

i. Conceptual Framework, Introduction & Background

Every day, soldiers serve in desert environments, forced to endure the intense heat and aridity inherent to these environments. Indeed, these environmental stressors may in fact pose a greater threat than that of enemy combatants. While billions of dollars have been spent on protecting soldiers from bullets, far less attention has been paid to the more insidious threat of heat and dehydration. What if there was a way to significantly enhance the performance and safety of our soldiers by reducing the physiological need for water, particularly in desert environments? While humans and most other mammals are exquisitely sensitive to dehydration, there are many animals, those having evolved in deserts, that are capable of living without ever drinking water. This basic science research proposal aims to understand the genetic and genomic underpinnings of survival without water in a desert-adapted rodent native to the Southwest United States, *Peromyscus eremicus*. Our long-term goal includes developing the ability to recapitulate the phenotype in non-desert adapted mammals, including humans.

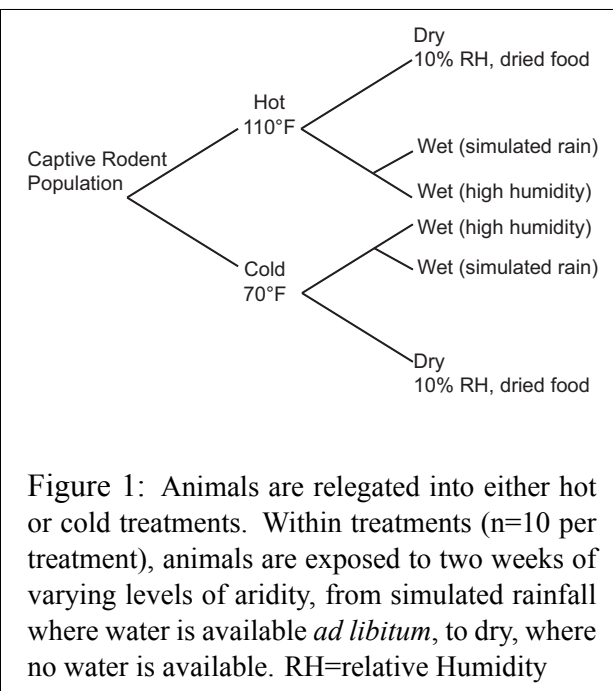
The maintenance of water balance in animals is one of the most important physiologic processes, and is critical to desert survival. Indeed, humans and other mammals are exquisitely sensitive to changes in osmolality, with slight derangement eliciting physiologic compromise. When the loss of water exceeds dietary intake, dehydration - and in extreme cases, death - can occur. Unlike most mammals, animals living in desert habitats are subjected to long periods of extreme heat and intense drought. As a result, desert animals have evolved mechanisms through which physiologic homeostasis is maintained despite severe and prolonged dehydration. **The proposed research uses a novel approach integrating physiology, genomics, and computational biology to better understand how animals thrive in what appears to be unsurvivable conditions.** This integrative research program will significantly enhance our understanding of the physiologic processes underlying osmoregulation in extreme environments, which is the critical 1st step in developing therapies that will enhance soldier performance and safety.

Specifically, I propose to study extreme physiologic water conservation and dehydration tolerance using a captive colony of *Peromyscus eremicus* rodents native to the desert Southwest United States. These rodents are housed in a specially designed walk-in desert chamber. This chamber will replicate the intense heat and aridity of the natural environment, while preserving the ability to manipulate the relevant variables (temperature, humidity, water availability) to meet experimental needs. For animals exposed to various experimental treatments (Figure 1, n=10 per treatment), I will collect and analyze physiologic data relevant to hydration status (e.g. serum electrolyte levels), and kidney function (e.g. BUN, Creatinine). In addition to this, because water requirements are related to metabolic rates I will collect data related to metabolism using a chamber designed to function in extreme desert conditions. Lastly, though most individuals are functionally anuric, urine will be collected when available and analyzed for specific gravity and osmolality. In addition to the thorough characterization of physiology related to desert survival, I will characterize the genomic changes that underlie changes in physiology.

The study of adaptation, or the process through which animals become fitted to their environment has intrigued researchers for decades (1; 2), though only recently have we had the ability to study the underlying genomic mechanisms. Interestingly, researchers interested in understanding the genetics of adaptation have the ability to ground modern studies of genetics on decades of work aimed at understanding the ecological context within which adaptation occurs. One particularly salient example of the connection between studies of ecology and natural history and modern ge-

nomics can be found in the study of physiologic adaptation to desert conditions. Here, remarkable physiologic, morphologic (3; 4; 5; 6) and behavioral (7) adaptation has been studied in the context of desert ecology. These studies provide a rich context for the current work, which aims to understand the links between physiology and genomics in rodents able to thrive in amongst the most harsh of conditions on Earth. Ultimately, this work aims to provide interventions that specifically enhance the performance and safety of our troops serving in desert conditions by reducing or eliminating the untoward effects of dehydration.

Lastly, the genomic processes related to desert survival have yet to be characterized. The few studies of genetics that have been done have focused on the role of single members of the Aquaporin gene family (but see (8)), which are large membrane-bound proteins that are critically involved in renal water transport (9; 10; 11; 12). These studies have shown that changes in Aquaporin (AQP) protein abundance and expression may be related to water availability (13; 14; 15). In addition to changes in expression, another study showed that the AQP4 pathway was completely lost in the desert rodent *Dipodomys merriami merriami* (16). 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.



ii. Experimental Plan and Preliminary Results

Specific Aim 1: To fully understand the physiological effects of desert conditions on rodents, I will relate multiple physiological variables to differences in temperature, relative humidity, and water availability. These experiments (and those described under Aim 2) are fundamentally linked to a series of environmental manipulations, described in Figure 1. I hypothesize that, as a result of unique mechanisms related to solute and water balance, that average 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.

From all animals in experimental treatment groups (n=10 * 6 treatments=60 animals), I will collect a urine sample, and measure specific gravity and urine osmolality using an Atago UG- α urine refractometer. Serum electrolytes including potassium, sodium, BUN, creatinine, calcium and bicarbonate ion concentration will be measured, as will carbon dioxide production and oxygen consumption.

Preliminary Data: To date, I have generated a physiology dataset that consists of serum electrolytes and urine osmolality from 5 animals housed in the 'cold/simulated rain' treatment group.

This small dataset allowed me to refine and perfect the protocols for data collection and analysis, and has already suggested interesting physiology. In the small subset of animals housed in the cool, wet, simulated rain conditions for a period of 3 weeks, all had serum Potassium levels >8.0, Creatinine <0.3, and Sodium >151, which is significantly different than 'normal' levels in humans and other mammals.

Specific Aim 2: I will understand the genetic response to extreme heat and aridity via a series of bisulfite sequencing and RNAseq experiments, and will link these patterns to individual physiologic state as defined in Aim 1. Differences in renal methylation patterns and mRNA gene expression will be tested using the 10 individuals per treatment group (n=60 total, specified above).

For analysis of the bisulfite sequence data, an accurate genome assembly is required. Using the existing draft genome (sequenced using startup funds) as a starting point, I will complete the genome assembly of the primary study organism, *Peromyscus eremicus*. This process will be significantly aided by leveraging the existing genome sequence data available from several other *Peromyscus* species. The genome will be validated by use of the CEGMA (20) and REAPR (21) algorithms. In general, this approach has been shown to be successful in producing high quality draft assemblies, though multiple assembly methods will be evaluated as per the findings of my previous work (22).

The renal transcriptome for each individual included in the physiology study in aim 1 will be sequenced. Specifically, we will map sequencing reads to the reference genome, and estimate expression for each transcript using a statistical model implemented in the software eXpress (?). We will evaluate differential expression and methylation using the software DESeq (?). Genes and regions that are significantly differentially expressed or methylated after appropriate statistical control will be further evaluated for allelic variation (SNPs) that may be potentially causative.

Preliminary Data: To date, I have generated a RNAseq dataset that consists of approximately 30M 150nt SE Illumina reads from the same 5 animals housed in the 'cold/simulated rain' treatment group from which I collected physiology data.

iii. Budget Description

Personnel: The project described above requires significant personnel needs. PI MacManes is responsible for aspects this proposal. Specifically, he will oversee the design and execution of the genetic components of aim 1 and 2, which involves extraction of RNA from tissue samples, generation of mRNA sequencing libraries, sequencing on the Illumina HiSeq 2500 platform, and all bioinformatic analysis. MacManes will train the graduate student, and will interact with other project participants (undergraduate students). 2 months of summer salary per year is requested for all three years of the project. Salary is increased annually by 3% to account for inflation, from a base rate of \$74000. In addition to the senior personnel, a PhD student and undergraduate student will be paid \$7000 and \$4500 respectively for 10 weeks of full time work on the proposed project.

Materials: Direct costs will sum to approximately \$60000 per year. These costs will be used for, in approximate order of expense: Illumina RNA sequencing, Illumina DNA sequencing (for studies of methylation), Illumina sequencing library generation, general lab consumables, computer hardware upkeep, publications charges and general office supplies that directly related to the proposed project.

Indirect Costs and Fringe Benefits: Facilities and Administration is charged at 47.5% of MTDC based UNH's federally negotiated rate. Fringe benefits for summer salaries are calculated at 7.7% per UNH's federally negotiated rates.

Total Project Costs: Total project costs are expected to be approximately \$416,000 over 3 years.

References

- [1] Darwin C (1859) *On the Origin of Species by Means of Natural Selection or the Preservation of Favored Races in the Struggle for Life* (J. Murray, London), 1st edition.
- [2] Fisher R (1930) *The Genetical Theory of Natural Selection* (Oxford University, London).
- [3] Dickinson H, Moritz K, Wintour EM, Walker DW, Kett MM (2007) A comparative study of renal function in the desert-adapted spiny mouse and the laboratory-adapted C57BL/6 mouse: response to dietary salt load. *American Journal of Physiology-Renal Physiology* 293:F1093-8.
- [4] Huntley AC, Costa DP, Rubin RD (1984) The contribution of nasal countercurrent heat exchange to water balance in the northern elephant seal, *Mirounga angustirostris*. *The Journal of experimental biology* 113:447--454.
- [5] Schmidt-Nielsen B, Schmidt-Nielsen K (1950) Pulmonary Water Loss in Desert Rodents. *American Journal of Physiology* 162:31--36.
- [6] Schmidt-Nielsen B (1952) Renal Tubular Excretion of Urea in Kangaroo Rats. *American Journal of Physiology* 170:45--56.
- [7] Nagy KA (1994) Seasonal Water, Energy and Food Use by Free-Living, Arid-Habitat Mammals. *Australian Journal of Zoology* 42:55--63.
- [8] Bartolo RC, Donald JA (2007) The distribution of renal hyaluronan and the expression of hyaluronan synthases during water deprivation in the Spinifex hopping mouse, *Notomys alexis*. *Comparative biochemistry and physiology Part A, Molecular & integrative physiology* 148:853--860.
- [9] Kwon TH, et al. (2009) Aquaporins in the kidney. *Handbook of experimental pharmacology* 190:95--132.
- [10] Verkman AS (2002) Physiological importance of aquaporin water channels. *Annals of medicine* 34:192--200.
- [11] Brown D, Katsura T, Kawashima M, Verkman AS, Sabolic I (1995) Cellular distribution of the aquaporins: A family of water channel proteins. *Histochemistry and Cell Biology* 104:1--9.
- [12] Nielsen S, Agre P (1995) The aquaporin family of water channels in kidney. *Kidney international*.
- [13] Bösel I, et al. (2009) Involvement of the V2 vasopressin receptor in adaptation to limited water supply. *PLOS ONE* 4:e5573.

- [14] Gallardo PA, Cortés A, Bozinovic F (2005) Phenotypic flexibility at the molecular and organismal level allows desert-dwelling rodents to cope with seasonal water availability. *Physiological and Biochemical Zoology* 78:145--152.
- [15] Bozinovic F, Gallardo PA, Visser GH, Cortés A (2003) Seasonal acclimatization in water flux rate, urine osmolality and kidney water channels in free-living degus: molecular mechanisms, physiological processes and ecological implications. *Journal of Experimental Biology* 206:2959--2966.
- [16] Huang Y, et al. (2001) Absence of aquaporin-4 water channels from kidneys of the desert rodent *Dipodomys merriami merriami*. *American Journal of Physiology-Renal Physiology* 280:F794--F802.
- [17] Maccallum I, et al. (2009) ALLPATHS 2: small genomes assembled accurately and with high continuity from short paired reads. *Genome Biology* 10:R103.
- [18] Simpson JT, Durbin R (2012) Efficient *de novo* assembly of large genomes using compressed data structures. *Genome Research* 22:549--556.
- [19] Miller JR, et al. (2008) Aggressive assembly of pyrosequencing reads with mates. *Bioinformatics (Oxford, England)* 24:2818--2824.
- [20] Parra G, Bradnam K, Korf I (2007) CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics (Oxford, England)* 23:1061--1067.
- [21] Hunt M, et al. (2013) REAPR: a universal tool for genome assembly evaluation. *Genome Biology* 14:R47.
- [22] Bradnam KR, et al. (2013) Assemblathon 2: evaluating *de novo* methods of genome assembly in three vertebrate species. *GigaScience* 2:10.
- [23] Haas BJ, et al. (2013) *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature protocols* 8:1494--1512.
- [24] Grabherr MG, et al. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29:644--652.
- [25] MacManes MD, Eisen MB (2013) Improving transcriptome assembly through error correction of high-throughput sequence reads. *PeerJ* 1:e113.
- [26] Cantarel BL, et al. (2008) MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Research* 18:188--196.