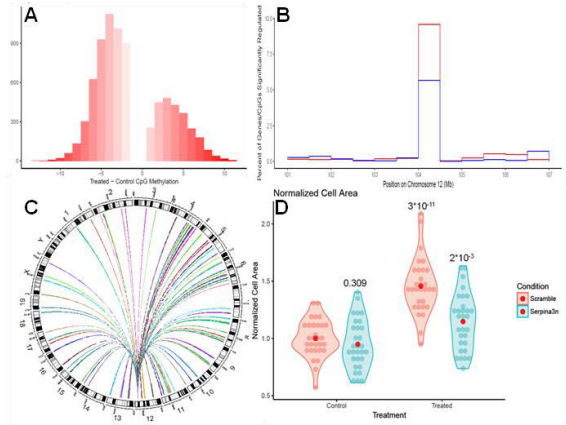
INTRODUCTION TO RESUBMISSION

I thank the reviewers for recognizing the numerous strengths of the proposal, especially the quality and productivity of the applicant, the superb mentoring team and clearly outlined CDP, the ideal institutional environment and the innovative research project supported by strong preliminary data. I am pleased to report that since my last submission, we have gathered even more data which strongly support the feasibility of the validation analyses proposed. Most notably, we have performed a pilot mQTL hotspot analysis in the subset (1%) of CpGs whose methylation status is significantly affected by catecholamine stimulation. Using this approach, we have identified a 500kb locus which is also implicated as a hotspot in the change in gene expression after catecholamine stimulation. Within this locus we have identified the gene *Serpina3n* as our likely candidate and have performed *in vitro* studies validating its effects on cardiomyocyte size and gene expression. Since my last submission, I have also published four additional manuscripts, two were first-author. One (***Methods Mol Biol 2017***, Rau et al) dealt with a suite of bioinformatic tools to prioritize candidate genes within loci which I use in my proposal, while the other (***Cell Systems 2017***, Rau et al) involves the analysis of the entire transcriptome and genome using a co-expression network algorithm. In this manuscript, we performed functional validation on a proposed driver gene, *Adamts2* and demonstrated that its role in the gene network was validated by qPCR analyses and (current) *in vivo* work. Coincidentally, *Serpina3n* plays an important role in one of these gene modules as well. This additional application of the heart failure HMPD panel, combined with the preliminary data outlined above, is strong evidence that the results of my proposed grant will be highly impactful in generating new hypothesis.



New Data. A) CpGs significantly altered by Isoproterenol. **B)** Overlapping 500kb hotspot containing *Serpina3n*. **C)** Genes whose expression is altered by *Serpina3n* locus. **D)** Effects on cardiomyocyte size by *Serpina3n* knockout.

The reviewers also identified several weaknesses in the proposal which have been rigorously addressed in the current submission. In particular, **1) The Research Plan is too discovery-oriented.** With the progress made since last submission, I have now added clear molecular hypotheses with newly discovered candidate genes. I will develop specific experiments to validate the function of *Mospd3* and *Serpina3n* as novel regulators of cardiac hypertrophy and epigenetics. The outcome will help to reveal novel mechanisms in epigenetic regulation by a cytosolic protein *Mospd3* and an extracellular localized protease *Serpina3n*. **2) A lack of mouse training in the mentored phase.** The high-confidence candidate genes will be functionally analyzed during the mentored phase of the project, giving me the opportunity to continue to acquire necessary skills and knowledge in cardiac pathophysiology under the mentorship of Dr. Wang. **3) No mentorship outside of UCLA.** I have established extensive collaborations with investigators outside of UCLA and I have included Dr. Noah Zaitlen, a professor at UCSF, as an advisor on this proposal. Dr. Zaitlen is an expert on bioinformatics methods development and has worked extensively with DNA methylation data in the past. Before the award begins, I will spend time in his lab at UCSF gaining training in bioinformatics. Dr. Zaitlen will be a valuable mentor to me going forward. **4) Concerns about separation from mentor.** I have expanded my explanation of how I will differentiate myself from Dr. Wang and how I can set up an HMDP at another institution. Several reviewers expressed concerns about the uniqueness of the HMDP to UCLA, yet the HMDP strains are all publically available from Jackson Laboratories, and numerous past trainees from the HMDP study group have achieved professorships and continued their studies, including Brian Parks (U Wisconsin), Charles Farber and Mete Civelek (U Virginia) and Brian Bennett (U North Carolina). Additionally, I have enjoyed excellent support from Dr. Wang, my primary mentor, who has clearly indicated that any results from these projects will be mine to pursue. I have also received assurances from Dr. Lusis that any data generated from the HMDP heart failure study will be mine to use as well. **5) Sex as a biological variable.** This study is feasible in the time frame of the K award due to the fact that it draws upon an already generated tissue bank from an HMDP study. I have expanded my explanation of this cohort to explain why it consists of only female mice, and why the replication of this cohort in males would have minimal training (and, based on the literature, scientific) value. However, I have included both sexes in the validation studies. **6) The Vertebrate Animal Section is underdeveloped.** I have expanded my vertebrate animal section as well as included more information (e.g. power calculations) into the research plan itself.

PROJECT SUMMARY

Heart failure (HF) is a growing concern among researchers, with rates expected to increase by 25% by 2030. HF is an incredibly complex disease with many pathological features including cardiomyocyte hypertrophy, contractile dysfunction and fibrotic remodeling. HF has complex etiologies, with common risk factors such as hypertension and diabetes being in and of themselves multi-factorial. Consequently, there is a tremendous amount of heterogeneity in human populations in both disease onset and progression which mask the demonstrated strong genetic component of common forms of HF. As a result of this heterogeneity, human genome-wide association studies have only been able to recover a handful of significant loci. We recently described a population of mice in which we identified over 30 loci for HF-related phenotypes and demonstrated that we had very significant overlap (50%) with significant or suggestive HF-associated loci in humans.

This proposal outlines an extension of this panel of mice to explore the epigenome of the heart both before and after catecholamine stimulation. Using novel experimental techniques and computational tools, we seek to identify important genes and pathways which control DNA methylation. We also seek to connect DNA methylation to changes in HF-related phenotypes. It also outlines an extensive career development plan for Dr. Rau to complete his training under the mentorship of Dr. Wang and transition to an independent academic position by establishing a multi-disciplinary research program in cardiovascular genetics and genomics.

During the K99 phase of this award, Dr. Rau will analyze the methylomes of 88 strains of mice both before and after catecholamine stimulation using reduced representation bisulfite sequencing. Research will focus on the use of methylation and phenotypic data to identify CpG-phenotype associations at epigenome-wide association study loci. Genotype-methylation associations will be examined to identify loci within which DNA mutations drive large differences in DNA methylation across the genome. By combining these loci with the significant amount of data previously gathered in this panel, we will predict causal genes and pathways implicated in HF. Based on preliminary results, during the K99 portion of the award the PI will also perform *in vivo* functional studies using the CRISPR/*Cas9* gene knockout system of two high-confidence candidate genes: *Mospd3*, which regulates heart weight and *Serpina3n*, which is implicated in the regulation of over 1800 CpGs. During the R00 portion, the PI will combine the genes identified during the K99 portion with his training in *in vivo* validation to identify novel genes and pathways which contribute to heart failure.

The overall goal of the proposed studies is to integrate systems biology, epigenetics and molecular analyses to lead to deeper understandings of the genetic pathways which regulate DNA methylation and HF-related phenotypes.

RELEVANCE TO PUBLIC HEALTH

Heart failure is a growing concern to researchers, with rates expected to rise by 25% by 2030. Understanding the biologic networks that underlie the complex interactions in the progression of heart failure is required for disease prevention, diagnosis and treatment.

SPECIFIC AIMS

Heart failure (HF) is a major burden on public health, affecting over 5 million patients with rates expected to increase by 25% and total cost of treatment is likely to double over the next 15 years. HF is highly heterogeneous in clinical manifestations among different patient populations. Consequently, there has been significant interest in understanding the genetic components which underlie common forms of HF. In humans, HF is a disease of the elderly with numerous different etiologies and inciting factors. This has made it very difficult to identify genome-wide significant candidate genes using population-based systems genetics tools such as genome-wide association studies (GWAS) which are easily confounded by environmental sources of variation. New methods are needed to study HF at the population level.

We recently implemented an ambitious study in which we examined HF in a large cohort of inbred mouse strains termed the Hybrid Mouse Diversity Panel (HMDP). We clearly demonstrated that combining a systems genetics approach with the HMDP, in which we fully control the environment, genetics and pathological stressor for HF, allowed us to capture significantly more of the underlying genetics which drive HF. We also demonstrated that using the β-adrenergic agonist isoproterenol (ISO) to mimic the catecholamine over-drive which is a common pathological trigger of HF allowed us to recover 50% of the suggestive human GWAS loci using fewer than 1000 mice. Despite these advances, much of the identified genetic heritability of HF in our mouse strains has not yet been recovered. We believe that this heritability is present, but acting through an alternate biological layer which we have not yet queried. In this proposal, we seek to elucidate novel contributors to heart failure by applying systems genetics approaches to the epigenome, an understudied but critical biological layer between genome and phenotype.

The epigenome consists of modifications of DNA and chromatin structures which are not dependent on the underlying sequence. The epigenome plays an important role in numerous crucial biological processes. Epigenetic modifications can be modulated by environmental factors, genetic backgrounds and, surprisingly, also passed down from generation to generation in a sequence-independent manner. Epigenetic changes can have a significant impact on the onset and progression of diseases and, therefore, may serve as an additional driver of HF. Several small-scale studies have begun to explore the methylome and its link to cardiac dysfunction; however population-level studies of the regulation of the methylome and its contribution to phenotypic change to the heart in mice have not been reported but are critically needed. In the same way that genome-wide studies provided an entirely new understanding of genomic variation and its regulation of complex phenotypes and diseases, so too will epigenome-wide studies open up novel and exciting mechanisms in DNA methylation and its effects on cardiac pathologies.

In this application, I propose to study the regulation of DNA methylation at a population-based level and its effects on cardiac hypertrophy and dysfunction using data and tissues gathered from the HMDP HF study. We have measured cardiac methylomes from left ventricular tissues across the strains of the HMDP-HF cohort both before and after isoproterenol stimulation. Our preliminary results demonstrate tissue-specific global DNA methylation patterns as well as significant changes in methylation status after ISO stimulation associated with differential phenotypic features. We have identified several highly promising candidate genes involved in global DNA methylation in the heart as well as ISO induced hypertrophy. These results support the feasibility of the proposed studies and lay the foundation for my current proposal as outlined in two aims.

Specific Aim 1: Systems Genetics Analyses of Cardiac DNA Methylome Regulation and its Effect on Heart Failure. My studies have generated a large amount of systems-wide data in the form of DNA polymorphisms, transcriptomes, comprehensive phenotyping and, now, DNA methylomes both at a basal state and in response to chronic catecholamine stimulation. This aim will focus on integrating these datasets to uncover novel insights into DNA methylation and HF. Integration of phenotypic data with methylome variation in epigenome-wide association studies will uncover DNA methylation changes which affect traits associated with HF. Mapping variable methylation sites to the genome will uncover methylation quantitative trait locus hotspots, where a single DNA polymorphism affects many CpG sites. Loci will then be queried *in silico* and *in vitro* to identify a prioritized list of candidate genes for *in vivo* validation and other mechanistic studies.

Specific Aim 2: Validation of the Role of High-Confidence Candidate Genes

Preliminary studies using subsets of the methylome data followed by *in vitro* analyses have identified the genes *Mospd3* and *Serpina3n* as high-confidence candidate genes which regulate DNA methylation and HF-associated phenotypes. This aim focuses on validation of these two candidate genes *in vivo* using CRISPR/*Cas9*-mediated knockout mice as well as laying out the strategy behind *in vivo* validation of several additional high-confidence candidate genes identified through Aim 1.

RESEARCH STRATEGY

1. SIGNIFICANCE

Heart Failure (HF) has resisted the downward trend of mortality seen in other cardiovascular disorders and its mortality is expected to rise by 25% by 2030[1]. HF is a complex disease with many pathological features including cardiomyocyte hypertrophy, contractile dysfunction and fibrotic remodeling. HF has complex etiologies, with common risk factors such as hypertension and diabetes arising due to complex, multi-factorial processes themselves[1–3]. The resulting heterogeneity seen in both disease onset and progression among human HF patients is, consequently, large. In addition to these external risks, there are also strong genetic factors that underlie common forms of HF[2–7]. The goal of this proposal is to uncover genetic and epigenetic drivers of common non-Mendelian forms of HF, which would lead to a better understanding of the fundamental pathogenic mechanisms of HF and aid in the development of precision medicine for patient populations.

Over the past decade, numerous groups have attempted to understand the genetic causes of complex, multi-factorial diseases through the use of **Systems Genetics** techniques, most notably the **Genome-Wide Association Study** (GWAS). Using this technique, scientists have identified over 10,000 genome-wide significant loci for over 1,500 traits in humans[8]. GWAS in heart failure, however, has only yielded a single locus in humans that reaches genome-wide significance as listed in the NHRGI GWAS catalogue[9]. Studies of specific features of heart failure have also had limited successes[10]. Unsurprisingly, the majority of the genetic underpinnings of HF have not yet been identified using GWAS. Recent research has suggested that an examination of other biological layers (e.g. the transcriptome) and component features (e.g. ejection fraction or left ventricular mass) of a disease may identify hidden genetic contributions to the overall phenotype[11].

There has been a recent explosion of interest in the **Epigenome**, a layer of biology in which changes are made to DNA or chromatin structures but not the underlying sequence. Researchers have demonstrated that changes to the epigenome may be modulated by environmental factors[12] as well as DNA variation[13] with certain degree of inheritability potential independent of DNA sequence[14]. Additionally, it has been demonstrated that epigenetic changes may contribute significantly to changes in phenotype, including the development and progression of HF. Our study would be the first systematic approach to investigate the influence of genetic variants to the cardiac epigenome in a controlled mouse population as well as the first to study on a population-based level how epigenome variants contribute to the manifestation of heart failure.

2. INNOVATION

***Conceptual Innovation*:** There are three major conceptual innovations in our proposal. 1. Systems genetics approaches to study the link between DNA methylation and cardiac disease using a large environmentally controlled animal population have not been performed. Our prior research has shown that such an approach may identify novel genes and pathways that affect phenotypic traits. 2. The epigenome is a vital yet understudied biological layer that must be linked to the genome and cardiac pathology. 3. New bioinformatic approaches addressing the distinctive realities of animal model population genetics are needed.

Methodological Innovation: A systems genetics approach and the integration of the epigenomic layer into our current analysis provides a unique opportunity to investigate heart failure in unprecedented depth and scope but also poses a tremendous technical challenge in integrating the large datasets obtained for genetic variants, epigenomic profiles, transcriptomic differences and phenotypic traits. The success of this ambitious proposal, however, is well supported by several key innovations in model systems and techniques, and developed and validated by extensive preliminary studies. These innovations include:

The Hybrid Mouse Diversity Panel Heart Failure Study: My work utilizes a recently developed panel of mice, the Hybrid Mouse Diversity Panel (HMDP), which consists of over 130 publicly available inbred strains of mice[15]. This cohort has several unique advantages that benefit the study of HF. First, environmental variations such as diet and climate can be eliminated while the pathological stressor can be accurately administered, minimizing confounding factors outside of defined genetic backgrounds. Second, physiological and molecular measurements can be assessed in the most relevant tissue, the heart, which is readily available in mice, making it feasible to conduct cardiac-centric systems studies. Third, since the HMDP mice are fully inbred and homozygous at each locus (i.e. genetically stable), the data gathered from different individual mice from the same strain and at different biological layers (transcriptome vs. epigenome) can be fully analyzed and integrated. We recently completed a study of 105 strains of female HMDP mice in which mice were treated with the beta adrenergic receptor agonist isoproterenol (ISO) for three weeks to mimic the chronic adrenergic stimulation seen in HF[16]. During this time, echocardiographic traits were analyzed weekly. Physical and clinical phenotypes (Figure 1) such as organ weights, cardiac function, tissue fibrosis, plasma lipids and glucose measurements were gathered along with left ventricular tissue for downstream analyses. We identified a number of significant HF-associated loci using GWAS in this population[16,17]. We observe

significant overlap between human and mouse GWAS loci for HF-related traits, suggesting that the underlying genetics that lead to HF are conserved between species[16]. Together, these facts suggest that the HMDP is a good model for systems-level analyses of HF.

Reduced Representation Bisulfite Sequencing of Mouse Heart Tissue: The DNA methylome consists of the methylation status of CpGs across the genome. Bisulfite conversion, in which sodium bisulfite is used to convert unprotected cytosines into uracils has been used to query DNA methylation status on a case-by-case basis for decades[18]. With recent advances in sequencing technology, it is now possible to explore DNA methylation across the entire genome using bisulfite conversion[18] by comparing the bisulfite converted sequence to a reference sequence. Whole-genome sequencing, however, is expensive and not suitable at this time to large-scale projects and other methods, such as meDIP-seq, are not high-throughput enough for our purposes. In contrast, Reduced Representation Bisulfite Sequencing (RRBS), where the restriction enzyme MspI is used to enrich the sample for CpG islands before sequencing[19], allows researchers to generate DNA methylation profiles enriched for phenotype-relevant CpGs for a fraction of the cost in a high-throughput manner. To date, no large studies have examined DNA methylation in mouse hearts from different genetic backgrounds, much less in hearts challenged by pathological stimulation. One of my collaborating advisors, Matteo Pellegrini (See Letter) has successfully analyzed hepatic DNA methylation profiles in the HMDP at a basal state[13]. Preliminary analysis supports the feasibility of successful acquisition of RRBS datasets from the HMDP-HF cohort, paving the way towards novel insights into the regulation of cardiac DNA methylation and its contribution to HF as outlines in the current proposal.

Table 1. Selected HMDP Phenotypes

Phenotype	Control	Treated
Total Heart Weight (mg)	98.5 +/- 14.3	133.4 +/- 24.2
Lung Weight (mg)	125.6 +/- 17.8	155.4 +/- 33.7
Fractional Shortening (%)	36.7 +/- 4.8	37.7 +/- 7.7
Cardiac Fibrosis (%)	2.02 +/- 1.7	8.71 +/- 8.84

Sophisticated Systems Biology Techniques: Successfully examining the results of the RRBS analysis in the HMDP HF study will require the use of cutting-edge bioinformatics tools. For instance, we will perform high-density GWAS using Fast-LMM on every variable methylation site to perform methylation Quantitative Trait Locus (mQTL) hotspot analyses in order to identify master regulators of cardiac methylation at a basal state or in response to pathological stressors. We will also use FastLMM-EWASher[20] to identify epigenetic variants associated with pathological traits of HF. This method was specifically designed to examine methylation data for epigenome-wide association studies, however it is possible that we will have to modify an existing method to work with a large population of mice that often require tools that make different assumptions to those made in human studies. A member of my mentorship team, Noah Zaitlen (see Letter) has extensive experience in the development and use of cutting-edge bioinformatics tools and has pledged to support this project. Preliminary analysis supports the successful implementation of these toolsets.

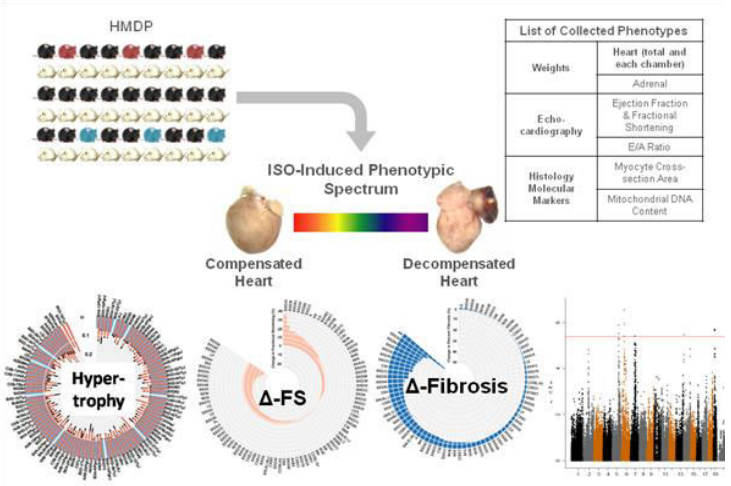


Figure 1. Phenotypic spectrum of cardiomyopathy in HMDP mice. 105 strains of HMDP mice was administered with 20 mg/kg isoproterenol for three weeks (average n=6 per strain). The insert table enlists the phenotypic traits collected from HMDP strains. The bottom circle plots illustrate the phenotypic differences among the HMDP strains in heart weight at basal (blue bars) and post-ISO (red bars), the changes in as well as an example of our GWAS results.

3. PRELIMINARY RESULTS

Phenotypic Spectrum of Catecholamine-Induced Cardiomyopathy in the HMDP. We completed our initial study in **105** strains of HMDP mice by comprehensively characterizing functional, structural and pathological changes following chronic stimulation with the clinically relevant β-adrenergic agonist isoproterenol (ISO)[16]. As shown in Figure 1 and Table 1, the pathological manifestations among the HMDP in response to this stressor are highly diverse. Analyses of inter-strain versus intra-strain variances indicate that genetics is responsible for between **64-84%** of the total observed phenotypic variation, which validates our central premise that subtle genetic variations among the HMDP majorly contribute to the pathogenesis of heart failure. The availability of HMDP mouse tissues at basal and post-ISO states and their associated cardiac phenotypic parameters gives us a unique opportunity to establish the contribution of epigenetics to HF as well its regulation in normal and diseased hearts.

Genetic Dissection of Heart Failure. Genome-wide association analyses were performed on the phenotypes gathered from the HMDP (Figure 1) using high-quality SNP profiling and the Efficient Mixed Model Algorithm (EMMA).[21] We identified **~25** significant loci contributing to cardiac hypertrophy and fibrosis[16] as well as **16** loci associated with cardiac structural and functional traits via echocardiographical analysis[17].

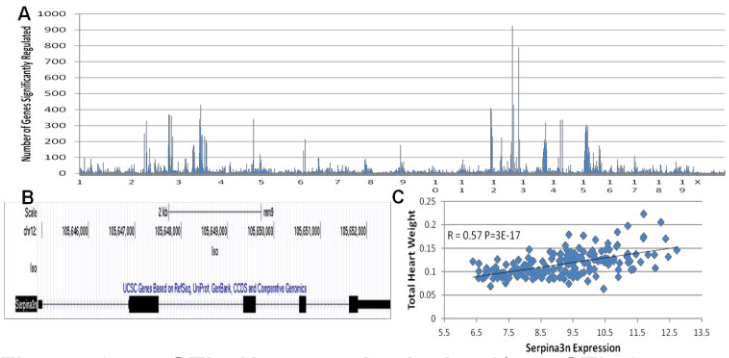


Figure 2. eQTL Hotspot Analysis A) eQTL hotspots across the genome. Each bar represents the number of genes that have a significant eQTL within a 500kb window. B) *Serpina3n* is the candidate gene underneath the chromosome 12 peak. C) *Serpina3n* has a strong and significant correlation with heart weight

We have further demonstrated[16] that roughly half of all the genetic loci with significant or suggestive association to human heart failure have been replicated in the HMDP HF study. This significant overlap strongly suggests that the underlying genetic pathways in both species are conserved. Our GWAS loci contain a number of candidate genes that are well established as significant contributors to HF such as *Ppp3ca*[22], *Pln*[23] and *Sgcd*[24], as well as a number novel genes, several of which have been validated using a combination of *in vitro* and *in vivo* techniques. These major advancements demonstrate that a systems genetics study in the HMDP-HF cohort is a tremendously powerful resource for gene discovery for HF and also a validated platform on which we will add epigenetic analyses for the current proposal. We will perform an epigenome-tailored version of GWAS as part of Aim 1 in this proposal.

eQTL Hotspot Analysis We previously demonstrated the feasibility of examining the role of individual loci on the genome that regulate, *in trans*, a large number of phenotypes. Prior work has focused on loci that regulate gene expression[25]. These loci, termed expression quantitative trait locus (eQTL) hotspots, often contain important mediators of tissue development and cellular processes and/or genes that respond to systemic stressors such as catecholamine stimulation. We performed an eQTL hotspot analysis in the HMDP HF study. The genome was partitioned into 4691 evenly spaced 500kb windows and eQTLs calculated via EMMA for 13,000 expressed genes. Each gene/window pairing was represented by the polymorphism with the lowest p-value for that pairing. Significant eQTL hotspots were determined by counting the genes in each window whose representative p-values were less than an FDR-identified cutoff followed by comparison to a bootstrapped distribution of enrichments to determine a final significance threshold (Figure 2A). We observed that many regions of the genome assist in the regulation of gene expression. Several loci, which we term 'master regulator loci,' stand out, each regulating over 5% of genes expressed in the heart. An examination of these loci reveals important genes for future study, such as *Drosha* and *Dicer*, the initiators and regulators of miRNA processing. We also identified the gene *Serpina3n (Figure 3B,C)*, which we have demonstrated *in vitro* regulates cardiac hypertrophy (Figure 4F) and which is dysregulated in dilated cardiomyopathy[26]. We will expand our hotspot approach to study DNA methylation in the heart to identify master regulator loci controlling global methylation patterns in Aim 1.

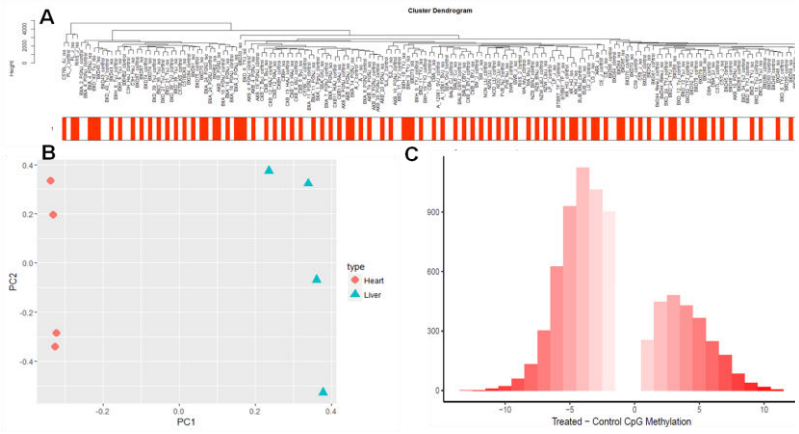


Figure 3. Global Analysis of RRBS Data. A) Hierarchical clustering analysis of the HMDP methylomes shows no preferential clustering between control (white) or treated (red) samples. B) Tissue type (Liver vs Heart) clearly divides a mixed pool of methylomes after principle component analysis. C) Significant changes to methylation between control and treated samples

Reduced Representation Bisulfite Sequencing of the HMDP HF Study: We performed multiplex 75-base pair RRBS analysis of left ventricular DNA samples using the Illumina GAIIx platform and the EpiTect bisulfite kit in 88 strains of female mice both with and without catecholamine stimulation from the HMDP HF Study. We chose to only perform RRBS in female mice as the generation of a male cohort of equivalent size would require at least two years of effort for minimal training value and is beyond the scope of this proposal. Prior research suggests that differences in DNA methylation due to sex are rare and typically nonsignificant[27].

Reads were aligned to the mouse genome (mm10) using the BS-Seeker2 algorithm[28] and counted using the methylKit and GenomicRanges R packages[29]. We recovered an average of 1.35 million CpG sites per sample (an average coverage of 60). Compared with a prior study performed on liver methylation[13], we observe similar proportions of CpGs within CpG islands and shores, suggesting that our data are of comparable quality.

Examining a subset of the data from both our study and the HMDP liver methylome study[13], we observed a striking difference between the two methylomes, with clear delineation between the tissues by principle component analysis (Figure 3B). We observe that 73k CpGs (~6%) vary significantly across the panel, supporting the assertion that global DNA methylation profiles are highly tissue- and strain-specific and it is necessary to investigate specific target tissues to establish relevant epigenetic profiles. In contrast (Figure 3A), we only observe 13K (~1%) CpGs that are significantly affected by isoproterenol, suggesting that global DNA methylation is largely stable and only a fraction of the CpGs are modulated by external stressors. These revelations are and will help us to focus our future analyses on these specific changes associated with cardiac pathology (Figure 3C).

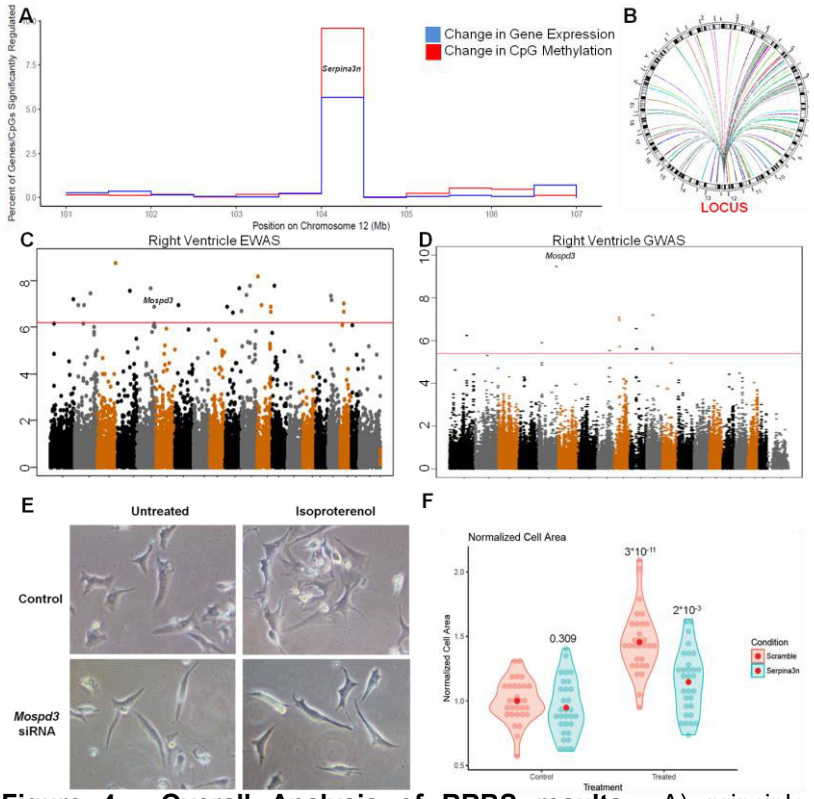


Figure 4. Overall Analysis of RRBS results A) principle component analysis of variable CpGs. B) GWAS on principle component 2 identifies a locus containing *Tox3*. C) EWAS on right ventricle weight identifies a locus containing *Mospd3*. D) This locus is also observed in a GWAS for the same phenotype. E) Representative photos of *in vitro* cardiomyocytes used to test *Mospd3*'s role in cellular hypertrophy. F) *Serpina3n* knockdown prevents cellular hypertrophy by isoproterenol

Pilot Studies*:* We performed two pilot experiments with our methylome data that further support the feasibility of our experimental design. We performed an mQTL hotspot analysis on the change in methylation status of the 13k CpGs that vary in response to catecholamine stimulation (Figure 3C). We observed a locus on chromosome 12 (Figure 4A) that regulated 9.7% of these variable CpGs and that perfectly overlapped the *Serpina3n* eQTL hotspot locus (Figure 2). Analysis of this site revealed significant associations with CpGs across the genome (Figure 4B). *Serpina3n* is a poorly described member of the serine protease inhibitor family that has been linked to improved wound healing and attenuation of muscular dystrophy symptoms[30,31], however its exact mechanism in cardiac tissue is unknown. We also performed EWAS in a single phenotype using the FastLMM-EWASher algorithm[20] (Figure 4C). We selected right ventricle (RV) weight as our phenotype of interest as we previously observed the largest number of GWAS hits for this trait and anticipated similar results for EWAS. We found 17 significant loci ($P<6.6*10^{-7}$) for RV weight after performing EWAS, including a locus on chromosome 5 ($P=1.37*10^{-7}$) that includes the gene *Mospd3***.** In our prior research, we identified a highly significant ($P=3.49*10^{-10}$) GWAS locus on chromosome 5 spanning roughly 200kb (Figure 4D) for RV weight. We identified the gene *Mospd3* as the strongest candidate at this locus as it contains multiple nonsynonymous coding variations among HMDP strains and is expressed in cardiac tissue. Prior research has suggested an RV-specific developmental defect in mice lacking *Mospd3*[32]*,* although its exact mechanism was not elucidated. Performing GWAS on the expression of *Mospd3* in ISO-treated mice shows a very strong association within 500kb ($P=2.6*10^{-8}$) of the gene, suggesting that *Mospd3* is linked to a *cis* genetic variant that strongly influences its expression. Observing *Mospd3* in both our GWAS and EWAS studies on the RV hypertrophy trait makes this gene a high priority for further study and experimentation.

For both *Serpina3n* and *Mospd3*, we have performed *in vitro* analyses using primary cardiomyocyte cultures (Figure 4E,F) that confirm that knockdown of both *Serpina3n* and *Mospd3* expression significantly ($P=1.5*10^{-4}$ and $P=5*10^{-10}$ respectively) affects NRVM hypertrophy and suggests that both genes play an

important role in cardiac hypertrophy and remodeling in response to ISO. We have begun the development of *in vivo* knockout mice for both of these genes and their validation will form the basis of part of Aim 2.

4. EXPERIMENTAL APPROACH

We propose to apply systems genetics approaches to understand the links connecting DNA variation, DNA methylation, and heart failure and to validate significant loci and novel genes associated with these functions **(approach outlined in Figure 5)**. **Aim 1** will integrate global genotypic, transcriptomic and methylomic datasets to identify gene candidates. **Aim 2** will perform focused studies of two high-confidence candidate genes associated in preliminary studies with cardiac dysfunction and begin study of novel genes identified in Aim 1. This proposal expands the research done on heart failure in Dr. Wang's lab from genetics to epigenetics, but relies on the same core principles that have driven our research in the past.

Specific Aim 1: Systems Genetics Analysis of DNA Methylome and Its Effects On Heart Failure

Rationale: The genetic factors that regulate DNA methylation and its subsequent effects on cardiac phenotypes are largely unknown due to limited accessibility of cardiac tissue in human studies. Previous work in the HMDP[16,17,33] demonstrated that it is well suited to studying multiple biological layers in HF and (in liver) DNA methylation[13]. Aim 1 will leverage the power of the HMDP to identify specific genetic variations that control the methylome and its subsequent role in the regulation of cardiac phenotypes. These candidates will be added to previously identified genes and examined in Aim 2. Each subaim adapts a powerful systems genetics technique for use in the analysis of DNA methylation.

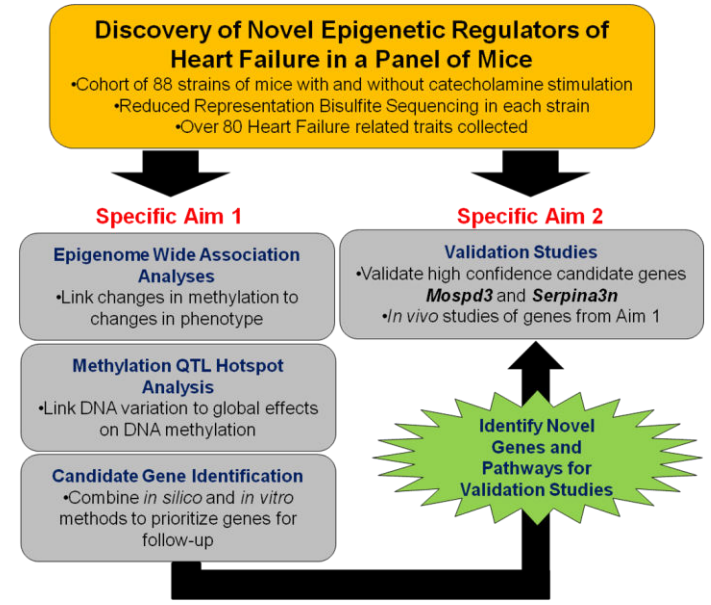


Figure 5. Overview of Experimental Approach

We generated RRBS methylomes from 88 strains with and without ISO stimulation (2 mice per strain and condition) from the HMDP HF study. Initial analyses (see preliminary results) suggest that these data are of high quality. Based on observed intrastrain variability in the HF HMDP, I expect to detect significant associations of individual CpGs that can explain down to 3% of the variation in a phenotype with more than 80% power. The experiments detailed below will be performed on CpGs that vary across the HMDP. Significantly varying CpGs will be identified using the MethylKit R package[29] using a previously detailed approach[13] where the median methylation for each CpG site will be calculated across the panel, and the absolute difference (Δ) calculated between each sample and the median. Variable methylation sites are defined as those in which at least 5% of the samples have a Δ greater than 50%.

Quality Control: Genotypic and transcriptomic data have been subjected to strict QC measures to avoid confounding effects as detailed in prior publications[16]. Methylome data will be treated similarly. CpGs that are also SNPs will be discarded, as will CpGs that are not present in at least 90% of samples.

Aim 1A: Identification of CpGs that regulate HF-related cardiac phenotypes: EWAS will be performed on the variable CpGs identified above using the FAST-LMM-EWASher algorithm[20]. EWASher is an extension of the FAST-LMM[34] algorithm that has been used to perform GWAS in the HMDP[35] and is designed to correct for sample relatedness, a major confounder of G/EWAS studies in mice. EWASher spectrally transforms methylome data and employs a linear mixed model to eliminate the effects of population structure from the analysis. We will perform EWAS on 12 phenotypes (Heart Weight, Each Ventricle Weight, Lung Weight, Liver Weight, Adrenal Weight, Left Ventricular Dimensions and Ejection Fraction). CpGs from untreated animals will be associated with both control phenotypes and treated phenotypes to identify CpGs that regulate phenotypes at a basal level and, importantly, any CpGs that may be used as a marker for future phenotypic outcome. CpGs will be examined

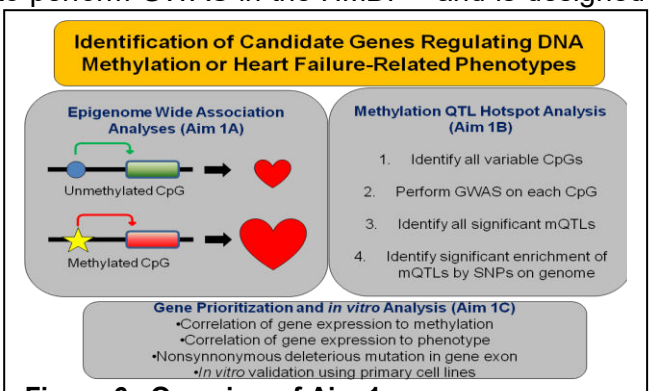


Figure 6. Overview of Aim 1

from treated animals to determine how they are linked to treated phenotypes. CpGs that pass a Bonferroni-corrected significance threshold ($6.6*10^{-7}$) will be analyzed further in Aim 1C to identify candidate genes, which, according to prior work[13], should lie within 500kb of the CpG.

Aim 1B: Identification of DNA variations that regulate large numbers of CpGs: Each variable methylation site will be treated as a separate quantitative trait and mapped to over 250,000 good-quality SNPs across the genome using the FAST-LMM[34] algorithm. A methylation quantitative trait locus (mQTL) is defined as one or more SNPs that are significantly ($4.2*10^{-6}$ [16]) associated with a difference in methylation across the strains of the HMDP. mQTL hotspots will be determined by binning the genome into 4961 evenly spaced 500kb bins and counting the number of variable methylation sites that possess at least one significant mQTL within each bin. This analysis will be performed three times, once on control Methylomes (to identify regions that regulate DNA methylation without perturbations), once on treated Methylomes (to identify loci that regulate CpGs relevant to HF-phenotypes) and once on the change in DNA methylation after ISO treatment (already completed in Preliminary Results). Permutation analysis will be used to determine bins with significantly more mQTLs than expected by chance and these bins will be analyzed further in Aim 1C to identify candidate genes involved in *en masse* regulation of cardiac DNA methylation.

Aim 1C: Prioritization of Candidate Genes Identified from Loci in Aims 1A and 1B: Candidate loci from Aims 1A and 1B will contain a large number of genes. In this subaim, I will use both *in silico* and *in vitro* approaches to prioritize these genes for validation in Aim 2C.

In silico methods: I have developed a suite of tools[36] that accelerates the *in silico* identification of candidate genes and expands on our group's validated approach previously used to prioritize genes for further analysis by leveraging all the information currently available to us in the HMDP HF Study[16,17,37]. For instance, special priority will be given to loci that overlap a locus previously identified in an HMDP study. *Loci identified in 1A:* All genes within 500kb (the average CpG linkage block size[13]) will be examined. Genes that are significantly correlated to both the tag CpG and the phenotype of interest will be prioritized and examined for nonsynonymous SNPs using the Wellcome Trust Mouse Genomes Resource[38]. Each gene's association with nearby SNPs will also be examined. Finally, literature analysis will be used to identify and prioritize genes that have relevant functions. The ideal candidate gene will show strong correlations between methylation, gene and phenotype, no *cis*-eQTL (unless that *cis* SNP is also associated with the mQTL for the CpG that is associated with the phenotype of interest) and, ideally, a gene function that is associated with a known cardiac signaling pathway. *Loci identified in 1B:* All genes within 2Mb of the center of the hotspot (the average SNP linkage block size in the HMDP[15]) will be examined and genes whose expression are suggestively ($P<.05$) correlated with at least 51% of the CpGs that map to that locus will be prioritized. The subsequent analysis is largely similar: examining each gene in the locus for eQTLs that map to SNPs in the locus, nonsynonymous mutations, and prior literature studies.

In vitro methods: Once we have identified our prioritized list of candidate genes via *in silico* approaches, we will examine the effects of siRNA-mediated knockdown of these genes in Neonatal Rat Ventricular Myocyte primary cell cultures as a preliminary screen of our candidates. Although each gene's specific protocol may vary (for instance, there is no need to examine ISO-treated cells if a gene is not implicated in an ISO-treated condition), each gene will be analyzed according to our screening protocol as previously described[33]. Briefly, NRVMs will be transfected using Lipofectamine RNAiMax, receiving either empty vector, scramble siRNA or one of two different siRNAs for the given gene. On day 2, transfected NVRMs will be treated with 60uM Isoproterenol for up to 3 more days. Cellular phenotypes to be analyzed include myocyte hypertrophy (cell sizes, protein synthesis) and hypertrophic gene regulation (marker genes including BNP, ANF, myh6/7). In all cases, these experiments will be repeated in triplicate (n=9 total for each gene/condition), which will give us sufficient power to validate the candidate gene. I plan to test 5 prioritized genes from each subaim. *Loci identified in Aim 1A:* Effects of gene expression on cell size, contractility rate, hypertrophic marker (Nppa, Nppb) expression and apoptosis using TUNEL staining will be measured. Differences across groups will be analyzed using ANOVA. Genes that show significant impact upon inactivation on cardiomyocyte hypertrophy, function and viability will be prioritized for *in vivo* studies (Aim 2C). *Loci identified in Aim 1B:* Effects of gene knockdown on cell parameters will be assessed as above and methylation-sensitive PCR will be performed using the Epimark Analysis Kit (NEB) on ten CpGs that are linked to the hotspot. If the methylation status of 7 or more of these CpGs are altered compared to unaltered NRVMs, the gene will be prioritized for *in vivo* studies (Aim 2C).

Expected Results, Potential Problems and Alternative Strategies: Prior analyses using GWAS and eQTL Hotspot analysis demonstrated that the HMDP is a powerful system capable of identifying causal genes for many complex traits. Although a methylome study is a novel approach for the heart, the results of a parallel

study in liver as well as our pilot studies in heart (see Preliminary Results) have already revealed over 20 loci, supporting the overall feasibility of main goals of our study. These results, paired with the validation analyses performed in Aim 1C showcased my significant experiences in both discovery and follow-up studies. I am highly confident that the outcome will provide a rich resource and strong pipeline for my future independent research projects going beyond the timeline of this K99/R00 proposal. In short, I do not anticipate problems in successfully completing Aim 1 during the K99 phase of this proposal. One criticism of this aim could be that it is not focused on testing a specific hypothesis. This is a limitation when the candidate genes have not been fully validated at functional level, leaving mechanistic hypothesis elusive at current stage of the discovery. However, population genetics studies, by their nature, are unbiased analyses of naturally occurring variation in the population and prior studies performed by myself, our lab and others have successfully identified key, relevant drivers of disease pathways that had not been identified using other techniques. I fully expect to identify novel and important regulators of heart failure and DNA methylation using these approaches, leading to novel hypotheses based on the molecular identity of the validated genes. The focused study in Aim 2 showcases the potential of our systems genetic approach in formulating novel hypothesis based the two newly validated genes. Another criticism is that it is possible that we may not detect any additional EWAS peaks or mQTL hotspots in our data, which would suggest that DNA methylation and its role in affecting HF-related phenotypes is more complex than we previously thought and we lack the power to see significant effects. While this is unlikely given the results of our pilot study, if necessary we can rely on the results of the pilot study alone, which have provided sufficient leads for Aim 2C and future experiments. If, conversely, the five genes we selected as our top priority candidates do not appear to affect phenotypes *in vitro*, then we can simply select genes further down the list of candidates, prioritizing those whose roles in DNA methylation or cardiac function are better understood. Finally, a criticism of our study design could be the decision to only study our previously generated female cohort and not expand into a male cohort. It is not possible to generate a male cohort *and* perform the proposed studies in the time frame of the K99 phase of the award due to logistic and budgetary constraints. Although this is a limitation, the reported role of sex on DNA methylation is relatively small[39], preliminary studies in the HMPD showed a stronger effect of catecholamines on female hearts and both sexes will be investigated in validation studies.

Specific Aim 2: Validation of Candidate Genes That Control DNA Methylation and Cardiac Traits

Rationale: My work has already identified two high-confidence candidate genes: one regulates DNA methylation *en masse* and another regulates heart weight. The goal of this aim is to validate these high-confidence candidates and begin investigations in the same vein on significant results obtained from Aim 1.

Aim 2A: The Role of *Mospd3* in Ventricular Growth in the Heart: We observed a shared GWAS and EWAS locus on chromosome 5 significantly associated with right ventricular weight following Iso treatment. *Mospd3* was selected as a top candidate for this locus based on the presence of a nonsynonymous SNP associated with small ventricle mass and an overall significant correlation of its expression with right ventricular weight ($R=.27$, $p<0.01$). *Mospd3* deficiency led to abnormal right ventricular development[32] however its role in adult hearts is unstudied and its molecular mechanism is unknown. In preliminary studies, (Figure 4) *Mospd3* knockdown in cultured myocytes significantly blunted ($p=5*10^{-10}$) cardiomyocyte hypertrophy. Based on these results, **I hypothesize that *Mospd3* expression is regulated by changes in cardiac DNA methylation in the heart and has a significant impact on stress induced cardiac hypertrophy**.

As *Mospd3* knockout mice have increased neonatal lethality[32], I will generate tissue-specific Mospd3 knockout mice by floxing the coding exons of *Mospd3*, and breeding it with aMHC-MCM mice expressing tamoxifen inducible Cre in cardiomyocytes in a C57BL6/J background[40], and confirm these genotypes via genomic DNA PCR. Mice of both sexes will be divided into six groups of 8 mice each: Two homozygous $Mospd3^{flox/flox/cre}$ knockout groups, two heterozygous $Mospd3^{flox/wt/cre}$ knockout groups and two floxed littermate $Mospd3^{flox/flox}$ control groups. Power calculations suggest that 8 mice per group will be sufficient to see phenotypic differences using ANOVA. Mice of both sexes will be maintained on standard diet until 9 weeks of age, and then given 5-day tamoxifen in their chew to initiate targeted *Mospd3* deletion. Cardiac function will be monitored by weekly echo following tamoxifen mediated gene inactivation (as previously described[16,17]). Cardiac hypertrophy will be measured based on tissue weight, gene expression and histological analysis (as described[41,42]). If basal cardiac phenotype is not affected by inducible KO of Mospd3, I will induce pressure-overload at 12 weeks (3 weeks post inducible KO) for each genotype via transaortic constriction (TAC) with sham surgery as a control. Progression to cardiac hypertrophy and HF will be demonstrated by serial weekly echocardiograms for six weeks. At the end of the study, mice will be sacrificed and relevant tissues (heart, lung, liver, adrenal) will be collected for hypertrophic marker gene expression and histology as previously described[32]. Additionally, the role of *Mospd3* in the cell is largely unknown. We will utilize mass spectrometry

to query protein-protein interaction partners and identify important functional motifs for *Mospd3* to inform future mechanistic studies. These experiments are expected to begin during the mentored (K99) phase and continue through the independent (R00) phase of the grant.

Aim 2B: The Role of *Serpina3n* in the Regulation of DNA Methylation after ISO Challenge eQTL and mQTL hotspot analyses were performed on the change in gene expression or DNA methylation after ISO challenge. An overlapping locus in both analyses spanning 500kb on chromosome 12 was identified. *Serpina3n* was selected as the top candidate in this locus due to its differential methylation, correlation with heart weight (R=.57, $P<1*10^{-16}$), and significant *cis*-eQTLs and *cis*-mQTLs for this locus, and a suggested role in muscular dystrophy[31] and viral myocarditis[43]. Finally, knockdown in cultured cardiomyocytes significantly blunts ($P=1.5*10^{-4}$) ISO induced cardiomyocyte hypertrophy (Figure 4). Based on these results, **I hypothesize that *Serpina3n* plays an important role in the regulation of cardiac DNA methylation during stress induced cardiac hypertrophy.**

We are currently obtaining a CRISPR/*Cas9*-mediated knockout line from the UC Davis Mouse Biology Program on a C57BL/6J background. The experimental design will be identical to aim 2A in terms of methods and numbers of mice in each group (*Serpina3n*$^{-/-}$ vs. WT). After detailed characterization at basal state, the sex and age-matched KO and WT mice will be subjected to either transaortic constriction or sham surgery at 12 weeks of age, and monitored for six weeks. Gene expression and RRBS will be performed on left ventricular tissue and genomic DNA using the Illumina GAIIx platform at the UCLA Genomic Core. Pair-ended RNA reads will be annotated using the Tuxedo suite[44] and differential gene expression will be determined via Cufflinks[44] and differential methylation determined via the MethylKit R package[29]. Effects of *Serpina3n* depletion will be determined for each gene/CpG predicted to be linked to *Serpina3n* expression and significance determined via ANOVA with appropriate FDR correction for multiple comparison testing. A significant enrichment (determined by permutation testing) for predicted genes will indicate that *Serpina3n* is acting as a causal gene for transcriptional regulation. Utilizing Mass Spec, we will identify protein-protein interaction partners that represent enzymatic targets for *Serpina3n*, allowing us to place the gene in the context of known pathways and inform future mechanistic studies. These experiments are expected to begin during the mentored (K99) phase and continue through the independent (R00) phase of the grant.

Aim 2C: Study Novel Genes Involved in DNA Methylation and Heart Failure Identified in Aim 1: We expect the systems genetics studies proposed in Aim 1 to provide a rich resource for identifying genes and pathways that regulate DNA methylation and HF-associated traits. For the scope of this proposal, I will analyze up to two candidate genes from both EWAS and mQTL hotspot analyses, modifying the approaches outlined in Aims 2A and 2B as appropriate to the phenotype and condition in which these candidates are observed. If appropriate, I will also utilize alternative *in vivo* or *in vitro* approaches. For example, the gene *Rspo1* shows significant methylation differences across the HMDP in response to ISO and is our top candidate for another differential methylation QTL locus that is currently undergoing *in vitro* testing. If *in vitro* analysis confirms the *in silico* prioritization of this gene, then it may be studied in a manner similar to that of *Serpina3n*. These ongoing studies will primarily be conducted during the independent (R00) phase of the grant.

Expected Results, Potential Problems and Alternative Strategies. We expect the results from Aim 2 to validate the significant loci associated with the regulation of DNA methylation and its effects on heart failure. Experiments to address *in vivo* effects are crucial for understanding the role of the gene in the context of the entire organism and will provide insight into the ways in which *Mospd3* and *Serpina3n* contribute to their respective phenotypes. A potential complication is that other genes within the associated loci might be contributing to the phenotype and, *in vivo*, our *in silico* and *in vitro* evidence will not be substantiated. In this case, future studies with other lower-priority genes within each locus may be warranted. Another potential complication is that both *Mospd3* and *Serpina3n* are expressed in multiple tissues in the body, and *Serpina3n*, specifically, has already been implicated in skeletal muscle regulation. Development of a tissue-specific conditional knockout for *Serpina3n* or *Mospd3* would be necessary. The techniques and approaches proposed in this aim are well established in my mentor's lab and there is significant experience in my mentorship team in studying heart failure in mice as well as any *in vitro* and *in vivo* studies appropriate to Aim 2C.

Proposed Timeline: All datasets (genotypes, transcriptomes and methylomes) from the HMDP HF study were collected before the start of this proposed study. Aim 1 will be completed during the K99 portion of the award. Functional characterization of *Serpina3n* and *Mospd3* (Aims 2A,B) will also begin during this portion. During the R00 phase, any additional *in vitro* validation will be completed (Aim 1C), molecular characterization at mechanistic level will continue for *Serpian3n and Mospd3*, and development of additional functional models (Aim 2C) will be prioritized. With the tremendous resources offered by Dr. Wang, Dr. Lusis, the UCLA

Cardiovascular Theme and my highly distinguished mentorship team, I anticipate completing the proposed studies without any major difficulties.

CANDIDATE BACKGROUND

I am fascinated by the ways in which mathematical and computational tools can be used to understand biological systems. I graduated with honors from Harvey Mudd College in 2007 with a Bachelors of Science degree in mathematical biology. Harvey Mudd is a nationally top-ranked undergraduate-only institution dedicated to the sciences where I was able to perform independent scientific research with tremendous freedom as early as my second year. While working in the lab of Dr. Mary Williams, I became my college's expert in microarray analysis, which was a novel technique at my small institution. I trained other students in microarray analysis and guest lectured in two courses regarding the science behind and analysis of transcriptome microarrays. My senior thesis was completed under Dr. Robert Drewell. I studied a series of enhancer elements which were, oddly, not conserved across the *drosophila* genus. Using my growing skills in bioinformatics, I found a conserved secondary structure in the enhancer which would now be called a lncRNA.

During my undergraduate studies, I also worked on the UC Santa Cruz Genome Browser under the mentorship of Dr. Todd Lowe. I developed an algorithm to identify conserved snoRNAs across multiple species. Working with the researchers at the genome browser solidified in my mind a desire to explore the big, genome-wide questions in biology where my love of math, programming and biology would be of great value.

I completed my dissertation research in the laboratory of Dr. Aldons "Jake" Lusis. My dissertation research resulted in four first author and seven co-author publications. I brought my growing skills in big-picture thinking, computation and genetics to bear on some of the large-scale projects being performed in his lab. My dissertation project used the Hybrid Mouse Diversity Panel to study the development of cardiac hypertrophy and failure after chronic catecholamine stimulation. I collected over 80 phenotypes in nearly 800 mice from over 100 strains to generate a massive body of research and tissues for further analysis. Working with an interdisciplinary team of collaborators, I combined high-density genotyping with these phenotypes to identify 46 genome-wide significant or suggestive associations between DNA polymorphisms and HF-related phenotypes. Through the use of transcriptomic analyses including expression quantitative trait loci (eQTLs), I identified the most likely candidate genes within these loci. Several of these identified candidates were previously implicated as signaling molecules in cardiac hypertrophy and pathology, validating the systems genetics approach I had taken. This effort led to the discovery of many novel regulators of HF. Several of these genes have been validated using *in vitro* or *in vivo* methods, such as *Miat*, a lncRNA and *Abcc6,* an orphan ABC transporter which, when knocked down, results in a dramatic increase in cardiac fibrosis after catecholamine stimulation. During my Ph.D. training, I received multiple predoctoral fellowship awards through competitive applications. More importantly, I extended my expertise in bioinformatics and genomics while gaining extensive training in experimental biology for cardiac pathology and physiology. This training motivated me to further develop my future research career in the area of the genomics of heart failure.

During my research into HF in mice, I began to collaborate closely with my current mentor, Dr. Yibin Wang, and we began to discuss exciting potential extensions of the work which I was doing which would mesh with his own expertise in murine HF. Dr. Wang is an internationally recognized leader in cardiac stress signaling and molecular cardiology research. His lab has developed sophisticated model systems to investigate stress signaling and gene regulatory pathways involved in cardiac pathogenesis. Together, we developed a tailored training strategy for me to broaden my research experience into new forms of analysis, new biological layers and a deeper understanding of cardiac biology while still harnessing the immense versatility of the resource I had generated and the broad and diverse interdisciplinary team which was present at UCLA. Consequently, I have become an expert in the cutting edge field of RNAseq, and have developed a novel co-expression network algorithm, which, when applied to the data generated from the HF HMDP study has revealed new and interesting candidates for further study, which we have validated in both *in vitro* and *in vivo* models. Recently, we finished the analysis of a pilot study on DNA methylation in the HF HMDP study and have just received the full methylation dataset from the entire panel of mice, which I plan to combine with the body of research already done on this panel of mice to identify novel regulators of DNA methylation and HF progression. This project is the focus of this grant application. As part of my postdoctoral work, I have presented my research at national and international conferences, and have already published eight first author and 18 co-author publications. Additionally, I have just completed an AHA Postdoctoral Fellowship.

The examination of complex, -omics-scale systems has been my passion since I was an undergraduate, and I am fully committed to transforming that passion into a life-long career studying the genetic basis of gene-environment interactions in cardiac disorders and beyond. The K99/R00 Pathway to Independence Award will provide me with the opportunity to advance my scientific training in several key areas and lead to my transition into an independent faculty position where I will set up an interdisciplinary research program based around systems genetics, gene-environment interactions and heart disease.

CAREER DEVELOPMENT AND TRAINING OBJECTIVES

The K99 funding opportunity provides for both research training and career development. Each of these components are critical for a successful career as an independent academic investigator. It is my goal to become proficient in the analysis of methylome data in the heart with the broader goal of being able to tackle systems genetics problems in both mouse and human research and both within the field of heart failure and beyond to phenotypes and diseases such as hypertension and atherosclerosis. This plan is, admittedly ambitious, so to support me in my development I have developed a two year mentored training plan. My overall goal during the next two-year period is the development of the skills needed to succeed as an independent investigator in the field of cardiovascular genetics. During my K99 phase, I seek to expand my knowledge of cardiovascular disorders and how to apply mathematical and computational tools to systems genetics problems relating to the development of disease as well as improve my bench science techniques to validate and pursue results from these computational approaches.

I have strengths in bioinformatics, cell biology and mouse genetics, and I will continue to develop myself in these fields. In order to succeed as a scientist in the competitive field of genetics of complex traits I have developed an intense career development strategy which will include technical courses, seminars, conferences and advisory committee meetings.

Formal Interactions with Mentor and Mentorship Team

The cornerstone of my career development plan is my mentorship team, which I have selected to both give me the skills and expertise I will need to succeed in the K99 and R00 phases of this proposal, but also enable me to gain important academic skills (such as grant preparation and teaching experience) to establish my own professional network through their extensive connections within the medical research community and to expand the scope and depth of future studies. In particular, all key members of my mentor team have demonstrated commitment and success in promoting young investigators to develop independent academic careers.

I meet with Dr. Wang for at least 1 hour every two weeks to discuss the progress and potential problems in my project. Dr. Wang is an expert in mouse genetics as it pertains to heart failure and has and will continue to train me in mouse genetics techniques and approaches.

In addition to my meetings with Dr. Wang, I also present my findings once every six weeks as part of our weekly lab meetings. I also present my work quarterly in the weekly Systems Genetics and Statistics Club, which I helped to organize. Additionally, constant interactions with my fellow postdocs and graduate students in my lab and others in the cardiovascular research laboratories (CVRL) group provide important feedback for my work.

I am fortunate to have a highly talented and distinguished panel of advisors to assist in my transition to independence. In addition to my primary mentor, Dr. Wang, my mentorship team includes Drs. James Weiss, Noah Zaitlen, Matteo Pellegrini and Paivi Pajukanta. Each has important skills and perspectives which they bring to my committee. Dr. Pellegrini is the leading expert in DNA methylation at UCLA and has worked with the model organism system (the HMDP) which I will be using. Dr. Zaitlen has extensive bioinformatic knowledge and his group has helped to develop many of the programs which I intend to use as part of this project. Both he and Dr. Pellegrini will also advise me on how to run a successful bioinformatics laboratory, including ways in which to set up fruitful collaborations. Dr. Weiss is the Chief of Cardiology at UCLA, with extensive knowledge of systems biology in cardiac diseases and heart failure. Dr. Pajukanta has extensive experience in GWAS and other cardiovascular traits. Additionally, Drs. Pajukanta and Weiss are experts in human studies and have connections within the medical community, both important to my desire to collaborate with others on human research later in my academic career.

Several members of my mentorship team attend the weekly Systems Genetics and Statistics Club meetings. In addition, I will meet specifically with my mentorship team twice a year to evaluate my progress. I will present the progress of my current research and career development in a 40 minute presentation and receive critical feedback. I expect to be supported as well as challenged by this distinguished group of scientists throughout the award period.

In addition to my time at UCLA, I will be spending a period of time in the laboratory of Dr. Zaitlen (see letter) at UC San Francisco to gain valuable experience in the development of bioinformatics techniques, in particular techniques relating to the analysis of DNA methylomes. This is a crucial part of my long-term career development and these novel skills will help distinguish me from Dr. Wang. I will return to UCLA before the beginning of the Mentored portion of the K99 with valuable perspectives and skills which will help me to complete the aims of both the mentored and independent portions of this grant.

Educational Activities:
I will continue to hone my skills in statistical analyses and bioinformatics by auditing the following classes:

Bioinformatics M260B: Algorithms in Bioinformatincs and Systems Blology
Biomath M271: Statistical Analysis of Incomplete Data
Biostats 266B: Advanced Biostatistics
Bioinformatics M252: Advanced Methods in Computational Biology

Career Development
I will attend workshops that will facilitate my transition to a tenure-track faculty position

1. **Seminar Series by Society of Postdoctoral Scholars:** This monthly seminar series focuses on topics that are relevant to finding a faculty position: CV preparation for job applications, preparation of research and teaching portfolios, giving job and chalk talks, and negotiating a job offer
2. **MORE Grant Writing Workshop**: National Institute of General Medical Sciences, New Orleans. Provides training in grant writing skills that will improve my ability to write effectively and compete for an R01 at the end of my career development award.
3. **Supervisory Principles Workshop:** UCLA Human Resources Department. This intensive, three day workshop provides an overview of supervision and skills needed to successfully lead and manage work groups such as an academic laboratory.
4. **UCLA Biosciences Postdoctoral Education Leadership Program**: I have completed this training program which involved observing faculty members while teaching, leading discussion groups, preparing exams and providing research mentoring to first year grad students.

Seminars

1. Attend Weekly Systems Genetics meetings. I will present at least quarterly.
2. Attend weekly CVRL meetings. I will present at this meeting once per year. When external speakers come, I frequently will have lunch with them or meet them independently to discuss research and career goals.
3. Attend Weekly Human Genetics Seminar Series. Internationally recognized geneticists present their work.
4. Attend Weekly IMED Seminars. A seminar series where nobel laureates and national academy members present their work to the UCLA community.
5. Attend Molecular Biology Institute Weekly Seminar. Rotating speakers from UCLA and beyond present on topics pertaining to members of the MBI. I will present once per year.
6. Attend Clinical and Translational Science Institute Seminars. National and internationally recognized scientists and doctors present on Translational science

Conferences

1. AHA Basic Cardiovascular Sciences Conference (Each Year)
2. Keystone Meeting(s) on DNA and RNA Methylation (Jan 2018) and Epigenetics and Human Disease (2019)
3. International Society for Heart Research (Year 2 Summer. Year 4 Summer)
4. International Conference on Pathways, Networks and Systems Medicine (Year 3 Summer)
5. American Heart Association Scientific Sessions (Year 1, Fall Year 4, Fall)
6. Greybill Conference on Statistical Genetics and Genomics (year 2 & Year 4, Summer)

Mentored Job Search

During the K99 phase of the award, I will look for entry-level tenure track faculty positions in research universities using multiple job search sites, such as Nature, Science and the Chronicle for Higher Education. My mentorship team have trained numerous graduate students and postdoctoral fellows who have successfully transitioned into faculty positions. Together, they have an extensive network of contacts at universities and institutions for possible tenure-track positions. In preparing applications for these positions, I will prepare my application with input from Dr. Wang and my mentorship team. I will also practice my job talk in our laboratory meeting, systems biology meeting and my advisory group meeting to receive critical feedback.

CAREER GOALS AND OBJECTIVES

My long-term career goal is to obtain a tenure-track academic professorship at a major research university and set up a laboratory to explore and dissect the complicated layers of interactions which lie between genomic variation and phenotypic expression, specifically in the context of heart disease. I will focus on creating a highly interdisciplinary research group that will reflect my shared interests in computational and wet-lab approaches to the study of disease. I will seek out highly collaborative, diverse departments with a focus on research relevant to human disease, as I firmly believe that the best research is done in a collaborative setting. I have always enjoyed sharing my scientific enthusiasm, knowledge and expertise with the next generation of scientists and have, over the course of my graduate and postdoctoral work mentored 14 undergraduate students and 2 graduate students, as well as provided several tutorials on bioinformatics techniques to my laboratory. I fully intend to continue to mentor individuals in my research laboratory and encourage their curiosity in science and their desire to pursue a career in biomedical research.

Understanding the genetic basis of complex disorders, such as heart failure, is a challenging problem, primarily due to the myriad number of interacting genetic, epigenetic and environmental factors. Each of these factors individually and synergistically contributes to disease pathology. In the past decade, there has been a significant influx of data generated from cardiac research in the form of systems-level analyses such as genotyping arrays, transcriptome microarrays, RNAseq, CHIPseq, proteomics and methylation analyses. Although each of these datasets can be analyzed individually, and single genes identified in each category of data may be tested in *in vitro* or *in vivo* models and their effects tested using standard molecular biology techniques, the potential applicable knowledge which can be derived from these data are far from being fully realized as much of it lies *between* these layers, rather than simply within one. As an independent researcher, I intend to devote my career to understanding the complex interactions between these intermediary steps between DNA variation and the phenome in order to unearth and elucidate the relationships which propagate and multiply small changes in individual layers into striking changes in the development and progression of disease. The advantage of this approach is that it allows information to flow naturally between layers of biology, subject to multiple perturbations and factors, rather than limiting analyses to a single layer of biology or a individual genetic perturbation (e.g. through CRISPR-mediated knockout mice). This approach will allow my lab to more easily identify high-quality candidate genes which can then be tested and studied in detail using traditional cell and molecular biology approaches.

I am also passionate about advancing the field through the development of novel bioinformatics tools. My postdoctoral fellowship has given me the opportunity to develop several algorithms and gain expertise in cutting-edge bioinformatics analysis methods. My mentorship committee includes both Dr. Matteo Pellegrini and Dr. Noah Zaitlen, both of whom have extensive backgrounds in the development of bioinformatics algorithms and who have agreed to mentor me in the development of additional tools in my own research.

My research has generated an enormous amount of high-dimensional, genome-wide data, including gene expression, DNA methylation and genotypes and over **80** clinically-relevant phenotypes. These data are a rich resource and will serve as an invaluable opportunity for further analyses and studies during not only the time frame of this award, but beyond. Drs. Lusis and Wang have agreed to allow me full and unrestricted access to these data even after I obtain my independent faculty position. Additionally, Dr. Wang has agreed to allow me to take any mice generated as a part of this study with me to perform studies during the independent (R00) phase of this award. The three of us agree that my blending of classical mouse genetics techniques in the field of heart failure and novel bioinformatic approaches will easily distinguish my lab from both the Lusis and Wang labs when I achieve an independent position. Furthermore, as demonstrated by earlier postdoctoral fellows at UCLA who have used the HMDP and established their own independent laboratories, the commercial availability of each HMDP strain, added to the assurances that I will be allowed full use of any data or animals generated over the course of my work at UCLA, makes me confident that I will be able to establish my own mouse genomics laboratory either at another institution or at UCLA. Finally, with the assistance of Drs. Weiss and Pajukanta, it is my goal to expand my research into human data and the broader cardiovascular field.

Through my graduate and postdoctoral research, I have established a strong background in genomics, epigenomics, cutting edge bioinformatics techniques and cardiac pathophysiology. The training which I receive during the K99 phase of this award will help me to perform the research proposed for the R00 phase. At the end of the award period, I hope to have generated novel and exciting findings which will serve as the basis for a R01 grant application.

EQUIPMENT RESOURCES

Wet Bench Equipment: The Division of Molecular Medicine has shared major equipment for all member faculties. They include Olympus laser scanning confocal microscope and commercially available (Metamorph, Imaris and Aitoquant) and custom-made software for image analysis including 3D deconvolution, colocalization measurements and 3D reconstruction; Vevo770 ultra-resolution imaging system (Visualsonics, Toronto) equipped with 30 Mhz mouse scan-head and cardiac function analysis suite capable of digitally recording and analyzing 2D images as well as M-mode line scanning images at 1000 Hz; ARIA Milllar pressure-volume catheter for hemodynamic measurement of cardiac function; ADI telemetry system for continuation EKG recording; Sorval high-speed and ultra-centrifuges; Real-time RT-PCR from Bio-Rad with analysis software; film developer in dark-room and machine shops; gel-documentation system. Genetics Core includes, among others, an Illumina GAIIx sequencing platform for use in miRNAseq and RRBS analyses. In addition, Dr. Wang's laboratory possesses light and fluorescent microscopes, a spectrophotometer, PCR machines, incubators, fume hoods, refrigerators, freezers and gel electrophoresis equipment. Also available is a fully stocked BSLII-certified tissue culture room.

Computational Resources: The proposed project requires large scale computational and data storage resources. Large-scale computations (mQTLs, EWAS) will be carried out on the Hoffman2 computation cluster at the Institute for Digital Research and Education which is located on the UCLA campus. This computational cluster consists of more than 13,000 processors. I have unrestricted access to these processors for up to 14 days of uninterrupted computational time. Data storage is also provided by the Hofman2 cluster and we have over 3 TB of storage space available for my data. These data are routinely backed up.

After these computations are completed, I will move the results to our laboratory's data repository computer and perform the analysis on this computer (32 core, 64 GB RAM, RAID-5 Linux-based) with over 12 TB of storage space. I have full access to a number of free and commercial software such as "R", "SAM" and "Origin." In addition, all HMDP data will be placed on the HMDP Server for an additional layer of backup security.

FACILITIES & RESOURCES

Laboratory. This work will be carried out in the laboratories of Dr. Wang, who has approximately 1200 square feet (BH 569) of laboratory space in the Center for Health Sciences. The lab shares several Core facilities provided by the Division of Molecular Medicine, including cold rooms, tissue culture rooms and animal surgery/recovery rooms.

Animal: Dr. Rau has direct access to a well-run, AAALAC-accredited vivarium which houses mice under specific pathogen-free conditions. All mouse work for this project will be conducted in a dedicated area which is maintained by staff and routinely inspected by University animal facility officials. Experiements are performed according to the policies and guidelines established by the UCLA Animal Research Committee. Furthermore, there is access at UCLA to Core facilities to assist in the development of transgenic and gene-targeted mice.

Computer: Dr. Rau has a personal computer with relevant computer programs for completing this grant. Furthermore, significant resources for the stroage/analysis of the extensive data collected for this project are available (see Equiptment Resources)

Office: Dr. Rau has a personal desk and file stroage area, with a telephone and high-speed internet. Administrative assistance is available from Dr. Wang's lab manager and the Division of Molecular Medicine

Safety Facilities: All facilities are available in the Lusis lab to ensure personal protection from any biohazardous or potentially dangerous substances. The Core Tissue Culture facility is BSL-2 certified and has the Environmental Health and Safety mandated equipment incuding dedicated bio-safety cabinets, incubators, centrifugation and cell stroage. There is a dedicated fume hood for experiments using volatile chemicals, as well as a dedicated bio-safety cabinet and incubator for work involving viruses

Scientific Environment: Specific activities within the Division of Molecular Medicine and the Cardiovascular Research Laboratories (CVRL) group include molecular and cell biology, systems biology, electrophysiology and genetics. The CVRL has a longstanding commitment to the pursuit of basic research approaches to futher our understanding of cardiac diseases. There are numerous faculty members with which Dr. Rau may seek out for collaborations and advice, including James Weiss, Alan Garfinkel, Tom Vondriska, Karen Reue, Matteo Pellegrini and many others. In addition to the support available within Dr. Wang's lab, the Division of Molecular Mediine fully supports beginning scientists. There are several organized peer groups that will be instrumental in facilitating Dr. Rau's cotninued intellectual growth. These include weekly meetings of the CVRL, the Biostatistics group and the weekly Human Genetics Seminar Series. As evidenced by Dr. Rau's application, these meetings provide a framework for growth, and an opportunity to discuss his work with experts in the field.

UCLA Graduate Division Academic Services, Office of Postdoctoral and Visiting Scholar Services:

UCLA Graduate Division is the primary administrative office for >1000 postdoctoral scholars at UCLA. Graduate Division processes postdoctoral appointments, offers orientation for all incoming postdocs, provides information on employment policies, benefits, counseling and wellness services, child care resources, housing, and social support services. Graduate Division also funds and collaborates on a variety of career and professional development training resources, facilitates funding support, and advocates for postdoctoral scholars, including sponsoring the Society of Postdoctoral Scholars at UCLA, a postdoctoral association that advocates for postdoctoral scholars, develops professional and social opportunities, provides funding for travel to professional conferences, and acts as an advisory board to the UCLA administration.

The Office of Postdoctoral Affairs for the Biomedical and Life Sciences: The David Geffen School of Medicine Office of Postdoctoral Affairs (Office) for the Biomedical and Life Sciences provides focused support to the 600+ postdoctoral scholars in biosciences disciplines; whether appointed to David Geffen School of Medicine, the College of Letters and Science, the Henry Samueli School of Engineering and Applied Science, or UCLA-affiliated schools centers and institutes. The Office works in conjunction with the UCLA Graduate Division to support postdoctoral scholars and their mentors by coordinating training to meet career and professional goals, facilitating funding applications, and advocating for postdoctoral trainees. The office offers a full slate of career and professional development resources for trainees, including skill building, career path exploration, network building, and direct training, all of them are made available through a central portal grad.ucla.edu/careerhub. The Office also coordinates the Responsible Conduct in Research Training for biosciences trainees, recruiting active faculty mentors to present and participate, and confirming and recording the participation and completion by trainees. In addition to the RCR courses, the Office provides a variety of training seminars on individual topics in scientific best practices. UCLA Graduate Division also hosts the Graduate and Postdoctoral Extramural Support (GRAPES) Database, a searchable database of private and publicly funding awards, grants and fellowships. The Office maintains a website of customized resources to

facilitate the development of funding proposals, including grant writing, proposal development, and funding searches.

Institutional Investment in Early Stage Investigators: As described in the Institutional Investment Letter, the UCLA Division of Molecular Medicine is deeply committed to the development of in-house talent and early stage investigators. The assembly of an advisory committee, consisting of **Drs. Yibin Wang, Matteo Pellegrini, Noah Zaitlen, James Weiss and Paivi Pajukanta** is just one example of the commitment of the institute to enabling progression and career development. This advisory committee will meet twice yearly and guide his career development as he transitions to independence. Furthermore, Dr. Rau has had full administrative support for this grant from the Division of Molecular Medicine, and will continue to have this after he transitions to the independent phase.

Yibin Wang, Ph.D. Professor
Chair, UCLA Cardiovascular Research Initiative
Vice Chair for Research, Anesthesiology
Chief, Division of Molecular Medicine
Departments of Anesthesiology, Physiology and Medicine
Cardiovascular Research Laboratory
Molecular Biology Institute
David Geffen School of Medicine at UCLA
Los Angeles, CA 90095
Tel: (310) 206-5197
Fax: (310) 206-5907
Email: yibinwang@mednet.ucla.edu

Re: Dr. Christoph Rau

Mentor's Statement

Dear Review Committee Members:

I am very pleased to serve as the Primary Scientific Mentor for Christoph Rau Ph.D. during his K99 Award. I have served as his postdoctoral advisor since 2014 after he completed his Ph.D training under the supervision of Dr. Lusis.

I am a tenured full Professor in the Department of Anesthesiology, Physiology and Medicine. My lab has a long term interest in cardiac biology and diseases, especially in molecular pathways and gene networks involved in stress-signaling and gene regulation for cardiac hypertrophy, dysfunction and remodeling. We have published more than 181 manuscripts in the leading journals of the field. In my capacity as the Vice Chair for Research and Division Chief for Molecular Medicine, I also have administrative responsibilities to supervise the research portfolio at the Department of Anesthesiology which is currently ranked at top 5 in NIH funding nationwide. I have mentored four junior faculties within the Division of Molecular Medicine who are now fully established Full or Associate Professors each with multiple NIH grants. I have an active role in education and research training at UCLA. As the Vice Chair for the Molecular, Cellular and Integrated Physiology (MCIP) Home Area in the Biomedical Sciences Program for Ph.D., and a course director for M252/262, the entry required course for MCIP students, I am responsible for student recruitment, curriculum development and other training activities of the Ph.D. program. I have directly supervised more than 12 Ph.D. students and numerous post-doc fellows, most of them have productive careers in academia and industry, including faculty at Emory, U of Arizona and Wuhan Uni. I have also participated in more than 30 Ph.D. thesis committees in the past decade. Finally, as the Chair of Cardiovascular Theme at UCLA School of Medicine, I am a key member of the UCLA leadership team to coordinate all cardiovascular research and educational efforts at UCLA in which Dr. Rau will play an important role. Therefore, I have adequate experience and credential to serve as his scientific mentor, and support his career development towards an independent scientist.

Rau has a unique situation who needs a specialized training strategy. Christoph joined my lab in 2013 as a postdoctoral fellow after completing his Ph.D. with Dr. Lusis as primary supervisor. During his Ph.D. thesis research, we have already collaborated extensively on a joint project of using HMDP and systems genetics approach to investigate the genetic basis of heart failure. Served as a driving force for this project, Christoph demonstrated extensive expertise in informatics and genomics, and the ability to integrate different datasets in order to discover new genes for heart failure and extrapolate important interactions between candidate genes and cardiac pathology. He has highly sophisticated skillsets in math, computer programming, modeling, informatics, genomics and large dataset analysis, and has gained extensive knowledge in cardiac physiology and pathology through first-hand collection and measurements from the HMDP mice. From these studies, Christoph has published extensively in the top journals of the field, including Cell Systems, Plos Genetics, and Circulation, Cardiovascular Genetics. In this process, he became highly interested in heart diseases as a main focus for his future independent research. Based on his unique strength and his career goals, we planned a training strategy which will allow him to explore novel biological layers where he will have ample chances to discover novel genes and pathways for future exploration. In the meantime, he will gain complementary training in molecular biology, molecular genetic tools, and cardiovascular pathophysiology. The ultimate goal is to prepare Christoph with an independent research portfolio with tremendous potential of future exploration, and help Christoph to establish necessary skills and know-how to perform functional

characterization of candidate genes in model systems at molecular, cellular and systems levels. In this K99 Postdoctoral award period, we will focus on the following specific goals: 1). I will continue to work on his analytical skills of scientific thinking and project development by evaluating current advancement in the field, and prepare to publish two to three review articles in the related area of his projects, namely, systems genetics methodological development and epigenetic regulation in cardiac development and diseases. 2). I will support and help him to develop a robust research project focusing on cardiac epi-genome profiling across the HMDP mice and exploring novel mechanisms involved in tissue-specific epigenetic programming and genetic regulation of cardiac epigenome. We plan to publish two high-impact papers directly derived from his discovery of epigenomic regulation in cardiac transcriptome and pathologies. 3). I will help and guide his effort in develop communication skills for data presentation, manuscript preparation as well as formulating research grants. I will support his effort to apply for AHA SDG grant, other NIH K01 and RO1 grants. I will support him to participate national and local scientific conferences and provide opportunity for networking and interaction with his peers.

Dr. Rau's proposed research is designed specifically to match his demonstrated expertise and interest in systems genetics of heart failure, and is explicitly prepared for him to continue in his own independent lab. In the past two years as a postdoc fellow under my direct mentorship, he has made impressive progress towards independence in experimental design, execution and data analysis. The research project has also evolved from hypothesis generating epigenomic profiling and discovery to more focused, hypothesis testing based on initial findings. In the current award period, I plan to support him, both during our regularly scheduled bi-weekly meetings as well as in general, to continue explore the epigenomic regulation at systems level, while help him to develop more focused studies based on these initial findings and offer in-depth mechanistic and functional insights for candidate genes for cardiac epigenetic regulation during stress induced pathological remodeling. Developing a balanced research portfolio, with hypothesis generating explorations transitioning combined with hypothesis testing, in-depth, and mechanism focused experiments, is the main goal of this transition period. I will mentor him to supervise research associates, student trainees and other research personnel, and share with him managerial skills in resource management, time management and personnel issues. I will also help him to be prepared for regulatory compliance issues and training for animal and potential human subject research.

Dr. Rau has generated an impressive body of preliminary data, and developed exciting new project based on initial analysis. He has published more than 26 papers, many in the top journals of systems biology and cardiovascular biology. We currently have two major publications under preparation, one focusing on the genetic basis of tissue specific epigenome regulation in heart at basal, diseased states and in response to ISO induction. Another paper will be focused on genes affected by DNA methylation (epigenetic regulation) in the pathogenesis of heart failure. From these series of studies, Christoph will have opportunity to explore further the mechanistic insights from the novel candidate genes for specific features of heart failure, including cardiac hypertrophy, fibrosis and contractile dysfunction. In two years, I will also help him to compete for an academic position in a leading institute by mentoring him the process of job interview and planning. It is gratifying to note that Dr. Rau has received extraordinary level of recognition from his peers and has been invited for seminars and paper reviews. In R00 phase of his proposal, I will continue to support his investigation by leveraging our current strengths in molecular signaling, genetic models and cardiac pathophysiology. I am fully committed that Dr. Rau will be leading the project further and I will not have any overlapping or competing projects derived from this line of research. Rather, I will continue to provide synergistic support to ensure the successful implementation of his proposed study.

My lab is current funded by four NIH grants with Dr. Wang serving in three of them as the PI or multi-PI. The lab is also supported by Cardiovascular Research Laboratories at UCLA and Department of Anesthesiology through dedicated Core supplements. Recently, I received a UCLA CTSI Scholar Award and two industry grant with substantial support towards genomic research in heart failure.

In short, I am fully committed and well prepared to support Dr. Rau as his primary mentor to ensure a successful and productive independent research career.

Yours sincerely,

Yibin Wang, Ph.D.
Professor and Director of Molecular Medicine
Chair of Cardiovascular Theme @ UCLA http://medschool.ucla.edu/cardiovascular

AUTHENTIFICATION OF KEY BIOLOGICAL AND CHEMICAL RESOURCES

All key resources for this proposal will be authenticated to enhance the reproducibility of our results, as appropriate and according to NIH policy.

Key Biological Resources that will be utilized in this proposal include:

Cell Lines: Primary neonatal rat ventricular myocytes
Chemicals: Isoproterenol, Lipofectamine
Vertebrate Animals: C57BL6/J mice, CRISPR-generated knockouts.
Nucleic Acids: siRNAs for cellular knockdown

Authentification of Cell Lines: Only primary cell lines derived by our lab will be used in this proposal. We will follow the same protocol for each derivation of cells as described in Brown et al 2005, and will confirm cellular identity concurrently with experimental qPCR using cardiomyocyte-specific cell markers (cardiac troponin I, myosin heavy chain, sarcomeric alpha-actinin).

Authentification of Chemicals:
Isoproterenol: 100 mg of Isoproterenol will be acquired from Sigma, which provides verification of its identity. Chemical will be stored at -80 degrees and aliquoted and diluted before *in vitro* experiments begin. Efficacy will be measured by examining the effect of Isoproterenol on wildtype cells run concurrently with knockdown samples as a positive control. If the chemical is unable to replicate the hypertrophy seen in prior experiments, it will be swapped out for a new aliquot, whose potency will be measured before *in vitro* experiments resume.
Lipofectamine: Lipofectamine RNAiMax will be obtained from Thermo Fisher, which provides verification of its identity.

Authentification of Vertebrate Animals:
Wildtype C57BL6/J animals will be purchased directly from Jackson Laboratories.
CRISPR-generated knockout founders will be purchased from the UC Davis Mouse Biology Program on a C7BL6/J background. Founder animals are genotyped by UC Davis, but will also be genotyped by us using tail tissue to confirm genetic manipulation and check for the documented *Nnt* mutation. These mice will then be bred to generate CRISPR-generated knockout mice. Each mouse used for experimentation will be genotyped against its littermates using tail tissue to confirm genetic manipulation.
Mouse Breeding and Pathogens: The department of laboratory animal medicine (DLAM) at UCLA has rigorous standards in place to prevent pathogens from entering vivariums, including rederivation of imported mouse lines, a separate clean breeding facility and a sentinel cage system for the researcher-accessible vivarium. In the case that a pathogen is detected in a sentinel cage, mice currently within that room will be tested for pathogens by DLAM and future mice transferred from the clean breeding facility will be housed in another room until the pathogen can no longer be detected in the original room. Additionally, we will test our mice for the presence of the pathogen and treat properly according to the recommendations of DLAM.

Authentification of Nucleic Acids:
siRNA: siRNAs will be obtained from Integrated DNA Technologies, which maintains a validated set of siRNAs for every gene in the mouse and rat genomes. siRNAs for each gene will be obtained and tested *in vitro* before actual experimentation until two *siRNAs* that show significant knockdown of the gene target via qPCR are identified. Scramble DNA and no siRNA will serve as negative controls.