UNIVERSITY OF CALIFORNIA

Los Angeles

# Translating Mouse Systems Genetics to Discovery in Human Disease

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Molecular, Cellular and Integrative Physiology

by

Pritha Payel Gupta

2017

ABSTRACT OF THE DISSERTATION

Translating Mouse Systems Genetics to Discovery in Human Disease

by

Pritha Payel Gupta

Doctor of Philosophy in Molecular, Cellular and Integrative Physiology

University of California, Los Angeles, 2017

Professor Aldons J. Lusis, Chair

This dissertation is the culmination of my graduate studies in the laboratory of Jake Lusis at UCLA. The research presented here utilizes systems genetics studies performed in mice to aid in discovery in human disease in three separate studies. A significant portion of disease-oriented research is performed in mice, but a major criticism from the medical community is that laboratory mice are generally inbred and thus have no genetic variation among individuals. Almost 15 years ago, the Lusis lab developed a novel genetic resource for association analysis in the mouse called the Hybrid Mouse Diversity Panel (HMDP). The HMDP is a panel of inbred mouse strains that was developed for performing association studies with adequate statistical power and resolution for mapping of complex traits. Mouse genome wide association studies (GWAS) studies are a powerful tool and can be performed relatively easily, but translating the data obtained from these studies to human disease is still in its infancy. My dissertation work

reveals three different novel approaches to the utilization of data from GWAS studies performed on the HMDP for translation into human disease processes, namely cardiovascular disease.

The first study utilizes novel genetic signatures in murine macrophages to predict disease incidence and survival in humans. The second study utilizes a traditional GWAS to candidate gene discovery to elucidate the mechanisms underlying cardiac remodeling in humans. Lastly, the third study utilizes mouse GWAS data for novel heart failure biomarker discovery in humans.

As an introduction to this dissertation, Chapter 1 briefly summarizes the history of GWAS in mice using the HMDP and GWAS in humans. Chapter 2 is a completed and accepted first-author manuscript entitled "Natural diversity reveals macrophage activation spectra predictive of inflammation and cancer survival." Chapter 3 explores the role of CD200, a candidate gene obtained from a large heart failure GWAS study in mice, and it's receptor, CD200R1 in cardiac homeostasis and injury. Chapter 4 describes a novel approach to biomarker discovery for human heart failure using data from a large heart failure GWAS study. Chapter 5 is a departure from mouse systems genetics. In this chapter, I describe the strengths and pitfalls of exome sequencing. In addition, I describe two cases of rare cardiovascular disease in which exome sequencing is utilized to find causal variants of disease.

Ultimately, I'd like to use what I've learned in my studies of mouse genetics and translate this to discovery in human disease.

In conclusion, this dissertation work contributes significant findings to the expanding knowledge of utilizing mouse GWAS for discovery in human disease.

The dissertation of Pritha Payel Gupta is approved.

Yibin Wang

Sherie Morrison

Yousang Gwack

Aldons J. Lusis, Committee Chair

University of California, Los Angeles

2017

*To my loving husband and my amazing co-fellows*

**Table of Contents**

**Acknowledgements**

I would like to make the following acknowledgements to recognize the people and organizations that made the work presented in this dissertation possible.

The work in Chapter 2 is made possible by Dr. Klaus Ley and his group at the La Jolla Institute for Allergy and Immunology in La Jolla, CA. *My contribution to this chapter was performing all of the experiments, helping Dr. Klaus Ley with the analysis, and writing the initial manuscript.* I would like to especially thank Dr. Konrad Buchner for his expertise and support that really took our data to the next level. I would also like to acknowledge and thank Mr. Erik Ehinger for his expertise in flow cytometry and his help with constructing the figures for the manuscript. I would like to recognize and thank Mr. Pramod Akula Bala and Calvin Pan for their help with computation and bioinformatics. I would like to recognize and thank Mr. George Tweet for his expertise in flow cytometry. Most importantly, thank you to Luz Orozco for producing the Hybrid Mouse Diversity Panel (HMDP) macrophage data that was the basis for this project. Last but not least, thank you to Nam Che for his assistance with the process of harvesting and culturing thioglycollate-elicited macrophages in our mice and optimizing the amino acid assays.

The work in Chapter 3 and 4 is made possible by Drs. Jessica Wang and Cristoph Rau, who originally produced the HMDP Heart Failure data*. My contribution to these two projects are experimental planning and execution for all experiments in Chapter 3, performing mouse echocardiograms (Chapter 3), data analysis for both mouse and human CD200 projects (Chapter 3), performing western blots to screen for plasma biomarkers in mice (Chapter 4), and plasma collection in control and Isoproterenol-treated mice (Chapter 4).* I would like to thank Dr. Jessica Wang for her unwavering support and assistance in the planning and execution of

these projects. I would like to thank Milagros Romay for her assistance with experimental planning. I would like to thank the Gorzynski Lab in Toronto, ON for providing the CD200 and CD200R KO mice. I would like to extend a very special thank you to both Zhiqiang Zhou and Shuxun (Vincent) Ren for their assistance with the mouse Isoproterenol pump surgeries. I would like to thank Adriana Huertas-Vasquez for leading this project to the advanced stages it is in today. In addition, I'd like to thank UCLA undergraduate students, Jeff Hsiao ad Gabriel Stolin, for their assistance in collecting mouse serum. Last but not least, thank you to Nam Che for his assistance with experimental technique and troubleshooting.

The work in Chapter 5 is made possible by Dr. Jessica Wang and the investigators of the Cardiovascular Disease Registry. I would also like to thank Ms. Jessica Rahman for her coordination of patient consent and sample collection. Library preparation, sequencing and data analysis were performed by Macrogen Clinical Lab in Rockville, MD. I would like to thank the technical support staff at Golden Helix for allowing me to use their program and for their assistance with the exome-sequencing pipeline. I would like to recognize Mindy H. Li et al. for her conceptualization of a schematic in whole exome sequencing that I adopted for a figure in this chapter.

Lastly, I would like to thank my thesis committee and especially Dr. Aldons (Jake) Lusis for his mentorship and support during my PhD work.

**BIOGRAPHICAL SKETCH**

| NAME | POSITION TITLE |
|---|---|
| Pritha Gupta, MD | Fellow in Cardiovascular Disease |

EDUCATION/TRAINING

| INSTITUTION AND LOCATION | DEGREE | YEAR(s) | FIELD OF STUDY |
|---|---|---|---|
| University of California, Los Angeles | BS | 06/05 | Biochemistry |
| New York University School of Medicine | MD | 06/10 | Medicine |
| University of Washington, Seattle | Residency | 07/13 | Internal Medicine |
| University of California, Los Angeles | Fellowship | Current | Cardiology |
| University of California, Los Angeles | Doctoral Candidate | Current | Molecular, Cellular and Integrative Physiology |

**A. Positions and Honors.**

**Positions**

2014- Graduate Student/Doctoral Candidate, Laboratory of Aldons J. Lusis, PhD, Department of Molecular, Cellular and Integrative Physiology, University of California, Los Angeles, CA

2013- Cardiovascular Disease Fellow, University of California, Los Angeles, CA

2010-2013 Internal Medicine Resident, University of Washington, Seattle, WA

**Honors**

01/2016 American College of Cardiology (ACC) and Merck Research Fellowship

09/2015 Tibian Fabor Award For Excellence in Cardiovascular Research

08/2015 American Heart Association (AHA) Women in Cardiology Trainee Award for Excellence

01/2015 UCLA Older Americans Independence Center (OAIC) and Clinical and Translational Institute (CTSI) Rapid Pilot Award Recipient ($10,000)

06/2014 Heart Failure Society of America (HFSA) Travel Grant

05/2009 Top 10 Finalist in 2009 Research Competition - American Association of Physicians of Indian Origin (AAPI)

09/2008 Association for American Physicians of Indian Origin (AAPI) Travel Award

09/2008 American Medical Student Association (AMSA) Local Project Grant Recipient

04/2008 Finalist in Best Scientific Poster Contest, American College of Cardiology

04/2007 AMSA National Primary Care Week MicroGrant Recipient ($1500)

04/2007 Glorney-Raisbeck Medical Student Grant in Cardiovascular Research ($3500)

06/2006 New York University School of Medicine Honors Research Grant ($4000)

**B. Selected peer-reviewed publications (in chronological order).**

1. Jadhav, N, **Gupta, P**, Calfon Press, M. Right Ventricular Marginal Branch Occlusion After Blunt Force Trauma to the Chest. (2015) *Trauma.*

2. Sparks, R. Salskov, AH, Chang, AS, Wentworth, KL, **Gupta, PP**, Staiger, TO, Anawalt, BD. Pocket Change: A Simple Educational Intervention Increases Hospitalist Documentation of Comorbidities and Improves Hospital Quality Performance Measures. (2015) *Q Manage Health Care.*

3. **Gupta, P.**, Fonarow, G., Horwich, T. (2014) Obesity and the Obesity Paradox in Heart Failure. *Canadian Journal of Cardiology.*

4. Hartman, M., Librande, JR., Medvedev,I. Ahmad,R., Moussavi-Harami,F**.**, **Gupta, PP.**, Chien, WM., Chin,MT. (2014) An Optimized and Simplified Model of Mouse Embryonic Stem Cell Cardiac Differentiation for the Assessment of Differentiation Modifiers. *PLOS One.*

5. Cubeddu, RJ, Don CW, Horvath, SA, Gupta, PP. Cruz-Gonzalez, I, Witzke, C. Inglessis, I., Palacios, IF. (2012) Left ventricular end-diastolic pressure as an independent predictor of outcome during balloon aortic valvuloplasty. Catheterization and Cardiovascular Interventions.

6. Don, C., **Gupta, P.**, Witzke, C., Kaserwani, M. Cubeddu, R., Herrero-Garibi, J., Pomerantsev, E., McCarty, D., Inglessis, I., Palacios, IF. (2011) Procedural and In-Hospital Outcomes of Percutaneous Balloon Aortic Valvuloplasty in Patients with Severe Calcific Aortic Stenosis and Left Ventricular End-diastolic Diameter < 4.0 cm. *Catheterization and Cardiovascular Interventions.*

7. *Kang, E., *Ponzio, M., **Gupta, P.**, Liu, F., Butensky, A., Gutstein, DE. (2009) Identification of binding partners for the cytoplasmic loop of Cx43: A novel interaction with β-tubulin. *Cell Communication and Adhesion.* *These authors contributed equally.

8. Kontogeorgis, A., Kaba, A., Kang, E., Feig, J., Gupta**, P.**, Ponzio, M., Liu, F., Rindler, M.,Wit, A., Fisher, E., Peters, N., and Gutstein, DE. (2008) Short-term pacing in the mouse alters cardiac expression of connexin43. *BMC Physiology*, Published online 2008, May 6.

9. Marmorino, M., **Gupta, P.** (2004) Surpassing the Temple Lower Bound. *Journal of Mathematical Chemistry* 35*,* 189-197.

**C. Research Support.**

**ACC Merck Research Fellowship** 07/01/16 – currently funded

**T32 Training Grant** T32 HL007895 Weiss (PI) 07/01/14-07/01/16 Role: Traine

*Chapter 1: Introduction*

## Introduction

It is an exciting era in medical genetics. In 2003, the sequencing of the human genome brought with it the promise of a genetic revolution in clinical medicine. Today, many envision a time when each person's genome will be sequenced and available for guidance in personalized approaches to health maintenance, disease prediction and prevention and treatment. Although this may sound like an impossible feat, resources are being developed for this purpose and medical centers around the nation are starting to genotype and whole exome sequence healthy individuals and banking the information for use at a later date. Without a doubt, large scale sequencing of humans in a clinical setting is still in its infancy as the infrastructure and manpower to handle this information is not yet present.

The obvious issues that will arise after collection of a massive amount of genetic data will be multi-faceted. If we were to sequence every patient that walked into the UCLA Medical Center today, we would quickly run out of space to securely store this sensitive information. More importantly, we would not have an adequate genetic workforce to interpret all of this information for research purposes, let alone deliver and explain these sensitive results to patients.

Although the study of disease in humans was my ultimate goal when starting this PhD, I quickly realized that due in part to reasons that are stated above, it is not easy to sequence thousands of humans to determine the underlying genetic causes of complex disease. This is not the case for rare, Mendelian disease, in which exome sequencing of patients and their affected family members is high yield. When studying complex disease, mouse models have come a long way in aiding these investigations from a molecular and genetics standpoint.

The laboratory mouse is considered the model organism for studying disease in humans, with whom they share 99% of their genes. Furthermore, mice and humans share most physiological and pathological features, including similarities in cardiovascular, nervous, endocrine, immune and other organ systems.[1] Despite being such a well-established model, the appropriateness of the mouse to recapitulate human disease and conditions was called into question in a 2013 study which reported poor correlation between human and mouse immune responses. [2] However, a subsequent study analyzed the same data using an arguably more rigorous and less biased methodology, and reported the exact opposite findings, which largely restored faith in the mouse as a model for human disease.[3]

The Lusis Lab utilizes the Hybrid Mouse Diversity Panel (HMDP), a collection of approximately 100 well-characterized inbred strains of mice, to analyze genetic and environmental factors underlying complex traits.[4] The use of a mouse model for mapping genetic loci has important advantages. Mice are used in the laboratory in a controlled setting and as a result, environmental factors can be controlled. In addition, relevant tissues are readily accessible for molecular phenotyping and because inbred strains are renewable, results from separate studies can be integrated. The Lusis Lab has utilized the HMDP to study traits relevant to many complex disease processes including obesity, diabetes, atherosclerosis, immune regulation, fatty liver disease, host-gut microbiota interactions and heart failure. These data are available and can be readily accessed. The majority of my work in the lab used data produced from the immune regulation and heart failure HMDP studies.

Mouse genome wide association studies (GWAS) are a powerful tool and can be performed relatively easily, but translating the data obtained from these studies to human disease is still a developing process. My dissertation work presents three different novel approaches to

the utilization of data from GWAS studies performed on the HMDP for translation into human disease processes, namely immunological and cardiovascular disease.

Using data from the immune regulation HMDP study, my first project utilizes novel genetic signatures in murine macrophages to predict disease incidence and survival in humans. Using data from the heart failure HMDP, my second project utilizes a traditional GWAS to candidate gene discovery method to elucidate the mechanisms underlying cardiac remodeling in humans by investigating a gene, CD200, thought to be involved in adverse left ventricular remodeling after myocardial injury. Lastly, my third project utilizes the heart failure HMDP data once more for novel heart failure biomarker discovery in humans.

The validity and success of my work using data from the HMDP studies were variable. My largest lesson from these projects pertained to model validity. Although it seems as though it should be an obvious point, the disease model used should be appropriate for the question being addressed. And although logistics can often limit the use of the correct model, an ideal disease model accurately mimics the human condition physiologically and pathologically. For example, a mouse model established to display the key motor symptoms seen in humans with amyotrophic lateral sclerosis (ALS) was utilized in the testing of promising drug candidates in preclinical studies; however these drugs ultimately failed in humans. [5,6] It was later shown that the mouse model used was a poor genetic and phenotypic model of the human condition.[7] This example illustrates how relevance to the human disease process being studied, supported by strong data validating the use of the model, is crucial for successful clinical translation.

In our case, it is possible that the Isoproterenol-induced heart failure model utilized in the heart failure HMDP study was not ideal in the recapitulation of the actual disease process in

humans. Although underlying increased beta-adrenergic tone is a significant part of the disease process and the cornerstone of beta-blocker therapy, it does not mimic some of the most important and common methods of myocardial injury. However, in the mice that underwent 3 weeks of Isoproterenol treatment, adverse cardiac remodeling was observed in some, but not in all, mice in the population. A contributing factor to detecting non-responders may have been the disease model that was used. Alternatively, this could simply be a demonstration of how genetic background plays a role in modulating the impact of chronic beta-adrenergic signaling in heart failure. In our effort to use the HMDP heart failure data in search for novel biomarkers, we were able to validate known biomarkers in HF using the ISO-treated mouse data; however it is not clear whether the elevation in known HF biomarkers was due to actual development of HF or if it was just a direct response to ISO treatment. The fact that we observed elevation in the same biomarkers in human hearts that were not treated with ISO argues that the biomarkers detected in mice were due to the actual development of HF.

Although my interests largely lie in complex metabolic and cardiovascular disease; I ultimately ended my dissertation work investigating rare cardiovascular disease via whole exome sequencing (WES). WES was used to find novel variants in 3 patients with rare cardiovascular disease. Although no known pathological variants were found in these patients, WES remains a useful tool in investigating cardiovascular disease.

The advent of next generation sequencing has changed the way we think about improving human health and is a significant driver in the movement towards personalized medicine. As mentioned above, the expense, ethical issues related to genetic data and the need for user-friendly software in the analysis of the raw sequences remain to be addressed. I hope to utilize

the knowledge and skills I've learned during my PhD to participate and aid in the effort towards personalized medicine in cardiovascular disease.

**References**


1. Rosenthal, N. and Brown, S. (2007). The mouse ascending: perspectives for human-disease models. Nat. Cell Biol. 9, 993-999.
2. Seok, J., Warren, H. S., Cuenca, A. G., Mindrinos, M. N., Baker, H. V., Xu, W., Richards, D. R., McDonald-Smith, G. P., Gao, H., Hennessy, L. et al. (2013). Genomic responses in mouse models poorly mimic human inflammatory diseases. Proc. Natl. Acad. Sci. USA 110, 3507-3512.
3. Takao, K. and Miyakawa, T. (2015). Genomic responses in mouse models greatly mimic human inflammatory diseases. Proc. Natl. Acad. Sci. USA 112, 1167-1172.
4. Bennett, B. J., C. R. Farber, L. Orozco, H. M. Kang, A. Ghazalpour, N. Siemers, M. Neubauer, I. Neuhaus, R. Yordanova, B. Guan, et al. 2010. A high-resolution association mapping panel for the dissec- tion of complex traits in mice. *Genome Res.* **20:** 281–290.
5. Wegorzewska, I., Bell, S., Cairns, N. J., Miller, T. M. and Baloh, R. H. (2009). TDP-43 mutant transgenic mice develop features of ALS and frontotemporal lobar degeneration. Proc. Natl. Acad. Sci. USA 106, 18809-18814.
6. Perrin, S. (2014). Preclinical research: make mouse studies work. Nature 507, 423-425.
7. Hatzipetros, T., Bogdanik, L. P., Tassinari, V. R., Kidd, J. D., Moreno, A. J., Davis, C., Osborne, M., Austin, A., Vieira, F. G., Lutz, C. et al. (2014). C57BL/6J congenic Prp-TDP43A315T mice develop progressive neurodegeneration in the myenteric plexus of the colon without exhibiting key features of ALS. Brain Res. 1584, 59-72.

*Chapter 2: Natural variation of macrophage activation as disease-relevant phenotype predictive of inflammation and cancer survival*

**Natural variation of macrophage activation as disease-relevant phenotype predictive of inflammation and cancer survival**

**Authors:**

P. Gupta [2†], K. Buscher [1†], E. Ehinger [1†], A. B. Pramod [1], D. Wolf [1], G. Tweet [1], C. Pan [2], C. D. Mills [3], A. J. Lusis [2], K. Ley [1*]

**Affiliations:**

[1] La Jolla Institute for Allergy and Immunology, Division of Inflammation Biology, La Jolla, California, USA.

[2] Departments of Medicine, Human Genetics, and Microbiology, Immunology, and Molecular Genetics, University of California, Los Angeles, California, USA.

[3] BioMedical Consultants, Marine on St. Croix, Minnesota, USA.

† These authors contributed equally to this work

* To whom correspondence should be addressed:

Klaus Ley, M.D., Division of Inflammation Biology, La Jolla Institute for Allergy & Immunology, 9420 Athena Circle Drive, La Jolla, CA 92037. (858) 752-6661 (tel), (858) 752-6985 (fax), klaus@lji.org

**One Sentence Summary:** Natural diversity of macrophage activation as inherent trait of human populations can by recapitulated by a mouse diversity panel, resulting in gene signatures that identify inter-individual risk profiles of disease susceptibility in humans.

**Abstract:**

Although mouse models exist for many immune-based diseases, insights often translate poorly into clinical outcomes. As a result, many clinical trials fail and medical resources are wasted. Most basic and translational studies utilize only a single inbred mouse strain. However, basal and diseased immune states in humans show vast inter-individual variability due to genetic and environmental factors, which seems critical for understanding systemic immune responses. Here, we addressed the lack of genetic diversity in mouse studies. We discovered a natural spectrum of macrophage activation propensity in response to lipopolysaccharide (LPS) in healthy humans. As a translational model of population diversity, a hybrid mouse diversity panel (HMDP) of 83 inbred strains was established. Exploiting the natural strain-specific repertoire, gene markers for LPS responsiveness were derived from peritoneal macrophages. In humans, they robustly identify susceptibility or resilience to several macrophage-related inflammatory and malignant disorders, including survival prediction across many tumors. These data introduce natural immune diversity of heterogeneous populations into mouse research, illustrate a mouse panel-

based approach to promote translation of basic research data, and identify a disease-relevant macrophage activation spectrum in humans.

**Introduction:**

Clinical and pharmaceutical researchers are concerned with the lack of relevance and low reproducibility of findings obtained in standard mouse models *(1–4)*. Many translational studies failed to confirm promising mouse data in patients *(5–9)* and clinical trial success rates continue to be low *(2, 3)*. Among the possible reasons, the large interindividual variation of immune system variables in heterogeneous human populations is cited *(10, 11)*. This natural variation has been shown to broadly impact on pathophysiology, e.g. disease resilience, tolerance *(12–14)* and vaccination responses *(10, 15)*. System level analyses of identical twins find both environmental and genetic components of variability *(16)*. However, both factors are intentionally removed in laboratory mice in order to reduce data variability. Recently, inbred mouse strains were placed in a more natural, „dirty“ environment leading to greater variation of immune cell populations *(17)*.

Macrophages are widely distributed throughout the body and are hence one of the first cells to react to a perturbation of homeostasis. Their functional programs are highly context-dependent and mostly influenced by pathogens, cellular origin (monocyte-derived or embryonic) and tissue cues (including cytokines) *(18, 19)*. Among a continuum of stimulus-dependent polarization states *(20)*, a pro-inflammatory phenotype with pathogen-killing abilities can be induced by lipopolysaccharide (LPS) via Toll-like receptors (TLR), resulting in activation of the central NF-κB signaling pathway *(21)*. These macrophages metabolize arginine to citrulline and nitric oxide through inducible nitric oxide synthase (iNOS, gene: Nos2). The production of the key macrophage cytokine Interleukin 12 (IL-12) promotes a T helper (Th) cell type 1 response *(22,*

*23)*. In contrast, homeostatic macrophages metabolize arginine to ornithine, a precursor of polyamine and hydroxyproline, and urea through Arginase (gene: Arg1), promote wound healing, angiogenesis, and a Th2 response *(22, 23)*. iNOS- and Arg1-centered metabolic pathways are largely exclusive due to cross-inhibition *(24)*. Importantly, genetic variation has been shown to affect macrophage activation by hierarchical function of lineage-determining transcription factors *(25)*.

The hybrid mouse diversity panel (HMDP) is a panel of about 100 inbred mouse strains developed for performing association studies with adequate statistical power and resolution for mapping of complex traits *(26, 27)*. It has been successfully used for investigating gene-environment interaction in activated macrophages *(28)*, insulin sensitivity *(29)*, and susceptibility to atherosclerosis *(30)*. Here, we employ the HMDP as a surrogate model for human immune diversity and investigate natural variation of macrophage activation in response to LPS in mice and humans.

**Results:**

*A spectrum of macrophage activation responses in human and mouse populations*

To determine variations of macrophage LPS responses in a diverse human population, we compared their transcriptional activation patterns. In alveolar macrophages of healthy humans exposed to LPS, large intrinsic variations in responses were detected in NF-κB and TLR pathway activities (Figure 1a,b), indicating different inter-individual sensitivities to LPS stimulation. To model the wide distribution of activation responses in mice, we employed 83 inbred and commercially available mouse strains of the HMDP. Inter-strain analysis of the transcriptome of

thioglycollate-elicited peritoneal macrophages at baseline and after LPS induction showed a similar large range of pathway activations as observed in humans (Figure 1c,d).

Next, the expression levels of selected macrophage activation genes were analyzed. Human alveolar macrophages show considerable differences among healthy volunteers under homeostatic conditions (Figure 1e, left column). After LPS treatment, Nos2, Arg1 and IL-12β were variably upregulated, with IL-12β showing the widest range (1.4-fold to 64-fold) (Figure 1e, bottom right). Similarly, using the HMDP, we found a large diversity of expression levels in peritoneal macrophage transcriptomes. At baseline, Arg1 expression showed a large variation across all strains, whereas Nos2 and IL-12β were expressed at low levels (Figure 1f, left column). After LPS treatment, IL-12β expression increased up to 64-fold change compared to baseline (Figure 1f, right column). Housekeeping genes like β-actin did not change (not shown).

We confirmed expression differences between mouse strains at the protein level. Therefore, a representative panel of 13 mouse strains was developed that mimics the dynamic range of gene expression in the entire HMDP (Supplemental fig. 1). IL-12 p70 (active form with subunits p35 and p40) levels in supernatants of LPS-stimulated peritoneal macrophages (Supplemental fig. 1) differed between strains. Intracellular iNOS and arginase protein expression also varied between strains as shown by flow cytometry (Supplemental fig. 1). Together, these data suggest that the inter-individual degree and direction of macrophage activation in response to LPS varies in human populations and can be recapitulated in a mouse diversity panel.

*Transcriptomic signatures of LPS responsiveness in 83 mouse strains*

Based on these findings, we established an activation factor that describes the distribution of LPS responsiveness in a diverse population. IL-12β and Arg1 are ideal candidates because of their

large transcriptomic variation and their differing role in resting and LPS-activated macrophages *(24)*. We ranked 83 mouse strains by the ratio of IL-12β/Arg1 mRNA expression and found a continuous spectrum of LPS-induced activation (Figure 2a), thus identifying inherent LPS low (like FVB/NJ) and -high (like KK/HIJ) responders (Supplemental table 1). As expected, LPS high-responder strains produced less ornithine (product of Arg1) and more citrulline (product of Nos2) (Supplemental fig. 3), supporting the biological significance of this ranking.

The macrophage transcriptomes of 83 strains were correlated with the activation factor, resulting in a highly significant ranked list of 1272 LPS-responder genes and 2615 LPS non-responder genes (Figure 2b) (Supplemental table 2). These lists were termed $M(LPS)^+$ and $M(LPS)^-$, according to a recent nomenclature proposal *(20)*. $M(LPS)^+$ genes were positively correlated with the activation factor, e.g. IRF5 *(31)*, and $M(LPS)^-$ genes were negatively correlated, e.g. SUMO3 *(32)* (Figure 2c,d). A functional classification showed enrichment for leukocyte migration, chemotaxis, inflammatory response and lymphocyte proliferation in $M(LPS)^+$ genes, whereas in $M(LPS)^-$ genes, apoptosis, DNA repair, cell cycle progression and metabolic genes were dominant (Figure 2e). Gene lists contained 227 and 42 transcription factors in LPS-responders and non-responders, respectively, providing a comprehensive map of the transcription factor landscape associated with LPS responsiveness across all HMDP strains. Known transcription factors such as IRF5, HIF1A, CEBPE, and STAT6 were validated *(33)* and new candidates with high significance are proposed, e.g. TRIM24, KDM5A and TP53 (Supplemental table 3) (Figure 2f).

*Robust correlation with human inflammatory disorders*

To apply the mouse findings to humans, we first tested whether the M(LPS)$^{+}$ and M(LPS)$^{-}$ gene lists sufficiently mapped to the human transcriptome. Human alveolar macrophages expressed about 70% of the mouse-derived gene lists (135 of the top 200 M(LPS)$^{+}$ genes; 142 of the top 200 M(LPS)$^{-}$ genes). As expected, resting human alveolar macrophages showed no significant M(LPS)$^{+}$ gene enrichment under baseline conditions, but a strong shift to M(LPS)$^{+}$ after LPS treatment (Supplemental fig. 4). Similar enrichment was observed using human CD14$^{+}$ monocyte-derived macrophages infected with Listeria monocytogenes (Supplemental fig. 4). Isolated healthy human synovial macrophages were enriched in M(LPS)$^{-}$ genes, whereas macrophages from rheumatoid arthritis patients were strongly skewed towards an M(LPS)$^{+}$ phenotype (Figure 3a). In lupus erythematodes patients, laser-dissected kidney glomerula from kidney biopsies were significantly M(LPS)$^{+}$ enriched in lupus nephritis, but not in healthy conditions (Figure 3b). Peripheral blood leukocytes of systemic inflammatory response syndrome (SIRS) and sepsis patients both at day 1 and 3 after clinical diagnosis showed gradually increasing M(LPS)$^{+}$ enrichment scores, demonstrating clinically relevant sensitivity to infection severity (Supplemental Fig. 4). In these datasets, gene enrichment analysis showed robustness to inter-individual transcriptomic differences between patients.

*Inter-individual prediction of cancer survival*

Macrophages play a decisive role in the tumor microenvironment by promoting or inhibiting tumor growth *(34)*. The monocyte/macrophage content in solid tumors can exceed more than 50% of all infiltrated leukocytes (Supplemental fig. 5) . We used deep transcriptome deconvolution to detect gene signatures in mixed transcriptomes. Under controlled conditions, the M(LPS) phenotypes are readily detectable in LPS- and IL-4 treated macrophages in vitro (Supplemental fig. 5). Gene expression data from tumor biopsies show enriched expression of

M(LPS)$^-$ genes, however, with great variation between patients (Figure 3c, Supplemental fig. 5). To test whether abundant M(LPS)$^+$ expression affects disease outcome, we employed the PRECOG database that ranks genes by correlation with overall survival in human cancers *(35)*. In data from 18,000 biopsies across 39 tumor types, we find that the M(LPS)$^+$ signature significantly correlates with survival, whereas the M(LPS)$^-$ signature correlates with cancer death (Figure 3d). This pattern was significant in many cancers of different ontological origin, e.g. osteosarcoma, melanoma, chronic lymphocytic leukemia, Burkitt lymphoma and large-cell lung carcinoma (Supplemental fig. 6). Patients with high expression of M(LPS)$^+$ or M(LPS)$^-$ genes show increased or decreased survival in multiple tumor entities, respectively (Figure 3e-g). Thus, individual macrophage LPS responsiveness is a predictor of cancer survival.

**Discussion:**

We show that introducing genetic diversity into mouse research reveals a much broader spectrum of innate immune responses with high translational relevance for disease outcomes. The robust correlation of many genes with the activation factor (IL12β/Arg1) across 83 mouse strains reveals that the LPS axis is a major player in macrophage biology. Out of 12,980 genes analyzed, 1272 (9.8%) and 2615 (20.2%) are positively and negatively correlated with the LPS response, respectively. This is not surprising per se. However, it is surprising that only 3 M(LPS)$^+$ genes (Dpep1, Gkn2 and Hoxd4) and 11 M(LPS)$^-$ genes (Arg1, Tpi1, Gpi1, Mif, Pkm, Zmynd8, Pcbp4, Myo6, Grk6, Egln1, Slc44a2) are found in resting (thioglycollate-elicited) peritoneal macrophages. Thus, it is the LPS challenge that brings out the individual propensity to activation rather than an a priori skewed phenotype at baseline. In mouse strains of the HMDP, this obviously originates from genetics, whereas in humans both genetic and environmental factors may play a role *(16)*.

The demonstrated concept of extracting robust and translatable data from the HMDP could be applied to many other fields of immune and cancer research. Importantly, a reduced number of 13 mouse strains were found sufficient to qualitatively represent the diversity in macrophage activation responses to LPS. Thus, testing hypotheses in a representative panel of HMDP strains (Supplemental table 1) may be cost effective and prudent before embarking on a clinical drug development program. All HMDP strains are fully inbred (homozygous at all loci) and commercially available *(27)*, providing immediate access for research facilities and allowing reproducibility studies. Using the HMDP as a translational tool to mimic natural immune variation in human populations could be valuable to increase translatability of mouse data.

A striking finding of this study is that the M(LPS) signatures robustly predict cancer survival or death from mixed-cell biopsy material, containing cancer, stromal and inflammatory cells. This makes these gene panels well suited for predictive tests in personalized medicine. Particularly, with new generation cancer treatments that manipulate tumor-associated macrophage polarization, new diagnostics are necessary to monitor treatment efficacy *(34, 36)*. A simple multiplex PCR or RNA sequencing run could harbor enormous predictive value, matching or exceeding the value of traditional biomarkers or histopathological cancer staging. Of note, in contrast to an increasing number of disease-specific genetic tests that were developed by gene-outcome association statistics, e.g. in breast cancer *(37)*, we extracted genes in a „bottom-up“ approach centering around macrophage biology. Given the ubiquitous disease-relevant role of macrophages, this yielded transcriptomic signatures with predictive value in a number of different disease entities. Furthermore, in contrast to conventional flow cytometry- or PCR-based estimations of macrophage polarization in a few dimensions, the developed gene panels allow a

gradual classification, thus enabling a population-based assessment in high resolution necessary for clinical applications.

Natural diversity of T helper cell differentiation has been initially observed in different mouse strains (e.g. C57BL/6J and Balb/c having a Th1 and Th2 bias, respectively). This was later confirmed in healthy humans, where a Th1 cell bias increases the risk of atherosclerosis *(38)*. In line with the interconnected nature of macrophage and T cell polarization *(22, 39)*, our data extend this concept to macrophages. As the quantitative LPS response of macrophages correlates with disease susceptibility, we posit that understanding inter-individual macrophage activation propensities will be a cornerstone of comprehensive immune phenotyping of health and disease.

**Materials and Methods:**

*Mice*

Male mice 6 - 10 weeks of age were obtained from Jackson Laboratories (Bar Harbor, ME, USA) and housed in pathogen-free conditions on chow diet (Ralston-Purina Co, St. Louis, Mo). Details on the hybrid mouse diversity panel (HMDP) have been reported before *(26, 28)* and data is accessible online (http://systems.genetics.ucla.edu/data). All mouse strains used in this study are listed in supplemental table 1. Experimental procedures were approved by the Institutional Care and Use Committee (IACUC) at the University of California, Los Angeles.

*Macrophage culture and activation*

Murine peritoneal macrophages were elicited with thioglycollate (BD, Sparks, MD, USA; same batch for all strains) for 4 days. Cells from up to 4 mice were pooled and plated at 4 x $10^6$ cells/ml in DMEM + 20% FBS + 1% streptomycin/penicillin in duplicates or triplicates. After

overnight culture, cells were washed in PBS and incubated with 2 ng/ml LPS (List Biological Inc., Campbell, CA, USA) or control in DMEM + 1% FBS for 4h before harvest. Viability of cultured macrophages was determined for some strains by staining with 2 μM calcein AM (Molecular Probes) and was greater than 90% with no difference between treated and untreated conditions *(28)*. Multiple replicates of some strains allowed determination of experimental variability as previously described *(28)*.

*Gene expression profiling*

Total RNA was obtained using RNeasy columns (QIAGEN, Valencia, CA, USA) with DNA digest according to manufacturer's instructions and subsequently hybridized to Affymetrix HT MG-430A chip arrays. Chip signals were transformed to robust multi-array average (RMA). Raw microarray data for peritoneal macrophages of the HMDP is deposited under the NCBI GEO accession number GSE38705 *(28)*.

*Amino acid detection*

For 13 selected mouse strains, peritoneal macrophages were cultured in the presence of 2 ng/ml Lipopolysaccharide in DMEM + 1% FBS media for 4 hours. Supernatant was collected after 0, 4, or 24 hours and stored at -80°C. Supernatant was analyzed for amino acid concentration using a Hitachi L-8900 Amino Acid Analyzer.

*Flow cytometry and cytometric bead array*

After LPS stimulation as described above, macrophages were harvested after removal of supernatant and washing with PBS. For flow cytometry, cells were stained with viability dye (Ghost UV450 or Ghost Red710 Tonbo Bioscience, San Diego) using an intracellular staining

protocol (IC fixation and permeabilization staining kit, eBioscience). Antibodies included iNOS (CXNFT, eBioscience), CD11b (M1/70, eBioscience), F4/80 (BM8, eBioscience), Arginase 1 (Polyclonal, R&D Systems). Macrophages were analyzed by flow cytometry on an LSRII instrument using FACS Diva software (BD Bioscience) and analyzed using FlowJo software (Tree Star, San Carlos, CA). The cytometric bead array (BD Bioscience) was used to determine cytokine concentrations in macrophage supernatant after LPS treatment according to manufacturers' instructions.

*Activation factor and correlated genes*

The activation factor describes the macrophage activation propensity in a diverse population based on Affymetrix-RMA gene expression values. For each mouse strain, IL-12β was divided by Arg1 and averaged to the population's baseline average (Equation 1). This was performed for both baseline and LPS-treated conditions. High and low numbers indicate high and low LPS responsiveness, respectively.

$$Activation\ factor_{IL12\beta} = \frac{IL12\beta}{Arg1} * \frac{Arg1_{mean}}{IL12\beta_{mean}}$$

*Equation 1:* Activation factor based on gene expression values ($log_2$). IL12β = Interleukin-12 beta, Arg1 = Arginase-1.

Each transcript of the Affymetrix chip (total of 39,000 probesets) was correlated to the activation factor in 83 mouse strains. Positive and negative correlations with $p < 0.0001$ (Pearson) and FDR

< 5% (Benjamini Hochberg) were selected as $M(LPS)^+$ and $M(LPS)^-$ gene signature lists, respectively (supplemental table 2).

*Gene enrichment analysis*

The gene set enrichment analysis (GSEA) tool *(40)* detects significant enrichment of a defined gene list in differentially expressed genes of two datasets (e.g. healthy vs. disease). This allows to determine whether genes in question are specifically up- or downregulated in either condition. While the algorithm only allows comparing two conditions, it can integrate many samples per conditions by averaging gene expression values. The algorithm calculates a running sum over the entire transcriptome and each match (shown as bar code under the graph) between a gene of the gene set and the transcriptome increases the sum as a function of DE-expression, resulting in a positive or negative deflection of the graph. Positive or negative deflection of the graph per se is not meaningful because this only reflects the orientation of the 2 datasets (A vs. B or B vs. A). Over- or under-representation is expressed as normalized enrichment score (NES) that takes gene set size into account. The first matches of genes are referred to as leading edge since these genes mainly drive the enrichment score. Here, as gene lists, we used signatures derived from the top correlated genes with the activation factor. In most analyses, the top 200 genes for M(LPS)+ or M(LPS)- were used unless otherwise indicated. Standard settings (weighted, 100 iterations) of GSEA version 2.0.1 in the Gene Pattern interface were used. For enrichment in a single dataset (e.g. PRECOG database), the pre-weighted GSEA algorithm was employed. Here, a single dataset is ranked by expression value and enrichment indicates over-representation in high or low expressed genes.

*Pathway enrichment and transcription factor determination*

Ingenuity's pathway analyzer (IPA, Qiagen) was used to analyze pathway enrichment in M(LPS) signatures. Activity (activation/inhibition) of canonical pathways is predicted based on differentially expressed genes in the input dataset. As input datasets, baseline and LPS treated macrophage transcriptomes were used in figure 1. Up- or down-regulation of pre-defined pathway genes are determined by IPA and the overlap is expressed as p-value overlap. A positive and negative z-score indicates pathway activation and inhibition, respectively. Transcription factors were extracted from the M(LPS) signatures and the p-value overlap and activation z-score was calculated using IPA's upstream regulator analysis. Z-scores greater than 2 and less than -2 are considered significant.

*RNA deconvolution*

For RNA deconvolution the CIBERSORT algorithm was used *(41)*. It allows to estimate the fraction of a cell type of interest in a bulk transcriptome, and was developed for high robustness and to discriminate closely related cell types. As input, it requires a transcriptome (mixture) to be deconvoluted and a signature consisting of several hundred genes with expression values. Based on a machine learning approach, the relative proportion of the signature is determined in the mixture as percentage and a global p value indicates the significance of correlation. It can be used for different units of gene expression such as RPKM (linear) or RMA ($log_2$) as long as the units in the mixture and signature are consistent. Core signatures used in this study included either the provided LM22 signature of several leukocyte subsets (e.g. naive B cells, memory B cells, plasma cells, naive CD4 T cells, CD4 memory cells, follicular helper T cells, γδ T cells, NK cells, monocytes, macrophages, dendritic cells, mast cells, eosinophils, neutrophils) or the HMDP-derived M(LPS) gene lists (top 200 genes). Settings used included 100 permutations and quantile normalization.

*Survival analyses*

For survival analyses, we performed gene set enrichment analysis on the PRECOG database that ranks genes by clinical survival either in a collapsed pan-cancer or in tumor-specific dataset *(35)*. Leading edge genes in gene set enrichment analyses are defined as most enriched genes in a particular condition, e.g. good or poor survival. Genes were picked based on the leading edge genes in gene enrichment analyses (Supplemental fig. 6). The study population was divided by the median of the mean expression of the tumor-specific gene list. Survival data was analyzed using PROGgene2 *(42)* and plotted as median without sub-cohort division in a Kaplan Meier format. Significance was calculated using the log-rank test.

*Statistics*

Statistical analysis was performed using GraphPad Prism (GraphPad Software Inc., San Diego). Affymetrix gene chip data was normalized using the robust multi-array average (RMA) method (= $\log_2$, background corrected, quantile normalized). Correlation analyses were based on Pearson's correlation unless otherwise indicated. The threshold for significant $M(LPS)^+$ or $M(LPS)^-$ genes was set as $p < 0.0001$ and FDR < 5% (Benjamini Hochberg) for the correlation between gene expression (RMA) and activation factor.

**Supplemental Data:**

Supplemental fig. 1:

Analysis of IL-12, Arginase and iNOS protein expression in a representative panel of HMDP strains.

Supplemental fig. 2:

Arg1, IL-12 and Nos2 gene expression in 83 strains of the HMDP.

Supplemental fig. 3:

Amino acid detection in a representative panel of HMDP strains.

Supplemental fig. 4:

Validation of M(LPS) gene signatures in human transcriptomes.

Supplemental fig. 5:

RNA deconvolution using M(LPS) gene signatures.

Supplemental fig. 6:

Gene set enrichment analysis of M(LPS) genes in human tumor transcriptomes.

**References:**

1. H. Ledford, Translational research: 4 ways to fix the clinical trial, *Nature* **477**, 526–528 (2011).

2. I. W. Mak, N. Evaniew, M. Ghert, Lost in translation: animal models and clinical trials in cancer treatment, *Am. J. Transl. Res.* **6**, 114–118 (2014).

3. J. Arrowsmith, Trial watch: Phase II failures: 2008–2010, *Nat. Rev. Drug Discov.* **10**, 328–329 (2011).

4. N. J. Schork, Personalized medicine: Time for one-person trials, *Nature* **520**, 609–611 (2015).

5. J. Mestas, C. C. W. Hughes, Of mice and not men: differences between mouse and human immunology, *J. Immunol. Baltim. Md 1950* **172**, 2731–2738 (2004).

6. H. B. van der Worp, D. W. Howells, E. S. Sena, M. J. Porritt, S. Rewell, V. O'Collins, M. R. Macleod, Can Animal Models of Disease Reliably Inform Human Studies?, *PLoS Med.* **7** (2010), doi:10.1371/journal.pmed.1000245.

7. Hackam DG, Redelmeier DA, Translation of research evidence from animals to humans, *JAMA* **296**, 1727–1732 (2006).

8. P. Pound, S. Ebrahim, P. Sandercock, M. B. Bracken, I. Roberts, Where is the evidence that animal research benefits humans?, *BMJ* **328**, 514–517 (2004).

9. C. R. Hooijmans, M. Ritskes-Hoitinga, Progress in Using Systematic Reviews of Animal Studies to Improve Translational Research, *PLOS Med* **10**, e1001482 (2013).

10. J. S. Tsang, P. L. Schwartzberg, Y. Kotliarov, A. Biancotto, Z. Xie, R. N. Germain, E. Wang, M. J. Olnes, M. Narayanan, H. Golding, S. Moir, H. B. Dickler, S. Perl, F. Cheung, Global analyses of human immune variation reveal baseline predictors of post-vaccination responses, *Cell* **157**, 499–513 (2014).

11. J.-L. Casanova, L. Abel, The human model: a genetic dissection of immunity to infection in natural conditions, *Nat. Rev. Immunol.* **4**, 55–66 (2004).

12. S. J. Chapman, A. V. S. Hill, Human genetic susceptibility to infectious disease, *Nat. Rev. Genet.* **13**, 175–188 (2012).

13. L. Råberg, D. Sim, A. F. Read, Disentangling genetic variation for resistance and tolerance to infectious diseases in animals, *Science* **318**, 812–814 (2007).

14. S. Davila, V. J. Wright, C. C. Khor, K. S. Sim, A. Binder, W. B. Breunis, D. Inwald, S. Nadel, H. Betts, E. D. Carrol, R. de Groot, P. W. M. Hermans, J. Hazelzet, M. Emonts, C. C. Lim, T. W. Kuijpers, F. Martinon-Torres, A. Salas, W. Zenz, M. Levin, M. L. Hibberd, International Meningococcal Genetics Consortium, Genome-wide association study identifies variants in the CFH region associated with host susceptibility to meningococcal disease, *Nat. Genet.* **42**, 772–776 (2010).

15. H. I. Nakaya, J. Wrammert, E. K. Lee, L. Racioppi, S. Marie-Kunze, W. N. Haining, A. R. Means, S. P. Kasturi, N. Khan, G.-M. Li, M. McCausland, V. Kanchan, K. E. Kokko, S. Li, R. Elbein, A. K. Mehta, A. Aderem, K. Subbarao, R. Ahmed, B. Pulendran, Systems biology of vaccination for seasonal influenza in humans, *Nat. Immunol.* **12**, 786–795 (2011).

16. P. Brodin, V. Jojic, T. Gao, S. Bhattacharya, C. J. L. Angel, D. Furman, S. Shen-Orr, C. L. Dekker, G. E. Swan, A. J. Butte, H. T. Maecker, M. M. Davis, Variation in the Human Immune System Is Largely Driven by Non-Heritable Influences, *Cell* **160**, 37–47 (2015).

17. L. K. Beura, S. E. Hamilton, K. Bi, J. M. Schenkel, O. A. Odumade, K. A. Casey, E. A. Thompson, K. A. Fraser, P. C. Rosato, A. Filali-Mouhim, R. P. Sekaly, M. K. Jenkins, V. Vezys, W. N. Haining, S. C. Jameson, D. Masopust, Normalizing the environment recapitulates adult human immune traits in laboratory mice, *Nature* **532**, 512–516 (2016).

18. T. A. Wynn, A. Chawla, J. W. Pollard, Macrophage biology in development, homeostasis and disease, *Nature* **496**, 445–455 (2013).

19. E. L. Gautier, T. Shay, J. Miller, M. Greter, C. Jakubzick, S. Ivanov, J. Helft, A. Chow, K. G. Elpek, S. Gordonov, A. R. Mazloom, A. Ma'ayan, W.-J. Chua, T. H. Hansen, S. J. Turley, M. Merad, G. J. Randolph, Gene expression profiles and transcriptional regulatory pathways underlying mouse tissue macrophage identity and diversity, *Nat. Immunol.* **13**, 1118–1128 (2012).

20. P. J. Murray, J. E. Allen, S. K. Biswas, E. A. Fisher, D. W. Gilroy, S. Goerdt, S. Gordon, J. A. Hamilton, L. B. Ivashkiv, T. Lawrence, M. Locati, A. Mantovani, F. O. Martinez, J.-L. Mege, D. M. Mosser, G. Natoli, J. P. Saeij, J. L. Schultze, K. A. Shirey, A. Sica, J. Suttles, I. Udalova, J. A. van Ginderachter, S. N. Vogel, T. A. Wynn, Macrophage activation and polarization: nomenclature and experimental guidelines, *Immunity* **41**, 14–20 (2014).

21. O. Takeuchi, S. Akira, Pattern recognition receptors and inflammation, *Cell* **140**, 805–820 (2010).

22. C. D. Mills, K. Kincaid, J. M. Alt, M. J. Heilman, A. M. Hill, M-1/M-2 macrophages and the Th1/Th2 paradigm, *J. Immunol. Baltim. Md 1950* **164**, 6166–6173 (2000).

23. S. K. Biswas, A. Mantovani, Macrophage plasticity and interaction with lymphocyte subsets: cancer as a paradigm, *Nat. Immunol.* **11**, 889–896 (2010).

24. M. Rath, I. Müller, P. Kropf, E. I. Closs, M. Munder, Metabolism via Arginase or Nitric Oxide Synthase: Two Competing Arginine Pathways in Macrophages, *Front. Immunol.* **5**, 532 (2014).

25. S. Heinz, C. E. Romanoski, C. Benner, K. A. Allison, M. U. Kaikkonen, L. D. Orozco, C. K. Glass, Effect of natural genetic variation on enhancer selection and function, *Nature* **503**, 487–492 (2013).

26. B. J. Bennett, C. R. Farber, L. Orozco, H. M. Kang, A. Ghazalpour, N. Siemers, M. Neubauer, I. Neuhaus, R. Yordanova, B. Guan, A. Truong, W. Yang, A. He, P. Kayne, P. Gargalovic, T. Kirchgessner, C. Pan, L. W. Castellani, E. Kostem, N. Furlotte, T. A. Drake, E. Eskin, A. J. Lusis, A high-resolution association mapping panel for the dissection of complex traits in mice, *Genome Res.* **20**, 281–290 (2010).

27. A. J. Lusis, M. Seldin, H. Allayee, B. J. Bennett, M. Civelek, R. C. Davis, E. Eskin, C. Farber, S. T. Hui, M. Mehrabian, F. Norheim, C. Pan, B. Parks, C. Rau, D. J. Smith, T. Vallim, Y. Wang, J. Wang, The Hybrid Mouse Diversity Panel: A Resource for Systems Genetics Analyses of Metabolic and Cardiovascular Traits, *J. Lipid Res.* (2016).

28. L. D. Orozco, B. J. Bennett, C. R. Farber, A. Ghazalpour, C. Pan, N. Che, P. Wen, H. X. Qi, A. Mutukulu, N. Siemers, I. Neuhaus, R. Yordanova, P. Gargalovic, M. Pellegrini, T. Kirchgessner, A. J. Lusis, Unraveling inflammatory responses using systems genetics and gene-environment interactions in macrophages, *Cell* **151**, 658–670 (2012).

29. B. W. Parks, T. Sallam, M. Mehrabian, N. Psychogios, S. T. Hui, F. Norheim, L. W. Castellani, C. D. Rau, C. Pan, J. Phun, Z. Zhou, W.-P. Yang, I. Neuhaus, P. S. Gargalovic, T. G.

Kirchgessner, M. Graham, R. Lee, P. Tontonoz, R. E. Gerszten, A. L. Hevener, A. J. Lusis, Genetic architecture of insulin resistance in the mouse, *Cell Metab.* **21**, 334–346 (2015).

30. B. J. Bennett, R. C. Davis, M. Civelek, L. Orozco, J. Wu, H. Qi, C. Pan, R. R. S. Packard, E. Eskin, M. Yan, T. Kirchgessner, Z. Wang, X. Li, J. C. Gregory, S. L. Hazen, P. S. Gargalovic, A. J. Lusis, Genetic Architecture of Atherosclerosis in Mice: A Systems Genetics Analysis of Common Inbred Strains, *PLoS Genet.* **11**, e1005711 (2015).

31. T. Krausgruber, K. Blazek, T. Smallie, S. Alzabin, H. Lockstone, N. Sahgal, T. Hussell, M. Feldmann, I. A. Udalova, IRF5 promotes inflammatory macrophage polarization and TH1-TH17 responses, *Nat. Immunol.* **12**, 231–238 (2011).

32. T.-H. Chang, S. Xu, P. Tailor, T. Kanno, K. Ozato, The small ubiquitin-like modifier-deconjugating enzyme sentrin-specific peptidase 1 switches IFN regulatory factor 8 from a repressor to an activator during macrophage activation, *J. Immunol. Baltim. Md 1950* **189**, 3548–3556 (2012).

33. T. Lawrence, G. Natoli, Transcriptional regulation of macrophage polarization: enabling diversity with identity, *Nat. Rev. Immunol.* **11**, 750–761 (2011).

34. Y. Komohara, M. Jinushi, M. Takeya, Clinical significance of macrophage heterogeneity in human malignant tumors, *Cancer Sci.* **105**, 1–8 (2014).

35. A. J. Gentles, A. M. Newman, C. L. Liu, S. V. Bratman, W. Feng, D. Kim, V. S. Nair, Y. Xu, A. Khuong, C. D. Hoang, M. Diehn, R. B. West, S. K. Plevritis, A. A. Alizadeh, The prognostic landscape of genes and infiltrating immune cells across human cancers, *Nat. Med.* **21**, 938–945 (2015).

36. A. Mantovani, P. Allavena, The interaction of anticancer therapies with tumor-associated macrophages, *J. Exp. Med.* **212**, 435–445 (2015).

37. J. A. Lynch, V. Venne, B. Berse, Genetic tests to identify risk for breast cancer, *Semin. Oncol. Nurs.* **31**, 100–107 (2015).

38. N. C. Olson, R. Sallam, M. F. Doyle, R. P. Tracy, S. A. Huber, T Helper Cell Polarization in Healthy People: Implications for Cardiovascular Disease, *J Cardiovasc. Transl. Res.* **6**, 772–786 (2013).

39. C. D. Mills, K. Ley, M1 and M2 macrophages: the chicken and the egg of immunity, *J. Innate Immun.* **6**, 716–726 (2014).

40. A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, J. P. Mesirov, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–15550 (2005).

41. A. M. Newman, C. L. Liu, M. R. Green, A. J. Gentles, W. Feng, Y. Xu, C. D. Hoang, M. Diehn, A. A. Alizadeh, Robust enumeration of cell subsets from tissue expression profiles, *Nat. Methods* **12**, 453–457 (2015).

42. C. P. Goswami, H. Nakshatri, PROGgeneV2: enhancements on the existing database, *BMC Cancer* **14**, 970 (2014).

**Funding:** This project was funded by NIH R01HL115232 to K.L., NIH HL28481 and HL30568 to A.J.L., and Deutsche Forschungsgemeinschaft (DFG) BU3247/1-1 to K.B.

**Author contributions:** K.B., E.E., and P.G. performed most experiments and analyses. A.B.P. and C.P. performed additional bioinformatics. E.E., G.T., and D.W. did flow cytometry experiments. K.L., K.B., C.D.M. and A.J.L. designed the study. K.L. and K.B. wrote the manuscript.

**Competing interests:** A patent on clinical application of M(LPS) gene lists has been filed.

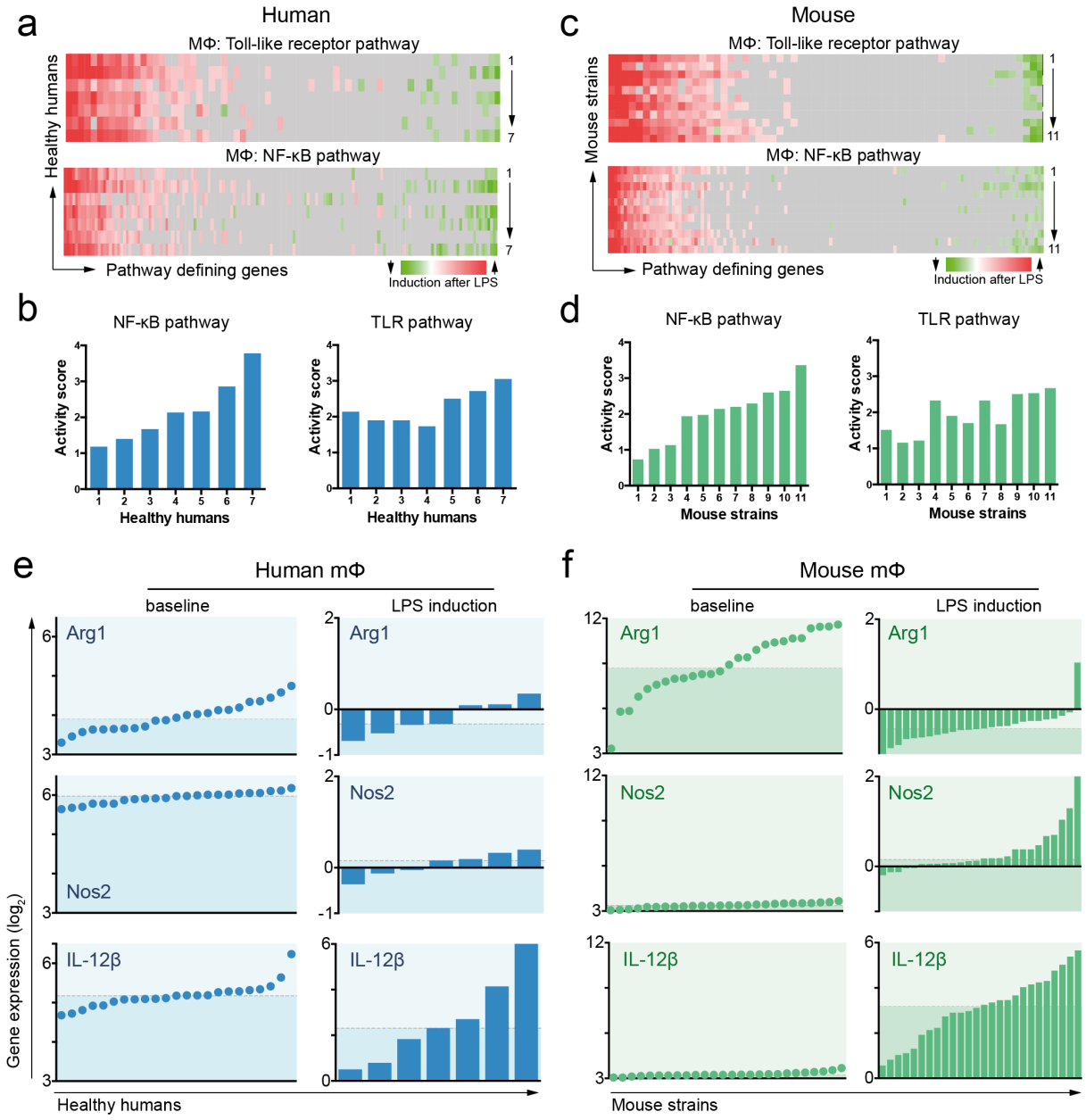



**Fig. 1. Natural variation of macrophage activation in response to LPS in humans and inbred mouse strains. a-f**, Human alveolar macrophages and thioglycollate-elicited peritoneal macrophages of the hybrid mouse diversity panel were sequenced after control and LPS treatment. **a-d**, Activity and induction after LPS treatment of the Toll like receptor- (total 69

genes) and NF-κB signaling (total 155 genes) pathways as determined by Ingenuity are shown (z-score). Data sorted ascendingly. **e,f**, Arg1 (arginase), Nos2 (iNOS), and IL-12β gene expression ($log_2$) and induction after LPS treatment (right columns). The vertical lines indicate the median. All data is sorted ascendingly. Human data from GSE27002 and GSE40885. Only some of 83 strains are shown for clarity (Supplemental fig. 2).

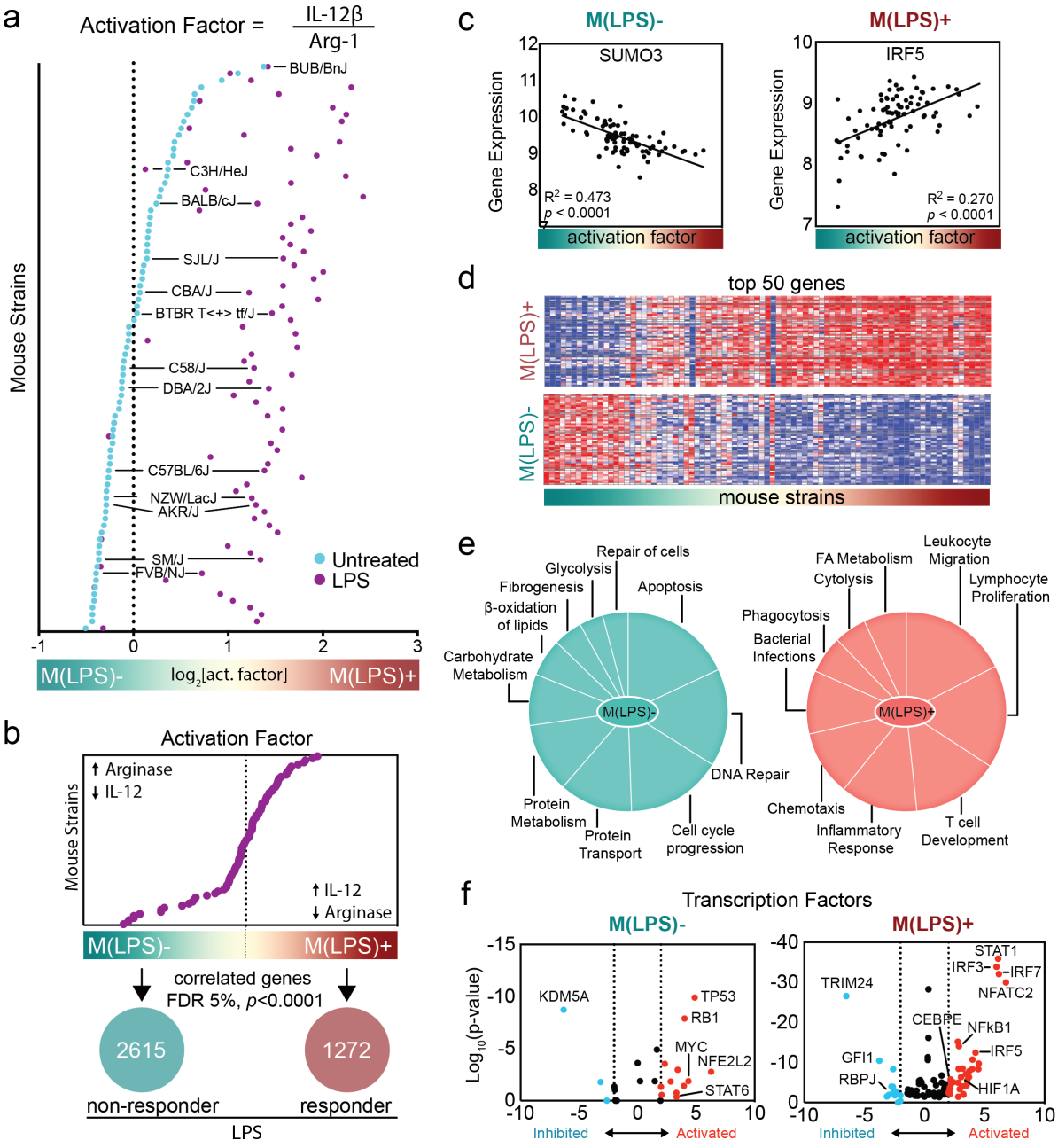


**Fig. 2. Core signatures of LPS responsiveness in macrophages of 83 mouse strains.**

**a,b**, Calculation of the activation factor ($log_2$) that represents the strain-specific LPS responsiveness of peritoneal macrophages based on a IL12β/Arg1 gene expression ratio before (blue) and after (purple) LPS treatment. Vertical line = median at baseline. The transcriptome was correlated with the activation factor (FDR < 5%, $p < 0.0001$). Positively and negatively correlated genes are designated $M(LPS)^+$ and $M(LPS)^-$, respectively. Ranking details in Supplemental Figure 1. Gene lists are provided in supplemental table 2. **c,d**, Correlation (Pearson) of genes with the activation factor that ranks strains from low to high LPS responsiveness. Each data point represents one mouse strain. Heatmap blue = down-, red = upregulation. **e**, Key biological processes enriched in M(LPS) gene signatures. The size of each segment indicates the significance of enrichment as $-log_{10}$(p-value). $p <= 0.001$. **f**, Transcription factors (TF) were extracted from M(LPS) lists and their activation score was calculated using Ingenuity. The default cutoff for activation is +/- 2 (dashed lines). Complete list in supplemental table 3.

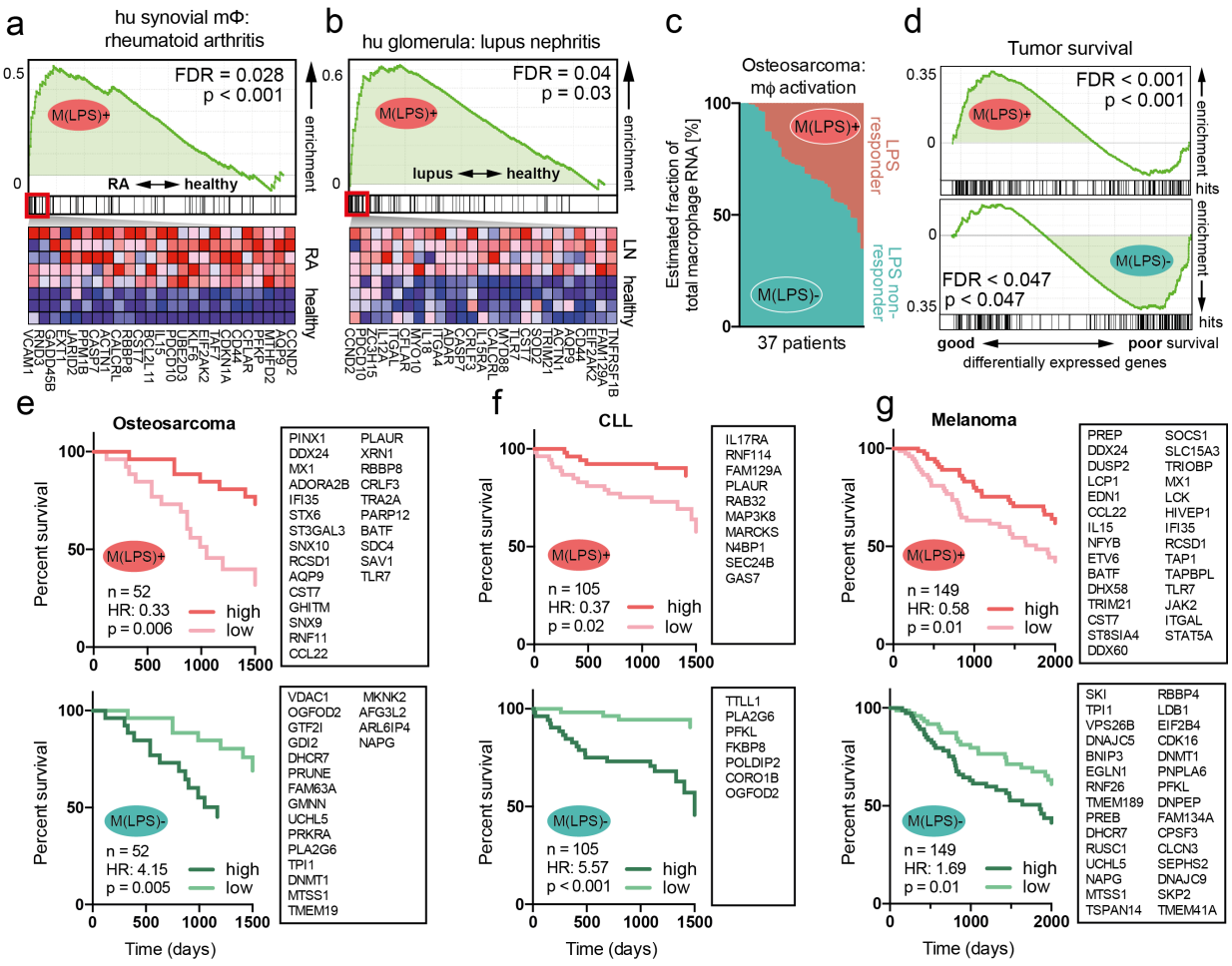



**Fig. 3. Robust translation of mouse-derived signatures to patients with diagnostic and prognostic value in inflammatory and malignant diseases. a,b**, Gene enrichment of the M(LPS)$^{+}$ signature in human synovial CD14+ macrophages from healthy (n=3) vs. rheumatoid arthritis (RA) patients (n=5) (GSE10500) and in glomerula from healthy controls (n=14) vs. patients with lupus nephritis (LN, n=32) (GSE32591). Heatmap shows 25 leading edge genes (red = up-, blue = downregulated). **c**, Fraction of M(LPS)$^{+}$ and M(LPS)$^{-}$ phenotypes in human

osteosarcoma biopsies determined by RNA deconvolution. Data from GSE39055. **d**, Gene enrichment of M(LPS) signatures in the PRECOG dataset (genes ranked by survival in collapsed data of ≈ 18.000 human tumor biopsies of 39 cancer entities *(35)*). **e-g**, Kaplan-Meier graphs for human cancer survival as a function of M(LPS) gene expression in the bulk biopsy transcriptome. The tumor-specific gene list (based on GSEA leading edge genes) is denoted in the adjacent boxes. Data from GSE21257 (Osteosarcoma), GSE22762 (CLL, chronic lymphocytic leukemia), and SKCM-TCGA (Melanoma).

*Chapter 3: The Role of CD200 and CD200R1 in Cardiac Homeostasis and Injury*

**Brief Overview of the Epidemiology of Heart Failure (HF)**

Over the past 50 years, there have been remarkable advances in the prevention, diagnosis, and management of cardiovascular disease (CVD). Overall, CVD-related deaths have declined by about two-thirds in industrialized nations (1). Mortality rates in acute coronary syndrome (ACS), valvular and congenital heart disease, uncontrolled hypertension and some arrhythmias have fallen considerably (1).

Unfortunately, heart failure (HF) is a notable exception to the otherwise hopeful and promising trends in CVD. Annual hospital discharges in patients with a primary diagnosis of HF have risen steadily since the 1970s, and in 2009, exceeded 1 million discharges per year (2,3). Discharges related to HF may finally be leveling off or even declining in the United States (4,5) **(Figure 1).** There have been modest improvements in HF mortality. Survival in HF has improved over the past 30 years. In addition, the age-adjusted death rate has declined and the mean age at death from HF is now higher (6,7). However, despite these encouraging trends, the 5-year mortality in HF is still 50%, rivaling or even worse than many deadly cancers (8).

HF primarily affects individuals above the age of 60 and its prevalence rises with age **(Figure 2).** The direct and indirect costs of HF in the Unites States are astounding. In 2010, they were estimated to be $39.2 billion (9). The estimated lifetime cost of HF for one individual patient is around $110,000 per year (10). In hospital care and multiple readmissions due to inadequate therapy are more than three-fourths of this cost. For example, among Medicare patients, the 30-day readmission rate after hospital discharge is 20% to 25% (11).

Although there are slow but progressive improvements in the prognosis of HF, the prevalence of this condition remains high. Some believe that it is the gradual improvement in management and the lack of a definitive therapy that keeps the prevalence climbing. No matter what the reason for the increasing number of individuals with this disease, with the continued aging of the population, HF will remain a major health problem in both industrialized nations and the developing world.

**Mechanisms of HF**

There are many mechanisms that lead to the hemodynamic compromise and resulting clinical picture of HF. In 1967, Eugene Braunwald and his colleagues defined HF as "a clinical syndrome characterized by well known symptoms and physical signs. It is the pathological state in which an abnormality of myocardial function is responsible for the failure of the heart to pump blood at a rate commensurate with the requirements of the metabolizing tissues during ordinary activity" (12).

The most important hemodynamic changes in HF are secondary to ventricular remodeling. Ventricular remodeling is common in patients with chronic dysfunction of the ventricular pump and it varies by HF type (13). For example, in patients with HF with preserved ejection fraction (HFPEF), the size and volume of the left ventricular (LV) cavity is typically normal, but the wall is thickened (14). In contrast, patients with HF with reduced ejection (HFREF), the LV cavity is typically dilated and the wall is typically normal in thickness or even thinned (15).

There are many accepted mechanisms or models of cardiac injury leading to ventricular remodeling, hemodynamic compromise and the clinical syndrome of HF, but two are thought to be the most prominent: the cardiorenal model and the neurohormonal model (16,17). The cardiorenal model emphasizes the close interplay between the heart and the kindey. In the cardiorenal model, renal sodium and water retention are fundamental components of the HF syndrome as they can lead directly to dyspnea and peripheral edema, which are two primary manifestations of clinical HF.

The neurohormal model is secondary to the activation of the adrenergic nervous system, which is an important regulator of cardiac performance during exertion in the healthy heart. The adrenergic nervous system allows the heart to increase myocardial contractility and reallocates cardiac output during exercise (16,17). In acute HF, adrenergic activation is necessary and life saving. Enhanced myocardial contractility in the failing heart is stimulated by the adrenergic nervous system and in addition, the stimulation of the renin-angiotensin-aldosterone system leads to increased concentrations of plasma angiotensin II. Angiotensin II is a potent vasoconstrictor of the renal efferent arterioles and systemic circulation. Systemic vasoconstriction raises the blood pressure and aids in the perfusion of vital organs.

However, prolonged activation of the adrenergic nervous system and of the renin-angiotensin-aldosterone system causes maladaptive remodeling of the ventricles, the LV in particular. Prolonged activation of these two systems can also lead to further myocardial injury, which then initiates the vicious cycle of the neurohormonal model of HF. The majority of HF therapies today target the neurohormonal pathways. Blockade of both the adrenergic nervous

system and the renin-angiotensin-aldosterone systems prolongs survival in patients with HF **(Figure 3).**

**HF Management and Available Therapies**

Over the last 30 years or so, the treatment of HFREF has revolutionized with the development of 3 classes of drugs: angiotensin Converting Enzyme (ACE) inhibitors/angiotensin II receptor blockers (ARBs), aldosterone antagonists, and beta-adrenergic blockers. In addition, the introduction of internal cardiac defibrillators (ICDs) and cardiac resynchronization therapy has substantially improved the quality of life for these patients.

In 1994, the US Food and Drug Administration approved the use of a pneumatically driven pulsatile-flow left ventricular assist device (LVAD) in advanced HF patients as a bridge to cardiac transplantation. The technology evolved rapidly to an electrical, nonpulsatile, continuous flow device that are much smaller, have only a single moving part and are more energy efficient. Newer generation devices impose a lower perioperative risk and result in more favorable long-term survival than their earlier predecessors (18).

A favorable and somewhat surprising effect of chronic LVAD support has been the substantial reverse remodeling of the heart that can occur (19,20). LV size and mass are reduced, ejection fraction (EF) rises, and there is sometimes regression of LV myocyte hypertrophy (21). Although the reverse remodeling seen with LVAD use is extremely encouraging, device-related thrombotic, hemorrhagic and infectious complications still exist and can be devastating or even fatal for our patients (22,23).

All of the abovementioned advances and therapies came about as a result of years of preclinical and clinical research culminating in large-scale, multicenter clinical trials. The results of these trials are then reflected in changes in the practice guidelines put forth by the American Heart Association (AHA) and the American College of Cardiology (ACC) (24-26), which have now become the standard of care for patients with HF.

While AHA/ACC guideline directed medical therapy improves 5-year mortality; cardiac transplant remains the only definitive therapy for patients with end-stage HF. Although survival after cardiac transplant has improved considerably over the past 30 years, it is far from a perfect solution due to many reasons, one being the limited supply of donor hearts (27). New therapies targeting novel, lesser-known pathways in myocardial injury and remodeling are needed to improve the care of the cardiovascular disease patient. One aspect of HF that has been investigated in the past, but has not been successfully harvested for its therapeutic potential is the inflammatory pathway in HF.

**Inflammation after Cardiac Injury**

Inflammatory pathways are critically involved in dilative and fibrotic remodeling of the infarcted heart and, therefore, drive key events in the pathogenesis of heart failure after myocardial injury. Repair of the injured myocardium can be described in three overlapping phases: the inflammatory phase, the proliferative phase, and the maturation phase (28). Irrespective of the cause of injury, the acute sudden death of cardiomyocytes in the injured heart rapidly activates innate immune pathways that trigger an intense, but transient, inflammatory reaction. This response clears the injured area of dead cells and extracellular matrix debris, and is then actively repressed to prepare the infarct for the proliferative phase of healing (29,30).

Numerous immune cell types including mast cells, dendritic cells, B cells, regulatory T cells and macrophages are found within cardiac tissue. The majority of immune cells found in the resting heart are resident macrophages, which are found primarily surrounding endothelial cells but are also seen in the interstitium amongst cardiomyocytes (31). Recent work has established that recruited monocyte-derived macrophages have a pathological role in the setting of sterile cardiac injury, while embryonically-derived resident macrophages have an important role in the tissue repair response (32-39).

There are two main subsets of circulating monocytes in mice: $LY6C^{hi}$ monocytes and $LY6C^{low}$ monocytes. Global transcriptional profiling has shown these monocyte subsets are conserved in humans and they have different roles *in vivo* (40). $LY6C^{low}$ monocytes adhere to and travel along the endothelium clearing damaged cells. In addition, $LY6C^{low}$ monocytes also trigger inflammatory responses without even entering the tissue (41,42). In contrast, $LY6C^{hi}$ monocytes are recruited to the myocardium during cardiac injury or stress (43-47). Monocyte recruitment is partly dependent on innate B cells, which are also recruited to the injured myocardium. Using CC-chemokine ligand 7 (CCL7), these B cells drive monocyte population expansion. Once the monocytes enter the injured myocardium, they begin to differentiate into macrophages (48).

Non-selective depletion strategies have shown that in the absence of both monocytes and macrophages, scar formation is impaired after ischemic injury of the myocardium (32,34). Experiments using apolipoprotein –E (ApoE) deficient mice reveal that uninhibited expansion of macrophages after myocardial injury can have a detrimental effect on infarct healing, leading to excessive inflammation, further myocardial damage, unwanted ventricular remodeling and impaired cardiac function (35). Another study using $Ccr2^{-/-}$ mice, which have significantly less

circulating monocytes, showed that monocyte recruitment and the associated inflammatory response leads to increased cardiac injury and decreased cardiac function (36-38). It has been shown that the loss of LY6C$^{hi}$ monocytes also prevents hypertension-induced cardiac fibrosis and improves left ventricular function after myocardial infarction (52). Taken together, these data support that these two monocyte subsets have very different roles in myocardial injury. Derivation of the therapeutic potential of these separate, very important monocyte subsets is yet to be realized.

**Inflammation as a Biomarker in HF**

In 1956, Eugene Braunwald and his colleagues described the elevation of C-reactive protein (CRP), an inflammatory biomarker, in HF (53). An increase in tumor necrosis factor-α (TNF-α) and interleukin-6 (IL-6) are also reported in advanced HF (54) **(Figure 4).** The presence and the levels of CRP, TNF-α and IL-6 are correlated with severity of HF. They are also considered to be independent predictors of outcome in HF and in the future, could become useful in testing novel anti-inflammatory therapies in HF patients.

**Therapies Targeting the Inflammatory Pathway of HF do not exist**

Inflammation is crucial to the healing myocardium after injury. It is also well accepted that inflammation in general, whether systemic or local, affects almost all HF patients. However, the cause and effect has not been proven; although one could make the argument that the inhibition of excessive inflammation in HF is likely to be beneficial.

Research in the early 1990s provided strong evidence implicating inflammatory activation as an important pathway in disease progression in chronic HF. Seminal clinical data found that elevated plasma cytokine levels correlated with HF in patients and subsequent animal

experiments suggested that certain anti-inflammatory therapies may be beneficial (55,56). This led to substantial efforts to translate these findings to humans through large clinical trials; however the results of these trials were largely disappointing due to either neutral findings or worsening HF.

Based on studies identifying a potential role of the proinflammatory cytokine TNFα as a mediator of disease progression in the failing heart, several randomized placebo-controlled trials of anti-TNFα therapies have been performed. The RENEWAL programme included two trials, RENAISSANCE and RECOVER, both of which tested etanercept, whereas the ATTACH trial used infliximab (57-60). None of these studies showed a beneficial effect and there were even indications of higher rates of mortality with such treatment.

In 2008, the Advanced Chronic Heart Failure Clinical Assessment of Immune Modulation Therapy (ACCLAIM) trial tested the effects of immunomodulation therapy in patients with chronic HF. The rationale was to have a non-specific but broad immunomodulatory effect by reducing pro-inflammatory and increasing anti-inflammatory cytokines (61,62). The study was based on a device where a blood sample is exposed to an *ex vivo* oxygen/ozone gas mixture at a temperature of 42.5°C for about 20 min, after which the treated blood sample is administered by intragluteal injection in an attempt to evoke beneficial immune responses (61). The trial did not find any significant reduction in mortality or cardiovascular hospitalization (61).

The underlying biological effects are also unclear for treatment based on the administration of intravenous immunoglobulins (IVIg). Most of the studies to date have been conducted in patients with myocarditis and dilated non-ischaemic cardiomyopathy (63-65). A beneficial effect of IVIg in HF was first reported in an uncontrolled study of 10 patients with myocarditis and acute cardiomyopathy (65). Next, a randomized trial was performed in 62

patients with recent onset cardiomyopathy or myocarditis, but this did not confirm a beneficial effect of IVIg compared with placebo: both IVIg and placebo treatments improved cardiac function to a similar extent. (66) Further clinical studies are necessary to define the possible utility of this therapy; however given the neutral results of this and many other trials investigating anti-inflammatory therapies, it may be some time before IVIg for heart failure is tested again.

My interest lies in the remodeling of the LV after injury to the myocardium. LV remodeling after myocardial injury manifests as ventricular dilation, fibrosis and decreased cardiac performance (67-71). Given the extensive involvement of matrix metalloproteinase (MMP) system in remodeling of the failing heart, anti-MMP therapy was considered. The first clinical trial in patients with acute MI, the PREMIER trial (71) failed to show a treatment benefit when using a synthetic MMP inhibitor. However, the MMP-inhibitor used targeted different MMPs, which may have resulted in unwanted side effects and off-target effects. Cardiac imaging or follow-up of plasma levels of MMP-inhibition was lacking, and therefore the therapeutic effects of MMP-inhibition remained unclear. There has been discussion of future clinical trials using specific MMP-blocking agents to target single MMPs during a limited time period, by preference in patients with acute HF where increased MMP-expression relates to acute dilatation and failure (72).

Chronic inflammation is a major player in the pathophysiology of heart failure. Acute inflammation also plays a role immediately after myocardial injury and can directly influence the size and shape of the LV. Although inflammatory pathways are seemingly good targets for HF therapy, the initial trials of strategies to target these processes have had limited success. The most successful trials have been those that were small and tested a very carefully selected group

of patients with specific types of HF, notably autoimmune related cardiomyopathies. Novel treatment options targeting inflammation are needed to combat morbidity and mortality in HF.

**Understanding HF through Genome-Wide Association Studies (GWAS)**

HF - and the pathologies associated with neighboring organ systems – is a heterogeneous and complex trait. There are likely numerous component genes that each contributes a small genetic effect that, along with environmental factors, cumulatively affect disease incidence (73,74). There is little doubt that HF is heritable and accordingly genetic studies are being utilized to illuminate the underlying mechanisms.

The completion of the Human Genome Project in 2003 (75-77) led to substantial advances in large-scale genotyping and computational technologies. It is now possible to systematically search the human genome for common variants that are associated with a particular phenotype. Although ultimately advancing technology will allow the examination of the whole genome to identify genetic variants, but currently the principles of linkage disequilibrium (LD) are utilized to identify a set of common variants in the human genome that are excellent statistical proxies for genetic variation at a particular frequency (78). Each variant can be tested for its individual effect using a simple case-control design of GWAS using thousands of samples. Although these associations do not infer causality, they provide to specific genetic hypotheses that can be tested. The unbiased GWAS design is well suited to detect the effects of common genetic variations in complex traits; yet a number of human HF GWAS performed to date have had limited success, likely due to the significant noise from environmental heterogeneity.

The Lusis Lab performs large GWAS studies in mouse populations. In the early 2000s, the Lusis Lab developed a powerful systems biology resource to discover genes and pathways

contributing to complex pathological features of HF termed the Hybrid Mouse Diversity Panel (HMDP). The HMDP is a panel of inbred mouse strains developed for performing association studies with adequate statistical power and resolution for mapping of complex traits (79).

Adverse ventricular remodeling as a result of uncontrolled inflammation leads to a dilated LV manifesting with increased left ventricular internal dimension in diastole (LVIDd). Using a genome-wide association approach in the HMDP, we identified a genomic locus on chromosome 16 that was associated with isoproterenol (ISO) -induced LV dilation. Based on bioinformatic analysis, we have prioritized *Cd200* as the top candidate in this region **(Figure 5).**

The Lusis Lab has previously characterized the changes in cardiac structure and function in response to chronic beta-adrenergic agonist ISO infusion using echocardiography in a mapping panel of 105 inbred mouse strains (80). We showed that cardiac structure and function, whether under normal or stress conditions, have a strong genetic component, with heritability estimates of 64% to 84%. Association analyses of cardiac remodeling traits, corrected for population structure and multiple comparisons, revealed 3 genome-wide significant loci and 13 suggestive loci, including several loci containing previously implicated genes. Cardiac tissue gene expression profiling, expression quantitative trait loci, expression-phenotype correlation, and coding sequence variation analyses were performed to prioritize candidate gene lists and to generate hypotheses for downstream mechanistic studies. This work has revealed over 16 loci contributing to cardiac remodeling phenotypes in HMDP due to beta-adrenergic overdrive, with promise to greatly expand our understanding of how common variations contribute to age-related pathological cardiac remodeling (80).

One of the genetic loci on chromosome 16, associated with adverse cardiac remodeling

characterized by left ventricular dilation at week 3 following ISO infusion, pointed to *Cd200* as a causal candidate gene (p-value $4.97x10^{-7}$). The region contained a total of 17 genes in linkage disequilibrium with the peak SNP. Out of these 17 genes, Cd200 expression in both control and ISO treated hearts showed a strong correlation with left ventricular dilation by week 3 of ISO treatment **(Figure 5).** In addition, mouse strains with the CC genotype at single nucleotide polymorphism (SNP) rs4181997 had a significantly higher Cd200 expression level in left ventricular tissues and less ventricular dilation. Mouse strains with the AA genotype at the same SNP had a lower Cd200 expression level and more ventricular dilation (increased LVIDd). Taken together the data strongly suggested that genetic variations in this region lead to altered Cd200 expression, which contributed to the variations observed in left ventricular dilation in the mice. As such, Cd200 was chosen as the candidate gene for this locus. We hypothesized that Cd200 expression in the heart plays a protective role in ISO-induced cardiac injury and remodeling.

**CD200 and its receptor CD200R1**

CD200, also known as OX-2, is a type-1 transmembrane glycoprotein expressed by a broad range of cell types. Its cognate receptor CD200R1 is a myeloid or lymphoid cell specific surface glycoprotein that, when activated, attenuates inflammatory cytokine production following lymphocyte stimulation. The first evidence that CD200 might have immunoregulatory function came from the Gorzynski Lab in Toronto, ON, Canada from studies in a murine renal allograft model (81). They showed that both mice (skin and renal allografts) and rats (intestinal allografts) with prolonged acceptance of vascular and non-vascular grafts had elevated levels of expression of CD200 and that anti-CD200 antibodies diminished prolonged graft survival in

these animals (81, 82). Increased survival was reversed following neutralization of CD200 by monoclonal antibodies and restored by adding CD200:Fc fusion protein (81). In addition, splenic cells taken from CD200:Fc treated mice showed cytokine production favoring type 2 cytokines (82-85). Type 2 cytokine production promotes antihelminth immunity, suppresses type 1-driven autoimmune disease, neutralizes toxins, maintains metabolic homeostasis, and regulates wound repair and tissue regeneration pathways following infection or injury. Lastly, they found that $F4/80^+$ $CD200R^+$ splenic cells that were mixed with CD200:Fc provided significantly more potent immunosuppressive activity via inhibition of alloreactivity in vitro or in vivo than that seen using either CD200:Fc or $CD200R^+$ cells alone (86).

The CD200:CD200R interaction is also important in autoimmune disease. A murine model of human rheumatoid arthritis (RA), collagen-induced arthritis, is suppressed by CD200:Fc treatment in DBA/1 mice (87), with these mice having decreased anti-collagen antibodies and decreased inflammatory type-1 cytokine production. Multiple sclerosis (MS) is another disease process in which CD200 may plan an important role. Experimental autoimmune encephalomyelitis (EAE) is induced by immunization with brain-derived peptides and it is an experimental model sharing many clinical and pathological features with human MS. The onset of EAE occurred 3 days earlier in $CD200^{-/-}$ mice than in normal controls, with concomitant increase in $CD200R^+$ cells and enhancement in their activity (88). EAE was attenuated by CD200:Fc (89). In a separate murine model for uveoretinitis, a $CD4^+$T cell mediated autoimmune disease, the lack of CD200 accelerated onset of this disease in mice (90).

As evidenced by some of the data presented above, dysregulation of CD200-CD200R1 interaction has been implicated in a number of age-related inflammatory conditions, including

cancer growth, autoimmune and allergic disorders such as rheumatoid arthritis, Parkinson's disease, Alzheimer's disease, infection, graft versus host disease in transplantation, bone development, and reproduction. However, the role of CD200-CD200R1 interaction in cardiac injury and remodeling has not been previously studied. Further understanding of the CD200-CD200R1 interaction in cardiac injury and remodeling has the promise to shed light on novel therapeutic options in a complex disease process like HF.

**Exploring the Role of the CD200:CD200R1 Interaction in HF**

Several years ago, we hypothesized that the cardiomyocyte CD200 and macrophage CD200R1 interaction plays an important role in appropriately suppressing the injury inflammatory response which in turn leads to protection against adverse LV remodeling and subsequent heart failure. Validation and characterization of a therapeutic role for CD200 and CD200R1 in heart failure and identification of downstream regulatory targets of this signaling pathway could lead to a critical discovery with great therapeutic promise.

**CD200 and CD200R1 Expression in the Heart**

First, we aimed to characterize CD200 and CD200R1 expression in the heart, which no group has previously done. C57BL/6J and DBA/2J mice, representing genotypes CC versus AA at peak single nucleotide polymorphism (SNP) rs4181997, are characterized by differential *Cd200* mRNA expression and LV dilation response to ISO. **(Figure 6A and 6B)** Our preliminary results also suggest that there is a trend towards *Cd200* mRNA downregulation at week 3 of ISO treatment (data not shown). It is not known where CD200 is expressed among cardiac cell populations and whether differences in *Cd200* mRNA expression translate to

differences in cell surface protein expression during injury.

We treated 8 to 10-week old female C57BL/6J and DBA/2J strains with the previously established ISO-infusion model (91). We obtained an echocardiogram at baseline (homeostasis), and day 21 of ISO infusion to assess changes in global cardiac function and cardiac remodeling. At each time point, mice were sacrificed and we examined the levels of CD200 and CD200R1 expression using Western Blot analysis.

I first confirmed that CD200 and CD200R1 are both expressed in the mouse heart (Figure 8A). In C57BL/6 whole heart lysate, both CD200 and CD200R1 were expressed and CD200 had about 5-fold higher expression than CD200R1.

**CD200 and CD200R1 Expression Among Immune Cell Populations in the Heart**

Macrophages play essential roles in the maintenance of tissue homeostasis and response to injury, depending on the micro-environment in which they are activated. When classically activated by pro-inflammatory cytokines, macrophages of the M1 phenotype effectively destroy infectious microorganisms. When alternatively activated by anti-inflammatory cytokines, macrophages of the M2 phenotype drive tissue repair, regeneration and fibrosis (34, 43, 44, 88). The heart contains several discrete macrophage populations with mixed ontological origins that constitute the primary immune cells that reside in the heart under resting conditions. Maintained by self-renewal, rather than through blood monocyte input, these embryonically-derived resident cardiac macrophages promote tissue repair and regeneration by up regulating a phagocytic receptor to internalize debris and downregulating pro-inflammatory cytokines, resembling alternatively activated M2 macrophages (43,44). During cardiac stress, monocyte-derived

macrophages express high levels of pro-inflammatory genes and promote adverse cardiac remodeling.

**CD200R1 expression on macrophages is believed** to be critical for immune homeostasis, limiting inflammatory amplitude and duration. CD200R1, which is expressed at high levels in alveolar and intestinal mucosal macrophages, has also been shown to be critical for immune homeostasis in the resting state as well as crucial in limiting inflammatory amplitude and duration in the lung (92). The relative expression of CD200R1 among cardiac resident tissue macrophages and cardiac myeloid populations in homeostasis is not known. Based on previous findings, we hypothesized that expression of CD200R1 marks embryonically-derived resident macrophages. We aimed to characterize the changes in CD200R1 expression within the resident and recruited immune cell population using flow cytometry on cells isolated from myocardium in homeostasis and following injury. To identify the cells that express CD2001 within the inflammatory cells of the myocardium, we used flow cytometry. We performed whole heart tissue enzymatic digestion using both a benchtop mechanical dissociator called gentleMACS by Miltenyi Biotec and a Langendorff preparation to isolate cardiac single-cell suspension from C57BL/6J and DBA/2J hearts at baseline and day 21 of ISO treatment. I then stained the cells using antibodies against CD200R1, as well as Cd11c, Cd11b, Ly6c, Ccr2 and Cd206 to identify populations of dendritic cells, monocyte-derived macrophages, recruited monocytes, peripheral monocytes and embryonically derived macrophages, respectively. Cd11c was not used in a few experiments for ease of flow cytometry detector use. This experiment was repeated several times and the findings were similar on each occasion. All detected cells were positive for CD200 and CD200R1; however none of the other stains were positive. According to this data, there were no

monocyte-derived cells in the myocardium. This would make sense in the resting state of the myocardium; however these results held true for digested tissue obtained from ISO treated hearts as well. After repeating this experiment several times in both control and ISO treated hearts, I concluded that it is possible that the digestion process mitigates or does not preserve monocyte-derived cells. I eliminated the possibility that the stains were not working by using a peritoneal macrophage control.

**Establishing a Role for CD200:CD200R1 Interaction in Myocardial Injury**

To demonstrate that *Cd200* expression is causally related to changes in LVIDd following ISO treatment, we obtained existing *Cd200* knockout (KO) and doxycycline (DOX)-inducible transgenic (Tg) mice from the Gorczynski lab (93). *Cd200* KO mice were found to have a normal appearance, fertility, and lifespan; however, their macrophage lineage cells exhibited an activated phenotype and were numerous (94). *Cd200* KO mice rejected organ allografts more vigorously than littermate controls and resisted metastatic growth of breast cancer cells (95). On the other hand, *Cd200* transgenic mice showed no obvious phenotype at baseline yet had less skin, cardiac, and renal allograft rejection and more metastatic growth of breast cancer cells (96). Unfortunately, the wait for these mice to arrive was long. We received the mice in January 2016 – a year and a half after first requesting them. After their arrival, the mice were genotyped to confirm their genotypes. The CD200 KO mouse was a heterozygous KO for CD200. The CD200R1 KO mouse was a heterozygous KO for CD200R1 and in addition, contained a CD200 transgene with a Tet activator in place. The CD200 transgenic mouse contained the transgene; however did not contain the Tet activator. Given the limited amount of time I had to complete the validation studies, I started to breed the CD200KO$^{-/+}$ mice to obtain a sizable cohort for my

first set of validation experiments. In late April 2016, I had a cohort of 10 CD200KO$^{-/+}$ mice that were around 8 weeks of age and ready for ISO treatment. These 10 mice, along with 10 age-matched C57BL6 control mice were treated with ISO at a concentration of 30ug/kg of body weight for 21 days (91). I then compared LV diameter by cardiac echocardiogram at baseline and day 21 to establish the phenotypic changes in LV dilation with normal and depleted *Cd200* expression. I also performed Mason's trichrome staining of cardiac tissue to measure the amount and distribution of fibrosis both at homeostasis and after ISO-induced HF in both CD200 KO$^{-/+}$ and control mice.

The first cohort of mice delivered promising results. At the end of the 21 days, we were left with 8 wild type mice and 9 CD200 KO$^{-/+}$ mice **(Figure 7).** 7 out of the 9 CD200 KO$^{-/+}$ mice had LV dilation, 6 of which dilated more than 0.5mm. The largest dilation was a change of 1.5mm from baseline. In contrast, only one mouse had a dilated LV with an increase in LV cavity size of around .25mm. The other 7 mice in this group had smaller LV cavity size, the smallest LV cavity was almost 1.5mm smaller than baseline. These results were significant by simple student's t-test ($p = 0.001$).

With these encouraging results, I decided to try a larger cohort to confirm these results. Unfortunately, the second cohort delivered mixed results with no definitive pattern of LV dilation in the CD200 KO$^{-/+}$ mice **(Figure 8).** There were initially 17 mice in each group; however only 14 mice survived in each group. The deaths seen prior to the end of the experiment were likely a result of post-surgical complications after placement of the ISO pump. In this cohort, only half or 7 out of 14 mice in the CD200 KO$^{-/+}$ group experienced dilation of the LV after ISO treatment. The largest LV dilation was only 1.3mm, in contrast to to 1.5mm in the first

cohort of mice. 4 out of the 14 mice in the control group dilated, the largest dilation of the LV being about 1.3mm. Unfortunately, these results were not significant by student's t-test ($p = .136$).

Given the lack of significance in the last cohort of mice tested for change in LV dilation after ISO treatment, we decided not to pursue further studies in CD200 validation at this time. However, we are currently breeding full CD200 KO mice and we may pursue similar validation studies in full KO mice.

**Drawbacks to the ISO induced HF Model**

Excessive doses of catecholamines produce diffuse myocardial tissue damage with cardiomyocyte necrosis and extensive fibrosis in both animals and humans [97-99]. The mechanism underlying myocardial damage is likely related to an imbalance between oxygen supply versus demand due to myocardial hyperactivity [100]. In mice, infusion of isoproterenol for 7 days has been shown to induce cardiac dysfunction [101]. In rats, subcutaneous administration of isoproterenol for 3 days leads to a dose-dependent impairment of cardiac function and neurohormornal activation [102,103], with cardiomyocyte necrosis and extensive LV hypertrophy and dilation and after 2 and 12 weeks, respectively [104, 105].

Although almost half of the mice tested showed some LV dilation after 21-day ISO treatment, none of the mice actually had appreciable cardiac dysfunction (by measurement of EF on echocardiogram). This simple fact makes it difficult to use the ISO model as a true model of HF.

**Soluble CD200 in Humans**

Soluble CD200 (sCD200) has been studied in human disease, mostly by the Gorzynski lab. It is not known whether sCD200 level differs from controls in chronic inflammatory conditions such as HF, coronary artery disease (CAD) and diabetes mellitus (DM) or correlates with disease severity or age. It is not known whether sCD200 can be used as a biomarker to predict HF severity or as a therapeutic agent for immune modulation in HF, CAD, and DM. Given our findings from the large HMDP HF GWAS study and from what is already known in the literature, we hypothesized that sCD200 in human serum increase with age, as inflammation generally increases with age. We also hypothesized that higher levels of sCD200 in human serum protect against incidence and progression of CAD.

Cd200:Cd200r1 interaction appears to be a negative regulator of inflammation and has thus far been demonstrated to be important in cancer growth, autoimmune and allergic disorders, infection, transplantation, bone development, and reproduction (106). In addition, sCd200 was shown to induce type 1 regulatory T cells (Tr1) to promote cardiac and skin allograft survival in mice and sCd200 at early times after engraftment may predict allograft survival (107). In a preliminary study from the University Health Network (UHN) in Toronto, cardiac transplant patients who had higher sCD200 levels in the immediate post-transplant period had fewer rejection episodes compared to those with lower sCD200 levels (personal communication of unpublished data). Arik et al. demonstrated recently that soluble CD200 (sCD200) level by enzyme-linked immunosorbent assay (ELISA) was elevated in diabetics versus healthy controls and that sCD200 was negatively correlated with diabetic foot ulcer severity (108).

Although CD200 has been extensively studied in a variety of inflammatory conditions, the role of CD200:CD200R1 interaction in cardiac injury and remodeling has not been

previously explored. Coronary artery disease, diabetes, and aging are chronic inflammatory states and are significant risk factors for cardiac injury and adverse cardiac remodeling. We proposed to ascertain the relationship between sCD200 and the presence and severity of CAD as we hypothesized that sCD200 plays a protective role in CAD.

We obtained human serum samples from the Indiana CTSI Fairbanks Institute. The inclusion criteria for the samples were individuals over the age of 65 who had either CAD (defined by coronary anatomy by coronary angiogram). The samples included 112 subjects with CAD and 50 age and gender matched controls. Our sample size is comparable to previous human studies of sCD200 levels demonstrating positive association with asthma and diabetic foot ulcer (108, 109). We had demographic information, coronary angiography data including number of diseased vessels and intervention (angioplasty vs. percutaenous coronary intervention vs. coronary artery bypass graft), Hgb A1C, brain natriuretic peptide, and echocardiogram parameters, with the possibility of obtaining further detailed data utilizing a data manager service at the institute. We measured both sCD200 and sCD200R1 using ELISA Kits for humans in the CAD samples and controls (162 total) in consultation with the Inflammatory Biology Core (IBC). We performed statistical analyses to test whether sCD200 and sCD200R1 levels in human serum are associated with age and predictive of severity of CAD. Our sample size was 162, which included CAD patients and age and gender matched controls. With a sample size of 162, the study had 80% power (at 2-sided alpha of 0.05) to detect unadjusted effects as small as 0.22 standard deviation (SD) decrement in CAD severity per SD increment in CD200 or CD200R1 serum level. In multiple linear regression with 5 covariates that explain 50% of the variance in the serum CD200 level, we had 80% power to detect adjusted (standardized) effects as small as

0.32. A similar study measured soluble CD200 in human serum of lupus patients. The correlation coefficient between CD200 level and lupus severity was 0.48 (110). If the effect of CD200 on CAD severity is as large as the effect on lupus severity, we will have had more than enough power to detect it in this study. In addition, we requested supplemental de-identified data to be extracted from the Indiana CTSI Fairbanks Institute to further elucidate echocardiographic traits and structural heart disease in these patients.

Unfortunately, when looking at the concentrations of CD200 and CD200R1 in CAD cases versus controls, there was no appreciable or significant difference between these two groups in the initial analysis. We felt that the lack of differences seen was at least in part due to the ELISA assay for CD200. This particular kit gave inconsistent results and the majority of the samples appeared to have a CD200 concentration below the lower limit (< 25 pg/mL). The samples that were detectable vary tremendously in their apparent CD200 concentration. For example, there was one sample that was diluted 20-fold and still appeared to be above the highest standard of 2400 pg/ml, resulting in a reported apparent value of ">48,000 pg/mL." Given the number of samples with undetectable levels of sCD200, we used sCD200 as a dichotomous variable of detectable or undetectable. We also characterized the data as (i) undetectable (<25 pg/mL), (ii) 25-2400 pg/mL, and (iii) off-scale high (>2400 pg/mL), and compared the distributions of these three variables between cases and controls, but again, there was no significant difference in sCD200 concentrations in the CAD cases versus control group. Although there were more samples with detectable levels of CD200R1; unfortunately, there was no significant difference in the concentrations of sCD200R1 between the CAD cases and controls. Due to the cost of large-scale ELISA assays, we did not pursue this further.

**Figure 1**

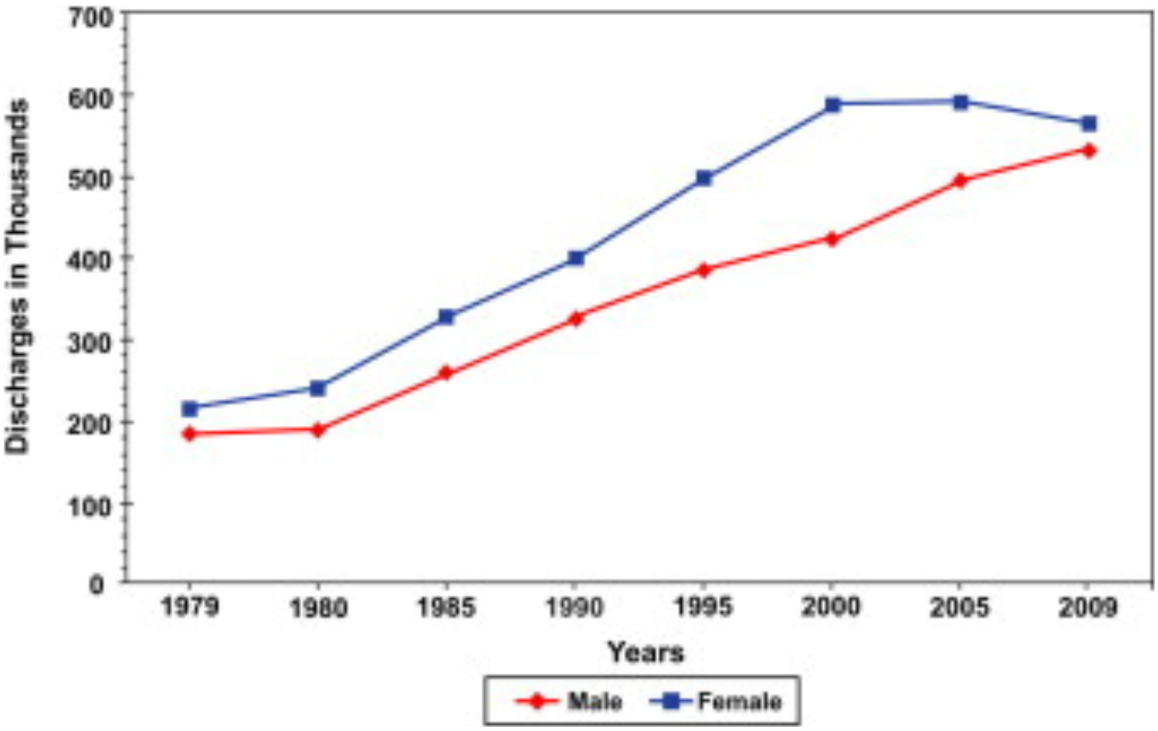



**Figure 1:** Discharges From Hospitalization Due to Heart Failure, by Sex (United States, 1979–2009)

*Used with permission from JACC Heart Failure. CHF. 2013;1(1):1-20. doi:10.1016/j.jchf.2012.10.002*

**Figure 2**

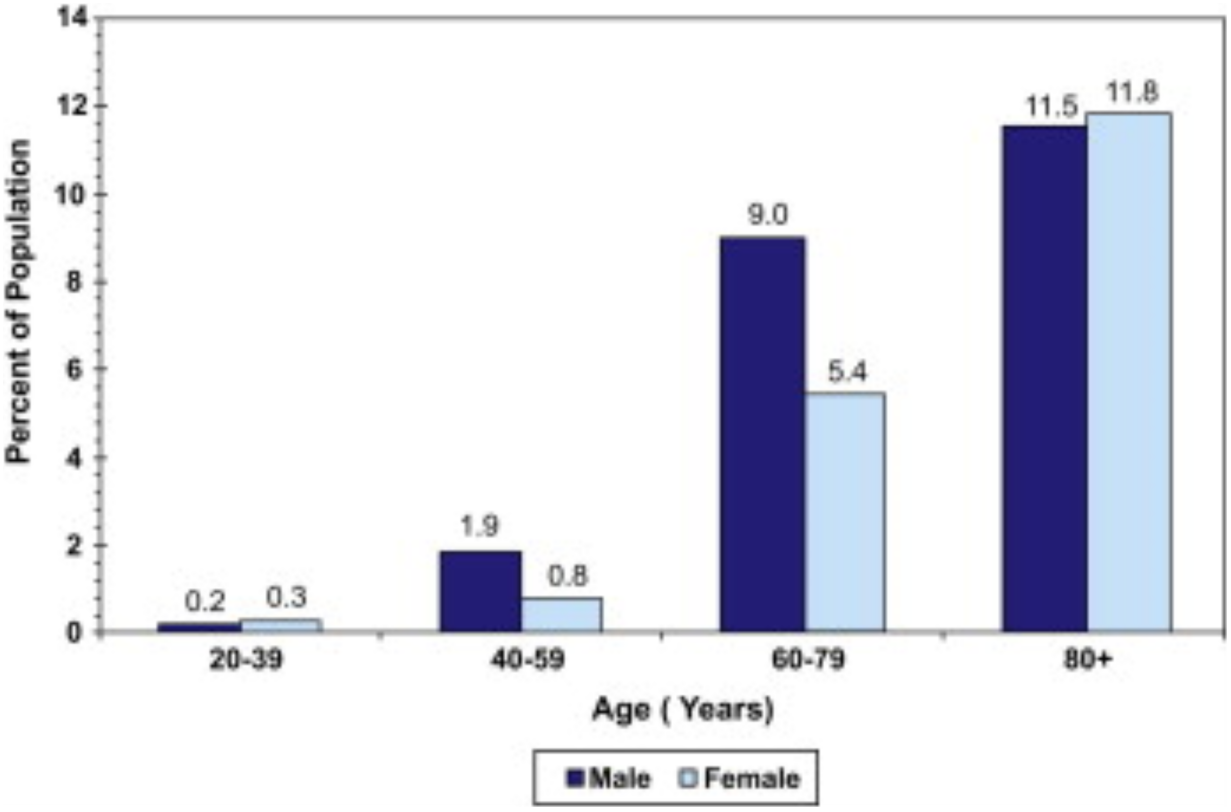



**Figure 2:** Prevalence of Heart Failure, by Sex and Age (National Health and Nutrition Examination Survey, 2005–2008)

*Used with permission from JACC Heart Failure. CHF. 2013;1(1):1-20. doi:10.1016/j.jchf.2012.10.002*

**Figure 3**

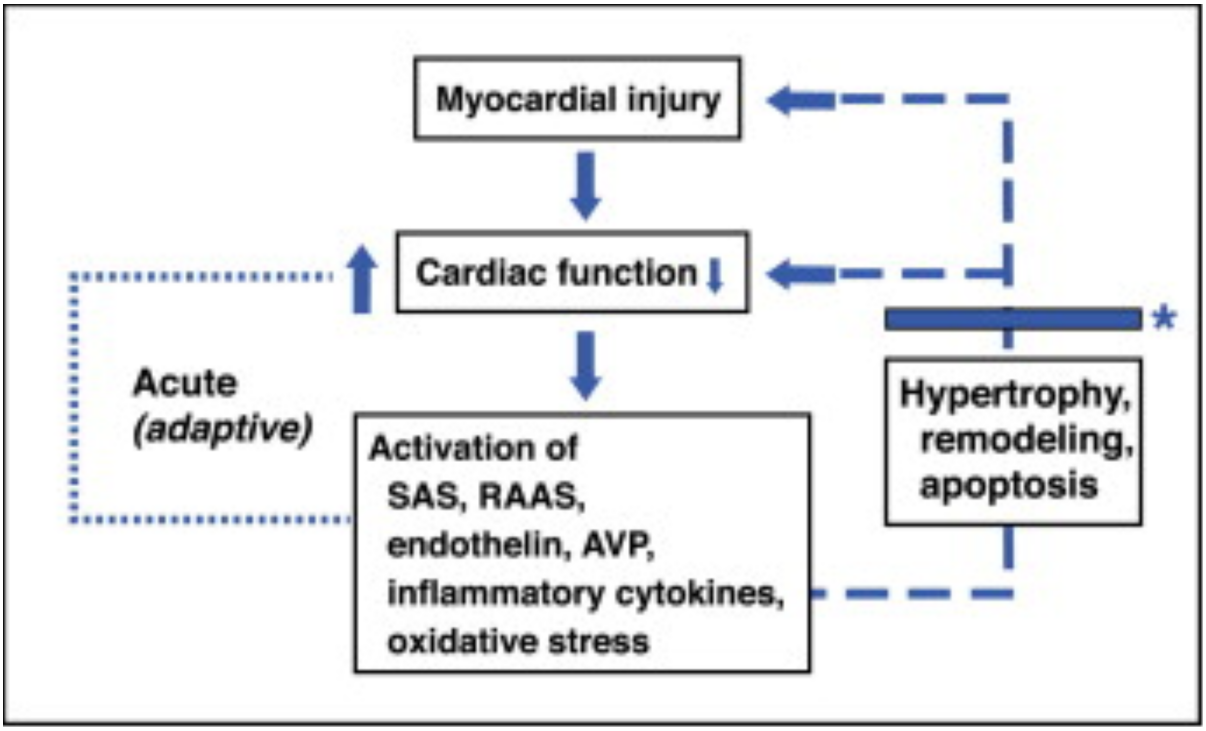



**Figure 3:** Interplay Between Cardiac Function and Neurohumoral and Cytokine Systems
Myocardial injury, which may have any of a number of causes, might depress cardiac function, which in turn may cause activation of the sympathoadrenal system and the renin-angiotensin-aldosterone system and the elaboration of endothelin, arginine vasopressin, and cytokines such as tumor necrosis factor (TNF)-α. In acute heart failure (left), these are adaptive and tend to maintain arterial pressure and cardiac function. In chronic heart failure (right), they cause maladaptive hypertrophic remodeling and apoptosis, which cause further myocardial injury and impairment of cardiac function. The horizontal line on the right shows that chronic maladaptive influences can be inhibited by angiotensin-converting enzyme inhibitors, β-adrenergic blockers, angiotensin II type 1-receptor blockers, and/or aldosterone antagonists.

*Used with permission from JACC Heart Failure. CHF. 2013;1(1):1-20. doi:10.1016/j.jchf.2012.10.002*

**Figure 4**

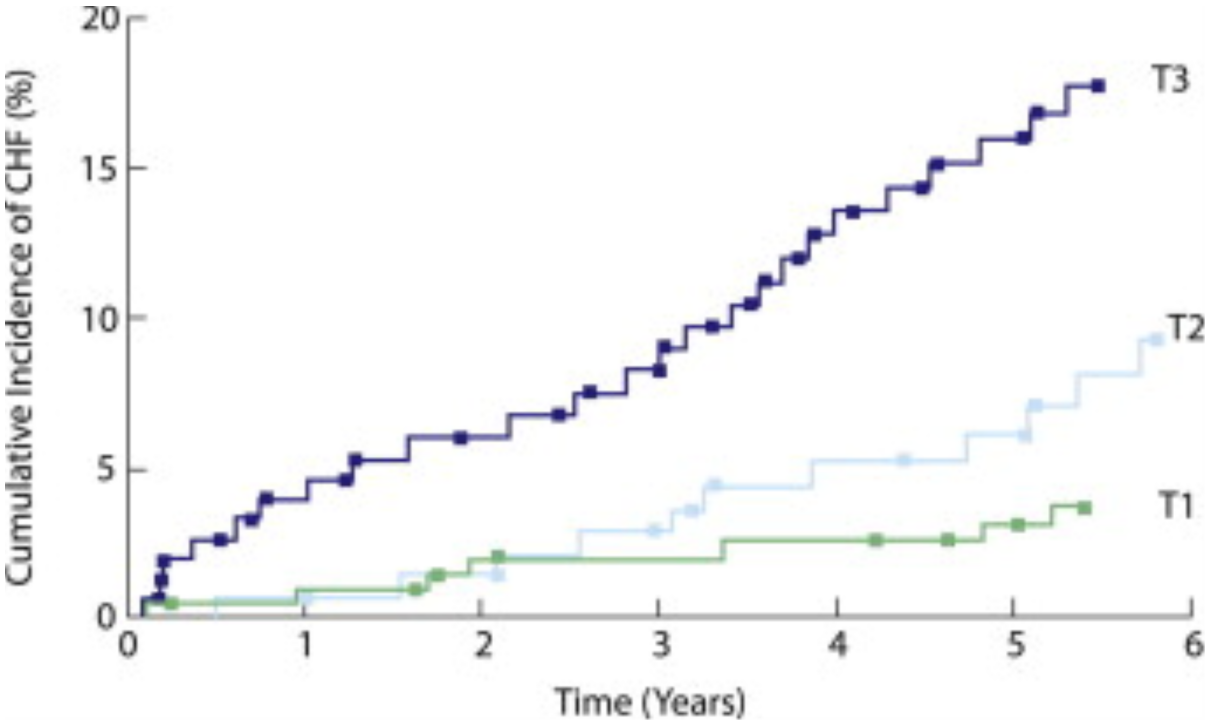
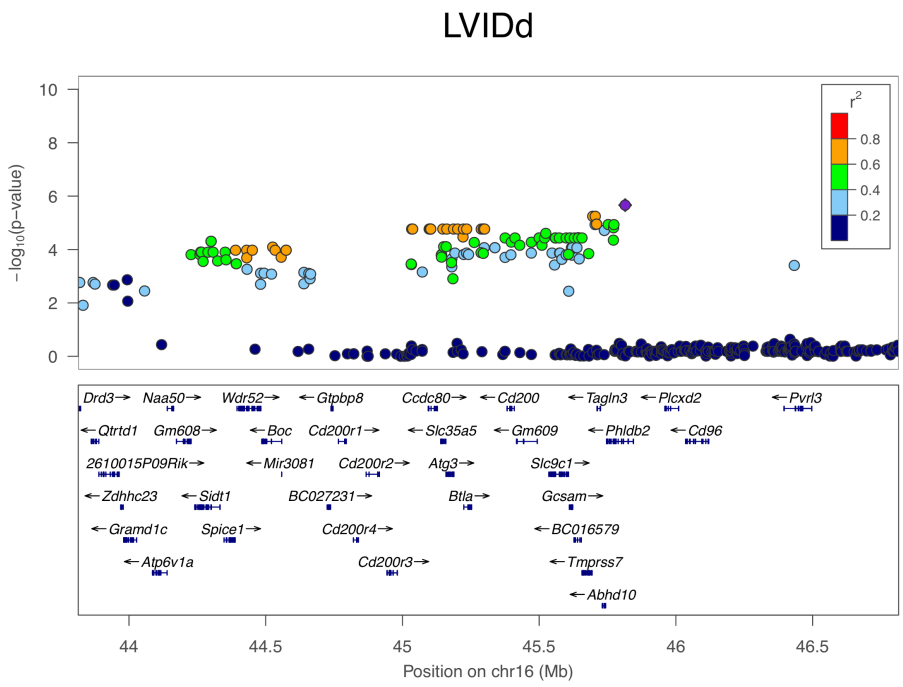
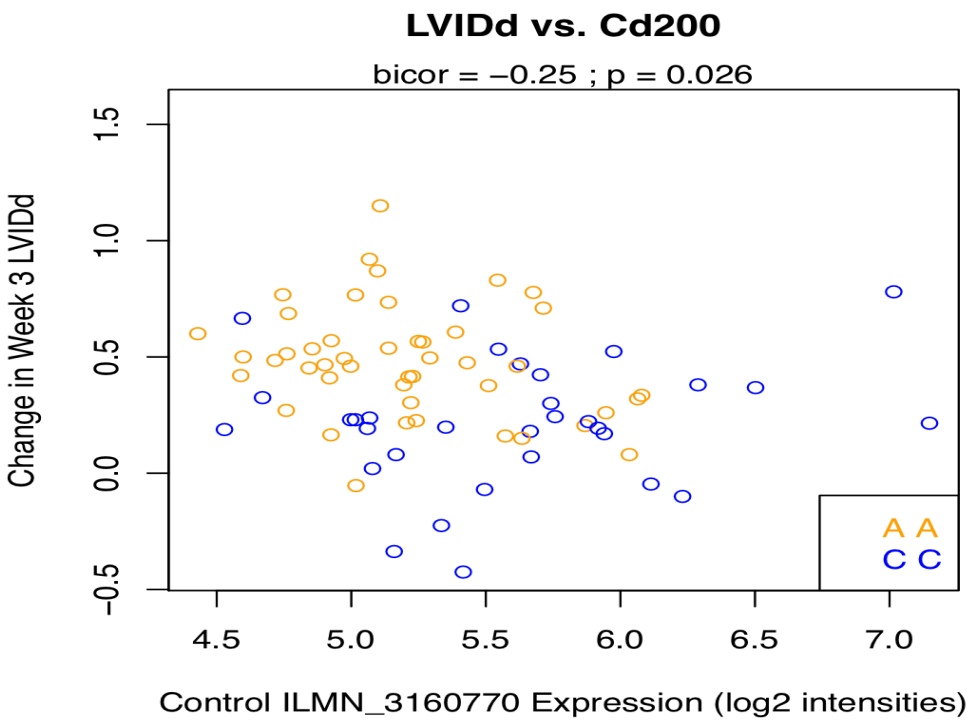



**Figure 4:** Circulating interleukin-6, an inflammatory cytokine, was prospectively related to heart failure incidence in a continuous, graded fashion among participants in the Framingham Heart Study. CHF = congestive heart failure. T1, T2, and T3 represent the lowest, middle, and highest tertiles of concentration.

Used with permission from JACC Heart Failure. CHF. 2013;1(1):1-20.
doi:10.1016/j.jchf.2012.10.002

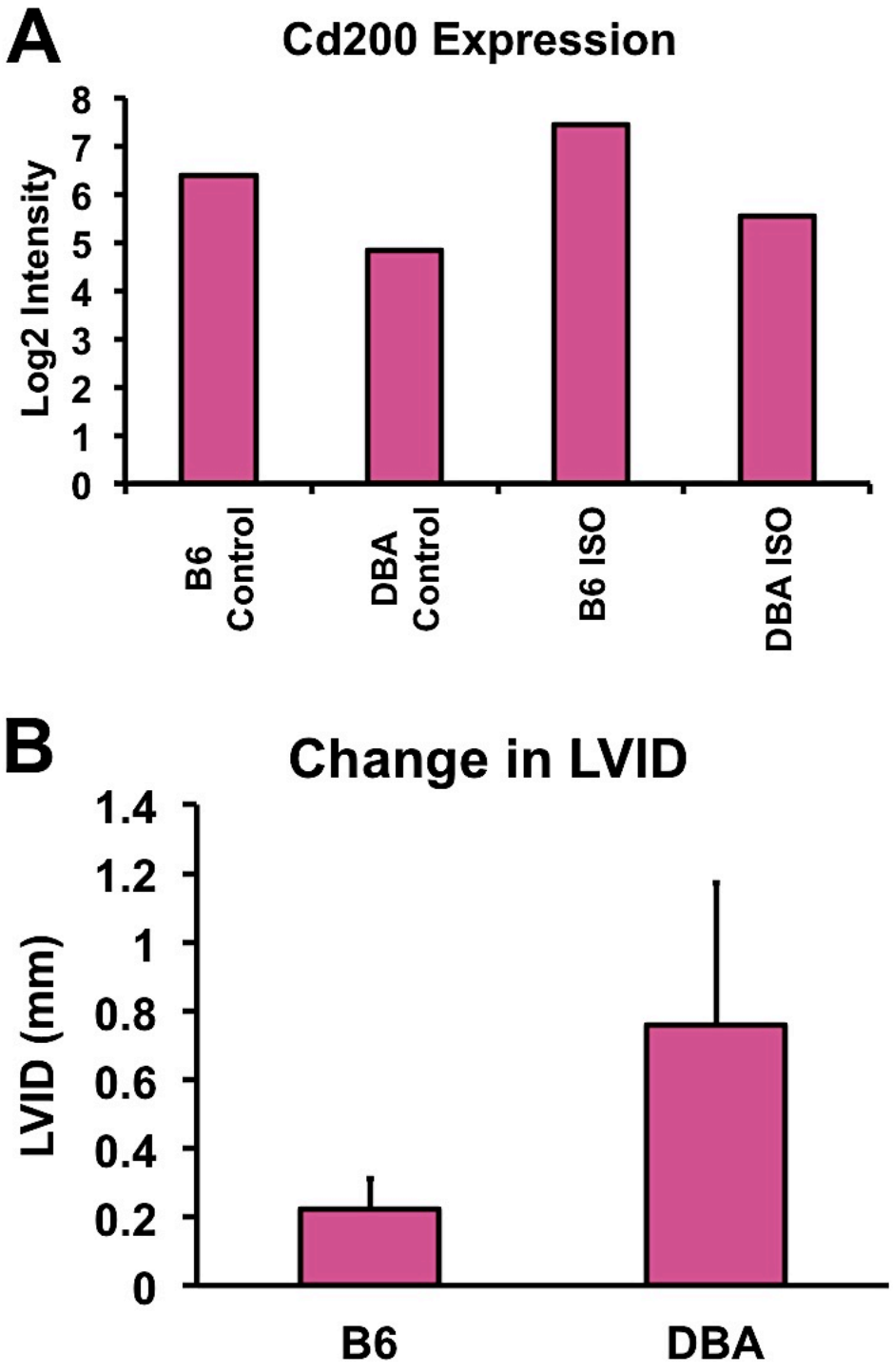
**Figure 5**

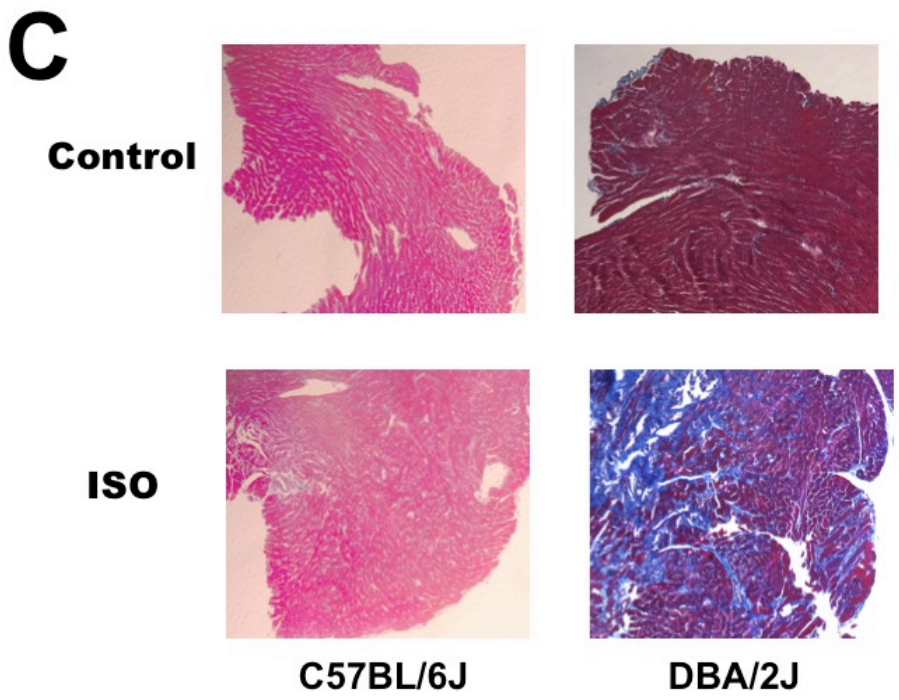

LVIDd
−log10(p−value)
10
8
6
4
2
0
r2
0.8
0.6
0.4
0.2
Drd3 Naa50 Wdr52 Gtpbp8 Ccdc80 Cd200 Tagln3 Plcxd2 Pvrl3
Qtrtd1 Gm608 Boc Cd200r1 Slc35a5 Gm609 Phldb2 Cd96
2610015P09Rik Mir3081 Cd200r2 Atg3 Slc9c1
Zdhhc23 Sidt1 BC027231 Btla Gcsam
Gramd1c Spice1 Cd200r4 BC016579
Atp6v1a Cd200r3 Tmprss7
Abhd10
44
44.5
45
45.5
46
46.5
Position on chr16 (Mb)




Chromosome 16 contains a significant association for LVIDd. **Top panel:** Manhattan plot of the Chromosome 16 locus for LVIDd. **Bottom panel:** Correlation plot of Cd200 expression with Change in LVIDd at week 3. AA and CC indicate the two possible genotypes at the Chromosome 16 locus.

**Figure 6**

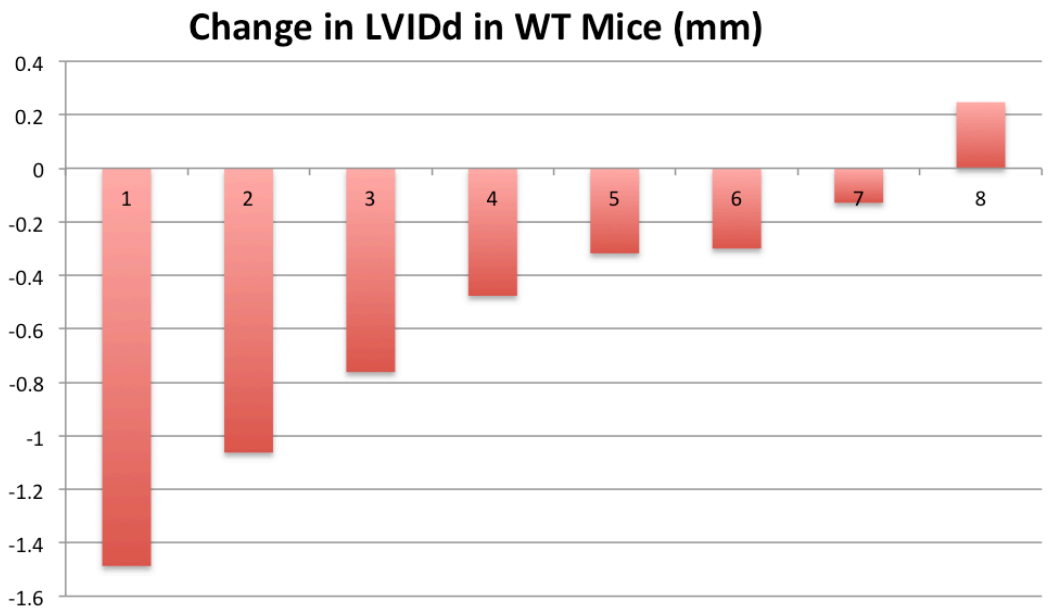

A
Cd200 Expression
Log2 Intensity
0
1
2
3
4
5
6
7
8
B6 Control
DBA Control
B6 ISO
DBA ISO
B
Change in LVID
LVID (mm)
0
0.2
0.4
0.6
0.8
1
1.2
1.4
B6
DBA



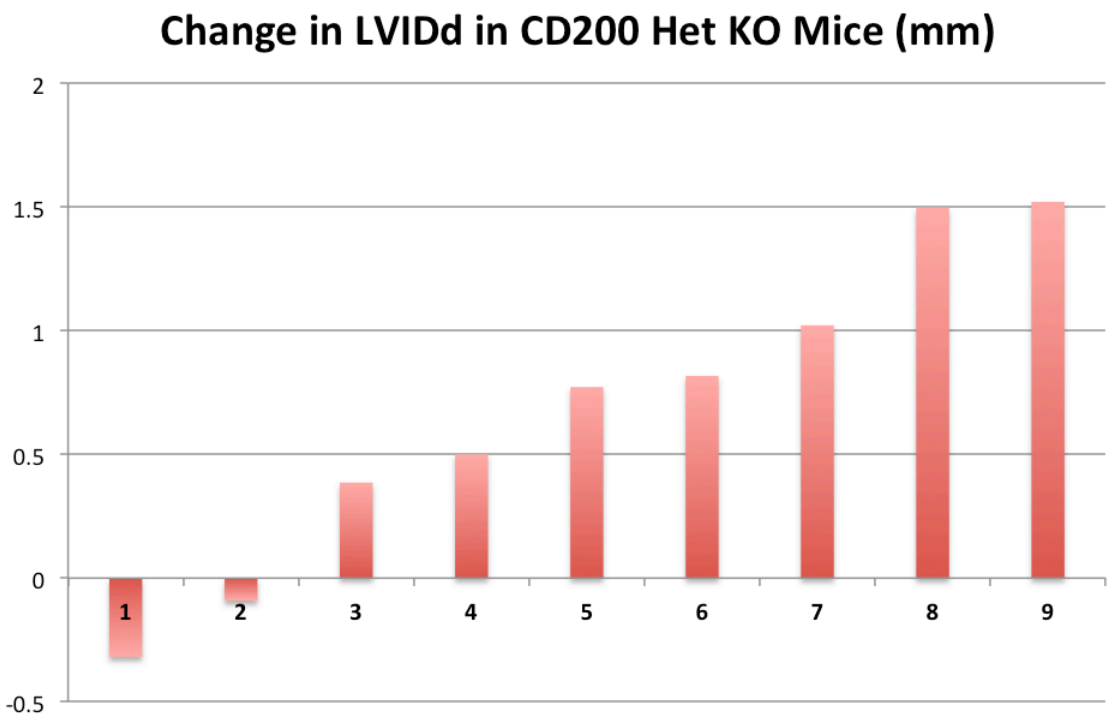

A: *Cd200* expression in C57BL/6J (B6) and DBA/2J (DBA) in control and ISO-treated mice. B: Change in left ventricular internal dimension (LVID) in diastole following three weeks of ISO treatment. C: Mason's trichrome staining in control and ISO- treated left ventricles.

**Figure 7**

**A.**


Change in LVIDd in WT Mice (mm)
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
-1.4
-1.6
1
2
3
4
5
6
7
8

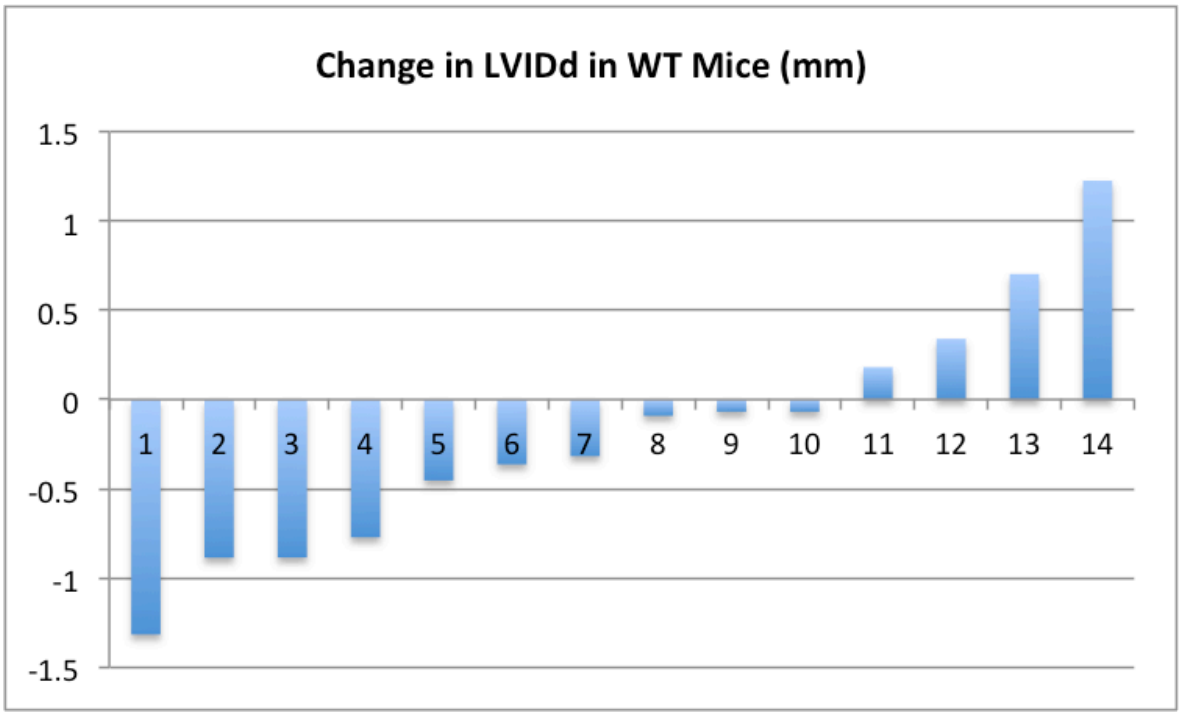
**B.**



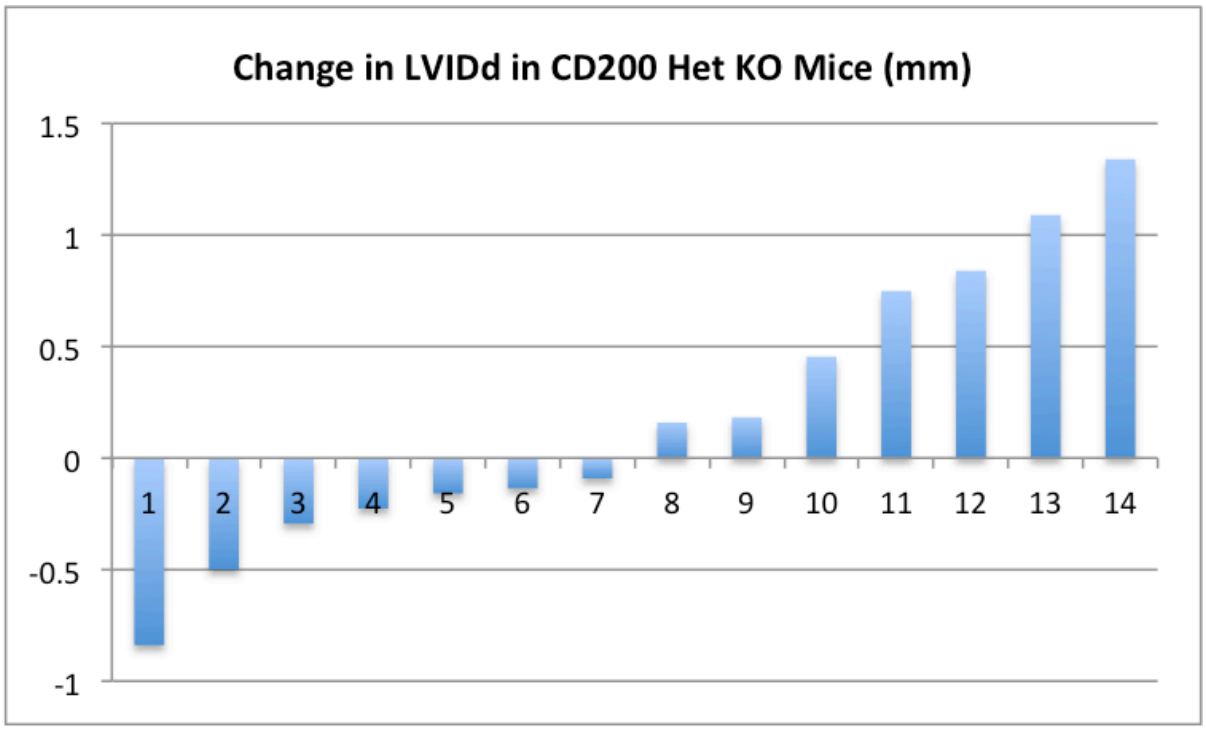

**A.** Change in LVIDd (mm) in WT mice from baseline to post ISO treatment. **B.** Change in LVIDd (mm) in CD200 KO mice from baseline to post ISO treatment.

**Figure 8**

**A.**


Change in LVIDd in WT Mice (mm)
1.5
1
0.5
0
-0.5
-1
-1.5
1
2
3
4
5
6
7
8
9
10
11
12
13
14

**B.**




**A.** Change in LVIDd (mm) in WT mice from baseline to post ISO treatment. **B.** Change in LVIDd (mm) in CD200 KO mice from baseline to post ISO treatment.

**References**


1. Nabel EG, Braunwald E. A tale of coronary artery disease and myocardial infarction. N Engl J Med 2012;366:54–63.

2. Roger VL, Go AS, Lloyd-Jones DM, et al. Heart disease and stroke statisticsd2012 update: a report from the American Heart Associa- tion. Circulation 2012;125:e2–220.

3. National Heart, Lung and Blood Institute. Morbidity and Mortality: 2012 Chart Book on Cardiovascular, Lung, and Blood Diseases. February 2012.

4. Chen J, Normand S-LT, Wang Y, Krumholz HM. National and regional trends in hear failure hospitalization and mortality rates for Medicare beneficiaries, 1998-2008. JAMA 2011;306:1669–78.

5. Levy D, Jenchaiah S, Larson MG, et al. Longterm trends in the incidence of and survival with heart failure. N Engl J Med 2002;347: 1397–402.

6. Laribi S, Aouba A, Nikolaou M, et al. Trends in death attributed to heart failure over the past two decades in Europe. Eur J Heart Fail 2012;14:234–9.

7. Roger VL. The heart failure epidemic. Int J Environ Res Public Health 2010;7:1807–30.

8. Askoxylakis V, Thieke C, Pleger ST, et al. Long-term survival of cancer patients compared to heart failure and stroke: a systematic review. BMC Cancer 2010;10:105.

9. Lloyd-Jones D, Adams RJ, Brown TM, et al. Heart disease and stroke statisticsd2010 Update. Circulation 2010;121:e46–215.

10. Dunlay SM, Shah ND, Shi Q, et al. Lifetime costs of medical care after heart failure diagnosis. Circ Cardiovasc Qual Outcomes 2011;4: 68–75.

11. Ross JS, Chen J, Lin Z, et al. Recent national trends in readmission rates after heart failure hospitalizations. Circ Heart Fail 2010;3: 97–103.

12. Braunwald E, Ross J Jr., Sonnenblick EH. Medical progress. Mechanisms of contraction of the normal and failing heart. N Engl J Med 1967;277:794–800, 853–63, 910–20, 962–71, 1012–22.

13. van Heerebeck L, Borbely A, Niessen HWM, et al. Myocardial structure and function differ in systolic and diastolic heart failure. Circulation 2006;113:1966–73.

14. Ohtani T, Mohammed SF, Yamamoto K, et al. Diastolic stiffness as assessed by diastolic wall strain is associated with adverse remodeling and poor outcomes in heart failure with preserved ejection fraction. Eur Heart J 2012;33:1742–9.

15. Borbely A, van der Velden J, Papp Z, et al. Cardiomyocyte stiffness in diastolic heart failure. Circulation 2005;111:774–81.

16. Braunwald E, Ross J Jr., Sonnenblick EH. Medical progress. Mechanisms of contraction of the normal and failing heart. N Engl J Med 1967;277:794–800, 853–63, 910–20, 962–71, 1012–22.

17. Braunwald E, Chidsey CA, Harrison DC, Gaffney TE, Kahler RL. Studies on the function of the adrenergic nerve endings in the heart. Circulation 1963;28:958–69.

18. Slaughter MS, Rogers JG, Milano CA, et al. Advanced heart failure treated with continuous-flow left ventricular assist device. N Engl J Med 2009;361:2241–51.

19. Ambardekar AV, Buttrick PM. Reverse remodeling with left ventricular assist devices. A

review of clinical, cellular, and molecular effects. Circ Heart Fail 2011;4:224–33.

20. Hall JL, Fermin DR, Birks EJ, et al. Clinical, molecular, and genomic changes in response to a left ventricular assist device. J Am Coll Cardiol 2011;57:641–52.
21. Zafeiridis A, Jeevanandam V, Houser SR, Margulies KB. Regression of cellular hypertrophy after left ventricular assist device support. Circulation 1998;98:656–62.
22. Park SJ, Milano CA, Tatooles A, et al. Outcomes in advanced heart failure patients with left ventricular assist devices for destination therapy. Circ Heart Fail 2012;5:241–8.
23. Eckman PM, John R. Bleeding and thrombosis in patients with continuous-flow ventricular assist devices. Circulation 2012;125: 3038–47.
24. Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure. ESC guidelines for the diagnosis and treatment of acute and chronic heart failure 2012. Eur Heart J 2012;33:1787–847.
25. Lindenfeld J, Albert NM, Boehmer JP, et al. HFSA 2010 comprehensive heart failure practice guideline. J Cardiac Fail 2010;16: e1–194.
26. Jessup M, Abraham WT, Caset DE, et al. 2009 focused update: ACCF/AHA guidelines for the diagnosis and management of heart failure in adults. Circulation 2009;119:1977–2016.
27. Lindenfeld J, Miller GG, Shakar SF, Zolty R, Lowes BD, Wolfel EE, Mestroni L, Page RL 2nd, Kobashigawa J. Drug therapy in the heart transplant recipient: part I: cardiac rejection and immunosuppressive drugs. *Circulation.* Dec 14 2004;110(24):3734-40.
28. Frangogiannis, NG. Regulation of the inflammatory response in cardiac repair. *Circ. Res.* 2012;110, 159–173.
29. Frangogiannis, NG. The inflammatory response in myocardial injury, repair, and

remodeling. *Nature Reviews Cardiology.* May 2014, 11, 255-265.

30. Yan, X. *et al.* Temporal dynamics of cardiac immune cell accumulation following acute myocardial infarction. *J. Mol. Cell Cardiol*. 2013;62, 24–35.
31. Epelman, S, Liu, P, Mann, DL. Role of innate and adaptive immune mechanisms in cardiac injury and repair. *Nature Reviews Immunology.* Feb 15 2015; 15: 117-129.
32. Nahrendorf, M. *et al.* The healing myocardium sequentially mobilizes two monocyte subsets with divergent and complementary functions. *J. Exp. Med.* 204, 3037–3047 (2007).
33. Frangogiannis, N. G. *et al.* Critical role of monocyte chemoattractant protein-1/CC chemokine ligand 2 in the pathogenesis of ischemic cardiomyopathy. *Circulation* 115, 584–592 (2007).
34. van Amerongen, M. J., Harmsen, M. C., van Rooijen, N., Petersen, A. H. & van Luyn, M. J. Macrophage depletion impairs wound healing and increases left ventricular remodeling after myocardial injury in mice. *Am. J. Pathol.* 170, 818–829 (2007).
35. Panizzi, P. *et al.* Impaired infarct healing in atherosclerotic mice with Ly-6C(hi) monocytosis. *J. Am. Coll. Cardiol.* 55, 1629–1638 (2010).
36. Dewald, O. *et al.* CCL2/monocyte chemoattractant protein-1 regulates inflammatory responses critical to healing myocardial infarcts. *Circ. Res.* 96, 881–889 (2005).
37. Leuschner, F. *et al.* Therapeutic siRNA silencing in inflammatory monocytes in mice. *Nature Biotech.* 29, 1005–1010 (2011).
38. Serbina, N. V. & Pamer, E. G. Monocyte emigration from bone marrow during bacterial infection requires signals mediated by chemokine receptor CCR2. *Nature Immunol.* 7, 311–317 (2006).

39. Zhou, L. *et al.* Monocyte chemoattractant protein-1 induces a novel transcription factor that causes cardiac myocyte apoptosis and ventricular dysfunction. *Circ. Res.* 98, 1177–1185 (2006)

40. Ingersoll, M. A. *et al*. Comparison of gene expression profiles between human and mouse monocyte subsets. *Blood* **115**, e10–e19 (2010).

41. Auffray, C. *et al*. Monitoring of blood vessels and tissues by a population of monocytes with patrolling behavior. *Science* **317**, 666–670 (2007).

42. Carlin, L. M. *et al*. Nr4a1-dependent Ly6C$^{low}$ monocytes monitor endothelial cells and orchestrate their disposal. *Cell* **153**, 362–375 (2013).

43. Epelman, S. *et al*. Embryonic and adult-derived resident cardiac macrophages are maintained through distinct mechanisms at steady state and during inflammation. *Immunity* **40**, 91–104 (2014).

44. Hashimoto, D. *et al*. Tissue-resident macrophages self-maintain locally throughout adult life with minimal contribution from circulating monocytes. *Immunity*. **38**, 792–804 (2013).

45. Jakubzick, C. *et al*. Minimal differentiation of classical monocytes as they survey steady-state tissues and transport antigen to lymph nodes. *Immunity*. **39**, 599–610 (2013).

46. Frangogiannis, N. G. *et al*. Critical role of monocyte chemoattractant protein-1/CC chemokine ligand 2 in the pathogenesis of ischemic cardiomyopathy. *Circulation* **115**, 584–592 (2007).

47. Hilgendorf, I. *et al*. Ly-6C$^{high}$ monocytes depend on Nr4a1 to balance both inflammatory and reparative phases in the infarcted myocardium. *Circ. Res.* **114**, 1611–1622 (2014).

48. Zouggari, Y. *et al*. B lymphocytes trigger monocyte mobilization and impair heart

function after acute myocardial infarction. *Nature Med.* **19**, 1273–1280 (2013).

49. Dewald, O. *et al*. CCL2/monocyte chemoattractant protein-1 regulates inflammatory responses critical to healing myocardial infarcts. *Circ. Res.* **96**, 881–889 (2005).
50. Leuschner, F. *et al*. Therapeutic siRNA silencing in inflammatory monocytes in mice. *Nature Biotech.* **29**, 1005–1010 (2011).
51. Serbina, N. V. & Pamer, E. G. Monocyte emigration from bone marrow during bacterial infection requires signals mediated by chemokine receptor CCR2. *Nature Immunol.* **7**, 311–317 (2006).
52. Epelman S, Liu P, Mann D. Role of innate and adaptive immune mechanisms in cardiac injury and repair. Nature Reviews Immunology. 2015: 15, 117-129.
53. Elster SK, Braunwald E, Wood HF. A study of C-reactive protein in the serum of patients with congestive heart failure. Am Heart J 1956; 51:533–41.
54. Vasan RS, Sullivan LM, Roubenoff R, et al. Inflammatory markers and risk of heart failure in elderly subjects without prior myocardial infarction. Circulation 2003;107:1486–91.
55. Rauchhaus M, Doehner W, Francis DP, Davos C, Kemp M, Liebenthal C, Niebauer J, Hooper J, Volk HD, Coats AJ, Anker SD. Plasma cytokine parameters and mortality in patients with chronic heart failure. Circulation. 2000;102:3060–3067.
56. Deswal A, Petersen NJ, Feldman AM, Young JB, White BG, Mann DL. Cytokines and cytokine receptors in advanced heart failure: an analysis of the cytokine database from the Vesnarinone trial (VEST) Circulation. 2001;103:2055–2059.
57. Gullestad L, Aukrust P. Review of trials in chronic heart failure showing broad-spectrum anti-inflammatory approaches. Am J Cardiol. 2005;95:17C–23C. discussion 38C–40C.

58. Mann DL, McMurray JJ, Packer M, Swedberg K, Borer JS, Colucci WS, Djian J, Drexler H, Feldman A, Kober L, Krum H, Liu P, Nieminen M, Tavazzi L, van Veldhuisen DJ, Waldenstrom A, Warren M, Westheim A, Zannad F, Fleming T. Targeted anticytokine therapy in patients with chronic heart failure: results of the Randomized Etanercept Worldwide Evaluation (RENEWAL) Circulation. 2004;109:1594–1602.

59. Mann DL. Inflammatory mediators and the failing heart: past, present, and the foreseeable future. Circ Res. 2002;91:988–998.

60. Anker SD, Coats AJ. How to RECOVER from RENAISSANCE? The significance of the results of RECOVER, RENAISSANCE, RENEWAL and ATTACH. Int J Cardiol. 2002;86:123–130.

61. Torre-Amione G, Anker SD, Bourge RC, Colucci WS, Greenberg BH, Hildebrandt P, Keren A, Motro M, Moye LA, Otterstad JE, Pratt CM, Ponikowski P, Rouleau JL, Sestier F, Winkelmann BR, Young JB. Results of a non-specific immunomodulation therapy in chronic heart failure (ACCLAIM trial): a placebo-controlled randomised trial. Lancet. 2008;371:228–236.

62. Voll RE, Herrmann M, Roth EA, Stach C, Kalden JR, Girkontaite I. Immunosuppressive effects of apoptotic cells. Nature. 1997;390:350–351.

63. Maisch B, Hufnagel G, Kolsch S, Funck R, Richter A, Rupp H, Herzum M, Pankuweit S. Treatment of inflammatory dilated cardiomyopathy and (peri)myocarditis with immunosuppression and i.v. immunoglobulins. Herz. 2004;29:624–636.

64. Gullestad L, Aass H, Fjeld JG, Wikeby L, Andreassen AK, Ihlen H, Simonsen S, Kjekshus J, Nitter-Hauge S, Ueland T, Lien E, Froland SS, Aukrust P. Immunomodulating therapy with intravenous immunoglobulin in patients with chronic

heart failure. Circulation. 2001;103:220–225.

65. McNamara DM, Rosenblum WD, Janosko KM, Trost MK, Villaneuva FS, Demetris AJ, Murali S, Feldman AM. Intravenous immune globulin in the therapy of myocarditis and acute cardiomyopathy. Circulation. 1997;95:2476–2478.
66. McNamara DM, Holubkov R, Starling RC, Dec GW, Loh E, Torre-Amione G, Gass A, Janosko K, Tokarczyk T, Kessler P, Mann DL, Feldman AM. Controlled trial of intravenous immune globulin in recent-onset dilated cardiomyopathy. Circulation. 2001;103:2254–2259.
67. Weber KT. Extracellular matrix remodeling in heart failure: a role for de novo angiotensin II generation. 1997;96
68. Mann DL. Basic mechanisms of disease progression in the failing heart: the role of excessive adrenergic drive. Prog Cardiovasc Dis. 1998;41:1–8.
69. Kurrelmeyer K, Kalra D, Bozkurt B, Wang F, Dibbs Z, Seta Y, Baumgarten G, Engle D, Sivasubramanian N, Mann DL. Cardiac remodeling as a consequence and cause of progressive heart failure. Clin Cardiol. 1998;21:I14–I19.
70. Janicki JS, Brower GL, Henegar JR, Wang L. Ventricular remodeling in heart failure: the role of myocardial collagen. Adv Exp Med Biol. 1995;382:239–245.
71. Hudson MP, Armstrong PW, Ruzyllo W, Brum J, Cusmano L, Krzeski P, Lyon R, Quinones M, Theroux P, Sydlowski D, Kim HE, Garcia MJ, Jaber WA, Weaver WD. Effects of selective matrix metalloproteinase inhibitor (PG-116800) to prevent ventricular remodeling after myocardial infarction: results of the PREMIER (Prevention of Myocardial Infarction Early Remodeling) trial. J Am Coll Cardiol. 2006;48:15–20.
72. Vanhoutte D, Schellings M, Pinto Y, Heymans S. Relevance of matrix metalloproteinases

and their inhibitors after myocardial infarction: a temporal and spatial window. Cardiovasc Res. 2006;69:604–613.

73. Mann DL, Bristow MR. Mechanisms and models in heart failure: the biomechanical model and beyond. Circulation. 2005;111(21):2837–49.
74. Seidman JG, Seidman C. The genetic basis for cardiomyopathy: from mutation identification to mechanistic paradigms. Cell. 2001;104(4):557–67.
75. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature **409**, 860–921 (2001)
76. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. Nature **431**, 931–945 (2004)
77. Venter, J. C., et al. The sequence of the human genome. Science **291**, 1304–1351 (2001)
78. Bush WS, Moore JH (2012) Chapter 11: Genome-Wide Association Studies. PLoS Comput Biol 8(12): e1002822. doi:10.1371/journal.pcbi.1002822
79. Ghazalpour A, Rau CD, Farber CR, et al. Hybrid mouse diversity panel: a panel of inbred mouse strains suitable for analysis of complex genetic traits. Mammalian genome: official journal of the International Mammalian Genome Society. 2012;23(9-10):680-692.
80. Wang JJ, Rau C, Avetisyan R, et al. Genetic Dissection of Cardiac Remodeling in an Isoproterenol-Induced Heart Failure Mouse Model. PLOS Genetics. 2016: e1006038.
81. Gorczynski, R. M., Z. Chen, X. M. Fu, and H. Zeng. 1998. Increased expression of the novel molecule OX-2 is involved in prolongation of murine renal allograft survival. *Transplantation* 65:1106-1114.
82. Gorczynski, R. M., M. S. Cattral, Z. Chen, J. Hu, J. Lei, W. P. Min, G. Yu, and J. Ni. 1999. An immunoadhesin incorporating the molecule OX-2 is a potent

immunosuppressant that prolongs allo- and xenograft survival. *J Immunol* 163:1654-1660.

83. Gorczynski, R. M., Z. Cohen, X. M. Fu, and J. Lei. 1999. Anti-rat OX-2 blocks increased small intestinal transplant survival after portal vein immunization. *Transplant Proc* 31:577-578.

84. Gorczynski, L., Z. Chen, J. Hu, Y. Kai, J. Lei, V. Ramakrishna, and R. M. Gorczynski. 1999. Evidence that an OX-2-positive cell can inhibit the stimulation of type 1 cytokine production by bone marrow-derived B7-1 (and B7-2)-positive dendritic cells. *J Immunol* 162:774-781.

85. Gorczynski, R. M., J. Bransom, M. Cattral, X. Huang, J. Lei, L. Xiaorong, W. P. Min, Y. Wan, and J. Gauldie. 2000. Synergy in induction of increased renal allograft survival after portal vein infusion of dendritic cells transduced to express TGFbeta and IL-10, along with administration of CHO cells expressing the regulatory molecule OX-2. *Clin Immunol* 95:182-189.

86. Kenick, S., R. P. Lowry, R. D. Forbes, and R. Lisbona. 1987. Prolonged cardiac allograft survival following portal venous inoculation of allogeneic cells: what is "hepatic tolerance"? *Transplant Proc* 19:478-480.

87. Gorczynski, R. M., Z. Chen, K. Yu, and J. Hu. 2001. CD200 immunoadhesin suppresses collagen-induced arthritis in mice. *Clin Immunol* 101:328-334.

88. Hoek, R. M., S. R. Ruuls, C. A. Murphy, G. J. Wright, R. Goddard, S. M. Zurawski, B. Blom, M. E. Homola, W. J. Streit, M. H. Brown, A. N. Barclay, and J. D. Sedgwick. 2000. Down-regulation of the macrophage lineage through interaction with OX2 (CD200). *Science* 290:1768-1771.

89. Liu, Y., Y. Bando, D. Vargas-Lowy, W. Elyaman, S. J. Khoury, T. Huang, K. Reif, and T. Chitnis. CD200R1 agonist attenuates mechanisms of chronic disease in a murine model of multiple sclerosis. *J Neurosci* 30:2025-2038.

90. Broderick, C., R. M. Hoek, J. V. Forrester, J. Liversidge, J. D. Sedgwick, and A. D. Dick. 2002. Constitutive retinal CD200 expression regulates resident microglia and activation state of inflammatory cells during experimental autoimmune uveoretinitis. *Am J Pathol* 161:1669-1677.

91. Rau CD, Wang J, Avetisyan R, et al. Mapping Genetic Contributions to Cardiac Pathology Induced by Beta-Adrenergic Stimulation in Mice. *Circ Cardiovasc Genet.* Dec 5 2014.

92. Snelgrove RJ, Goulding J, Didierlaurent AM, et al. A critical function for CD200 in lung immune homeostasis and the severity of influenza infection. *Nat Immunol.* Sep 2008;9(9):1074-1083.

93. Yu K, Chen Z, Gorczynski R. Effect of CD200 and CD200R1 expression within tissue grafts on increased graft survival in allogeneic recipients. *Immunol Lett.* Jan 2013;149(1-2):1-8.

94. Lindenfeld J, Miller GG, Shakar SF, Zolty R, Lowes BD, Wolfel EE, Mestroni L, Page RL 2nd, Kobashigawa J. Drug therapy in the heart transplant recipient: part I: cardiac rejection and immunosuppressive drugs. *Circulation.* Dec 14 2004;110(24):3734-40.

95. Cohn, JN., Ferrari, R.,Sharpe, N. on behalf of an International Forum on Cardiac Remodeling. Cardiac remodeling—concepts and clinical implications: a consensus paper from an international forum on cardiac remodeling. *J. Am. Coll. Cardiol.* 2000*;*35, 569–582.

96. White HD. *et al.* Left ventricular end-systolic volume as the major determinant of survival after recovery from myocardial infarction. *Circulation.*1987*;*76, 44–51.

97. Padmanabhan M, Prince PS (2006) Preventive effect of S-al- lylcysteine on lipid peroxides and antioxidants in normal and isoproterenol-induced cardiotoxicity in rats: a histopathological study. Toxicology 224:128–137.

98. Padmanabhan M, Rajadurai M, Prince PS (2008) Preventive effect of S-allylcysteine on membrane-bound enzymes and glycoproteins in normal and isoproterenol-induced cardiac tox- icity in male Wistar rats. Basic Clin Pharmacol Toxicol 103:507–513.

99. Teerlink JR, Pfeffer JM, Pfeffer MA (1994) Progressive ven- tricular remodeling in response to diffuse isoproterenol-induced myocardial necrosis in rats. Circ Res 75:105–113.

100. El-Demerdash E, Awad AS, Taha RM, El-Hady AM, Sayed- Ahmed MM (2005) Probucol attenuates oxidative stress and energy decline in isoproterenol-induced heart failure in rat. Pharmacol Res 51:311–318.

101. Oudit GY, Crackower MA, Eriksson U, Sarao R, Kozieradzki I, Sasaki T, Irie-Sasaki J, Gidrewicz D, Rybin VO, Wada T, Steinberg SF, Backx PH, Penninger JM (2003) Phosphoinositide 3-kinase gamma-deficient mice are protected from isoprotere- nol-induced heart failure. Circulation 108:2147–2152.

102. Grimm D, Elsner D, Schunkert H, Pfeifer M, Griese D, Bruckschlegel G, Muders F, Riegger GA, Kromer EP (1998) Development of heart failure following isoproterenol adminis- tration in the rat: role of the renin-angiotensin system. Cardio- vasc Res 37:91–100

103. Zbinden G, Moe RA (1969) Pharmacological studies on heart muscle lesions induced by

isoproterenol. Ann N Y Acad Sci 156:294–308

104. Halapas A, Papalois A, Stauropoulou A, Philippou A, Pissi- missis N, Chatzigeorgiou A, Kamper E, Koutsilieris M (2008) In vivo models for heart failure research. In Vivo 22:767– 780

105. Zbinden G, Bagdon RE (1963) Isoproterenol-induced heart necrosis, an experimental model for the study of Angina Pectoris and Myocardial Infarct. Rev Can Biol 22:257–263

106. Gorczynski RM. Cd200:Cd200r-mediated regulation of immunity. In: Orentas R, Puccetti P, eds. *Isrn immunology*. 2012.

107. Gorczynski R, Chen Z, Khatri I, Yu K. Scd200 present in mice receiving cardiac and skin allografts causes immunosuppression in vitro and induces tregs. *Transplantation*. 2013;95:442-447

108. Arik HO, Yalcin AD, Celik B, Seyman D, Tetik G, Gursoy B, Kose S, Gumuslu S. Evaluation of soluble cd200 levels in type 2 diabetic foot and nephropathic patients: Association with disease activity. *Med Sci Monit*. 2014;20:1078-1081

109. Yalcin AD, Tural Onur S, Celik B, Gumuslu S. Evaluation of d-dimer, cxcl8, homocysteine, eosinophil cationic peptide, 25(oh)-vitamin d, and immunomodulatory ox-2 levels in allergic patients. *J Asthma*. 2014:1-20

110. Yang, L, Zhao, L, Tong, L, Qian, S, Ren, Y, Zhang, L, Ging, X, Wang, Y, Zhang, W, Lipsky, P. Aberrant CD200/CD200R1 expression and function in systemic lupus erythematosus contributes to abnormal T-cell responsiveness and dendritic cell activity. *Arthritis Research & Therapy.* 2012.

*Chapter 4: Low Glycoprotein NMB (GPNMB) levels are associated with Heart Failure*

**Low Glycoprotein NMB (GPNMB) levels are associated with Heart Failure**

Lin L-Y[1,2,3], Chang S[1], Gupta P[1], Deng M[1], Laakso M[4], Sinsheimer JS[7-8], Yibin Wang[6], Lusis AJ[1,5-7], Wang J[1*], Huertas-Vazquez A[1*]

1. Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, USA. 2. Division of Endocrinology and Metabolism, Department of Medicine, Taipei Veterans General Hospital, Taipei, Taiwan. 3. Faculty of Medicine, National Yang-Ming University, Taipei, Taiwan 4. Institute of Clinical Medicine, Internal Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland. 5. Department of Microbiology, Immunology and Molecular Genetics, David Geffen School of Medicine, University of California, Los Angeles, USA 6. Departments of Anesthesiology, Physiology, and Medicine, Cardiovascular Research Laboratories, David Geffen School of Medicine, University of California, Los Angeles, USA 7. Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, USA. 8. Department of Biomathematics, David Geffen School of Medicine, University of California, Los Angeles, USA.

**Background** We had previously examined the cardiac transcriptome across control versus isoproterenol-treated inbred mouse strains and observed significant differential regulation of cardiac transcripts which had previously been identified as heart failure plasma biomarkers, such as *Lgals3* (galectin-3) and *Timp1* (TIMP metallopeptidase inhibitor 1). We had noted the Glycoprotein non-metastatic melanoma *(Gpnmb)* to be similarly regulated and postulated it to be a candidate for a novel heart failure biomarker. GPNMB, also known as osteoactivin, is a type I transmembrane glycoprotein, whose ectodomain is cleaved and shed by a disintegrin and metallopeptidase domain-containing protein 10 (ADAM10), into the blood stream. We aimed to determine the association between GPNMB levels and heart failure in mouse models and human samples.

**Methods** We examined the cardiac transcriptome between control and isoproterenol-treated inbred mouse strains and identified candidate transcripts, including *Gpnmb*, that were significantly differentially regulated. We confirmed protein expression of GPNMB in cardiac tissue of C57BL/6J treated with isoproterenol. Using ELISA assays, plasma levels of GPNMB were measured in the transverse aortic constriction (TAC) and the isoproterenol-induced heart failure mouse models. Plasma levels of GPNMB were measured in 119 heart failure cases and 272 controls from the METabolic Syndrome In Men (METSIM) study.

**Results** Similar to our finding in the cardiac transcriptome, we observed an increase in GPNMB protein expression in isoproterenol-treated C57BL/6J hearts. A significant decrease in the plasma levels of GPNMB was observed in the TAC (6.09±1.65 ng/mL in sham vs. 1.86±1.90 ng/ml in TAC, p=0.028) and the isoproterenol-treated (5.96±2.66 ng/mL in control vs. 3.18±1.08 in ISO, p=0.007) heart failure models. Similar to our observation in mice, GPNMB levels were lower in

the plasma of heart failure cases compared with the controls (19.1±10.8 ng/mL in control vs. 8.4±8.7 in heart failure, $p<0.001$) among 391 human subjects.

**Conclusions** Our results in mice and humans provide compelling evidence that circulating GPNMB levels are associated with HF. Interestingly, while the cardiac transcript and protein expression levels were increased in the heart in mice, the circulating plasma levels of GPNMB were decreased in the diseased state in both mice and humans. The mechanism underlying this association and the relevance role of GPNMB and its cleavage products in heart failure warrant further investigation.

**Key words:** Heart failure, cardiac transcriptome, biomarkers, glycoprotein NMB/Osteoactivin.

## Introduction

Heart failure (HF) is a major public health problem with an alarming annual incidence of approximately 6 million [1], and yet current methods of prediction and prevention are insufficient. Although several circulating biomarkers of diagnosis, underlying pathophysiology and risk stratification for HF, have been identified [2-10], their relative contribution on HF is unclear. More accurate prediction of disease manifestation and progression can provide valuable information to guide patient counseling, life planning and patient selection for heart transplantation. Given the high prevalence of HF in combination with its high rates of mortality, as well as the lack of reliable predictors, the discovery of novel biomarkers for diagnosis and risk assessment remains a priority.

By assessing cardiac transcriptome between mice sustaining chronic injury due to systemic isoproterenol (ISO) treatment and controls from the Hybrid Mouse Diversity Panel (HMDP), an inbred population of mice with natural genetic diversity, we identified the Glycoprotein non-metastatic melanoma B (GPNMB) as a novel candidate biomarker for HF. Because the HMDP mice were raised in a controlled environment and given a uniform cardiac injury, we were able to generate data that were much less confounded by other factors compared to studies in human populations [11, 12]. GPNMB is a type 1 transmembrane protein [13] that has been involved in several cancers and sarcomas [14-17], osteoblast differentiation and function [18-20] and most recently in myocardial infarction (MI) where *Gpnmb* was found to be upregulated after MI in murine hearts [21]. In the present study, we aim to determine whether plasma levels of GPNMB are associated with HF. We investigated plasma levels of GPNMB in two alternative mouse models for HF, the transverse aortic constriction (TAC) and the

isoproterenol models. Finally, we examined plasma levels of GPNMB in human subjects with and without HF from the Metabolic Syndrome in Men (METSIM) study [22, 23].

## Methods

### Cardiac Transcriptome analysis in the HMDP

The cardiac transcriptome data from the Heart Failure-HMDP study was published previously [24]. Briefly, RNA was extracted from frozen left ventricular tissues using RNeasy columns (QIAGEN, Valencia, CA, USA). RNA quality was assessed using the Bioanalyzer RNA kits (Agilent Technologies, Santa Clara, CA, USA). Expression profiling of pooled biological replicates was performed using Illumina MouseRef-8 v2.0 Expression BeadChip arrays (Illumina, Inc., San Diego, CA, USA). Expression profiles for 90 control and 91 ISO strains were included. Moderated t-statistics and associated p values were calculated using the R package limma [25]. Correction for multiple testing was performed using the Benjamini-Hochberg procedure. Probes with log2-fold change > 0.2 and adjusted p-value < 0.05 were considered significantly differentially expressed. A total of 1,502 of the 18,335 probes were differentially expressed at > 15% change at an adjusted p-value < 0.05.

### Selection of GPNMB as a candidate biomarker for HF

Based on our cardiac transcriptome results between controls and mice sustaining chronic injury due to systemic isoproterenol (ISO) treatment (Table 1), we chose to investigate GPNMB as an attractive novel biomarker for HF based on differential expression data (logFC=1.6, p-value=5.21x10-$^{26}$). *Gpnmb* was the third most significantly upregulated transcript after *Lgals3*

(galectin-3) and *Timp1* (TIMP metallopeptidase inhibitor 1), whose circulating plasma counterparts have already been well-established as HF plasma biomarkers in humans [26-28].

**Western blot analysis of gpnmb in heart tissues of ISO treated mice and control**

Proteins from heart tissue were harvested in buffer (50mM HEPES [pH7.4], 150mM NaCl, 1% NP-40, 1mM EDTA, 1mM EGTA, 1mM glycerophosphate, 2.5mM sodium pyrophosphate 1mM Na3VO4, 20mM NaF, 1 mM phenylmethylsulfonyl fluoride, 1 μg/mL of aprotinin, leupeptin, and pepstatin). Equal amounts of protein were separated on 4-12% Bis-Tris gels (Invitrogen) using an electroblotting apparatus (Bio-Raand Laboratories, Hercules, CA, USA) and transferred onto a nitrocellulose blot (Amersham). The blot was probed with the indicated primary antibodies using the polyclonal anti-Gpnmb (R&D Systems) and anti-Gadph (Invitrogen). Protein signals were detected using HRP conjugated secondary antibodies (Cell Signaling Technologies) and enhanced chemiluminescence (ECL) western blotting detection regents (Amersham, GH Healthcare).

**Mouse Models of Heart Failure**

We assessed circulating GPNMB levels in 2 well-established mouse HF models representing cardiac insults caused by pressure overload (transverse aortic constriction) or by chronic β-adrenergic stimulation induced cardiac hypertrophy (isoproterenol continuous infusion). For the TAC model, midsternal incision was made to expose transverse aorta between truncus anonymous and the left carotid artery. With 6-0 silk suture, a ligature is tied around the transverse aorta against a 26-gauge needle, as previously described [29]. Mice were divided to

Sham surgery group, which received midsternal incision to expose transverse aorta only. At the end of the experiments, animals were euthanized and the hearts and lungs were removed and weighed. Hearts were dissected and tissues were either immediately immersed into 4% buffered formaldehyde or quickly frozen in liquid nitrogen for further experiments. For the ISO model, mice were anesthetized with intraperitoneal ketamine as a surgical anesthetic agent, and osmotic minipumps (Alzet 1004, Durect, Cupertino, CA) were implanted subcutaneously under aseptic conditions, as previously described [24], to deliver isoproterenol (30 mg/kg/day). Mice were divided into two experimental groups as control and ISO-treated for 21 days. Both HF models were performed in 10 week-old C57BL/6J mice, and plasma samples were collected by retro-orbital puncture at the time of euthanization from TAC mice (n=4) 4 weeks after intervention and ISO-infusion mice (n=10) for 3 weeks. Blood was collected in a BD Microtainer blood collection tube using a microhematocrit heparinized capillary and retro-orbital bleeding, centrifuged at maximum speed for 10 minutes, transferred to a microcentrifuge tube, and then flash frozen in liquid nitrogen until measurement. The UCLA Institutional Animal Care and Use Committee (IACUC) approved all animal studies.

**Cross-sectional study of the METSIM cohort**

The Metabolic Syndrome in Men (METSIM) Study comprises 10,197 Finnish men aged from 45 to 74 years (mean±SD 58±7 years; mean body mass index [BMI] 27±4 kg/m$^2$), and randomly selected from the population register of Kuopio town, Eastern Finland. The METSIM study and its methods have been described in detail elsewhere [22, 23]. The cross-sectional study performed in 2005–2010 was approved by the Ethics Committee of the University of Eastern Finland and Kuopio University Hospital and conducted in accordance with the Helsinki

Declaration. Each study participant gave written informed consent. Congestive HF cases were identified by screening medical records for the HF diagnostic code and by querying the Finnish medication reimbursement database for HF medications from the METSIM cohort. A total of 119 subjects with congestive HF were identified. A total of 272 control subjects with similar age range with no previous diagnosis of HF or EF> 50% or current clinical or biochemical indication of cardiovascular diseases or other chronic disease were determined to be controls. Blood samples were collected at out-patient clinics in East Finland after a minimum of 10 hours fasting. Plasma glucose was measured by enzymatic hexokinase photometric assay (Konelab Systems reagents, Thermo Fisher Scientific; Vantaa, Finland). Low-density lipoprotein cholesterol (LDL-c) and high-density lipoprotein cholesterol (HDL-c) were measured by enzymatic colorimetric tests (Konelab Systems Reagents). Plasma creatinine was measured by standard methods (Konelab Systems Reagents) and estimated glomerular filtration rate (eGFR) were calculated using MDRD formula. The immunoturbidimetric assay by Beckman Coulter (Fullerton, CA, USA) was used for the plasma high-sensitivity C-reactive protein (Hs-CRP) measurements.

**GPNMB measurements in mice and humans**

Plasma GPNMB levels in mice and human samples were assayed using commercial enzyme-linked immunosorbent assay Kits from R&D (R&D systems, Minneapolis MN) (Catalogue numbers: DY2330 and DY2550, respectively).

**Statistical analysis**

The t-test statistic was used to examine differences between ISO treated and control plasma GPNMB levels and protein levels of mice heart tissues, with statistical significance

considered for p-value < 0.05. Clinical characteristics of HF cases and non-HF controls were compared using t-tests for continuous variables and Pearson chi-square or Fisher exact tests for categorical variables (Table 2). Plasma samples from 119 subjects with HF and 272 non-HF controls from METSIM were included to evaluate the effect of GPNMB levels on risk of HF using HF /non-HF controls as the outcome. We incorporated known risk factors of cardiovascular outcomes as covariates, such as age, gender, blood pressure and history of MI. We used a multivariate logistic regression to estimate the effect of GPNMB levels on risk of HF.

## Results

### GPNMB in mouse models of HF

The complete list of up and down-regulated genes between control and ISO treated mice form the HMDP have been published previously [24]. The top 20 up and down regulated genes are shown in Table 1. Up-regulated genes with a fold change of 1.5 or higher were observed for the TIMP Metallopeptidase Inhibitor 1 (timp1), log FC=20.4, $p=1.1X10^{-22}$; galectin 3 (lgals3) logFC=1.82, $p=1.6x10^{-22}$, and Glycoprotein NMB (GPNMB/Osteoactivin), logFC=1.6, $p=1.26x10^{-10}$. In addition, GPNMB was highly correlated with left ventricular internal dimension (LVID) (bcor=0.363, p value= 0.00052). Since TIMP1 and Galectin-3 are known biomarkers for HF, we hypothesized that further investigation of the differentially expressed genes, such as GPNMB, may reveal additional biomarkers of HF and elucidate underlying biology.

We measured GPNMB levels in two alternative mouse models for HF, the TAC and the ISO models. In the ISO model, after 3 weeks of continuous infusion of ISO, plasma levels of GPNMB were lower in the ISO-treated group than in the control group (5.96±2.66 ng/mL in control vs. 3.18±1.08 in ISO, p=0.007) (Figure 1A). Similar findings were found in the TAC

model. As compared with Sham surgery group animals, the TAC animal models of HF showed significant decrease in GPNMB levels (6.09±1.65 ng/mL in sham vs. 1.86±1.90 ng/ml in TAC, p=0.028) at 4 weeks after surgery (Figure 1B). Finally, to elucidate the potential mechanism, we examined the protein expression of GPNMB in heart lysates in the ISO-treated HF model. We observed that ISO-treatment resulted in increased GPNMB expression (the ratio of GPNMB/Gadph 0.710 ± 0.062 in ISO vs. 0.452 ± 0.080 in control, $p < 0.05$, n = 3) (Figure 2), suggesting that there is increased GPNMB cardiac expression in the HF mice group.

**GPNMB levels in human HF from the METSIM study**

To further investigate the relationship between plasma GPNMB levels and HF, we measured plasma GPNMB levels in 119 HF subjects and 272 non-HF control subjects from the METSIM study. Patients' baseline characteristics are listed in Table 2. Significant differences between HF subjects and controls included age, body mass index (BMI), systolic blood pressure (SBP), history of hypertension (HTN), diabetes (T2DM), coronary artery disease (CAD), glucose, LDL-c, HDL-c, high sensitivity CRP and eGFR. Those with HF were older, had higher BMI, higher SBP and possessed risk factors such as HTN, T2DM, CAD, hyperlipidemia and decreased renal function.

As observed in mice, there were significantly lower plasma GPNMB levels in patients with HF compared with non-HF controls (19.1±10.8 ng/mL in control vs. 8.4±8.7 in heart failure, p<0.001) (Figure 3). A multiple linear regression analysis revealed that the presence of HF was significantly associated with the decreased plasma levels of GPNMB (Table 3). The distribution of plasma GPNMB did not reveal normality in both control and HF groups (Supplementary figure 2). In a subset of HF cases, where proBNP levels were available, there

was no significant correlation between log transformed plasma GPNMB and proBNP ($r^2$<0.001, p=0.863) (Figure 4), which argues in support of it being a novel biomarker that provides additional information than simply being a surrogate marker for proBNP. A negative association between plasma log GPNMB and high sensitivity C-reactive protein in METSIM population (r=-0.1503, *p*=0.0032) was observed (Supplementary figure 3).

**Discussion**

In the present study, we analyzed cardiac transcriptome data from the Heart Failure-HMDP study as an innovative strategy to identify novel plasma biomarkers of heart failure (HF). We found that transcripts of established HF plasma biomarkers including TIMP1 and LGALS3 as well as novel transcripts were differentially expressed in our study. Based on differential expression data from the ISO-induced HF model in the HMDP, we identified GPNMB as an attractive potential biomarker for HF (logFC=1.6, p-value=5.21x10-$^{26}$) with subsequent validation in mice and human plasma samples. First, we confirmed the presence of circulating GPNMB in the plasma of mice treated with ISO continuous infusion. We then measured plasma GPNMB levels in a HF mouse model with cardiac injury induced by TAC. We found significantly lower levels of circulating GPNMB in both HF mouse models when compared to control samples (Figure1). Furthermore, we investigated whether low levels of circulating GPNMB were associated with chronic HF in humans in the outpatient setting. As observed in mice, circulating plasma GPNMB levels were also lower in patients with HF from the METSIM study compared to the control group.

Although the function of GPNMB in HF is unknown, previous studies have suggested that GPNMB inhibits the activation of T lymphocytes by binding syndecan 4 [30], a

proteoglycan that is up-regulated in chronic HF [31] and has been previously shown to adversely influence cardiac remodeling [32]. In addition, GPNMB is known to be involved in inflammation and fibrosis after ischemic heart injury. Inflammatory pathways are critically involved in dilative and fibrotic remodeling of the infarcted heart and, therefore, drive key events in the pathogenesis of HF after myocardial injury. If GPNMB were to serve as an inflammatory stop signal in HF, it is possible that increased consumption of GPNMB or lower circulating levels of GPNMB could be indicative of more severe HF or compensation of HF in order to preserve heart function. A previous study has shown that upon macrophage activation, increased GPNMB is associated with a reduction of IL-6, suggesting that GPNMB may play a role in mitigating inflammatory responses [33]. Interestingly, we found a negative association between plasma GPNMB levels and high sensitivity C-reactive protein, a marker of inflammation, in the METSIM population ($r=-0.1503$, $p=0.0032$). Recently, Jarve et al. have also reported that in mice with lack functional GPNMB due to a natural spontaneous point mutation, GPNMB deficiency was associated with preserved cardiac function and less ventricular dilatation after MI [21].

Here, we observed an overexpression of GPNMB in cardiac transcriptome data and protein level in HF setting, but lower protein concentrations of GPNMB in plasma in vivo study. It is well known that many steps in protein transcription, translation and degradation can affect mRNA expression and protein abundances [34]. The exact mechanism by which GPNMB exerts its effect on HF are still unknown. GPNMB is a glycoprotein localized at the cell surface that results from ectodomain shedding by ADAM Metallopeptidase Domain 10 (ADAM10) [35]. Interestingly, we found that ADAM10 expression was inversely correlated with GPNMB in the HMDP cardiac transcriptome data (Supplementary Figure 1), suggesting ADAM10 could play a

role in GPNMB regulation.

Finally, we found that lower GPNMB levels in plasma were not associated with elevated proBNP levels, suggesting that GPNMB may be predictive of outcomes based on properties that are dissimilar to the most commonly used biomarker for HF. Additional biomarkers that ascertain various properties of HF may be important addition to the full evaluation of HF susceptibility.

**Limitations**

We have previously identified an association between GPNMB and HF in female mice. A limitation of our study is that only men were included in the METSIM study, and therefore, we do not know whether the observed results will be applicable to women. However, we speculate that the same findings will reproduce in female patients. Despite the gender restriction, we consider the METSIM study sample as a valuable resource that provides a unique opportunity to investigate potential biomarkers in a very well characterized population. Second, the predictive value of GPNMB for HF outcomes requires a follow-up study. Whether the plasma GPNMB is a predictor for HF hospitalization and cardiovascular mortality was not assessed in this study. Finally, HF in humans is a complex disease with multiple causes such as CAD, HTN and DM. In our study, we investigated two established HF mice models induced by pressure overload and cardiomyopathy to elucidate the association between circulating GPNMB and HF. However, these two HF mouse models could not completely represent the whole pathogenesis of HF in humans.

In conclusion, our results in mice and humans provide compelling evidence that circulating GPNMB could be an informative biomarker for HF. Further studies need to be performed to investigate the mechanisms underlying GPNMB and its role in HF.

**Acknowledgements**

The support of Professor Yii-Der Ida Chen is greatly appreciated.

**Sources of Funding**

This work was supported in part by research grant Taipei Veterans General Hospital-National Yang-Ming University Excellent Physician Scientists Cultivation Program, No. 105-V-A-008 to L.Y.L..

**Bibliography**


1. Mozaffarian, D., et al., *Heart Disease and Stroke Statistics-2016 Update: A Report From the American Heart Association.* Circulation, 2015.
2. van Kimmenade, R.R., et al., *Utility of amino-terminal pro-brain natriuretic peptide, galectin-3, and apelin for the evaluation of patients with acute heart failure.* J Am Coll Cardiol, 2006. **48**(6): p. 1217-24.
3. Shah, R.V., et al., *Galectin-3, cardiac structure and function, and long-term mortality in patients with acutely decompensated heart failure.* Eur J Heart Fail, 2010. **12**(8): p. 826-32.
4. van der Velde, A.R., et al., *Prognostic value of changes in galectin-3 levels over time in patients with heart failure: data from CORONA and COACH.* Circ Heart Fail, 2013. **6**(2): p. 219-26.
5. Kakkar, R. and R.T. Lee, *The IL-33/ST2 pathway: therapeutic target and novel biomarker.* Nat Rev Drug Discov, 2008. **7**(10): p. 827-40.
6. Manzano-Fernandez, S., et al., *Usefulness of soluble concentrations of interleukin family member ST2 as predictor of mortality in patients with acutely decompensated heart failure relative to left ventricular ejection fraction.* Am J Cardiol, 2011. **107**(2): p. 259-67.

7. Rehman, S.U., T. Mueller, and J.L. Januzzi, Jr., *Characteristics of the novel interleukin family biomarker ST2 in patients with acute heart failure.* J Am Coll Cardiol, 2008. **52**(18): p. 1458-65.

8. Fonarow, G.C., et al., *Admission B-type natriuretic peptide levels and in-hospital mortality in acute decompensated heart failure.* J Am Coll Cardiol, 2007. **49**(19): p. 1943-50.

9. Bettencourt, P., et al., *N-terminal-pro-brain natriuretic peptide predicts outcome after hospital discharge in heart failure patients.* Circulation, 2004. **110**(15): p. 2168-74.

10. Peacock, W.F.t., et al., *Cardiac troponin and outcome in acute heart failure.* N Engl J Med, 2008. **358**(20): p. 2117-26.

11. Ghazalpour, A., et al., *Hybrid mouse diversity panel: a panel of inbred mouse strains suitable for analysis of complex genetic traits.* Mamm Genome, 2012. **23**(9-10): p. 680-92.

12. Lusis, A.J., et al., *The Hybrid Mouse Diversity Panel: a resource for systems genetics analyses of metabolic and cardiovascular traits.* J Lipid Res, 2016. **57**(6): p. 925-42.

13. Selim, A.A., *Osteoactivin bioinformatic analysis: prediction of novel functions, structural features, and modes of action.* Med Sci Monit, 2009. **15**(2): p. MT19-33.

14. Kuan, C.T., et al., *Glycoprotein nonmetastatic melanoma protein B, a potential molecular therapeutic target in patients with glioblastoma multiforme.* Clin Cancer Res, 2006. **12**(7 Pt 1): p. 1970-82.

15. Rose, A.A., et al., *Osteoactivin promotes breast cancer metastasis to bone.* Mol Cancer Res, 2007. **5**(10): p. 1001-14.

16. Oyewumi, M.O., et al., *Osteoactivin (GPNMB) ectodomain protein promotes growth and invasive behavior of human lung cancer cells.* Oncotarget, 2016. **7**(12): p. 13932-44.

17. Maric, G., et al., *GPNMB cooperates with neuropilin-1 to promote mammary tumor growth and engages integrin alpha5beta1 for efficient breast cancer metastasis.* Oncogene, 2015. **34**(43): p. 5494-504.

18. Selim, A.A., et al., *Anti-osteoactivin antibody inhibits osteoblast differentiation and function in vitro.* Crit Rev Eukaryot Gene Expr, 2003. **13**(2-4): p. 265-75.

19. Abdelmagid, S.M., et al., *Osteoactivin, an anabolic factor that regulates osteoblast differentiation and function.* Exp Cell Res, 2008. **314**(13): p. 2334-51.

20. Moussa, F.M., et al., *Osteoactivin promotes osteoblast adhesion through HSPG and alphavbeta1 integrin.* J Cell Biochem, 2014. **115**(7): p. 1243-53.

21. Jarve, A., et al., *Adverse left ventricular remodeling by glycoprotein nonmetastatic melanoma protein B in myocardial infarction.* FASEB J, 2017. **31**(2): p. 556-568.

22. Stancakova, A., et al., *Changes in insulin sensitivity and insulin release in relation to glycemia and glucose tolerance in 6,414 Finnish men.* Diabetes, 2009. **58**(5): p. 1212-21.

23. Laakso, M., et al., *The Metabolic Syndrome in Men study: a resource for studies of metabolic and cardiovascular diseases.* J Lipid Res, 2017. **58**(3): p. 481-493.

24. Wang, J.J., et al., *Genetic Dissection of Cardiac Remodeling in an Isoproterenol-Induced Heart Failure Mouse Model.* PLoS Genet, 2016. **12**(7): p. e1006038.

25. Ritchie, M.E., et al., *limma powers differential expression analyses for RNA-sequencing and microarray studies.* Nucleic Acids Res, 2015. **43**(7): p. e47.

26. de Boer, R.A., et al., *Galectin-3: a novel mediator of heart failure development and progression.* Eur J Heart Fail, 2009. **11**(9): p. 811-7.

27. Ho, J.E., et al., *Galectin-3, a marker of cardiac fibrosis, predicts incident heart failure in the community.* J Am Coll Cardiol, 2012. **60**(14): p. 1249-56.

28. Goldbergova, M.P., et al., *The association between levels of tissue inhibitor of metalloproteinase-1 with acute heart failure and left ventricular dysfunction in patients with ST elevation myocardial infarction treated by primary percutaneous coronary intervention.* Genet Test Mol Biomarkers, 2012. **16**(10): p. 1172-8.

29. Sun, H., et al., *Catabolic Defect of Branched-Chain Amino Acids Promotes Heart Failure.* Circulation, 2016. **133**(21): p. 2038-49.

30. Chung, J.S., et al., *DC-HIL is a negative regulator of T lymphocyte activation.* Blood, 2007. **109**(10): p. 4320-7.

31. Takahashi, R., et al., *Serum syndecan-4 is a novel biomarker for patients with chronic heart failure.* J Cardiol, 2011. **57**(3): p. 325-32.

32. Kojima, T., et al., *Plasma levels of syndecan-4 (ryudocan) are elevated in patients with acute myocardial infarction.* Thromb Haemost, 2001. **85**(5): p. 793-9.

33. Ripoll, V.M., et al., *Gpnmb is induced in macrophages by IFN-gamma and lipopolysaccharide and acts as a feedback regulator of proinflammatory responses.* J Immunol, 2007. **178**(10): p. 6557-66.

34. Vogel, C. and E.M. Marcotte, *Insights into the regulation of protein abundance from proteomic and transcriptomic analyses.* Nat Rev Genet, 2012. **13**(4): p. 227-32.

35. Rose, A.A., et al., *ADAM10 releases a soluble form of the GPNMB/Osteoactivin extracellular domain with angiogenic properties.* PLoS One, 2010. **5**(8): p. e12093.

**Figure legends**

**Figure 1.** GPNMB levels are lower in the isoproterenol (ISO) infusion (A) and transverse aortic constriction (TAC) (B) heart failure mice models.

**Figure 2.** Western blot analysis of GPNMB expression in C57BL/6J mouse heart lysates after isoproterenol (ISO) treatment.

**Figure 3.** Box plots showing differences in GPNMB values between patients with heart failure compared with those without heart failure in METSIM study.

**Figure 4.** Non-significant correlation between log GPNMB and log proBNP levels in a subgroup of heart failure subjects (n=48).

**Table 1.** List of the top 20 up and down regulated genes from the cardiac transcriptome analysis in mice treated with ISO-proterenol vs control.

| PROBE_ID | SYMBOL | logFC | AveExpr | t | adj.P.Val | B |
|---|---|---|---|---|---|---|
| ILMN_3103896 | Timp1 | 2.040682 | 8.750288 | 15.27314 | 1.01E-22 | 48.61835 |
| ILMN_2769918 | Timp1 | 2.025033 | 8.69029 | 15.09472 | 1.65E-22 | 47.90656 |
| ILMN_1223317 | Lgals3 | 1.822672 | 8.651887 | 15.12024 | 1.62E-22 | 48.00865 |
| ILMN_2648669 | Gpnmb | 1.608681 | 6.968906 | 8.243321 | 1.26E-10 | 17.94402 |
| ILMN_2690603 | Spp1 | 1.314258 | 5.357506 | 7.563696 | 2.09E-09 | 14.86774 |
| ILMN_2997494 | Lox | 1.296259 | 6.909805 | 10.63671 | 6.01E-15 | 28.84176 |
| ILMN_1232261 | Catnal1 | 1.177767 | 10.99792 | 16.91307 | 7.12E-25 | 54.9607 |
| ILMN_2975345 | Cdo1 | 1.167495 | 7.843213 | 7.515948 | 2.54E-09 | 14.65316 |
| ILMN_2666312 | BC025833 | -1.0907 | 8.773729 | -12.7884 | 1.09E-18 | 38.33156 |
| ILMN_2844820 | Angptl7 | 1.089466 | 7.797022 | 5.603739 | 5.27E-06 | 6.387533 |
| ILMN_3091003 | Ms4a7 | 1.052859 | 7.410336 | 13.43802 | 9.46E-20 | 41.09684 |
| ILMN_2968211 | Lgals4 | -1.02901 | 9.78074 | -9.34586 | 1.31E-12 | 22.97708 |
| ILMN_2626235 | Lgals4 | -0.98579 | 8.280642 | -10.5643 | 8.16E-15 | 28.51499 |
| ILMN_1225835 | Mfap5 | 0.97532 | 10.30179 | 11.94393 | 3.18E-17 | 34.66261 |
| ILMN_2769795 | Phkg1 | 0.943491 | 8.426488 | 9.829841 | 1.78E-13 | 25.1846 |
| ILMN_1249000 | 1500015O10Rik | 0.924078 | 6.323165 | 4.624987 | 0.000171 | 2.569494 |
| ILMN_2946088 | Panx1 | 0.850637 | 6.586325 | 11.20643 | 6.32E-16 | 31.39737 |
| ILMN_2753809 | Mmp3 | 0.845797 | 7.024605 | 7.140418 | 1.18E-08 | 12.97523 |

| | | | | | | |
|---|---|---|---|---|---|---|
| ILMN_2692316 | Dpep2 | 0.822808 | 5.599864 | 10.22991 | 3.21E-14 | 27.00254 |
| ILMN_2687872 | Col1a1 | 0.820139 | 9.02824 | 9.19732 | 2.45E-12 | 22.29851 |
| ILMN_2680770 | Cysltr1 | 0.813838 | 5.817484 | 13.26209 | 1.65E-19 | 40.35302 |
| ILMN_2711163 | Ctsk | 0.813322 | 7.99203 | 9.999707 | 8.44E-14 | 25.95743 |
| ILMN_3117381 | Fhl1 | 0.807907 | 12.7992 | 15.0029 | 2.06E-22 | 47.53863 |
| ILMN_2713285 | Fhl1 | 0.807591 | 12.86946 | 15.26067 | 1.01E-22 | 48.56876 |

**Table 2.** Clinical characteristics of patients with heart failure and control subjects

| | Control (n=272) | Heart failure (n=119) | P value |
|---|---|---|---|
| Age (years) | 57.3±4.9 | 64.3±5.7 | <0.001 |
| Body mass index (kg/m$^2$) | 26.8±3.7 | 30.3±6.0 | <0.001 |
| Former/current smokers, n (%) | 160 (58.8) | 85 (71.4) | 0.092 |
| Systolic blood pressure (mmHg) | 134.6±14.9 | 143.3±17.7 | <0.001 |
| Diastolic blood pressure (mmHg) | 86.9±8.9 | 85.2±10.1 | 0.084 |
| Hypertension, n (%) | 34 (12.5) | 55 (46.2) | <0.001 |
| Diabetes mellitus, n (%) | 0 (0.0) | 41 (34.5) | <0.001 |
| Coronary artery disease, n (%) | 0 (0.0) | 63 (52.9) | <0.001 |
| Arrhythmia, n (%) | 2 (0.7) | 15 (12.6) | <0.001 |
| NYHA Fc III-IV, n (%) | | 79 (66.4) | |
| Antiplatelets, n (%) | 3 (1.1) | 58 (48.7) | <0.001 |

| | | | |
|---|---|---|---|
| ACEi or ARB, n (%) | 30 (11.0) | 74 (62.2) | <0.001 |
| Beta-blocker, n (%) | 9 (3.3) | 85 (71.4) | <0.001 |
| Diuretics, n (%) | 19 (7.0) | 58 (48.7) | <0.001 |
| Statins, n (%) | 45 (16.5) | 79 (66.4) | <0.001 |
| Nitrates, n (%) | 1 (3.7) | 30 (25.2) | <0.001 |
| Left ventricular ejection fraction (%) | | 42.4±14.7 | |
| Glucose (mg/dL) | 103.8±10.8 | 117.3±29.6 | <0.001 |
| eGFR (mL/min/1.73 $m^2$) | 88.2±12.4 | 82.2±20.4 | <0.001 |
| LDL cholesterol (mg/dL) | 134.7±29.0 | 105.9±38.6 | <0.001 |
| HDL cholesterol (mg/dL) | 58.0±15.8 | 50.0±13.7 | <0.001 |
| ProBNP (pg/mL) | | 3900.3±3785.4 | |
| Hs-CRP (mg/L) | 1.80±2.08 | 3.16±3.47 | <0.001 |

NYHA Fc=the New York Heart Association functional classification, ACEi=angiotensin-converting enzyme inhibitor, ARB=angiotensin II receptor blocker, eGFR=estimated glomerular filtration rate, LDL=low-density lipoprotein, HDL= high-density lipoprotein, ProBNP=pro-B-type natriuretic peptide, Hs-CRP=high sensitivity C-reactive protein

**Table 3**. Multiple linear regression analysis for plasma GPNMB

| Independent Variable | Regression Coefficient | SE of Regression Coefficient | P value |
|---|---|---|---|
| Heart failure, yes/no | -10.699 | 1.123 | <0.001 |
| Age, years | 0.149 | 0.105 | 0.155 |
| Diabetes mellitus, yes/no | -0.793 | 2.042 | 0.698 |
| Hypertension, yes/no | 1.021 | 1.420 | 0.456 |
| Coronary artery disease, yes/no | -0.134 | 1.918 | 0.944 |
| Body mass index, $kg/m^2$ | 0.225 | 0.119 | 0.060 |
| eGFR, mL/min/1.73 $m^2$ | -0.044 | 0.035 | 0.204 |
| LDL-c, mg/dL | 0.026 | 0.017 | 0.126 |

eGFR: estimated glomerular filtration rate; LDL-c: low density lipoprotein cholesterol. $R^2$ = 0.211.

**Figure 1.**

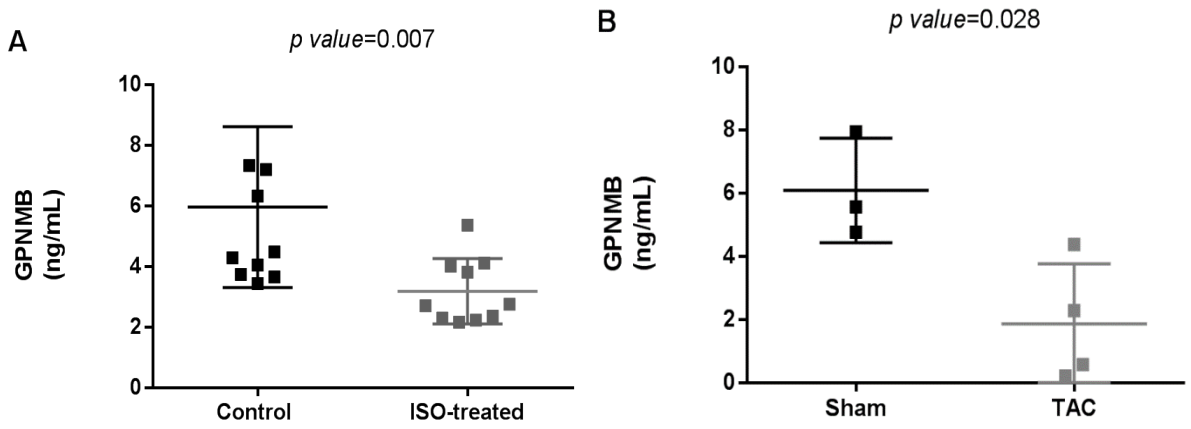
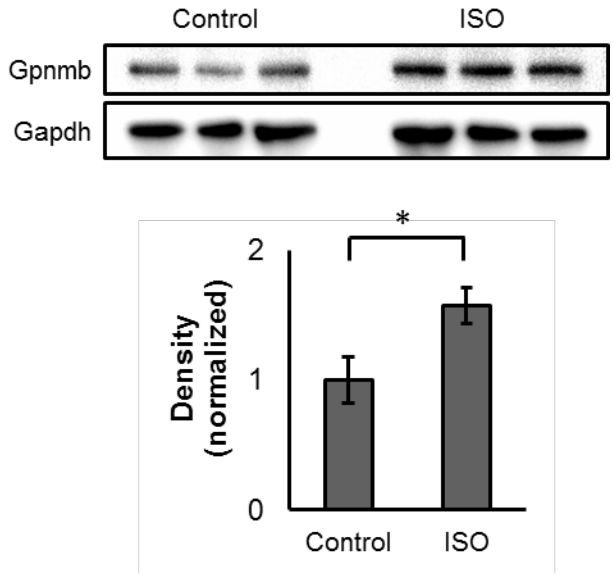

A
p value=0.007
GPNMB (ng/mL)
10
8
6
4
2
0
Control
ISO-treated
B
p value=0.028
GPNMB (ng/mL)
10
8
6
4
2
0
Sham
TAC

**Figure 2.**




Isoproterenol (ISO) treatment continuous infusion for 3 weeks (normalized density in a.u., arbitrary units). Gapdh used as loading control. n=3 per group, *$p<0.05$ for student's t-test.

**Figure 3.**

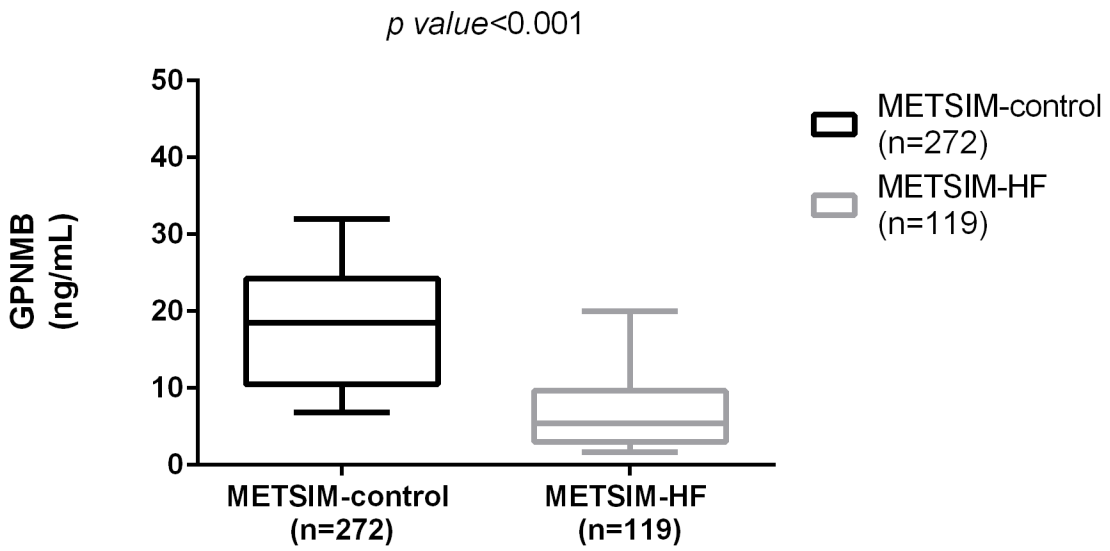
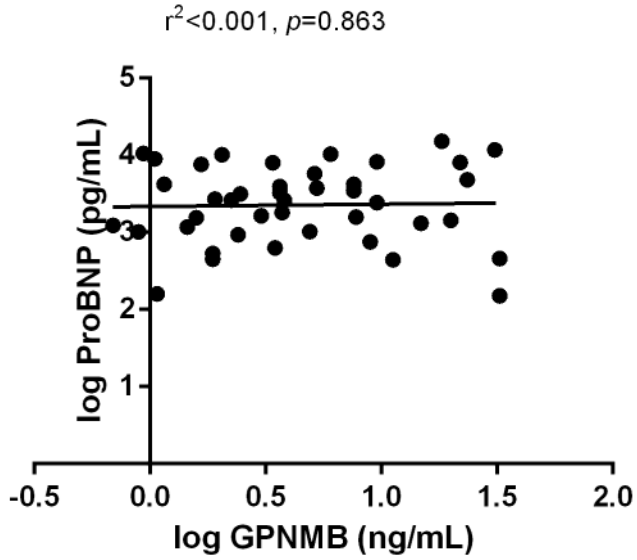
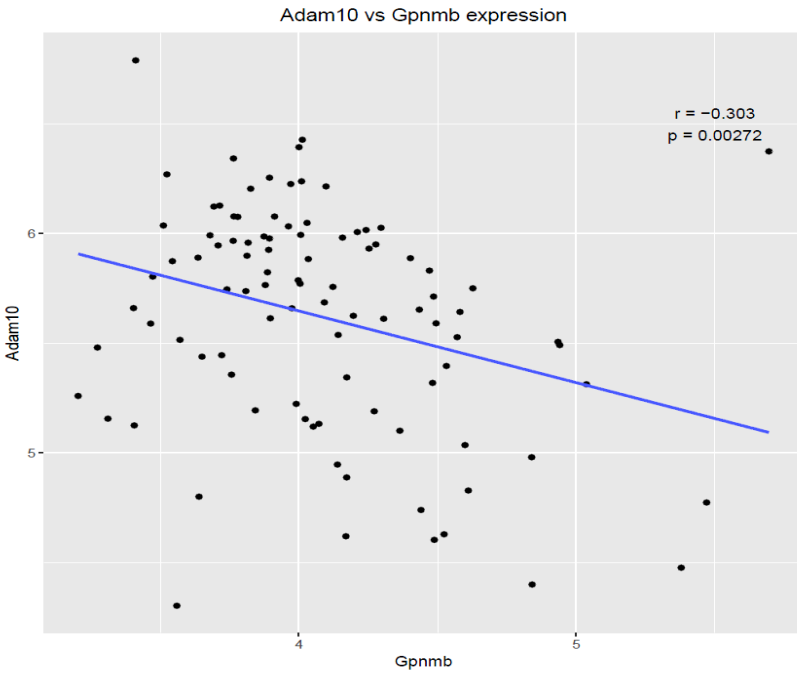

p value<0.001
GPNMB (ng/mL)
50
40
30
20
10
0
METSIM-control (n=272)
METSIM-HF (n=119)
METSIM-control (n=272)
METSIM-HF (n=119)

**Figure 4.**



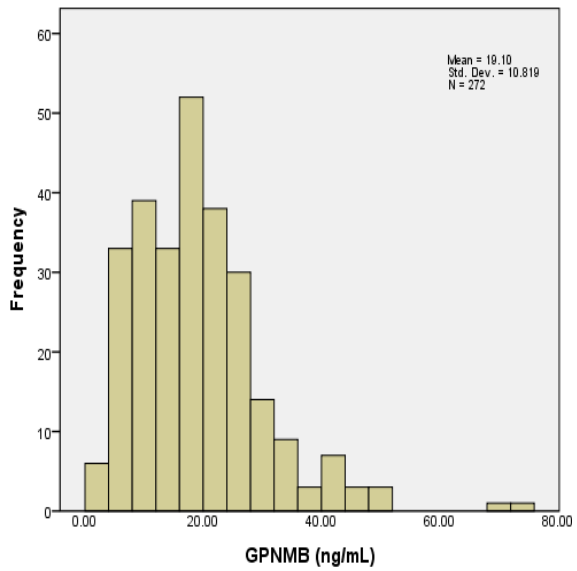

The total number of subjects with ProBNP and GPNMB levels was n=48

**Supplementary figure 1.** Negative correlation between GPNMB and ADAM10 in cardiac transcriptome data in HF-HMDP.

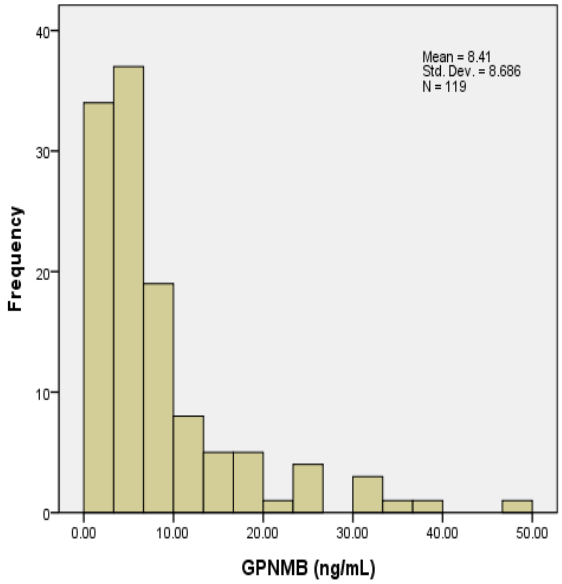


**Supplementary figure 2.** The distribution of plasma GPNMB in METSIM study. (A) control group(n=272) (B) heart failure group(n=119).

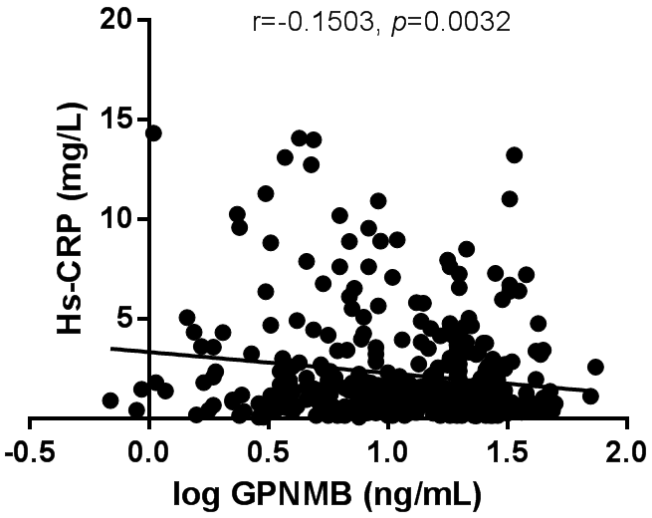
A




B



**Supplementary figure 3.** Negative association with log GPNMB and high sensitivity C-reactive protein in METSIM study (n=391).



*Chapter 5: Exome Sequencing in Rare Cardiovascular Disease*

**Exome Sequencing: Background**

In the era of precision medicine, DNA sequencing is of major importance and viewed to be one of the pillars of personalized clinical care. The Human Genome Project (HGP) was established using an older method of DNA sequencing called Sanger sequencing.[1] Sanger sequencing is a method of DNA sequencing developed by Frederick Sanger in 1977, based on the selective incorporation of chain-terminating dideoxynucleotides by DNA polymerase during in vitro DNA replication.[1] It was the most widely used sequencing method for approximately 39 years. Eventually, newly developed technologies replaced the traditional methods for whole-genome sequencing (WGS) and whole-exome sequencing (WES) with methods that revolutionized sequencing technologies, collectively called next-generation sequencing (NGS).[2-5] The advent of next generation sequencing has changed the way we think about improving human health and is a significant driver in the movement towards personalized medicine; although the expense, ethical issues related to produced genetic data and the need for user-friendly software in the analysis of the raw sequences remain to be addressed.

NGS has improved our understanding of the genetic pathology of disease. Ng et al. published the first report of selectively sequencing of whole exome in 2009. They reported parallel sequencing of the exomes of 12 humans, including 8 individuals previously characterized by HapMap and the Human Genome Structural Variation Project as well as 4 unrelated individuals affected by a rare dominantly inherited disorder called Freeman-Sheldon syndrome.[6] Freeman-Sheldon syndrome (FSS) is caused by mutations in the *MYH3* gene (FSS; MIM193700). Rare and common variants were identified and 13,347 variants appeared to be novel. Subsequent filtering of these variants against dbSNO (v129) or those found in the

HapMap samples defined *MYH3* as the disease-causing gene in FSS patients.[6] This is a prime example of the way WES can help define causal variants of disease. Many studies like this one have been performed to uncover the genetic basis of disease, which could eventually lead to better treatment and management options for patients.

Cardiovascular disease (CVD) is a major cause of death in the United States. From a geneticist's standpoint, there are two forms of CVD: the rare form that is monogenic (Mendelian: includes structural cardiomyopathies, channelopathies and familial dyslipidemias) and the common form that often has polygenic and multifactorial causes. Defining which genes are causal in more rare or Mendelian forms of CVD provides the groundwork for the discovery of potential therapies.

**Exome Sequencing at UCLA**

A group of investigators at UCLA received a grant from the Cardiovascular Discovery Fund to perform WES on patients and their families of interest. In the case of rare disease, WES is the most efficient way to identify genetic variants in the entire coding genome as opposed to SNP arrays that can only detect shared genetic variants common to many individuals. In the initial stages of this investigation, the group has exome sequenced 10 patients with varying and rare disorders of the heart. These investigators are taking a phenotype first approach. Instead of genotyping or sequencing all comers, they focus on a particular disease state of interest. There are some singular patients who underwent WES, but ideally multiple family members who share a phenotype of interest undergo WES in hopes of finding genetic variants that may be causal.

**Case One: Background**

The first case I studied was an unfortunate woman of 43 years of age who is now deceased. I was unable to interview her directly due to her mental status during her hospitalization, but I was able to gather phenotypic information from her family members and prior medical records.

As mentioned, the patient was a 43-year-old female with a past cardiac history that was significant for syncope, which was thought to be vasovagal in nature as a result of dehydration before exercise, and frequent premature ventricular contractions (PVCs) that were evaluated several years prior to her most recent presentation to the hospital. Her other non-cardiac medical history included hyponatremia, which was treated with high-salt diet, Raynaud's phenomenon, eczema and self-reported muscle stiffness and incoordination for over 10 years. Immediately prior to her most recent presentation, she was seen by an orthopedic specialist who did not find significant neurological deficits on examination and recommended continued physical therapy and additional imaging for aid with diagnosis.

On May 19$^{h}$, 2016, the patient presented to the hospital after she collapsed during a spin class at the gym. The patient exercises daily and had taken this particular spin class many times before. Around 6:55 pm, the patient told her friend next to her that she didn't feel well. She collapsed during the spin class shortly after that. Bystander CPR was started and EMS arrived around 7:20pm. When the paramedics arrived, the patient was in PEA arrest. An airway was placed and one round of Epinephrine was given with return of spontaneous circulation (ROSC). According to paramedic report, the patient was pulseless again during the ride to the hospital and CPR was performed during transport. After arrival to the Emergency Department, she was noted to be pulseless with a rhythm of ventricular fibrillation. 3 defibrillator shocks were delivered

along with 3 rounds of CPR and 1 one round of epinephrine. Her temporary airway was exchanged for a formal endotracheal tube. A central line was placed for easier venous access and a Levophed drip at 0.5mg/hour was started. She was subsequently transferred to the Medical Intensive Care Unit (MICU).

Later that evening, a head CT was performed and revealed evidence of hypoxemic brain injury. The patient's neurological exam showed further evidence of poor prognosis with complete unresponsiveness and myoclonic jerking. Keppra was started for seizure prophylaxis. Versed and Propofol were also started, but only after initial neurological examination was performed.

The patient's records had some helpful information regarding her past cardiac history; however nothing that pointed to a clear diagnosis or etiology of her cardiac arrest. In 2008, a doctor evaluated the patient for frequent palpitations. A 24-hour Holter showed frequent PVCs; however the percentage of PVC burden was not noted. Also in 2008, an exercise treadmill test (ETT) and a stress echocardiogram were performed and neither of those examinations showed any evidence of ischemia. She was noted to have a normal ejection fraction on echocardiogram (55-60%). In addition, also in 2008, a coronary computed tomography with angiogram (CTA) revealed normal coronary arteries without any significant calcification. She was noted to have a right-dominant system. At that time, her heart did not show any morphological signs of arrythmogenic right ventricular dyplasia (ARVD).

Her health remained relatively stable until 2014. In 2014, she was admitted to a different hospital in Santa Monica, CA for one day for syncope during exercise. According to the patient, she felt lightheaded and also experienced palpitations prior to loss of consciousness for several

minutes during exercise. She reported that she had been very dehydrated and had not eaten anything prior to exercise that day. During her admission, serial troponin levels and electrocardiograms (EKGs) were unremarkable. An echocardiogram was also unremarkable and revealed an ejection fraction of 55-60%. Her syncope was thought to be vasovagal in etiology due to severe dehydration.

Of note, the patient was a lifelong non-smoker. She drank around 15 glasses of wine per week. She was sexually active with her husband only and used condoms for birth control. She was not on any medications prior to her cardiac arrest.

The patient's baseline EKG (last performed in the UCLA system on 7/17/2008) revealed sinus bradycardia with a heart rate (HR) of 45 beats per minute (bpm), right axis deviation (RAD) and T wave inversions (TWI) inferiorly and laterally. On 5/19/16, her EKG around 19:35 revealed normal sinus rhythm (NSR) with premature atrial contractions (PACs), RAD and TWI inferiorly and laterally with non-specific ST changes anteriorly that did not meet ST elevation criteria. Her EKGs slowly evolved over the next few days to reveal sinus bradycardia and eventually AV conduction block and isorhythmic AV dissociation. The patient's echocardiogram on 5/20/16 revealed normal left ventricular (LV) size, normal wall thickness and severely reduced global systolic function with an ejection fraction of 25-30%. There was mild LV diastolic dysfunction. (Grade I impaired relaxation). The right ventricle was normal in size and mildly reduced in systolic function. There was mid tricuspid regurgitation and moderately elevated pulmonary artery pressure at 42 mmHg.

Upon admission, the patient was placed on the hypothermia protocol to 33F. She was also under continuous electroencephalogram (EEG) monitoring. The patient was noted to have some

myoclonus and epileptiform discharges. After rewarming, her myoclonus improved, however she was noted to have bilateral blown pupils, and the following day (5/23/16) she was noted to have completely absent brain stem reflexes with flat activity on EEG. Initial head CT on admission showed diffuse loss of gray-white differentiation and hypodensity in the basal ganglia concerning for hypoxic ischemic injury. CT head on 5/22/16 showed marked interval progression of changes related to diffuse hypoxic ischemic brain injury. Brain magnetic resonance imaging (MRI) on 5/24/16 showed extensive diffuse hypoxic ischemic injury, with diffuse brain edema, compression of brainstem, herniation of intracranial contents, and absence of vascular flow voids, which were findings in keeping with brain death. Unfortunately, the patient was declared brain dead by two attending physicians at 1700 on 5/24/16.

The etiologies of her PEA and VF arrest were never determined, but an arrythmogenic disorder was highly suspect. The differential diagnosis included long QT syndromes, idiopathic ventricular tachycardia (VT) and Brugada syndrome. There was low suspicion for acute coronary syndrome as an etiology for her cardiac arrest given lack of risk factors and clean coronaries on CTA in 2008. Patient had a known history of hyponatremia and per some observers in her spin class she was noted to be seizing, so hyponatremia was also considered a possible cause of her arrest, although her sodium on admission was 127. Given the lack of answers and no clear cause for her sudden death, the patient underwent autopsy as a coroner's case.

**Differential Diagnosis for Sudden Cardiac Death**

Structural heart disease or coronary artery disease and resultant ischemia are by far the most common causes of sudden cardiac death.[7] Once obvious structural and/or coronary disease

has been ruled out in the cardiac arrest survivor, the differential diagnosis shifts to include primary electrical disease or latent structural causes.

The conditions that are on the list of differential diagnoses are those causing abnormalities in cardiac depolarization or repolarization, usually due to inherited, drug-, metabolic-, or electrolyte-induced ion channel dysfunction. These disorders have been termed channelopathies and they represent primary electrical diseases. Another category of diseases included in the differential are subclinical structural disease such as myocarditis, coronary vasospasm, arrythmogenic right ventricular cardiomyopathy (ARVC), and sarcoidosis.[8] At times the structural abnormalities associated with these diseases can be easily detected with advanced imaging; however sometimes these abnormalities may be subtle or even undetectable with standard testing and even advanced imaging early in the disease course, requiring a high index of suspicion to discern.

Survivors of cardiac arrest without overt heart disease require a multidisciplinary team of physicians and specialists, including an electrophysiologist, a heart failure specialist and a geneticist. This team works with survivors of cardiac arrest and their family members to treat and screen appropriate individuals for the disease in question.[9-12] Screening multiple family members of a sudden death survivor is often termed cascade family screening.[10,13] It should be recognized that in some cases, the known causes of cardiac arrest without overt heart disease have been excluded with in-depth testing and a definitive cause of cardiac death in these patients will remain unexplained.[8]

**Investigation of the Sudden Cardiac Arrest Survivor**

Unfortunately, in our particular patient's case, a comprehensive clinical investigation of her cardiac disease was not performed due to her rapidly declining neurological status. She was pronounced brain dead within a few days of admission. However, in an instance when a survivor of cardiac arrest is neurologically intact, the first step to evaluation should be a comprehensive review of family and drug history. Family history should inquire not only about sudden death, but also about events such as drowning, fatal single-vehicle accidents, sudden infant death syndrome and frequent miscarriages; all potential signs of an inherited predisposition to sudden death. In our patient's case, there were no such red flags present in her family or personal history. Baseline electrolyte and metabolic testing was also performed as mentioned above; however did not reveal a direct cause for cardiac channel instability. In normal circumstances in which the patient's prognosis is good, further biochemical, immunologic, and serological testing should be performed to rule out systemic and/or infiltrative cardiac disease such as amyloid, sarcoid and viremia.

**Case One: Methods**

**Sample acquisition and pre-test sample processing**

Once we determined that our sudden cardiac death patient was of interest and clinically appropriate for WES, the patient's family approached for consent to collect a blood sample and extract DNA for sequencing. After the patient's family agreed to testing, a whole blood sample was collected in an EDTA or lavender top tube. Genomic DNA was extracted using standard methods (QIAcube Qiagen, Valencia, CA). DNA quantity was measured.

**Exome capture and sequencing**

Library preparation, sequencing and data analysis were performed by Macrogen Clinical Lab in Rockville, MD. For the generation of standard exome capture libraries, Macrogen Clinical Lab utilized the Agilent SureSelect Target Enrichment protocol for Illumina paired-end sequencing library (v. 2.0.1, May 2010) together with 1 μg input DNA. In all cases, the SureSelect Human All Exon V4 probe set was used. The quantification of DNA and the DNA quality was measured by PicoGreen and Nanodrop. Fragmentation of 1ug of genomic DNA was performed using adaptive focused acoustic technology (AFA; Covaris). The fragmented DNA was repaired, an 'A' was ligated to the 3′ end. Agilent adapters were then ligated to the fragments. Once ligation had been assessed, the adapter ligated product was PCR amplified. The final purified product was then quantified using qPCR according to the qPCR Quantification Protocol Guide and qualified using the Caliper LabChip High Sensitivity DNA (PerkinElmer). For exome capture, 250 ng of the DNA library was mixed with hybridization buffers, blocking mixes, RNase block and 5 μl of SureSelect. All exome capture was performed according to the standard Agilent SureSelect Target Enrichment protocol. Hybridization to the capture baits was conducted at 65°C using heated thermal cycler lid option at 105°C for 24 hours on PCR machine.

The captured DNA was then amplified. The final purified product was then quantified using qPCR according to the qPCR Quantification Protocol Guide and qualified using the TapeStation DNA screentape (Agilent). Sequencing was performed using the HiSeq™ 4000 platform (Illumina,San Diego, USA). Followings are details of sequencing: HiSeq control software: 3.3.52

cBOT: 2.0.29.0, Cluster generation: HiSeq 3000/4000 SR Cluster Kit, Sequencing reagent kit: HiSeq 3000/4000 SBS Kit and Read length: 100bp paired-ends.

**Sequence data analysis**

**Variant filtration and interpretation**

Annotated variants for our patient were deposited into Golden Helix VarSeq 1.4.3 database for filtering. The first step of variant filtration was selecting only the variants with a sequencing depth or read depth greater than 10. Read depth describes the number of times that a given nucleotide in the genome has been read during sequencing. Next, variants were filtered by genotype qualities (GQ) or Phred quality score. A Phred quality score is a measure of the quality of the identification of the nucelobases generated by automated DNA sequencing.[14,15] Variants with GQ > 20 were chosen. Next, variants were filtered by zygosity. The VarSeq algorithm computes zygosity of a sample by using the individual's genotype combined with frequency information available in a variant frequency catalog. The catalogs used in our analysis were ExAC and 1000 genomes. Both heterozygous and homozygous minor variants were chosen. The next step of variant filtering was filtering out common variants with minor allele frequency (MAF) of 1%. MAF was estimated using a combined dataset incorporating all available data from the 1000 genomes project (1Kg), NHLBI exome sequencing project (ESP), NIEHS environmental genome project (EGP) and HapMap, without the distinction of ethnic background. Any variant not observed in any of these datasets was considered "novel" for future analysis. We further filtered the annotated variants by sequence ontology and included those that were the following: disruptive inframe deletion and insertion, frameshift variant, inframe deletion and insertion, missense variant, splice acceptor variant, splice donor variant, splice region variant, stop gained, stop lost, stop retained variant.

**Case One: Results**

No variants were found after WES in this patient. Please see discussion section for further discussion regarding false positive and negative results with exome sequencing.

**Case Two: Background**

The second case studied was that of two sisters, both born with Tetralogy of Fallot (TOF). Sister number one is a 47-year-old female with a congenital diagnosis of TOF. TOF is a congenital heart defect that is present at birth.[16] A pedigree is provided in Figure 1. Important symptoms in a newborn are cyanosis, heart murmur, finger clubbing and early tiring upon breastfeeding.[17] The cause is typically not known, but known risk factors include a mother who uses alcohol during pregnancy, has diabetes, is over the age of 40 or gets rubella during pregnancy. It may also be associated with Down syndrome.[18] Classically, there are four defects: a ventricular septal defect (VSD), pulmonary stenosis (PS), right ventricular hypertrophy (RVH), and an overriding aorta, which allows blood from both ventricles to enter the aorta.[16] TOF is typically treated by open heart surgery in the first year of life; however timing of surgery can vary from patient to patient depending on their symptoms and size. The patient underwent a Waterston anastomosis at the age of 2. A Waterston anastomosis is a joining of the ascending aorta and right pulmonary artery in an effort to increase pulmonary blood flow in infants born with severe obstructive lesions of the right side of the heart.[19] The patient then underwent intracardiac repair of her TOF at 4 years of age.. The exact details of the patient's surgical repair are unknown; however in general, surgical repair in TOF aims to improve blood flow to the lungs and correct the flow of arterial and venous blood. Typically, the pulmonary valve is widened or replaced, and the passage from the right ventricle to the pulmonary artery is enlarged, improving blood flow to the lungs. Additionally, the VSD is closed with a patch. The patch stops

arterial and venous blood from mixing between the left and right ventricles. Unfortunately, although intracardiac repair can be curative of the original congenital defect, it comes with the possibility of multiple complications down the line. For example, the patient has longstanding and severe pulmonary regurgitation (PR), which results in non-sustained ventricular tachycardia, which has been noted on prior stress test in 1994. Due to this complication, the patient underwent re-operation with pulmonary valve replacement with a 31-mm Hancock procine bioprosthesis in 1994. These patients also suffer from atrial arrhythmias of the heart. In April of 2015, the patient suffered exertional dizziness and was found to have continuous atrial flutter at rates of 34 to 273 beats per minute (bpm). She subsequently underwent successful ablation of her atrial flutter on 5/15/15. Due to persistent PR, even after her re-operation in 1994, the patient underwent transcatheter 23mm Sapien 3 valve placement in the pulmonary position, with resolution of PR. At this time, she also developed atrial flutter, which was again ablated successfully on 5/19/16.

The patient's sister is a 45-year-old woman with congenital TOF. She underwent intracardiac repair at the age of 4. The details of her operation are also unknown; however are likely similar to those described above. She also has longstanding and severe pulmonary regurgitation; however she has not needed intervention or re-operation, as she is asymptomatic. She continues to do excellent from a cardiac standpoint.

**Case Two: Methods**

**Sample acquisition and pre-test sample processing**

Once we determined that the probands, the two sisters, were of interest and clinically appropriate for WES, the patients as well as their parents were approached for consent to collect

a blood sample and extract DNA for sequencing. After the patients and their parents agreed to testing, a whole blood sample was collected in an EDTA or lavender top tube. Genomic DNA was extracted using standard methods (QIAcube Qiagen, Valencia, CA). DNA quantity was measured.

**Exome capture and sequencing**

Library preparation, sequencing and data analysis were performed by Macrogen Clinical Lab in Rockville, MD. Please see the section under Exome capture and sequencing for Case One for methods.

**Sequence data analysis**

**Variant filtration and interpretation**

Annotated variants for our patients and their parents were deposited into Golden Helix VarSeq 1.4.3 database for filtering. The first step of variant filtration was selecting only the variants with a sequencing depth or read depth greater than 10. Read depth describes the number of times that a given nucleotide in the genome has been read during sequencing. Next, variants were filtered by genotype qualities (GQ) or Phred quality score. A Phred quality score is a measure of the quality of the identification of the nucelobases generated by automated DNA sequencing.[14,15] Variants with GQ > 20 were chosen. Next, variants were filtered by zygosity. The VarSeq algorithm computes zygosity of a sample by using the individual's genotype combined with frequency information available in a variant frequency catalog. The catalogs used in our analysis were ExAC and 1000 genomes. Both heterozygous and homozygous minor variants were chosen. The next step of variant filtering was filtering out common variants with

minor allele frequency (MAF) of 1%. MAF was estimated using a combined dataset incorporating all available data from the 1000 genomes project (1Kg), NHLBI exome sequencing project (ESP), NIEHS environmental genome project (EGP) and HapMap, without the distinction of ethnic background. Any variant not observed in any of these datasets was considered "novel" for future analysis. We further filtered the annotated variants by sequence ontology and included those that were the following: disruptive inframe deletion and insertion, frameshift variant, inframe deletion and insertion, missense variant, splice acceptor variant, splice donor variant, splice region variant, stop gained, stop lost, stop retained variant.

There was one high quality *de novo* variant found in both probands and inherited from the mother. The gene name is DSPP.

**Case Two: Results**

TOF occurs in about 1 in 2,000 newborns. Males and females are affected equally.[16] Untreated, TOF rapidly results in progressive right ventricular hypertrophy due to the increased resistance caused by narrowing of the pulmonary trunk. This progresses to heart failure and dilated cardiomyopathy. Mortality rate depends on the severity of the condition. If left untreated, TOF carries a 35% mortality rate in the first year of life, and a 50% mortality rate in the first three years of life.[16]

The cause is thought to be a combination of environmental and genetic factors. Embryologic studies show that TOF is a result of anterior malalignment of the aorticopulmonary septum, resulting in the hallmark clinical combination of a VSD, pulmonary stenosis, and

overriding aorta. Due to this anatomy and the resulting hemodynamics, RVH eventually develops.

It is known to be associated with chromosome 22 deletions and DiGeorge syndrome. Specific genetic associations include: JAG1[20], NKX2-5[21], ZFPM2[22], and VEGF[23]; however no causal mutation or variant has been identified thus far.

The DSPP or Dentin sialophosphoprotein gene codes for a secreted protein, also called DSPP. The DSPP gene encodes two principal proteins of the dentin extracellular matrix of the tooth. The preproprotein is secreted by odontoblasts and cleaved into dentin sialoprotein and dentin phosphoprotein. Dentin phosphoprotein is thought to be involved in the biomineralization process of dentin. Mutations in this gene have been associated with dentinogenesis imperfecta-1; a disorder of tooth development that most often causes the teeth to be discolored and translucent. In some individuals, dentinogenesis imperfecta occurs in combination with an autosomal dominant form of deafness. Allelic differences due to repeat polymorphisms have been found for this gene.[24]

After thorough literature search, there does not seem to be any link between DSPP and congenital cardiac disease; however no studies have been performed to rule out this possibility. However, it is noted that DSPP is synthesized in both the mesenchyme and epithelium at varying stages of tooth development. DSPP is also expressed in ureteric bud branches of embryonic metanephric kidney and alveolar epithelial buds of the developing lung. However, DSPP is not expressed in the heart; therefore a link between this gene and congenital cardiac disease is unlikely.[25-29]

**Discussion**

The rapid development of genomic sequencing technologies such as massively parallel next-generation sequencing has decreased cost, improved efficiency, and increased the use of genetic testing in the clinical setting.[30-31] Exome sequencing is being increasingly utilized in the diagnosis of rare disease and disorders of unknown genetic etiology.[32] As costs continue to decline, it is likely the use of whole genome-sequencing will increase.[33] The role of genetic testing as a tool to investigate cardiovascular disease has had increased focus in recent years.[34]

Studies report varying diagnostic exome sequencing success rates, ranging from 22.8 % to 50 %.[35,36] Although there is increasing evidence for the utility of exome sequencing, the majority of the sequencing performed in patients presenting with a presumed Mendelian genetic trait does not produce novel variants. Exome sequencing raises many challenges both from a technical and diagnostic standpoint,[37,38] which can contribute to the lack of answers from testing.

Compared with targeted gene sequencing and comprehensive panels specific for disease, which are currently available,[39] there are important differences to consider with exome sequencing. Importantly, only 1-2 % of the human genome contains protein–coding sequences; thus, these regions must undergo an exon-targeting "*capture*" process before being sequenced.[32,33,39,40] Traditional sequencing methods for individual genes or panels do not require this step. Target capturing can sometimes miss exons. As a result, many genes with clinical relevance are not well-covered by standard exome sequencing. Several genes are completely absent from exome sequences, and many more genes are only partially covered; partly due to lack of targeting by exome platforms and high guanine/cytosine content. This may be a large contributing factor to the absence of any novel variants in our sudden cardiac death case. Figure

2 outlines a basic schematic of a hypothetical analysis of patients carrying pathogenic sudden cardiac death variants via exome sequencing and the steps that potentially can lead to missed variants. Table 1 displays a current list of 103 genes associated with sudden cardiac death.

By design, exome sequencing primarily targets genomic regions that code for protein; therefore variants located in untranslated regions (UTRs), intronic, promoter, and intergenic regulatory regions are missed. Although it is often difficult to interpret novel variants in these regions, any disease-causing variants that may lie in these regions cannot be detected by standard exome sequencing platforms. In addition, exome sequencing may miss structural variants as most platforms limit variant identification to single nucleotide variants (SNVs) and small indels thereby incompletely interrogating genes for structural variation.

Another consideration is determining the "*coverage*" of captured regions. Coverage or depth of sequencing refers to how many times a nucleotide meeting criteria for being a high-quality base is represented in a random collection of raw sequences.[41] This helps differentiate sequencing errors from a true sequence variant. The higher the coverage or depth of sequencing, the more likely that the captured base is accurate and not a false read due to technical errors. For example, a captured base with 20x coverage means that base is represented at least 20 times at that position in multiple raw sequences. Of note, "coverage", in addition to meaning *depth* of sequencing, may also refer to the general *proportion* of bases covered in genomic sequence at a specific depth.[42] For example, an exon with 90 % coverage at 20x means 90 % of the bases in that exon are represented at least 20 times on multiple raw sequences. Although sequencing depth or coverage is a quality control measure that is commonly used; average depth does not necessarily mean better sensitivity to detect variants, as some regions of a gene may have

excessive coverage while others are completely missed. In addition, it is important to account for variation in coverage for a particular gene from sample to sample. If this variance is great, there is a chance that a gene with good average coverage across the gene may not be well covered and represented in a patient's sample. Disease-causing variants can be missed if the coverage for a particular gene is not sufficient across the whole gene consistently in every sample.

There can also be limitations in variant identification due to missing information in the human reference genome. Robust gene representation in the reference genome is critical to analyzing whole exome data; however most reference genomes are missing regions of the genome, including whole genes. The reference genome is updated often and there are over 650 medically important genes that may be affected by an assembly update. For example, *SHANK2* is a gene associated with autism spectrum disorders for which clinical testing is currently available. Two exons of this gene were not included in the 2014 human reference sequence, GRCh37. This means that sequence reads from this region were not included in exome alignments and therefore variants in these exons were not detected through exome sequencing.

**Future Directions**

At UCLA, our hope is to utilize whole exome sequencing in all incidences of rare cardiovascular disease without an already established genetic cause. In many cases, even those diseases with a known genetic cause may be worth investigating.

The sequencing of the human genome has the promise of a genetic revolution for clinical medicine, especially for rare disease. Many already envision a time in the near future when each

patient's genome will be sequenced and available to provide guidance in a personalized approach to health maintenance, early diagnosis of disease prevention and treatment.

The potential benefits of whole exome sequencing could be substantial for a health-care environment that is moving away from the one-size-fits-all paradigm of evidence-based medicine.[43] Some believe that whole exome sequencing for every patient will increasingly emphasize a more patient-centered, personalized and preventative approach to wellness.

As of now, we do not plan to utilize whole exome sequencing in healthy individuals; however physician groups around the country, including some large hospital systems, have already started to pilot programs that sequence healthy individuals on a voluntary basis. One such project is called the MedSeq Project, two randomized trials of integrating whole genome sequencing into clinical medicine.[44] This pair of randomized trials compares the use of whole genome sequencing to standard of care in two clinical settings; a cardiomyopathy clinic in which disease-specific genomic medicine can be utilized and in a primary care clinic in which general genomic medicine can be utilized. The investigators recruited both cardiologists and primary care physicians and approximately 200 of their patients. The investigators continue to recruit participants. The current patient participants in both the cardiology and primary care trials were randomly assigned to receive a family history assessment with or without whole genome sequencing. All physicians were non-geneticists. Physicians were provided educational curriculum regarding whole genome sequencing and a hotline to genetics professionals for guidance in interpreting and managing their patients' genome reports. Using surveys, semi-structured interviews and review of clinical data, the investigators hope to gauge the impact of emerging sequencing technologies on patient care. The results from the cardiomyopathy portion

of the trial are not yet published or available. However, initial results in the primary care portion of the trial lead us to believe that there is no clear indication for sequencing in primary care patients and all findings, not surprisingly, were unanticipated (personal communication, J. Vassy). The variants reported provided information about monogenic disease, recessive carrier status, pharmacogenomics traits and polygenic risk estimates in the "healthy" primary care population. One in five patients had a monogenic disease variant. The investigators and physicians participating in the study only reported variants that were likely pathogenic or pathogenic to study patients. The information reported led to increased testing including laboratory, imaging and cardiac testing (EKG, echocardiogram and cardiac MRI), but only increased cardiac testing reached statistical significance (personal communication, J. Vassy). Only 2 out of over 100 patients that have been studied thus far showed phenotypic evidence of variants that were found through whole genomic sequencing one with fundus albipunctatus (white spots on fundus) and another with variegate porphyria (rash). When following this group of patients for 6 months, there was a consistent trend of greater number of laboratory testing, imaging and cardiac testing, but again, only cardiac testing reached statistical significance.

There are many investigators around the United States and the world performing similar studies to determine whether genomic sequencing of healthy adults provides any useful information. Initial results reveal that genomic sequencing in generally healthy adults will reveal unexpected findings and these findings may result in increased clinical action. These results will vary greatly depending on variant filtering, interpretation and choice of reporting pipeline.

Although UCLA has not started to use whole genome sequencing on healthy patients, it may be the way of the future and I believe is worth exploring. The important question that

remains is if the genetics workforce is sufficient to handle all of the information that may result from sequencing healthy individuals. Strengthening our genetic data storage and management systems as well as proper training of non-geneticist physicians and staff is essential for the future of personalized and precision medicine.

**Figure 1 Pedigree for Case 2 (Proband 1 and Proband 2 shown)**

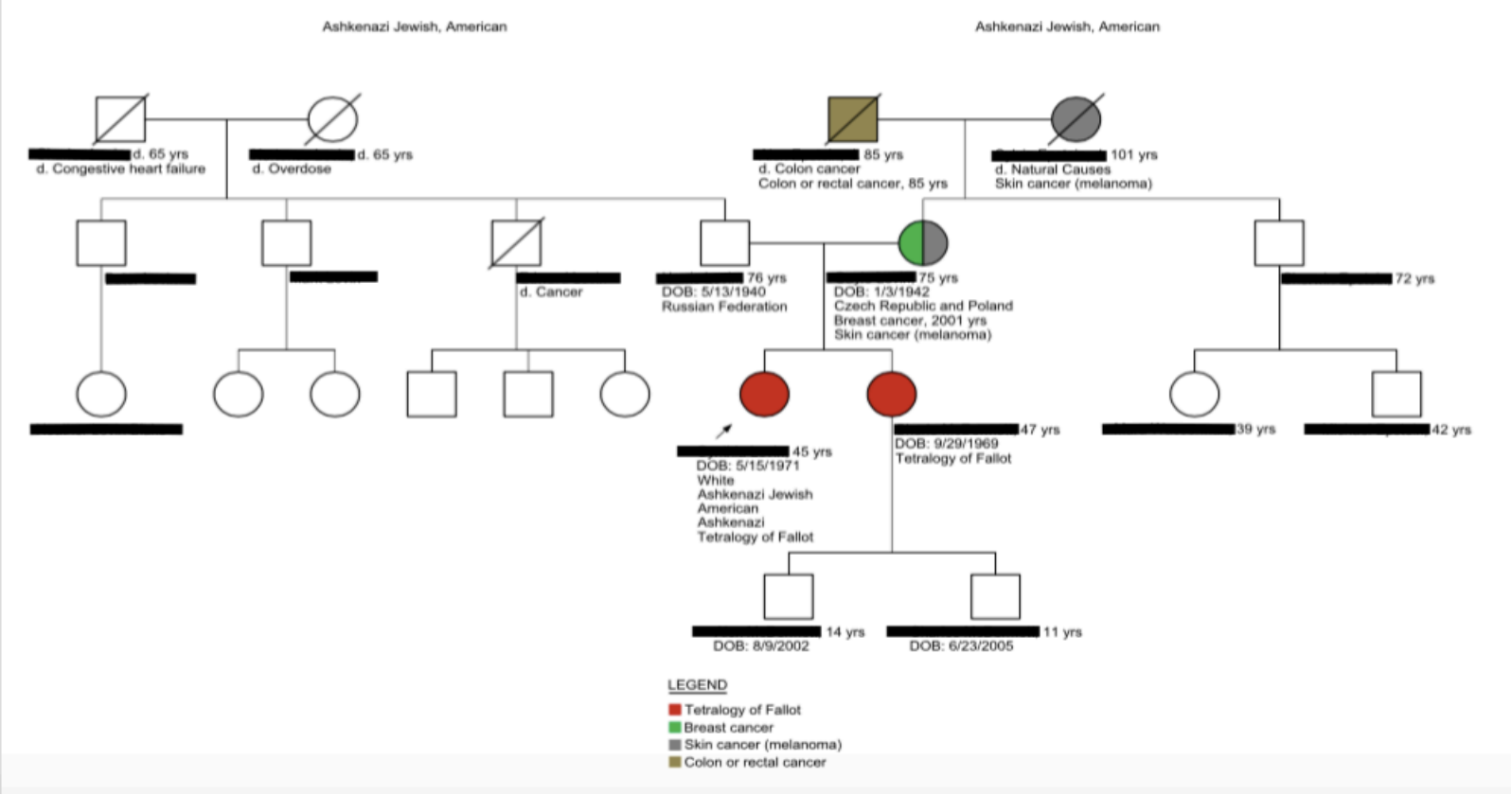
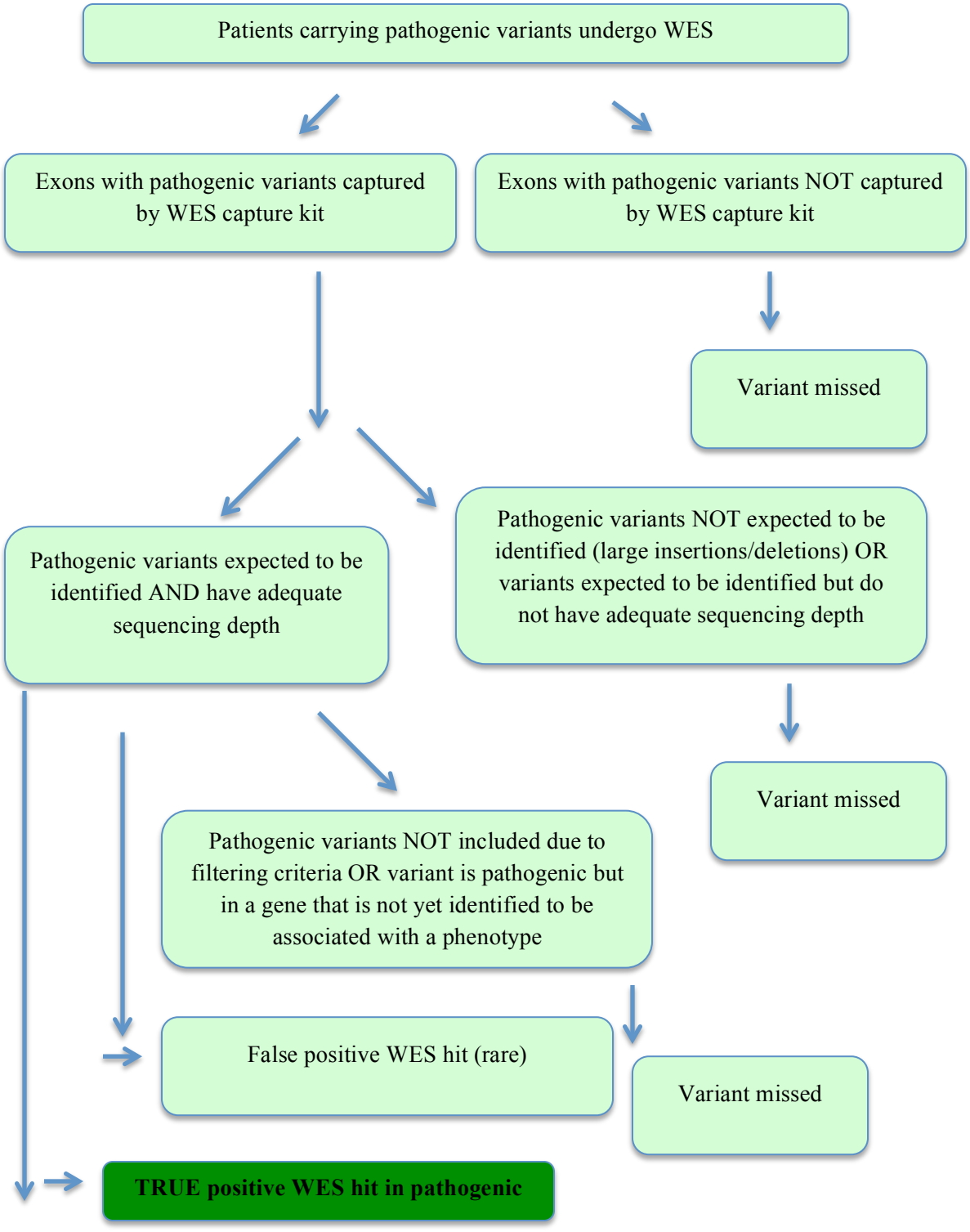


**Figure 2 Schematic of patient with pathogenic variant undergoing WES**



**References**


1. Sanger F; Coulson AR (May 1975). "A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase". *J. Mol. Biol*. **94** (3): 441–8.
2. P. Mayer et al., presented at the Fifth International Automation in Mapping and DNA Sequencing Conference, St. Louis, MO, USA (October 7–10, 1998). DNA colony massively parallel sequencing ams98 presentation "A very large scale, high throughput and low cost DNA sequencing method based on a new 2-dimensional DNA auto-patterning process"
3. Karl V. Voelkerding; Shale A. Dames & Jacob D. Durtschi (2009). "Next-Generation Sequencing: From Basic Research to Diagnostics". *Clinical Chemistry*. **55** (4): 641–658.
4. Matthew W. Anderson; Iris Schrijver (2010). "Next Generation DNA Sequencing and the Future of Genomic Medicine,". *Genes*. **1** (1): 38–69.
5. Tracy Tucker; Marco Marra & Jan M. Friedman (Aug 2009). "Massively Parallel Sequencing The Next Big Thing in Genetic Medicine". *Am J Hum Genet*. **85** (2): 142–54.
6. Ng S, Turner EH, Robertson PD et al. Targeted capture and massively parallel sequencing of 12 human exomes. Nature 461, 272-276 (2009).
7. Huikuri HV, Makikallio TH, Raatikainen MJ, Perkiomaki J, Castellanos A, Myerburg RJ. Prediction of sudden cardiac death: appraisal of the studies and methods assessing the risk of sudden arrhythmic death. *Circulation*. 2003;108:110–115.
8. Krahn AD, Healey JS, Chauhan V, Birnie DH, Simpson CS, Champagne J, Gardner M, Sanatani S, Exner DV, Klein GJ, Yee R, Skanes AC, Gula LJ, Gollob MH. Systematic assessment of patients with unexplained cardiac arrest: Cardiac Arrest Survivors With

Preserved Ejection Fraction Registry (CASPER). *Circulation*. 2009;120:278–285.

9. Behr E, Wood DA, Wright M, Syrris P, Sheppard MN, Casey A, Davies MJ, McKenna W. Cardiological assessment of first-degree relatives in sudden arrhythmic death syndrome. *Lancet*. 2003;362:1457–1459.
10. Hofman N, Tan HL, Alders M, van Langen IM, Wilde AA. Active cascade screening in primary inherited arrhythmia syndromes: does it lead to prophylactic treatment? *J Am Coll Cardiol*. 2010;55:2570 –2576.
11. van der Werf C, Hofman N, Tan HL, van Dessel PF, Alders M, van der Wal AC, van Langen IM, Wilde AA. Diagnostic yield in sudden unex- plained death and aborted cardiac arrest in the young: the experience of a tertiary referral center in the Netherlands. *Heart Rhythm*. 2010;7: 1383–1389.
12. van der Werf C, van Langen IM, Wilde AA. Sudden death in the young: what do we know about it and how to prevent? *Circ Arrhythm Electro- physiol*. 2010;3:96–104.
13. Schwartz PJ. Cascades or waterfalls, the cataracts of genetic screening are being opened on clinical cardiology. *J Am Coll Cardiol*. 2010;55: 2577–2579.
14. Ewing B; Hillier L; Wendl MC; Green P. (1998). "Base-calling of automated sequencer traces using phred. I. Accuracy assessment". *Genome research*. **8** (3): 175–185. doi:10.1101/gr.8.3.175. PMID 9521921.
15. Ewing B, Green P (1998). "Base-calling of automated sequencer traces using phred. II. Error probabilities". *Genome research*. **8** (3): 186–194. doi:10.1101/gr.8.3.186. PMID 9521922.
16. "What Is Tetralogy of Fallot?". *NHLBI*. 1 July 2011. Retrieved 2 October 2016.

17. "What Are the Signs and Symptoms of Tetralogy of Fallot?". *NHLBI*. 1 July 2011. Retrieved 2 October 2016.
18. "What Causes Tetralogy of Fallot?". *NHLBI*. 1 July 2011. Retrieved 2 October 2016.
19. Pickering D, Trusler GA, Lipton I, et al Waterston anastomosis Thorax 1971;26:457-459.
20. Eldadah ZA, Hamosh A, Biery NJ, et al. (January 2001). "Familial Tetralogy of Fallot caused by mutation in the jagged1 gene". *Hum. Mol. Genet*. **10** (2): 163–9. doi:10.1093/hmg/10.2.163. PMID 11152664.
21. Goldmuntz E, Geiger E, Benson DW (November 2001). "NKX2.5 mutations in patients with tetralogy of Fallot". *Circulation*. **104** (21): 2565–8. doi:10.1161/hc4601.098427. PMID 11714651.
22. Pizzuti A, Sarkozy A, Newton AL, et al. (November 2003). "Mutations of ZFPM2/FOG2 gene in sporadic cases of tetralogy of Fallot". *Hum. Mutat*. **22** (5): 372–7. doi:10.1002/humu.10261. PMID 14517948.
23. Lambrechts D, Devriendt K, Driscoll DA, et al. (June 2005). "Low expression VEGF haplotype increases the risk for tetralogy of Fallot: a family based association study". *J. Med. Genet*. **42** (6): 519–22. doi:10.1136/jmg.2004.026443. PMC 1736071. PMID 15937089.
24. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei D, Astashyn A, Badretdin A, Bao Y, Blinkova O, Brover V, Chetvernin V, Choi J, Cox E, Ermolaeva O, Farrell CM, Goldfarb T, Gupta T, Haft D, Hatcher E, Hlavina W, Joardar VS, Kodali VK, Li W, Maglott D, Masterson P, McGarvey KM, Murphy MR, O'Neill K, Pujar S, Rangwala SH, Rausch D, Riddick LD, Schoch C, Shkeda A, Storz SS, Sun H, Thibaud-Nissen F, Tolstoy I, Tully RE,

Vatsan AR, Wallin C, Webb D, Wu W, Landrum MJ, Kimchi A, Tatusova T, DiCuccio M, Kitts P, Murphy TD, Pruitt KD. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 2016 Jan 4;44(D1):D733-45

25. Molecular cloning of a human dentin sialophosphoprotein gene. (PMID: 10706475) Gu K. … Rutherford R.B. Eur. J. Oral Sci. 2000 3 4 22 64
26. Human dentin phosphophoryn nucleotide and amino acid sequence. (PMID: 9879917) Gu K. … Ritchie H.H. Eur. J. Oral Sci. 1998 3 4 22 64
27. Dentin phosphoprotein and dentin sialoprotein are cleavage products expressed from a single transcript coded by a gene on human chromosome 4. Dentin phosphoprotein DNA sequence determination. (PMID: 8995371) MacDougall M. … Gu T.T. J. Biol. Chem. 1997 2 3 22 64
28. A DSPP mutation causing dentinogenesis imperfecta and characterization of the mutational effect. (PMID: 23509818) Lee S.K. … Kim J.W. Biomed Res Int 2013 3 4 64
29. Rough endoplasmic reticulum trafficking errors by different classes of mutant dentin sialophosphoprotein (DSPP) cause dominant negative effects in both dentinogenesis imperfecta and dentin dysplasia by entrapping normal DSPP. (PMID: 22392858) von Marschall Z. … Fisher L.W. J. Bone Miner. Res. 2012 3 4 64
30. Jamal SM, Yu J, Chong JX, Dent KM, Conta JH, Tabor HK, Bamshad MJ. Practices and Policies of Clinical Exome Sequencing Providers: Analysis and Implications. Am J Med Genet A. 2013;161A:n/a-n/a.
31. Interpreting Secondary Cardiac Disease Variants in an Exome Cohort. Circ Cardiovasc Genet. 2013;6:337–46.

32. Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, et al. Exome sequencing as a tool for Mendelian disease gene discovery. Nat. Rev. Genet. 2011;12:745–55. doi: 10.1038/nrg3031.

33. Wang Z, Liu X, Yang B, Gelernter J. The Role and Challenges of Exome Sequencing in Studies of Human Diseases. Front Genet. 2013;4:160.

34. Arndt A, MacRae CA. Genetic testing in cardiovascular diseases. Curr. Opin. Cardiol. 2014;29:235–40.

35. Atwal PS, Brennan M, Cox R, Niaki M, Platt J, Homeyer M, et al. Clinical whole-exome sequencing: are we there yet? Genet. Med. 2014;16:717–9.

36. Clinical application of exome sequencing in undiagnosed genetic conditions. J Med Genet. 2012;49:353–61.

37. Yang Y, Muzny DM, Reid JG, Bainbridge MN, Willis A, Ward PA, et al. Clinical Whole-Exome Sequencing for the Diagnosis of Mendelian Disorders. N Engl J Med. 2013;369:1502–11.

38. Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, Lee C, et al. Targeted capture and massively parallel sequencing of 12 human exomes. Nature. 2009;461:272–6.

39. Krawitz P, Mundlos S. Strategies for exome and genome sequence data analysis in disease-gene discovery projects. Clin. Genet. 2011;80:127–32.

40. Wilde AAM, Behr ER. Genetic testing for inherited cardiac disease. Nat Rev Cardiol. 2013;10:571–83.

41. Raymond C, Raymond C, Aravind L. Initial sequencing and analysis of the human genome. Nature. 2001;409:860–921.

42. Sims D, Sudbery I, Ilott NE, Heger A, Ponting CP. Sequencing depth and coverage: key considerations in genomic analyses. Nat. Rev. Genet. 2014;15:121–32.

43. Goldberger JJ, Buxton AE: Personalized medicine vs guideline-based medicine. JAMA 2013, 309:2559–2560.

44. Vassey JL, Lautenbach DM, McLaughlin HM, et al. The MedSeq Project: a randomized trial of integrating whole genome sequencing into clinical medicine. Trials 2014, 15:85.

*Chapter 6: Conclusions and Future Directions*

**Conclusion and Future Directions**

The central theme of this body of work, with exception to Chapter 5, is utilization of mouse systems genetics, more specifically data derived from the Hybrid Mouse Diversity Panel (HMDP), for translational research pertaining to human disease. Mouse genome wide association studies (GWAS) studies are a powerful tool, but translating the data obtained from these studies to human disease is still a developing process. The translation to human disease can be difficult for several reasons. Specifically, in immune-based disease, although mouse models recapitulating disease such as Lupus or Multiple Sclerosis (MS) exist, most studies utilize only a single inbred mouse strain. However, basal and diseased immune states in humans show vast inter-individual variability. Although diversity in the humoral immune response to the influenza virus in BALB/c mice has been previously reported (Staudt et al, 1983), this diversity does not reach anywhere near the degree of diversity in the human immune response.

One of the most important conclusions made from my dissertation work is that the HMDP is an important and viable source for recapitulation of the human immune system, and human genetic diversity in general. The M1/M2 project (Chapter 2) demonstrates how a spectrum of macrophage phenotypes in many different inbred mouse strains can be used to extract gene signatures of lipopolysaccharide (LPS) responsiveness. Since the analysis of commonly LPS regulated genes in all strains failed, we established a surrogate marker based on *IL-12β and Arg1* gene expression. Compared to published signatures, the resulting gene lists are unique, yet more robustly predictive of many human inflammatory and malignant disorders. This suggests that accounting for immune diversity in a heterogeneous population increases the translatability of mouse data. While the human immune system is shaped by both environmental

factors and genetics, only the latter plays a role in laboratory inbred mice. It remains to be investigated what exact mechanisms account for the inter-strain differences. One possibility is inter-strain differences in the Interleukin-6/Interleukin-10/STAT3 pathway affecting macrophage activation. Future studies on peritoneal and other tissue-resident macrophages in response to various stimuli will more fully characterize macrophage immune diversity in mice and humans. The tool to measure polarization factor developed in our study may be useful in such future studies.

Although not my original intent, the two most impactful projects in this body of work pertain to precision medicine, a field I've been interested in since its introduction into popular media several years ago. A striking finding of the M1/M2 study is that the genetic signatures developed from mouse data are robust enough to predict cancer survival or death from mixed-cell biopsy material in humans. It is then very conceivable that the genetic signatures developed in our study are well suited for predictive tests in personalized medicine, especially with the new generation of cancer treatments focused on manipulation of tumor-associated macrophage polarization (Mantovani et al, 2015). In addition, in contrast to conventional flow cytometry or PCR-based estimations of macrophage polarization in ordinal scale, the developed genetic signature allows a more continuous classification, thus enabling a population-based assessment in higher resolution necessary for a vast range of clinical applications.

Perhaps the most exciting application for the developed genetic signature for macrophage polarization is in prediction of the susceptibility of a particular individual to immune-mediated disease (M1) vs. cancer (M2) at birth. This is a gross oversimplification, because it is well known that macrophage polarization is a spectrum, rather than two distinct states of "fight" or

"heal." However as this tool is developed further over time, it is conceivable that in the near future, we may be able to predict to an exact percentage of risk how susceptible an individual is to atherosclerotic disease, autoimmune disorders such as Lupus or MS, cancer or even response to the common cold. Assuming this information would be obtained at birth, each individual could leave the hospital with a personalized health prevention plan fit to his or her particular susceptibilities in their first few days of life. Our version of this kind of prediction and preventative planning today is the family history portion of a medical interview. A genetic signature for disease prediction could be the ultimate tool in personalized medicine and disease prevention.

The remainder of my dissertation work presents two additional novel approaches to the utilization of data from GWAS studies performed on the HMDP for translation into human disease processes (Chapters 3 and 4). Using data from the heart failure HMDP, my second project utilized a traditional GWAS to candidate gene discovery method to elucidate the mechanisms underlying cardiac remodeling in humans by investigating a gene, CD200, thought to be involved in adverse left ventricular remodeling after myocardial injury. Although the genetic and eQTL data appeared promising, the validation studies are initially negative and could be explored further. My third project utilized the heart failure HMDP data once more for novel heart failure biomarker discovery in humans. Our results from the mouse cardiac transcriptome data was validated by matching to human heart failure data, and provide compelling evidence that circulating GPNMB levels are associated with heart failure. The mechanisms underlying this association and the role of GPNMB in heart failure are yet to be determined.

The last 6 months of my time in the lab was spent learning about next generation sequencing techniques and exome sequencing several patients with rare cardiovascular disease. This could not have been a better segue into my career as a clinical cardiologist with an emphasis on cardiovascular genetics. In July, I will be joining the UCLA Faculty as a General Cardiologist. I plan to stay very involved in Cardiovascular Genetics, including recruitment of patients with rare cardiovascular disease for exome sequencing and actually performing the sequence data analysis, variant filtration and interpretation. I will also be seeing patients every Friday in the Cardiovascular Genetics Clinic and hope to help build our program so that we have a large patient population and referral base within the UCLA Community. Last but not least, I would like to be a part of the effort to recruit cardiovascular disease patients for genotyping as a part of UCLA's precision medicine initiative to genotype 50,000 individuals per year for the next three years.

I am extremely satisfied that after several very informative and valuable years of work in mouse genetics, that I am able to realize my dream of participating and carrying on patient-centered research in genetics. This was my initial goal. The background I've gained in systems genetics and mouse genetics during my PhD will certainly help me achieve my goals of further characterizing the genetics underlying rare or extreme phenotypes in cardiovascular disease.