Fellowship Training

I. Title and Abstract:

Title: Integrated metabolomic and genomic approach to metabolic variation across horse breeds

Rationale: Equine Metabolic Syndrome (EMS) is defined by a clustering of clinical signs, namely hyperinsulinemia, insulin resistance and adiposity, which predispose horses to the development of laminitis. Certain breeds appear to be more susceptible to EMS, while other breeds seem to be at lower risk. Genetic selection in horses has resulted in metabolic and athletic phenotypes that allowed horses to efficiently perform different types of work. These underlying metabolically efficient, or "thrifty" alleles, in particular those that regulate energy intake, storage and use, coupled with changes in equine husbandry practices in recent times, including dense high caloric feed and limited exercise, may explain much of the increasing prevalence of equine metabolic disease phenotypes in modern environments.

Hypothesis/Objectives: We ***hypothesize*** that breed differences in key metabolic phenotypes are due to high frequencies of alleles that modify metabolic traits. Our ***objectives*** are to **1)** further dissect the metabolic differences between breeds at the molecular level using total serum metabolite profiling; and **2)** use these breed-specific molecular metabolic profiles to identify candidate genes underlying breed metabolic differences.

Study Design: In **Objective 1** serum metabolite profiles will be analyzed before and after an oral sugar test in 274 horses from 5 breeds with distinct metabolic phenotypes. These data will be used to **1)** identify the metabolites and metabolic pathways that are significantly different between breeds using functional annotation and mapping of metabolites to known pathways, pathway and metabolite set enrichment analysis, network analysis, and pathway activity profiling; and **2)** to correlate the metabolite/pathway differences with the previously identified breed differences in key hormonal and biochemical measurements. In **Objective 2** candidate genes responsible for metabolic differences between breeds will be identified by using high-density SNP genotype data to locate genomic regions and specific haplotypes that are highly differentiated between breeds. Metabolites and metabolic pathways that are significantly different between breeds will then be used to provide context for narrowing the focus to specific candidate genes within these genomic regions.

Preliminary Data: Our data demonstrate significant differences among Morgans, Arabians, Welsh ponies, Tennessee Walking Horses and Quarter Horses in EMS-defining metabolic traits (e.g. insulin dynamics, lipid metabolism, adipokines) that mirror EMS risk. We have also demonstrated the use of SNP genotype data to identify regions of breed differentiation/selection across a wide breed panel, and across Quarter Horses and Welsh Ponies, and provided examples in the *GYS1* and *MSTN* genes that selection for certain performance traits results in near-fixation of alleles that alter energy metabolism. Finally, we have demonstrated the potential for serum metabolomic data to lead to insight into the metabolic differences between horses and clinical phenotypes.

Expected Results: We expect to detect major breed differences in metabolites/metabolic pathways that are correlated to previously identified biochemical and hormonal differences. We also expect to identify and prioritize candidate genes within genomic regions of interest that influence a spectrum of metabolic traits, particularly the susceptibility to metabolic syndrome.

Budget and Timeline: [redacted] Approximately one year is required to achieve each objective.

Potential Impact for Animal Health: Elucidating the evolution of the genetic basis of metabolic efficiency and metabolic syndrome is a novel, unexploited approach to the study of the genetic basis of obesity, energy dysregulation and EMS. This project will provide novel insights into disease biology, allowing the identification of new therapeutic targets, and increasing our understanding of the pathophysiology of EMS and its associated clinical features. Moreover, the identification of genes underlying the EMS phenotype will also directly impact equine health by allowing for the development of genetic tests to identify horses at risk for the development of obesity and laminitis prior to the onset of clinical disease.

July 15th, 2015

To the Scientific Advisory Board – Fellowship Training Proposal

I am a Postdoctoral Research Associate in the Equine Genetics and Genomics Laboratory at the University of Minnesota under the supervision of Drs. Molly McCue and James Mickelson. I have been working in this laboratory since the Fall semester of 2014, with the intent of taking further steps towards a successful career in animal research. My goal is to obtain a tenure-track position at a major institution where I can establish a laboratory to conduct research in the field of large animal genetics and genomics, as well as be involved in teaching. My professional vision is to be part of a strong and collaborative program that provides me the opportunity to perform cutting-edge research in animal health, disease and performance and to learn constantly from my peers. I also have a passion for teaching, and would like to continue working with students both in the classroom and in the laboratory, throughout my career.

I obtained my Bachelor's degree in Biological Sciences from the University of Brasilia (Brazil) in 2005, and my Ph.D. in Biomedical Sciences from Texas A&M University in 2014. I understood the importance of research early on in my career, so while I was an undergraduate in Brazil, I was an intern at the Laboratory of Animal Molecular Genetics at EMBRAPA, Brazil's Federal Agricultural Research Institute. During that period, I worked on the genetics of bovine and sheep reproduction, and this experience not only consolidated my passion for animal research, but also provided me with knowledge that laid the foundation for my subsequent endeavors. After that, I did another internship at the Laboratory of Uterine Biology and Pregnancy at Texas A&M University, where I continued working on the genetics of sheep reproduction. Then, I decided to come back to the United States to pursue a graduate degree in animal genetics.

My doctoral studies, conducted in the Molecular Cytogenetics and Genomics Laboratory at Texas A&M University under the supervision of Dr. Terje Raudsepp, included generating a whole genome integrated map for the alpaca (*Lama pacos*). This research project, funded by Morris Animal Foundation [redacted] allowed us to develop a genome-wide set of molecular markers that successfully integrated the alpaca genome sequence assembly with the physical chromosomal maps for this species. As part of this research project, and also with funds obtained from a *Veterinary Student Scholars Program Grant* from Morris Animal Foundation awarded to me in 2011, we successfully mapped candidate genes involved in deafness associated with depigmentation in alpacas. Therefore, it is safe to say that Morris Animal Foundation played an instrumental role in the successful completion of my doctorate, and for that I am very grateful.

During my undergraduate and graduate careers, I was either the author or co-author of a total of 24 publications in national and international conference proceedings, as well as 12 articles in international peer-reviewed journals. Moreover, I was fortunate to win several honors and awards including *Outstanding Graduate Student* and *High Impact Achievement* awards from the Texas A&M College of Veterinary Medicine, and the *U.S. Senator Phil Gramm* fellowship for excellence in research, teaching and as a scholar-mentor, among others.

Up until this point, my research background has been very diverse, ranging from the annotation of transposable elements in a nematode species to my current research in the genetic basis of EMS, including bovine, sheep and camelid studies. I believe that my experience has provided me with a skillset that will allow me to pursue excellence in all aspects of my postdoctoral career, including research, grant and manuscript writing, and communicating my findings to fellow researchers and to the public.

My postdoctoral research project aims at identifying candidate genes associated with Equine Metabolic Syndrome, or EMS. The opportunity to work with equine genetics provided me with invaluable knowledge on the potentially harmful outcomes of genetic diseases in horses. Moreover, I was able to understand the importance of preventative approaches to genetic diseases, especially in the case of EMS, which can lead to fatal consequences in affected horses. The identification of candidate genes associated with this syndrome is an important first step towards early diagnostics and preventative measures that can be taken prior to the development of clinical signs. I

believe it also lays the foundation for the discovery of the pathophysiology of the disease, as well as of targets for novel therapies that can greatly improve the health and welfare of potentially affected individuals.

Therefore, the research study presented in this Fellowship Proposal aims at identifying candidate genes associated with the EMS phenotype and breed-specific metabolic profiles. In order to accomplish that, we will identify genome-wide signatures of positive selection in four breeds with distinct metabolic profiles (Arabian, Morgan, Quarter Horse and Welsh Pony) using genotyping data for 2 million SNP markers. Then, we will use a comprehensive set of phenotypic, environmental and epidemiologic measurements, as well as breed-specific global metabolic profiles derived from serum metabolomics analysis, to refine and prioritize the prioritize genes and pathways associated with EMS phenotypes and distinct serum metabolic profiles in horses and ponies.

We anticipate that these findings will be critical to gain further knowledge on the genetic basis of this syndrome, and for our long-term goal of developing genetic tests for early EMS diagnosis. I am confident that our laboratory, which is well-known for its solid research history and use of cutting-edge technologies, will provide me with the ideal facilities and human resources to conduct the proposed research study. My mentor, Dr. Molly McCue, is a very successful DVM/PhD scientist who has dedicated years to studying the genetics of equine metabolic diseases, having published her numerous findings in high impact journals. My co-mentor, Dr. James Mickelson, is a well-established researcher who helped establish the field of equine genetics and genomics, having played a significant role in developing and using the first ever genetic analysis tools in the horse.

With the help of previous Morris Animal Foundation funding, my interest and passion for animal research have been translated into scientific knowledge and experience that allowed me to develop my career up until this point. Continued support from Morris Animal Foundation will allow me to pursue my newfound interest in the genetic basis of EMS, as well as support me in building my skills in equine genetics and taking further steps towards developing my career as an independent investigator. Thank you in advance for your consideration of my proposal.

Sincerely,

Felipe Avila, Ph.D.
Equine Genetics and Genomics Laboratory
College of Veterinary Medicine
University of Minnesota

UNIVERSITY OF MINNESOTA

Twin Cities Campus

Veterinary Population Medicine
College of Veterinary Medicine

225 Veterinary Medical Center
1365 Gortner Avenue
St. Paul, MN 55108
Office: 612-625-7755

Dear members of the Large Animal Scientific Advisory Board,

It is my distinct pleasure to write this mentor letter for Dr. Felipe Avila in support of his Morris Animal Foundation Postdoctoral Fellowship proposal. I had the opportunity to recruit Felipe to our research group in the fall of 2014. Since that time I've had the opportunity to become familiar with Felipe's strengths and have watched him step in and become an integral part of our research team.

Felipe came to our group with some background in reproductive genetics and cytogenetics but with genuine interest in learning how to take advantage of high dimensional data sets and 'omic technologies. The work done in our lab and the focus of Felipe's research thus far has required Felipe to acquire an entirely new set of skills in the last ten months. His motivation and dedication to this task has been inspiring. It is not necessarily an easy transition for a bench top scientist that has been heavily invested in molecular work to step away from pipettes and into the world of computational biology. This transition has required Felipe to learn core computing skills including some basic programming. While this transition results in a steep learning curve with many daily opportunities for frustration, Felipe has maintained a remarkably positive attitude and has been tenacious in his efforts.

Felipe has taken two large projects previously overseen by a senior postdoc in the laboratory and he has moved both of these projects forward substantially, presenting at the International Plant and Animal Genome meeting barely 4 months into his post-doc and having an abstract accepted for the Havemeyer Equine Genomics meeting in July 2015. These projects, which focus on the identification of genomic regions highly differentiated between breeds or undergoing positive selection, form the basis of the genomic preliminary data in Felipe's post-doc proposal. As part of this work, Felipe has learned how to manipulate and analyze single nucleotide polymorphism (SNP) genotyping data and the analysis of whole genome sequence data, and has developed the genomic analysis pipeline outlined in objective 2a. The work outlined in this proposal builds on the skills he has developed thus far, and his preliminary data, by adding a functional high dimensional data set in the form of non-targeted serum metabolomic data. This functional data will give Felipe the opportunity to explore new statistical analyses and to tie molecular phenotypes to his genomic regions of interest; which will round out Felipe's post-doctoral training.

At the completion of his postdoctoral training, Felipe will truly have a diverse research portfolio, with experiences ranging from genome annotation and FISH mapping to high dimensional data analysis. Further, Felipe will have had the opportunity to perform genetic research in a range of large animal species which fits with his long-term goal of establishing a research lab focusing on large animal genetics and genomics. In addition to his diverse training, Felipe has several other strengths which make him well-suited to become an independent investigator. As mentioned above he is hard-working and tenacious, which is demonstrated not only by his work in my laboratory, but also by his stellar academic record which includes primary or co-authorship on 12 articles in international peer-reviewed journals. He also has excellent people skills, and has shown a talent for teaching/mentoring while working with visiting/rotating PhD students and veterinary students in the lab.

As the primary mentor for Felipe's postdoctoral fellowship I am committed to mentoring Felipe not only on his research projects but also in his professional development. This includes supporting Felipe as he develops his technical writing skills through mentored manuscript and grant writing experiences (such as this one). Further I will support Felipe and his passion for teaching by identifying opportunities for him to become involved in teaching during his postdoctoral training. Finally, as Felipe has mentioned in his training plan, I am committed to helping him network in the professional community through opportunities to speak and professional contacts at meetings, as well as through a number of visiting scientists that have visited/will visit the lab during his tenure here. Our laboratory group meets almost weekly for 1 to 2 hours to discuss current literature, new data analyses, or data from a specific student's project. In addition, I meet with each student in my laboratory including Felipe weekly in a one-on-one meeting to discuss data and/or troubleshoot ongoing projects.

My previous mentoring experience includes being a graduate faculty member in 3 graduate programs at the University of Minnesota, and serve as a faculty mentor for both T32 and T35 awards, and I am the co-director of our post-doctoral T32 in Comparative Medicine and Pathology. I have served/am serving as the primary mentor for 6 PhD students and 3 MS students, 3 post-doctoral trainees, and as the primary clinical mentor for 5 internal medicine residents. I am also currently the co-advisor for 2 PhD students and have served on an additional 10 graduate committees (5 PhD and 5 MS).

Student	Degree	Role	Dates	Current position
[redacted]				

This proposal is the result of Felipe's intellectual contribution, with input from Dr. Mickelson and myself in formulating the research objectives. Felipe produced the initial completed draft of the proposal and we worked together to edit content for clarity and to fit with in the 5 allotted pages—which is a challenge for this novel and sophisticated idea! As the primary mentor I am committed to funding the research portion of this project. All samples have been collected and the money is in-hand to generate metabolomics data from more than half the number of outlined horses. I have pending grant proposals to cover the remaining data generation, and in the event that these pending proposals are not funded, I will commit internal dollars to ensure that Felipe has data to complete this project.

In summary, I would like to reiterate that Felipe is a bright, hard-working and capable scientist who has a bright future ahead of him. He is very deserving of this award and will take full advantage of the opportunities that this fellowship would provide.

Sincerely,

Molly McCue DVM, MS, PhD, Diplomate, American College of Veterinary Internal Medicine
Associate Professor, University of Minnesota College of Veterinary Medicine

UNIVERSITY OF MINNESOTA

Twin Cities Campus

Department of Veterinary and Biomedical Sciences
College of Veterinary Medicine

295 An Sci / Vet Med
1988 Fitch Avenue
St. Paul, MN 55108

612-624-2700
Fax: 612-625-0204

July 6, 2015

Dear Review Committee members:

It is my great pleasure to provide you with this letter in support of Dr Felipe Avila's postdoctoral fellowship application. I have known Felipe for approximately 9 months as a co-mentor of his outstanding work using dense genotyping data from multiple breeds and populations to locate genomic loci of interest in a number of different phenotypes in the horse. I have organized this letter according to your wishes.

Candidate accomplishments, perceived strengths, motivation, academic abilities and potential. Felipe came to us having been trained by Terje Raudsepp who you know as one of the world's most accomplished animal geneticists from her work on horse and alpaca genomics and the genetics of reproduction and infertility. During his time with Terje Felipe made advances in the alpaca genome and solved a problem in their karyotype that puzzled cytogeneticists for decades. Felipe has authorship on 12 publications, with three of them as lead author from his PhD, and I will leave it to you to count the many abstracts and presentations. It all adds up to a beautiful portfolio presenting the first phase of his training towards becoming an independent scientist.

Felipe easily stepped into our group and made an immediate impact. He is very motivated to succeed and learns new technologies very rapidly, enabling him to quickly go from a primarily laboratory based scientist at TAMU to a computational biologist at Minnesota with an entirely different set of tools and an entirely different type of data. He now has a handle on all the necessary approaches to working with dense genotyping data, analyzing it to localize regions of reduced variation, scan these regions for underlying genes and develop databases to catalog gene functions. And, I have no concerns about his ability to work with the metabolomic data that will be generated in his fellowship proposal.

His current work is coming together nicely now in the areas of identifying loci of interest for metabolic traits which in some ways is moving Felipe into a new area that relates more to animal health. We are thrilled that he has already been invited to give platform presentations of his preliminary work at the Plant and Animal Genome conference (Jan 2015) and the Havemeyer Horse Genomics Workshop (July 2015). I wish I had space to tell you about more of the other types of loci he is identifying, such as regions of selection in different sub-types of Quarter Horses that have different performance traits, loci under selection in Thoroughbreds and Standardbreds, and potential size loci in drafts, ponies and miniature horses.

Lastly, here, Felipe is incredibly kind, calm, good-natured, generous, and polite. He is extremely well-organized, logical and well-spoken. All of these traits are apparent in his one on one meetings, his oral presentations to our group meetings, and his talk at the PAG meeting. They also come through in his day to day work where to make progress in the very complex projects he is undertaking requires on organized dedicated workflow with multiple stages going on at once and incredible patience to deal with time consuming computations on computer networks.

To conclude Felipe has all the tools necessary to become an outstanding independent researcher, able to utilize all types of "omics" data, and I have no doubts whatsoever that he will achieve this goal.

Mentoring relationship. I am a co-mentor to Felipe with Molly McCue. Together Molly and I maintain a dynamic, productive and internationally-recognized group of 6 – 8 grad students/postdocs/faculty/techs/vet students in basic and applied equine genetics and genomics. Molly is Felipe's primary mentor with clear strengths in current methods of genome analysis and understanding of equine pathobiology. I consider my role as

providing independent opinions on genomic results and hypotheses in discussions and adding context to results from my earlier expertise in genetics with the basic biochemistry and biology behind many of the phenotypes we study. Our (usually weekly) group meetings mainly consist of directed study on major papers in current topics of genome analysis. In addition Molly and I have weekly joint one on one 30 – 60 min sessions with each student or fellow. I have a similar role in a canine genomics group which can be beneficial as both groups use similar approaches and encounter similar problems and can benefit from each other's work.

History of mentoring, including list of mentees and outcomes. My history of primary and co-mentoring is evidenced below. Many of my trainees hold the DVM degree, all have found exciting positions in academia, government or industry, with a great many of them as independent scientists doing work in animal health and production at research universities.

Trainee	Program	Role	Current Position
[redacted]			

In addition, I have played more than simple PhD thesis committee roles for Eva Furrow (DVM), Raffaella Texiera (DVM), Annette McCoy (DVM), who have received their PhDs and are now in faculty positions at Colleges of Veterinary Medicine, and Nichol Shultz (DVM), Elaine Norton (DVM) and Samantha Beeson (DVM student), who are working towards their PhDs.

Roles in the preparation of this fellowship application and status of research project funding. As with all Fellowship applications this document is a joint product of the trainee and the advisors. Felipe assembled the first draft of the proposal from his readings of the literature, other proposals we have submitted, and his own drafting of his preliminary work. The advisors worked on the first draft (mainly to shorten it!) and ensure clarity and flow. Current funding has generated all data necessary for Objective 2. As indicated in the Support pages, multiple proposals are submitted to complete funding for Objective 1.

Jim Mickelson, PhD, Professor of Biochemistry and Genetics

Training Plan

I have several training objectives that will support my career development and provide me with the necessary skills for success in this proposed study, as well as in my future in animal research. They include development of technical research skills relative to metabolomics and genomics, and non-technical aspects such as writing, networking, developing critical thinking skills, and academic competences.

Research training: based on previous training and experience, my skillset reflects that of a trained molecular cytogeneticist and geneticist. Therefore, I seek to broaden my knowledge by incorporating different 'omics' tools to my scientific repertoire. This fellowship will allow me to further develop my skills on computational biology applied to genomics and dense genotyping data, which I have been learning since I started working in this laboratory, in the Fall of 2014. Moreover, I am thrilled to gain a more in depth understanding of Equine Metabolic Syndrome and its clinical phenotypes, as well as the use of metabolomics to address the pathophysiology of this disease. I plan to learn and successfully apply the techniques listed in this research study, such as: the use of bioinformatics tools to identify genomic signatures of selection; collection and analysis of various metabolic measurements in horses; statistical analyses applied to 'omics' data, among others. For that, I plan to attend workshops at the University of Minnesota's Supercomputing Institute on handling and analyzing big genotyping data and at the Center for Mass Spectroscopy and Proteomics on analysis of metabolomic data.

Technical writing: during this fellowship I will have to opportunity to write scientific reports, abstracts and manuscripts stemming from our findings, with the guidance and input of my mentors, whom have extensive experience with scientific writing. In addition to mentored writing of abstracts, manuscripts, and reports, I plan to receive formal training in grant and manuscript writing through short courses offered at the University of Minnesota, including the "Getting Started" and "Write Winning Grants" workshops, and the "Scientific Writing Series" sponsored by the Center for Translational Science Institute.

Networking and development of critical thinking skills: I believe that the Equine Genetics and Genomics Laboratory will provide me with the ideal environment to establish networks with other scientists, as well as critical thinking skills. Drs. McCue and Mickelson mentor a large group of undergraduate and graduate students, and laboratory technicians from various backgrounds including DVMs and veterinary medicine students. We hold weekly meetings in which we discuss research articles, as well as the progress of our research projects. Being part of a talented, multidisciplinary group is critical to learn various aspects of my project from veterinarians and researchers alike, and to develop critical thinking skills by discussing outcomes and pitfalls, and by incorporating inputs and/or suggestions to my work. Moreover, my mentors have an extensive array of collaborators, with whom we exchange ideas, knowledge, and technologies. I also plan to attend and present my findings at national and international scientific meetings and workshops (e.g. Plant and Animal Genome, Havemeyer Equine Genome Workshop). My mentors will serve as networking facilitators at such meetings, helping me develop independent collaborations which will be useful in my career as a researcher. Finally, it is worth to mention that I will have the opportunity to work with Dr. Susan Van Riper at the University of Minnesota Informatics Institute, a nationally recognized expert on statistical analysis of metabolomic data.

Mentoring plan: Dr. McCue D.V.M., M.S., Ph.D., a board-certified large animal internal medicine specialist (DACVIM), will serve as my primary mentor. Dr. McCue is internationally recognized as a leader in equine genetics and genomics. Her primary research focus is the genetics of neuromuscular and metabolic diseases affecting the horse. Dr. McCue will provide guidance for questions related to the clinical phenotypes and various metabolic measurements, analysis of the large scale datasets (SNP and metabolomic data), statistics, and biologic interpretation of results. She will also provide guidance and critical evaluation for reports, abstracts, and publications related to this work. Dr. Mickelson M.S., Ph.D., as my co-mentor, will provide independent opinions on data acquisition and analysis, adding context to results from his extensive expertise in genetics, biochemistry, and biology. It is worth to mention that the technology used to genotype the animals in this study - the 2 million equine SNP array - was developed by Drs. McCue and Mickelson.

Institutional environment: apart from the human resources available from the talented, multidisciplinary group of mentors and fellow students which I am part of, the laboratory is equipped with all the necessary, cutting edge resources for the completion of this research study. Computing resources available in the lab include 7 Windows computers with a dual installed Linux virtual machine, containing all the software needed for the proposed analyses. Additionally, the laboratory maintains an account with the University of Minnesota Supercomputing Institute (MSI) which includes access to the Computational Genetics and Basic Sciences computing laboratories. Finally, all metabolomics analytical work will be performed at the University of Minnesota Center for Mass Spectroscopy and Proteomics.

UNIVERSITY OF MINNESOTA

Twin Cities Campus

Department of Veterinary Population Medicine
College of Veterinary Medicine

235 Veterinary Medical Center
1365 Gortner Avenue
St. Paul, MN 55108
Phone: 612-625-7755
Fax: 612-625-6241

July 10, 2015

To the Large Animal Scientific Review Board,

It is my distinct pleasure to provide you with this letter in support of Dr. Felipe Avila's postdoctoral fellowship application.

Felipe proposes to continue his postdoctoral training with Drs. McCue and Mickelson, who comprise an established and very successful research team with a history of mentorship of both graduate and postdoctoral students. The team they have formed is a leader in the field of applying state-of-the-art high-throughput genomics approaches to identify and understand genes and mutations that underlie important health, disease and performance traits in horses. Their team is integrated, highly productive, and scientifically rigorous. This team invests heavily in preparing students for technical proficiency, basic sciences comprehension, and grantsmanship. Dr. Avila could not have selected a better program in which to continue his training in a career leading him to a position as an independent scientist in animal health and disease.

Our College's research program emphasizes three Signature Research programs in Comparative Medicine, Emerging and Zoonotic Disease and Population Systems. Animal genomics research (both large and small animal) occupies a key position in the Comparative Medicine program. I would also like to note that the University of Minnesota as a whole has invested heavily in support of genomics, proteomics, metabolomics and computational biology work through its core facilities and infrastructure that has enabled groups such as the equine genomics group to have established themselves and propose higher order and complex projects such as Dr. Avila's. Further, Dr. McCue, the primary mentor for this proposal, is uniquely positioned to help Dr. Avila access these University resources with her intimate involvement with the newly formed University of Minnesota Informatics Institute (UMII) as both a member of the UMII Scientific Advisory Board and as a UMII Transdisciplinary Research Faculty Fellow.

As the chair of Dr. McCue's home department I will do everything possible to ensure that Dr. Avila's training is well-structured and he has the physical resources necessary to achieve his training goals.

To conclude, the college is strongly supportive of this training fellowship, which brings together a strong candidate and a solid mentoring team to work on an important disease in the horse. Please do not hesitate to contact me with any questions.

Sincerely,

Thomas Molitor, PhD
Distinguished Teacher
Professor and Chair,
Veterinary Population Medicine
College of Veterinary Medicine

UNIVERSITY OF MINNESOTA

Twin Cities Campus

Veterinary Population Medicine
College of Veterinary Medicine

225 Veterinary Medical Center
1365 Gortner Avenue
St. Paul, MN 55108
Office: 612-625-7755

July 10, 2015

RE: Salary Verification for Morris Animal Foundation Fellowship Proposal

To Whom It May Concern:

Dr. Felipe Avila is a Postdoctoral Research Associate who joined the Equine Genetics and Genomics Laboratory, Department of Veterinary Population Medicine, in August of 2014. We are requesting an annual salary of [redacted] indirect costs for his position.

The requested salary rate is in line with salaries paid to postdocs in the College of Veterinary Medicine [redacted] This salary rate is also below NIH guidelines for a fellow with his equivalent experience.

Respectfully,

Natalie L. Dillon
Administrative Director
University of Minnesota
Veterinary Population Medicine

III. Resubmission Summary

N/A

IV. Name, Institution and Email Address

Fellow:
Felipe Fagundes de Avila, PhD
Postdoctoral Research Associate
Department of Veterinary Population Medicine
College of Veterinary Medicine
University of Minnesota
[redacted]

Mentor:
Molly McCue, DVM, MS, PhD, Diplomate of ACVIM
Associate Professor
Department of Veterinary Population Medicine
College of Veterinary Medicine
University of Minnesota
[redacted]

Co-Mentor:
James Mickelson, MS, PhD
Professor
Department of Veterinary Biosciences
College of Veterinary Medicine
University of Minnesota
[redacted]

1. Hypothesis and Objectives. Equine Metabolic Syndrome (EMS) is defined by a clustering of clinical signs, namely hyperinsulinemia, insulin resistance and adiposity, which predispose horses to the development of laminitis[1,2]. Certain breeds (Morgan, Arabian, Welsh Pony, Tennessee Walking Horse) appear to be more susceptible to EMS, while other breeds (Quarter Horse) seem to be at lower risk[3]. Our preliminary data demonstrate significant breed differences in EMS-defining metabolic traits (e.g. insulin dynamics, lipid metabolism, adipokine concentrations) that mirror EMS risk. We ***hypothesize*** that breed differences in these key metabolic phenotypes are due to high frequencies of alleles that modify metabolic traits within breeds. The ***goals of this proposal*** are to **1)** further dissect the metabolic differences between breeds at the molecular level using total serum metabolite profiling; and **2)** use these breed-specific molecular metabolic profiles to identify candidate genes underlying breed metabolic differences.

Objective 1: Determine the molecular basis of breed metabolic variation by connecting metabolic pathways to the hormonal and biochemical differences between breeds. Global serum metabolite profiles will be analyzed before and after an oral sugar test in 274 horses from 5 breeds with distinct metabolic phenotypes. These data will be used to **1)** identify the metabolites and metabolic pathways that are significantly different between breeds using functional annotation and mapping of metabolites to known pathways, pathway and metabolite set enrichment analysis, network analysis and pathway activity profiling; and **2)** correlate the metabolite/pathway differences with the previously identified breed differences in key hormonal and biochemical measurements.

Objective 2: Identify candidate genes responsible for metabolic differences between breeds. In ***objective 2a***, genomic regions and specific haplotypes that are highly differentiated between breeds will be identified using high-density SNP genotype data from the breeds in objective 1. In ***objective* 2b**, metabolites and metabolite pathways that are significantly different between breeds will be used to provide functional context for narrowing the focus to specific candidate genes within these genomic regions.

Combined, this project will expand our knowledge of the metabolic and genomic bases for across-breed differences in metabolic profiles and how they relate to EMS physiological and clinical phenotypes, and as a result greatly increase our understanding of the complex biology of EMS.

2. Justification, Significance and Literature Review. EMS refers to a cluster of clinical abnormalities minimally represented by three criteria: documented or suspected insulin resistance (IR), obesity and/or increased regional adiposity, and a predisposition to laminitis[1,2]. Other reported abnormalities include hypertriglyceridemia, dyslipidemia, increased low-density lipoprotein concentrations, hyperleptinemia, arterial hypertension, and increased systemic inflammatory markers[4-7], although there is disagreement between studies[8]. Using morphometric, biochemical and hormonal phenotypes combined with epidemiologic and environmental data from 610 horses/ponies, we have demonstrated that variability in metabolic phenotypes is due to both environmental and individual factors, including genetics (**preliminary data**). In addition, these data highlight the fact that while breeds can share key metabolic features of EMS, they can also differ in the magnitude of these responses or of other features, such as fasting insulin, triglyceride, non-esterified fatty acid and adipokine concentrations. We *hypothesize* that the differences in the severity and secondary features of the EMS phenotype between breeds are the result of the variable frequencies of genetic risk alleles within breeds. Thus, the heterogeneity of the EMS phenotype across individuals and breeds is a result of the combination of underlying genetic alleles, the interactions between these alleles and the environment, and the resulting molecular pathophysiology. In this proposal, we build upon our ongoing work on the genetics of individual metabolic variation to focus on the molecular and genetic differences underlying breed-specific metabolic profiles.

Intense selective breeding in domestic animals leads to the fixation of phenotypic traits within breeds. Since domestication, selection in horses has included metabolic and athletic phenotypes that allowed horses to efficiently perform different types of work. Thus, many modern breeds likely have alleles segregating at high frequency or approaching fixation (i.e., 100% frequency) that modulate energy utilization by mechanisms such as enhanced anaerobic and/or aerobic metabolism, increased muscle mass and strength, increased cardiovascular and respiratory fitness, localized accumulation of adipose tissue, and improved economy of locomotion, among others[9].These metabolically efficient or "thrifty" alleles, in particular those that regulate energy intake, storage and use, coupled with changes in equine husbandry practices in recent times including dense high caloric feed and limited exercise, may help explain the increasing prevalence of equine metabolic disease phenotypes in modern environments. Our proposal is based on the hypothesis that genetic predisposition to the development of

metabolic syndrome phenotypes in certain breeds is due to modern environments interacting with genotypes selected for under different environmental pressures. Studies demonstrating the overlap between genetic loci under selection in the human genome and loci conferring susceptibility to metabolic traits including obesity and type 2 diabetes lend validity to this hypothesis[10,11]. Similarly, in the horse, we have identified selection for a mutation in the *GYS1* gene in Belgian draft horses, resulting in excess skeletal muscle glycogen associated with Polysaccharide Storage Myopathy Type 1 in modern environments[13] and selection at the myostatin (*MSTN*) gene locus in Quarter Horses, which alters muscle fiber type proportions and which we hypothesize also alters many aspects of muscle and adipose tissue energy metabolism (**preliminary data**). Elucidating the evolution of the genetic basis of metabolic efficiency and metabolic syndrome across horse breeds is a novel, unexploited approach to the study of the genetic basis of obesity, energy dysregulation and EMS.

The current lack of information regarding the genetic basis of and variation in EMS within and across breeds restricts understanding of the pathophysiology, as well as limits the ability to predict disease risk and to identify individuals who can benefit from management changes and/or therapeutic intervention prior to the onset of disease/laminitis. To further untangle this complex phenotype, we propose collection of global serum metabolomics data both before and after an oral glucose challenge in our large across-breed cohort of horses and ponies. Non-targeted studies such as this provide unbiased data on the presence/absence and proportions of metabolites contained within a sample, and will provide a snapshot into the metabolic state of these horses both before and after a dynamic challenge (i.e., OST). Biologic differences between breeds will be explored by **1)** identification of breed differences in metabolite concentrations; **2)** functional annotation, mapping and visualization of metabolites within known predefined metabolic pathways; **3)** pathway overrepresentation analysis of significant metabolite lists; **4)** metabolite set enrichment analysis; **5)** the construction of metabolite networks and network topology analysis; and **6)** pathway activity profiling. Identified metabolic differences between breeds will be linked to putative candidate genes through a set of analyses designed to detect regions of low genetic diversity and positive selection in breeds of varying metabolic and EMS phenotypes.

We expect the major contribution of the proposed research to be the detection of breed differences in metabolites/metabolic pathways that are correlated to previously identified biochemical and hormonal differences, and the identification and prioritization of candidate genes that influence a spectrum of metabolic traits, particularly the susceptibility to metabolic syndrome. These findings will provide novel insights into disease biology, allowing the identification of new therapeutic targets, and increasing our understanding of the pathophysiology of EMS and its associated clinical features. Moreover, the identification of genes underlying the EMS phenotype will also directly impact equine health by allowing for the development of genetic tests to identify horses at risk for the development of obesity and laminitis prior to the onset of clinical disease.

3. Preliminary Data.

Clinical and breed variation in metabolic phenotype. We have collected 11 phenotypes, epidemiologic and environmental data from a total of 610 horses/ponies from 166 farms. These data include individuals from 5 breeds with distinct metabolic phenotypes (Arabians, Morgans, Quarter Horses,Tennessee Walking Horses and Welsh Ponies). Using these data, we have been able to rigorously assess several key aspects of variation in metabolic phenotypes and EMS, including: **1)** how metabolic measures differ in horses when parsed by clinical status (i.e., non-obese, non-laminitic [NO-NL], non-obese, laminitic [NO-L], obese, non-laminitic [O-NL], and obese, laminitic [O-L] and **2)** how *individual* (i.e., age, **breed,** gender, etc) and *environmental* (i.e., dietary adaptation, exercise, etc) variables impact trait measures (**Table 1**).

Our data indicate that fasting insulin, triglycerides, and insulin 75 minutes post-oral sugar challenge (OST) are consistently elevated, whereas adiponectin is consistently decreased in individuals with a history of laminitis, regardless of obesity status, age, gender or breed (**Table 1**). In addition to differences between clinical groups, we have also identified significant differences in the relevant biochemical measurements between breeds regardless of clinical phenotype (**Table 2**). For example, Quarter Horses have significantly lower fasting insulin and insulin OST, and lower leptin in relation to other breeds, whereas Welsh Ponies tend to have higher insulin OST. Moreover, Welsh Ponies have significantly higher fasting serum triglyceride and NEFA concentrations than horses as a whole (data not shown), even after correction for obesity and laminitis status. Heritability estimates from whole genome SNP genotype data in Morgans indicate many of these traits are highly heritable (**Table 1**).

Table 1. Key findings and estimated heritability for phenotypes in 610 horses and ponies.							
Response variables	**Predicted mean trait values by clinical group[€]**					**Predictor (explanatory) variables significantly associated with outcome measures (p< 0.001)**	**Estimated heritability (Morgans)**
	NO-NL	NO-L	O-NL	O-L	Overall p-value		
N:H RATIO	0.64^{c}	0.66^{a}	0.68^{ab}	0.70^{b}	**$4.60x10^{-19}$**	**breed**, gender, obesity, laminitis	21.24%
G:H RATIO	1.19^{a}	1.19^{a}	1.24^{b}	1.24^{b}	**$2.10\ x10^{-17}$**	gender, obesity	10.48%
GLU (mg/dl)	77.5^{a}	79.8^{a}	78.2^{a}	78.6^{a}	0.44	season	NA*
INS (uIU/ml)	5.1^{c}	9.5^{ab}	7.8^{a}	13.3^{b}	**$1.40\ x10^{-19}$**	**breed**, age, obesity, laminitis	32.19%
GLU OST (mg/dl)	98.3^{a}	100.7^{a}	101.9^{a}	102.8^{a}	0.15	dietary starch, hours in stall	NA*
INS OST (uIU/ml)	19.5^{b}	35.2^{a}	32.6^{a}	42.8^{a}	**$2.90\ x10^{-11}$**	**breed**, gender, age, obesity, laminitis	34.39%
TG (mg/dl)	23.8^{b}	33.7^{a}	26.8^{b}	34.4^{a}	**$6.40\ x10^{-08}$**	laminitis, season, thyroxine	NA*
NEFA (mmol/L)	0.20^{a}	0.21^{a}	0.20^{a}	0.21^{a}	0.96	hours grazing, hours in stall, season	NA*
LEP (ng/ml)	4.2^{a}	4.0^{a}	6.1^{b}	6.0^{b}	**$2.32\ x10^{-11}$**	**breed**, age, latitude, season	54.50%
APN (ng/ml)	4240^{b}	2495^{a}	3752^{b}	2418^{a}	**$1.15\ x10^{-08}$**	**breed**, gender, obesity, season	13.58%
ACTH (pg/ml)	26.9^{a}	29.3^{a}	28.1^{a}	27.7^{a}	0.48	gender, laminitis, season	55.65%

[€]Values are adjusted for breed, sex, age, month, latitude, diet, exercise, and L-thyroxine supplementation. Different letters within a row indicate a significant pair-wise difference (p <0.05, Holm adjustment for multiple comparison). *NA indicates heritability estimates <2%. Abbreviations: neck- [NH] and girth- [GH] to height ratio, fasting glucose (GLU), insulin (INS), adrenocorticotropic hormone (ACTH), triglyceride (TG), non-essential fatty acid (NEFA), leptin (LEP) and adiponectin (APN) concentrations; and insulin (INS OST) and glucose (GLU OST) post oral sugar challenge.

Breed-specific genomic regions. Our use of low-density SNP genotype data to identify regions of breed differentiation/selection across a wide breed panel has been demonstrated earlier[12-15]. Here we propose a similar approach in the relevant 5 breeds using ~1.8 million SNP markers, with genomic signatures of selection identified by calculating d_i statistics in non-overlapping 10 kb windows across the genome[15,16] (**Figure 1**). The highest *di* values for the Quarter Horse (QH) are located within a 3-Mb window on chromosome 18 around the myostatin gene (*MSTN*), which we have previously reported as a signature of selection in this breed[15]. We have recently demonstrated that a SINE insertion in the *MSTN* promoter region results in altered skeletal muscle fiber type proportions[17] and that the SINE alters metabolic phenotype in QH. Specifically, QHs with the SINE allele have less apparent adiposity (lower G:H RATIO), lower fasting and OST insulin, and leptin concentrations. In contrast, serum adiponectin concentrations are positively associated with the SINE (*manuscript in preparation*). In the Welsh Pony (WP), the highest *di* values were found to be on ECA6 around the interleukin-1 receptor-associated kinase 3 (*IRAK3*) gene, which is down-regulated in association with obesity and metabolic syndrome in humans[18]. These data demonstrate the feasibility of identifying breed-specific genomic regions associated with metabolic phenotypes in our cohort.

Table 2. Breed differences in metabolic trait measures. Values are adjusted for clinical status, sex, age, month, latitude, diet, exercise, and L-thyroxine supplementation.

	Morgan	Arab	Pony	QH	TW	Overall p-
N:H	0.69^{b}	0.65^{a}	0.69^{b}	0.68^{ab}	0.68^{ab}	**$3.10x10^{-06}$**
INS	$9.6b^{c}$	5.9^{ab}	11.7^{c}	5.5^{a}	10^{bc}	**$4.00x10^{-06}$**
INS OST	34.3^{a}	25^{ab}	41.8^{b}	16.1^{a}	45.6^{b}	**$1.50x10^{-06}$**
TG	29.9^{ab}	26.8^{ab}	32.3^{ab}	24.1^{b}	34.5^{ab}	**$2.60x10^{-03}$**
LEP	6.2^{a}	4.5^{ab}	5.1^{a}	3.2^{b}	5.8^{a}	**$1.79x10^{-06}$**
ACTH	26.4^{b}	30.1^{ab}	35.9^{a}	23.8^{b}	22.6^{b}	**$6.60x10^{-05}$**

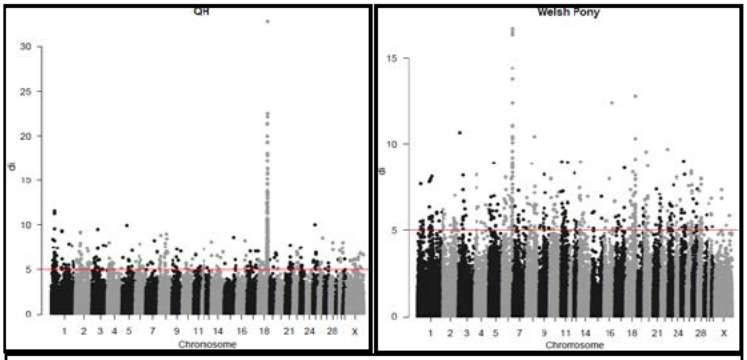

Figure 1. Genome-wide *di* values for QH and WP. Regions of putative positive selection (*di* values above the red line) were those in the top 99% of each empirical distribution

Characterization of serum metabolites and metabolic pathways. Metabolomic analysis, completed by a commercial laboratory (Metabolon®) using a combination of GC-MS and LC-MS, was performed on serum samples obtained from 20 Welsh ponies before and at 75 min during an OST. Ponies were classified as healthy [CON] (n = 10, insulin < 30 mU/L) or having insulin dysregulation [ID] (n = 10, insulin > 60 mU/L) at 75 min post OST. The serum

metabolomic profiles comprised 646 biochemical compounds, of which 506 were of known identity. Statistically significant differences between ID and CON, or between baseline and post OST, were identified for > 130 biochemicals, including hexoses (glucose, mannose, fructose), metabolites involved in the tricarboxylic acid cycle (citrate, fumarate, malate), fatty acid metabolism (palmitoleate, eicosapentanoate, palmitolycarnitine, laurylcarnitine), and branched-chain amino acid oxidation (4-methyl-2-oxopentanoate, 3-methyl-2-oxoalerate and isovalerylglycine). These data demonstrate the potential for serum metabolomic data to lead to insight into the metabolic differences between horses and clinical phenotypes.

4. Experimental Methods and Design.

Objective 1: Determine the molecular basis of breed metabolic variation by connecting metabolic pathways to the hormonal and biochemical differences between breeds.

Rationale. Global serum metabolomics data collected both before and after an oral glucose challenge in an across-breed cohort of horses and ponies will provide a *comprehensive* and *unbiased* approach to understanding the molecular physiology/pathophysiology of breed phenotypic variation/EMS by providing data on the proportions of all measureable metabolites. Further, annotation and assembly of metabolite data into pathways and networks and correlation of these pathways/networks to hormonal and biochemical differences between breeds can potentially pinpoint the molecular alterations leading to, or resulting from, hormonal/biochemical differences (e.g., the effects of hyperinsulinemia, etc). Metabolites, metabolite ratios, and networks/pathways will also provide a more precise foundation for investigation of the genetic component of metabolic trait variation in **objective 2**.

Experimental Methods and Design. Fasting and 75 min post oral glucose challenge (OST) serum samples from 274 previously-phenotyped individuals from 5 breeds (39 Arabians, 75 Morgans, 46 Quarter Horses, 39 Tennessee Walking Horses and 75 Welsh Ponies) will be used. Based on power calculations performed in Metaboanalyst for multivariate models using preliminary metabolomic data, to determine group (i.e. breed) differences with a specificity of 0.95 and a sensitivity of 0.85, the number of needed individuals per group is 38-75 depending on the desired confidence interval (i.e., 90% vs 95%)[19-21]

All analytical work will be performed at the U of M Center for Mass Spectroscopy and Proteomics. Samples will be deproteinized and aliquots lyophilized[22]. An initial QC analysis to assay technical, sample preparation and biologic variability will be used to set acceptable QC limits. Then, raw data will be pre-processed to provide structured data and appropriate format for subsequent analysis using Progenesis QI[23] (for UHPLC-MS) or XCMS[24] (for GC-MS) software. Background subtraction will be performed as indicated by initial QC investigation. Data will be auto-scaled (i.e., mean centered and divided by the sample standard deviation) to avoid skewing due to dynamic range differences, and log-transformed to ensure normality, if necessary (i.e., for regression-based analyses). Sample and feature outliers will be removed and missing values will be imputed[22]. Unidentified metabolites of statistical interest will be subjected to secondary fragmentation and resulting patterns searched against existing spectral libraries containing annotated fragmentation patterns.

Statistical differences between breeds and correlations between metabolic features and continuous measured outcome variables (i.e., INS, INS OST, TG, GHR, APN, etc)[26] at each time point (baseline and post-OST) will be performed in MetaboAnalyst[19-21]. Comparison of temporal profiles (i.e. fasting and post-OST) will be evaluated using the MetATT tool using multivariate empirical Bayes time series analysis (MEBA)[27]. These analyses will identify differences between breeds, time points, and the interactions between breed and time point.

Biological patterns, functions and metabolic pathways differences will be further explored by functional annotation, mapping and visualization of metabolites within known predefined metabolic pathways using MBRole[28] and Interactive Pathways Explorer (iPath)[29], with data from KEGG[30], HMDB[31], PubChem (pubchem.ncbi.nlm.nih.gov), and ChEBI[32] databases. Pathway overrepresentation analysis and metabolite set enrichment analysis will be performed using MESA. The construction of metabolite networks and network topology analysis and pathway activity profiling will be performed with MetPA and PAPi, respectively. The differences in metabolites, pathways, networks and pathway activity estimates between breeds will highlight key metabolic functions that are differentiated by breed.

Expected outcomes and potential pitfalls. We expect to identify up to several thousand metabolites, as observed in the human serum metabolome project[26]. Several dozen to hundreds of metabolites may differ significantly by

breed or temporal sampling, and several metabolites and metabolic pathways are likely to be correlated to phenotypic measurements such as INS, INS OST, APN, LEP, TG, etc.

Objective 2. Identify candidate genes responsible for metabolic differences between breeds.

Rationale. Domestic animal breeds originate through artificial selection in which breeders work to fix desirable phenotypic traits. In the horse, there is evidence that this selection process has focused on metabolic and athletic traits that have resulted in several breeds that are metabolically efficient, and by extension at high risk for the development of EMS. When the alleles underlying these traits reach high frequency or become fixed with in a breed, genetic association studies lose their power to identify these loci. To identify high allele frequency loci harboring alleles responsible for breed differences in metabolic traits, we first identify genomic regions within each breed where there is evidence of decreased genetic diversity and haplotype differentiation relative to other breeds, which together suggest that the region harbors genes/alleles responsible for breed phenotypic differentiation. Then, we identify the genes present in these genomic regions of interest (ROIs) (**objective 2a**). In **objective 2b**, we capitalize on the breed differences in metabolic phenotypes identified in objective 1 to prioritize genes within these ROIs. Gene lists from ROIs will be prioritized for further association investigation using the metabolic pathways that are different between breeds and correlated with key hormonal measures. Identification of putative metabolic pathways and candidate genes that result in breed phenotypic differences constitutes a key step in linking certain phenotypes to causal genotypes and is critical to understanding the complex biology underlying EMS.

Experimental Methods and Design - Objective 2a. Genomic ROIs will be identified in all breeds using an Fst-based statistic (d_i) calculated in non-overlapping 10 kb regions across the genome, with regions in the top 1% of the empirical data picked as putative regions of genomic differentiation (**preliminary data**) [15,16]. Haplotype sharing within and specific haplotypes within ROIs will be delineated by the hapQTL method of quantifying local haplotype sharing[33] for which we will choose significant SNP markers with –log (Bayes Factor) > 5 with at least 2 significant markers overlapping the above regions. Finally, extended haplotype homozygosity analysis will be conducted to identify core haplotypes containing putative causative alleles[34,13]. The BioMart software suite (www.biomart.org) will be used to query the Ensembl database and create lists of genes located within the haplotype boundaries identified above. The equine genome (EquCab2) and syntenic regions of the human genome sequence (GRCh37) will be queried to ensure candidate genes are not missed due to inaccurate annotation of EquCab2. Gene ids, gene structure information including start and end positions and exonic regions will be retrieved. All genes will be annotated with gene ontology (GO) terms to facilitate identification of genes with relevant biologic function.

Experimental Methods and Design - Objective 2b. Mapping and annotation of metabolites to known metabolic pathways in objective 1 using MBRole and iPath allow for changes in metabolite concentrations to be linked to the gene products (and subsequently genes) that are involved in metabolic pathways. Genes in ROIs for a given breed will be prioritized as candidate genes if the gene's product plays a role in the formation of the metabolites or in pathways that are different within the breed. Metabolite correlation networks (MetPA) can also be used to prioritize genes when metabolites do not map to canonical pathways. In this instance, breed-specific sub-networks containing genes identified within ROIs will be extracted from the metabolite networks. The biological coherency of the subnetwork will be assessed by calculating the average metabolite interaction weight between all possible interactions[21], and used to reconstruct biochemical pathways related to breed metabolic differences.

Expected results and potential pitfalls. As shown in preliminary results, there may be many dozens of loci showing differentiation between the 5 breeds which contain compelling genes in pathways highlighted by metabolomic analysis. We anticipate that the myostatin locus on ECA18 will be highlighted as a major driver of the metabolic differences in QHs as compared to other breeds, and as such should serve as a 'proof of principle' positive control and that the ECA6 locus in Welsh ponies may provide another major target. It is possible that we do not identify any genes within genomic ROIs that are strongly correlated with breed-specific networks. In this circumstance, we will simply use differential network analysis to identify candidate genes for further investigation, regardless of ROIs.

Timeline. Objective 1: Sample preparation and submission; months 1-6. Serum metabolite profile analysis; months 6-12. **Objective 2a**: Identification of signatures of selection; months 12-14. O**bjective 2b**: Candidate gene prioritization; months 14-20. Manuscript(s) preparation: months 20-24.

Morris Animal Foundation

Animal Involvement Justification

(From the proposal guidelines, single-spaced, no page limit)

Morris Animal Foundation (MAF) is dedicated to funding scientifically sound, relevant and humane studies that specifically address the health and well-being of animals. All studies receiving funding must follow MAF's Health Study Policy for Animals Involved in Research (adopted October 18, 2008), which was written to ensure that each and every animal involved in a MAF funded health study receives excellent, compassionate care throughout the study. MAF shall not fund health studies which require euthanasia as an endpoint or the induction of disease or injury, unless the nature of the disease or condition to be studied is of such significance for improving animal health that such means are justified, and that meaningful information can be obtained in no other way. Furthermore, MAF will not fund any study that induces or allows pain or distress unless such pain or distress can be controlled by appropriate anesthetic, analgesic, tranquilizing drugs, or nursing care. Click here for the full Health Study Policy.

A. If this study does not involve live animals please indicate here by N/A: ____**NA**_____

B. Does this study involve biological samples, tissues, etc.? ____**YES**_____

If yes, describe in detail what samples will be used and where & how they will be (or were) acquired. Note: Morris Animal Foundation reserves the right to request a copy of the Institutional Animal Care and Use Committee (IACUC) application/approval and other relevant applications/approvals (e.g., wildlife permit) covering the original collection of samples, including archived samples. MAF reserves the right to request IACUC (or equivalent) review and approval for any Foundation study ***regardless of the Institution's requirements***. This would include the use of archived samples as well as clinical trials.

ALL SAMPLES HAVE BEEN COLLECTED AS PART OF PRIOR STUDIES. Information regarding prior collection is provided below. Blood samples were collected before and 75 minutes after an oral sugar test from the jugular vein.

C. If this study involves live animals, succinctly address the following: (please restate the questions and directives).

1. What species will be studied? **EQUINE**
2. State the status of your IACUC application/approval. All recipients of MAF funding will be required to submit the entire IACUC protocol and document. A copy of the IACUC approval should not be included with the application, but it is required before funding can be awarded.

   Samples were collected under the University of Minnesota IACUC protocol # 1109B04448 approved 10/16/11 (M McCue PI)

3. List the USDA category for pain and distress (B, C, D, E): __**C**____

   Note: Any study beyond category C will require review by MAF's Animal Welfare Advisory Board (AWAB). In general MAF does not fund studies beyond category C (category D studies will only be considered if they conform with MAF's Health Study Policy, category E studies will not be considered).

4. Does this proposal involve client-owned animals? ___**YES**_____

   If yes, the protocol for client-owned animals must be approved by the appropriate peer review committee before the project is funded. *If this proposal involves client-owned animals, an informed client consent form must be submitted with this proposal. For a suggested list of items to be considered in an informed client consent form, click here.*

   Informed client consent attached.

5. Explain how animals will be acquired (e.g., client-owned, USDA licensed breeder, institutional "herds" or "colonies") and verify that the animals are suitable for the study (e.g., have no physiologic, physical or

pharmacologic issues that would interfere with results)

6. How many animals will be used? **__274 (for this study)____**

   a. Summarize numerical justification **Based on power calculation –see experimental design**

7. Does this study induce disease, injury, pain or distress in animals? Note: any study requiring the induction of disease, injury, pain, or distress will have an additional evaluation by MAF's AWAB.

   If yes,

   a. Defend the necessity of experimental design

   Collection of blood samples is necessary for accurate phenotyping of horses and ponies for metabolic traits. Samples have been previously collected (under the UMN IACUC referenced above as well as under IACUC protocol #0804A30547, Approved 4/20/2010, renewed 4/21/2011).

   b. Explain how pain and/or distress will be controlled

   Jugular veniapuncture was performed by licensed veterinarians, and caused minimal distress.

   c. Justify that no alternative, including clinical studies, can be used to accomplish study objectives and the disease/condition to be studied is of such significance for improving the health of the species.

   An alternative method to characterize the serum metabolites in horses of different breeds/ metabolic phenotype without the collection of blood samples is unavailable.

8. Explain the environment and housing conditions (quality of life) in which the animals will live (address species-appropriate exercise, enrichment, socialization, veterinary care, etc.)

   Animals are client owned and living in their normal environment.

9. What will happen to the animals upon completion of the study? **N/A, animals are client owned**

   If adoption, explain the adoption process. Provide assurance that whenever possible and when in the animal's best interest, investigators shall make companion animals available for adoption at the end of the study or return the animals to the owner/responsible agency in an environment that promotes animal welfare and excellent quality of life.

10. If euthanasia, provide the following additional information (note: any study requiring euthanasia as an endpoint will have an additional evaluation by a MAF's AWAB.

    i. Total number that will be euthanized and justification for numbers

    ii. Method of euthanasia

    iii. Justification that no alternatives can be used to accomplish study goal(s) and that the disease/condition to be studied is of such significance for improving the health of the species that a terminal endpoint is deemed necessary.

    iv. Reason for euthanasia in lay language (this wording may be shared with staff, donors and media)

    v. Provide objective criteria for determining when euthanasia is appropriate or necessary (note: Morris Animal Foundation wants assurance that an animal will not be allowed to suffer and that monitoring for pain and suffering is adequate)

    Note: Morris Animal Foundation does not consider the use of CO2 alone to be an appropriate method of euthanasia

Please note:

1. If an animal is used in an invasive study, MAF may require that a guarantee is provided, through principal investigator and institutional signatures that the animal will not participate in any future invasive study or procedure
2. MAF does not allow inclusion of ancillary data in MAF funded research that includes animal use

protocols not in agreement with our Health Study Policy, even if it is obtained using other funding sources.

3. Morris Animal Foundation considers euthanasia acceptable when an animal develops unanticipated illness or injury that results in pain and suffering that cannot be alleviated with standard veterinary interventions.

Informed Client Consent Equine Metabolic Syndrome (EMS) Study

We are delighted that you have agreed to allow us to examine and obtain blood samples from your horses/ponies, in addition to obtaining hay/pasture/feed samples, in order to determine if your horse(s) have a group of risk factors for laminitis that have been termed Equine Metabolic Syndrome (EMS). We will use the results of this study to determine early predictors for laminitis in horses and ponies. We will also save some of your horse's genetic material (DNA) so we can identify the underlying genetic causes that predispose horses to EMS. The results will be made available to you. Identification of risk factors in your horses will allow you to identify high risk horses and change their management, and hopefully prevent future episodes of laminitis. The results of our study will be published in scientific journals **without** the identity of the farms involved or horses involved being disclosed. Records will be kept confidential indefinitely.

The examination will consist of a physical examination of your horse, the determination of a body condition score (BCS), measurements of height, circumference of the neck and girth and weight (estimated by weight tape). We will also evaluate your horse for any signs of lameness/laminitis and will collect historical data regarding prior episodes of laminitis. Digital photographs will be taken to allow assessment of BSC by a second researcher. We will collect blood samples before and after an oral sugar test for measurement of glucose, triglyceride, free fatty acid, insulin, GLP-1, C-peptide, and ACTH concentrations. The oral sugar test will be performed by administering 50 g/kg dextrose as syrup by mouth using a dose syringe. All blood samples will be collected after a short fasting period (approximately 6 hours). A portion of the blood samples will be used to extract DNA for future genetic analysis. Serum and/or plasma will also be frozen for future studies on EMS phenotyping.

All of these procedures are routine and non-invasive. We do not anticipate any complications. Horses may experience a slight discomfort from the needle when blood is sampled. If you notice any swelling at this site, please contact us, and well will arrange for medical care if necessary and cover any costs incurred. Dextrose is a non-structural carbohydrate and insulin resistant horses are more sensitive to non-structural carbohydrates (NSC) and at higher risk for laminitis development when excess levels of NSC are consumed in the diet over a period of time. However, *the oral glucose tolerance test will only involve a one-time administration of dextrose in an amount that is small enough to not place an insulin resistant horse at a higher risk of developing laminitis.* The amount of dextrose being administered is approximately equivalent to the amount of NSC in a typical grain meal. We have performed this procedure in more than 500 horses without any complications.

Once you are enrolled in the study, the examination and blood collection will take approximately 90 minutes. We may contact you in the future to follow-up with you regarding any bouts of laminitis your horse has experienced after our visit.

The contacts for this study are Dr. Molly McCue (612) 624-9320 or Dr. Ray Geor (517) 355-9593. Either the researchers or the owner of the farms involved in this study have the right to withdraw at any time. The cost of the examinations, blood measurements, and hay/pasture/feed analysis will be covered entirely by the research team.

Informed consent: I, _____ understand that there may be unforeseen risks involved in any research activity. If I have any concerns about the performance of this study I can contact the Department Chairman at The University of Minnesota Dr. Tom Molitor at (612) 625-7755 or the Institution animal care and use committee (IACUC).

_____ _____

Owner signature and date Investigator signature and date

Revised October 2010

VII. Recombinant DNA/Biohazards:

NA

VIII. Cited References

1. Geor, R. & Frank, N. (2009). Metabolic syndrome - From human organ disease to laminar failure in equids. *Vet.Immunol.Immunopathol.* 129: 151-154.
2. Frank, N., Geor, R.J., Bailey, S.R., Durham, A.E., Johnson, P.J. (2010). Equine metabolic syndrome. *J.Vet.Intern.Med.* 24: 467-475.
3. Frank, N., Geor, R.J., Bailey, S.R., Durham, A., Johnson, P.J. (2010). Equine Metabolic Syndrome: ACVIM 2009 Large Animal Consensus Statement. *J.Vet.Intern.Med.* 24:467–475.
4. Treiber, K.H., Kronfeld, D.S., Hess, T.M., Byrd, B.M., Splan, R.K., Staniar, W.B. (2006). Evaluation of genetic and metabolic predispositions and nutritional risk factors for pasture-associated laminitis in ponies. *J.Am.Vet.Med.Assoc.* 228: 1538-1545.
5. Carter, R.A., Treiber, K.H., Geor, R.J., Douglass, L., Harris, P.A. (2009). Prediction of incipient pasture-associated laminitis from hyperinsulinaemia, hyperleptinaemia and generalised and localised obesity in a cohort of ponies. *Equine Vet.J.* 41: 171-178.
6. Frank, N., Elliott, S.B., Brandt, L.E., Keisler, D.H. (2006). Physical characteristics, blood hormone concentrations, and plasma lipid concentrations in obese horses with insulin resistance. *J.Am.Vet.Med.Assoc.* 228: 1383-1390.
7. Bailey, S.R., Habershon-Butcher, J.L., Ransom, K.J., Elliott, J., Menzies-Gow, N.J. (2008). Hypertension and insulin resistance in a mixed-breed population of ponies predisposed to laminitis. *Am.J.Vet.Res.* 69: 122-129.
8. **McCue, M.E.**, Geor, R.J., Schultz, N. (2015). Equine metabolic syndrome: a complex disease influenced by genetics and the environment. *J Equine Vet. Sci*. 35(5):367-375.
9. Johnson, P.J. (2002). The equine metabolic syndrome: peripheral Cushing's syndrome. *Vet Clin North Am Equine Pract.* 18(2):271-93.
10. Pan, W. (2008). Network-based model weighting to detect multiple loci influencing complex diseases. *Hum Genet.* 124: 225-234.
11. Ritchie, M.D., Holzinger, E.R., Li, R., Pendergrass, S.A., Kim, D. (2015). Methods of integrating data to uncover genotype-phenotype interactions. *Nature Reviews Genetics* 16: 85-97.
12. Petersen, J.L., **Mickelson, J.R**., Valberg, S.J., **McCue, M.E.** (2015). Genome-wide SNP data show little differentiation between the Appaloosa and other American stock horse breeds. *Anim Genet.* Epub ahead of print.
13. McCoy, A.M., Schaefer, R., Petersen, J.L., Morrell, P.L., Slamka, M.A., **Mickelson, J.R**., Valberg, S.J., **McCue, M.E**. (2014). Evidence of positive selection for a glycogen synthase (*GYS1*) mutation in domestic horse populations. *J Hered.*105(2):163-72.
14. Petersen, J.L., **Mickelson, J.R.,** Cothran, E.G., Andersson, L.S., Axelsson, J., Bailey, E., Bannasch, D., Binns, M.M., Borges, A.S., Brama, P., da Câmara Machado, A., Distl, O.,Felicetti, M., Fox-Clipsham, L., Graves, K.T., Guérin, G., Haase, B., Hasegawa, T., Hemmann, K., Hill, E.W., Leeb, T., Lindgren, G., Lohi, H., Lopes, M.S., McGivney, B.A., Mikko, S., Orr, N., Penedo, M.C., Piercy, R.J., Raekallio, M., Rieder, S., Røed, K.H., Silvestrelli, M., Swinburne, J., Tozaki, T., Vaudin, M.M., Wade, C., **McCue, M.E**. (2013). Genetic diversity in the modern horse illustrated from genome-wide SNP data. *PLoS One.* 8(1):e54997
15. Petersen, J.L., **Mickelson, J.R.**, Rendahl, A.K., *et. al.*, **McCue, M.E.** (2013). Genome-wide analysis reveals selection for important traits in domestic horse breeds. *PLoS Genet*. 9(1):e1003211.
16. Akey, J.M., Ruhe, A.L., Akey, D.T., Wong, A.K., Connelly, C.F., Madeoy, J., Nicholas, T.J., Neff, M.W. (2010). Tracking footprints of artificial selection in the dog genome. *Proc Natl Acad Sci USA.* 107(3):1160-5.

17. Petersen, J.L., Valberg, S.J., **Mickelson, J.R., McCue, M.E.** (2014). Haplotype diversity in the equine myostatin gene with focus on variants associated with race distance propensity and muscle fiber type proportions. *Anim Genet.* 45(6):827-35.
18. Hulsmans, M., Geeraert, B., De Keyzer, D., Mertens, A., Lannoo, M., Vanaudenaerde, B., Hoylaerts, M., Benhabilès, N., Tsatsanis, C., Mathieu, C., Holvoet, P. (2012). Interleukin-1 receptor-associated kinase-3 is a key inhibitor of inflammation in obesity and metabolic syndrome. *PLoS One* 7(1):e30414
19. Xia, J., Wishart, D.S. (2011). Metabolomic data processing, analysis, and interpretation using MetaboAnalyst. *Curr Protoc Bioinformatics* Chapter 14: Unit 14.10.
20. Xia, J., Wishart, D.S. (2011). Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst. *Nat Protoc.* 6(6):743-60.
21. Xia, J., Mandal, R., Sinelnikov, I.V., Broadhurst, D., Wishart, D.S. (2012). MetaboAnalyst 2.0--a comprehensive server for metabolomic data analysis. *Nucleic Acids Res.* (Web Server issue):W127-33.
22. Dunn, W.B., Broadhurst, D., Begley, P., Zelena, E., Francis-McIntyre, S., Anderson, N., Brown, M., Knowles, J.D., Halsall, A., Haselden, J.N., Nicholls, A.W., Wilson, I.D., Kell, D.B., Goodacre, R. (2011). Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat.Protocols* 6: 1060-1083.
23. Da Qi, Philip Brownridge, Dong Xia, Katherine Mackay, Faviel F. Gonzalez-Galarza, Jenna Kenyani, Victoria Harman, Robert J. Beynon, and Andrew R. Jones. (2012). A Software Toolkit and Interface for Performing Stable Isotope Labeling and Top3 Quantification Using Progenesis LC-MS. *OMICS.* 16(9): 489–495.
24. C. A. Smith, E. J. Want, G. O'Maille, R. Abagyan, and G. Siuzdak. (2006). XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal. Chem.*, 78:779–787.
25. Redestig, H., Kusano, M., Fukushima, A., Matsuda, F., Saito, K., Arita, M. (2010). Consolidating metabolite identifiers to enable contextual and multi-platform metabolomics data analysis. *BMC bioinformatics* 11: 214.
26. Dunn, W.B., Broadhurst, D.I., Atherton, H.J., Goodacre, R., Griffin, J.L. (2011). Systems level studies of mammalian metabolomes: the roles of mass spectrometry and nuclear magnetic resonance spectroscopy. *Chemical Society Reviews* 40: 387-426.
27. Xia, J., Sinelnikov, I.V., Wishart, D.S. (2011). MetATT: a web-based metabolomics tool for analyzing time-series and two-factor datasets. *Bioinformatics* 27(17):2455-6.
28. Chagoyen, M., Pazos, F. (2011). MBRole: enrichment analysis of metabolomic data. *Bioinformatics* 27: 730-731.
29. Takuji Yamada, Ivica Letunic, Shujiro Okuda, Minoru Kanehisa, and Peer Bork (2011). iPath2.0: interactive pathway explorer. *Nucleic Acids Res.* 39(Web Server issue): W412–W415.
30. Kanehisa, M., Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research* 28: 27-30.
31. Wishart, D.S., Tzur, D., Knox, C., Eisner, R., Guo, A.C., Young, N., Cheng, D., Jewell, K., Arndt, D., Sawhney, S. (2007). HMDB: the human metabolome database. *Nucleic Acids Research* 35: D521-D526.
32. Degtyarenko, K., De Matos, P., Ennis, M., Hastings, J., Zbinden, M., McNaught, A., Alcantara, R., Darsow, M., Guedj, M.l., Ashburner, M. (2008). ChEBI: a database and ontology for chemical entities of biological interest. *Nucleic Acids Research* 36: D344-D350.
33. Xu, H., Guan, Y. (2014). Detecting local haplotype sharing and haplotype association. *Genetics.* 197(3):823-38.
34. Sabeti, P.C., Reich, D.E., Higgins, J.M., Levine, H.Z., Richter, D.J., Schaffner, S.F., Gabriel, S.B., Platko, J.V., Patterson, N.J., McDonald, G.J., Ackerman, H.C., Campbell, S.J., Altshuler, D., Cooper, R., Kwiatkowski, D., Ward, R., Lander, E.S. (2002). Detecting recent positive selection in the human genome from haplotype structure. *Nature.* 419(6909):832-7.

PROPOSAL BUDGET

Note: First Award – complete year 1, year 2 and total only. **Pilot Study** – complete year 1 only. **Fellowship Training** – complete salary, fringe benefits, indirect costs and total for year 1 and year 2 only. For Established Investigator, First Award and Pilot Study; Fellowship Training lines can be removed.

Category	Year 1	Year 2	Year 3	Total
Personnel: 1. Principal investigator (name)* 2. Co-investigator #1 (name) 3. Co-investigator #2 (name) 4. Technician Salary (X%) Fringe benefits (Y%) 5. Student Assistant Salary (X%) Fringe benefits (Y%) Total Salaries & Wages **Fellowship Training Only:** Salary (90.5%) Fringe benefits (9.5%)	[redacted]	[redacted]	[redacted]	[redacted]
Supplies & Expenses: 1. 2. 3. 4. Provide justification in the designated section. Total Supplies & Expenses:				
Animal Use & Care: Animal Purchase: Animal Per diem: Total Animal Care:				
Subtotal of All Categories:	[redacted]	[redacted]		
Maximum of 8% - Indirect Costs: ** (8%)	[redacted]	[redacted]		
Grand Total Requested from MAF:	[redacted]	[redacted]		

* Salary requests for principal investigators must be clearly defined and justified in the following budget justification section. You may request salary for technicians, residents, graduate students, and postdoctoral fellows, based on their percentage of time involved in the project

** Indirect costs may be claimed only if you are charged for indirect costs by your institution for work carried out in this proposal. **You must make this calculation yourself.** If your institution charges less than 8%, claim only that amount and indicate the percentage.

X. Itemized Budget Justification (one-page limit): Salaries, supplies and animal care costs not justified may be deleted from the budget of an approved/funded proposal. The role of each investigator should be clearly defined. Investigator salary requests will be thoroughly scrutinized. Indicate and justify a percent effort on this grant for all individuals, including technicians, graduate students, etc., for whom MAF salary funds are requested.