

SPECIFIC AIMS (Mechanisms for Dehydration Resistance, PI MacManes)

The maintenance of water balance is critical for survival. Humans are exquisitely sensitive to changes in osmolality, with slight derangement eliciting physiologic compromise. When the loss of water exceeds dietary intake, dehydration - and in extreme cases, death - can occur. Far from uncommon, millions of people die every year as a direct result of dehydration. In contrast to humans, animals living in desert habitats thrive without water and endure extreme heat and intense drought, as a direct result of unique adaptations. These adaptations allow them to survive conditions fatal to humans and most other animals. Despite being a well-known ecological phenomenon with obvious implications for human health, we know very little of the underlying mechanisms that allow for survival in desert environments. **The proposed research uses an innovative approach integrating physiology, evolutionary genomics, and computational biology to better understand how animals survive in what appear to be non-survivable conditions.** This proposal represents the foundational steps toward developing the cactus mouse (*Peromyscus eremicus*) as a model system for the study of physiologic water conservation. Indeed, this model offers the scientific community a unique opportunity to gain a deep understanding into the physiology and genomics of osmoregulation in extreme environments – a critically important insight that is impossible to achieve using a traditional model system like *Mus* that, like humans, die when subjected to these conditions. While not a part of this proposal, this project lays the groundwork for our long-term research goal – to identify the causal links between phenotype and genotype, using emerging technologies like the CRISPR-Cas9 system. Ultimately, understanding the mechanisms underlying extreme osmoregulation may suggest novel treatment strategies for conditions (e.g. diarrhea) resulting in acute dehydration in humans.

SPECIFIC AIM 1: To characterize the physiologic and genomic response (differential gene expression, patterns of methylation or isoform use) to extreme water restriction and heat.

The working hypothesis is that while desert-adapted mice may demonstrate genome wide expression patterns suggestive of stress (e.g. activation of heat shock protein, vasopressin responsive pathways) during dehydration, these responses function to preserve normal physiology and thus serum chemistry will be similar to mice with unrestricted access to water.

SPECIFIC AIM 2: To determine the ontogeny of extreme osmoregulatory ability, from the neonatal period during which fluid (milk) intake is obligate through weaning, when oral fluid intake is exceptionally rare.

The hypothesis here is that patterns of renal gene expression during fetal development through weaning will resemble patterns of gene expression, isoform use, and methylation typical of adult mice when water is freely available.

The proposed project aims to integrate studies of physiology, genomics, and computational biology to gain a deep understanding of a fundamental physiological problem – how to conserve water when intake is limited. Although dehydration is both common and dangerous, the biology underlying its physiological effects is currently invisible to researchers using traditional mammalian models of disease that lack the eco-evolutionary history present in desert-adapted mice. This project will fill a critically important gap in our understanding, which is in support of the research aims of the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), and specifically, of the Kidney Basic Research program, which supports fundamental research on the normal development, structure, and function of the kidney.

i. Significance

Dehydration, whether caused by exposure to extreme environmental conditions, water deprivation, or by infection (e.g. diarrheal illnesses) represents a significant threat to human life. In spite of modern medicine, millions of people die every year from dehydration. Compounding issues of exposure and illness are public health issues regarding the delivery of safe drinking water. With global climate change, these challenges are thought to become only more severe and as a result, research providing insight into the mechanisms underlying physiologic resistance to acute dehydration is urgently needed. The response to acute dehydration in humans and traditional mammalian models is generally maladaptive and may include death - this response limits our ability to develop novel insights into this important cause of human mortality. As such, the study of dehydration-tolerant mammalian models will significantly enhance our understanding, and will provide fodder for novel treatments. **The proposed work aims to study extreme osmoregulation in a uniquely suited novel desert-adapted model organism.**

While the mechanisms underlying physiological compromise in dehydration are well characterized (Roberts et al., 2010), some animals possess the ability, much unlike humans, to osmoregulate despite extreme heat and a complete lack of extrinsic water intake (Nagy and Gruchacz, 1994). Specifically, highly adapted desert mice may never drink water, produce an extremely viscous urine, or no urine at all, and excrete urea in the form of uric acid crystals in the feces (Schmidt-Nielsen and Schmidt-Nielsen, 1952). This phenotype results in an animal that is very resistant to dehydration-related physiologic compromise, and is in stark contrast to the phenotype of humans and traditional model organisms (e.g. *Mus* and *Rattus*). Although model organisms are attractive targets for study, they lack the requisite biology which may limit insight. In contrast with traditional model organisms, non-model desert-adapted organisms may provide a unique opportunity to study dehydration tolerance, though they typically lack many of the genomic and physiologic tools characteristic of model organisms. Despite this, renal gene expression has been characterized for several genes in desert animals, and was shown to be highly derived in some (e.g. *Dipodomys* (Huang et al., 2001)), but not in others (e.g. *Notomys* Weaver et al. (1994)). No studies characterizing genome-wide patterns of gene expression, methylation or isoform use in desert-adapted water stressed animals have been done and therefore the extent to which differences in these parameters underlie phenotype remains unknown. The proposed work effectively integrates the power of a model organism with the unique biology of a desert-adapted rodent, the cactus mouse (*Peromyscus eremicus*), to generate insights into extreme osmoregulation not current possible.

ii. Innovation

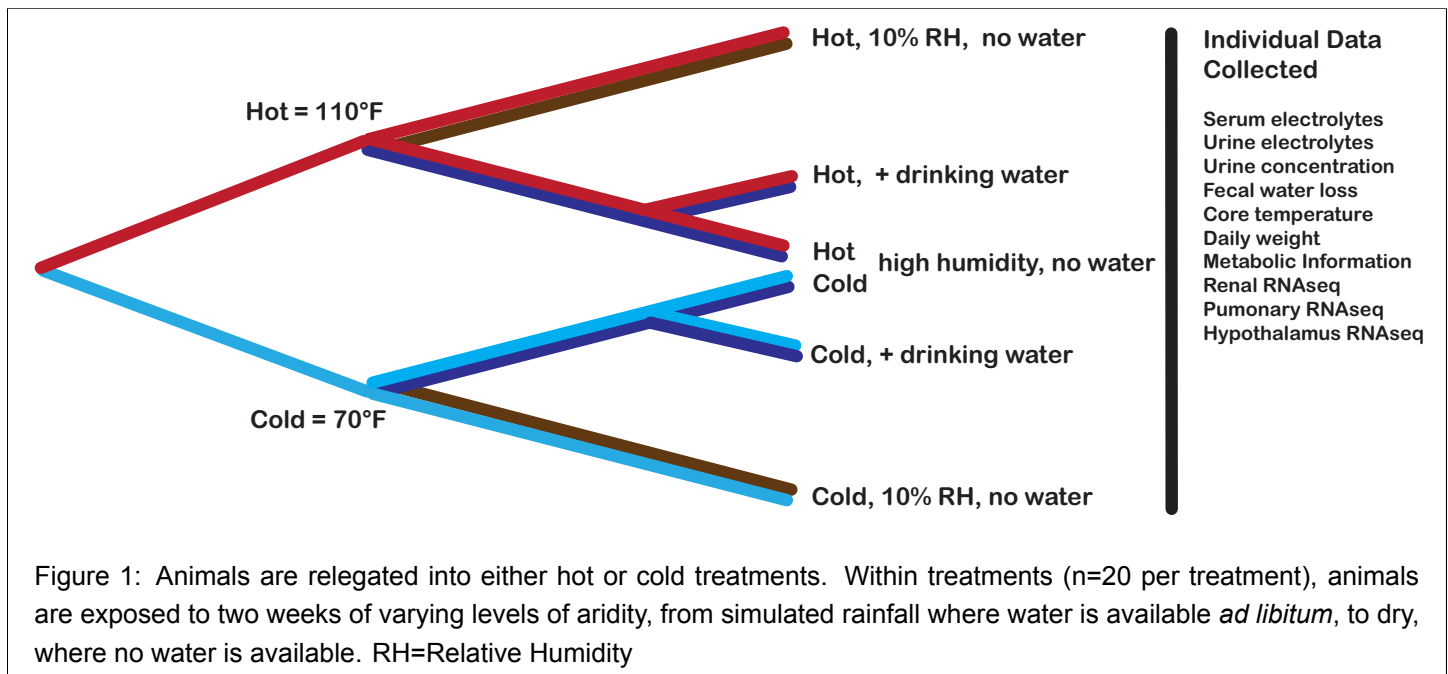
The proposed work recognizes that successful treatment requires an appropriate model, and while traditional models are powerful, they lack the biology (extreme osmoregulation) upon which more successful interventions may be modeled. The desert-adapted rodent *P. eremicus* retains many of the beneficial characteristics of model organisms, while enhancing opportunity to assay interesting biological phenomenon. In addition to this fundamental innovation, the project is innovative in a number of other ways.

- Experimental, conceptual, theoretical, and technical innovation: The proposed project leverages unprecedented control over environmental conditions using an ideally suited novel model organism and unique analytical methods to understand the physiologic and genomic response to water deprivation.

iii. Approach

Aim 1: To characterize the physiologic and genomic response (differential gene expression, patterns of methylation or isoform use) to extreme water restriction and heat.

To better understand the physiologic and genomic response underlying dehydration resistance, a series of experiments that will allow us to understand how differences in temperature, relative humidity, and water availability affect the desert-adapted rodent *Peromyscus eremicus* will be conducted. These experiments are fundamentally a series of environmental manipulations, described in Figure 1. The experimental design is fully factorial, meaning that the focal experimental parameter (e.g. water availability) will be tested in the context of the full range of other conditions (e.g. humidity, temperature). *Environmental parameters are chosen to match the most extreme (hottest and driest) conditions faced by wild animals in nature*, while cool/moist conditions effectively represent a less challenging control environment. Animals are exposed to each experimental condition for 14 days. Animal care is standardized between experiments and includes measures to reduce the water content of food and bedding materials. Both of these will be dried in a standard desiccation oven to less than 1% water/volume. Twenty individuals per treatment will be included - power analyses suggest this sample size will allow for detection of statistical support for patterns with small to medium effect sizes. Together, this design will make it possible to tease-apart the physiologic and genomic response to the various conditions.



For each experiment in Aim 1, physiologic and genomic data will be collected. In the context of limited water intake, how animals achieve electrolyte balance is unknown. Electrolytes are both easy to assay and are critical to physiological well being. Indeed, proper electrolyte balance is fundamental to all other physiological processes like neuronal signal transduction and muscle (including cardiac) contractility. Here, serum electrolytes will be measured using the VetScan VS2 critical care panel which includes ALT, BUN, Cl, CRE, GLU, K, Na, bicarbonate ion in a 100uL sample volume.

In addition to assaying electrolytes themselves, measures of urine electrolytes and specific gravity will be collected, as the urinary system represents that major pathway through which these chemicals are lost. These pa-

rameters will be measured using an Atago UG- α urine refractometer and tests conducted at the IDEXX reference lab. Lastly, animals will be weighed to the nearest 0.1gm every other day, including the day of sacrifice. Body temperature will be assayed with weighing using a digital thermometer and probe designed by World Precision Instruments (Sarasota, FL). In connection with this, feces will be collected and water content will be measured using standard methods.

Key metabolic parameters such as carbon dioxide production and oxygen consumption that may influence water consumption will be collected. In addition, the change in relative humidity within the metabolic chamber will be assayed, which will allow us to understand the rate of pulmonary water loss (or gain). These tests will be measured during a twenty four hour period at the end of the experimental manipulation, just prior to euthanasia, using a metabolic chamber (Sable Inc.) modified for use in the desert chamber. Together, these data will represent a uniquely rich characterization of the physiological state of a desert rodent held in captivity but more importantly, exposed to conditions typical of the natural environment. Of note, all procedures involving vertebrate animals conform to the guidelines provided in (Sikes et al., 2011) and have been approved by the University of New Hampshire Animal Care and Use Committee.

Aim 1a: Determine the physiologic response to drinking-water deprivation, extreme temperature, and humidity in the desert-adapted rodent *P. eremicus*. It is hypothesized that, as a result of unique mechanisms related to solute and water balance, average serum electrolyte concentrations will remain relatively constant throughout various experimental manipulations, but the variance in measured levels between individuals will increase in the most extreme conditions. These differences will be echoed in differences in urine electrolytes and concentration. Predictions regarding other parameters are detailed in Table 1.

Background: The human body consist of 60% water (Jéquier and Constant, 2009). Far from a static reservoir, proper physiologic function requires water for countless processes including nutrient transport (Haussinger, 1996), signal transduction, pH balance, thermal regulation (Montain et al., 1999) and the removal of metabolic waste. To accomplish these functions, approximately 2 liters of fluid are used daily - these fluids are lost mainly via the gastrointestinal and genitourinary systems, and by evaporative loss, which is accelerated greatly in extremes of heat and aridity (Cheuvront et al., 2010). These losses must be matched by intake (Jéquier and Constant, 2009), mainly in the form of oral fluid intake. Though the body possesses limited reserves, when loss exceeds intake over even a short period of time, dehydration and in extreme cases, death can occur. Humans and most other animals are exquisitely sensitive to dehydration, and possess limited compensatory mechanisms. In contrast, desert rodents survive in extreme environmental conditions, often without fluid intake. Understanding the mechanisms underlying this remarkable phenotype requires we understand the physiology that accompanies it. The work described here aims to characterize the physiology of dehydration resistance in desert adapted rodents.

While the prolonged absence of drinking water is invariably fatal for humans and many other animals, one potentially mitigating effect may be the acquisition of water (or limitation of loss) via the pulmonary vasculature, which is known to be variably permeable to water (Berger et al., 2011; Goralski et al., 2010). While pulmonary water acquisition has not been quantified in humans or in mammalian models, the pulmonary vasculature is ideally positioned to retain water from inspired air. Following this, relative humidity - the amount of extractable water present in respired air may be important to overall hydration status. The design described above incorporates two different levels of humidity to begin to disentangle the effects of drinking water from water acquisition via the pulmonary system.

Although water stress is obviously important to the survival of desert rodents - a phenotype which is relevant to human health and wellness, extreme temperatures represent another way in which physiological processes may be challenged. While desert animals may thrive in extreme heat, humans cannot. The physiological response is characterized in model organisms, but not in other animals adapted to these conditions. Genes like the heat-shock proteins are protective in humans, but no record of their activity on desert rodents is known.

Research Plan:To accomplish this aim, physiologic data from animals held with and without drinking water will be gathered, factorial with respect to the other conditions (e.g. temperature and humidity). The specific experiments described in Figure 1 will allow us to tease apart the effects of water deprivation from other parameters. Though the data we propose to collect is described above, in brief, we plan to collect blood and urine electrolytes and urine specific gravity. We will collect data on fecal water content, animal weight and temperature, as well as a battery of metabolic parameters. The specific predictions regarding several of these parameters are described in Table 1.

The statistical treatment of the data will include a multivariate regression (either linear or non-linear) to establish the relationships between the data. Many of these analyses will be conducted with non-parametric tests, as data

are often non-normally distributed nor independent. One of the most interesting comparisons will be to understand the relationship between serum sodium and urine sodium, urine concentration, fecal water content, and changes in body weight. Ultimately (e.g. Aim 1b) these data will be linked with patterns of gene expression, methylation, and isoform use to gain a synthetic understanding of dehydration resistance.

		Serum E-lytes	Urine Conc.	Fecal Water	Weight
Hot/Dry Hot/Wet		Mean Na, Cl, K ↑ to =, Var ↑↑	↑↑↑↑↑	↓↓↓↓↓	↓↓↓↓↓
	Rain	Mean Na, Cl, K ↑ to =	↑↑	↓↓	↓↓
	Humidity	Mean Na, Cl, K ↑ to =	↑↑↑	↓↓↓↓	↓↓↓↓
Cold/Dry Cold		Mean Na, Cl, K ↑ to =, Var ↑↑	↑↑↑↑↑	↓↓↓↓↓	↓↓↓↓↓
	Rain	Mean Na, Cl, K ↑ to =	↑	↓	↓
	Humidity	Mean Na, Cl, K ↑ to =	↑↑	↓↓	↓↓

Table 1: Predicted response given specific experimental manipulations. The number of arrows indicate the predicted relative magnitude of the response.

Preliminary data: The electrolyte profile of 2 individuals housed at 70F, 50% RH, water *ad lib* and two individuals housed in identical conditions except that drinking water was withheld has been characterized. Despite being housed in typical laboratory conditions, these animals have remarkably unusual electrolyte panel. For instance, mean serum sodium is 152 mmol/L, chloride 105 mmol/L, potassium in an un-hemolyzed sample is unusually high at 8.1 mmol/L, while Creatinine is low, with a mean measurement of 0.25mg/dL. mean blood urea nitrogen (BUN) is 47mg/dL. In contrast, animals without *ad lib* water were obviously dehydrated, with a mean serum sodium of >170 mmol/L and chloride 126 mmol/L. Interesting severe dehydration was not complicated by renal impairment as evidenced by a mean serum creatinine of 0.3 mg/dL and BUN of 59 mg/dL. Animals lost a remarkable amount

of weight, on average 28% of total body weight. Despite this decline in weight and electrolyte derangement, animals were active as per usual. **These results are shockingly distinct from human response to dehydration, and warrant further study.**

Aim 1b: Define patterns of gene expression, isoform use, and methylation given differences in environmental condition. We will understand the genetic response to extreme heat and aridity via a series of Illumina bisulfite, Illumina mRNA sequencing, and PacBio mRNA sequencing experiments, and will link these patterns to individual physiologic state as defined in Aim 1. We hypothesize that genes responsible for water and solute transport will be particularly active in the most extreme conditions in renal and pulmonary tissues, while genes involved in the activation of the hypothalamic-neurohypophysial system will be differentially regulated in the hypothalamus.

Background: Broadly speaking, genes underlie the vast majority of observable phenotypes. Whether this relationship is mediated by patterns of expression (e.g. Teets et al. (2012)), which itself may be mediated by differences in methylation (Brenet et al., 2011), or by use of alternative splice isoforms (Yukutake and Yasui, 2010), linking genotype to phenotype is extremely difficult. In addition to these mechanisms, function (=phenotype) may be determined by post-translational modifications like phosphorylation of specific sites (Moeller et al., 2009). The identification of these mechanisms is important, not only because in doing so we gain a deeper understanding of evolution, but also because these molecular mechanisms may be later used as targets for drug development or other therapeutic intervention. With regards to resistance to dehydration, the development of novel therapies is critical, as millions of people die yearly as a consequence.

In model organisms, dehydration precipitates a physiological response that is largely driven by the neuroendocrine system. Very much simplified, the cascade begins with the stimulation of osmoreceptors (Arsenijevic and Baertschi, 1985), which in turn stimulates neurons located in the paraventricular and supraoptic nuclei of the hypothalamus to release anti-diuretic hormone (ADH) (Zingg et al., 1986). ADH then binds to vasopressin-responsive receptors located in the renal medulla, resulting in aquaporin movement to the surface of the collecting duct (Nielsen et al., 1995) which encourages water re-uptake. In addition to the aquaporins, the renin-angiotensin-aldosterone system (Gubler and Antignac, 2009), natriuretic peptides (Totsune et al., 1994), the SLC and mTOR families (Ortells et al., 2012), and potentially other yet to be discovered pathways are important to water balance. Far from canonical, each stage in these cascades is dynamic and therefore pathways revealed in *Mus* and humans may not be equivalent to pathways in uniquely adapted desert animals, particularly given radically different phenotypes.

The genomic processes related to dehydration resistance in desert animals has yet to be characterized. The few studies of genetics that have been conducted have focused on the role of expression of single members of the aquaporin gene family (but see Bartolo and Donald (2007)), which are large membrane-bound proteins that are critically involved in renal water transport (Kwon et al., 2009; Verkman, 2002; Brown et al., 1995; Nielsen and Agre, 1995). These studies have shown that changes in Aquaporin (AQP) protein abundance and expression may be related to water availability (Böselt et al., 2009; Gallardo et al., 2005; Bozinovic et al., 2003). In addition to changes in expression, another study showed that the AQP4 pathway was completely lost in the desert rodent *Dipodomys merriami merriami* (Huang et al., 2001). Despite these studies, we have a limited understanding of the genomics of renal water and solute regulation in desert animals. While AQPs are functionally important, water and solute balance is extraordinarily complex, and therefore single-gene studies are necessarily limited in their purview. A more complete understanding of this phenotype and its mechanistic underpinnings will require a

sophisticated genome-level approach, which will be the outcome of the proposed research. In contrast to the limited amount known about patterns of renal gene expression, much less is known about gene expression in other tissues, and absolutely nothing about differential methylation or isoform use, even though we know that these complexities are mechanistically important to this specific function (Yukutake and Yasui, 2010; Silberstein et al., 2004).

Research Plan: The analysis of the genome wide patterns of response to dehydration will be conducted using the same individuals for which we collected physiology data. To accomplish this goal, RNAseq libraries for each individual and tissue ($n=120$ animals * 3 tissues) will be constructed at the HCGS. Because of the large number of samples to be prepared, this work critically depends on the enhanced capacity of the GEBRI - specifically the automated liquid handler. Sequencing will be conducted on the HCGS HiSeq 2500. We aim to generate approximately 10-20 million 100nt paired-end sequences per sample, which corresponds to 15 high-output HiSeq lanes using 24-way multiplexing.

RNAseq reads derived from kidney, lung, and hypothalamus will be mapped to the existing annotated draft genome, which was sequenced using startup funds. This phase of the project will be accomplished using the short read aligner BWA (Li, 2013) and best practices previously established (MacManes, 2014). Differential expression will be evaluated via the Cufflinks package (Trapnell et al., 2012), while evidence for coordinated changes in large numbers of genes will be detected using the software wcn (Langfelder and Horvath, 2008). The MacManes lab has demonstrated expertise in this area. Pipelines previously developed by PI MacManes combined with GEBRI-enhanced computational infrastructure makes analyses of this scale feasible.

Accurate isoform reconstruction is notoriously difficult using high-throughput short read sequence data such as that produced by Illumina HiSeq platform (Pyrkosz et al., 2013; Hiller et al., 2009), despite the advent of longer read lengths and newer analytical techniques (LeGault and Dewey, 2013; Jiang and Wong, 2009). In projects like this, where differential isoform use may be critical to phenotype, a different approach may be warranted. For instance, the sequencing technology available from Pacific Biosystems (PacBio) is suggested to provide a resolution to the isoform reconstruction problems (Au et al., 2013), specifically because it involves a long-read single molecule sequencing strategy (Eid et al., 2009). To identify patterns of differential isoform use, we will sequence poly-A selected mRNA samples using PacBio technology available to us via one of our regional GEBRI partners at the University of Delaware. Reads will be error corrected using the program LSC (Au et al., 2012), and isoforms will be identified using methods contained in Au et al. (2013). Because PacBio throughput is relatively low, which may limit the precision with which quantitation can be achieved, we will explore alternative ways to accurately estimate isoform specific expression. One previously unexplored approach involves estimating expression in the program eXpress (Roberts and Pachter, 2013) using only those reads that map uniquely and unambiguously to a specific isoform. Because this approach is uncharacterized, it will be validated using a set of isoform specific PCR primers that will allow us to estimate isoform-specific expression using qPCR.

Lastly, aside from differences in expression of isoform use, patterns of methylation could be important in the development of extreme osmoregulation - indeed, methylation has been shown to be important to many other complex phenotypes including behavior (Lyko et al., 2010), metabolism (Foret et al., 2012), and physiologic stress (including heat stress) response (Sonna et al., 2002). To understand patterns of methylation, a large bisulfite sequence dataset will be generated, which will contain information from every individual included in the mRNAseq experiments, described above. This dataset will allow for the understanding of another layer of genomic complexity not typically available to researchers conducting RNAseq experiments in isolation. Importantly, in addition to

enhancing our understanding of the mechanisms underlying dehydration tolerance, phenotypes related to differential methylation may be prime therapeutic targets.

Preliminary Data: To date, the lab has generated a RNAseq dataset that consists of approximately 30M 150nt SE Illumina reads from the same 2 animals housed in the 'cold/simulated rain' treatment group from which physiology data was collected. We have generated *de novo* transcriptome assemblies as well as mapped to the reference genome. Though the scope of the analyses is preliminary, the results are interesting. 99.7% of the RNAseq reads map to the genome, with over 73% mapping concordantly. This suggests that the content of the draft genome is complete and genomic contiguity is high. We have recovered many of the aquaporin genes, as well as many other critical genes including vasopressin and its receptor, Renin, Angiotensin, Angiotensin Converting Enzyme, as well as the genes that code for the natriuretic peptides. We have estimated expression for all transcripts. Interestingly, within the aquaporin genes, Aquaporins 1 and 2 had the highest expression, while expression of Aquaporin 5, 9, 10 and 12 were undetectable, though they are present in the genomic reference.

Expected Outcome: Upon completion of Aim 1, we will have a synthetic understanding of the physiologic and genomics patterns associated with extreme osmoregulation. These data will allow us to generate a list of genes, genomic regions, isoforms, and methylation states putatively linked to the phenotype of interest. This list is critical, and will form the basis for our first R01 submission, which will propose the development of a system where manipulation of specific genes is possible (e.g. the CRISPR/CAS9 transgenic system), thus moving the work from correlation to causation. This grant will be developed and submitted during the second year of the COBRE tenure. In addition to this, the completion of Aim 1 will allow us to become more proficient in the collection and bioinformatic analysis of physiology data. Lastly, part of Aim1b involved the development of a novel pipeline for the identification of differential isoform use using PacBio RNA sequence data. This skill will be useful to the investigator's broader scientific goals, as well as to the broader scientific community.

Regarding dissemination, the work will be published in open access journals, after rapid release using preprint servers. We envision several papers that are a direct result of this work, include papers describing the physiological and metabolic response to water deprivation as well as their genomic responses. In addition, we aim to publish a more methods-oriented paper surrounding the study of isoform using PacBio data. Aside from peer-reviewed publication, results will be disseminated via social media, the PI's blog, and at the annual meeting of the Society for the Study of Evolution.

Aim 2: Given the transition from the obligate intake of fluids as infants, to its complete absence later in life, the ontogeny of physiologic water conservation will be elucidated.

Background: Given that desert adapted mice, capable of surviving without water are as neonates dependent on liquid intake, the study of the ontogeny of physiologic water conservation is extremely interesting and relevant to the current work. The study of gene expression in renal, pulmonary and hypothalamus tissue types along the transition from the intrauterine environment through birth in the context of differences in oral fluid intake is remarkably novel and will yield unique insights into physiologic water conservation. Although several studies have assayed renal gene expression in neonates, these studies have typically been limited to a small number of genes in a specific context (e.g. hypertension (Sampson et al., 2012; Shanmugam, 1996)).

In addition to the transition from lactation-dependence through weaning, an even more fundamental tran-

sition happens at birth, which is accompanied by substantial changes in renal physiology. In utero, the large volumes of dilute urine are typically produced (Wintour and Moritz, 1997) while post-birth, relatively small volumes of concentrated urine are typical. While this trend appears to be canonical in well-characterized mammal models, whether the fetuses of water-stressed desert-adapted mice follow this trend is unknown, and may be extremely revealing in the context of dehydration resistance. The proposed work aims to use a genome wide approach to characterize these transitions in desert-adapted rodents. This work will provide novel insights into fundamental biological processes, bearing hard upon dehydration resistance, a phenotype which could, if translated to biomedical intervention, save millions of lives annually.

Research Plan: This phenomenon will be explored using fetal and neonatal mice whose mothers are exposed to treatments and an abbreviated set of methods listed in Aim 1. Many of the physiological measurements (e.g. blood and urine analyses) will be impossible to collect in very young animals secondary to sample volume requirements, though a full battery of genomic tests will be possible. To evaluate the ontogeny, five fetal and neonatal mice will be culled per treatment at four different time-points (immediately prior to birth, 2 hours after birth, mid-lactation (approximately 10 days after birth), 1 day after weaning). These time-points have been chosen as together they will allow us to assay the breadth of developmental stages. We hypothesize that patterns of gene expression, methylation, and isoform use will resemble those common in conditions where water is available *ad lib*, though the novelty of this aspect of the study limits firm predictions.

Expected Outcome: Upon completion of Aim 2, we will have a synthetic understanding of the genomics patterns associated with the ontogeny of extreme osmoregulation. These data, together with the data associated with Aim 1 will allow us to generate a list of genes, genomic regions, isoforms, and methylation states putatively linked to the phenotype of interest. This list is critical, and will form the basis for the first R01 submission, which will propose the development of a system where manipulation of specific genes is possible (e.g. the CRISPR/CAS9 transgenic system), thus moving the work from correlation to causation. This grant will be developed and submitted during the second year of the COBRE tenure.

Table 2: Timeline

Activity	FY2015	FY2016	FY2017
Recruit PDF, grad students, undergraduates	X		
Increase Colony Size & ID animals for experiments	X		
Conduct Physiology Experiments -- AIM 1A	X		
Collect & Analyze expression data -- AIM 1B	X		
Analyze Bisulfite and PacBio data -- AIM 1B		X	
Animal breeding in prep for Aim 2		X	
Collect & Analyze genomic data -- AIM 2		X	X
Write papers & submit		X	X
Present results at international conference		X	X
Prepare R01 & and resubmit as needed		X	X
Train Undergrad, Grad students, & PDF	X	X	X
Disseminate info	X	X	X

iv. Relationship to GEBRI program

The theme of the proposed COBRE (Genome-Enabled Biomedical Research Institute) focuses on the development of genomic resources for novel systems with important relevance to critical issues in human health. Obviously, the work proposed here fits ideally into this theme. The most important functions for the GEBRI to the long term success of this work include:

1. The formation of a collaborative Center bringing together faculty with shared interests in the application of genomics of novel systems, including opportunities to interact with colleagues with and outside UNH. PI MacManes is already involved in overseeing the weekly genomics group meetings, and has lead a group charged with creating a bioinformatics computer core. In addition, MacManes is serving on 4 graduate student thesis committees, providing genomics and bioinformatics expertise for each.
2. The opportunities for mentoring described below represent a valuable resource. The PI is already participating in the Research Engagement Academy and looks forward to participation in the Up-2-NIH program as well as the ongoing mentorship of PIs Thomas and Cote. This mentorship will greatly enhance the efficiency and productivity of PI MacManes.
3. Access to high quality genomics and bioinformatics infrastructure and expertise. PI MacManes brings significant experience and expertise in bioinformatics to the GEBRI. To this end, the MacManes lab has already released a transcriptome assembly pipeline (<http://sourceforge.net/projects/tamrs/>) and automated quality control software (<http://sourceforge.net/projects/qcpro/>). In addition to this, the PI MacManes is an active developer of the transcriptome assembly program Trinity (Haas et al., 2013) and annotation software Trinotate (<http://trinotate.sourceforge.net/>). MacManes also has recognized expertise in genome assembly, having participated in an international research consortium charged with improving methods surrounding genome assembly (Bradnam et al., 2013). MacManes will interact with other GEBRI participants (especially PI Jeong and PI Plachetzki) as well as internal and external partners specifically focusing on *de novo* genome and transcriptome assembly.

v. Mentorship Plan

The mentorship program is divided into three layers with an overlapping temporal sequence:

Research and Engagement Academy

Goals: The Office of the Senior Vice Provost for Engagement and Academic Outreach (Julie Williams) and the Office of the Senior Vice Provost for Research (Jan Nisbet) offer the Research and Engagement Academy at UNH. This program is designed to advance faculty scholarship through external funding which is aligned with the UNH strategic plan, including a specific focus on multidisciplinary opportunities. The Research and Engagement Academy supports UNH faculty (tenure-track, extension and research) interested in enhancing their scholarly agenda through training and critical feedback.

The Academy is a semester-long experience which includes a series of workshops, coaching through the grant writing process, and interacting with colleagues about successful strategies with federal agencies and foundations. To be considered for the Academy, the faculty member must secure a nomination from their Dean/Director and submit an online application. Faculty members from all academic disciplines are encouraged to apply. In the last three years, 63 UNH faculty have participated in the Academy. More information and a link to the online

application can be found at www.unh.edu/engagement/research.

Curriculum: Seven (4-6 hours) workshops over the course of the spring semester including: Interaction with successful faculty on panels and in presentations, coaching through the grant writing process, proposal planning and review prior to submission, interaction with program officers and/or Program Officials, and mock panels. Expected outcomes include the submission of a competitive proposal. Evaluation of the program is accomplished by a longitudinal study to assess the impact of programmatic interventions

Up-2-NIH

Up-2-NIH is a UNH program to support selected faculty interested in pursuing funding from the National Institutes of Health. The justification to establish this NIH specific program run parallel with COBRE mission to increase NIH funded research programs in our state. It was recognized that many UNH faculty have research programs in health, biomedical and behavioral sciences, yet their success in competing for extramural funding from the National Institutes of Health (NIH) has been limited. To address UNH's relative lack of success in securing NIH awards, Senior Vice Provost for Research (SVPR) staff and a representative group of UNH faculty with NIH experience have proposed a strategy to increase our faculty's competitiveness to receive NIH funding for their research. This program takes selected faculty with interests in securing funding from NIH. They receive focused support from the Research Office over a 12-month period. This support includes a small seed grant, NIH-experienced faculty mentors and a full year program of workshops, and projects focused on the development of a competitive proposal including mock NIH panels. Participants are selected through an application process that occurs in the spring and we expect all COBRE faculty to participate. We anticipate full participation of our junior COBRE faculty. Applications are submitted in the spring for fall enrollment.

Mentorship in the Genome-Enabled Biomedical Research Institute

While our experience with the first two layers has been excellent and will play an important role in the development of our faculties programs, the goals of our COBRE program include a purposeful multidisciplinary approach. To accomplish this, we will pair our junior faculty with mentors that represent the breadth of discipline that underpin their programs. The challenges faced by multidisciplinary research can be daunting. These include the need to design and target research proposals to the appropriate institutes while maintaining the intellectual integrity of the science. In this last layer of mentoring, these faculty will interact with their mentors with a focus on opportunities for enriching their knowledge. We specifically expect that the mentors will work with the PIs to hone their research strategies, including frequent review of progress with discipline-specific feedback on publication strategies, target institutes, and meeting and workshop opportunities.

The expected outcome of this NIH specific mentoring program is the submission of an NIH grant under one of the funding mechanisms. Explicit in the mentorship program are a set of milestones focused on the submission of competitive applications. The development of the individual research projects for this proposal is the first milestone expected from participation in the Research Engagement Academy. The next milestone will be the submission of an NIH grant one year after the fall. Depending upon the development of preliminary data and specific progress of current aims, the mechanism may vary from R01 to R15 or R21 mechanisms.

Expectation of mentors: Within each of these programs, mentors are expected to provide written feedback

shared with the mentees. Conversely, we will seek evaluations of mentorship progress each month as a standard component of our monthly meetings.

Research Design: The administrative framework for the Research Engagement Academy and Up-2-NIH will be coordinated with the office of the Vice Provost for Research and the Vice Provost for Engagement. These programs have established administrative support provided by the institution. The PI Thomas and Co-Investigator Cote both participate in the REA and will coordinate efforts with that administrative staff.

COBRE specific mentor program will be evaluated based on a set of explicit milestones established by each investigator in the context of their projects and the mentoring plan. These include: (1) Preliminary data that will significantly improve future grant proposals. (2) Significant new skills and demonstrated abilities. (3) Publications (4) Participation in national and international meetings. (5) Grant submissions.

vi. Milestones

In addition to the primary scientific goals described above, the proposal has significant implications for the investigators professional development and preparedness for the submission of an R01. Specifically, this proposal will allow us to collect the critical data that will make for a competitive R01 application. Here, detailed physiological data combined with a deep characterization of the genomic underpinnings will set the stage for functional testing using emerging technologies like the CRISPR/CAS9 system. We aim to complete a substantial amount of the data collection and analysis by year 2, such that a competitive R01 will be submitted.

Regarding skills, this proposal will provide opportunity to develop new skills mostly as a product of cross-pollination between the various GEBRI investigators as well as internal and external partners. Specifically, interacting with PIs Thomas and Cote, experts in bioinformatics and molecular biology respectively will provide the investigator with important technical skills that will facilitate the next stages of his research program, the functional testing of specific candidate genes. Regarding this, collaboration with PI Plachetzki will allow us to be better prepared to implement the CRISPR/CAS9 transgenic systems in our organism.

Dissemination of work is something about which PI MacManes cares about deeply. As a scientist of Native American Descent, I recognize the links between inclusion and dissemination. To this end, the lab publishes exclusively in open access journals, and engage the general public prior to publication via a blog, twitter, and via the use of a preprint server. Using these modes of communication, several publications will result including descriptions of the physiology, of gene expression, isoform use, and methylation in water restricted individuals. With regards to the data produced as a result of the proposed work, we maintain an open access policy- all data and code used in analyses are immediately released under a CC0-BY license. Lastly, in addition to written dissemination, the lab attends the international Evolution meeting yearly, and given available funds, will attend the annual Society for Molecular Biology and Evolution meeting.