RNA sequencing

We also performed RNA sequencing (RNA-seq) and obtained the transcriptome landscapes in all samples. Briefly, RNA was isolated using the RNeasy Micro Kit (Qiagen) or ISOLATE II RNA Mini Kit (Bioline) according to the manufacturer's instructions. Sample quality was assessed using the 2100 Bioanalyzer with an RNA 6000 Pico Kit (Agilent), and sample quantity was measured using Qubit Fluorometer with an HS RNA Assay (Thermo Fisher). Afterward, libraries were prepared using the NEXTflexTM Rapid RNA-seq Kit (Bioo Scientific) and sequenced by the Nextseq500 platform (Illumina). Sequenced reads were aligned to the human reference genome GRCh37 using STAR v2.4.2a [34]. Reads per kilobase million reads sequenced (RPKMs) were calculated with edgeR's RPKM function [35]. To identify a list of differentially expressed genes between HCM and control hearts at Padj.<0.05, we employed d DESeq2 to process all the raw counts per sample per group in the Galaxy environment [30].

Allele-specific expression of MYBPC3 in ChIP-seq and RNA-seq data

RNA-seq and ChIP-seq reads were processed with TrimGalore (version 0.6.5) to detect and clip off adaptor sequences, reads with a remaining length of 20 or more nucleotides, and an average read quality q>20 were selected. The reads were aligned to the human genome (GRCh37) with transcript annotation from Ensembl (version 74) using the STAR aligner (version 2.7.1a). To facilitate the equal alignment of reads from wild-type and mutant alleles, alignments were made without clipping off end sequences (STAR option alignEndsType set to EndToEnd), and only best scoring alignments were selected (STAR option outSAMprimaryFlag set to AllBestScore). Alignments were selected to remove duplicated reads by an in-house HTSeq based python (version 2.7.10) script [36]. Reads that aligned to the same genomic interval and that aligned with identical bases to the two non-indel MYBPC3 SNPs (rs397516082 and rs387907267) of interest, were considered duplicates. ChIP-seq reads matching to one of the three SNPs of interest were assigned to the wildtype or mutant allele based on the following rules: for the single base polymorphic SNPs (rs397516082 and rs387907267), based on having the wildtype or mutant nucleotide aligned to the SNP position; for the indel SNP (rs397515963) based on having the exact wildtype or mutant sequence of the SNP with 10 surrounding bases in a read aligned to the SNP. In addition, ChIP-seq reads with wildtype or mutant assignment were required to have >90% of the bases aligned to the genome. For RNA-seq reads, the assignment to be wildtype or mutant allele derived was complicated by changed splice patterns for two of the three SNPs. For rs387907267 (c.2827C>T) that has no changed splice patterns, reads were classified based on having the wildtype or mutant nucleotide aligned to the SNP position. For rs397515963 (c.2373dupG) that disrupts the correct splicing of intron 23, reads covering the splice-junction with the correct intron spliced out were counted as wildtype reads, whereas reads using either the correct intron donor or acceptor site, but not both, were counted as mutant reads. For rs397516082 (c.927-2A>G) that disrupts the correct splicing of intron 11, reads covering the splice-junction with the correct intron spliced out were counted as wildtype reads, whereas reads that were running through the intron 11 - exon 12 splice site and into the SNP position were counted as wildtype or mutant based on having the wildtype or mutant nucleotide aligned to the SNP position.

Integrating ChIP-seq and RNA-seq data with Two-way Orthogonal Partial Least Squares

To find common parts between RNA-seq and ChIP-seq data simultaneously across all genes and regions, a data integration approach is considered using Two-way Orthogonal Partial Least Squares (O2PLS) [37]. O2PLS decomposes both RNA-seq and ChIP-seq datasets into joint, omic-specific, and residual parts. The joint subspaces contain variations that are correlated to one another. The Joint Principal Components

(JPCs) that span the joint subspaces are obtained by finding linear combinations of genes and regions that maximize the covariation. Omic-specific subspaces capture variation unrelated to another omics dataset, enabling JPCs to better estimate the underlying system. Here, the rows of the ChIP-seq and RNA-seq data should represent the same samples. Note that O2PLS uses all genes and regions in the datasets, and does not rely on prior information about the position or function of these features. Furthermore, O2PLS is unsupervised, and its algorithm is implemented in the OmicsPLS R package and freely available from CRAN [38]. Prior to the analysis, genes with expression lower than 10 counts in at least 22 samples were removed. Samples in both datasets were matched and 23 overlapping samples were retained. The expression data is normalized. Both datasets were log-transformed, and quantile normalized across samples. The dimensionality of the preprocessed datasets is 23 by 15,882 (RNA-seq) and 23 by 33,642 (ChIP-seq).

Proteomics

We performed proteomics using cardiac samples from the same patients (n=11) and 8 non-failing donor samples from the Sydney heart bank (Supplementary Table 1). Briefly, proteins were loaded to a 4-12% NuPAGE Novex Bis-Tris 1.5 mm mini gel (Invitrogen) for separation, followed by fixation (50% ethanol and 3% phosphoric acid) and staining using 0.1% Coomassie brilliant blue G-250 solution. In-gel digestion was performed and the samples were concentrated in a vacuum centrifuge as described previously [39]. Nano-LC-MS/MS was performed as described previously [40]. Briefly, separated peptides were separated using an Ultimate 3000 Nano LC-MS/MS system (Dionex LC-Packings, Amsterdam, The Netherlands) and trapped. Eluting peptides were ionized into a Q Exactive mass spectrometer (Thermo Fisher, Bremen, Germany), and intact masses were measured in the orbitrap. Among all, the top 10 peptide signals were selected and analyzed using the MS/MS. MS/MS spectra were searched against a Uniprot human reference proteome FASTA file (Swissprot_2017_03_human_canonical_and_isoform.fasta, 42161 entries) using MaxQuant version 1.5.4.1. Differentially expressed proteins between HCM and control hearts at $P<0.05$ were identified using DESeq2 [30].

Functional enrichment analysis

Gene Ontology (GO) enrichment analysis: To study the enriched biological functions, GO enrichment analysis was performed using the ToppFun ToppGene Suite under the default settings (FDR correction, p-value cutoff of 0.05, and gene limits between 1 and 2,000) [41].

Gene set enrichment analysis (GSEA): Established gene sets involved in the most enriched biological functions in HCM versus control hearts were collected from Molecular Signature Database v7.1 (Supplementary Table 8) [42,43]. Each gene set per biological function was studied for its positive or negative correlation with genes annotated from the altered acetylated levels in the ChIP-seq data, genes with altered mRNA levels in the RNA-seq data, and protein-coding genes with altered protein levels in the proteomic data in this study.

Protein-protein interaction (PPI) networks: Protein networks were performed using the STRING Version 11.0 under the following settings: the meaning of network edges: confident, minimum required interaction score: medium confidence (0.400) [44].

Western blot

Western blot was performed as described previously [11]. Primary antibodies, including mouse anti-HSPA2 (heat shock-related 70 kDa protein 2, 66291-1, Proteintech Group), mouse anti-GAPDH (10R-G109a, Fitzgerald Industries International), anti-ATP2A2 (sarcoplasmic reticulum Ca2+-ATPase 2, also known as SERCA2), and anti-β-actin, were used.

Immunohistochemistry and immunofluorescence (IF) staining

Tissue sections were thawed and left at RT for 20 min inside a closed box. Then the sections were treated with peroxidase blocking solution (3% H2O2 in MeOH) and blocked with 1% BSA. HSPA2 (Anti-HSPA2 rabbit-Polyclonal antibody Prestige Antibodies HPA000798) primary antibodies were incubated for 1 hour at RT (1:100 for IF and 1:200 for IHC). Sections were washed and incubated with the secondary antibodies for 30 min. Vector Vectastain Universal Elite ABC Kit (PK-6200) was used to enhance the brightfield staining (IHC) and WGA (for cell membranes) and DAPI (nuclei) were added to counterstain fluorescence staining. The brightfield slides were washed, treated with DAB, counterstained with Mayers Hematoxylin, dehydrated, and mounted with DPX. Fluorescence staining was mounted with Mowiol. Images were acquired with the Vectra Polaris Scanner (brightfield/IHC) or Confocal Nikon A1 (IF). The analysis was performed with QuPath, Fiji, and NIS Nikon software.

Availability of data and materials

All relevant data are available within the article and the supplementary files. Because of the sensitive nature of the data collected for this study (13 patient samples and 10 control samples), requests to access the raw sequencing dataset from qualified researchers trained in human subject confidentiality protocols may be sent to the corresponding authors. It is important to note that processed RNA-seq and H3K27ac ChIP-seq data of 11 patient samples and 7 control samples are published in a GWAS study [7] and available in Supplementary Table 3 and Supplementary Table 5, respectively. Raw proteomics data within the article can be found at the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD012467 [45]. Currently, the paper is accepted by Circulation: Heart Failure and will become publicly available soon.

ACKNOWLEDGMENT

We would like to thank Pedro Espinosa for performing immunohistochemistry and immunofluorescent stainings. Grateful thanks to Cris dos Remedios and the Sydney Heart Bank for providing non-failing donor tissue.

Funding

This work was supported by the Netherlands Foundation for Cardiovascular Excellence (to C.C.), the NWO VENI grant (no. 016.176.136 to M.H.), two NWO VIDI grants (no. 91714302 to C.C. and no. 016096359 to M.C.V), ZonMW-NWO VICI grant 91818902 (to J.V.), the Erasmus MC fellowship grant (to C.C.), the RM fellowship grant of the UMC Utrecht (to C.C.), Wilhelmina Children's Hospital research funding (no. OZF/14 to M.H.), the Netherlands Cardiovascular Research Initiative: An initiative with the support of the Dutch Heart Foundation (CVON2014-40 DOSIS to J.V., M.H., F.W.A., CVON2014-11 RECONNECT to C.C., M.C.V., Dekker 2015T041 to A.F.B., and Queen of Heart to C.C. and M.C.V.), UCL Hospitals NIHR Biomedical Research Centre (to F.W.A.).

Supplementary Files – Online

Supplementary Table 1: An overview of the clinical characteristics in all cardiac samples.

Supplementary Table 2: (A) Differentially acetylated regions between HCM and control hearts from the H3K27ac ChIP-seq data. **(B)** Genes annotated to differentially acetylated regions using the 5 kb window of the transcription start site. **(C)** Genes annotated to differentially acetylated regions using the 50 kb window of the transcription start site. **(D)** Enriched GO terms and pathways by genes annotated to the hyperacetylated regions (50 kb window). **(E)** Enriched GO terms and pathways by genes annotated to the hypoacetylated regions (50 kb window).

Supplementary Table 3: (A) Differentially expressed genes between HCM and control hearts from the RNA-seq data. **(B)** Enriched GO terms and pathways by up-regulated genes in HCM versus control hearts. **(C)** Enriched GO terms and pathways by down-regulated genes in HCM versus control hearts.

Supplementary Table 4: (A) Differentially expressed proteins between HCM and control hearts from the proteomics data. **(B)** Enriched GO terms and pathways by up-regulated proteins in HCM versus control hearts. **(C)** Enriched GO terms and pathways by down-regulated proteins in HCM versus control hearts.

Supplementary Table 5: (A) Unsupervised analysis (O2PLS) ranking DNA regions that contributed to the separation between HCM and control hearts. **(B)** Unsupervised analysis (O2PLS) ranking genes that contributed to the separation between HCM and control hearts.

Supplementary Table 6: (A) The overlapping genes between differentially expressed genes using DESeq2 and the top 2,000 genes that discriminate HCM from control hearts using O2PLS. **(B)** Enriched GO terms and pathways by the overlapping genes. **(C)** Enriched GO terms and pathways by a subset of the overlapping genes, which showed the same changing direction at their protein levels in HCM hearts when compared with controls.

Supplementary Table 7: Enrichment results of identified transcription factors that were supported by both ChIP-seq and RNA-seq data in HCM versus control hearts.

Supplementary Table 8: Gene sets collected from the Molecular Signature Databases.

Supplementary Figure 1: Examples of differentially acetylated regions between HCM and control hearts.

Supplementary Figure 2: (A) Protein-protein interactions between up-regulated proteins in HCM versus control hearts. **(B)** Protein-protein interactions between down-regulated proteins in HCM versus control hearts.

Supplementary Figure 3: (A) Gene set enrichment report showing the correlation between extracellular matrix (ECM) related genes and genes annotated to differentially acetylated regions. **(B)** Gene set enrichment report showing the correlation between ECM-related genes and differentially expressed genes. **(C)** Gene set enrichment report showing the correlation between ECM-related genes and genes encoding differentially expressed proteins. **(D)** Gene set enrichment report showing the correlation between cardiac muscle contraction related genes and genes annotated to differentially acetylated regions. **(E)** Gene set enrichment report showing the correlation between cardiac muscle contraction related genes and differentially expressed genes. **(F)** Gene set enrichment report showing the correlation between cardiac muscle contraction related genes and genes encoding differentially expressed proteins. **(G)** Gene set enrichment report showing the correlation between fatty acid metabolism-related genes and genes annotated to differentially acetylated regions. **(H)** Gene set enrichment report showing the correlation between fatty acid metabolism-related genes and differentially expressed genes. **(I)** Gene set enrichment report showing the correlation between fatty acid metabolism-related genes and genes encoding differentially expressed proteins.

Supplementary Figure 4: (A) Enriched transcription factor binding motifs in the hyperacetylated regions in HCM versus control hearts. **(B)** Enriched transcription factor binding motifs in the hypoacetylated regions in HCM versus control hearts.

Supplementary Figure 5: The acetylation and mRNA levels of the wildtype and mutant MYBPC3 alleles in all samples.

REFERENCES

1. Elliott PM, Anastasakis A, Borger MA, Borggrefe M, Cecchi F, Charron P, et al. 2014 ESC Guidelines on diagnosis and management of hypertrophic cardiomyopathyThe Task Force for the Diagnosis and Management of Hypertrophic Cardiomyopathy of the European Society of Cardiology (ESC). Eur Heart J. Oxford Academic; 2014;35:2733–79.

2. Llam N, Bollen IAE, Niessen HWM, Dos Remedios CG, Michels M, Poggesi C, et al. Sex-specific cardiac remodeling in early and advanced stages of hypertrophic cardiomyopathy. PLoS One. 2020;15:e0232427–e0232427.

3. Sabater-Molina M, Pérez-Sánchez I, Hernández Del Rincón JP, Gimeno JR. Genetics of hypertrophic cardiomyopathy: A review of current state. Clin Genet. 2018;93:3–14.

4. Wijnker PJM, van der Velden J. Mutation-specific pathology and treatment of hypertrophic cardiomyopathy in patients, mouse models and human engineered heart tissue. Biochim Biophys Acta Mol Basis Dis. 2020;165774.

5. Liu C-F, Tang WHW. Epigenetics in Cardiac Hypertrophy and Heart Failure. JACC Basic Transl Sci. 2019;4:976–93.

6. Zhang W, Song M, Qu J, Liu G-H. Epigenetic Modifications in Cardiovascular Aging and Diseases. Circ Res. 2018;123:773–86.

7. Hemerich D, Pei J, Harakalova M, van Setten J, Boymans S, Boukens BJ, et al. Integrative Functional Annotation of 52 Genetic Loci Influencing Myocardial Mass Identifies Candidate Regulatory Variants and Target Genes. Circ Genom Precis Med. 2019;12:e002328.

8. Manduchi E, Hemerich D, Setten J, Tragante V, Harakalova M, Pei J, et al. A comparison of two workflows for regulome and transcriptome‐based prioritization of genetic variants associated with myocardial mass [Internet]. Genetic Epidemiology. 2019. Available from: http://dx.doi.org/10.1002/gepi.22215

9. Low TY, Mohtar MA, Ang MY, Jamal R. Connecting Proteomics to Next-Generation Sequencing: Proteogenomics and Its Current Applications in Biology. Proteomics. 2019;19:e1800235.

10. Kresin N, Stücker S, Krämer E, Flenner F, Mearini G, Münch J, et al. Analysis of Contractile Function of Permeabilized Human Hypertrophic Cardiomyopathy Multicellular Heart Tissue. Front Physiol. 2019;10:239.

11. Dorsch LM, Schuldt M, dos Remedios CG, Schinkel AFL, de Jong PL, Michels M, et al. Protein Quality Control Activation and Microtubule Remodeling in Hypertrophic Cardiomyopathy. Cells [Internet]. 2019;8. Available from: http://dx.doi.org/10.3390/cells8070741

12. Skelly DA, Squiers GT, McLellan MA, Bolisetty MT, Robson P, Rosenthal NA, et al. Single-Cell Transcriptional Profiling Reveals Cellular Diversity and Intercommunication in the Mouse Heart. Cell Rep. 2018;22:600–10.

13. Gladka MM, Molenaar B, de Ruiter H, van der Elst S, Tsui H, Versteeg D, et al. Single-Cell Sequencing of the Healthy and Diseased Heart Reveals Cytoskeleton-Associated Protein 4 as a New Modulator of Fibroblasts Activation. Circulation. 2018;138:166–80.

14. New perspectives on the role of SERCA2's Ca2+ affinity in cardiac function. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research. Elsevier; 2006;1763:1216–28.

15. Somura F, Izawa H, Iwase M, Takeichi Y, Ishiki R, Nishizawa T, et al. Reduced Myocardial Sarcoplasmic Reticulum Ca(2+)-ATPase mRNA Expression and Biphasic Force-Frequency Relations in Patients With Hypertrophic Cardiomyopathy. Circulation [Internet]. Circulation; 2001 [cited 2020 Jul 10];104. Available from: https://pubmed.ncbi.nlm.nih.gov/11489771/

16. Ranek MJ, Stachowski MJ, Kirk JA, Willis MS. The role of heat shock proteins and co-chaperones in heart failure. Philos Trans R Soc Lond B Biol Sci [Internet]. The Royal Society; 2018 [cited 2020 Apr 9];373. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5717530/

17. Birket MJ, Raibaud S, Lettieri M, Adamson AD, Letang V, Cervello P, et al. A Human Stem Cell Model of Fabry Disease Implicates LIMP-2 Accumulation in Cardiomyocyte Pathology [Internet]. Stem Cell Reports. 2019. p. 380–93. Available from: http://dx.doi.org/10.1016/j.stemcr.2019.07.004

18. Li W, Rong R, Zhao S, Zhu X, Zhang K, Xiong X, et al. Proteomic Analysis of Metabolic, Cytoskeletal and Stress Response Proteins in Human Heart Failure. J Cell Mol Med [Internet]. J Cell Mol Med; 2012 [cited 2020 Jul 10];16. Available from: https://pubmed.ncbi.nlm.nih.gov/21545686/

19. Gonthier K, Poluri RTK, Audet-Walsh É. Functional Genomic Studies Reveal the Androgen Receptor as a Master Regulator of Cellular Energy Metabolism in Prostate Cancer. J Steroid Biochem Mol Biol [Internet]. J Steroid Biochem Mol Biol; 2019 [cited 2020 Jul 5];191. Available from: https://pubmed.ncbi.nlm.nih.gov/31051242/

20. Subbarayan V, Mark M, Messadeq N, Rustin P, Chambon P, Kastner P. RXRα overexpression in cardiomyocytes causes dilated cardiomyopathy but fails to rescue myocardial hypoplasia in RXRα-null fetuses [Internet]. Journal of Clinical Investigation. 2000. p. 387–94. Available from: http://dx.doi.org/10.1172/jci8150

21. Fisch S, Gray S, Heymans S, Haldar SM, Wang B, Pfister O, et al. Kruppel-like factor 15 is a regulator of cardiomyocyte hypertrophy. Proc Natl Acad Sci U S A. National Academy of Sciences; 2007;104:7074–9.

22. van der Velden J, Tocchetti CG, Varricchi G, Bianco A, Sequeira V, Hilfiker-Kleiner D, et al. Metabolic changes in hypertrophic cardiomyopathies: scientific update from the Working Group of Myocardial Function of the European Society of Cardiology. Cardiovasc Res. Oxford University Press; 2018;114:1273.

23. van Driel BO, van Rossum AC, Michels M, Huurman R, van der Velden J. Extra energy for hearts with a genetic defect: ENERGY trial. Neth Heart J. Springer; 2019;27:200.

24. Montag J, Kowalski K, Makul M, Ernstberger P, Radocaj A, Beck J, et al. Burst-Like Transcription of Mutant and Wildtype -Alleles as Possible Origin of Cell-to-Cell Contractile Imbalance in Hypertrophic Cardiomyopathy. Front Physiol. 2018;9:359.

25. Parbhudayal RY, Garra AR, Götte MJW, Michels M, Pei J, Harakalova M, et al. Variable cardiac myosin binding protein-C expression in the myofilaments due to MYBPC3 mutations in hypertrophic cardiomyopathy. J Mol Cell Cardiol. 2018;123:59–63.

26. Sanchez GJ, Richmond PA, Bunker EN, Karman SS, Azofeifa J, Garnett AT, et al. Genome-wide dose-

dependent inhibition of histone deacetylases studies reveal their roles in enhancer remodeling and suppression of oncogenic super-enhancers. Nucleic Acids Res. Oxford University Press; 2018;46:1756.

27. Freitas P, Ferreira AM, Arteaga-Fernández E, de Oliveira Antunes M, Mesquita J, Abecasis J, et al. The amount of late gadolinium enhancement outperforms current guideline-recommended criteria in the identification of patients with hypertrophic cardiomyopathy at risk of sudden cardiac death. J Cardiovasc Magn Reson. BioMed Central; 2019;21:1–10.

28. Cardiomyocytes facing fibrotic conditions re-express extracellular matrix transcripts. Acta Biomater. Elsevier; 2019;89:180–92.

29. Pei J, Harakalova M, Treibel TA, Thomas Lumbers R, Boukens BJ, Efimov IR, et al. H3K27ac acetylome signatures reveal the epigenomic reorganization in remodeled non-failing human hearts. Clin Epigenetics. BioMed Central; 2020;12:1–18.

30. Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Cech M, et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res. 2018;46:W537–44.

31. Ouyang N, Boyle AP. TRACE: transcription factor footprinting using DNase I hypersensitivity data and DNA sequence [Internet]. bioRxiv. 2019 [cited 2020 Jun 29]. p. 801001. Available from: https://www.biorxiv.org/content/10.1101/801001v1.abstract

32. Rosenbloom KR, Sloan CA, Malladi VS, Dreszer TR, Learned K, Kirkup VM, et al. ENCODE data in the UCSC Genome Browser: year 5 update. Nucleic Acids Res. 2013;41:D56–63.

33. McLeay RC, Bailey TL. Motif Enrichment Analysis: a unified framework and an evaluation on ChIP data [Internet]. BMC Bioinformatics. 2010. Available from: http://dx.doi.org/10.1186/1471-2105-11-165

34. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15–21.

35. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26:139–40.

36. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31:166–9.

37. Bouhaddani S el, el Bouhaddani S, Houwing-Duistermaat J, Salo P, Perola M, Jongbloed G, et al. Evaluation of O2PLS in Omics data integration [Internet]. BMC Bioinformatics. 2016. Available from: http://dx.doi.org/10.1186/s12859-015-0854-z

38. Bouhaddani S el, el Bouhaddani S, Uh H-W, Jongbloed G, Hayward C, Klarić L, et al. Integrating omics datasets with the OmicsPLS package [Internet]. BMC Bioinformatics. 2018. Available from: http://dx.doi.org/10.1186/s12859-018-2371-3

39. Warmoes M, Jaspers JE, Pham TV, Piersma SR, Oudgenoeg G, Massink MP, et al. Proteomics of Mouse BRCA1-deficient Mammary Tumors Identifies DNA Repair Proteins With Potential Diagnostic and Prognostic Value in Human Breast Cancer. Mol Cell Proteomics [Internet]. Mol Cell Proteomics;

2012 [cited 2020 Jul 10];11. Available from: https://pubmed.ncbi.nlm.nih.gov/22366898/

40. Piersma SR, Broxterman HJ, Kapci M, de Haas RR, Hoekman K, Verheul HM, et al. Proteomics of the TRAP-induced Platelet Releasate. J Proteomics [Internet]. J Proteomics; 2009 [cited 2020 Jul 10];72. Available from: https://pubmed.ncbi.nlm.nih.gov/19049909/

41. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. Nucleic Acids Res. 2009;37:W305–11.

42. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. National Academy of Sciences; 2005;102:15545–50.

43. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell systems. NIH Public Access; 2015;1:417.

44. Snel B. STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene [Internet]. Nucleic Acids Research. 2000. p. 3442–4. Available from: http://dx.doi.org/10.1093/nar/28.18.3442

45. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. 2019;47:D442–50.

Chapter 6

Proteomic and functional studies reveal detyrosinated tubulin as treatment target in sarcomere mutation-induced hypertrophic cardiomyopathy

Maike Schuldt[1], Jiayi Pei[2], Magdalena Harakalova[2], Larissa M. Dorsch[1], Saskia Schlossarek[3,4], Michal Mokry[5], Jaco C. Knol[6], Thang V. Pham[6], Tim Schelfhorst[6], Sander R. Piersma[6], Cris dos Remedios[7], Michiel Dalinghaus[8], Michelle Michels[9], Folkert W. Asselbergs[2,10,11], Marie-Jo Moutin[12], Lucie Carrier[3,4], Connie R. Jimenez[6], Jolanda van der Velden[1‡*], Diederik W.D. Kuster[1‡*].

‡shared last author, *corresponding author

[1]Amsterdam UMC, Vrije Universiteit Amsterdam, Department of Physiology, Amsterdam Cardiovascular Sciences, Amsterdam, The Netherlands
[2]Department of Cardiology, Division Heart and Lungs, University Medical Center Utrecht, Utrecht, The Netherlands
[3]Institute of Experimental Pharmacology and Toxicology, Cardiovascular Research Center, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
[4]DZHK (German Centre for Cardiovascular Research), partner site Hamburg/Kiel/Lübeck, Hamburg, Germany
[5]Department of Pediatric Gastroenterology, Wilhelmina Children's Hospital, University Medical Center Utrecht, Utrecht, The Netherlands
[6]Amsterdam UMC, Vrije Universiteit Amsterdam, Department of Medical Oncology, OncoProteomics Laboratory, VUmc-Cancer Center Amsterdam, Amsterdam, The Netherlands
[7]Sydney Heart Bank, Discipline of Anatomy, Bosch Institute, University of Sydney, Sydney, Australia
[8]Department of Pediatric Cardiology, Erasmus Medical Center Rotterdam, Rotterdam, The Netherlands
[9]Department of Cardiology, Thorax Center, Erasmus Medical Center Rotterdam, Rotterdam, The Netherlands
[10]Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College London, London, United Kingdom
[11]Health Data Research UK and Institute of Health Informatics, University College London, London, United Kingdom
[12]Grenoble Institut des Neurosciences (GIN), Université Grenoble Alpes, Grenoble, France

Accepted by Circulation: Heart Failure

Abstract

Background: Hypertrophic cardiomyopathy (HCM) is the most common genetic heart disease. While ~50% of HCM patients carry a sarcomere gene mutation (HCM_{SMP}), the genetic background is unknown in the other half of the patients (sarcomere mutation-negative, HCM_{SMN}). Genotype-specific differences have been reported in cardiac function. Moreover, HCMSMN patients have later disease onset and better prognosis than HCM_{SMP} patients. To define if genotype-specific derailments at the protein level may explain the heterogeneity in disease development, we performed a proteomic analysis in cardiac tissue from a clinically well-phenotyped HCM patient group.

Methods: A proteomics screen was performed in cardiac tissue from 39 HCM_{SMP}, 11 HCM_{SMN} patients and 8 non-failing controls. HCM patients had obstructive cardiomyopathy with left ventricular outflow tract obstruction and diastolic dysfunction. A novel *MYBPC3*$_{2373insG}$ mouse model was used to confirm functional relevance of our proteomic findings.

Results: In all HCM patient samples we found lower levels of metabolic pathway proteins and higher levels of extracellular matrix proteins. Levels of total and detyrosinated α-tubulin were markedly higher in HCM_{SMP} than in HCM_{SMN} and controls. Higher tubulin detyrosination was also found in two unrelated *MYBPC3* mouse models and its inhibition with parthenolide normalized contraction and relaxation time of isolated cardiomyocytes.

Conclusion: Our findings indicate that microtubules and especially its detyrosination contribute to the pathomechanism of HCM_{SMP} patients. This is of clinical importance since it represents a potential treatment target to improve cardiac function in HCM_{SMP} patients, whereas a beneficial effect may be limited in HCM_{SMN} patients.

Introduction

Hypertrophic cardiomyopathy (HCM) is characterized by asymmetric left ventricular (LV) hypertrophy and diastolic dysfunction, which lead to LV outflow tract obstruction (LVOTO) in the majority of cases[1]. Mutations in genes encoding sarcomere proteins cause HCM and are identified in more than half of the patients (sarcomere mutation-positive, HCM_{SMP}). The heterogeneity in genetic background of HCM is large with more than 1500 identified HCM-causing mutations[2]. Approximately 80% of mutations are located in *MYH7* and *MYBPC3*. Less frequent are mutations in *TNNT2* and *TNNI3*[3, 4]. Previous research in HCM mouse models and humans showed genotype-specific differences in cellular characteristics and cardiac remodeling and function. Gene-specific differences in cellular redox and mitochondrial function were reported in mice harboring a *MYH7* or *TNNT2* mutation[5]. In accordance with studies in HCM mouse models, studies on patient myectomy samples reported gene-specific differences in the response to calcium, ADP, protein kinase A, and length-dependent activation of myofilaments compared to non-failing cardiomyocytes[6, 7]. Also, a comparison of two different patient-specific induced pluripotent stem cell-derived cardiomyocyte cell lines, carrying either a mutation in *MYBPC3* or *TPM1*, showed differences in calcium handling and electrophysiological properties[8]. These in vitro studies are strengthened by clinical patient studies which revealed a more severe decline in myocardial efficiency in *MYH7* than in *MYBPC3* mutation carriers[9], accompanied by a different response to therapy[10]. Notably, there is also a large patient population in which a disease-causing mutation cannot be identified, the so-called sarcomere mutation-negative patients (HCM_{SMN}). While the cause of the disease in these patients is unknown, they present with the same clinical phenotype as HCM_{SMP} patients albeit at older age[3]. Moreover, recent data from the SHaRe registry indicate that HCM_{SMP} have a 2-fold greater risk of adverse outcomes than HCM_{SMN}[11]. Whereas LVOTO can be invasively corrected by surgical myectomy, other symptoms can only be managed by pharmacological therapies, which do not halt or reverse cardiac disease[1]. Knowledge about the cellular changes that cause cardiac dysfunction and hypertrophy in HCM patients is needed to design new therapies.

The main goal of this study was to define HCM- and genotype-specific derailments at the protein level, which may explain the heterogeneity in cardiac characteristics and disease initiation and progression. Therefore, we used an unbiased proteomics approach in a large number of myectomy samples from a clinically well-characterized HCM patient group with (HCM_{SMP}) and without (HCM_{SMN}) sarcomere mutations. We show that lower levels of metabolic pathway proteins and higher levels of extracellular matrix (ECM) proteins are the most prominent genotype-independent HCM-specific disease

characteristics. However, abundance and detyrosination of □-tubulin was significantly higher in HCM_{SMP} than in non-failing (NF) controls, with intermediate levels in HCM_{SMN}. Recent studies in human heart failure identified a central role for detyrosinated microtubules in regulating cardiomyocyte function and demonstrated the functional benefit upon reversal of this modification[12, 13]. Our study in a European HCM patient cohort and genetic HCM mouse models strengthens the concept that targeting the microtubule network represents a therapeutic strategy to correct impaired function, and extends it to HCM caused by sarcomere mutations.

Methods

Human cardiac samples

Tissue of the IVS of HCM patients was obtained during myectomy surgery to relieve LVOTO or after heart transplantation (1 sample, HCM 164). Samples of IVS from 8 healthy NF donors with no history of cardiac abnormalities was obtained from the Sydney Heart Bank (HREC Univ Sydney 2012/030) and served as controls. The clinical parameters of the HCM and NF groups are summarized in *Table S1*. In this table we organized HCM patients based on their genotype into 5 sub-groups: Patients with the Dutch *MYBPC3* founder mutation (2373insG), where the truncating mutation resulted in MYBPC3 haploinsufficiency[14]; Patients with *MYBPC3* mutations other than the 2373insG mutation of which 81.8% were truncating mutations as well; Patients with *MYH7* mutations; Patients with mutations in less frequently affected sarcomere genes (*TNNT2*, *TNNI3* and *MYL2*); and HCM_{SMN} patients. In line with studies in other cohorts[15], almost all mutations in *MYBPC3* were truncating mutations, whereas mutations in other sarcomeric proteins were missense except the truncating mutation c.814C>T in *TNNT2*.

Extended methods section in the supplementary materials.

Results

No major genotype-specific protein changes

We performed an unbiased proteomics approach to compare the protein expression profile of 50 interventricular septum (IVS) samples from HCM patients with 8 NF_{IVS} samples (*Table S1*; patient characteristics). We identified 3,811 proteins of which we included 2,127 into our analysis after applying our inclusion criteria of an average normalized count of >1.4 and the protein being detected in >25% of

samples. This filtered out low-level proteins that could not be reliably quantified. Samples were grouped based on their genotype and protein expression changes were analyzed in different group-wise comparisons. *Table S2* lists all group-wise comparisons and the corresponding number of significantly deregulated proteins that contributed to the subsequent cluster and gene ontology (GO) analysis. An unbiased principal component analysis (PCA) of the protein expression data revealed separate clustering of the NF_{IVS} and HCM samples, indicating different protein expression profiles between controls and patients (*Figure 1A*). Supervised hierarchical clustering of a multi-group comparison revealed separate clusters for NF_{IVS} and HCM samples (*Figure 1B*). However, in both analysis the HCM samples did not cluster according to their genotypes, indicating that the protein profile at the time of myectomy surgery is relatively homogeneous with differences between genotypes that are not sufficiently large to distinguish them with cluster analysis. Sample HCM 83 did not cluster with any of the other HCM or NF_{IVS} samples and turned out to have a very high serum albumin content. This implicated contamination with blood and therefore we excluded this sample from all further analyses. Sample HCM 164 also showed a unique protein expression pattern. It may reflect the infant proteome due to the young age (2 months), or the very severe disease stage since this is the only sample obtained from a heart transplantation. Since the variation in this sample is due to a biological reason, we did not exclude it.

To validate our experimental approach, we checked the expression of proteins involved in pathways that are known to be altered in HCM (*Figure S1, Table S3*). Fibrosis, characterized by an increase in ECM components, is a well-established feature of HCM as shown by data from myectomy biopsies and cardiac magnetic resonance imaging of patients[16, 17]. In line with the presence of fibrosis in patients, we found increased levels of fibronectin, thrombospondin 4 and periostin (*Figures S1A-C*). Since hypertrophy is a morphological hallmark of HCM, we checked FHL2 expression as a negative regulator of hypertrophy[18]. In line with pro-hypertrophic signaling, reduced FHL2 protein expression was observed (*Figure S1D*). The hypertrophic phenotype was also evident from reduced α-MHC encoded by *MYH6*, and increased CSRP3 expression in HCM samples (*Figures S1E-F*)[19, 20]. Finally, MYBPC3 haploinsufficiency was confirmed in samples with a mutation in *MYBPC3* irrespective of whether the mutation is a truncation or missense mutation (*Figures S1G-H, Table S4*)[14, 21].

Additionally, we performed RNA sequencing in a subset of samples. Similar to the proteomics data, the PCA plot of the RNA data did not show clustering based on genotype (*Figure S2A*). The MA-plot depicts all genes with the differentially expressed genes in red (*Figure S2B*).Gene set enrichment analysis of the proteomics and transcriptomics data revealed a significant correlation of upregulated proteins with genes that showed higher expression levels in HCM compared to NF_{IVS} (FDR<0.001), while less expressed

proteins are positively correlated with genes that showed lower expression levels in HCM (FDR=0.008). This observation suggested that protein expression changes were well correlated with the mRNA expression changes between HCM and NF_{IVS} (*Figure S2C-D*).

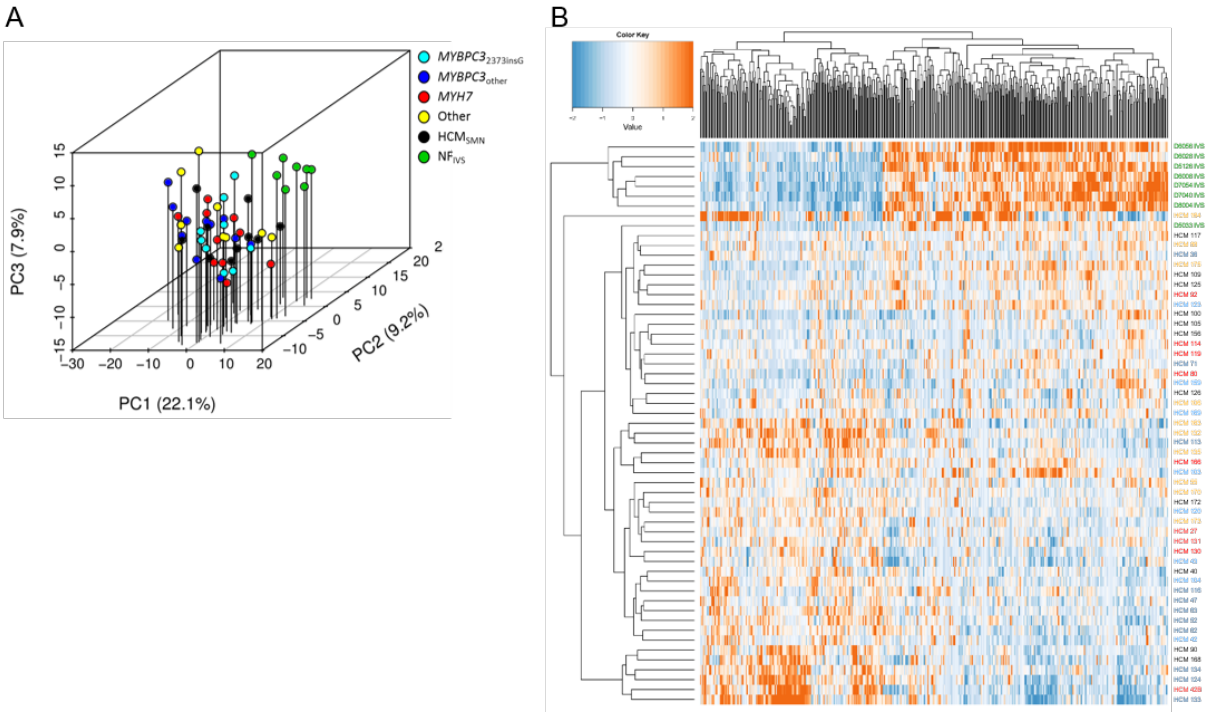


Figure 1: Clustering of proteome of controls and patient samples. (A) Principal component (PC) analysis of the filtered protein expression data. (B) Hierarchical clustering of a multi-group comparison of all proteins with p<0.05 when comparing HCM_{all} with NF_{IVS}.

HCM-specific protein changes

To identify HCM disease-specific protein expression patterns we grouped all HCM samples (HCM_{all}) and compared them to NF_{IVS}. We identified clusters of interacting proteins for higher and lower expressed proteins separately.

Analysis of all significantly lower expressed proteins of HCM_{all} compared to NF_{IVS} uncovered the highest enrichment in oxidative phosphorylation, generation of precursor metabolites and energy, NAD metabolic process, translation, fatty acid catabolic process, regulated exocytosis and neutrophil degranulation (*Figure 2A*, more extensive list in *Figure S3*). This shows that the major changes in tissue of HCM patients are in metabolic pathways related to energy metabolism including both glucose and fatty acid metabolism. We selected all significantly lower expressed proteins annotated to the pathways and created a heatmap displaying the log2-fold change for a detailed visualization of protein changes in the pathway. We included RNA expression data of the corresponding genes to check if the protein changes

coincide with changes at the RNA level. Strikingly, most of these proteins are not changed on RNA level (*Figure S4*).

The top clusters of the higher expressed proteins are associated with the GO terms ECM organization, actin filament-based process, myofibril assembly, muscle contraction, post-translational protein modification, protein folding and microtubule cytoskeleton organization (*Figure 2B*, more extensive list in *Figure S5*). The biggest cluster of more abundant proteins is ECM organization representing different collagens that are components of fibrotic tissue. Interestingly, many of these genes showed unaltered RNA levels compared to NF_{IVS}. The same was observed for the other top clusters of higher expressed proteins, namely actin filament-based process, myofibril assembly and muscle contraction, where RNA levels were unaltered or even lower than in NF_{IVS} (*Figure S6*).

Since HCM is characterized by pronounced hypertrophy of the myocardium, we expected a protein cluster related to hypertrophy among the significantly increased proteins. Surprisingly, we did not find this. This could be because the myocardium is not in a state of active hypertrophic growth at the time of sample collection, or because we did not detect low abundant hypertrophy promoting signaling proteins. It is also known that several signaling proteins are regulated by post-translational modifications rather than by abundance on which our proteomics screen was based. Western blot analysis of the expression and phosphorylation of AKT and ERK, two well-known inducers of hypertrophy[22, 23], showed higher AKT and ERK phosphorylation in HCM_{all} samples, indicating activation of hypertrophic signaling (*Figure S7*).

Subsequently, we repeated the cluster and GO analysis restricted to proteins that are higher and lower expressed if all 5 genotype HCM groups are compared individually to NF_{IVS} to extract the most consistent and robust changes. Venn diagrams (*Figure S8 and S9*) were made to identify overlapping protein changes (76 higher and 92 lower expressed), which were subsequently used as analysis input.

The results confirmed the findings from the initial analysis showing that the major HCM-specific protein changes include reduced metabolism, increased ECM remodeling, and pathways including muscle related processes.

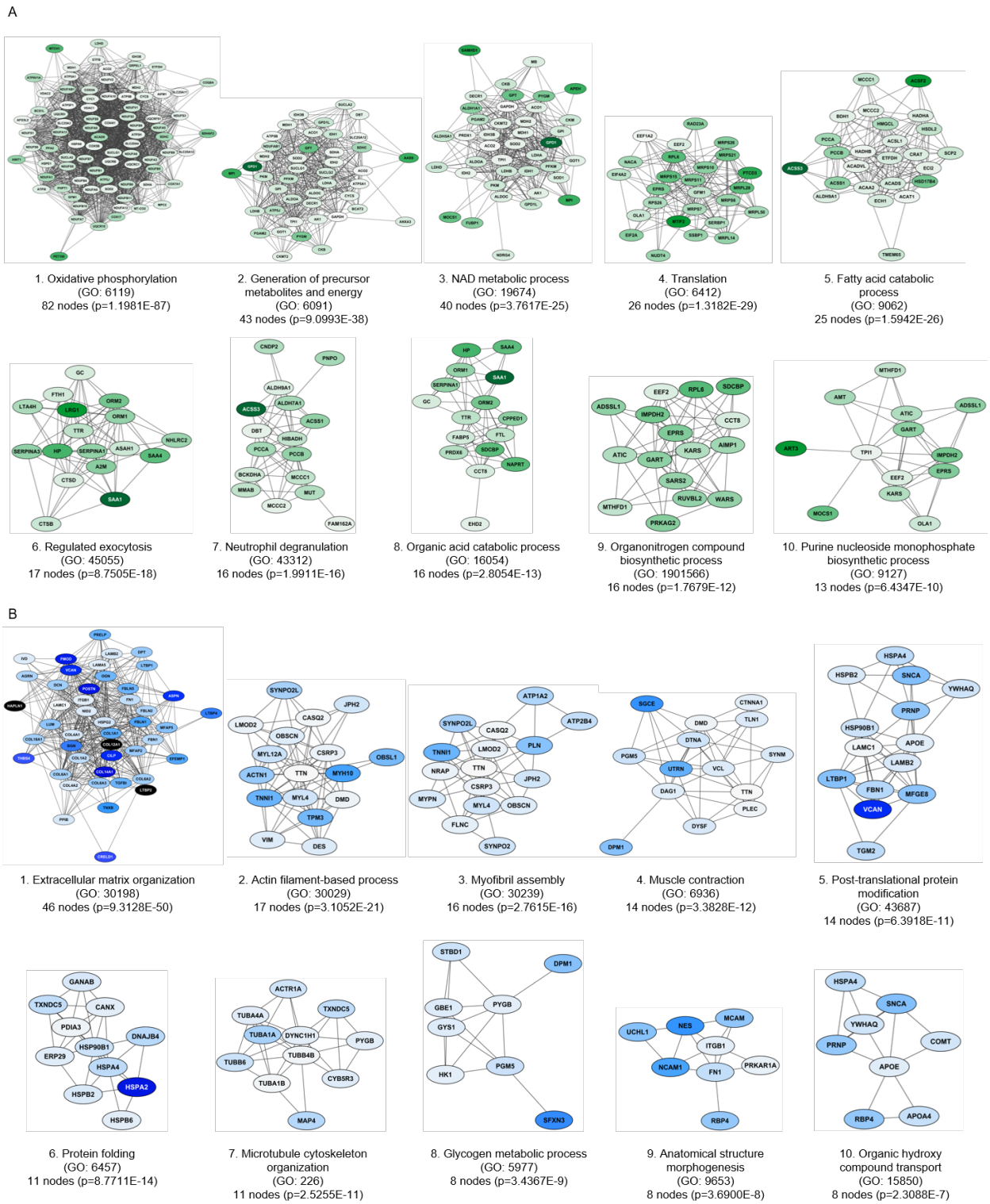


Figure 2: HCM-specific changes in biological processes. Protein interaction cluster of significantly different proteins between HCM_{all} and NF_{IVS} were identified and are displayed with the most significant corresponding gene ontology (GO) term. (A) Top 10 downregulated protein interaction cluster based on cluster size with the most significant biological process related to this cluster. (B) Top 10 upregulated protein interaction cluster based on cluster size with the most significant biological process related to this cluster. The color gradient from light to dark indicates an increase in fold change.

Specific changes for HCM_{SMP} and HCM_{SMN}

To study the pathways that are characteristic for HCM_{SMP} or HCM_{SMN} and might explain sarcomere mutation-specific changes in cardiac function, we created a Venn diagram of significantly different proteins in HCM_{SMP} and HCM_{SMN} compared to NF_{IVS}. The Venn diagrams (*Figures 3 and 4*) illustrate that the majority of significantly different proteins are similar in HCM_{SMP} and HCM_{SMN} (191 lower and 121 higher expressed proteins). However, a substantial number of proteins is only changed in either HCM_{SMP} or HCM_{SMN}.

The 130 proteins that are specifically less abundant in HCM_{SMP} overlap to a large degree with the biological processes of the 194 proteins that are significantly less abundant in both HCM_{SMP} and HCM_{SMN} (*Figure S10*). Analysis of the 62 proteins that are specifically less expressed in HCM_{SMN} results in clusters related to stress granule assembly, translational initiation and protein folding (*Figure 3*).

The majority of the shared higher expressed proteins involve ECM organization (*Figure S11*). GO analysis of the 68 proteins that are specifically more abundant in HCM_{SMP} identified amongst others a protein clusters related to microtubule cytoskeleton organization (*Figure 4*).

This cluster representing tubulin subunits is solely more abundant in HCM_{SMP}. Since the tubulin network is highly regulated by post-translational modifications, we determined the levels of total α-tubulin, tyrosinated and detyrosinated tubulin by western blot (*Figure 5, Figure S12*). In line with the proteomics data we found an increase in total α-tubulin which is more prominent in HCM_{SMP} than in HCM_{SMN} (*Figure 5A*). A post-translational modification of tubulin that leads to increased stiffness of cardiomyocytes is detyrosination of α-tubulin[12]. We found markedly elevated detyrosinated tubulin only in HCM_{SMP} samples (*Figure 5C*), while levels of tyrosinated tubulin were slightly increased in both HCM_{SMP} and HCM_{SMN} (*Figure 5B*). Our data show that high levels of tubulin in HCM_{SMP} represent mostly the detyrosinated form. We also determined desmin protein levels by Western blot since this protein is associated with microtubules in cardiomyocytes. Accordingly, desmin levels were elevated in HCM with the largest increase in HCM_{SMP} (*Figure 5D*).

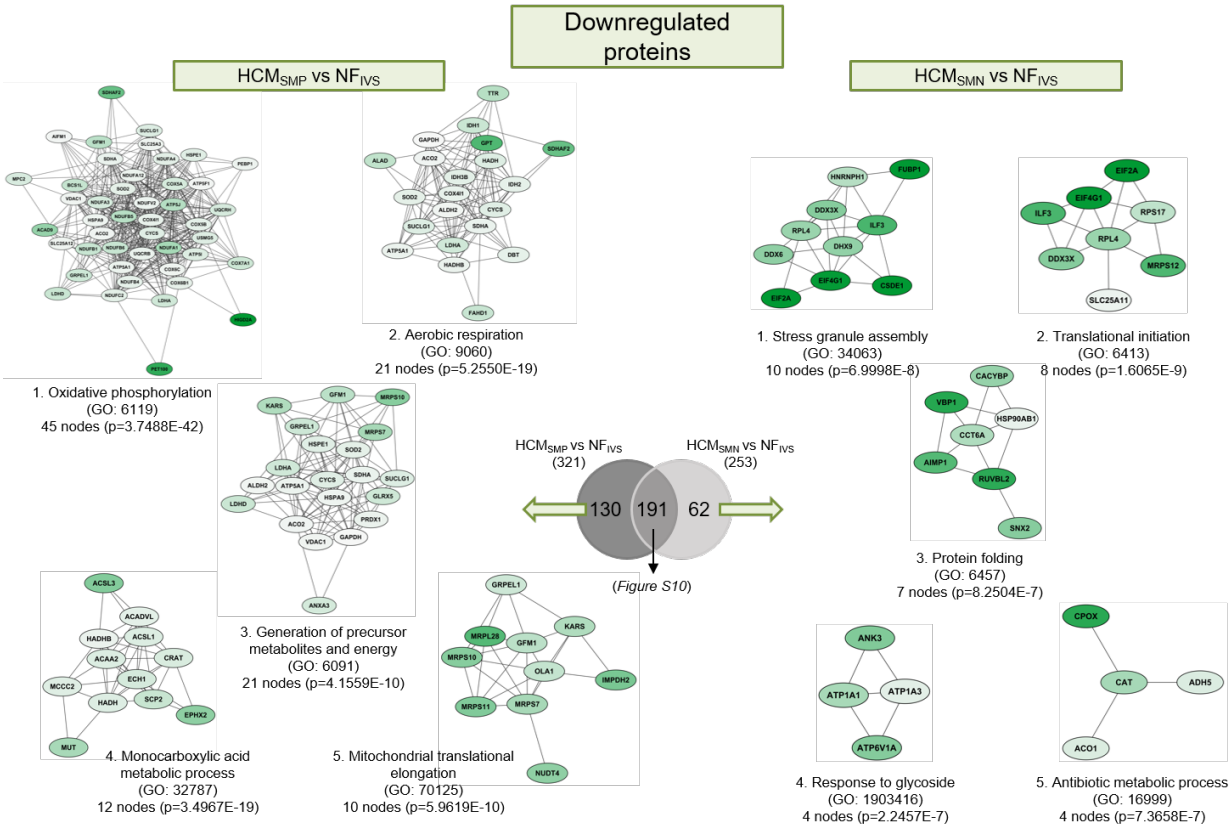


Figure 3: Differences in downregulated proteins between HCM_{SMP} and HCM_{SMN}. Protein interaction cluster of proteins that are only significantly downregulated for the HCM_{SMP} vs NF_{IVS} or the HCM_{SMN} vs NF_{IVS} comparison were identified and are displayed with the most significant corresponding gene ontology (GO) term. The top 5 protein interaction clusters of downregulated proteins are displayed. The color gradient from light to dark indicates an increase in fold change.

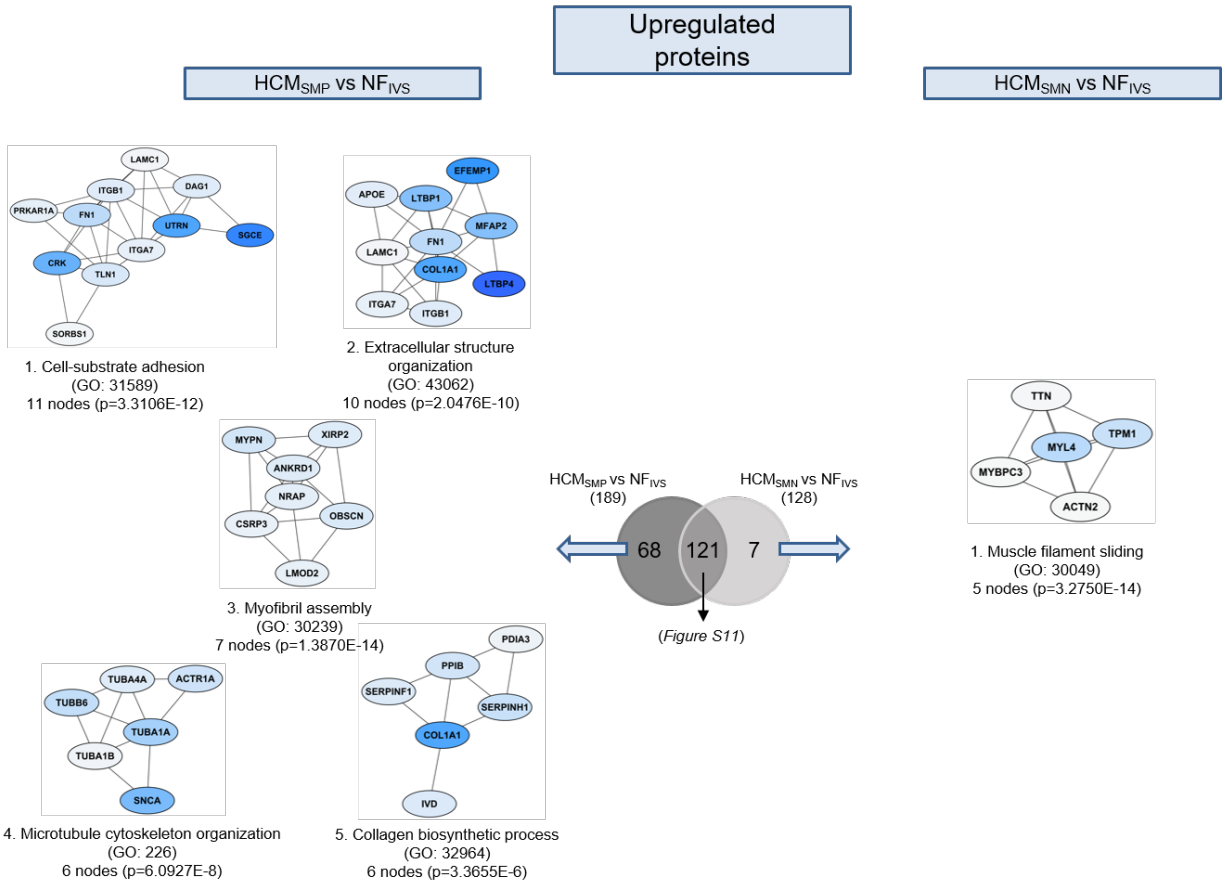


Figure 4: Differences in upregulated proteins between HCM_{SMP} and HCM_{SMN}. Protein interaction cluster of proteins that are only significantly upregulated for the HCM_{SMP} vs NF_{IVS} or the HCM_{SMN} vs NF_{IVS} comparison were identified and are displayed with the most significant corresponding gene ontology (GO) term. The top 5 protein interaction clusters of upregulated proteins are displayed The color gradient from light to dark indicates an increase in fold change.

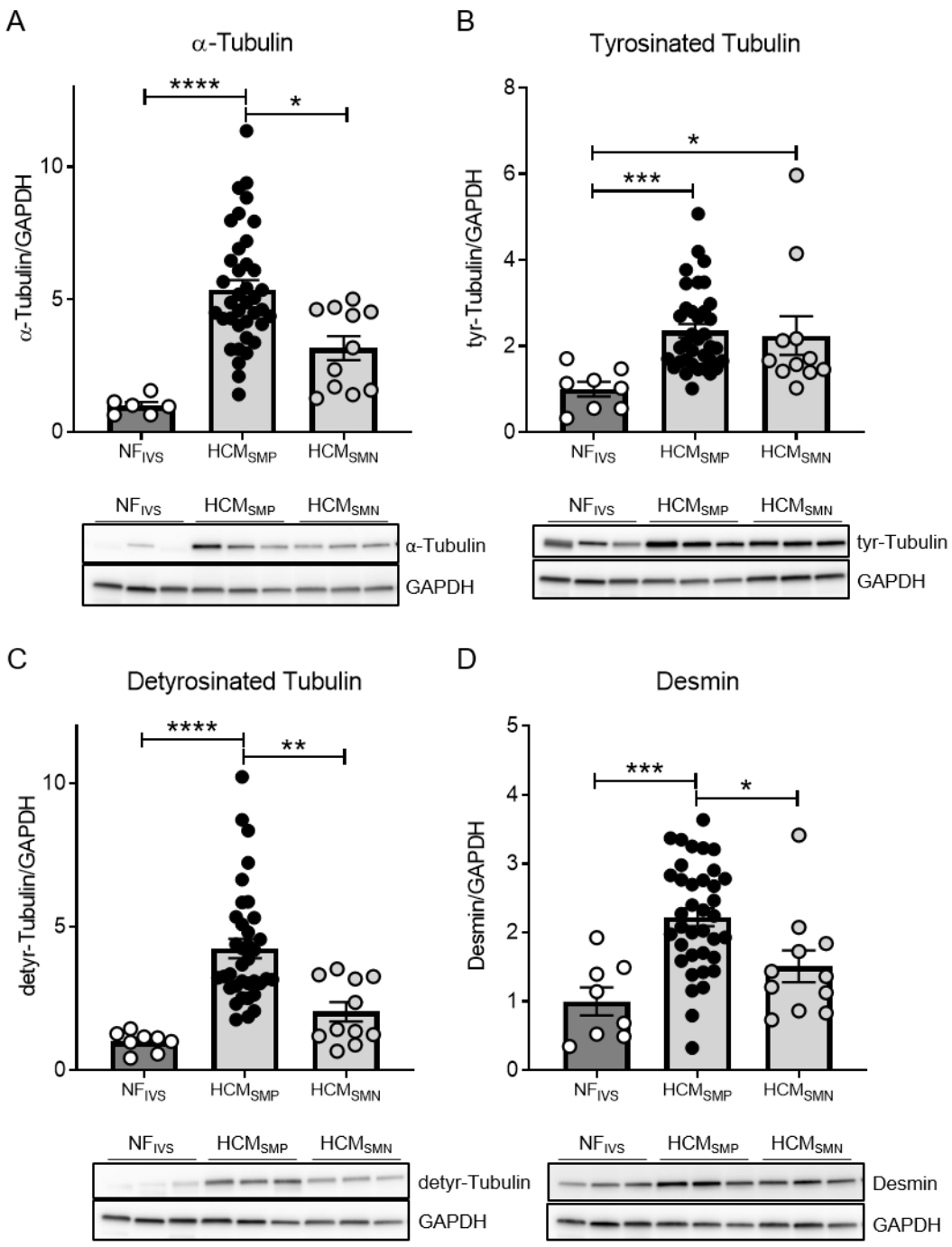


Figure 5: Tubulin expression and posttranslational modifications in HCM patients. Protein levels of α-tubulin (A), its tyrosinated (B) and detyrosinated forms (C) and desmin (D), all normalized to GAPDH, in tissue of HCM patients. Kruskal-Wallis test with Dunn's multiple comparisons test, **** p<0.0001, *** p=0.0003 in (B), *** p=0.0009 in (D), ** p=0.0074, * p=0.0437 in (A), * p=0.0499 in (B) and * p=0.0357 in (D). Average of the control group is set to 1. n(NF_{IVS}/HCM_{SMP}/HCM_{SMN})=6/38/11 for (A), 8/36/11 for (B-C), 8/37/11 for (D).

Inhibition of tubulin detyrosination corrects cardiomyocyte dysfunction in *MYBPC3*$_{2373insG}$ mice

Based on previous studies in human heart failure[13], our data in human myectomy samples indicate that microtubules may represent a treatment target to correct cardiac dysfunction in HCM. To provide proof for a role of tubulin in modulating cardiomyocyte function in HCM caused by a sarcomere gene mutation, we generated a HCM knock-in (KI) mouse model of the Dutch founder mutation c.2373insG in *MYBPC3* (*Figures S13A-B*). This mutation introduces a new splice donor site in exon 25 leading to a frameshift and premature stop codon resulting in an expected truncated protein of 95 kDa[24]. As no truncated protein is found in HCM patients carrying this mutation at the heterozygous state[14], degradation of mutant mRNA and/or protein is likely. Accordingly, Western blot analysis did not reveal any truncated MYBPC3 protein in homozygous *MYBPC3*$_{2373insG}$ mice (*Figure S13C*).

Levels of total and tyrosinated α-tubulin did not differ between the groups, whereas the detyrosinated tubulin level was markedly higher in *MYBPC3*$_{2373insG}$ than in wildtype (WT) mice (*Figures 6A-C)*. To evaluate whether these findings were specific to this model, we assessed the levels of total and detyrosinated tubulin in a second HCM mouse model carrying a different *MYBPC3* mutation[25]. Homozygous *MYBPC3*$_{772G>A}$ mice showed a strong accumulation of α-tubulin and no difference in tyrosinated tubulin compared to their WT littermates. In addition, a trend to higher levels of detyrosinated tubulin was observed (*Figures 6D-F)*. Overall, these mouse models consistently show tubulin changes in cardiomyopathy caused by *MYBPC3* gene mutations.

MYBPC3$_{2373insG}$ mice had a severe cardiac phenotype characterized by higher ventricular weight to body weight ratio, increased LV anterior wall diameter, lower ejection fraction and longer isovolumetric relaxation time than WT mice (*Figures 7A-D*). Single cardiomyocytes from *MYBPC3*$_{2373insG}$ hearts showed contractile deficits compared to WT as shown by an increase in time to peak of contraction and an impaired relaxation shown by an increase in time to baseline (*Figures 7E-F*). Inhibition of detyrosination by treatment with parthenolide (PTL) normalized the contraction and relaxation times in *MYBPC3*$_{2373insG}$ mice to baseline WT levels (*Figures 7E-F*), whereas it had no effect on calcium release and reuptake time in *MYBPC3*$_{2373insG}$ mice (*Figure S14A-B*), indicating a direct effect on myofilament function.

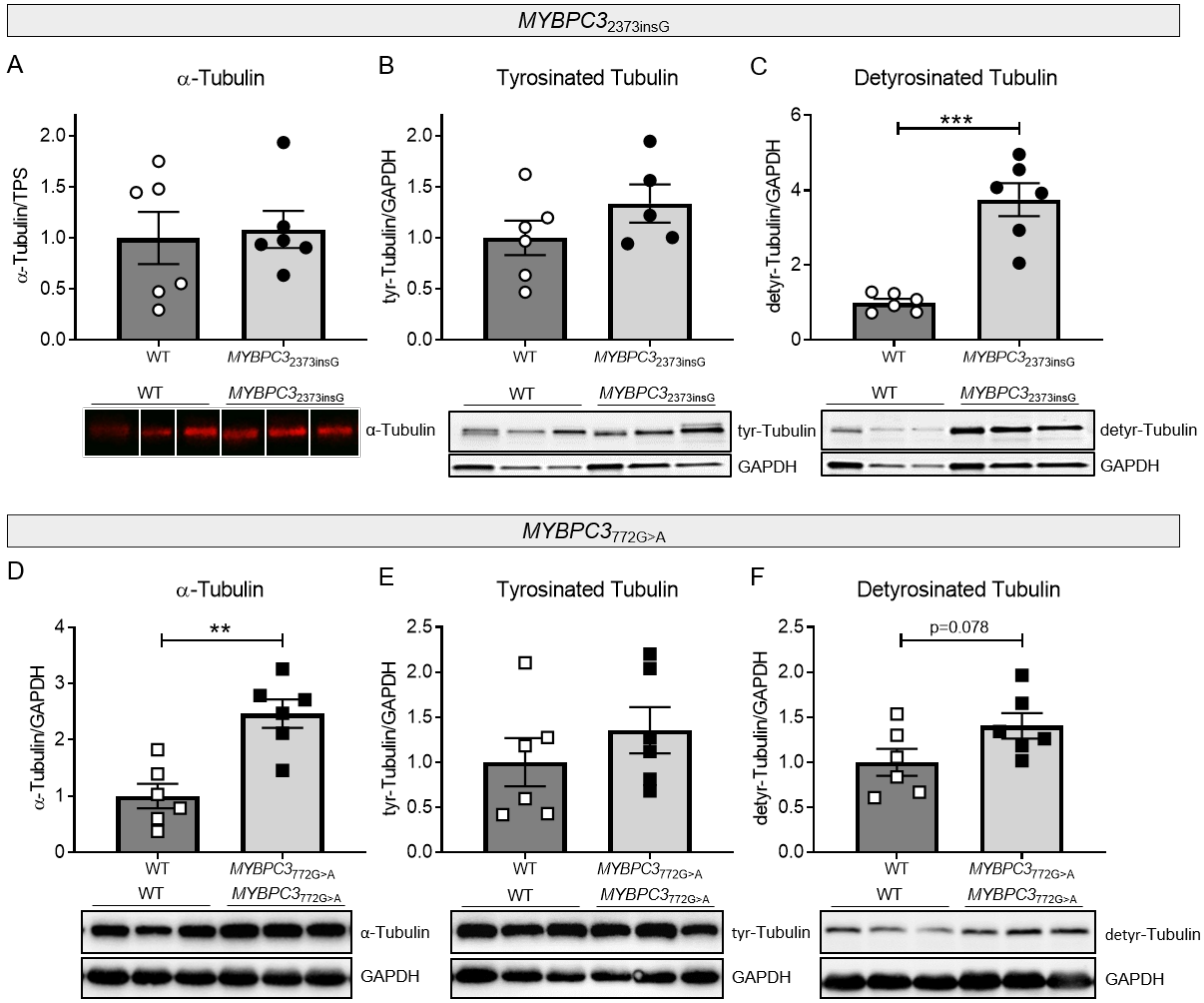


Figure 6: Tubulin composition in *MYBPC3*$_{2373insG}$ and *MYBPC3*$_{772G>A}$ mouse models. Quantification and representative western blot images of (A) α-tubulin, (B) tyrosinated tubulin and (C) detyrosinated tubulin in *MYBPC3*$_{2373insG}$ mice and (D-F) in *MYBPC3*$_{772G>A}$ mice, respectively. (A) is normalized to total protein stain (TPS), (B-F) are normalized to GAPDH. Lanes in (A) were run on the same gel but were noncontiguous. n(WT/*MYBPC3*$_{2373insG/772G>A}$)=6/6 (4 females, 2 males/3 females, 3 males of 20-27 weeks for *MYBPC3*$_{2373insG}$ and the corresponding WT; 2 females, 4 males for of 55-59 weeks for *MYBPC3*$_{772G>A}$ and the corresponding WT) for (A, C-F) and 6/5 (4 females, 2 males/3 females, 2 males; 20-27 weeks) for (B), unpaired two-tailed t-test, *** p=0.0001, ** p=0.0014.

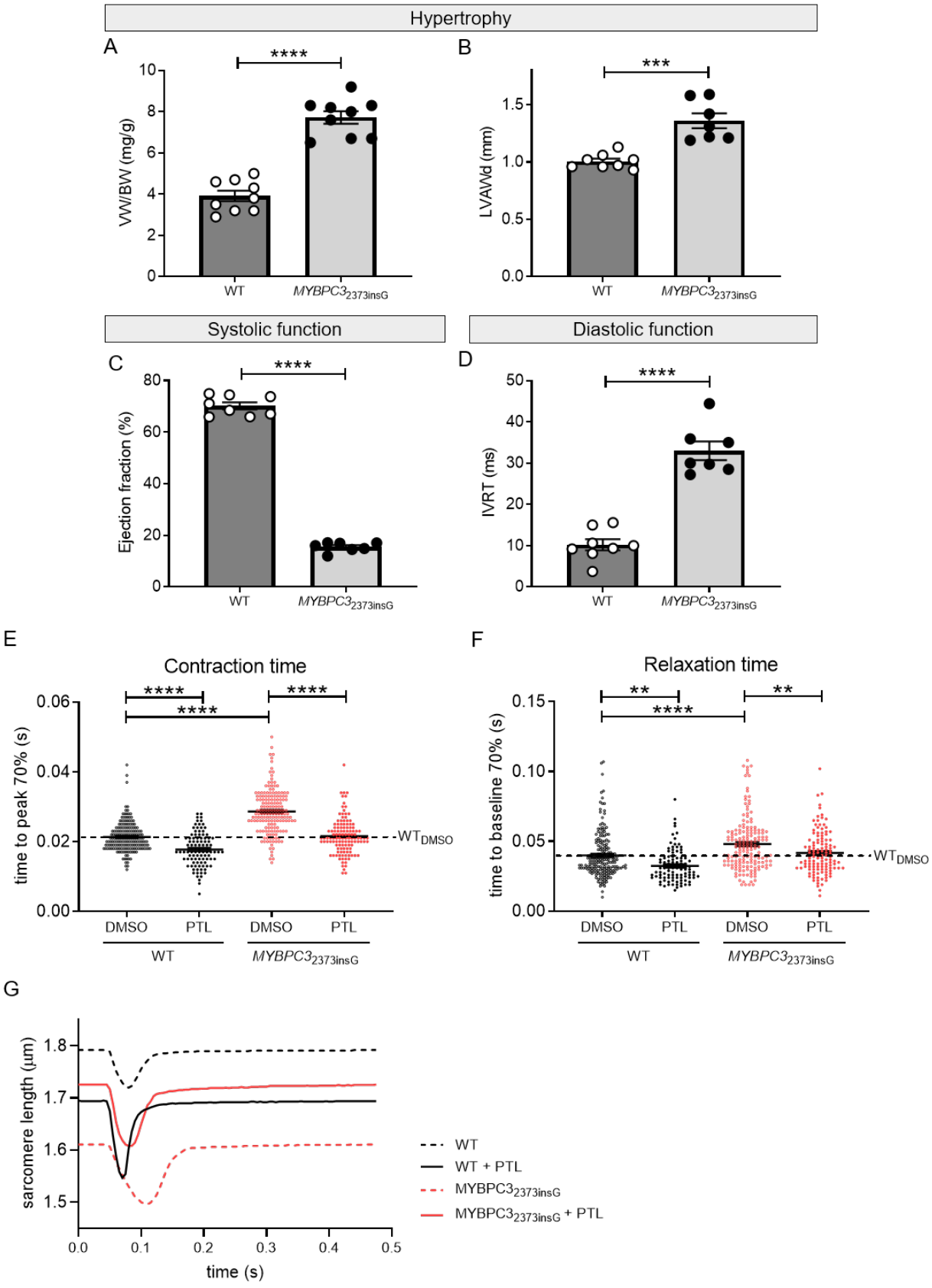

Hypertrophy
A
****
VW/BW (mg/g)
10
8
6
4
2
0
WT
MYBPC3 2373insG
B
***
LVAWd (mm)
1.5
1.0
0.5
0.0
WT
MYBPC3 2373insG
Systolic function
C
****
Ejection fraction (%)
80
60
40
20
0
WT
MYBPC3 2373insG
Diastolic function
D
****
IVRT (ms)
50
40
30
20
10
0
WT
MYBPC3 2373insG
E
Contraction time
****
****
****
time to peak 70% (s)
0.06
0.04
0.02
0.00
WT DMSO
DMSO
PTL
DMSO
PTL
WT
MYBPC3 2373insG
F
Relaxation time
**
****
**
time to baseline 70% (s)
0.15
0.10
0.05
0.00
WT DMSO
DMSO
PTL
DMSO
PTL
WT
MYBPC3 2373insG
G
sarcomere length (µm)
1.8
1.7
1.6
1.5
0.0
0.1
0.2
0.3
0.4
0.5
time (s)
WT
WT + PTL
MYBPC3 2373insG
MYBPC3 2373insG + PTL

Figure 7: Morphometric and phenotypic analysis of $MYBPC3_{2373insG}$ mice and contractile function of isolated $MYBPC3_{2373insG}$ cardiomyocytes upon inhibition of tubulin detyrosination. Quantification of the hypertrophy parameters (A) ventricle weight (VW)/body weight (BW) ratio and (B) wall thickness measured by echocardiography as left ventricular anterior wall diameter in diastole (LVAWd). (C) shows systolic function measured by ejection fraction and (D) diastolic function assessed by isovolumetric relaxation time (IVRT). n(WT/$MYBPC3_{2373insG}$)=9/9 (4 females, 5 males; 20-27 weeks) for (A) and 8/7 (3 females, 5 resp. 4 males; 25-27 weeks) for (B-D), unpaired two-tailed t-test, **** p<0.0001, *** p=0.0001. (E) displays the effect of tubulin detyrosination inhibition by PTL on the contractile parameter time to peak 70% and (F) on the diastolic parameter time to baseline 70%. The dotted line visualizes the WT baseline level. (G) shows example force transients for each condition of the single cell measurements. For (E-F) N(WT mice)=4 (2 females and 2 males, 13-33 weeks) with total n(cells DMSO/PTL)=191/99 and N(*$MYBPC3_{2373insG}$* mice)=6 (2 females and 4 males, 13-35 weeks) with total n(cells DMSO/PTL)=169/123. (E) and (F) were analyzed by 2way-ANOVA, ** p<0.01, **** p<0.0001.

Discussion

In this study we compared the protein profile of cardiac tissue from HCM patients with different disease-causing gene mutations to identify common HCM disease changes as well as genotype-specific protein changes. The majority of detected protein changes were common for all HCM samples and independent of the underlying gene mutation. Our approach revealed different protein profiles in the presence or absence of a sarcomere gene mutation. While hypertrophic remodeling in HCM_{SMP} is characterized by an increase in the levels of proteins involved in microtubule cytoskeleton organization, HCM_{SMN} samples show reduced levels of proteins involved in protein translation.

Deregulated energy metabolism proteome

Our analysis revealed that the majority of deregulated proteins are related to energy metabolism and show a consistent lower expression of proteins involved in oxidative phosphorylation, glycolysis and fatty acid oxidation. Especially subunits of mitochondrial respiratory chain complex I, which are part of cluster 1 in Figure 2A, are consistently lower expressed. This is in line with another recent proteomics study of HCM tissue samples in which reduced levels of energy metabolism proteins was one of the main findings[26]. It also matches observations of energy deficiency in animal models[27] and human studies[28, 29]. Energy deficiency has been proposed as the primary mutation-induced pathomechanism leading to compensatory hypertrophy[30], which is supported by the fact that even asymptomatic mutation carriers without hypertrophy display reduced cardiac energetic status[9, 29]. The reduced cardiac efficiency was larger in *MYH7* mutation carriers compared to *MYBPC3* mutation carriers pointing towards genotype-specific functional differences[9]. Our proteomic analyses did not show gene-dependent differences in proteins involved in cardiac energy metabolism. This indicates that the deregulated energy metabolism proteome is a secondary consequence due to cellular stress that is similar in patients with advanced HCM.

Lack of direct correlation between changes in protein and mRNA level in selected pathways

Interestingly, most of the significantly different proteins did not show the same changes at the RNA level (*Figure S4, S6*), despite an overall correlation of proteomics and transcriptomics data (*Figures S2C-D*).

A partial mismatch of RNA and protein expression data has also been observed in an integrated omics analysis of heart failure and sudden cardiac death[31], demonstrating that this finding is not unique. The divergence of RNA and protein expression may partly be explained by post-translational regulation and different half-life of proteins and RNAs with RNAs being less stable than proteins. The mRNA expression data can give indications about ongoing processes in the cell as reaction to the pathological stimulus, while changes at the protein level reflect the actual functional response.

Protein homeostasis differs between HCM_{SMP} and HCM_{SMN}

Interestingly, proteins involved in protein translation are less abundant, particularly in HCM_{SMN} patient tissue implying that the protein translation system is either more impaired or differently regulated in HCM_{SMN}. Among the more abundant proteins, we observed a protein cluster related to protein folding. Protein folding proteins (chaperones) are needed for correct folding of de novo synthesized proteins as well as for refolding of misfolded mutant or damaged proteins. Since our data show downregulation of protein translation, it is unlikely that protein folding proteins are upregulated for folding of newly synthesized proteins. Instead, we speculate that their expression is upregulated to repair or remove misfolded (mutant) proteins[32]. This is in line with recent data describing an upregulation of protein folding proteins specifically in HCM_{SMP} samples[33]. Protein folding proteins represent potential treatment targets since boosting their expression has already shown beneficial effects in animal models of other cardiac diseases and in cardiomyopathies[34-36]. Furthermore, activation of autophagy in *MYBPC3*$_{772G>A}$ mice rescued cardiac dysfunction[37]. Induction of heat shock proteins also produced beneficial effects in a model for atrial fibrillation[38] and myocardial infarction[39]. In the latter study, the induction of mitochondrial heat shock proteins preserved mitochondrial function, which would be a desirable result in HCM patients to counteract the prominent mitochondrial dysfunction.

Impact of increased tubulin network on cardiomyocyte function in HCM

A striking observation was the specific upregulation of microtubule subunits and post-translational modification detyrosination in HCM_{SMP} compared to HCM_{SMN}. Previous studies indicated important regulatory roles of microtubules in cardiomyocyte function[40]. Recent reports showed that the translation

of the sarcomere proteins is localized to the myofilaments which points to a role of microtubules in the transportation of mRNA to the myofilament[41, 42]. Increased expression of tubulin subunits strengthens the microtubule network and facilitates the transportation of mRNAs to the sarcomere and concomitant increasing cardiomyocyte stability. In addition to tubulins, desmin protein level was significantly higher in HCM_{SMP} than in HCM_{SMN}. In the healthy heart, desmin is localized at the Z-discs forming a striated pattern and playing a central role in cardiomyocyte mechanical stability[43]. Overall, these protein changes suggest a compensatory mechanism of the cell to ensure sarcomere stability in the presence of a sarcomere gene mutation.

Furthermore, microtubules and their post-translational modifications play an important role in cardiomyocyte mechanics, especially in regulating cardiomyocyte stiffness[44, 45]. Detyrosinated tubulin stabilizes microtubules by inhibiting disassembly[46] and can anchor microtubules to the Z-discs of the sarcomere, most likely via desmin, and enhance stability and stiffness of the myofilaments and the microtubular network[44]. Tubulin detyrosination is enzymatically regulated by the tubulin tyrosine ligase (TTL) and tubulin carboxypeptidases (TCPs) that have detyrosinating activity. Recently, vasohibins have been identified as the first tubulin detyrosinating enzymes[47, 48], and act in complex with small vasohibin-binding protein (SVBP). In accordance with findings by Robison et al. in explanted hearts of heart failure and cardiomyopathy patients[12], we observed a specific upregulation of tubulins and enhanced detyrosination in our genetically well-characterized HCM myectomy samples. Notably, we found the increase in tubulin detyrosination specific for HCM_{SMP} samples. Chen et al. have already demonstrated the reversibility of tubulin detyrosination in isolated cardiomyocytes from explanted HCM hearts associated with an improvement of contractile function[13, 49]. Here we performed proof-of-concept studies in HCM mouse models to define the impact of tubulin detyrosination in the presence of a sarcomere mutation. Homozygous $MYBPC3_{2373insG}$ mice replicated the tubulin detyrosination without an increase in total α-tubulin levels, whereas homozygous $MYBPC3_{772G>A}$ mice display an increase in total α-tubulin with a trend to increased tubulin detyrosination. In $MYBPC3_{2373insG}$ mice we showed that the increase in tubulin detyrosination is accompanied by reduced contraction and relaxation kinetics in isolated intact cardiomyocytes. Inhibition of tubulin detyrosination by PTL restored contraction and relaxation kinetics to WT levels in $MYBPC3_{2373insG}$ cardiomyocytes. The used concentration of PTL has already been proven to be sufficient to reduce detyrosination of tubulin[12, 50] and Chen et al. and Robison et al. have demonstrated that the PTL-induced effects on kinetics of contraction are the same as obtained by overexpression of TTL[12, 13], which specifically lowers tubulin detyrosination. The positive effect of PTL on cardiomyocyte

function of *MYBPC3*$_{2373insG}$ cardiomyocytes is therefore explained by a decrease in cardiomyocyte stiffness due to lower levels of detyrosinated tubulin.

Overall, our findings in a European HCM patient cohort strengthens the evidence that increased detyrosination of microtubules contributes to cardiomyocyte stiffness and dysfunction in HCM. Importantly, we show that this is especially true in the presence of a sarcomere mutation.

Study limitations and clinical implications

The number of genotype-specific protein changes was low, which is likely due to the fact that all samples were obtained from obstructive HCM patients. Genotype-specific differences might have occurred at earlier disease stages and declined with disease progression. Due to unavailability of myocardial biopsies from asymptomatic mutation carriers, animal models and human-derived cardiomyocyte muscle models should provide insight into gene-specific pathomechanisms at early disease stages.

Most of the patients in this study have been on drug therapy, therefore we cannot exclude that some of the proteomic changes are caused by medication. However, some of the main findings, e.g. energy deficiency, have also been observed in animal models[27], and the number of patients on the most commonly used drugs does not differ between HCM_{SMP} and HCM_{SMN} (*Table S7)*. Therefore it is more likely that the observed changes are driven by disease rather than medication.

Although we detected and quantified a large number of proteins in this study, we are limited to more abundant cardiac proteins why our list of deregulated proteins is incomplete. Also, we did not assess post-translational modifications in the proteomics study which have important regulatory functions. The changes at the protein level reflect whole tissue alterations and cannot be solely assigned to the mutation-carrying cardiomyocytes but might also arise from other cell types. However, cardiomyocytes are responsible for most of the tissue volume and therefore drive most of the protein changes.

Homozygous *MYBPC3*$_{2373insG}$ mice do not resemble a perfect model of HCM since human patients carry heterozygous mutations. But due to the lack of phenotype in heterozygous mice they provide a suitable genetic model of loss of MYBPC3 protein levels.

On the basis of our proteomic screen in a large set of human HCM samples and functional studies in a novel HCM mouse model, we propose that an increase in detyrosinated tubulin contributes to the clinical and cellular differences that we see between HCM_{SMP} and HCM_{SMN} samples (*Figure 8*). Detyrosinated microtubules may represent a target for therapeutic intervention in genetic heart disease because reducing detyrosination improves contractile function in isolated cardiomyocytes. Therefore, a specific inhibitor of detyrosination needs to be developed due to the known off-target effects of PTL. As

sarcomere mutation carriers are identified before disease onset, targeting the microtubules may represent a preventive treatment option.

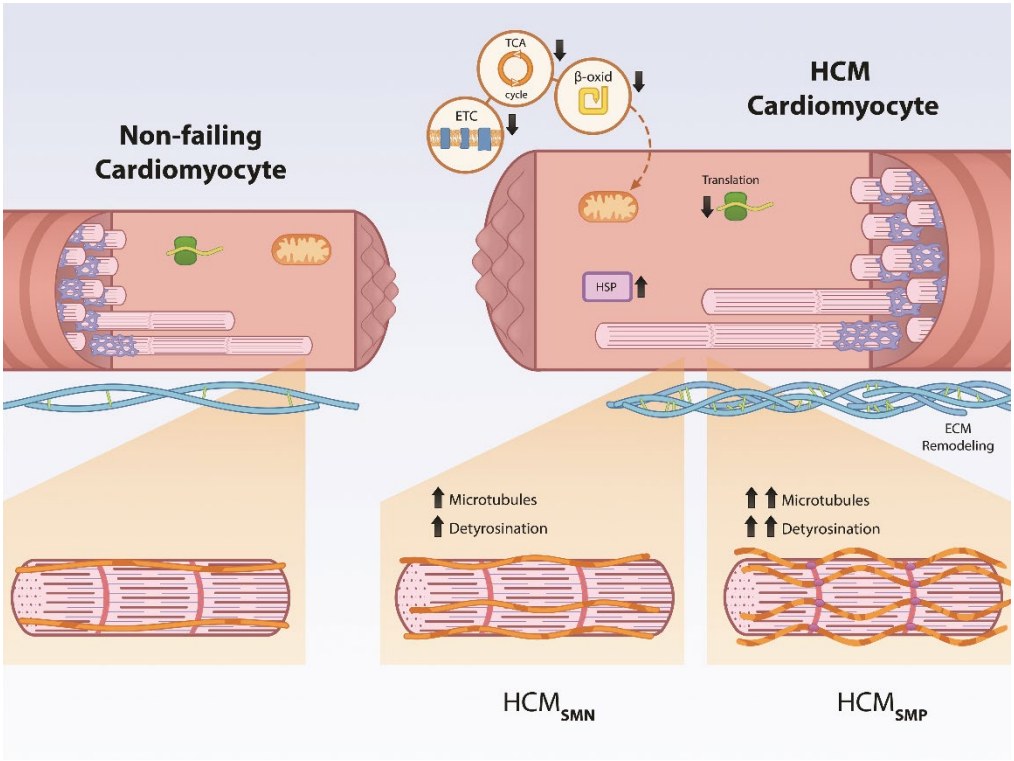


Figure 8: Schematic representation of genotype-independent changes in HCM and genotype-specific differences in the microtubular system. Our analysis shows that all HCM patients display downregulation of metabolic pathways (electron transport chain (ETC), tricarboxylic acid (TCA) cycle, β-oxidation (β-oxid)) and ribosomal proteins (Translation), as well as an upregulation of protein folding proteins (heat shock proteins, HSPs) and extracellular matrix (ECM) proteins. HCM_{SMP} patients have a large increase in microtubules and levels of its detyrosinated form, whereas HCM_{SMN} patients only have a slight increase compared to non-failing controls.

Acknowledgements

We would like to thank Ruud Zaremba, Valentijn Jansen and Max Goebel for technical assistance. We thank Feng Zhang for providing the Addgene plasmid #42230.

Sources of Funding

We acknowledge the support from the Netherlands Cardiovascular Research Initiative: An initiative with support of the Dutch Heart Foundation, CVON2014-40 DOSIS and NWO (NWO-ZonMW; 91818602 VICI grant to Jolanda van der Velden). Folkert Asselbergs is supported by UCL Hospitals NIHR Biomedical

Research Centre. Magdalena Harakalova is supported by NWO VENI grant (no. 016.176.136). Lucie Carrier and Saskia Schlossarek are supported by the German Centre for Cardiovascular Research (DZHK), the German Ministry of Research Education (BMBF), the Deutsche Herzstiftung and the Helmut und Charlotte Kassau Stiftung.

Disclosures

None.

Supplementary Files – Online

Supplementary Material. Extended section on methods and materials.

Supplementary Table 1. Patient characteristics.

Supplementary Table 2. Descriptive matrix of the different group comparisons.

Supplementary Table 3. Expression of expected deregulated proteins in HCM patient tissue.

Supplementary Table 4. Expression of MYBPC3 in the $MYBPC3_{2373insG}$ patient group.

Supplementary Table 5. The full list of detected proteins and their expression levels in HCM and control samples.

Supplementary Table 6. Statistics of RNA sequencing results.

Supplementary Table 7. Chi-square test comparing the number of patients on the three most commonly used drugs in HCM_{SMP} and HCM_{SMN}.

Supplementary Figure 1. Validation of proteomics analysis with known target proteins.

Supplementary Figure 2. RNA sequencing results.

Supplementary Figure 3. HCM-specific changes in biological processes in downregulated proteins.

Supplementary Figure 4. Top 4 downregulated and functionally non-repetitive GOs showing all significantly deregulated proteins and RNAs comparing HCM_{all} with NF_{IVS} annotated to the corresponding GO term.

Supplementary Figure 5. HCM-specific changes in biological processes in upregulated proteins.

Supplementary Figure 6. Top 4 upregulated and functionally non-repetitive GOs showing all significantly deregulated proteins and RNAs comparing HCM_{all} with NF_{IVS} annotated to the corresponding GO term.

Supplementary Figure 7. Cardiac hypertrophy markers.

Supplementary Figure 8. Venn diagram of significantly downregulated proteins of all 5 mutation groups compared to NF_{IVS}.

Supplementary Figure 9. Venn diagram of significantly upregulated proteins of all 5 mutation groups compared to NF_{IVS}.

Supplementary Figure 10. Protein interaction cluster of all significantly downregulated proteins that are common between HCM_{SMP} compared to NF_{IVS} and HCM_{SMN} compared to NF_{IVS}.

Supplementary Figure 11. Protein interaction cluster of all significantly upregulated proteins that are common between HCM_{SMP} compared to NF_{IVS} and HCM_{SMN} compared to NF_{IVS}.

Supplementary Figure 12. α-Tubulin (α-3A1) western blot data. Protein levels of α-tubulin with the α-3A1 antibody normalized to GAPDH in tissue of HCM patients.

Supplementary Figure 13. $MYBPC3_{2373insG}$ mouse model.

Supplementary Figure 14. Calcium kinetics of isolated $MYBPC3_{2373insG}$ cardiomyocytes upon inhibition of tubulin detyrosination.

Supplementary Figure 15. Images of gels that were used for proteomics experiment with slicing scheme.

References

1. Michels M, Olivotto I, Asselbergs FW and van der Velden J. Life-long tailoring of management for patients with hypertrophic cardiomyopathy : Awareness and decision-making in changing scenarios. *Neth Heart J*. 2017;25:186-199. doi: 10.1007/s12471-016-0943-2
2. Ingles J, Burns C, Barratt A and Semsarian C. Application of Genetic Testing in Hypertrophic Cardiomyopathy for Preclinical Disease Detection. *Circ Cardiovasc Genet*. 2015;8:852-9. doi: 10.1161/CIRCGENETICS.115.001093
3. Authors/Task Force, Elliott PM, Anastasakis A, Borger MA, Borggrefe M, Cecchi F, Charron P, Hagege AA, Lafont A, Limongelli G, Mahrholdt H, McKenna WJ, Mogensen J, Nihoyannopoulos P, Nistri S, Pieper PG, Pieske B, Rapezzi C, Rutten FH, Tillmanns C and Watkins H. 2014 ESC Guidelines on diagnosis and management of hypertrophic cardiomyopathy: the Task Force for the Diagnosis and Management of Hypertrophic Cardiomyopathy of the European Society of Cardiology (ESC). *Eur Heart J*. 2014;35:2733-79. doi: 10.1093/eurheartj/ehu284
4. Richard P, Charron P, Carrier L, Ledeuil C, Cheav T, Pichereau C, Benaiche A, Isnard R, Dubourg O, Burban M, Gueffet JP, Millaire A, Desnos M, Schwartz K, Hainque B, Komajda M and Project EHF. Hypertrophic cardiomyopathy: distribution of disease genes, spectrum of mutations, and implications for a molecular diagnosis strategy. *Circulation*. 2003;107:2227-32. doi: 10.1161/01.CIR.0000066323.15244.54
5. Vakrou S, Fukunaga R, Foster DB, Sorensen L, Liu Y, Guan Y, Woldemichael K, Pineda-Reyes R, Liu T, Tardiff JC, Leinwand LA, Tocchetti CG, Abraham TP, O'Rourke B, Aon MA and Abraham MR. Allele-specific differences in transcriptome, miRNome, and mitochondrial function in two hypertrophic cardiomyopathy mouse models. *JCI Insight*. 2018;3. doi: 10.1172/jci.insight.94493
6. Sequeira V, Wijnker PJM, Nijenkamp LLaM, Kuster DWD, Najafi A, Witjas-Paalberends ER, Regan Ja, Boontje N, Ten Cate FJ, Germans T, Carrier L, Sadayappan S, van Slegtenhorst Ma, Zaremba R, Foster DB, Murphy AM, Poggesi C, Dos Remedios C, Stienen GJM, Ho CY, Michels M and van der Velden J. Perturbed length-dependent activation in human hypertrophic cardiomyopathy with missense sarcomeric gene mutations. *Circulation research*. 2013;112:1491-505. doi: 10.1161/CIRCRESAHA.111.300436
7. Sequeira V, Najafi A, Wijnker PJ, Dos Remedios CG, Michels M, Kuster DW and van der Velden J. ADP-stimulated contraction: A predictor of thin-filament activation in cardiac disease. *Proc Natl Acad Sci U S A*. 2015;112:E7003-12. doi: 10.1073/pnas.1513843112
8. Ojala M, Prajapati C, Polonen RP, Rajala K, Pekkanen-Mattila M, Rasku J, Larsson K and Aalto-Setala K. Mutation-Specific Phenotypes in hiPSC-Derived Cardiomyocytes Carrying Either Myosin-Binding Protein C Or alpha-Tropomyosin Mutation for Hypertrophic Cardiomyopathy. *Stem Cells Int*. 2016;2016:1684792. doi: 10.1155/2016/1684792
9. Witjas-Paalberends ER, Guclu A, Germans T, Knaapen P, Harms HJ, Vermeer AM, Christiaans I, Wilde AA, Dos Remedios C, Lammertsma AA, van Rossum AC, Stienen GJ, van Slegtenhorst M, Schinkel AF, Michels M, Ho CY, Poggesi C and van der Velden J. Gene-specific increase in the energetic cost of contraction in hypertrophic cardiomyopathy caused by thick filament mutations. *Cardiovasc Res*. 2014;103:248-57. doi: 10.1093/cvr/cvu127
10. Ho CY, Lakdawala NK, Cirino AL, Lipshultz SE, Sparks E, Abbasi SA, Kwong RY, Antman EM, Semsarian C, Gonzalez A, Lopez B, Diez J, Orav EJ, Colan SD and Seidman CE. Diltiazem treatment for pre-clinical hypertrophic cardiomyopathy sarcomere mutation carriers: a pilot randomized trial to modify disease expression. *JACC Heart Fail*. 2015;3:180-8. doi: 10.1016/j.jchf.2014.08.003
11. Ho CY, Day SM, Ashley EA, Michels M, Pereira AC, Jacoby D, Cirino AL, Fox JC, Lakdawala NK, Ware JS, Caleshu CA, Helms AS, Colan SD, Girolami F, Cecchi F, Seidman CE, Sajeev G, Signorovitch J, Green EM and Olivotto I. Genotype and Lifetime Burden of Disease in Hypertrophic Cardiomyopathy:

Insights from the Sarcomeric Human Cardiomyopathy Registry (SHaRe). *Circulation*. 2018;138:1387-1398. doi: 10.1161/CIRCULATIONAHA.117.033200
12. Robison P, Caporizzo MA, Ahmadzadeh H, Bogush AI, Chen CY, Margulies KB, Shenoy VB and Prosser BL. Detyrosinated microtubules buckle and bear load in contracting cardiomyocytes. *Science*. 2016;352:aaf0659. doi: 10.1126/science.aaf0659
13. Chen CY, Caporizzo MA, Bedi K, Vite A, Bogush AI, Robison P, Heffler JG, Salomon AK, Kelly NA, Babu A, Morley MP, Margulies KB and Prosser BL. Suppression of detyrosinated microtubules improves cardiomyocyte function in human heart failure. *Nat Med*. 2018;24:1225-1233. doi: 10.1038/s41591-018-0046-2
14. van Dijk SJ, Dooijes D, dos Remedios C, Michels M, Lamers JM, Winegrad S, Schlossarek S, Carrier L, ten Cate FJ, Stienen GJ and van der Velden J. Cardiac myosin-binding protein C mutations and hypertrophic cardiomyopathy: haploinsufficiency, deranged phosphorylation, and cardiomyocyte dysfunction. *Circulation*. 2009;119:1473-83. doi: 10.1161/CIRCULATIONAHA.108.838672
15. Alfares AA, Kelly MA, McDermott G, Funke BH, Lebo MS, Baxter SB, Shen J, McLaughlin HM, Clark EH, Babb LJ, Cox SW, DePalma SR, Ho CY, Seidman JG, Seidman CE and Rehm HL. Results of clinical genetic testing of 2,912 probands with hypertrophic cardiomyopathy: expanded panels offer limited additional sensitivity. *Genetics In Medicine*. 2015;17:880. doi: 10.1038/gim.2014.205
16. Nijenkamp L, Bollen IAE, van Velzen HG, Regan JA, van Slegtenhorst M, Niessen HWM, Schinkel AFL, Kruger M, Poggesi C, Ho CY, Kuster DWD, Michels M and van der Velden J. Sex Differences at the Time of Myectomy in Hypertrophic Cardiomyopathy. *Circ Heart Fail*. 2018;11:e004133. doi: 10.1161/CIRCHEARTFAILURE.117.004133
17. Maron MS, Appelbaum E, Harrigan CJ, Buros J, Gibson CM, Hanna C, Lesser JR, Udelson JE, Manning WJ and Maron BJ. Clinical profile and significance of delayed enhancement in hypertrophic cardiomyopathy. *Circ Heart Fail*. 2008;1:184-91. doi: 10.1161/CIRCHEARTFAILURE.108.768119
18. Purcell NH, Darwis D, Bueno OF, Muller JM, Schule R and Molkentin JD. Extracellular signal-regulated kinase 2 interacts with and is negatively regulated by the LIM-only protein FHL2 in cardiomyocytes. *Mol Cell Biol*. 2004;24:1081-95. doi: 10.1128/mcb.24.3.1081-1095.2004
19. Lowes BD, Minobe W, Abraham WT, Rizeq MN, Bohlmeyer TJ, Quaife RA, Roden RL, Dutcher DL, Robertson AD, Voelkel NF, Badesch DB, Groves BM, Gilbert EM and Bristow MR. Changes in gene expression in the intact human heart. Downregulation of alpha-myosin heavy chain in hypertrophied, failing ventricular myocardium. *J Clin Invest*. 1997;100:2315-24. doi: 10.1172/JCI119770
20. Nakao K, Minobe W, Roden R, Bristow MR and Leinwand LA. Myosin heavy chain gene expression in human heart failure. *J Clin Invest*. 1997;100:2362-70. doi: 10.1172/JCI119776
21. Marston S, Copeland O, Gehmlich K, Schlossarek S and Carrier L. How do MYBPC3 mutations cause hypertrophic cardiomyopathy? *J Muscle Res Cell Motil*. 2012;33:75-80. doi: 10.1007/s10974-011-9268-3
22. Heineke J and Molkentin JD. Regulation of cardiac hypertrophy by intracellular signalling pathways. *Nat Rev Mol Cell Biol*. 2006;7:589-600. doi: 10.1038/nrm1983
23. Matsui T, Nagoshi T and Rosenzweig A. Akt and PI 3-kinase signaling in cardiomyocyte hypertrophy and survival. *Cell Cycle*. 2003;2:220-3. doi:
24. Moolman Ja, Reith S, Uhl K, Bailey S, Gautel M, Jeschke B, Fischer C, Ochs J, McKenna WJ, Klues H and Vosberg HP. A newly created splice donor site in exon 25 of the MyBP-C gene is responsible for inherited hypertrophic cardiomyopathy with incomplete disease penetrance. *Circulation*. 2000;101:1396-1402. doi: 10.1161/01.CIR.101.12.1396
25. Vignier N, Schlossarek S, Fraysse B, Mearini G, Kramer E, Pointu H, Mougenot N, Guiard J, Reimer R, Hohenberg H, Schwartz K, Vernet M, Eschenhagen T and Carrier L. Nonsense-mediated mRNA decay and ubiquitin-proteasome system regulate cardiac myosin-binding protein C mutant levels in cardiomyopathic mice. *Circ Res*. 2009;105:239-48. doi: 10.1161/CIRCRESAHA.109.201251

26. Coats CJ, Heywood WE, Virasami A, Ashrafi N, Syrris P, Dos Remedios C, Treibel TA, Moon JC, Lopes LR, McGregor CGA, Ashworth M, Sebire NJ, McKenna WJ, Mills K and Elliott PM. Proteomic Analysis of the Myocardium in Hypertrophic Obstructive Cardiomyopathy. *Circ Genom Precis Med*. 2018;11:e001974. doi: 10.1161/CIRCGEN.117.001974
27. Luedde M, Flogel U, Knorr M, Grundt C, Hippe HJ, Brors B, Frank D, Haselmann U, Antony C, Voelkers M, Schrader J, Most P, Lemmer B, Katus HA and Frey N. Decreased contractility due to energy deprivation in a transgenic rat model of hypertrophic cardiomyopathy. *J Mol Med (Berl)*. 2009;87:411-22. doi: 10.1007/s00109-008-0436-x
28. Guclu A, Knaapen P, Harms HJ, Parbhudayal RY, Michels M, Lammertsma AA, van Rossum AC, Germans T and van der Velden J. Disease Stage-Dependent Changes in Cardiac Contractile Performance and Oxygen Utilization Underlie Reduced Myocardial Efficiency in Human Inherited Hypertrophic Cardiomyopathy. *Circ Cardiovasc Imaging*. 2017;10. doi: 10.1161/CIRCIMAGING.116.005604
29. Crilley JG, Boehm EA, Blair E, Rajagopalan B, Blamire AM, Styles P, McKenna WJ, Ostman-Smith I, Clarke K and Watkins H. Hypertrophic cardiomyopathy due to sarcomeric gene mutations is characterized by impaired energy metabolism irrespective of the degree of hypertrophy. *J Am Coll Cardiol*. 2003;41:1776-82. doi:
30. Ashrafian H, Redwood C, Blair E and Watkins H. Hypertrophic cardiomyopathy:a paradigm for myocardial energy depletion. *Trends Genet*. 2003;19:263-8. doi: 10.1016/S0168-9525(03)00081-7
31. Foster DB, Liu T, Kammers K, O'Meally R, Yang N, Papanicolaou KN, Talbot CC, Jr., Cole RN and O'Rourke B. Integrated Omic Analysis of a Guinea Pig Model of Heart Failure and Sudden Cardiac Death. *J Proteome Res*. 2016;15:3009-28. doi: 10.1021/acs.jproteome.6b00149
32. Tarone G and Brancaccio M. Keep your heart in shape: molecular chaperone networks for treating heart disease. *Cardiovasc Res*. 2014;102:346-61. doi: 10.1093/cvr/cvu049
33. Dorsch LM, Schuldt M, dos Remedios CG, Schinkel AFL, de Jong PL, Michels M, Kuster DWD, Brundel B and van der Velden J. Protein Quality Control Activation and Microtubule Remodeling in Hypertrophic Cardiomyopathy. *Cells*. 2019;8. doi: 10.3390/cells8070741
34. Dorsch LM, Schuldt M, Knezevic D, Wiersma M, Kuster DWD, van der Velden J and Brundel B. Untying the knot: protein quality control in inherited cardiomyopathies. *Pflugers Arch*. 2018. doi: 10.1007/s00424-018-2194-0
35. Sanbe A, Daicho T, Mizutani R, Endo T, Miyauchi N, Yamauchi J, Tanonaka K, Glabe C and Tanoue A. Protective effect of geranylgeranylacetone via enhancement of HSPB8 induction in desmin-related cardiomyopathy. *PLoS One*. 2009;4:e5351. doi: 10.1371/journal.pone.0005351
36. Bhuiyan MS, Pattison JS, Osinska H, James J, Gulick J, McLendon PM, Hill JA, Sadoshima J and Robbins J. Enhanced autophagy ameliorates cardiac proteinopathy. *J Clin Invest*. 2013;123:5284-97. doi: 10.1172/JCI70877
37. Singh SR, Zech ATL, Geertz B, Reischmann-Dusener S, Osinska H, Prondzynski M, Kramer E, Meng Q, Redwood C, van der Velden J, Robbins J, Schlossarek S and Carrier L. Activation of Autophagy Ameliorates Cardiomyopathy in Mybpc3-Targeted Knockin Mice. *Circ Heart Fail*. 2017;10. doi: 10.1161/CIRCHEARTFAILURE.117.004140
38. Brundel BJ, Shiroshita-Takeshita A, Qi X, Yeh YH, Chartier D, van Gelder IC, Henning RH, Kampinga HH and Nattel S. Induction of heat shock response protects the heart against atrial fibrillation. *Circ Res*. 2006;99:1394-402. doi: 10.1161/01.RES.0000252323.83137.fe
39. Marunouchi T, Inomata S, Sanbe A, Takagi N and Tanonaka K. Protective effect of geranylgeranylacetone via enhanced induction of HSPB1 and HSPB8 in mitochondria of the failing heart following myocardial infarction in rats. *Eur J Pharmacol*. 2014;730:140-7. doi: 10.1016/j.ejphar.2014.02.037
40. Grimes KM, Prasad V and McNamara JW. Supporting the heart: Functions of the cardiomyocyte's non-sarcomeric cytoskeleton. *J Mol Cell Cardiol*. 2019;131:187-196. doi: 10.1016/j.yjmcc.2019.04.002

41. Lewis YE, Moskovitz A, Mutlak M, Heineke J, Caspi LH and Kehat I. Localization of transcripts, translation, and degradation for spatiotemporal sarcomere maintenance. *J Mol Cell Cardiol*. 2018;116:16-28. doi: 10.1016/j.yjmcc.2018.01.012
42. Scholz D, Baicu CF, Tuxworth WJ, Xu L, Kasiganesan H, Menick DR and Cooper Gt. Microtubule-dependent distribution of mRNA in adult cardiocytes. *Am J Physiol Heart Circ Physiol*. 2008;294:H1135-44. doi: 10.1152/ajpheart.01275.2007
43. Thottakara T, Friedrich FW, Reischmann S, Braumann S, Schlossarek S, Kramer E, Juhr D, Schluter H, van der Velden J, Munch J, Patten M, Eschenhagen T, Moog-Lutz C and Carrier L. The E3 ubiquitin ligase Asb2beta is downregulated in a mouse model of hypertrophic cardiomyopathy and targets desmin for proteasomal degradation. *J Mol Cell Cardiol*. 2015;87:214-24. doi: 10.1016/j.yjmcc.2015.08.020
44. Robison P and Prosser BL. Microtubule mechanics in the working myocyte. *J Physiol*. 2017;595:3931-3937. doi: 10.1113/JP273046
45. Zile MR, Koide M, Sato H, Ishiguro Y, Conrad CH, Buckley JM, Morgan JP and Cooper Gt. Role of microtubules in the contractile dysfunction of hypertrophied myocardium. *J Am Coll Cardiol*. 1999;33:250-60. doi:
46. Infante AS, Stein MS, Zhai Y, Borisy GG and Gundersen GG. Detyrosinated (Glu) microtubules are stabilized by an ATP-sensitive plus-end cap. *J Cell Sci*. 2000;113 (Pt 22):3907-19. doi:
47. Aillaud C, Bosc C, Peris L, Bosson A, Heemeryck P, Van Dijk J, Le Friec J, Boulan B, Vossier F, Sanman LE, Syed S, Amara N, Coute Y, Lafanechere L, Denarier E, Delphin C, Pelletier L, Humbert S, Bogyo M, Andrieux A, Rogowski K and Moutin MJ. Vasohibins/SVBP are tubulin carboxypeptidases (TCPs) that regulate neuron differentiation. *Science*. 2017;358:1448-1453. doi: 10.1126/science.aao4165
48. Nieuwenhuis J, Adamopoulos A, Bleijerveld OB, Mazouzi A, Stickel E, Celie P, Altelaar M, Knipscheer P, Perrakis A, Blomen VA and Brummelkamp TR. Vasohibins encode tubulin detyrosinating activity. *Science*. 2017;358:1453-1456. doi: 10.1126/science.aao5676
49. Kerr JP, Robison P, Shi G, Bogush AI, Kempema AM, Hexum JK, Becerra N, Harki DA, Martin SS, Raiteri R, Prosser BL and Ward CW. Detyrosinated microtubules modulate mechanotransduction in heart and skeletal muscle. *Nat Commun*. 2015;6:8526. doi: 10.1038/ncomms9526
50. Swiatlowska P, Sanchez-Alonso JL, Wright PT, Novak P and Gorelik J. Microtubules regulate cardiomyocyte transversal Young's modulus. *Proc Natl Acad Sci U S A*. 2020;117:2764-2766. doi: 10.1073/pnas.1917171117

Chapter 7

Integrative Functional Annotation of 52 Genetic Loci Influencing Myocardial Mass Identifies Candidate Regulatory Variants and Target Genes

Daiane Hemerich[1,2], Jiayi Pei[3], Magdalena Harakalova[1], Jessica van Setten[1], Sander Boymans[4], Bas J. Boukens[5], Igor R. Efimov[6], Michelle Michels[15], Jolanda van der Velden[7], Aryan Vink[8], Caroline Cheng[3], Pim van der Harst[9], Jason H. Moore[10], Michal Mokry[11], Vinicius Tragante[1], and Folkert W. Asselbergs[1,12,14]

1Dept of Cardiology, UMC Utrecht, Utrecht University, Utrecht, NL
2CAPES Foundation, Ministry of Education of Brazil, Brasília, Brazil
3Dept of Nephrology & Hypertension, Cancer Genomics Netherlands, Dept of Genetics, UMC Utrecht, Utrecht
4Center for Molecular Medicine, Cancer Genomics Netherlands, Dept of Genetics, UMC Utrecht, Utrecht
5Dept of Medical Biology, Academic Medical Center, Amsterdam, NL
6Dept of Biomedical Engineering, The George Washington University, Washington, DC
7Dept of Physiology, Amsterdam Cardiovascular Sciences, Amsterdam University Medical Center, Amsterdam
8Dept of Pathology, UMC Utrecht, Utrecht University, Utrecht, NL
9Dept of Cardiology, UMC Groningen, NL
10Dept of Biostatistics & Epidemiology, Institute for Biomedical Informatics, University of Pennsylvania, PA
11Dept of Pediatrics, Wilhelmina Children's Hospital, UMC Utrecht, Utrecht University, Utrecht, NL
12Durrer Center for Cardiogenetic Research, ICIN-Netherlands Heart Institute, Utrecht, NL
13Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College London, London, UK
14Health Data Research UK London & Institute of Health Informatics, University College London, London, UK
15Dept of Cardiology, Erasmus MC, Rotterdam, NL

Circulation: Genomic and Precision Medicine
DOI: 10.1161/CIRCGEN.118.002328

Abstract

Background: Regulatory elements may be involved in the mechanisms by which 52 loci influence myocardial mass, reflected by abnormal amplitude and duration of the QRS complex on the electrocardiogram (ECG). Functional annotation thus far did not take into account how these elements are affected in disease context.

Methods: We generated maps of regulatory elements on hypertrophic cardiomyopathy (HCM) patients (ChIP-seq N=14, RNA-seq N=11) and non-diseased hearts (ChIP-seq N=4, RNA-seq N=11). We tested enrichment of QRS-associated loci on elements differentially acetylated and/or directly regulating differentially expressed genes between HCM patients and controls. We further performed functional annotation on QRS-associated loci using these maps of differentially active regulatory elements.

Results: Regions differentially affected in disease showed a stronger enrichment ($p=8.6\times10^{-5}$) for QRS-associated variants than those not showing differential activity (p=0.01). Promoters of genes differentially regulated between HCM patients and controls showed more enrichment (p=0.001) than differentially acetylated enhancers (p=0.8) and super-enhancers (p=0.025). We also identified 74 potential causal variants overlapping these differential regulatory elements. Eighteen of the genes mapped confirmed previous findings, now also pinpointing the potentially affected regulatory elements and candidate causal variants. Fourteen new genes were also mapped.

Conclusions: Our results suggest differentially active regulatory elements between HCM patients and controls can offer more insights into the mechanisms of QRS-associated loci than elements not affected by disease.

Keywords: ECG; heart failure; epigenetics; gene regulation; genetics; bioinformatics; functional annotation; OMICS data integration

Introduction

The QRS complex on the electrocardiogram (ECG) represents cardiac depolarization and conduction of the electrical signal through the ventricular muscle. Duration and amplitude of the QRS complex is used as a proxy for left ventricular mass[1, 2]. Abnormalities of the QRS complex are associated with an increased risk of cardiovascular (CV) mortality and morbidity[3, 4].

A recent large-scale genome-wide association study (GWAS) meta-analysis of four correlated and clinically used QRS traits (Sokolow-Lyon, Cornell, 12-lead-voltage duration products (12-leadsum), and QRS duration) identified 52 independent loci at $p<1\times10^{-8}$ [5]. However, the identification of causal variants, their target genes and disturbed mechanisms remain an important challenge. In a given GWAS locus, the SNP with the most significant association with the disease (lowest p-value) is usually reported as the 'lead' SNP. This lead SNP is not necessarily the causal variant, and a SNP in high linkage disequilibrium (LD) with the lead SNP may be the causal one[6]. Only three of the 52 SNPs reported in the original study are located in coding regions of the genome[5]. A non-synonymous variant in high LD with a high deleteriousness score, such as CADD score > 12.37[7], can also be considered candidate for causality, and that is the case of 14 variants in nine QRS-associated loci (Supplemental Table 1). Potential mechanisms for causality in the remaining loci include disturbance of regulatory elements, genomic regions that play a crucial role in transcriptional regulation. Indeed, increasing evidence shows multiple GWAS variants regulate transcription[8, 9]. Non-coding variants often affect gene expression in a cell-type specific manner by altering the function of enhancer and promoter elements[10, 11]. Results from in silico analyses have suggested strong enrichment of QRS-associated variants in specific chromatin states associated with active enhancers, promoters, and transcription in the human heart[12]. By contrast, no enrichment was observed for transcriptionally repressive histone marks[5, 13]. Although using relevant cell-types, these previous studies did not include regulatory information obtained from tissue in disease state. Highlighting possible differences in transcription from control to diseased myocardial tissue may help understand the underlying mechanisms of changes in left ventricular mass. To gain insight into these mechanisms, we integrate genome, regulome and transcriptome information to test whether regions differentially affected in disease can be more informative than the regulatory landscape of non-diseased tissue. We use regulatory information obtained from non-diseased myocardium and diseased tissue from patients with hypertrophic cardiomyopathy (HCM). We also co-localize candidate causative QRS-associated variants and regulatory regions that show differential activity (proxied by H3K27ac levels or by gene expression) in HCM tissues compared to controls, in order to identify regulatory elements potentially altered by the

variation. Taken together, our results may enhance our understanding on the regulatory mechanisms underlying increased myocardial mass.

Methods

The procedures for obtaining human samples were approved by the scientific advisory board of the biobank of the University Medical Center Utrecht (protocol number 12/387), the Washington University School of Medicine Ethics Committee (Institutional Review Board) and the local ethics committee of the Erasmus MC, and written consent was obtained. Biopsies on HCM patients are septal myomectomy specimens. Control samples were obtained from donor hearts not used for transplantation. Mutation and clinical data of patients are given in Table 1.

Because of the sensitive nature of the data collected for this study, requests to access the dataset from qualified researchers trained in human subject confidentiality protocols may be sent to the corresponding author at F.W.Asselbergs@umcutrecht.nl. The methods of our study are detailed in the supplemental data.

Results

In this study, we functionally annotated and fine-mapped 52 QRS-associated loci. We used regulatory information (promoters from RNA-seq experiments, H3K27ac regions highlighting active promoters and enhancers, and super-enhancers from ChIP-seq experiments) obtained from non-diseased myocardium and diseased tissue from HCM patients (Table 1, Supplemental Table 2). Categorizing as expressed genes with average RPKM > 0.5 across samples of each group, 12,136 genes were expressed on HCM group and 12,540 on control group (Supplemental Tables 2 and 3). Differential expression analysis between the 11 HCM patients and 11 controls resulted in 1557 up-regulated and 1202 down-regulated genes (Supplemental Table 4). Differential acetylation analysis between 14 HCM patients and 4 controls resulted in 4068 up-acetylated and 2983 down-acetylated regions (Supplemental Table 5). From the total set of super-enhancers identified, 1048 were unique to the HCM group, not overlapping with the control group (Supplemental Table 6). Promoter, enhancer and super-enhancer regions were further narrowed down to those overlapping regions of open chromatin, in order to retain only sequences accessible to TF binding. We used these regulatory features to test the enrichment of QRS-associated variants and vicinity in LD. In

order to identify an optimal LD cutoff, we identified variants in LD with the 52 QRS-associated SNPs, and divided them into bins, as described in the Methods section. We overlapped each bin with regulatory elements differentially regulated/acetylated in HCM, and performed 10k permutation tests to calculate enrichment (Supplemental Figure 1). Although LD thresholds showed fluctuations on enrichment, we observed a continuous decrease in enrichment as the LD threshold becomes more lenient. We defined LD r^2>0.5 as cutoff to expand the set of candidate causal QRS-associated variants, from 52 lead SNPs to 4620 SNPs. We co-localized these candidate causal QRS-associated variants and regulatory regions that show differential activity in disease, in order to identify regulatory elements potentially altered by the variation. We also investigated which set of regulatory features showed more enrichment for QRS-associated variants, and thus has the potential to be more informative for fine-mapping efforts.

Disease-affected regulatory regions show more enrichment for QRS-associated variants than all tissue-specific regulatory regions identified in control and disease patients

We performed enrichment tests aiming to identify which set of regulatory regions is more informative for fine-mapping efforts. Regulatory features differentially regulated/acetylated in HCM were more enriched ($p=8.6\times10^{-5}$) than those not showing differential activity (p=0.01), suggesting these can offer more insight into mechanisms altered by genetic variation than general tissue-specific regulatory elements (Figure 1A).

Promoters of differentially expressed genes show more enrichment than regions highlighted by differential acetylation

Given the strong enrichment of QRS-associated variants in differentially regulated promoters and differentially acetylated regulatory elements in HCM (Figure 1A), we overlapped candidate causative SNPs and enriched regulatory features to identify variants that might be altering the function of these elements. Of the 4620 candidate QRS-associated SNPs, 74 co-localized with differential regulatory features (Figures 1 and 2, Supplemental Table 1), more than expected by chance (p=8.6e-5). These variants show more enrichment in differentially expressed promoters from RNA-seq experiments (p=0.001) than differentially acetylated regulatory elements highlighted by H3K27 from ChIP-seq experiments (enhancers p=0.8, super-enhancers p=0.02) (Figure 1B). These results highlight the potential of integrating regulatory elements that show differential behavior in disease, especially promoter regions.

Table 1. Patient characteristics.

Sample name	ChIP-seq	RNA-seq	Sex	Cardiac region	Age (at operation)	Mutation	Type of mutation	BMI	Ejection Fraction (%, via Echo)	Cause of death
HCM_1	x	x	M	Septum	32	MYBPC3	truncation			
HCM_2	x		M	Septum	60	MYBPC3	truncation			
HCM_3	x	x	M	Septum	17	MYBPC3	truncation			
HCM_4	x		M	Septum	26	MYBPC3	truncation			
HCM_5	x	x	M	Septum	33	MYBPC3	truncation			
HCM_6	x	x	M	Septum	27	MYBPC3	truncation			
HCM_7	x	x	F	Septum	24	MYBPC3	truncation			
HCM_8	x	x	M	Septum	33	MYBPC3	truncation			
HCM_9	x	x	M	Septum	49	MYBPC3	truncation			
HCM_10	x	x	F	Septum	21	MYBPC3	truncation			
HCM_11	x	x	F	Septum	53	MYBPC3	truncation			
HCM_12	x	x	M	Septum	53	MYBPC3	truncation			
HCM_13	x		M	Septum	48	MYBPC3	truncation			
HCM_14	x	x	M	Septum	22	MYBPC3	truncation			
CONTROL_ChIP_1	x		F	Septum	NA			NA	NA	
CONTROL_ChIP_2	x		M	LV	NA			NA	NA	
CONTROL_ChIP_3	x		M	Septum	NA			NA	NA	
CONTROL_ChIP_4	x		F	Septum	NA			NA	NA	
CONTROL_RNA_1		x	M	RV	46			32.9	60-65	Intracranial hemorrhage/Stroke
CONTROL_RNA_2		x	M	LV	46			32.9	60-65	Intracranial hemorrhage/Stroke
CONTROL_RNA_3		x	M	RV	50			27.4	65	Intracranial hemorrhage/Stroke
CONTROL_RNA_4		x	F	RV	56			31.4	60-65	Intracranial hemorrhage/Stroke
CONTROL_RNA_5		x	F	LV	56			31.4	60-65	Intracranial hemorrhage/Stroke
CONTROL_RNA_6		x	F	LV	59			22.8	45-50	Head trauma
CONTROL_RNA_7		x	M	RV	46			25.4	55	Head trauma

CONTROL_RNA_8	x	M	LV	46		25.4	55	Head trauma
CONTROL_RNA_9	x	M	RV	NA		NA	NA	NA
CONTROL_RNA_10	x	F	RV	61		40.9	NA	Cerebrovascular/stroke
CONTROL_RNA_11	x	F	LV	61		40.9	NA	Cerebrovascular/stroke

Fine-mapping pinpoints potential causal variants and candidate genes

The 74 variants overlapping regulatory elements affected in HCM are spread through 20 QRS-associated loci (Supplemental Figures 2–21). We investigated which genes are the potential targets of the 74 QRS-associated variants. We retrieved the three genes nearest to each of the 74 variants. We found 13 of the nearest genes are down-regulated and 12 are up-regulated in HCM versus controls (Supplemental Table 1). eQTL mapping with LV tissue from GTEx portal[14] confirmed the involvement of these differentially regulated genes with QRS-associated candidate regulatory variants in three loci (ACP2/MADD, FADS1/FADS2, PROCR/EDEM2). Eighteen of the 68 candidate genes that were identified by the original study were confirmed with this new approach[5]. The functional annotation performed in this study pinpoints the likely regulatory elements affected by the variation that in turn affect expression of these genes, as well as candidate causative SNPs (Supplemental Tables 1 and 7). The function of the new fourteen candidate genes potentially involved in myocardial mass is described in Table 2. Visualization and further description of each of the 20 QRS-associated loci overlapping regulatory regions of interest can be found on the Supplemental Material.

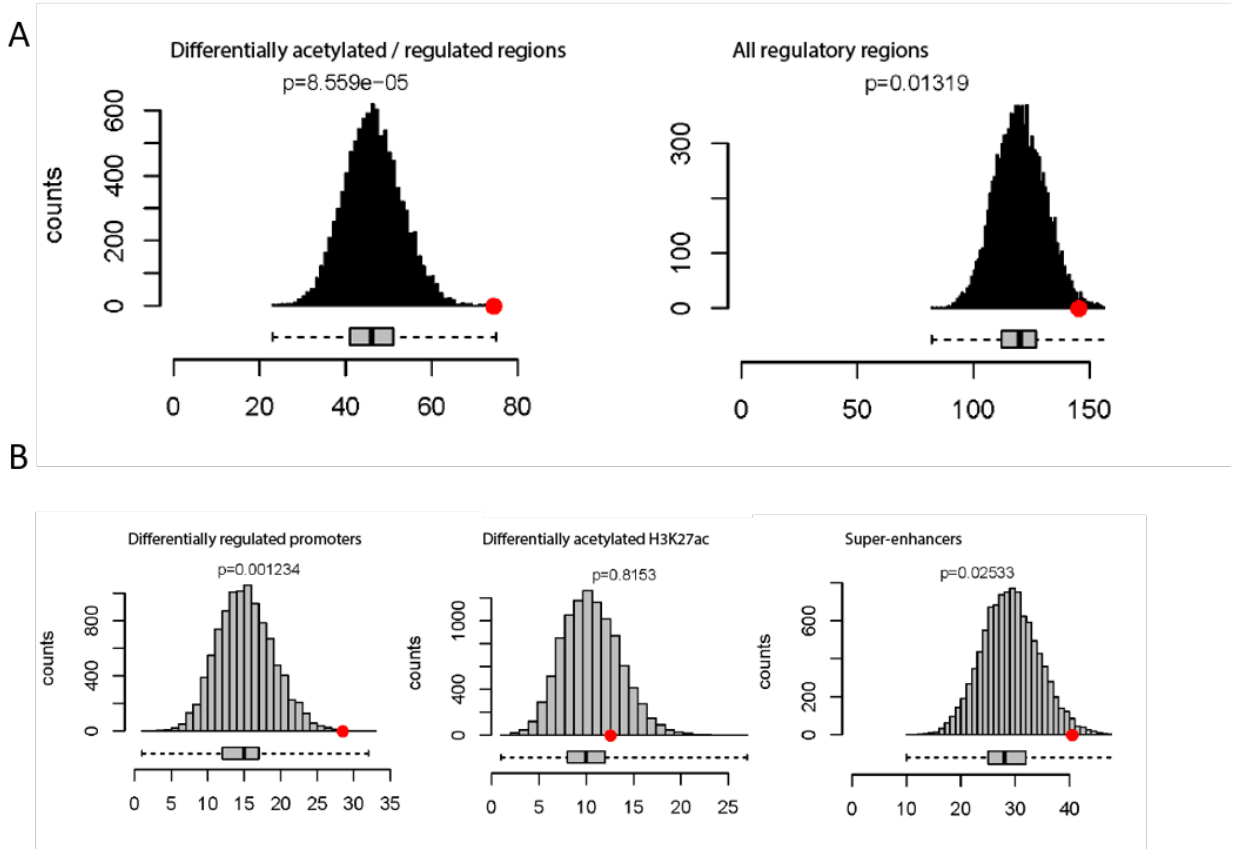


Figure 1. A) Mean number of regulatory elements overlapping with QRS-associated loci (red circle) compared with 10,000 matched control sets (gray bars). Differentially acetylated / regulated elements in HCM show more

enrichment (p=8.6e-5) for QRS-associated candidate causal variants than those not showing differential activity (p=0.01). B) Promoters of genes differentially regulated between HCM patients and controls showed more enrichment (p=0.001) for QRS-associated variants than differentially acetylated enhancers (p=0.8) and super-enhancers (p=0.025).

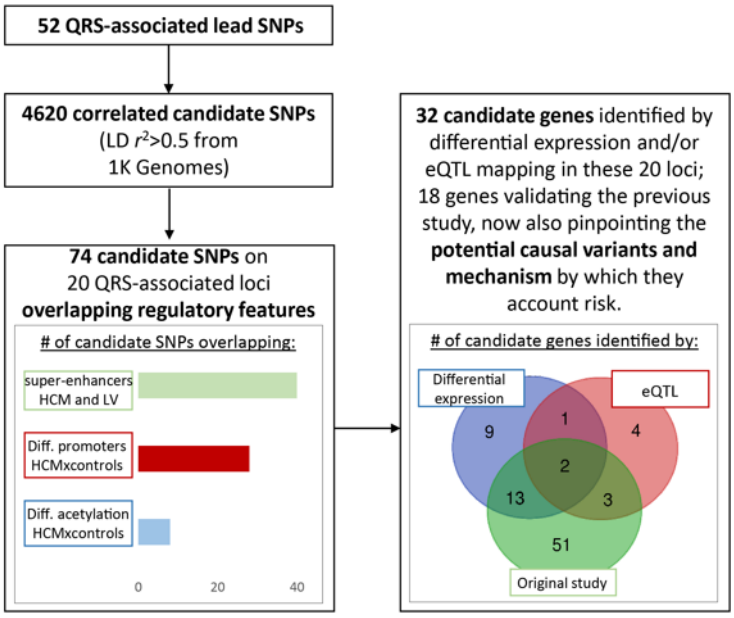


Figure 2. Overview of functional annotation pipeline. We expanded the search from 52 lead SNPs to include variants in LD (r2>0.5), leading to 4620 candidate variants. Of these, 74 overlapped one or more regulatory feature of interest: differentially acetylated (H3K27ac) regions between HCM and control patients, differentially expressed promoter regions between HCM and control patients, and super-enhancers specific to HCM or LV. In the 20 loci where these overlaps were observed, differential expression analysis on HCM patients and controls and or/eQTL mapping in LV identified 32 candidate genes. Eighteen of these validate the findings of the previous study 5, now also pinpointing causal variants and the potential mechanisms by which they account risk. In addition, 14 new candidate genes were identified, some with cardiac function and previously implicated in other studies on ECG traits (Table 2, Supplemental Tables 1 and 7).

Table 2. Description of new candidate genes identified mapped to QRS-associated loci overlapping regulatory regions with abnormal expression / acetylation between HCM and control patients.

LIPH - Lipase H	Lipase H is a phosphatidic acid-selective phospholipase A1 (PLA1) that produces 2-acyl lysophosphatidic acid (LPA). LPA is a lipid mediator with diverse biologic properties, including induction of platelet aggregation, smooth muscle contraction, and stimulation of cell proliferation[15].

CPNE5 - Copine 5	This gene is one of several genes that encode a calcium-dependent protein containing two N-terminal type II C2 domains and an integrin A domain-like sequence in the C-terminus. Has been implicated in previous GWAS in heart rate[16].
DPY19L1 - Dpy-19 Like C-Mannosyltransferase 1	Function unknown. Automated annotations supported by experiments on knockout mouse models associate this gene to abnormal heart morphology and enlarged heart[17].
TTC39A - Tetratricopeptide repeat domain 39A	Function unknown. Automated annotations supported by experiments on knockout mouse models associate this gene to increased heart weight and increased circulating phosphate level[18].
PROCR - Endothelial protein C receptor	The protein encoded by this gene is a receptor for activated protein C, a serine protease activated by and involved in the blood coagulation pathway. The encoded protein is an N-glycosylated type I membrane protein that enhances the activation of protein C. Mutations in this gene have been associated with venous thromboembolism, myocardial infarction and coronary artery disease[19, 20].
PRKCD - Protein kinase C delta	Protein kinase C (PKC) is a family of serine- and threonine-specific protein kinases that can be activated by calcium and the second messenger diacylglycerol. Studies both in human and mice demonstrate that PRKCD kinase is involved in B cell signaling and in the regulation of growth, apoptosis, and differentiation of a variety of cell types[21].
CORO6 - Coronin 6	Involved in actin filament binding. Overexpressed in heart and skeletal muscle[22]. Has been implicated in previous GWAS in coronary artery disease[19].
TCEA3 - Transcription Elongation Factor A3	This gene is associated to other measure of amplitude of the electrocardiogram, ST-T wave[23] and QT- interval[24]. Overexpressed in heart[22].
FUT11 - Fucosyltransferase 11	Function unknown.
CENPA - Centromere protein A	Histone H3-like variant which exclusively replaces conventional H3 in the nucleosome core of centromeric chromatin at the inner plate of the kinetochore[25]. Required for recruitment and assembly of kinetochore proteins, mitotic progression and chromosome segregation. May serve as an epigenetic mark that propagates centromere identity through replication and cell division[26, 25, 27, 28].
AGBL5 - ATP/GTP binding protein- like 5	Function unknown.

KCNK3 - Potassium channel, subfamily K, member 3	PH-dependent, voltage-insensitive, background potassium channel protein. Rectification direction results from potassium ion concentration on either side of the membrane. Acts as an outward rectifier when external potassium concentration is low[29].
KHK - ketohexokinase	Catalyzes the first step of metabolism of dietary fructose, conversion of fructose to fructose-1- phosphate. It has been shown that myocardial hypoxia influences fructose metabolism in human and mouse models of pathologic cardiac hypertrophy through hypoxia-inducible factor 1-alpha activation of *SF3B1* and *SF3B1*-mediated splice switching of *KHK-A* to *KHK-C*[30, 31]. In mice, heart-specific depletion of *SF3B1* or genetic ablation of *KHK* suppressed pathologic stress-induced fructose metabolism, growth, and contractile dysfunction[31].
FADS1 - Fatty Acid Desaturase 1	A member of the fatty acid desaturase (FADS) gene family. Component of a lipid metabolic pathway that catalyzes the biosynthesis of highly unsaturated fatty acids from precursor essential polyunsaturated fatty acids, linoleic acid, and alpha-linolenic acid. SNPs of the *FADS* gene cluster have been associated with polyunsaturated fatty acids in a cohort of patients with cardiovascular disease[32].

Discussion

We performed functional annotation of 52 loci influencing myocardial mass, considering regulatory elements with differential activity between diseased (HCM) heart samples and control donors may act as mediators between genetic variation and gene expression. The 52 QRS-associated variants showed enrichment on the set of differential regulatory elements in comparison to non-differential regulatory elements. In addition, differentially expressed promoters identified by RNA-seq in HCM and controls showed more enrichment of QRS-associated variants, than the other differential regulatory elements analyzed. These results can aid the study design of future fine-mapping studies.

Our functional annotation showed twenty loci showed overlapping with the enriched differential regulatory regions. We identified the variants in LD with the lead QRS-associated SNP overlapping regulatory features, thus pinpointing the potential causative variants and mechanisms through which these variants account risk for disease. We observed that some loci showed enrichment of more than one candidate SNP in the same regulatory feature. Studies indeed show that several variants can be causal

inside a disease-associated locus[33, 34]. In addition, more than one gene can be causal inside the same loci[35], and we indeed identified more than one candidate gene per locus in some cases. Further assays, such as chromatin conformation capture[36–40], can add evidence on the interaction between QRS-candidate regulatory SNPs and each candidate gene, preferably also taking into account the disease context[41]. These experiments can also reveal whether regions identified as promoters also function as enhancers[42–44]. Moreover, genes potentially influenced by the regulatory elements affected by QRS-associated variants that are not in the immediate vicinity can also be detected[45].

We observed that 18 genes mapped by the previous study were confirmed by differential expression and/or eQTL mapping, adding evidence for the role of these genes in natural variance in myocardial mass. We also identified 14 genes not previously mapped to the QRS-associated variants, some with known cardiac function and previously implicated in studies on ECG parameters and/or heart phenotypes. Definitive assignment of function(s) to putative cis-regulatory elements requires perturbation of these elements[46]. A promising approach is deletion and manipulation of nucleases through CRISPR/Cas9, which was successful in recent studies[47, 10].

Future efforts can unravel the mechanisms of the remaining QRS-associated loci by also taking into account other regulatory elements apart from enhancers and promoters, such as silencers, highlighted by different histone modifications. In addition, other histone modifications can also help identify additional enhancer elements. A subset of active enhancers has been identified in mouse embryonic stem cells that lack H3K27ac but are marked by H3K122ac and/or H3K64ac, H3K4me1 and enhancer RNAs (eRNAs)[48]. This evidence shows that although H3K27ac highlights enhancer activity, enhancers that lack H3K27ac are not necessarily inactive, and other histone marks can carry out the activating role of H3K27ac[49]. Moreover, as chromatin states are spatially and temporally dynamic along disease progression, regulatory information is needed in more different conditions including different stages of disease progression. This will allow a more precise and complete fine-mapping and functional annotation. It can also help answer whether QRS-associated SNPs promote aberrant function of regulatory elements, or the SNPs alter the activity of regulatory elements following disease establishment. Finally, the low number of control heart samples included in our ChIP-seq experiments might partly explain the lower degree of enrichment for enhancer/super-enhancer regions as compared to the promoter regions identified from RNA-seq experiments. Increasing sample sizes may help unravel more regulatory regions and their enrichment for QRS-associated variants.

Taken together, our results show the importance of using differential regulatory elements in disease tissue on fine-mapping studies, since influence from cell conditions are crucial for epigenetic modifications, chromatin accessibility, TF binding and consequently gene regulation[50]. We also expand the view of the epigenetic and regulatory aspects involved in myocardial mass by pinpointing potential causal variants, mechanisms potentially affected by genetic variation, and candidate genes. Validation of the identified candidate causal variants through methods such as genome-editing technologies will broaden our knowledge on how QRS-associated variants contribute to disease.

Acknowledgments

This work was supported by the National Institute of health (NIH) grant LM010098. DH is supported by the National Council for the Improvement of Higher Education (CAPES) and Science without Borders Project, process nº 13259/13–0. M.H. is supported by Wilhelmina Children’s Hospital research funding OZF/14 and by personal research fund VENI ZonMW/NWO2016‐016.176.136. FWA is supported by UCL Hospitals NIHR Biomedical Research Centre. We acknowledge support from the Netherlands Cardiovascular Research Initiative—an initiative with support of the Dutch Heart Foundation, CVON: The Netherlands CardioVascular Research Committee, CVON2014–40 DOSIS. We gratefully acknowledge the contribution of Noortje van den Dungen and Nico Lansu from the Department of Paediatrics, UMCU, as well as Joyce van Kuik and Erica Sierra-de Koning from the Department of Pathology, UMCU, and the Utrecht Sequencing Facility. We also acknowledge Gregory Crawford Lab from Duke Center For Genomic And Computational Biology and the ENCODE consortium for the public dataset of open chromatin in heart used in this study.

Source of Funding

This work was supported by the National Institute of health (NIH) grant LM010098 (United States of America). DH is supported by the National Council for the Improvement of Higher Education (CAPES) and Science without Borders Project, process nº 13259/13‐0 (Brazil). M.H. is supported by Wilhelmina Children’s Hospital research funding OZF/14 and by personal research fund VENI ZonMW/NWO2016‐016.176.136 (The Netherlands). FWA is supported by UCL Hospitals NIHR Biomedical Research Centre. UCL Hospitals NIHR Biomedical Research Centre (United Kingdom). We acknowledge support from the

Netherlands Cardiovascular Research Initiative—an initiative with support of the Dutch Heart Foundation, CVON: The Netherlands CardioVascular Research Committee, CVON2014–40 DOSIS (The Netherlands).

Supplementary Files – Online

Supplementary Material. Extended Methods and sequencing data.

References

1. Levy D, et al. Determinants of sensitivity and specificity of electrocardiographic criteria for left ventricular hypertrophy. Circulation. 1990;81:815–820. [PubMed: 2137733]

2. Okin PM, et al. Time-voltage QRS area of the 12-lead electrocardiogram: detection of left ventricular hypertrophy. Hypertension. 1998;31:937–942. [PubMed: 9535418]

3. Kannel WB, et al. Left ventricular hypertrophy by electrocardiogram. Prevalence, incidence, and mortality in the Framingham study. Ann Intern Med. 1969;71:89–105. [PubMed: 4239887]

4. Usoro AO, et al. Risk of mortality in individuals with low QRS voltage and free of cardiovascular disease. Am J Cardiol. 2014;113:1514–1517. [PubMed: 24630386]

5. van der Harst P, et al. 52 Genetic Loci Influencing Myocardial Mass. J Am Coll Cardiol. 2016;68:1435–1448. [PubMed: 27659466]

6. Corradin O, et al. Enhancer variants: evaluating functions in common disease. Genome Med. 2014;6:85. [PubMed: 25473424]

7. Kircher M, et al. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014;46:310–315. [PubMed: 24487276]

8. Maurano MT, et al. Systematic localization of common disease-associated variation in regulatory DNA. Science. 2012;337:1190–1195. [PubMed: 22955828]

9. Schaub MA, et al. Linking disease associations with regulatory information in the human genome. Genome Res. 2012;22:1748–1759. [PubMed: 22955986]

10. Gupta RM, et al. A Genetic Variant Associated with Five Vascular Diseases Is a Distal Regulator of Endothelin-1 Gene Expression. Cell. 2017;170:522–533 e515. [PubMed: 28753427]

11. Yao L, et al. Functional annotation of colon cancer risk SNPs. Nat Commun. 2014;5:5114. [PubMed: 25268989]

12. Roadmap Epigenomics C, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015;518:317–330. [PubMed: 25693563]

13. Parmar PG, et al. International Genome-Wide Association Study Consortium Identifies Novel Loci Associated With Blood Pressure in Children and Adolescents. Circ Cardiovasc Genet. 2016;9:266– 278. [PubMed: 26969751]

14. Consortium GT, et al. Genetic effects on gene expression across human tissues. Nature. 2017;550:204–213. [PubMed: 29022597]

15. Sonoda H, et al. A novel phosphatidic acid-selective phospholipase A1 that produces lysophosphatidic acid. J Biol Chem. 2002;277:34254–34263. [PubMed: 12063250]

16. den Hoed M, et al. Identification of heart rate-associated loci and their effects on cardiac conduction and rhythm disorders. Nat Genet. 2013;45:621–631. [PubMed: 23583979]

17. “Gene: Dpy19l1.”. Retrieved 6 11, 2018, from http://www.mousephenotype.org/data/genes/MGI:1915685.

18. “Gene: Ttc39a.” Retrieved 6 11, 2018, from http://www.mousephenotype.org/data/genes/MGI:2444350.

19. van der Harst P, et al. Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. Circ Res. 2018;122:433–443. [PubMed: 29212778]

20. Dennis J, et al. The endothelial protein C receptor (PROCR) Ser219Gly variant and risk of common thrombotic disorders: a HuGE review and meta-analysis of evidence from observational studies. Blood. 2012;119:2392–2400. [PubMed: 22251481]

21. O’Leary NA, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 2016;44:D733–745. [PubMed: 26553804]

22. Fagerberg L, et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. Mol Cell Proteomics. 2014;13:397–406. [PubMed: 24309898]

23. Verweij N, et al. Twenty-eight genetic loci associated with ST-T-wave amplitudes of the electrocardiogram. Hum Mol Genet. 2016;25:2093–2103. [PubMed: 26962151]

24. Arking DE, et al. Genetic association study of QT interval highlights role for calcium signaling pathways in myocardial repolarization. Nat Genet. 2014;46:826–836. [PubMed: 24952745]

25. Orthaus S, et al. Assembly of the inner kinetochore proteins CENP-A and CENP-B in living human cells. Chembiochem. 2008;9:77–92. [PubMed: 18072184]

26. Shelby RD, et al. Assembly of CENP-A into centromeric chromatin requires a cooperative array of nucleosomal DNA contact sites. J Cell Biol. 1997;136:501–513. [PubMed: 9024683]

27. Sullivan BA, et al. Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. Nat Struct Mol Biol. 2004;11:1076–1083. [PubMed: 15475964]

28. Black BE, et al. Structural determinants for generating centromeric chromatin. Nature. 2004;430:578–582. [PubMed: 15282608]

29. Plant LD, et al. SUMOylation silences heterodimeric TASK potassium channels containing K2P1 subunits in cerebellar granule neurons. Sci Signal. 2012;5:ra84. [PubMed: 23169818]

30. Mirtschink P, et al. Hypoxia-driven glycolytic and fructolytic metabolic programs: Pivotal to hypertrophic heart disease. Biochim Biophys Acta. 2016;1863:1822–1828. [PubMed: 26896647]

31. Mirtschink P, et al. HIF-driven SF3B1 induces KHK-C to enforce fructolysis and heart disease. Nature. 2015;522:444–449. [PubMed: 26083752]

32. Malerba G, et al. SNPs of the FADS gene cluster are associated with polyunsaturated fatty acids in a cohort of patients with cardiovascular disease. Lipids. 2008;43:289–299. [PubMed: 18320251]

33. Trynka G, et al. Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. Nat Genet. 2011;43:1193–1201. [PubMed: 22057235]

34. Stahl EA, et al. Bayesian inference analyses of the polygenic architecture of rheumatoid arthritis. Nat Genet. 2012;44:483–489. [PubMed: 22446960]

35. Flister MJ, et al. Identifying multiple causative genes at a single GWAS locus. Genome Res. 2013;23:1996–2002. [PubMed: 24006081]

36. Dekker J, et al. Capturing chromosome conformation. Science. 2002;295:1306–1311. [PubMed: 11847345]

37. Simonis M, et al. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat Genet. 2006;38:1348–1354. [PubMed: 17033623]

38. Zhao Z, et al. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nat Genet. 2006;38:1341–1347. [PubMed: 17033624]

39. Belton JM, et al. Hi-C: a comprehensive technique to capture the conformation of genomes. Methods. 2012;58:268–276. [PubMed: 22652625]

40. Javierre BM, et al. Lineage-Specific Genome Architecture Links Enhancers and Non-coding Disease Variants to Target Gene Promoters. Cell. 2016;167:1369–1384 e1319. [PubMed: 27863249]

41. Rosa-Garrido M, et al. High-Resolution Mapping of Chromatin Conformation in Cardiac Myocytes Reveals Structural Remodeling of the Epigenome in Heart Failure. Circulation. 2017;136:1613–1625. [PubMed: 28802249]

42. Catarino RR, et al. Promoting transcription over long distances. Nat Genet. 2017;49:972–973. [PubMed: 28656981]

43. Dao LTM, et al. Genome-wide characterization of mammalian promoters with distal enhancer functions. Nat Genet. 2017;49:1073–1081. [PubMed: 28581502]

44. Diao Y, et al. A tiling-deletion-based genetic screen for cis-regulatory element identification in mammalian cells. Nat Methods. 2017;14:629–635. [PubMed: 28417999]

45. Sandhu KS, et al. Large-scale functional organization of long-range chromatin interaction networks. Cell Rep. 2012;2:1207–1219. [PubMed: 23103170]

46. Elkon R, et al. Characterization of noncoding regulatory DNA in the human genome. Nat Biotechnol. 2017;35:732–746. [PubMed: 28787426]

47. Xia Q, et al. The type 2 diabetes presumed causal variant within TCF7L2 resides in an element that controls the expression of ACSL5. Diabetologia. 2016;59:2360–2368. [PubMed: 27539148]

48. Pradeepa MM, et al. Histone H3 globular domain acetylation identifies a new class of enhancers. Nat Genet. 2016;48:681–686. [PubMed: 27089178]

49. Atlasi Y, et al. The interplay of epigenetic marks during stem cell differentiation and development. Nat Rev Genet. 2017;18:643–658. [PubMed: 28804139]

50. Liu L, et al. Modeling the relationship of epigenetic modifications to transcription factor binding. Nucleic Acids Res. 2015;43:3873–3885. [PubMed: 25820421]

Chapter 8

Epigenetic profiling reveals disrupted lipid metabolism in failing hearts and iPSC-cardiomyocytes with a pathogenic phospholamban mutation

Jiayi Pei[1,2,3]*, Renee Maas[1,2]*, Ema Nagyova[2,4], Johannes M.I.H. Gho[1], Christian Snijders Blok[1,2], Iris van Adrichem[1,2], Jorg Calis[1,2], René van Es[1], Shahrzad Sepehrkhouy[5], Dries Feyen[6], Noortje van den Dungen[4], Nico Lansu[4], Marc P. Buijsrogge[1], Nicolaas de Jonge[1], Hester M. den Ruijter[1], Gerard Pasterkamp[1], Manon M. Huibers[5], Roel A. de Weger[5], Linda van Laake[1], Marianne Verhaar[2,3], Peter van Tintelen[1], Frank G. van Steenbeek[1,2,7], Alain van Mil[1,2], Jan Willem Buikema[1], Boudewijn Burgering[8], Ioannis Kararikes[6], Mark Mercola[6], Pieter A. Doevendans[1,9], Joost Sluijter[1,2], Aryan Vink[5], Caroline Cheng[2,3], Michal Mokry[2,4,10]*, Folkert W. Asselbergs[1,9,11,12]*#, Magdalena Harakalova[1,2*#]

*equal contribution; #corresponding authors

[1]Department of Cardiology, Division Heart & Lungs, UMC Utrecht, University of Utrecht, Utrecht, The Netherlands
[2]Regenerative Medicine Center Utrecht, UMC Utrecht, Utrecht, The Netherlands
[3]Department of Nephrology and Hypertension, Division of Internal Medicine and Dermatology, UMC Utrecht, The Netherlands
[4]Epigenomics facility, UMC Utrecht, Utrecht, The Netherlands
[5]Department of Pathology, UMC Utrecht, Utrecht, The Netherlands
[6]Stanford Cardiovascular Institute and Department of Medicine, Stanford University, Stanford, CA, USA
[7]Department of Clinical Sciences, Faculty of Veterinary Medicine, University of Utrecht, NL
[8]Center for Molecular Medicine, Molecular Cancer Research Section, UMC Utrecht, The Netherlands
[9]Netherlands Heart Institute, Utrecht, The Netherlands
[10]Wilhelmina Children's Hospital, Department of Pediatric Gastroenterology, Division Child Health, Utrecht, the Netherlands
[11]Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College London, London, United Kingdom
[12]Health Data Research UK and Institute of Health Informatics, University College London, London, UK.

Manuscript in preparation

ABSTRACT

Background: Carriers of the R14del mutation in the phospholamban (PLN) gene develop cardiomyopathy and have an increased risk for life-threatening ventricular arrhythmias potentially due to adipocyte infiltration and fibrosis in the myocardium. The precise pathophysiological mechanisms underlying the severe fat infiltration in PLN R14del mutation carriers remain unknown. This study explores the role of epigenetic changes and their subsequent biological influences on metabolic changes in cardiac tissues and programmed cardiomyocytes from PLN R14del carriers.

Results: Using cardiac tissue of PLN R14del patients and donors, we identified differentially acetylated promoters and enhancers (H3K27ac ChIP-seq) with annotated enriched transcription factor (TF) binding motifs located in those regions and identified differentially expressed genes (RNA-seq). In line with previous studies showing the fibrofatty replacement in PLN R14del hearts at the histological levels, here we observed the downregulation of PPARA and KLF15, two key TF regulators in fatty acid oxidation (FAO) metabolisms in our multi-omics analyses. We further used a human induced pluripotent stem cell-derived cardiomyocytes (iPSC-CMs) from a PLN R14del patient and a healthy donor to study their transcriptional regulations using RNA-seq. Consistently, differentially expressed genes between PLN and wild type iPSC-CMs were highly enriched for biological processes involved in cellular metabolism. Seahorse assays were performed comparing the mitochondrial FAO function and glycolysis of PLN and wild type iPSC-CMs cultured in media with a variable composition of lipid and glucose, the two major sources of energy substrate for healthy cardiomyocytes. PLN R14del iPSC-CMs showed an enhanced preference for glucose utilization using the glycolysis pathway over FAO, and limited flexibility to switch between energy substrates as compared to the wild type. Using Nile red staining, we also detected an increased amount of the lipid droplet formation in PLN R14del iPSC-CMs when compared with the wild type allele. Interestingly, CRISPR-Cas9 corrected isogenic iPSC-CMs reversed the lipid droplet phenotype, suggesting a tight association between PLN R14del and the altered FAO.

Conclusion: Using multi-omics approaches, we obtained a unique list of chromatin regions and genes, including TF coding genes, which played important roles in the metabolism-related signaling in PLN R14del hearts. Furthermore, we confirmed these findings obtained from the end-stage cardiac tissue in the early stage PLN iPSC-CMs when compared to the wild type. Combined, the switch of the metabolic substrate preference in PLN R14del cardiomyocytes from mitochondrial fatty acid oxidation to cytoplasmic glycolysis provides several options for biomarkers to monitor disease progression as well as treatment strategy in PLN R14del, and beyond.

Keywords: phospholamban, heart failure, DNA regulatory region, chromatin regulation, H3K27ac, transcription factor binding motif

INTRODUCTION

The R14del mutation in the phospholamban (*PLN*) gene is associated with dilated cardiomyopathy (DCM) and/or arrhythmogenic (biventricular) cardiomyopathy (ACM) with a high risk for life-threatening ventricular arrhythmias.[1,2] To date, there is no effective treatment available for PLN R14del patients. This mutation explains a large proportion of Dutch DCM and ACM cases, 15% and 12%, respectively, which is linked to a shared haplotype originated around the year 1400 A.D. in the northern parts of The Netherlands.[3,4] Additional families carrying the PLN R14del mutation have been detected in other European countries, the United States, and Canada.[5] PLN is a small phosphoprotein located in the cardiomyocyte sarcoplasmic reticulum (SR) acting as the major regulator of SERCA2a/ATP2A2 activity and calcium (Ca2+) cycling. Previous research in PLN R14del cardiomyocytes confirmed disturbed Ca2+ cycling and abnormal cytoplasmic distribution of the PLN protein.[1,6] Aggregation, aggresome formation, and autophagy of the mutant PLN R14del protein in patient cardiac tissue have also been described.[7] Macroscopically, the end-stage PLN R14del heart shows a biventricular subepicardial fibrofatty tissue replacement of the myocardium, which is accompanied by profound interstitial fibrosis, transmural adipocyte infiltration, and islands of isolated cardiomyocytes between the adipocytes.[8–10]

Adipocyte infiltration and fibrosis in the myocardium serve as an anatomical barrier and thus create a substrate for fatal arrhythmias resulting from re-entry of the electrical signal.[11] The distribution of fibrofatty tissue replacement in PLN R14del carriers measured by late gadolinium enhancement can distinguish mutation carriers with and without ventricular arrhythmias.[12] Many diseases, such as the various types of acquired (e.g. ischemic) and genetic (e.g. due to mutations in *PKP2, DSG2, DSP, DSC2, JUP, SCN5A*, and *TNNT2)* cardiomyopathies, myotonic dystrophy, obesity, and atrial fibrillation, exhibit fibrofatty infiltration as one of the pathophysiological hallmarks.[11,13–15] However, it is still unclear which cell type(s) or mechanisms are responsible for the adipocyte infiltration by either the activation of the already existing pool of adipocytes or transdifferentiation of (cardiac) cells into adipocytes.[11] Interestingly, next to the extracellular adipocyte infiltration, abnormalities in intracellular cardiomyocyte lipid metabolism have been observed in arrhythmogenic right ventricular cardiomyopathy (ARVC) caused by mutations in *PKP2*. The production of lipid droplets and adipogenic markers in *PKP2* mutant cardiomyocytes have also been reported.[16] Induced pluripotent stem cell-derived cardiomyocytes (iPSC-CMs) with a homozygous *PKP2* mutation showed aggressive lipogenesis, elevated apoptosis, and reduced metabolic flexibility.[17] Furthermore, there are several forms of childhood cardiomyopathies caused by mutations in the mitochondrial fatty acid oxidation (FAO) pathway genes, such as *HADHA, HADHB, CPT2,*

ACADVL, which shown intracellular cardiomyocyte lipid droplet storage and adipocyte infiltration in the myocardium and other organs. Given the fact that lipid droplets and adipocyte infiltration is observed in various types of (genetic) forms of cardiac remodeling, a metabolism-related mechanism could be responsible for the progression of cardiomyopathies in general.

During heart failure, stressed cardiomyocytes suffer from an impaired capacity to oxidize fatty acids.[18] The failing heart subsequently switches to glycolysis for energy production,[19] resembling the energy balance of the fetal heart. However, the compensatory increase of glucose oxidation in the fetal-like energy management manner is not sufficient to meet adult energy consumption needs and will disturb the capacity of FAO-based ATP synthesis, leading to further starvation of the heart.[20,21] Two key regulators of glucose and fatty acid metabolism on a molecular level, glucose transporter 1 (*GLUT1*) and peroxisome proliferator-activated receptor alpha (*PPARA*), respectively, are both downregulated in failing hearts,[22] indicating insufficient energy production. Accumulation of lipid droplets has been shown in murine hearts by suppressing *PPARA* expression.[23] Myocardial adipogenesis in patients with cardiomyopathies, including PLN R14del carriers, is considered as a form of aberrant remodeling that is associated with the myocardial loss.[24] Nevertheless, the metabolic adaptation and the precise pathophysiological mechanisms in PLN R14del mutation carriers involved in severe fat infiltration remain unknown.

Here, we used multiple next-generation-sequencing approaches in a hypothesis-free manner to study the compact myocardial area of PLN R14del patients and healthy control hearts, in which the fatty epicardial layers were removed. We first profiled the activity of H3K27ac assayed by chromatin immunoprecipitation and sequencing (ChIP-seq) in human cardiac tissue from PLN R14del cardiomyopathy patients in comparison to healthy donor hearts. Based on the list of differentially acetylated regulatory regions between PLN patients and controls, we performed *in silico* region-to-gene annotations and TFBMs enrichment analysis. We also obtained the transcriptional changes in PLN versus control hearts using RNA sequencing (RNA-seq). By integrating the information at the DNA and RNA levels, we identified a set of TFs, which played a critical role in the end-stage PLN R14del hearts, enriched for the energy metabolism. To further characterize the identified TFs and their regulated metabolic signaling in the early stage at the cellular level, we used iPSC-CMs harboring the PLN R14del mutation and examined their metabolic changes in comparison to wild type iPSC-CMs at the RNA and the protein levels. Multiple *in vitro* assays were performed to elucidate the early-stage pathological phenotype in relation to the late-stage *ex vivo* results.

METHODS

Study design and samples

This study was approved by the Biobank Research Ethics Committee, University Medical Center Utrecht, Utrecht, The Netherlands, under protocol number 12-387 (cardiac tissues), and 14-513 (skin biopsies). Written informed consent was obtained or in certain cases waived by the ethics committee when obtaining informed consent was not possible due to the death of the individual. Heart samples collected at autopsy or transplantation were obtained from a homogeneous cohort of patients whom all carried the same pathogenic PLN R14del mutation (n=6). Four control hearts obtained from rejected organ donors (n=3) or autopsy (n=1) were used as a reference. To further elaborate on PLN R14del-specific changes, hearts from patients with ischemic cardiomyopathy (n=4) and non-ischemic cardiomyopathy based on mutations in genes encoding sarcomeric proteins (n=6) were also included. An overview of cardiac tissues is presented in Table S1A. Cardiac tissues used for ChIP-seq and RNA-seq were obtained from regions halfway between the atrioventricular valves and the apex and were stored at -80°C. From each individual, an adjacent block of tissue from the same biopsy was used for ChIP-seq and was paraffin-embedded and stained with Masson's trichrome (Figure S1A). High resolution systematic digital histological quantification of fibrosis and fatty tissue in trichrome stained slides was used to create schematic overviews showing mean fibrosis or adipose tissue in the PLN and control groups (Figure S1B). If applicable, macroscopically visible regions of subepicardial fat or myocardial fibrofatty replacement were removed from frozen tissue blocks and 12-30 tissue slices of 10 µm thick were cut to achieve a comparable amount of myocardial tissue in the sequenced material. For all the other samples ten 10 µm thick frozen slices were collected.

Dermal fibroblasts were obtained from a 4mm skin biopsy under informed consent from one PLN R14del patient and one healthy donor, followed by the reprogramming of iPSC. Wild type and PLN R14del iPSC cell lines were kindly provided by Dr. Joseph Wu (Stanford Medicine) and Prof. Mark Mercola (Stanford Medicine), respectively. Wild type cardiomyocytes (CVI-111) and PLN R14del cardiomyocytes (D4) were differentiated from three different iPSC colonies (n=3) per group, and they were included in all cell-based experiments as shown below. In terms of the immunofluorescence staining, an additional cardiomyocyte cell line (C31), in which the PLN R14de mutation was corrected using CRISPR/Cas9 from the PLN R14del iPSC cell lines, was included. An overview of iPSC-CMs is presented in Table S1B.

Chromatin H3K27ac immunoprecipitation and sequencing of human cardiac tissues

Chromatin was isolated from frozen cardiac tissues using the MAGnify™ Chromatin Immunoprecipitation System kit (Life Technologies) according to the manufacturer's instructions. In brief, the obtained cardiac tissue was crosslinked with 1% formaldehyde and the crosslinking was stopped by adding 1.25 M glycine. Cells were lysed using the kit-provided lysis buffer and nuclei were sonicated using Covaris microTUBE (duty cycle 5%, intensity 2, 200 cycles per burst, 60s cycle time, 10 cycles). Sheared chromatin was diluted based on the expected number of isolated cells and was incubated with an anti-H3K27ac antibody (ab4729, Abcam) pre-coupled to magnetic beads for 2 hours at 4°C. Beads were extensively washed and crosslinking was reversed by the kit-provided reverse crosslinking buffer with 20 mg/mL Proteinase K. DNA was purified using ChIP DNA Clean & Concentrator kit (Zymo Research). Isolated DNA was additionally sheared, end-repaired, sequencing adaptors were ligated and the library was amplified by PCR using primers with sample-specific barcodes according to our modification to manufacturer's recommendations. After PCR, the library was purified and checked for the proper size range and for the absence of adaptor dimers on a 2% agarose gel and sequenced on SOLiD Wildfire sequencer.

H3K27ac ChIP-seq analyses

Sequencing reads were mapped against the reference genome (hg19 assembly, GRCh37) using the BWA package (-c, –l 25, -k 2, -n 10).[25] Multiple reads mapping to the same location and strand were collapsed to single reads and only uniquely placed reads were used for peak/region calling. Regions were called using Cisgenome 2.0 (–e 150 -maxgap 200 –minlen 200).[26] Next, to obtain a common reference, region coordinates from all PLN and control samples were stretched to at least 2000 base pairs and collapsed into a single common list. Overlapping regions were merged based on their outmost coordinates. Only the autosomal regions supported by at least 2 independent datasets were further analyzed. Sequencing reads from each ChIP-seq library were overlapped with the common region list, to set the H3K27ac occupancy for every region-sample pair. Obtained regions were further examined in 4 analyses (Table S2).

Detection and gene annotation of differentially acetylated regions (Analysis 1): Regions with differential H3K27ac occupancy between PLN and control hearts were identified using DESeq2 standard settings ($p<0.05$ as calculated by Wald test)[27] and are referred to as 'differentially acetylated regions'. Supervised hierarchical clustering was performed with quantile normalized (limma::normalizeQiantiles() function in R), log2 transformed and median centered read counts per common region. To avoid the log2 transformation of zero values, one read was added to each region. Region to gene annotation was performed in silico using a conservative window of +/-5kb from the transcription start site (TSS).

Enrichment of TF binding motifs (TFBM) in differentially acetylated regions (Analysis 2): A total of 3396 cardiac DNAse hypersensitivity sites (DHS) obtained from the ENCODE database (Heart_OC, Primary frozen heart tissue from NICHD donor ID:1104, Male, Caucasian, 35 years old)[28] overlapping with differentially acetylated regions in the PLN vs. control group (both hypo- and hyperacetylated) were used for this analysis. The genomic sequence of DHS was repeat masked and the enrichment of TFBM was calculated against the shuffled sequences using the Analysis Motif of Enrichment (AME tool) of the MEME Suite[29] with the following settings: motif database: human (HOCOMOCO v9), background model sequence set to 0.29182,0.20818,0.20818,0.29182, pseudo count added to a motif column: 0.25, Wilcoxon rank-sum test (quick), $p<0.05$), number of multiple tests for Bonferroni correction: #Motifs× #PartitionsTested = 426×1 = 426. The functional annotation of the enriched TFs was performed as described above.

PLN-specificity analysis of differentially acetylated regions (Analysis 3): Sequencing reads from each ChIP-seq sample (PLN (n=6), control (n=4), ischemic (n=4), and sarcomeric (n=6) groups) were compared to the common differentially acetylated region list to set the H3K27ac occupancy for every region-sample pair. Raw read counts were quantile normalized (limma::normalizeQiantiles() function in R), log2 transformed and median centered (to avoid log2 transformation of zero values, one read was added to each region). The median value from each sample group was used to construct an n x k table where n = 4 (one value per each sample type) and k represent the number of differentially acetylated regions. The k-means (nstart = 200) function in R was used to partition the regions into 12 different clusters. To enable the reproducibility of identified clusters set.seed(10) R command was called before the clustering. Annotation of clusters with PLN-specific patterns was performed as described above.

Pathway analysis of genes annotated to differentially acetylated regions (Analysis 4): ToppFun and STRING were used for gene list enrichment analysis and candidate gene prioritization based on functional annotations and protein interactions networks.[30,31] For ToppFun, the list of hyper- and hypoacetylated genes was tested using probability density function p-value calculation method, FDR B&H correction, p-value cut-off of 0.05, and gene limit of 1-2,000 genes per term. Since the pre-build gene/protein networks integrated into ToppFun were not created using the same criteria and the annotated number of genes varies significantly, we reported also genes belonging to known disease pathways even below the p-value threshold (where indicated). For protein network interaction visualization STRING v10.0 was used with a minimum required interaction score at the highest confidence setting for all differentially acetylated peaks.

Transcriptome analysis of human cardiac tissues using RNA sequencing

RNA was isolated using ISOLATE II RNA Mini Kit (Bioline) according to the manufacturers' instructions with minor adjustments. After the selection of mRNA, libraries were prepared using the NEXTflexTM Rapid RNA-seq Kit (Bioo Scientific). Libraries were sequenced on the Nextseq500 platform (Illumina), producing single-end reads of 75bp. Reads were aligned to the human reference genome GRCh37 using STAR v2.4.2a.[32] Picard's AddOrReplaceReadGroups v1.98 (http://broadinstitute.github.io/picard/) was used to add read groups to the BAM files, which were sorted with Sambamba v0.4.5 and transcript abundances were quantified with HTSeq-count v0.6.1p1 using the union mode.[33,34] Subsequently, reads per kilobase per million mapped reads (RPKMs) were calculated with edgeR's RPKM function.[35] DESeq2 was used to identify differentially expressed regions using the cutoff of $p<0.05$ in the Galaxy environment (default settings).[36]

Differentiation of iPSC into cardiomyocytes

Reprogrammed Sendai (kit) accordingly as described before.[37] Briefly, human iPSCs (>p20 <p50) were grown to ~90% confluence in 6 wells format and maintained in E8 medium for at least three passages before starting cardiac lineage differentiation. Upon differentiation, the medium was changed to medium without insulin (INS- medium), as insulin inhibits cardiac mesoderm formation.[52] On day zero (day 0) 3 ml/well medium with insulin (INS+ medium) was supplemented with 6-8 µM CHIR99021 (Selleck Chemicals). After one night incubation, an additional 2 ml of the INS^- medium was carefully added (day 1), and an additional 1 ml of the INS- medium was added on day 2. On day 3, the medium was replaced by 3 ml/well INS^- medium supplemented with 2 µM Wnt-C59 (Tocris Bioscience). Afterward, the culturing medium was changed every other day and contracting cells were generally seen between day 7 and day 9. After day 9, the INS- medium was replaced by 3ml/well INS+ medium. To metabolically select and purify iPSC-CMs, INS+ was replaced with a purification medium (without D-glucose) for 2 days. Purified iPSC-CMs were continuously cultured in INS+ for 2 days to recover. On day 15 iPSC-CMs were detached by 10 mins incubation of 10X TrypLE select (Thermo Fisher) at 37°C. Detached cells in TrypLE solution were flushed using INS+ medium and cell pellets were obtained by centrifuging at 300g for 3 mins. After aspiration, iPSC-CMs were resuspended in INS+ medium supplemented with ROCK inhibitor and KO-serum (Thermo Fisher) and filtered through a 100 µm pore filter (Corning). Approximately three million iPSC-CMs were re-plated into Matrigel-coated wells in a 6-well plate and further cultured for 2 days in INS+ medium for

recovery. To ensure the purity of iPSC-CMs, they were cultured in purification medium for 2 days before all cell-based experiments in this study (e.g. RNA-seq and Seahorse assays).

Media used for the cultivation of iPSC-CMs

To mature iPSC-CMs, we used the maturation medium developed in the Mercola lab (D. Feyen, et al. 2020, under revision, Cell Reports). After the 2nd purification as indicated above, INS+ medium was replaced by the maturation medium (Table S3) for a minimum of 3 weeks, which metabolically matured iPSC-CMs. Next, to study the metabolic capacity of iPSC-CMs, they were continuously cultured in glucose-rich medium or lipid-rich medium (Table S3) for another 3 weeks before the seahorse assays and the immunofluorescence analyses.

RNA isolation and transcriptome analysis of iPSC-CMs using RNA sequencing

Total RNA was isolated from cells using TriZol (Ambion) according to the manufacturer's instructions. DNase treatment was carried out using the RNase-Free DNase Set (Qiagen). The quality of total RNA was first examined using Bioanalyzer 6000 Pico Kit (Agilent), and mRNA was selected. Libraries were prepared and sequenced as indicated above. Raw counts per gene per sample were imported to the Galaxy server and differentially expressed genes between PLN and wild type iPSC-CMs were identified using DESeq2 as indicated above.[36]

Measure iPSC-CMs metabolism using seahorse assay

Metabolic activity of purified iPSC-CMs was determined by mitochondrial fatty acid β-oxidation (FAO) and cytosolic glycolysis using the Seahorse XFe24 Extracellular Flux Analyzer (Seahorse Bioscience). Briefly, long-term cultured iPSC-CMs (77-159 days old) were first seeded to Matrigel-coated Seahorse XFe24 assay plates at a density of 50,000 cells/well and cultured in the standard culture medium for 2 days. Afterward, the standard culture medium was replaced by glucose-rich or lipid-rich medium and refreshed every other day (Table S3). To measure the oxygen consumption rate (OCR, pmol/min/μg protein) and extracellular acidification rate (ECAR, mpH/min/μg of protein), iPSC-CMs were first washed three times with 0.5 ml fresh seahorse medium (Agilent Technologies/Seahorse Bioscience) supplemented with 2% B27, 1%

Chemically Defined Lipid Concentrate (Gibco), 4mM L-glutamine (Gibco), and 10mM glucose (Agilent, 103577-100) and then kept in the same medium for 1 hour in a non-CO^2 incubator at 37 °C. Afterward, 100 µM etomoxir (ETO, a specific irreversible inhibitor of carnitine palmitoyltransferase 1, Agilent)[16] was added to determine FAO dependency by measuring the oxygen consumption rate (OCR, pmol/min/µg protein), followed by the addition of 50 mM 2-Deoxy-D-glucose (2DG competitive glycolytic inhibitor, Sigma-Aldrich)[16] to determine cytosolic glycolysis by measuring extracellular acidification rate (ECAR, mpH/min/µg of protein). Three measurements were taken before and after each injection and mixing cycle. Both OCR and ECAR were compared between PLN R14del and wild type iPSC-CMs.

Seahorse mitochondrial FAO assay by measuring OCR: Each parameter for % OCR from baseline were measured and processed as follows: (1) Baseline: the last measurement of basal respiration before ETO injection (time point 3) was normalized to 100%, and the base measurement was defined by combining three measurements normalized to time point 3. (2) FAO: the three measurements after ETO injection normalized to time point 3. (3) Glycolysis: the three measurements after 2-DG injection normalized to time point 3. (4) FAO dependency: the percentage of fatty acids dependence was calculated by quantifying the change in basal OCR after ETO-inhibited FAO compared to the total mitochondrial function from other substrate oxidation. (5) Metabolic flexibility: the maximal respiration due to the utilization of other substrates was determined by the OCR after 2-DG-inhibited glycolysis.

Seahorse glycolysis assay by measure ECAR: Each parameter for % ECAR from baseline were measured and processed as follows: (1) Baseline: the last measurement of baseline before ETO injection (time point 3) was normalized to 100%, and the base measurement was defined by combining three measurements normalized to time point 3. (2) Glycolysis: the three measurements after ETO injection normalized to time point 3. Glycolysis was quantified as the maximum percentage increase of ECAR over baseline. (3) Glycolysis inhibition: the three measurements after 2-DG injection normalized to time point 3. (4) Glycolysis dependency: the degree of glycolysis dependency was determined by the reduction of glycolysis after 2-DG injection. (5) Glycolytic reserve was determined by the average ECAR level after ETO injection substrated with the average ECAR level after 2-DG injection.

Statistical analysis of OCR and ECAR measurements: The experiment was done in 3 biological replicates, with each replicate consisting of 5–12 technical repeats per condition. OCR and ECAR rates were normalized to the non-glycolytic acidification rate per well and all values were further normalized to the *in situ* nuclear staining cell counts (Mean ± SD, n=3 counts/well). To compare OCR and ECAR between PLN R14del and wild type iPSC-CMs, an unpaired t-test or one-way ANOVA (biological replicates n=3, n=5-12

wells per replicate) was used. All data was represented as Mean ± SD and significance was displayed by *p < 0.05; **p < 0.01; ***p < 0.001, 0.001 and ****P ≤.0001. All ANOVA tests were subsequently analyzed using Tukey's post hoc test. Statistical analysis was performed using the GraphPad Prism 8.3 software.

Morphological characterization of iPSC-CMs

iPSC-CMs were cultured on Matrigel-coated coverslips, fixed in paraformaldehyde (4%), and permeabilized in blocking/permeabilization buffer (5% BSA/0,3% Triton-X-100 in PBS) for 30 minutes. Primary antibodies were added as shown in Table S4 and incubated in 1:5 diluted blocking/permeabilization buffer DPBS overnight at 4°C. Cells were washed four times with PBS (5 mins each time) and incubated with Alexa-conjugated secondary antibodies (Life Technologies) diluted in 1:5 blocking/permeabilization buffer (Table S4) in the dark at room temperature for 1 hr. Cells were washed as previously described and nuclei were stained using 1 µg/ml Hoechst (Life Technologies) for 15 mins. Coverslips were mounted using Fluoromount-G (Southern Biotech) and images were acquired using a Leica SP8X confocal microscope. Immunofluorescence stainings of Nile-red were quantified by the lipid mask macro of Image J, displaying the number of lipids/nuclei and were quantified by using one-way ANOVA (n=3/3 coverslips/images per well, n=3 biological replicates, n=>10 wells per condition).

RESULTS

The histone acetylome of PLN R14del cardiac tissue differs from the controls

Firstly, using an unbiased ChIP-seq approach on cardiac tissues we identified 28,149±9,538, and 25,721±8,460 regions with differential H3K27ac binding in PLN and control hearts, respectively (Table S2). Next, we combined regions that were identified in at least two independent samples into a set of 23,356 regions to assess differentially acetylated regions between two groups. In total, 2,107 autosomal regions showed differential H3K27ac levels (Figure 1A). Among these, 958 and 1,149 regions were identified as hyper- and hypoacetylated in PLN hearts as compared to the controls, respectively (Table S5). Although we used end-stage cardiac tissues from PLN R14del hearts, a reasonable number of differentially acylated regions were identified in PLN versus control hearts, followed by several *in silico* analyses.

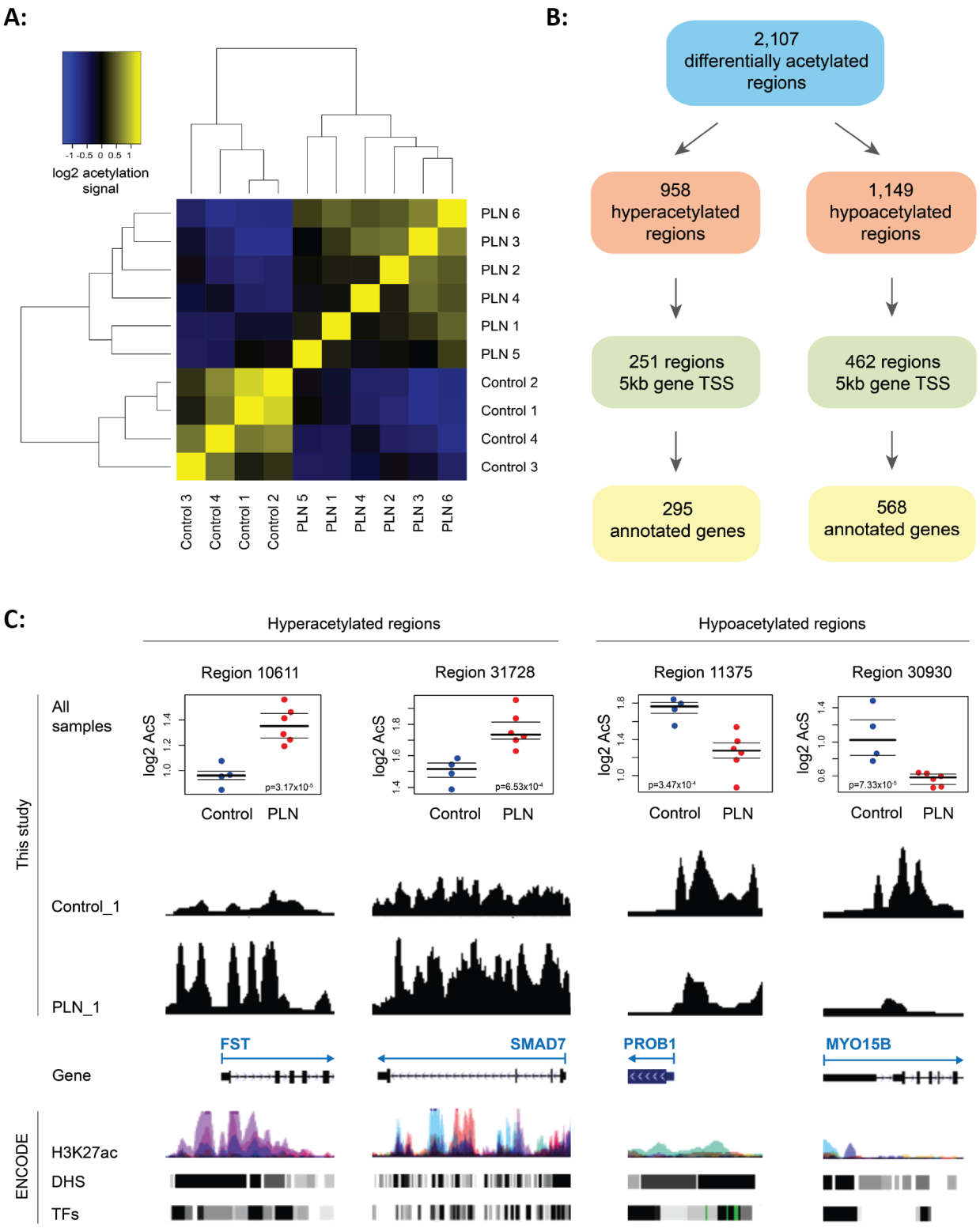


Figure 1: A) Differentially acetylated regions between PLN and control hearts. **B)** An overview of annotated genes from the hyper- and the hypoacetylation regions. **C)** Selected examples of differentially acetylated regions depicted using the UCSC Genome Browser. In each example the region overlaps with the putative promoter (upstream region, 5'UTR, first exons in region tracks) and shows a significant difference between PLN patients and controls (dot plots).

Arrow beginning indicates gene transcription start site (TSS). More region details can be found in Table S3 and S4. AcS = acetylation signal. ENCODE = publicly available ENCODE consortium data default display. H3K27ac = layered H3K27ac ChIPseq data in 6 cell types (GM12878 - red, H1-hESC – orange, HSMM – green, HUVEC – light blue, K562 – dark blue, NHEK – purple, NHLF – pink). DHS – DNaseI hypersensitivity clusters in 125 cells, TFs = ChIPseq for 161 TFs.

Genes annotated to differentially acetylated regions in PLN R14del cardiac tissue are enriched in remodeling and metabolic pathways

Next, we focused on differentially acetylated regions in close vicinity to gene bodies. We have chosen a conservative cutoff and followed only genes with differentially acetylated regions using a window of ±5kb between the peak center and transcription start site (TSS) of a gene. In total, 295 genes were identified in the close vicinity to 251 hyperacetylated peaks, including 233 protein-coding genes, 2 pseudogenes, 38 long and 22 short non-coding RNA genes. We have also detected 568 genes close to 462 hypoacetylated peaks, including 476 protein-coding genes, 5 pseudogenes, 53 long and 34 short non-coding RNA genes (Figure 1B, Table S5). Selected examples of regions in the vicinity of 4 cardiac genes (*FST*, *SMAD7*, *PROB1*, and *MYO15B*) compared to H3K27ac signals from the ENCODE project are shown in Figure 1C. To predict altered biological processes in PLN hearts when compared with controls, we performed gene enrichment analysis of 863 genes annotated to differentially acetylated regions using ToppFun (padj<0.05). The analysis of 295 hyperacetylated genes resulted in several enriched GO terms and pathways of various sources related to three main groups: 1. fibrosis (extracellular matrix components, collagen synthesis, TGF-beta signaling pathway, cell adhesion, and immune system), 2. (cardiovascular) development and 3. chromatin assembly (histone function, nucleosome, cell senescence and proliferation, Table S6A). Significantly enriched processes related to chromatin assembly were mostly based on the core histone H1 cluster on chromosome 6 based on 5 neighboring significantly hyperacetylated regions. On the other hand, gene enrichment analysis of hypoacetylated 568 genes resulted in two main groups: 1. (lipid) metabolism (fatty acid beta-oxidation, transferase activity) and 2. mitochondrial function (Table S6B).

The predicted transcription factor binding sites located in differentially acetylated regions in PLN R14del cardiac tissue are enriched in remodeling and metabolic pathways

Next, we studied TFBMs that were enriched in the differentially acetylated regions, which might be involved in the pathogenesis of the end-stage heart failure. We employed the AME program to test the enrichment of TFBMs using the DNA sequence of differentially acetylated peaks and detected enrichment

in 202 TFBMs and annotated them to 200 TF-encoding genes (Table S7A). In line with previous results, gene enrichment analysis of annotated TFs pointed to similar biological processes, among others, those involved in adipogenesis, myogenesis, mitochondrial structure, TGF beta signaling in fibrosis, heart development, and cell death (Table S7B). Protein interactions of several TFs linked to notable heart failure pathways are shown in Figure 2A, including the well-described key TFs involved in heart development, such as *GATA4*, *TBX5,* and *MEF2C*.[38] Interestingly, the regulatory regions in the close vicinity of gene bodies encoding seven TFs with enriched TFBM also showed differential acetylation level: *ESR1, ESRRA, KLF3, MAFG, MYC, NFATC2, and ZFHX3*. Based on the enriched TFBMs inside differentially acetylated regions, we thus provide evidence that point to TFs working together with genes with differentially acetylated regulatory regions in the development of heart failure. Examples of several TFs with their enriched TFBMs are shown in Figure 2B.

Metabolic pathways are overrepresented in PLN R14del cardiomyopathy as compared to other types of cardiomyopathies

To further characterize specific pathways involved in PLN R14del and to exclude general events linked to end-stage heart failure, we also compared H3K27ac profiles of patients with other genetic types of ischemic cardiomyopathy (n=4) and non-ischemic dilated cardiomyopathy (sarcomeric group, n=6). K-mean analysis was used to partition the PLN R14del-specific regions into 12 distinct clusters, which were then compared to clusters of regions linked to control or other cardiomyopathies (Table S8A). Based on this analysis, we focused on the main clusters showing a PLN R14del-specific pattern (Figure 3A). We have identified two clusters with hyperacetylation (cluster 1 and 12) linked to 50 genes and two clusters with hypoacetylation (clusters 3 and 4) linked to 158 genes that are unique in PLN 14del when compared with the other groups (Table S8B). We show PLN R14del-specific patterns of eight regions in the vicinity of genes, such as *LINC01140, WIPF1, OSR1, IFFO2, LTBP1, GCGR, or CNNM4* (Figure 3B). At least 13 genes from the hyperacetylated group and 55 genes from the hypoacetylated group are known to be involved in various pathways related to (phospho)lipid and lipoprotein synthesis and metabolism, including genes such as *DGKZ, HADHA/HADHB, INPP5J, MLYCD, PLCD3, and PLPP1*. Thus, by comparing differentially acetylated regions in PLN patients to other types of dilated cardiomyopathy, we detected region clusters that distinguishes PLN R14del from other types of cardiomyopathies.

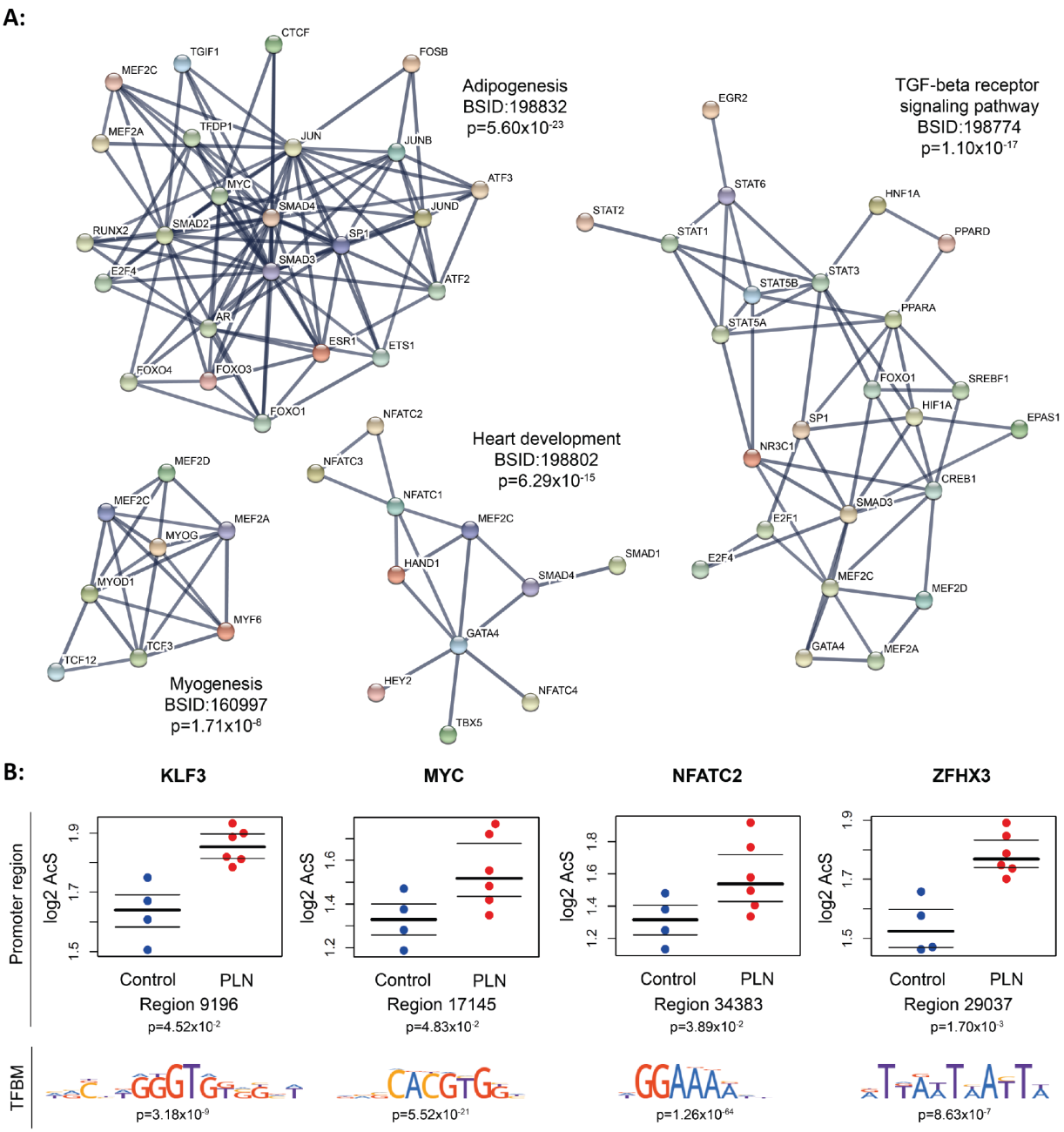


Figure 2: Enrichment of TF binding motifs in differentially acetylated regions. **A)** Selected examples of STRING protein database output used to detect interacting TFs annotated by both hyper- and hypoacetylated regions. Only the highest confidence interactions are displayed. Disconnected nodes were removed from the network image. **B)** Selected examples of TFs with both differentially acetylated promoter as well as enrichment of binding motif in the differentially acetylated regions. Dot plots represent the acetylation signal (AcS) measured for each sample.

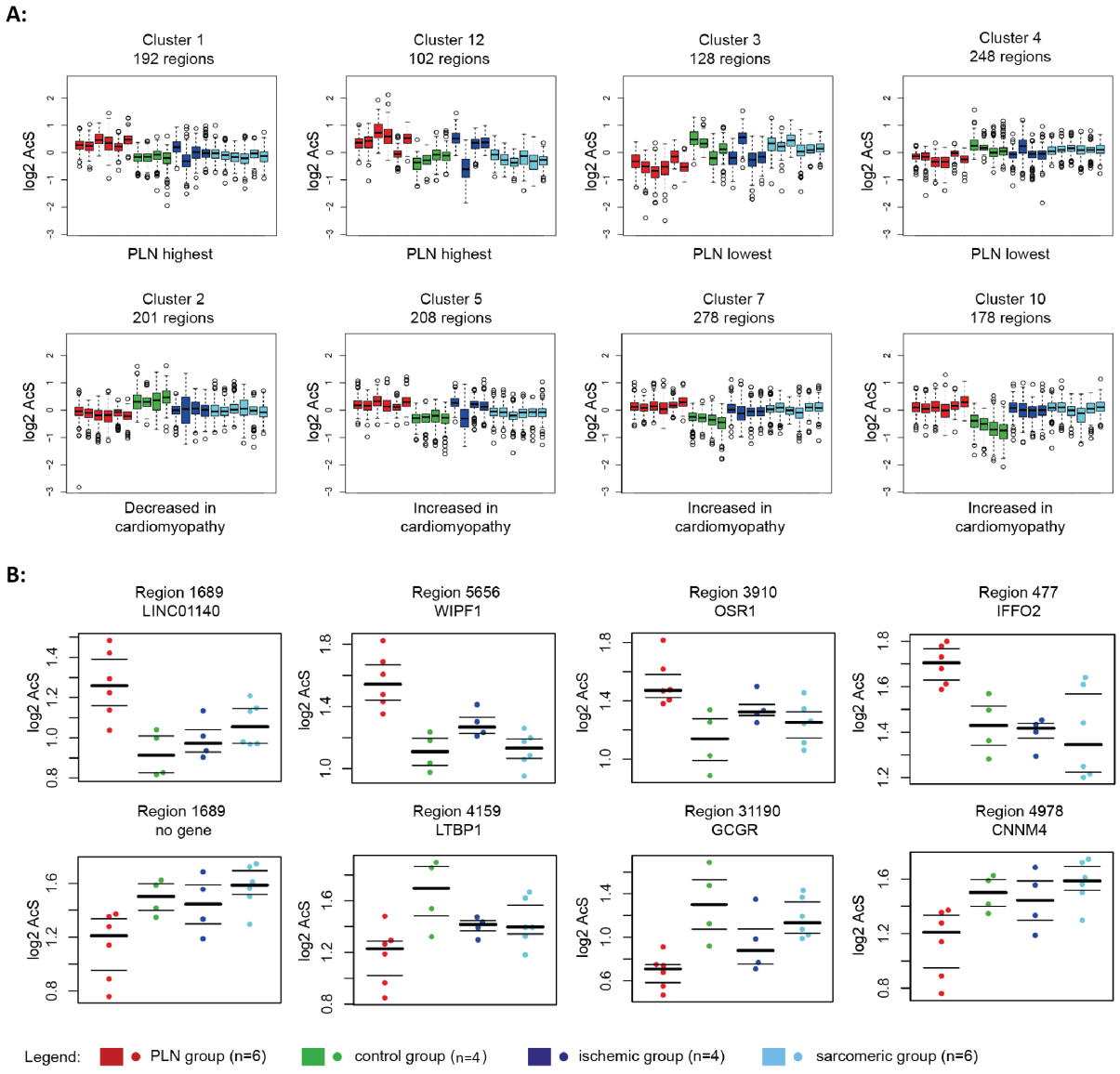


Figure 3: K-mean clusters of PLN R14del cardiomyopathy (red), healthy controls (green), ischemic cardiomyopathy (dark blue) and sarcomeric non-ischemic cardiomyopathy (light blue) based on H3K27ac signal was user to partition the regions into 12 different clusters. Selected were clusters considered to be specifically hyperacetylated in PLN R14del **(A)** and hypoacetylated in PLN R14del **(B)** as compared to other conditions. (n=number of detected regions).

Differentially expressed genes in PLN R14del cardiac tissue are enriched in remodeling and metabolic pathways

Using RNA-seq on cardiac tissues from exactly the same PLN and control hearts samples that were used for epigenetic analysis, we obtained 3,541 differentially expressed genes between PLN and control hearts. Among these, 1,668 and 1,873 genes were identified as up- or downregulated genes in the PLN group as

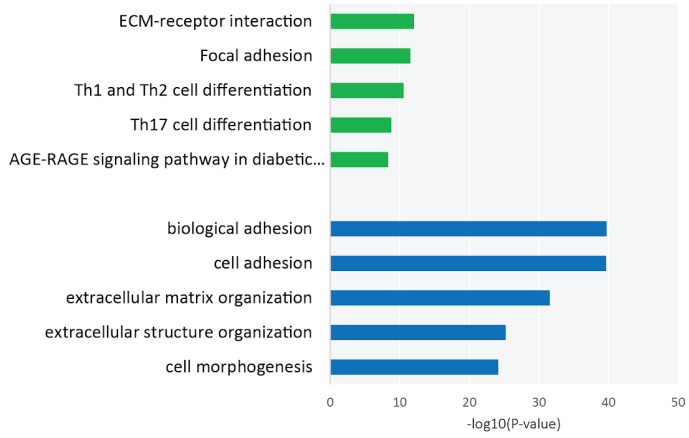
compared with the control group, respectively (Table S9). Similar to the isolated chromatin data, mRNAs isolated from end-stage heart failure tissue of explanted hearts also provided a reasonable number of differentially expressed genes for further annotation.

To study the biological processes that occur in PLN hearts as compared with the controls, we examined the pathways enriched by the differentially expressed genes. Consistent with the results using annotated genes from differentially acetylated regions, the enriched GO terms and pathways by 1,668 upregulated genes were mostly involved in fibrosis (extracellular matrix organization, cell adhesion, collagen binding, TGF-beta signaling pathway), and cardiovascular development (circulatory system development, blood vessel development, Figure 4A, and 4B, Table S10A). Similarly, gene set enrichment analysis of 1,873 downregulated genes was dominantly involved in energy metabolism (oxidative phosphorylation, ATP synthesis coupled proton transport, acetyl-CoA C-acyltransferase activity) and mitochondrial function (Table S10B).

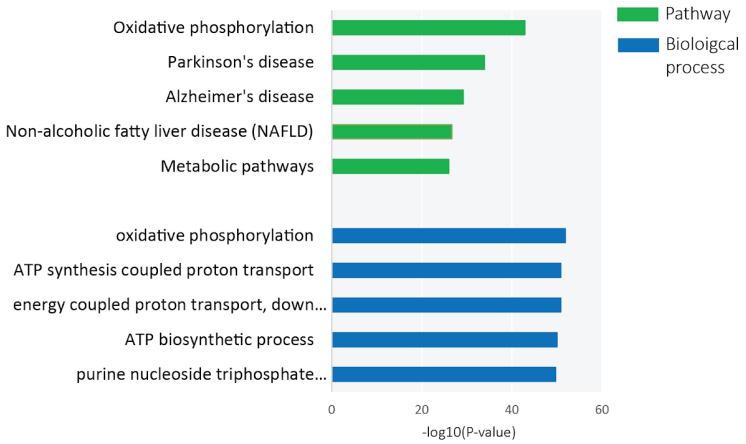
Integrative analysis of the ChIP-seq and RNA-seq data in PLN versus control hearts further validates enrichment of identified TFs involved in remodeling and metabolic pathways

To investigate the upstream players that could play a major role in regulating the subsequent biological processes, we performed an integrative analysis using both the ChIP-seq and the RNA-seq data. We collected all 200 TF-encoding genes, which were annotated from the enriched TFBMs in differentially acetylated regions using the ChIP-seq data, and examined their mRNA expression changes between the PLN and the control groups using the RNA-seq data. As a result, 26 and 13 annotated TF coding genes also showed significantly up- and downregulated mRNA expression levels in PLN versus control hearts (Figure 4C). Notably, several overlapping TF coding genes are well-known to play a critical role in the most enriched GO terms and pathways that were identified using the ChIP-seq and/or RNA-seq data, such as ECM-related *SMAD3* (upregulated) and metabolism-related *KLF15* (downregulated).

A: Enrichment analysis of upregulated genes



B: Enrichment analysis of downregulated genes



C:

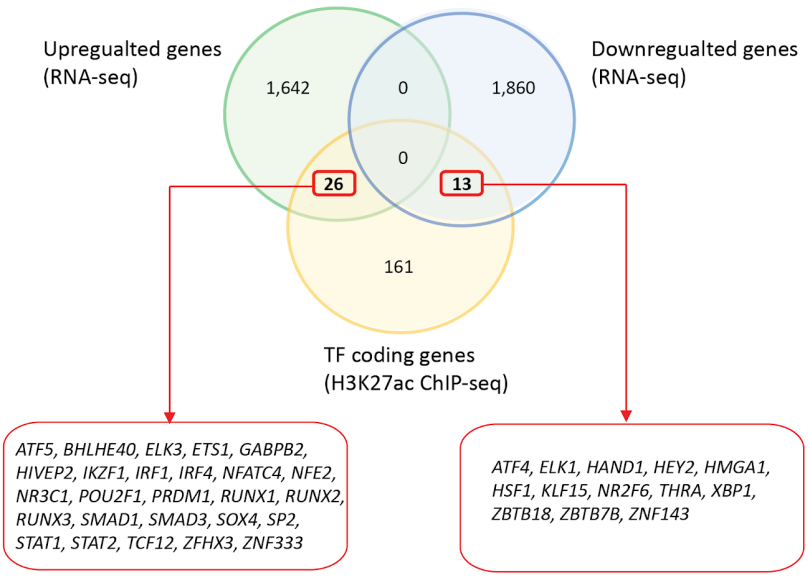


Figure 4: A) Top 5 enriched GO terms by upregulated genes in PLN versus control hearts; **B)** Top 5 enriched GO terms by downregulated genes in PLN versus control hearts; **C)** TF coding genes, which were predicted to bind to enriched motifs in H3K27ac ChIP-seq data, with altered mRNA expression levels between PLN and control hearts.

Differentially expressed genes in PLN iPSC-cardiomyocytes are enriched in remodeling and metabolic pathways

Besides the findings obtained at the tissue level using the end-stage PLN hearts, we also studied the transcriptome profiles in PLN R14del iPSC-CMs when compared with wild type iPSC-CMs. In total, we obtained 963 differentially expressed genes between PLN and wild type iPSC-CMs (Table S11A), which were both long-term cultured (77-159 days old) in maturation medium (From now referred to as early-stage (PLN) iPSC-CMs). Among these, 727 and 226 genes were identified as up- and downregulated genes in PLN versus the wild type iPSC-CMs, respectively. We also examined the transcriptome profiles of PLN and wild type iPSC-CMs cultured in glucose-rich medium or lipid-rich medium to further investigate their energy metabolism. In the glucose-rich medium, we obtained 1,321 differentially expressed genes between the PLN and the wild type groups (Table S11B), composed of 419 upregulated and 902 downregulated genes. In contrast, in the lipid-rich medium, we obtained 2,104 differentially expressed genes between the PLN and the wild type groups (Table S11C), which included 680 upregulated and 1,424 downregulated genes.

Next, we performed enrichment analyses using differentially expressed genes between PLN and wild type iPSC-CMs (Figure 5). Interestingly, the most enriched GO terms and pathways defined by upregulated genes were involved in fibrosis, regardless if they were exposed to maturation medium, the lipid-rich medium, or glucose-rich medium (Table S12A-C). Likewise, the most enriched GO terms and pathways defined by downregulated genes were involved in energy metabolism, similarly regardless of the type of their culture media. Nevertheless, it is important to note that the more detailed biological processes in the energy metabolism differed among the three culturing conditions. Consistent with the results obtained from the cardiac tissues, the general metabolic processes, such as oxidative phosphorylation and ATP synthesis coupled proton transport, remained the top enriched GO terms and pathways under the maturation condition (Table S12A). Enriched metabolic processes under the glucose-rich condition were more related to contraction regulation, such as cardiac muscle contraction and actin filament-based movement (Table S12B). Enriched metabolic processes under the lipid-rich condition were more associated with the endoplasmic reticulum activities, such as protein localization to the endoplasmic reticulum (ER) and protein targeting to ER (Table S12C). Strikingly, glycolysis, a glucose-based metabolic pathway, was also significantly enriched by the downregulated genes under the lipid-rich condition.

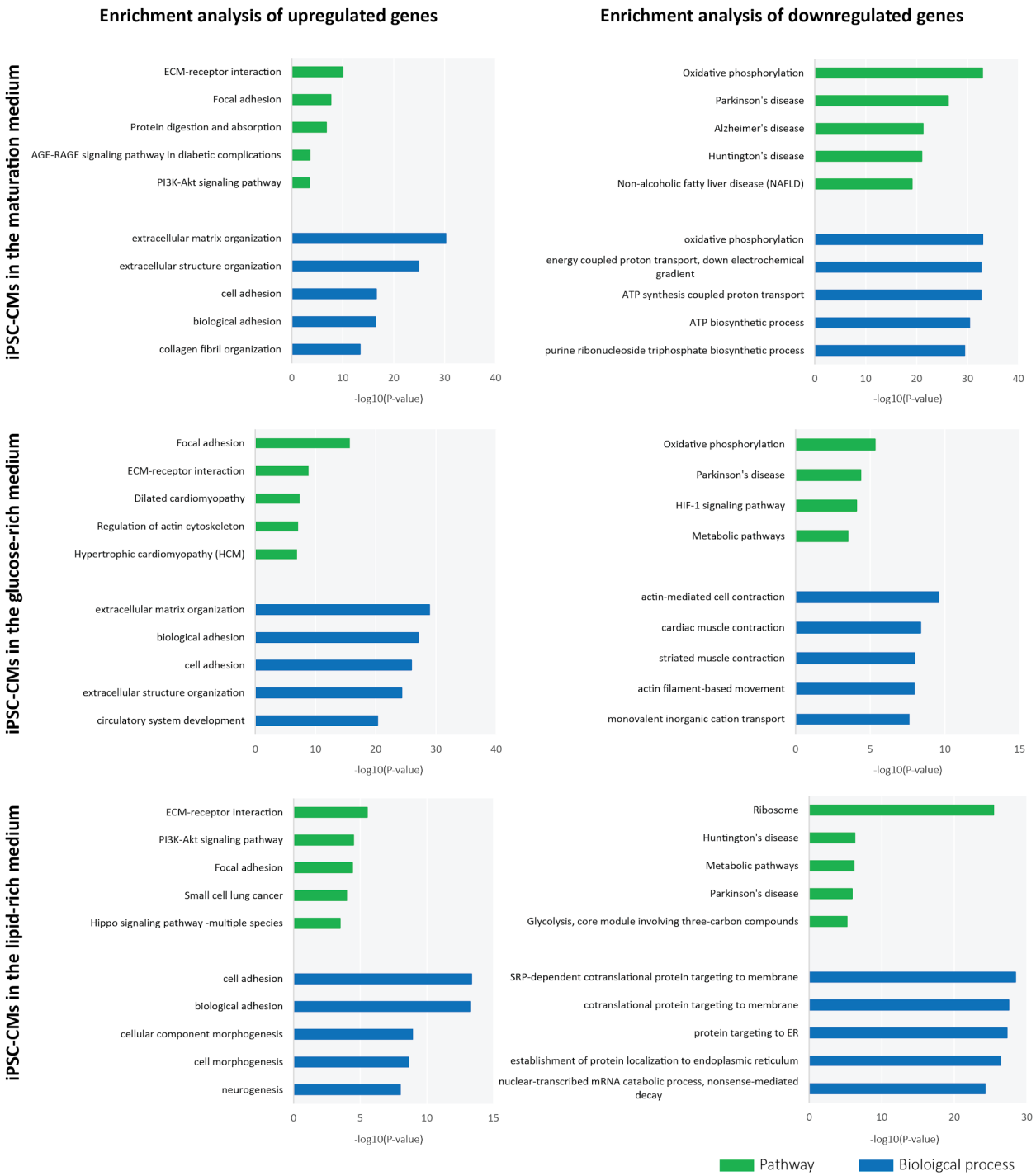


Figure 5: Top enriched GO terms by differentially expressed genes between PLN and wild type iPSC-CMs cultured in the maturation medium, the glucose-rich medium, and the lipid-rich medium.

Fatty acid oxidation (FAO) and metabolic flexibility are impaired in PLN R14el iPSC-cardiomyocytes

Next we used Seahorse analysis to investigate mitochondrial FAO at the functional level, using a protocol that combines irreversible ETO-FAO blockade with 2-DG glycolysis blockade to evaluate mitochondrial respiration (Figure 6A). First, we compared the basal OCR of PLN and wild type iPSC-CMs cultured in maturation medium, which represents the mitochondrial function at the baseline level. We observed a comparable basal OCR between the PLN and the wild type groups (Figure 6B and 6C). After blocking FAO using ETO to drive iPSC-CMs to rely on non-fatty acid oxidation, the OCR rates remained comparable between the two groups (Figure 6C and 6D), suggesting a similar FAO-dependence. By blocking glycolysis with 2-DG, we observed some increase in OCR of wild type iPS-CMs, whereas the OCR of PLN iPSC-CMs continued to decrease and became significantly lower than the OCR of wild type iPSC-CMs (Figure 6B and 6E). This observation suggests that the wild type iPSC-CMs are less dependent on glycolysis and have more metabolic flexibility when cultured under maturation medium conditions.

Similarly, in the glucose-rich medium, both the basal OCR and the decrease in OCR after ETO-inhibited FAO were comparable between the PLN and the wild type groups (Figure 6F, 6G, and 6H). After 2-DG-inhibited glycolysis, the OCR was significantly higher in wild type iPSC-CMs than PLN iPSC-CMs, once again suggesting more metabolic flexibility in wild type iPSC-CMs (Figure 6F, 6G, and 6I).

In the lipid-rich medium, the basal OCR was significantly lower in PLN iPSC-CMs than in the wild type (Figure 6J and 6K), implying that FAO was impaired in the PLN iPSC-CMs. However, FAO dependency was not significantly different. OCR of both PLN and wild type iPSC-CMs continued to decrease after ETO-inhibited FAO and 2-DG-inhibited glycolysis, but wild type iPSC-CMs consistently showed higher OCR rate when compared with PLN iPSC-CMs (Figure 6K, 6L, and 6M). These findings indicate that in conditions of excess lipids in the medium, PLN iPSC-CMs are less efficient in FAO, are more glycolysis dependent and metabolic flexibility is lower in PLN deficient iPSC-CMs similar to all the other medium conditions.

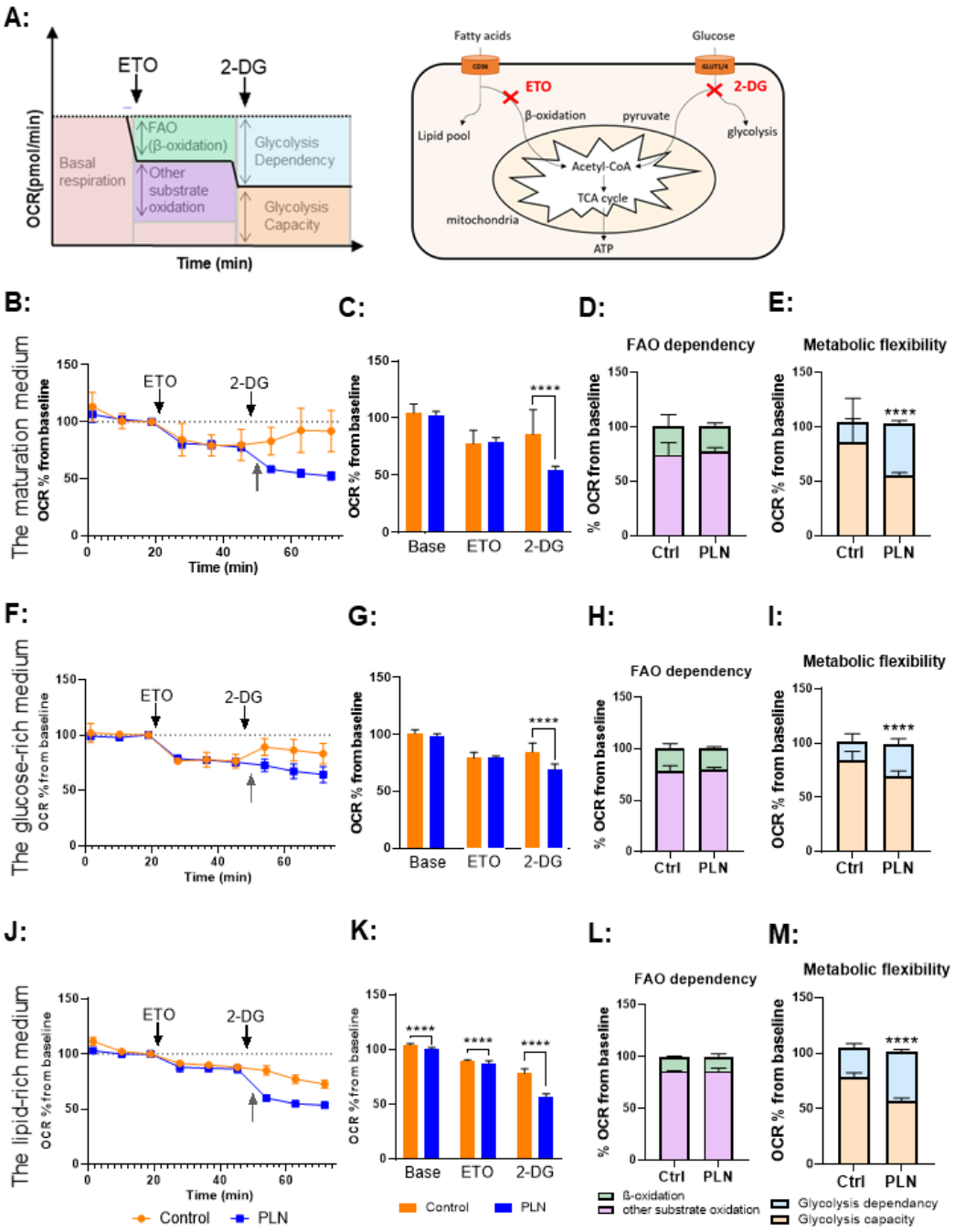


Figure 6: A) Seahorse XF24 Extracellular Flux assay measuring the oxygen consumption rate (OCR) to study the activity of the fatty acid oxidation (FAO) in PLN and wild type iPSC-CMs by manipulating the fatty acid oxidation and glucose metabolism using etomoxir (ETO) and 2-deoxyglucose (2-DG). **B)** OCR of iPSC-CMs cultured in the maturation medium (grey arrow indicates a switch in energy substrates). **C)** Quantification of basal, β-oxidation and glycolysis OCRs in PLN and wild type iPSC-CMs normalised to nuclei count. **D)** FAO dependency of cells cultured in the maturation medium. The degree of FAO dependency is determined by the reduction of mitochondrial function after ETO injection. **E)** Metabolic flexibility of cells cultured in the maturation medium. The degree of metabolic flexibility is determined by the OCR after 2-DG injection. **F)** OCR of iPSC-CMs cultured in the glucose-rich medium. **G)** Quantification of basal, β-oxidation and glycolysis OCRs in PLN and wild type iPSC-CMs normalised to nuclei count.

H) FAO dependency of cells cultured in the glucose-rich medium. **I)** Metabolic flexibility of cells cultured in the glucose-rich medium. **J)** OCR of iPSC-CMs cultured in the lipid-rich medium. **K)** Quantification of basal, β-oxidation and glycolysis OCRs in PLN and wild type iPSC-CMs normalised to nuclei count. **L)** FAO dependency of cells cultured in the lipid-rich medium. **M)** Metabolic flexibility of cells cultured in the lipid-rich medium. Data are presented as mean (±SD), biological replicates are 3 individual differentiations with each N= 5-10 wells. All P-values were calculated by using one-way ANOVA or Student's t-tests. *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001.

PLN iPSC-CMs show higher glucose dependency

We also studied glucose metabolism by comparing ECAR between the two groups. Similarly, we first drove the cells towards glucose utilization by blocking FAO using ETO, ensuring that the increased level of ECAR represents the maximum level of glycolysis. Next, we depleted glycolysis using 2-DG. The reduction of ECAR from the maximum level represents the glycolysis capacity (Figure 7A).

In the maturation medium, we showed that the ECAR rate was significantly lower in PLN versus wild type iPSC-CMs after either ETO-inhibited FAO or 2-DG-inhibited glycolysis (Figure 7B and 7C). After 2-DG-inhibited glycolysis, the decrease in ECAR in PLN versus wild type iPSC-CMs was particularly profound (Figure 7 and 7C). This indicates that compared to the wild type, PLN iPSC-CMs has a lower glycolysis capacity, and a higher dependency on glycolysis in the absence of other energetic substrates (Figure 7D). PLN iPSC-CMs also showed a higher glycolytic reserve than wild type iPSC-CMs ((Figure 7E). Combined, these data suggested that an increased glucose dependency and a decreased metabolic flexibility in PLN iPSC-CMs.

In glucose-rich medium, ECAR levels remained comparable between two groups after either ETO-inhibited FAO or 2-DG-inhibited glycolysis (Figure 7F and 7G), However, the decrease in ECAR after glycolysis inhibition was more profound in PLN versus wild type iPSC-CMs, once again indicating a higher glycolysis dependence and a higher glycolytic reserve of the PLN iPSC-CMs (Figure 7H and 7I).

In lipid-rich medium, PLN iPSC-CMs had a higher ECAR level after ETO-inhibited FAO and showed a significant ECAR reduction upon glycolysis inhibition when compared with the wild type (Figure 7J and 7K) resulting in a higher glycolysis dependency in PLN iPSC-CMs (Figure 7L). PLN iPSC-CMs also had a higher glycolytic reserve than the wild type (Figure 7M). These findings suggested that in the presence of lipids and glucose, PLN iPSC-CMs favored glucose as the substrate for mitochondrial respiration.

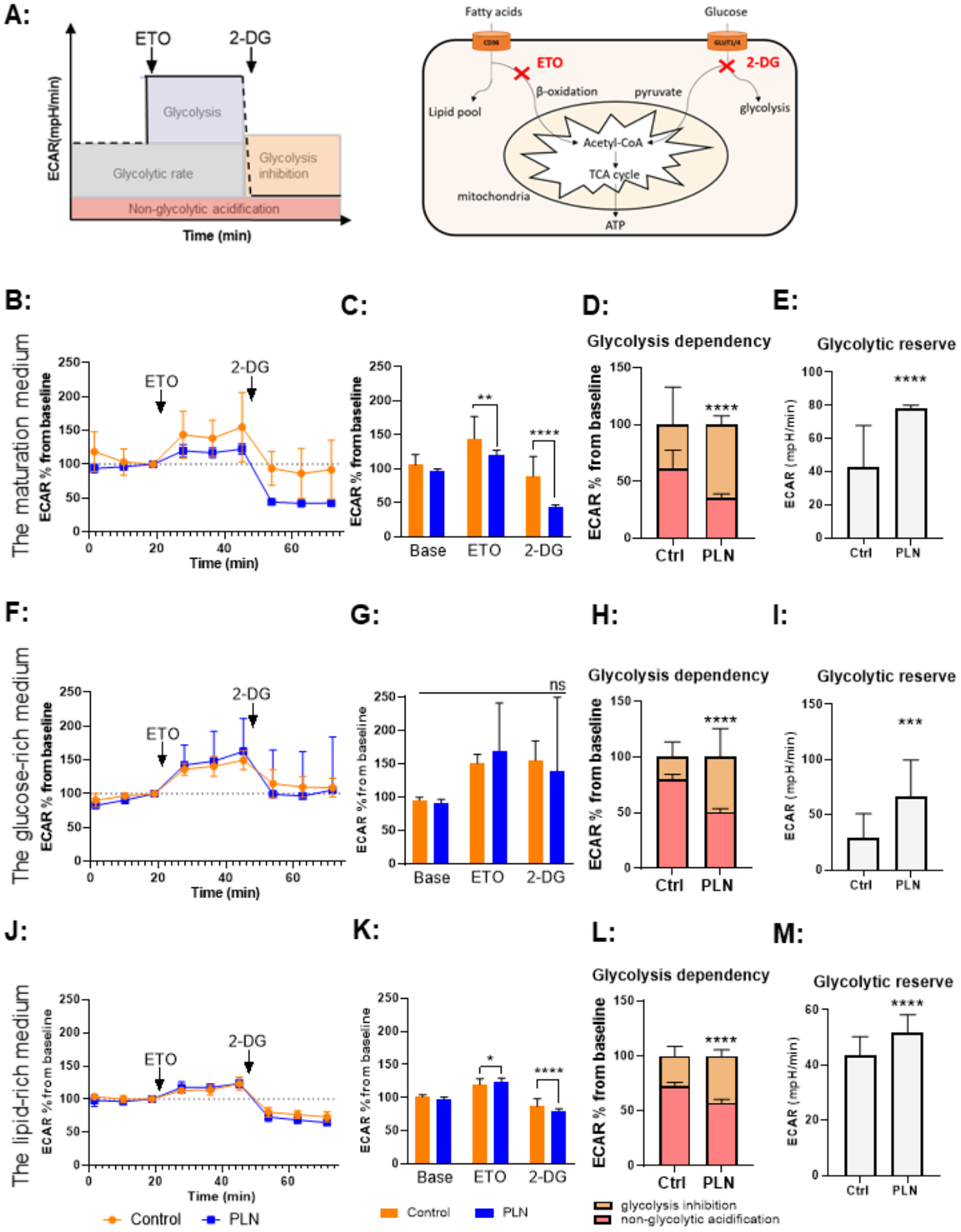


Figure 7: A) Seahorse XF24 Extracellular Flux assay measuring extracellular acidification rate (ECAR) to study the

activity of the glycolytic pathway in PLN and wild type iPSC-CMs by manipulating the fatty acid oxidation and glucose metabolism using etomoxir (ETO) and 2-deoxyglucose (2-DG). **B)** ECAR of iPSC-CMs cultured in the maturation medium (grey arrow indicates a switch in energy substrates). **C)** Quantification of ECAR at the baseline level, after ETO injection, and after 2-DG injection, values were normalized to nuclei count. **D)** Glycolysis dependency of cells cultured in the maturation medium. The degree of glycolysis dependency is determined by the total glycolysis minus the non-glycolytic acidification (after 2-DG injection). **E)** Glycolytic reserve of cells cultured in the maturation medium. The glycolytic reserve ability is determined by ECAR after ETO injection minus ECAR after 2-DG injection. **F)** ECAR of iPSC-CMs cultured in the glucose-rich medium. **G)** Quantification of ECAR at the baseline level, after ETO injection, and after 2-DG injection, values were normalized to nuclei count. **H)** Glycolysis dependency of cells cultured in the glucose-rich medium. **I)** Glycolytic reserve of cells cultured in the glucose-rich medium. **J)** ECAR of iPSC-CMs cultured in the lipid-rich medium. **K)** Quantification of ECAR at the baseline level, after ETO injection, and after 2-DG injection, values were normalized to nuclei count. **L)** Glycolysis dependency of cells cultured in the lipid-rich medium. **M)** Glycolytic reserve of cells cultured in the lipid-rich medium. Data are expressed as mean ±SD, biological replicates are 3 individual differentiations with each N= 5-12 wells, *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001, One-way ANOVA with Tukey's post-hoc comparison or Student's t-tests were used to compute statistical significance from the combined experiments.

Intracellular lipid droplet accumulation is a feature of PLN R14del iPSC-CMs

We also examined lipid accumulation in PLN and wild type iPSC-CMs. First, we demonstrated the presence of cardiomyocyte sarcomeres and mitochondrial in both groups cultured using α-actinin (green) and Tomm20 (red, Figure 8A), respectively. The nuclei of iPSC-CMs were stained by Hoechst and shown in blue. Both PLN and wild type iPCS-CMs showed well-developed cellular structure. Next, we examined the lipid accumulation in PLN and wild type iPSC-CMs cultured in maturation, glucose rich, and lipid rich media. We observed more accumulated lipids in PLN iPSC-CMs when compared with the wild type ones regardless of the type of culture medium (Figure 8B). However, this lipid accumulation in PLN iPSC-CMs became much more extensive when the cells were cultured in the glucose-rich medium and the lipid-rich medium. Notably, we further included a PLN-corrected isogenic iPSC-CMs as an additional control group to study the direct association between PLN R14del and lipid accumulation. We found that the level of lipid accumulation was more profound in PLN iPSC-CMs when compared to the wild type and the corrected isogenic iPSC-CMs (Figure 8C and 8D).

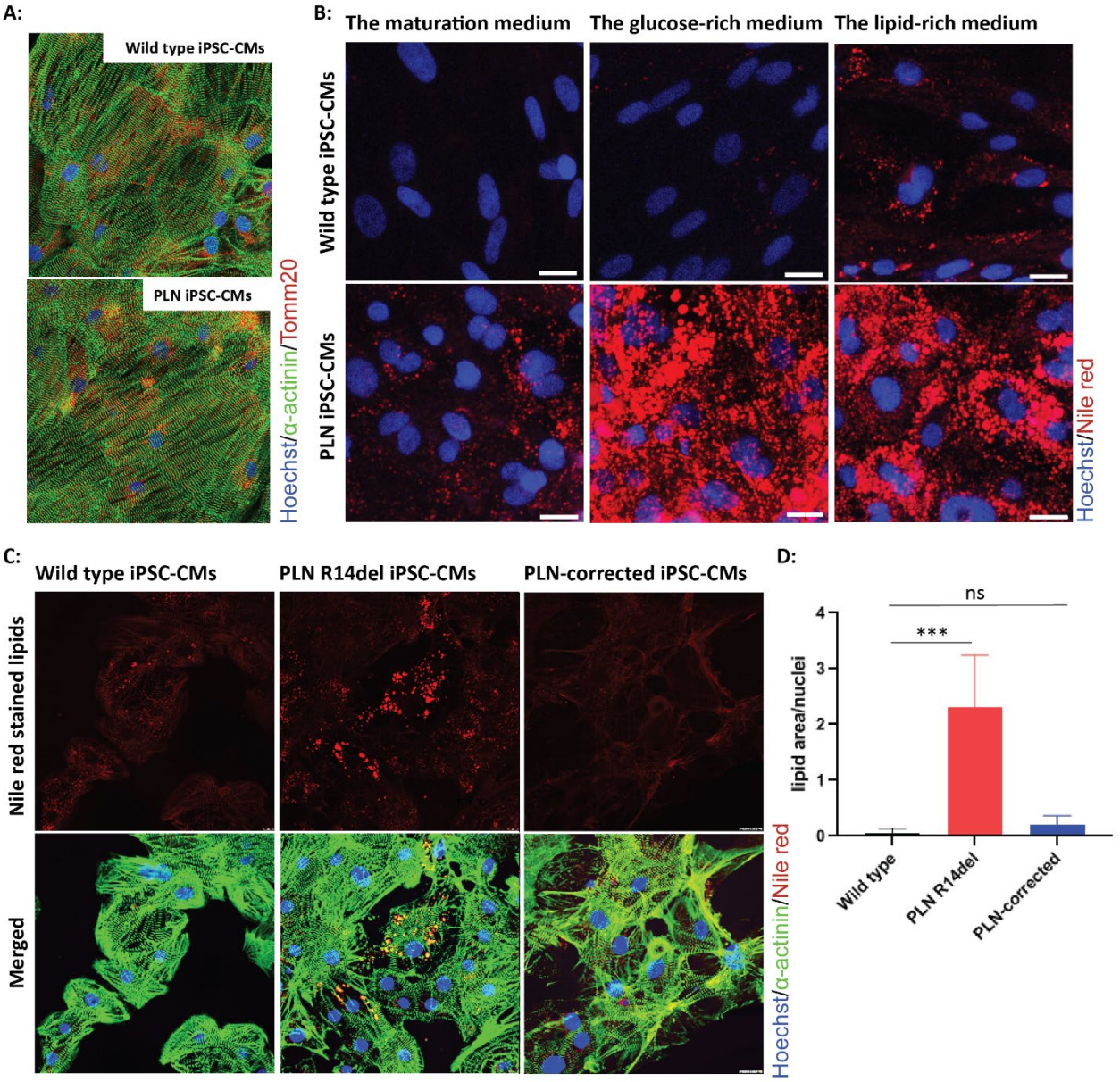


Figure 8: A) Examination of the iPSC-CMs structure. **B)** Nile red staining showing the lipid accumulation in PLN iPSC-CMs and wild type iPSC-CMs cultured in three different media. **C)** Nile red staining showing the lipid accumulation in PLN iPSC-CMs, wild type iPSC-CMs, and PLN-corrected isogenic iPSC-CMs. **D)** Quantification of intracellular lipid formation in wild type, PLN R14el, and PLN-corrected isogenic iPSC-CMs (n ≥ 4 per group). Immunofluorescence images were taken at 63x magnification. Data were normalized to nuclei and shown as mean ±SD, and one-way ANOVA was used to compute statistical significance (***P<0.001, ns= not significantly different).

DISCUSSION

This study provides unprecedented information on changes of chromatin activity and global transcriptional regulation in the end-stage hearts from cardiomyopathy patients carrying PLN R14del mutation when compared with the healthy donor hearts and hearts with other types of cardiomyopathies. Integrating these multi-omics data in a hypothesis-free manner, we showed an inhibited lipid metabolism in PLN hearts and identified several key transcription factors (e.g. *PPARA* and *KLF15*), which may play critical roles in regulating the altered metabolism. Besides the information obtained from the end-stage hearts, we compared the transcriptome profiles between PLN and control iPSC-CMs, which allowed us to study the biological changes due to PLN R14del in the early stage. Consistently, we observed the decrease in metabolic-related processes in PLN iPSC-CMs when compared to wild type iPSC-CMs. To further characterize the signaling at the functional level, we performed multiple Seahorse assays to measure the mitochondrial FAO function and the glycolysis pathways in PLN and wild type iPSC-CMs. We showed that PLN iPSC-CMs had a lower FAO and showed a preference for using glucose over fatty acids when compared with the wild type cells. Additionally, the metabolic flexibility was lower in PLN versus wild type iPSC-CMs. We also detected an increased amount of lipid droplet formation in PLN iPSC-CMs than the wild type using Nile red staining. Strikingly, PLN-corrected isogenic iPSC-CMs showed attenuation of lipid accumulation.

It has been shown that regulatory regions display a high level of tissue specificity and that enrichment of GO terms based on the annotation to genes with the closest TSS similarly demonstrate cell-specific patterns.[39,40] Here we present for the first time, data on cardiac gene regulation in patients with a PLN R14del mutation. Based on our analysis of the chromatin activity in heart tissue from both annotated genes in the vicinity of differentially acetylated regions as well as their TFBM enrichment, the findings point towards known pathways involved in end-stage cardiomyopathies, such as fibrosis, cell death or impairment of muscle function and lipid (energy substrate) metabolism.[41,42,43] We have annotated several notable genes to differentially acetylated regions. For example, the regulatory region in the vicinity of *FST* had the strongest observed hyperacetylation status in our dataset. It is associated with extracellular matrix-related and calcium-binding proteins and its expression is altered during heart failure.[44] Differentially acetylated regions were detected also within ±5kb from TSS of *CHKB, DES, FHL2, HADHA/HADHB, KCNQ1, MLYCD, MYH7, MYL3, PDLIM3, PLEC*, and *PNPLA2*, and binding motifs were enriched for TFs *GATA4* and *POU1F1*, which are genes previously were shown to cause genetic forms of dilated cardiomyopathy.[45] It has to be noted we have not detected a significant difference between the

level of H3K27ac binding in the PLN gene itself, nor its binding partners such as *ATP2A1, ATP2A2, PRKG1, PRKG2* or *EDA* (Source: Gene Cards, STRING).

By comparing PLN R14del with diseased hearts without PLN R14del-related cardiomyopathies, we identified a list of PLN R14del-specific regions and annotated several notable genes in these regions, including many known genes involved in lipid metabolism (e.g. *AGPAT2, GLIPR1, HADHA/HADHB, MLYCD, MSMO1, PLPP1, PTGDS*, and *INPP5J21*). In addition, notable extracellular matrix-associated genes involved in fibrosis formation have been found, such as *COL6A3, COL8A1, COL14A1, CFT1, LAMB2, LUM*, and *MMP223*, in addition to genes crucial for sarcomeric contraction: *CNN3, DES, MB, MYH7, MYL3, MYL7*, and *MYLPF19*. We also identified differentially expressed genes in PLN versus control hearts using RNA-seq and showed similar enrichment results, including activated fibrosis formation and suppression of energy metabolism. Notably, many annotated genes from PLN R14del-specific regions also showed significantly differentially expressed mRNA expression between PLN and control hearts, such as *HADHA, HADHB, PTGDS, COL14A1*, and *LUM*.

By integrating H3K27ac ChIP-seq and RNA-seq data, we identified 39 TF coding genes which could play critical roles in regulating the pathological mechanisms in PLN R14del hearts. *KLF15*, one of the identified TF coding gene from the integrative analysis with a suppressed mRNA level, is a key regulatory in the cardiac lipid metabolism.[46] A decrease in FAO and increase in glycolysis are observed in KLF15-insufficient hearts. Furthermore, the direct targets and the direct cooperator of *KFL15*, such as *PDK4* and *PPARA*,[46] in the regulation of FAO have also been identified in our H3K27ac ChIP-seq and/or RNA-seq data.

Apart from these findings in cardiac tissues, *in vitro*, a set of downregulated genes PLN versus wild type iPSC-CMs was able to confirm suppression of specific process involved in metabolic regulation, regardless of the availability of fatty acids or glucose in the culture medium. It is important to note that despite the excessive access to fatty acids in the lipid-rich medium, PLN iPSC-CMs still showed a lower FAO profile than the wild type at the transcriptional level. We could further confirm the decrease in FAO in PLN versus iPSC-CMs at a functional level by measuring mitochondrial respiration, and by demonstrating increased lipid accumulation in PLN iPSC-CMs. Similarly, previous studies also showed inhibition of FAO and accumulated lipids in iPSC-CMs carrying *PKP2* mutation when compared with controls.[47,16] FAO changes serve as the indicator for early adapted or maladapted metabolic response.[48] To meet the energy demand of the heart, fatty acids, especially the free circulating fatty acids, are transported to cardiomyocytes via the translocases like *CD36*, whereas the activated key metabolic regulator *AMPK* promotes the mitochondrial FAO.[49] In impaired cardiomyocytes responding to e.g. saturated fatty acid treatment to

mimic metabolic cardiomyopathy, FAO decreases whereas intracellular lipid accumulation increases, resulting in the lipotoxicity of cardiomyocytes and cell death.[50,51] Interestingly, by activating Cpt1b-mediated FAO, the fatty acid clearance is significantly improved, thereby attenuating the lipotoxicity and related apoptosis.[50] Combined, we hypothesize a vicious loop between the increase in energy cost for Ca2+ regulation due to PLN R14del mutation, the decrease in FAO in the cardiomyocytes, and enhanced cell death due to the intracellular lipid accumulation.

Next to FAO suppression, we showed that PLN R14del iPSC-CMs exhibited a preference for glucose utilization during oxidation. Although healthy adult cardiomyocytes prefer fatty acids over other oxidative substrates (e.g. glucose and lactase) during the energy production,[52] the increase in glucose utilization has been previously shown in cardiomyocytes from hypertrophic cardiomyopathies and failing hearts.[53] It is worth noting that increased glucose metabolism in return inhibits FAO by malonyl coenzyme A-mediated signaling.[54] Two key transporters, *GLUT1* and *GLUT4*, are responsible for the exogenous glucose uptake, followed by rapid phosphorylation of intercellular glucose for glycolysis or being stored as glycogen.[55] To generate the same amount of ATP, glucose metabolism requires less oxygen than fatty acid metabolism, indicating a higher efficiency of using glucose for energy production.[55] However, in the long term, glucose-dependent energy metabolism has been associated with the progression of cardiac dysfunction, suggesting that the switch from fatty acid to glucose utilization might be an early pathological process in the diseased heart.[53] Additionally, like the lipotoxicity, high glucose levels also have deleterious effects on cardiomyocytes by increasing the oxidative stress-related cell death and decreasing the contractile force.[56,57]

Cardiomyocytes show the flexibility in switching energy subtracts under different conditions, such as from glucose-dependent fetal condition into fatty acids-dependent adult condition and from fatty acids/glucose-dependent non-fasting condition to ketones-dependent fasting condition.[53] Similar to PLN R14del related mitochondrial dysfunction, Bround *et al.* showed that cardiomyocytes with reduced mitochondrial metabolism due to the depletion of *RYR2* had lower mRNA levels of key regulators in the fatty acid metabolism (e.g. Sirt1, Foxo1, Ppargc1a, Pparα, Pparγ, and Klf15) as well as glucose metabolism (e.g. Glut4 and Pck1), suggesting impaired flexibility in switching substrates.[58] Thus, reduced metabolic flexibility may be a critical factor for maintaining normal cardiac function and preventing cardiac maladaptation in response to the disease or other stressors.[59]

This metabolic impairment remains difficult to study in PLN R14del cardiomyopathy. A previous study using a murine model demonstrated that DCM hearts carrying another mutation (p.Arg9Cys) in *PLN* showed increasing mRNA levels of profibrotic cytokines and ECM markers from the early to the severe stages. These include TGFβ, connective tissue growth factor, periostin, and thrombospondin.[60] Notably, transcriptional changes in cardiomyocytes from the early stage DCM hearts showed significant repression of genes in aerobic metabolism (e.g. FAO and oxidative phosphorylation) and activation of genes in glucose metabolism when compared with wild type cardiomyocytes, indicating the switch from fatty acid to glucose utilization occurs at a early stage. Furthermore, when comparing cardiomyocytes from the end-stage DCM hearts and controls, PPAR signaling was the most enriched metabolic pathway in the downregulated genes set, including Ppara, Ppargc1a, Rxrg, Klf15. Several PPARA agonists that regulate fatty acid metabolism have been approved by the FDA,[61,62] and have shown protective effects by reducing the progression of cardiomyopathies.[63] Many of these metabolic-related genes also showed suppressed acetylation levels and/or mRNA levels in PLN R14 del hearts and iPSC-CMs in the present study.

Although fat is a normal component of the heart, especially in the epicardium,[64] the precise pathological mechanism underlying fat infiltration related cardiomyopathies remains unclear.[65] A previous study showed the possibility of differentiating adipose-derived stromal vascular fraction cells into cardiomyocyte-like cells.[66] Another study showed the transduction of cardiac mesenchymal progenitors into adipocytes.[67] Additionally, fibroblasts can also be differentiated into adipocytes directly and indirectly, by promoting the differentiation of adipose progenitor cells.[68,69] These studies highlight the complex networks that exist between cardiomyocytes and non-myocyte cell types during metabolic (mal)adaption. Future studies are should focus on investigating cell-to-cell interactions during the cardiac energy rearrangement in PLN R14del cardiomyopathy.

Taken together, we have cataloged valuable information on the chromatin acetylation activities and transcriptome regulations in PLN R14del hearts when compared with controls. We also revealed the transcriptome regulations in PLN R14del iPSC-CMs when compared with controls. Combining these multi-omics data altogether, we identified disturbed energy metabolism in cardiac cells with a PLN R14del mutation, which we further validated using multiple functional assays and have highlighted several key metabolic regulators involved in the process. The in-depth analysis of integrating data using human end-stage PLN hearts combined with *in vitro* findings in early-stage PLN iPSC derived cardiomyocytes presented in this study revealed for the first time the link between PLN R14del mutation, the impairment of FAO, and intracellular lipid accumulation. Additionally, the identified TFs and genes involved in the

metabolic pathway could contribute to the pathological progression of PLN R14del cardiomyopathy, and may serve as promising therapeutic targets for future treatments.

ACKNOWLEDGMENTS

We are very thankful to the patients and their families for providing valuable cardiac tissue for research purposes. We further thank Joyce van Kuik and Erica Sierra-de Koning for technical assistance with tissue processing.

SOURCES OF FUNDING

This work was supported by Dutch PLN patient organization grant Stichting PLN/2015_01 (MH, JMIGH, FWA), PLN foundation (MH, FWA), Leducq grant (CURE_PlaN to JC, DF, IK, MM, PAD, MH, FWA), the NWO VENI grant (no. 016.176.136 to MH), National Institute of Health Ro1 grant LM010098 (MH, FWA), Netherlands Heart Foundation DOSIS grant CVON2014-40 (MH, FWA), Wilhelmina Children's Hospital research funding (MH, MM), Netherlands Foundation for Cardiovascular Excellence (CC), two NWO VIDI grants (no. 91714302 to CC and no. 016096359 to MCV), the ErasmusMC fellowship grant (CC), the RM fellowship grant of the UMC Utrecht (CC), the Netherlands Cardiovascular Research Initiative: An initiative with support of the Dutch Heart Foundation (MCV, CC), the Dutch Heart Foundation (Queen of Hearts to GP, HdR, JP and CC), Netherlands Heart Foundation Dekker scholarship Junior Staff Member grant 2014T001 (FWA), and UCL Hospitals NIHR Biomedical Research Centre grant BRC86A (FWA).

DISCLOSURES

None

Supplementary Files – Online

Supplementary Table 1: A) An overview of cardiac tissues, **B)** An overview of iPSC-CMs

Supplementary Table 2: An overview of H3K27 ChIP-seq data obtained from cardiac tissues in 4 analyses

Supplementary Table 3: Media used for culturing iPSC-CMs

Supplementary Table 4: Antibodies used in iPSC-CMs

Supplementary Table 5: Differentially acetylated regions between PLN and control hearts and their annotated genes.

Supplementary Table 6: A) Gene set enrichment analysis of annotated genes from hyperacetylated regions in PLN versus control hearts, **B)** Gene set enrichment analysis of annotated genes from hypoacetylated regions in PLN versus control hearts

Supplementary Table 7: A) Enriched transcription factor binding motifs in differentially acetylated regions and TFs coding genes, **B)** Gene set enrichment analysis of transcription factor coding genes

Supplementary Table 8: A) Altered regulatory regions that were specific in PLN R14del hearts when compared with the controls and hearts with ischemic or non-ischemic dilated cardiomyopathy, **B)** Annotated genes from the hyper- and the hypoacetylated regions that were PLN-specific when compared with the other groups

Supplementary Table 9: Differentially expressed genes between PLN and control hearts

Supplementary Table 10: A) Gene set enrichment analysis of upregulated genes in PLN versus control hearts, **B)** Gene set enrichment analysis of downregulated genes in PLN versus control hearts

Supplementary Table 11: A) Differentially expression genes between PLN and wild type iPSC-CMs cultured in the maturation medium, **B)** Differentially expression genes between PLN and wild type iPSC-CMs cultured in the glucose-rich medium, **C)** Differentially expression genes between PLN and wild type iPSC-CMs cultured in the lipid-rich medium

Supplementary Table 12: A) Gene set enrichment analysis of differentially expressed genes between PLN and wild type iPSC-CMs cultured in the maturation medium, **B)** Gene set enrichment analysis of differentially expressed genes between PLN and wild type iPSC-CMs cultured in the glucose-rich medium,

C) Gene set enrichment analysis of differentially expressed genes between PLN and wild type iPSC-CMs cultured in the lipid-rich medium

Supplementary Figure 1: A) Masson's trichrome staining of tissues, **B)** Histological quantification of fibrosis and fatty tissues

REFERENCES

1. Haghighi, K. *et al.* A mutation in the human phospholamban gene, deleting arginine 14, results in lethal, hereditary cardiomyopathy. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 1388–1393 (2006).
2. van Rijsingen, I. A. W. *et al.* Outcome in phospholamban R14del carriers: results of a large multicentre cohort study. *Circ. Cardiovasc. Genet.* **7**, 455–465 (2014).
3. van der Zwaag, P. A. *et al.* Phospholamban R14del mutation in patients diagnosed with dilated cardiomyopathy or arrhythmogenic right ventricular cardiomyopathy: evidence supporting the concept of arrhythmogenic cardiomyopathy. *Eur. J. Heart Fail.* **14**, 1199–1207 (2012).
4. Hof, I. E. *et al.* Prevalence and cardiac phenotype of patients with a phospholamban mutation. *Neth. Heart J.* **27**, 64–69 (2019).
5. Doevendans, P. A., Glijnis, P. C. & Kranias, E. G. Leducq Transatlantic Network of Excellence to Cure Phospholamban-Induced Cardiomyopathy (CURE-PLaN). *Circ. Res.* **125**, 720–724 (2019).
6. Karakikes, I. *et al.* Correction of human phospholamban R14del mutation associated with cardiomyopathy using targeted nucleases and combination therapy. *Nat. Commun.* **6**, 6955 (2015).
7. Te Rijdt, W. P. *et al.* Phospholamban p.Arg14del cardiomyopathy is characterized by phospholamban aggregates, aggresomes, and autophagic degradation. *Histopathology* **69**, 542–550 (2016).
8. Gho, J. M. I. H. *et al.* High resolution systematic digital histological quantification of cardiac fibrosis and adipose tissue in phospholamban p.Arg14del mutation associated cardiomyopathy. *PLoS One* **9**, e94820 (2014).
9. Te Rijdt, W. P. *et al.* Distinct molecular signature of phospholamban p.Arg14del arrhythmogenic cardiomyopathy. *Cardiovasc. Pathol.* **40**, 2–6 (2019).
10. Sepehrkhouy, S. *et al.* Distinct fibrosis pattern in desmosomal and phospholamban mutation carriers in hereditary cardiomyopathies. *Heart Rhythm* **14**, 1024–1032 (2017).
11. Justus M. B. Anumonwo, T. H. Fatty Infiltration of the Myocardium and Arrhythmogenesis: Potential Cellular and Molecular Mechanisms. *Front. Physiol.* **9**, (2018).
12. Te Rijdt, W. P. *et al.* Myocardial fibrosis as an early feature in phospholamban p.Arg14del mutation carriers: phenotypic insights from cardiovascular magnetic resonance imaging. *Eur. Heart J. Cardiovasc. Imaging* **20**, 92–100 (2019).
13. Kumar, D. & Elliott, P. *Principles and Practice of Clinical Cardiovascular Genetics*. (Oxford University Press, USA, 2010).
14. Molina, P. *et al.* Arrhythmogenic cardiomyopathy with left ventricular involvement versus ischemic heart disease: lessons learned from the family study and the reviewed autopsy of a young male. *Forensic Sciences Research* **4**, 274 (2019).
15. Morales, A. *et al.* Late onset sporadic dilated cardiomyopathy caused by a cardiac troponin T mutation. *Clin. Transl. Sci.* **3**, 219–226 (2010).

16. Kim, C. *et al.* Studying arrhythmogenic right ventricular dysplasia with patient-specific iPSCs. *Nature* **494**, 105 (2013).

17. Shah, K. *et al.* Modeling Arrhythmogenic Right Ventricular Dysplasia/Cardiomyopathy with Patient-Specific iPSCs. in *Human iPS Cells in Disease Modelling* 27–43 (Springer, Tokyo, 2016).

18. Osorio, J. C. *et al.* Impaired Myocardial Fatty Acid Oxidation and Reduced Protein Expression of Retinoid X Receptor-α in Pacing-Induced Heart Failure. *Circulation* vol. 106 606–612 (2002).

19. Carvajal, K. & Moreno-Sánchez, R. Heart Metabolic Disturbances in Cardiovascular Diseases. *Archives of Medical Research* vol. 34 89–99 (2003).

20. Dávila-Román, V. G. *et al.* Altered myocardial fatty acid and glucose metabolism in idiopathic dilated cardiomyopathy. *J. Am. Coll. Cardiol.* **40**, 271–277 (2002).

21. Kolwicz, S. C., Jr, Purohit, S. & Tian, R. Cardiac metabolism and its interactions with contraction, growth, and survival of cardiomyocytes. *Circ. Res.* **113**, 603–616 (2013).

22. Kolwicz, S. C., Jr & Tian, R. Metabolic Therapy at the Crossroad: How to optimize myocardial substrate utilization? *Trends Cardiovasc. Med.* **19**, 201 (2009).

23. Son, N. H. *et al.* PPARγ-induced Cardiolipotoxicity in Mice Is Ameliorated by PPARα Deficiency Despite Increases in Fatty Acid Oxidation. *J. Clin. Invest.* **120**, (2010).

24. Austin, K. M. *et al.* Molecular mechanisms of arrhythmogenic cardiomyopathy. *Nat. Rev. Cardiol.* **16**, 519–537 (2019).

25. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).

26. Ji, H. *et al.* An integrated software system for analyzing ChIP-chip and ChIP-seq data. *Nat. Biotechnol.* **26**, 1293–1300 (2008).

27. Huber, W. *et al.* Orchestrating high-throughput genomic analysis with Bioconductor. *Nat. Methods* **12**, 115–121 (2015).

28. Rosenbloom, K. R. *et al.* ENCODE data in the UCSC Genome Browser: year 5 update. *Nucleic Acids Res.* **41**, D56–63 (2013).

29. McLeay, R. C. & Bailey, T. L. Motif Enrichment Analysis: a unified framework and an evaluation on ChIP data. *BMC Bioinformatics* vol. 11 (2010).

30. Chen, J., Bardes, E. E., Aronow, B. J. & Jegga, A. G. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res.* **37**, W305–11 (2009).

31. Franceschini, A. *et al.* STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* **41**, D808–15 (2013).

32. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

33. Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. & Prins, P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* vol. 31 2032–2034 (2015).

34. Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).

35. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

36. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

37. Sharma, A. *et al.* High-throughput Screening of Tyrosine Kinase Inhibitor Cardiotoxicity With Human Induced Pluripotent Stem Cells. *Sci. Transl. Med.* **9**, (2017).

38. McCulley, D. J. & Black, B. L. Transcription Factor Pathways and Congenital Heart Disease. *Current Topics in Developmental Biology* 253–277 (2012) doi:10.1016/b978-0-12-387786-4.00008-7.

39. Sonawane, A. R. *et al.* Understanding Tissue-Specific Gene Regulation. *Cell Rep.* **21**, 1077 (2017).

40. Lee, B.-K. *et al.* Cell-type specific and combinatorial usage of diverse transcription factors revealed by genome-wide binding studies in multiple human cells. *Genome Res.* **22**, 9 (2012).

41. Fukushima, A., Milner, K., Gupta, A. & Lopaschuk, G. D. Myocardial Energy Substrate Metabolism in Heart Failure : from Pathways to Therapeutic Targets. *Curr. Pharm. Des.* **21**, 3654–3664 (2015).

42. Harvey, P. A. & Leinwand, L. A. Cellular mechanisms of cardiomyopathy. *The Journal of Cell Biology* vol. 194 355–365 (2011).

43. Louzao-Martinez, L. *et al.* Characteristic adaptations of the extracellular matrix in dilated cardiomyopathy. *Int. J. Cardiol.* **220**, 634–646 (2016).

44. Lara-Pezzi, E. *et al.* Expression of Follistatin-Related Genes Is Altered in Heart Failure. *Endocrinology* vol. 149 5822–5827 (2008).

45. Harakalova, M. *et al.* A systematic analysis of genetic dilated cardiomyopathy reveals numerous ubiquitously expressed and muscle-specific genes. *Eur. J. Heart Fail.* **17**, 484–493 (2015).

46. Prosdocimo, D. A. *et al.* Kruppel-like Factor 15 Is a Critical Regulator of Cardiac Lipid Metabolism. *J. Biol. Chem.* **289**, (2014).

47. Caspi, O. *et al.* Modeling of Arrhythmogenic Right Ventricular Cardiomyopathy With Human Induced Pluripotent Stem Cells. *Circ. Cardiovasc. Genet.* **6**, (2013).

48. Rech, M. *et al.* Assessing fatty acid oxidation flux in rodent cardiomyocyte models. *Sci. Rep.* **8**, 1–6 (2018).

49. Energetic metabolism in cardiomyocytes: molecular basis of heart ischemia and arrhythmogenesis. https://vpjournal.net/article/view/2339 doi:10.20517/2574-1209.2017.34.

50. Haffar, T., Bérubé-Simard, F. & Bousette, N. Impaired Fatty Acid Oxidation as a Cause for Lipotoxicity in Cardiomyocytes. *Biochem. Biophys. Res. Commun.* **468**, (2015).

51. Palmitate-mediated Alterations in the Fatty Acid Metabolism of Rat Neonatal Cardiac Myocytes. *J. Mol. Cell. Cardiol.* **32**, 511–519 (2000).

52. Piper, H. M., Siegmund, B. & Spahr, R. Energy Metabolism of Adult Cardiomyocytes. in *Myocardial Energy Metabolism* 171–180 (Springer, Dordrecht, 1988).

53. Florencia Pascual, R. A. C. Fuel Availability and Fate in Cardiac Metabolism: A Tale of Two Substrates. *Biochim. Biophys. Acta* **1860**, 1425 (2016).

54. van Weeghel, M. *et al.* Increased Cardiac Fatty Acid Oxidation in a Mouse Model With Decreased malonyl-CoA Sensitivity of CPT1B. *Cardiovasc. Res.* **114**, (2018).

55. Targeting mitochondrial oxidative metabolism as an approach to treat heart failure. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1833**, 857–865 (2013).

56. Davargaon, R. S., Sambe, A. D. & S, M. V. V. Toxic Effect of High Glucose on Cardiomyocytes, H9c2 Cells: Induction of Oxidative Stress and Ameliorative Effect of Trolox. *J. Biochem. Mol. Toxicol.* **33**, (2019).

57. Taegtmeyer, H., Sen, S. & Vela, D. Return to the fetal gene program. *Annals of the New York Academy of Sciences* vol. 1188 191–198 (2010).

58. Bround, M. J. *et al.* Cardiomyocyte ATP Production, Metabolic Flexibility, and Survival Require Calcium Flux through Cardiac Ryanodine Receptors in Vivo. *J. Biol. Chem.* **288**, 18975 (2013).

59. Taegtmeyer, H., Golfman, L., Sharma, S., Razeghi, P. & van Arsdall, M. Linking gene expression to function: metabolic flexibility in the normal and diseased heart. *Ann. N. Y. Acad. Sci.* **1015**, 202–213 (2004).

60. Burke, M. A. *et al.* Molecular profiling of dilated cardiomyopathy that progresses to heart failure. *JCI Insight* **1**, (2016).

61. Hong, F., Xu, P. & Zhai, Y. The Opportunities and Challenges of Peroxisome Proliferator-Activated Receptors Ligands in Clinical Drug Discovery and Development. *Int. J. Mol. Sci.* **19**, (2018).

62. Home, P. Safety of PPAR Agonists. *Diabetes Care* **34**, S215 (2011).

63. Khuchua, Z., Glukhov, A. I., Strauss, A. W. & Javadov, S. Elucidating the Beneficial Role of PPAR Agonists in Cardiac Diseases. *Int. J. Mol. Sci.* **19**, (2018).

64. Rabkin, S. W. Epicardial Fat: Properties, Function and Relationship to Obesity. *Obes. Rev.* **8**, (2007).

65. Pantanowitz, L. Fat infiltration in the heart. *Heart* **85**, 253–253 (2001).

66. Yang, G. *et al.* Obtaining spontaneously beating cardiomyocyte-like cells from adipose-derived stromal vascular fractions cultured on enzyme-crosslinked gelatin hydrogels. *Sci. Rep.* **7**, 1–11 (2017).

67. Kami, D., Kitani, T., Kawasaki, T. & Gojo, S. Cardiac mesenchymal progenitors differentiate into adipocytes via Klf4 and c-Myc. *Cell Death Dis.* **7**, e2190–e2190 (2016).

68. Wu, W., Jin, Y.-Q. & Gao, Z. Directly reprogramming fibroblasts into adipogenic, neurogenic and hepatogenic differentiation lineages by defined factors. *Exp. Ther. Med.* **13**, 2685 (2017).

69. Ejaz, A. *et al.* Fibroblast feeder layer supports adipogenic differentiation of human adipose stromal/progenitor cells. *Adipocyte* **8**, 178–189 (2019).

Chapter 9

Summary and General discussion

Summary and General discussion

This chapter summarizes the findings presented in the previous chapters and discusses their application for a better understanding of the pathological mechanisms and identifying the driving factor(s) per heart failure (HF) etiology.

Summary

The major risk factors concerning acquired HF were addressed in **Chapters 2**, **3**, and **4**. Chronic kidney disease (CKD) occurs in 26% to 53% of the HF with preserved ejection fraction (HFpEF) population, and the subclinical diastolic dysfunction appears to be the most common echocardiographic feature in asymptomatic CKD patients on hemodialysis,[1,2] indicating a tightly linked cardiac and kidney functions. Furthermore, elderly females with increased urinary albumin excretion or cystatin C are more vulnerable to develop HFpEF than the rest population.[3] In **Chapter 2**, we focused on the microvasculature, which plays an important role in the kidney and cardiac interactions, and conducted a meta-analysis on the protective impact of estradiol (E2) on the endothelium against the uremic toxins (UTs) induced deleterious effects due to the kidney disease. By collecting the transcriptional changes of human-derived endothelial cells (ECs) treated by E2, indoxyl sulfate (IS), or a high concentration of inorganic phosphate (HP) from published studies using the microarray approach, we identified 188 genes that were increased by UTs (either IS and/or HP) and decreased by E2 and 572 genes that were decreased by UTs and increased by E2. Gene set enrichment analyses using the identified genes suggested that E2 could protect the (myo)endothelium by inhibiting UTs-induced inflammation and facilitating angiogenesis in injured vessels.

In **Chapter 3**, we looked into the IS-treated ECs and studied alterations in the transcriptome using RNA sequencing (RNA-seq), which is a more sensitive technique in revealing the gene expression profile than microarray analysis.[4] Enriched pathways that were identified by affected genes were involved in altered vascular formation and cell metabolism. We further examined the effect of IS on angiogenesis and cell metabolism at the functional level. Consistent with the transcriptome analysis, we also observed activated angiogenic response, decreased viability, and increased cell senescence in IS-treated ECs. Additionally, we showed that by blocking the mRNA expression level of *CYP1B1*, the top 1 up-regulated gene in IS-treated ECs, the pro-angiogenic response of IS was attenuated. IS activated reactive oxygen species (ROS) production, which promotes angiogenesis,[5] was also diminished after silencing *CYP1B1* expression. Combined, these findings suggested a critical role of *CYP1B1* in IS-activated angiogenic response.

Chapter 4 profiles the histone acetylome changes in remodeled non-failing human hearts due to severe aortic stenosis. We presented a novel dataset of chromatin regulatory regions that showed different acetylation levels between the disease hearts and the healthy controls. Furthermore, we identified transcription factor (TF) binding motifs that were enriched in these differentially acetylated regions and annotated a list of TF that bind to these motifs. Out of these, we also showed that 64 TF-encoding genes were differentially expressed at the mRNA level using RNA-seq. Therefore, they could present as promising upstream regulators in mediating biological signals upon myocardial remodeling.

Chapters 5 and **6** focus on inherited HCM carrying the *MYBPC3* mutations from multiple perspectives, namely the DNA, RNA, and protein levels. In **Chapter 5**, we first revealed the histone acetylation landscape and the transcriptome profile in the HCM versus the control hearts. Next, we employed two independent analyses to process the DNA and the mRNA datasets. By integrating both analyses, we identified a list of genes that played a key role in regulating the biological processes in HCM hearts. Likewise, we also compared differentially expressed proteins between the HCM and the control groups. Connecting the changed proteins to the list of genes obtained from the integrative analysis, we provided 53 candidates that could serve as promising targets to restore the altered cardiac functions in HCM hearts.

Chapter 6 further focused on the genotype-to-phenotype relationship in HCM at the proteomics and functional levels. In line with the findings from our previous studies and other published studies, differentially expressed proteins in HCM *versus* the healthy control hearts indicated activated fibrotic formation and inhibited metabolic pathways. Notably, within HCM hearts, we compared those which carry sarcomere mutations with those without known mutations. Increased tubulin and an altered posttranslational modification pattern were observed in mutation-positive *versus* mutation-negative HCM hearts. We further investigated the altered tubulin using the HCM mouse model and showed that decreased tubulin detyrosination normalized the contraction and relaxation of cardiomyocytes. Combined, these findings suggest that the tubulin network could play a role in the pathological mechanism in HCM hearts with a sarcomere mutation.

In Chapter 7, we studied the association of 52 general genetic loci, which affect myocardial mass and lead to an abnormal QRS complex on the electrocardiogram,[6] with the chromatin histone regulation in HCM hearts. By integrating genome, regulome, and transcriptome information, we observed that QRS-related loci had a higher correlation with differentially acetylated regions, particularly the promoter regions, than the non-differentially acetylated regions. Next, we collected a set of 4,620 candidate QRS-related causative single nucleotide polymorphisms (SNPs) in high linkage disequilibrium (LD) with the 52 loci. Out

of these, 74 SNPs co-localized with the leading loci in the differentially acetylated regions. By overlapping genes located next to these 74 SNPs with the transcriptome profile, we identified 25 genes that were differentially expressed between HCM versus control hearts, highlighting this integrative analytic approach in identifying potential causative SNPs and their influences on gene transcription.

Chapter 8 investigated the chromatin transcriptional regulation in inherited DCM carrying the PLN-R14del mutation. Using cardiac tissue of PLN-R14del patients and healthy donors, we first identified differentially acetylated regions (including promoters and enhancers) and studied the enriched TF binding motifs in these regions. We also obtained differentially expressed genes between the PLN-R14del and control hearts. By integrating the histone regulome and the transcriptome data, we identified key TF regulators in the fatty acid oxidation (FAO) metabolism. Interestingly, previous studies from our group and others also demonstrated the typical fibrofatty replacement in PLN-R14del hearts, suggesting an altered FAO metabolism. Next, we validated the inhibited metabolism and accumulated lipid droplets in human induced pluripotent stem cell-derived cardiomyocytes (hiPSC-CMs) from a PLN-R14del patient versus hiPSC-CMs from a healthy donor at the transcriptional and the protein levels. Furthermore, we performed multiple Seahorse assays to compare the glycolysis and FAO in the PLN-R14del and the control hiPSC-CMs cultured in media with a variable composition of lipid and glucose, two major sources of energy substrate for healthy cardiomyocytes. We observed that PLN-R14del hiPSC-CMs had higher glycolysis dependent and less metabolic flexibility in switching the energy substrates as compared to the controls. Combined, these findings provide several options for biomarkers to monitor disease progression as well as treatment in PLN R14del, and beyond.

Discussion

Based on the knowledge learned in all chapters, the possible improvements and future directions in the field of HF are discussed briefly, which may benefit the next-generation clinics for individuals who have HF or at a high risk of developing HF.

From the tissue to the single-cell level

We performed bulk sequencing to study the regulome, transcriptome, and/or proteome changes in diseased versus healthy hearts (**Chapter 4-6** and **Chapter 9**). The heart is composed of multiple cell types, among which cardiomyocytes, endothelial cells (ECs), and fibroblasts are the most abundant cell types in the healthy adult hearts.[7] A previous study showed 24 cell types with distinct transcriptome profiles in the mouse heart using single-cell RNA sequencing (scRNA-seq).[8] Another scRNA-seq study revealed 9 major cell types and 20 subclusters of cell types in the human heart based on their cellular transcriptome profiles.[9] Multiple subtypes have also been identified in the same cell type in the heart, including cardiomyocytes and ECs.[10,11] In the diseased hearts, myocardium remodels and cellular composition changes, resulting in for example fewer cardiomyocytes (myocyte necrosis and apoptosis) and more fibroblasts (increased fibrosis).[12] Therefore, dissecting the cell-type-specific signals could provide valuable insight into the diseases.

Using scRNA-seq and H3K27ac ChIP-seq, a previous study identified the upstream TFs ELK1 and NRF1/2 that played an important role in regulating the cardiomyocyte hypertrophy in murine hearts with pressure overload.[13] In line with this finding, we also showed that both *ELK1* and *NRF1* were annotated to the enriched TFBMs by the hyperacetylated chromatin regions in human pressure overloaded hearts caused by aortic stenosis when compared with controls (**Chapter 4**). Martini and colleagues employed scRNA-seq to study the immune responses in pressure overloaded hearts, and they found activated genes that were specific for the immune cell subpopulations, including the upregulated *CCR2* in macrophages, *CXCR5* in B cells, *PDCD1* in T cells, and *IL6* in mast cells. We also obtained enriched inflammatory responses in the upregulated genes of aortic stenosis-related hearts versus controls. However, we could not identify signals that indicated the activation of specific immune cell subpopulations in our bulk sequencing data (**Chapter 4**). A recent scRNA-seq study showed DNA damage response served as the hallmark in failing cardiomyocytes from DCM hearts versus normal cardiomyocytes from control hearts.[14] Consistently, enrichment in DNA damage response was identified in our bulk sequencing data comparing DCM hearts

against controls (**Chapter 8**). Apart from cardiomyocytes, Li *et al.* employed scRNA-seq and identified a subpopulation of cardiac-specific resident ECs that were leading the angiogenic responses after myocardial infarction.[15] We observed pro-angiogenic responses in ECs treated with uremic toxins (**Chapter 3**). By connecting the transcriptome profile of the cardiac-specific resident ECs revealed by their study with the transcriptome profile shown in ours, we may obtain a list of genes that are highly relevant for mediating the neovasculogenesis in injured vessels from patients with the cardiorenal syndrome for future studies.

Other research based on the scRNA-seq approach has shown the novel pro-fibrotic marker *CAP4* (also known as *CASP8*) to be exclusively upregulated in active fibroblasts when compared with other cell types in ischemic hearts.[16] The mRNA expression level of *CASP8* was significantly higher in hypertrophic hearts due to aortic stenosis and dilated hearts with PLN R14del when compared with controls (**Chapter 4** and **Chapter 8**). Whereas, the mRNA expression level of *CASP8* was compatible between hypertrophic hearts carrying mutated MYBPC3 and controls (**Chapter 5**). Strikingly, aortic stenosis-related hypertrophic hearts and PLN-related dilated hearts showed a more extensive formation of fibrosis than MYBPC3-related hypertrophic hearts, further confirming the involvement of the active fibroblasts.

Since bulk sequencing of whole tissue could easily miss the cell type-specific information that characterizes the (mal)adaption of myocardial remodeling in HF, piling studies in HF are now frequently employing single-cell sequencing techniques. However, it is important to note that most of these scRNA-seq studies are limited to analysis of heart tissues from disease-mimicking animal models, due to easier handling and quick accessibilities.[16,17] The conservation of global gene expression between human and other mammal species is still debatable.[18,19,20,21] Thus, the single cellular responses in disease hearts as observed in animal models may not fully translate into the cellular responses in human hearts. Tucker and colleagues recently performed single nuclear RNA-seq in isolated nuclei from healthy human hearts and revealed 9 main clusters and over 20 subclusters of cell types.[9] However, this technique neglects a significant part of the mRNA signals in the cytoplasm. Until now, dissecting human hearts on a single-cell level without damaging the cell integrity in the wet lab procedure remains difficult, and a highly sensitive and flexible technique to capture single cells with various sizes and shapes in a consistently high throughput bulk approach is still lacking.[22] To tackle these methodical limitations, several studies have employed *in situ* sequencing to analyze the transcriptome landscape in the cardiac tissue, thus circumventing the typical problems associated with cell dissociation procedures.[23,24] Nevertheless, this new technique of *in situ* sequencing is still limited by the available antibodies and/or probes to detect

gene expression at the genome-wide scale, and further developments are needed to fully investigate cellular changes on a single-cell level in the (human) heart.

Beyond the central dogma

Since transcriptome profiling shows great potential in understanding the disease pathologies and providing clinical indications,[25,26] we studied the acquired HF with a focus on the mRNA expression level in **Chapters 2** and **3**. Besides the transcriptome profiling, we focused on the histone acetylation level in acquired HF in **Chapter 4**, because histone acetylation plays a major role in regulating gene transcription.[27] To further broaden our understanding of inherited HF, we integrated informative from the DNA, mRNA, and/or protein levels in **Chapter 5-8**. We noticed that the histone acetylation level correlates with the mRNA expression level in the same sample, however, the correlation between the global acetylation level, mRNA level, and/or protein level after the pairwise comparison becomes less promising.

Although the central dogma of cellular information transition from DNA, RNA, to protein is a key concept of molecular biology, increasing evidence has indicated that this linear mechanism is oversimplified.[28,29] For example, only a minor portion of our genome consists of the sequences that encode all proteins. In addition, a large portion of sequences encodes for reverse transcriptases, proteins which reverse the transcription, indicative of the role of additional machinery required during the translation of mRNA to protein.[30] Small molecules (e.g. polyketides and riboswitches), which are synthesized in the cells, can modulate the biological processes and thereby contribute to the genotype-to-phenotype variations.[31] Similarly to our studies, a previous study showed that the correlation between polymerase II occupancy and the mRNA expression level was weaker than the correlation between polymerase II occupancy and the TF occupancy at the genome-wide scale.[32] Another study showed a modest correlation between mRNA and protein expression levels.[33] Furthermore, variations among the correlation between DNA copy number and mRNA expression, mRNA expression and protein expression, and DNA copy number and protein expression have also been revealed.[34] The degradation of mRNA is one of the major factors that result in the poor correlation between DNA copies and mRNA expression. A rapid turnover rate of mRNA, defined as the ratio of mRNA synthesis and degradation, is essential to facilitate the dynamic biological processes in response to various stimuli.[35] Cytoplasmic mRNA can be degraded from 3' to 5' untranslated regions via PAN2/PAN3 and CCR4/NOT deadenylation complexes or from 5' to 3' untranslated regions via the decapping enzyme DCP2 or NUDT16.[36,37,38] The interaction between mRNA and specific RNA-binding

proteins, such as Puf proteins, can result in the degradation of mRNA.[39] The cytoplasmic foci in mammalian cells capture non-translating mRNAs and recruit the deadenylation complexes or the decapping enzymes, leading to mRNA degradation.[40] Small RNAs, such as microRNAs (miRNAs) and short interfering RNAs (siRNAs), bind to mRNA and repress the post-translational gene expression.[41] Naturally occurring genetic variants also influence the mRNA expression in the hearts.[42] Likewise, protein degradation could hamper the correlation between mRNA expression and protein expression. The degradation of misfolded and/or unfolded proteins is most commonly regulated by the ubiquitin-proteasome pathway (UPP) and autophagy.[43] In UPP, ubiquitin targets proteins directly and directs them towards proteasome for degradation in an ATP-dependent manner.[44] During autophagy, proteins are degraded by lysosomes.[45] However, both UPP and autophagy can regulate proteolysis via the inhibition of the mammalian target of rapamycin (mTOR).[46] A recent study indicated a new machinery in which the translating ribosome regulates the backtracked elongation complexes and subsequently reactivate the transcription, which further contributes to the complexity of the correlation between mRNA and protein expressions.[47] Ribosome profiling also shows a better ability in predicting the final protein level when compared to RNA-seq.[42]

Taken together, the correlation between DNA, RNA, and protein expression levels from our studies and others could be due to numerous factors, including the post-translational cleavage, mRNA, and/or protein degradation. A review from our group already highlighted the importance of proper multi-omics data integration in providing deep insights into the disease etiology.[48] In conclusion, an optimal way to integrate multi-omics data rather than a simple correlative and the discovery of new types of machinery beyond the central dogma is needed in future studies.

Clinical application of the extensive multi-omics data

The advantage of high-throughput sequencing techniques is its ability in generating large data even from a small sample size at a single level, such as the global gene expression.[49] More and more studies now are employing integrative analysis, which crosses multi-omics data from different levels, to characterize the disease subtyping, understand the complex mechanism underlying the diseases, and identify predictive biomarkers.[50] Furthermore, connecting the clinical parameters to multi-omics data shows great potential in creating a prediction model of the disease outcome.[51,52] Specific target that is obtained by integrating multi-omics data can serve as a potential therapeutic target and/or biomarker.[53] Besides, the biological

network that is supported by multi-omics data also facilitates drug discovery.[54] Clinical trials in HF have already been launched to study the precision medicine using multi-omics approaches (i.e. ClinicalTrials.gov Identifier: NCT04196842). Our studies show comprehensive information in both acquired and inherited HF at different levels using multi-omics profiling. For example, we performed integrative analysis and showed that the mRNA and protein levels of *MCAM* (also known as CD146) were significantly higher in HCM hearts than the controls (**Chapter 5**). Interestingly, *MCAM* has been proposed as a biomarker for endothelial damage,[55] and endothelial dysfunction is highly associated with HCM and HF.[56,57] The inhibition of fatty acid metabolism that we identified in the PLN R14del hearts using the integrative analysis (**Chapter 8**) has also been shown in failing hearts,[58,59,60] and its metabolite serves as the biomarker for adverse cardiac events.[61] Additionally, we presented a novel approach to study QRS-associated loci that affect LV mass by integrating these loci with altered chromatin regulation regions and genes in hypertrophic hearts versus controls (**Chapter 7**). Besides these, we also showed other promising and novel candidates that play a critical role during the HF progression. Comparing the altered biological processes that are revealed in aortic stenosis-related hypertrophic hearts (HFpEF), mutated *MYBPC3*-related hypertrophic hearts (HFpEF), and *PLN* R14del-related dilated hearts (HFrEF), we can further present the general pathways in HF and the disease-specific pathways.

Besides *ex vivo* hearts and *in vitro* hiPSC-CMs, future studies using engineered heart tissues and cardiac organoids are also needed to investigate promising candidates and their involved biological functions. Their 3-dimensional structures allow us to further elucidate cell-to-cell and cell-to-matrix networks, which resemble the tissue-level organization have more clinical relevance when compared to the traditional 2-dimensional cell culture.[62,63] Furthermore, unlike the limited access to cardiac samples from the early stage diseased hearts, hiPSC-CMs, engineered heart tissues, and cardiac organoids can mimic the chronological progression of HF,[64,65,66] which subsequently provide valuable insight to the identification of early disease markers and the understanding of the heterogeneity among patients. By connecting the information obtained from multi-omics approaches and *in situ* sequencing-based single-cell imaging in various models, we can dissect the regulome, the transcriptome, and the proteome landscape of each cell in the heart and integrate the druggable genome and study the potentially druggable genes per cell type.[9] These will further benefit the tailored therapeutic strategies per disease etiology in HF.

Taken together, this thesis presents the unique and comprehensive data in HF and provides key regulators/factors that may drive the pathological mechanism(s) underlying the disease. Changes of

chromatin regulatory region activities in diseased hearts compared to control hearts provide valuable insights into identifying TF binding sites and related TFs, which regulate the downstream transcription. Changes of the transcriptional landscape further elucidate on the disease-driving and/or disease-related biological processes and pathways. Changes of proteome profile show the final products of the affected genes, which subsequently reveals the complex protein-protein networks and serve as potential clinical or diagnostic biomarkers to monitor the disease onset and disease progression. Integrative analysis of these data at DNA, mRNA, and protein levels strengthens our confidence in identifying the key factors underlying the disease. These identified candidates and their regulated biological networks can be further investigated in early disease models to elucidate their roles during the disease progression and targeted using pharmacological treatments, which are of great interest for further translational steps towards clinical implementation.

References

1. Pecoits-Filho, R., Bucharles, S. & Barberato, S. H. Diastolic heart failure in dialysis patients: mechanisms, diagnostic approach, and treatment. *Semin. Dial.* **25**, 35–41 (2012).

2. Sharma, K. & Kass, D. A. Heart failure with preserved ejection fraction: mechanisms, clinical features, and therapies. *Circ. Res.* **115**, 79–96 (2014).

3. Brouwers, F. P. *et al.* Incidence and epidemiology of new onset heart failure with preserved vs. reduced ejection fraction in a community-based cohort: 11-year follow-up of PREVEND. *European Heart Journal* vol. 34 1424–1431 (2013).

4. Rao, M. S. *et al.* Comparison of RNA-Seq and Microarray Gene Expression Platforms for the Toxicogenomic Evaluation of Liver From Short-Term Rat Toxicity Studies. *Front. Genet.* **9**, (2018).

5. Ushio-Fukai, M. & Nakamura, Y. Reactive oxygen species and angiogenesis: NADPH oxidase as target for cancer therapy. *Cancer Lett.* **266**, 37–52 (2008).

6. van der Harst, P. *et al.* 52 Genetic Loci Influencing Myocardial Mass. *J. Am. Coll. Cardiol.* **68**, 1435–1448 (2016).

7. Pingzhu Zhou, W. T. P. Recounting cardiac cellular composition. *Circ. Res.* **118**, 368 (2016).

8. Wolfien, M. *et al.* Single-Nucleus Sequencing of an Entire Mammalian Heart: Cell Type Composition and Velocity. *Cells* **9**, 318 (2020).

9. Tucker, N. R. *et al.* Transcriptional and Cellular Diversity of the Human Heart. doi:10.1101/2020.01.06.896076.

10. Suryawanshi, H. *et al.* Cell atlas of the foetal human heart and implications for autoimmune-mediated congenital heart block. *Cardiovasc. Res.* (2019) doi:10.1093/cvr/cvz257.

11. Hu, P. *et al.* Single-nucleus transcriptomic survey of cell diversity and functional maturation in postnatal mammalian hearts. *Genes Dev.* **32**, 1344–1357 (2018).

12. Izhak Kehat, J. D. M. Molecular pathways underlying cardiac remodeling during pathophysiologic stimulation. *Circulation* **122**, (2010).

13. Nomura, S. *et al.* Cardiomyocyte gene programs encoding morphological and functional signatures in cardiac hypertrophy and failure. *Nat. Commun.* **9**, 1–17 (2018).

14. Ko, T., Hatano, M. & Komuro, I. Single-Cell RNA-seq of Human Cardiomyocytes Revealed DNA Damage Response as a Novel Predictor for Therapeutic Prognosis in Heart Failure Patients with Dilated Cardiomyopathy. *The Journal of Heart and Lung Transplantation* vol. 39 S22–S23 (2020).

15. Li, Z. *et al.* Single-cell transcriptome analyses reveal novel targets modulating cardiac neovascularization by resident endothelial cells following myocardial infarction. *Eur. Heart J.* **40**, 2507–2520 (2019).

16. Gladka, M. M. *et al.* Single-Cell Sequencing of the Healthy and Diseased Heart Reveals Cytoskeleton-Associated Protein 4 as a New Modulator of Fibroblasts Activation. *Circulation* **138**, 166–180 (2018).

17. Li, G. *et al.* Single cell expression analysis reveals anatomical and cell cycle-dependent transcriptional shifts during heart development. *Development* **146**, (2019).

18. Monaco, G., van Dam, S., Casal Novo Ribeiro, J. L., Larbi, A. & de Magalhães, J. P. A comparison of human and mouse gene co-expression networks reveals conservation and divergence at the tissue, pathway and disease levels. *BMC Evol. Biol.* **15**, 259 (2015).

19. Uosaki, H. & Taguchi, Y.-H. Comparative Gene Expression Analysis of Mouse and Human Cardiac Maturation. *Genomics Proteomics Bioinformatics* **14**, 207–215 (2016).

20. Byrne, G. W., Du, Z., Sun, Z., Asmann, Y. W. & McGregor, C. G. A. Changes in cardiac gene expression after pig-to-primate orthotopic xenotransplantation. *Xenotransplantation* **18**, 14–27 (2011).

21. He, J.-J. *et al.* Global Transcriptome Profiling of Multiple Porcine Organs Reveals Toxoplasma gondii-Induced Transcriptional Landscapes. *Front. Immunol.* **10**, (2019).

22. Wolfien, M. *et al.* Single nuclei sequencing of entire mammalian hearts: strain-dependent cell-type composition and velocity. *Cardiovasc. Res.* (2020) doi:10.1093/cvr/cvaa054.

23. A Spatiotemporal Organ-Wide Gene Expression and Cell Atlas of the Developing Human Heart. *Cell* **179**, 1647–1660.e19 (2019).

24. Maïno, N. *et al.* A microfluidic platform towards automated multiplexed in situ sequencing. *Sci. Rep.* **9**, 1–10 (2019).

25. Casamassimi, A., Federico, A., Rienzo, M., Esposito, S. & Ciccodicola, A. Transcriptome Profiling in Human Diseases: New Advances and Perspectives. *Int. J. Mol. Sci.* **18**, (2017).

26. Lee, H. *et al.* Diagnostic utility of transcriptome sequencing for rare Mendelian diseases. *Genet. Med.* **22**, 490–499 (2019).

27. Verdone, L., Agricola, E., Caserta, M. & Di Mauro, E. Histone acetylation in gene regulation. *Brief. Funct. Genomics* **5**, 209–221 (2006).

28. Sarah Franklin, T. M. V. Genomes, Proteomes and the Central Dogma. *Circ. Cardiovasc. Genet.* **4**, 576 (2011).

29. Kassem, H. S., Girolami, F. & Sanoudou, D. Molecular genetics made simple. *Global Cardiology Science & Practice* **2012**, (2012).

30. Henikoff, S. Beyond the Central Dogma. *Bioinformatics* **18**, (2002).

31. Schreiber, S. L. Small molecules: the missing link in the central dogma. *Nat. Chem. Biol.* **1**, 64–66 (2005).

32. Mokry, M. *et al.* Integrated genome-wide analysis of transcription factor occupancy, RNA polymerase II binding and steady-state RNA levels identify differentially regulated functional gene classes. *Nucleic Acids Res.* **40**, 148 (2012).

33. Payne, S. H. The utility of protein and mRNA correlation. *Trends Biochem. Sci.* **40**, 1 (2015).

34. Myhre, S. *et al.* Influence of DNA copy number and mRNA levels on the expression of breast cancer related proteins. *Mol. Oncol.* **7**, 704 (2013).

35. Yamada, T. & Akimitsu, N. Contributions of regulated transcription and mRNA decay to the dynamics of gene expression. *Wiley Interdiscip. Rev. RNA* **10**, e1508 (2019).

36. Wahle, E. & Winkler, G. S. RNA decay machines: deadenylation by the Ccr4-not and Pan2-Pan3 complexes. *Biochim. Biophys. Acta* **1829**, 561–570 (2013).

37. Arribas-Layton, M., Wu, D., Lykke-Andersen, J. & Song, H. Structural and functional control of the eukaryotic mRNA decapping machinery. *Biochim. Biophys. Acta* **1829**, 580–589 (2013).

38. Song, M.-G., Li, Y. & Kiledjian, M. Multiple mRNA decapping enzymes in mammalian cells. *Mol. Cell* **40**, 423–432 (2010).

39. Miller, M. A. & Olivas, W. M. Roles of Puf proteins in mRNA degradation and translation. *Wiley Interdiscip. Rev. RNA* **2**, 471–492 (2011).

40. Teixeira, D. Processing bodies require RNA for assembly and contain nontranslating mRNAs. *RNA* vol. 11 371–382 (2005).

41. Valencia-Sanchez, M. A., Liu, J., Hannon, G. J. & Parker, R. Control of translation and mRNA

degradation by miRNAs and siRNAs. *Genes Dev.* **20**, 515–524 (2006).

42. van Heesch, S. *et al.* The Translational Landscape of the Human Heart. *Cell* **178**, 242–260.e29 (2019).

43. Sha, Z., Zhao, J. & Goldberg, A. L. Measuring the overall rate of protein breakdown in cells and the contributions of the ubiquitin-proteasome and autophagy-lysosomal pathways. *Methods Mol. Biol.* **1844**, 261 (2018).

44. Hanna, J., Guerra-Moreno, A., Ang, J. & Micoogullari, Y. Protein Degradation and the Pathologic Basis of Disease. *Am. J. Pathol.* **189**, 94–103 (2019).

45. Vilchez, D., Saez, I. & Dillin, A. The role of protein clearance mechanisms in organismal ageing and age-related diseases. *Nat. Commun.* **5**, 1–13 (2014).

46. Zhao, J., Zhai, B., Gygi, S. P. & Goldberg, A. L. mTOR Inhibition Activates Overall Protein Degradation by the Ubiquitin Proteasome System as Well as by Autophagy. *Proc. Natl. Acad. Sci. U. S. A.* **112**, (2015).

47. Stevenson-Jones, F., Woodgate, J., Castro-Roa, D. & Zenkin, N. Ribosome reactivates transcription by physically pushing RNA polymerase out of transcription arrest. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 8462–8467 (2020).

48. Harakalova, M. & Asselbergs, F. W. Systems analysis of dilated cardiomyopathy in the next generation sequencing era. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **10**, e1419 (2018).

49. Wu, C. *et al.* A Selective Review of Multi-Level Omics Data Integration Using Variable Selection. *High-Throughput* **8**, (2019).

50. Subramanian, I., Verma, S., Kumar, S., Jere, A. & Anamika, K. Multi-omics Data Integration, Interpretation, and Its Application. *Bioinform. Biol. Insights* **14**, (2020).

51. Athreya, A. *et al.* Augmentation of Physician Assessments With Multi-Omics Enhances Predictability of Drug Response: A Case Study of Major Depressive Disorder. *IEEE Comput. Intell. Mag.* **13**, (2018).

52. Mankoo, P. K., Shen, R., Schultz, N., Levine, D. A. & Sander, C. Time to Recurrence and Survival in Serous Ovarian Tumors Predicted From Integrated Genomic Profiles. *PLoS One* **6**, (2011).

53. Recent Advances in In Vitro and In Vivo Multi-omics Analyses of Extracellular Vesicles: Therapeutic Targets and Biomarkers | Frontiers Research Topic. https://www.frontiersin.org/research-topics/13173/recent-advances-in-in-vitro-and-in-vivo-multi-omics-analyses-of-extracellular-vesicles-therapeutic-t.

54. Turanli, B. *et al.* A Network-Based Cancer Drug Discovery: From Integrated Multi-Omics Approaches to Precision Medicine. *Curr. Pharm. Des.* **24**, (2018).

55. Muccilli, A. *et al.* MCAM in Multiple Sclerosis: a biomarker of disease activity? (P1.381). *Neurology* **88**, (2017).

56. Dimitrow, P. P., Undas, A., Bober, M., Tracz, W. & Dubiel, J. S. Plasma Biomarkers of Endothelial Dysfunction in Patients With Hypertrophic Cardiomyopathy. *Pharmacol. Rep.* **59**, (2007).

57. Gogiraju, R., Bochenek, M. L. & Schäfer, K. Angiogenic Endothelial Cell Signaling in Cardiac Hypertrophy and Heart Failure. *Front. Cardiovasc. Med.* **6**, (2019).

58. Karwi, Q. G., Uddin, G. M., Ho, K. L. & Lopaschuk, G. D. Loss of Metabolic Flexibility in the Failing Heart. *Front. Cardiovasc. Med.* **5**, (2018).

59. Lionetti, V., Stanley, W. C. & Recchia, F. A. Modulating fatty acid oxidation in heart failure. *Cardiovasc. Res.* **90**, 202 (2011).

60. Abdurrachim, D. *et al.* Good and bad consequences of altered fatty acid metabolism in heart failure:

evidence from mouse models. *Cardiovasc. Res.* **106**, 194–205 (2015).

61. Ruiz, M. *et al.* Circulating acylcarnitine profile in human heart failure: a surrogate of fatty acid metabolic dysregulation in mitochondria and beyond. *Am. J. Physiol. Heart Circ. Physiol.* **313**, H768–H781 (2017).

62. Mills, R. J. *et al.* Drug Screening in Human PSC-Cardiac Organoids Identifies Pro-proliferative Compounds Acting via the Mevalonate Pathway. *Cell Stem Cell* **24**, 895–907.e6 (2019).

63. Goldfracht, I. *et al.* Generating ring-shaped engineered heart tissues from ventricular and atrial human pluripotent stem cell-derived cardiomyocytes. *Nat. Commun.* **11**, 1–15 (2020).

64. Song, J.-P. *et al.* Elevated plasma β-hydroxybutyrate predicts adverse outcomes and disease progression in patients with arrhythmogenic cardiomyopathy. *Sci. Transl. Med.* **12**, (2020).

65. Nugraha, B., Buono, M. F., von Boehmer, L., Hoerstrup, S. P. & Emmert, M. Y. Human Cardiac Organoids for Disease Modeling. *Clin. Pharmacol. Ther.* **105**, 79–85 (2019).

66. Extracellular Matrix From Hypertrophic Myocardium Provokes Impaired Twitch Dynamics in Healthy Cardiomyocytes. *JACC: Basic to Translational Science* **4**, 495–505 (2019).

Appendix

Acknowledgements

I am so grateful to get to work with all the wonderful people during this journey, though it is impossible to name them all and express my gratitude.

First, I would like to thank my promotor, **Prof. Marianne C Verhaar**, thank you for all the advice and information you gave me during our meetings. And thanks for the lovely birthday wishes you sent me every year.

I would also like to thank my co-promotor, **Prof. Folkert W Asselbergs**, thanks a lot for your support throughout the way, and I very much appreciated your caring about my happiness at work and all your efforts in creating such a cozy and rocking team. Meetings with you are always enlightening and motivating.

Dr. Caroline Cheng, my supervisor, thank you for having me on the team. I have learned a lot from you during the past few years, your knowledge of the vasculature and cell biology, your critical insights, your concrete plans for all the projects, etc.

My co-supervisor, **Dr. Magdalena Harakalova**, thank you for teaching me all the omics magics. What a great pleasure to work with you, your creative passion and your full understanding mean a lot to me. Apart from work, thanks for taking me to Slovakia and I had an amazing time there with your family (cannot forget the rice cake).

Another magician, **Michal**, thank you for helping me with the difficult experiments and the sequencing data. I still remembered how you optimized a 10-page-long protocol for my first ChIP-seq experiment. Thanks for all the beautiful plots you made out of the massive data and your comments and suggestions during my presentations as well as in all our projects.

Thanks to all the colleagues in Caroline's group and in the Nephrology Department. **Laura**, how amazing it was to start our PhD exactly the same day in the same group. So happy for your progress and achievements during your PhD. And thanks for all the delicious cookies you brought me from Spain, I am so happy that our tradition of exploring new restaurants in Utrecht continues after your graduation and already looking forward to our next foodie-hunt. It had always been a great time with you, either in the lab and outside work, **Femke**. Thanks for all the nice suggestions you gave me during the tea & chocolate breaks. It was also very enjoyable to run with you near UMC after a full working day. I hope we can still do some sportive activities together (what about the mud run and survival run) and wish you all the successes with your new adventure in clinics. **Dienty**, you are extremely helpful and practical and it was

such a blessing to have you in the group. Besides, big thanks for taking care of me with all the violently amazing cookies and cakes you made (chocolate bonbons, coconut balls, macarons, cinnamon rolls, etc., I probably need another chapter to list all). Now you moved to Heidelberg and I am sure you will rock there as well. **Merle, Isabel, and Laura**, I really enjoyed our projects together at work and our 4th-year-PhD dinners. **Chris, Maarten, Paul, Tobia, Joost, Jaap,** and **Petra**, thank you all for the help and the nice chit-chats during the coffee breaks.

Huge appreciation to everyone in Folkert's group and in the Cardiology Department. **Jantine**, thank you for scheduling all the meetings and arranging all kinds of forms and paperwork for me. **Jessica**, I had a great time with you during the ESHG meeting in Milan and I hope we can have more outings and other activities in the coming future. **Annette** and **Mark**, thank you both for all the useful feedback you provided during our weekly Wednesday meetings. **Jorg**, thanks for the great analyses you performed in our projects and introducing the nice places to visit during the CURE PLaN meeting in Philadelphia. **Frank**, thanks for your practical and critical advice and your amusing vibe, I wish I had known you and worked with you way before. **Renee**, the master of iPSC-cardiomyocytes, thank you so much for teaching me how to culture cardiomyocytes and for all the amazing readouts (calcium handling, metabolic rate, etc) you have shown me. **Christian**, you are always helpful and thanks for making my life in the lab much easier, I cannot imagine the lab without you now. **Floor**, I first got to know you during your internship in the Pathology Department and am so happy that now you joined our group. You always bring the sunshine to the lab and I am sure we will have many more cheerful moments working together. Of course, thank **all the rockstars** in the group for the brainstorming meetings, the sailing trips, the BBQ, and so on and so forth.

Ema, my partner in crime, cannot count how many hours we spent together in the lab cutting the cardiac tissues, isolating RNAs, preparing sequencing samples, etc. You are so organized and structured at work and I learned a lot from you. I also enjoyed playing squash with you.

Noortje, thank you big time for showing me around the lab, helping me with all the technical issues, and checking up on me every now and then. You are an amazing person to work with.

I am also grateful to everyone in the Pathology Department. **Aryan**, **Roel, Manon**, **Erica**, and **Joyce**, you have helped me massively, from cutting the precious human cardiac tissues, performing Qubit and bioanalyzer measurements, to optimizing the manuscript content. My thesis would not be possible without you all.

Anyan, my best friend and the other me on this planet, 谢谢你永远懂我在说什么又想要做什么，和我一起"变扭"的活着。这 10 年成长的路，我们走的有些幸苦，但也都挺直脊梁骨拼尽全力了；接下来的路，很期待和你继续走下去。

Lastly, endless gratitude to my family. 谢谢我爱的家人们对我超过宇宙容量的支持，理解和疼爱。每次我提出什么想法，爸爸总是："特别棒，太好了，去尝试吧。" 妈妈总是："好呀，你想好了就去做吧。" 从小到大，爷爷奶奶对我所有的期待就是健康和快乐；而外公外婆对我唯一的要求，就是别委屈自己。没有比拥有有你们更棒的事情了，感谢你们一直都在。

List of Publications

Pei J, Harakalova M, den Ruijter H, Pasterkamp G, Duncker DJ, Verhaar MC, Asselbergs FW, Cheng C. Cardiorenal disease connection during post-menopause: The protective role of estrogen in uremic toxins induced microvascular dysfunction. Int J Cardiol. 2017 Jul 1;238:22-30. doi: 10.1016/j.ijcard.2017.03.050.

Pei J, Juni R, Harakalova M, Duncker DJ, Asselbergs FW, Koolwijk P, Hinsbergh VV, Verhaar MC, Mokry M, Cheng C. Indoxyl Sulfate Stimulates Angiogenesis by Regulating Reactive Oxygen Species Production via CYP1B1. Toxins (Basel). 2019 Aug 2;11(8):454. doi: 10.3390/toxins11080454.

Pei J*, Harakalova M*, Treibel TA, Lumbers T, Boukens BJ, Efimov IR, van Dinter JT, González A, López B, El Azzouzi H, van den Dungen NAM, van Dijk CGM, Krebber MM, den Ruijter HM, Pasterkamp G, Duncker DJ, Nieuwenhuis EES, de Weger R, Huibers MM, Vink A, Moore JH, Moon JC, Verhaar MC, Kararigas G, Mokry M*, Asselbergs FW*, Cheng C. H3K27ac acetylome signatures reveal the epigenomic reorganization in remodeled non-failing human hearts. Clin Epigenetics. 2020 Jul 14; doi: 10.1186/s13148-020-00895-5.

Hemerich D, **Pei J**, Harakalova M, van Setten J, Boymans S, Boukens BJ, Efimov IR, Michels M, van der Velden J, Vink A, Cheng C, van der Harst P, Moore JH, Mokry M, Tragante V, Asselbergs FW. Integrative Functional Annotation of 52 Genetic Loci Influencing Myocardial Mass Identifies Candidate Regulatory Variants and Target Genes. Circ Genom Precis Med. 2019 Feb;12(2):e002328. doi: 10.1161/CIRCGEN.118.002328.

Schuldt M, **Pei J**, Harakalova M, Dorsch LM, Schlossarek S, Mokry M, Knol JC, Pham TV, Schelfhorst T, Piersma SR, dos Remedios C, Dalinghaus M, Michels M, Asselbergs FW, Moutin M, Carrier L, Jimenez CR, van der Velden1 J*, PhD, Kuster1 DWD*. Proteomic and functional studies reveal detyrosinated tubulin as treatment target in sarcomere mutation-induced hypertrophic cardiomyopathy. Accepted by Circulation: Heart Failure.

Huethorst E, Krebber MM, Fledderus JO, Gremmels H, Xu YJ, **Pei J**, Verhaar MC, Cheng C. Lymphatic Vascular Regeneration: The Next Step in Tissue Engineering. Tissue Eng Part B Rev. 2016 Feb;22(1):1-14. doi: 10.1089/ten.TEB.2015.0231.

Chrifi I, Hermkens D, Brandt MM, van Dijk CGM, Bürgisser PE, Haasdijk R, **Pei J**, van de Kamp EHM, Zhu C, Blonden L, Kros JM, Duncker DJ, Duckers HJ, Cheng C. Cgnl1, an endothelial junction complex protein, regulates GTPase mediated angiogenesis. Cardiovasc Res. 2017 Dec 1;113(14):1776-1788. doi: 10.1093/cvr/cvx175.

Brandt MM, van Dijk CGM, Chrifi I, Kool HM, Bürgisser PE, Louzao-Martinez L, **Pei J**, Rottier RJ, Verhaar MC, Duncker DJ, Cheng C. Endothelial loss of Fzd5 stimulates PKC/Ets1-mediated transcription of Angpt2 and Flt1. Angiogenesis. 2018 Nov;21(4):805-821. doi: 10.1007/s10456-018-9625-6. Epub 2018 May 29.

Manduchi E, Hemerich D, van Setten J, Tragante V, Harakalova M, **Pei J**, Williams SM, van der Harst P, Asselbergs FW, Moore JH. A comparison of two workflows for regulome and transcriptome-based prioritization of genetic variants associated with myocardial mass. Genet Epidemiol. 2019 Sep;43(6):717-726. doi: 10.1002/gepi.22215.

van Dijk CGM, Brandt MM, Poulis N, Anten J, van der Moolen M, Kramer L, Homburg EFGA, Louzao-Martinez L, **Pei J**, Krebber MM, van Balkom BWM, de Graaf P, Duncker DJ, Verhaar MC, Luttge R, Cheng C. A new microfluidic model that allows monitoring of complex vascular structures and cell interactions in a 3D biological matrix. Lab Chip. 2020 May 19;20(10):1827-1844. doi: 10.1039/d0lc00059k.

de Boer RA, Nijenkamp LLAM, Silljé HHW, Eijgenraam TR, Parbhudayal R, van Driel B, Huurman R, M Michels M, **Pei J**, Harakalova M, van Lint FHM, Jansen M, Baas AF, Asselbergs FW, van Tintelen JP, Brundel BJJM, Dorsch LM, Schuldt M, Kuster DWD, van der Velden J, DOSIS consortium. Strength of patient cohorts and biobanks for cardiomyopathy research. Neth Heart J. 2020 Aug;28(Suppl 1):50-56. doi: 10.1007/s12471-020-01456-4.

Manuscripts in preparation

Pei J*, Schuldt M*, Nagyova E*, Gu Z, Bouhaddani SE, Jansen M, Calis J, Dorsch LM, van den Dungen NAM, Lansu N, Bouken BJ, Efimov IR, Michels M, Verhaar MC, de Weger R, Vink A, van Steenbeek FG, Baas AF, Uh HW, Kuster DWD, Cheng C, Mokry M, van der Velden J*, Asselbergs FW*, Harakalova M*. Multi-omics integration delineates the key regulators and pathomechanisms in hypertrophic cardiomyopathy due to truncating MYBPC3 mutations

[Manuscript under submission]

Pei J*, Maas R*, Nagyova E, Gho JMIH, Blok CS, van Adrichem I, Calis J, van Es R, Sepehrkhouy S, Feyen D, van den Dungen NAM, Lansu N, Buijsrogge MP, de Jonge N, den Ruijter HM, Pasterkamp G, Huibers MM, de Weger RA, van Laake L, Verhaar MC, van Tintelen P, van Steenbeek FG, van Mil A, Buikema JW, Burgering B, Kararikes I, Mercola M, Doevendans PA, Sluijter J, Vink A, Cheng C, Mokry M*, Asselbergs FW*#, Harakalova M*. Epigenetic profiling reveals disrupted lipid metabolism in failing hearts and iPSC-cardiomyocytes with a pathogenic phospholamban mutation.

[Manuscript in preparation]

*equal contribution

Curriculum Vitae

Jiayi Pei was born on August 21st, 1991 in Jingdezhen, China. In 2012, she obtained a Bachelor of Science degree in Pharmaceutical Engineering from Jilin University Zhuhai College and a second bachelor's degree in Business English from Jilin University in China. She continued with her Master's degree in Molecular Life Science at Wageningen University, the Netherlands. During her Master's studies, she conducted the first internship at the Department of Nutrition and Pharmacology from the Wageningen University under the supervision of Prof. dr. Renger Witkamp and Dr. Jocelijn Meijerink and investigated the possible receptor(s) for DHA serotonin as an anti-inflammatory compound in macrophages. Afterward, she completed the second internship at the Medicine Faculty from the Southampton University (United Kingdom) under the supervision of Prof. dr. Philip Calder and Dr. Elizabeth A Miles and studied the impact of vitamin D deficiency during pregnancy on the immune system in the offspring at birth. She returned to the Netherlands and started a Ph.D. under the supervision of Dr. Caroline Cheng and Prof. dr. Marianne Verhaar at the department of Nephrology and Hypertension as well as the supervision of Dr. Magdalena Harakalova and Prof. dr. Folkert Asselbergs at the department of Cardiology from the University Medical Center Utrecht, which results in this thesis. Since July 2020, Jiayi works as a postdoc researcher in Prof. dr. Folkert Asselbergs's group at the University Medical Center Utrecht and focuses on the altered metabolism and allelic imbalance in relation to the cell-to-cell and genotype-to-phenotype variations in cardiomyopathies.