BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. **DO NOT EXCEED FIVE PAGES.**

NAME: Jeremy J Berg

eRA COMMONS USER NAME (credential, e.g., agency login): JEREMYBERG

POSITION TITLE: Postdoctoral Researcher

EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, include postdoctoral training and residency training if applicable. Add/delete rows as necessary.)*

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	Start Date MM/YYYY	Completion Date MM/YYYY	FIELD OF STUDY
University of Wisconsin – Madison	B.S.	09/2007	12/2010	Biology - Evolution
University of California, Davis	PhD	09/2011	09/2016	Population Biology

A. Personal Statement

Broadly speaking, my research interests lie at the intersection of quantitative and population genetics. The primary focus of my research program over the near to mid-term is in understanding 1) the genetic architecture of complex traits, 2) the ways in which natural selection acts on such traits, and in particular 3) how these factors influence one another to produce the patterns of variation in complex traits that we observe in human and other populations.

I received my undergraduate degree in biology with an emphasis on evolution from the University of Wisconsin. There I received my first research experience performing wet lab work under the supervision of Dr. David Baum investigating the evolutionary and developmental genetics of petal spot formation in the California wildflower *Clarkia gracilis*, and was second author on the ensuing publication (Martins et al 2013). For my PhD work I made the transition to dry lab mathematical and statistical evolutionary genetics, working under the supervision of statistical population geneticist Dr. Graham Coop. I worked on two major topics: 1) the development and application of methods for using genome wide association study data to detect signals of polygenic adaptation (Berg and Coop 2014), and 2) coalescent modeling of selective sweeps originating from standing variation after an environmental change (Berg and Coop 2015). Through these projects I gained valuable expertise in probabilistic, computational and statistical approaches to making inference from population genetic data. This experience will be particularly valuable in the development of inference tools to understand how the parameters of the mathematical models I develop relate to population genetic datasets (i.e. in inferring factors responsible for governing the architecture of complex disease).

The work proposed here aims to develop a first principles approach to constructing models of complex disease evolution. The entirety of my work since beginning my graduate program has focused on adaptation, so the focus on disease is a new direction for me. This project will also involve a mathematical modeling component that is substantially more rigorous and involved than anything I have undertaken before, and this experience will help me to further develop the broad modeling skill set that I will need in order to develop my own independent research program. My sponsor, Dr. Guy Sella, is a widely recognized leader in the application of mathematical modeling to problems in population genetics and evolutionary biology, and is therefore an

optimal choice as a collaborator in this work and as a mentor to help me develop these skills.

1. Martins TR, Berg JJ, Blinka S, Rausher MD, Baum DA. Precise spatio-temporal regulation of the anthocyanin biosynthetic pathway leads to petal spot formation in Clarkia gracilis (Onagraceae). New Phytologist; 2013; 197:3. PMCID: PMC3540125
2. Berg JJ and Coop G. A population genetic signal of polygenic adaptation. PLOS Genetics 2014; 10:8. PMCID: PMC4125079
3. Berg JJ and Coop G. A coalescent model for a sweep of a unique standing variant. Genetics 2015; 201:2. PMCID: PMC4596678

B. Positions and Honors

Positions and Employment

01/2013 - 03/2013	Teaching Assistant, Genetics, University of California, Davis
09/2016 -	Postdoctoral Researcher, Columbia University

Honors

10/2013-09/2016	NSF Graduate Research Fellowship
2016	Society for the Study of Evolution Hamilton Award Finalist
2016	Merton Love Award for Outstanding Dissertation in Ecology and Evolution

C. Contributions to Science

The selective sweep paradigm has been a dominant force in efforts to understand the genetic of adaptation for the last quarter of a century. In this paradigm, one supposes that populations adapt when a new mutation arises that fulfills some pre-existing adaptive requirement not presently met by mutations segregating in the population, and that mutation subsequently “sweeps” to fixation, erasing genetic diversity in the surrounding area via the “hitchhiking effect”, and thereby leaving behind a signature which can be readily observed in population genetic data. While efforts to identify sites in the human genome where these selective sweeps events might have taken place found some examples, but it is widely believed that the number found are too few to account for the extent of adaptation taking place in natural populations. It has been proposed that instead, populations may frequently draw on genetic variation which is already present in the population at the time of an environmental change. Detection of these sort of events depends on the development of statistical methods which recognize qualitatively different patterns in population genetic data from those upon which the identification of classic selective sweeps is based. My graduate research contributions focused on characterizing these statistical patterns for two modes of adaptation which fall within this broader umbrella of “adaptation from standing variation”. These two phenomena are generally known as “polygenic adaptation” and “sweeps from standing variation”.

Polygenic Adaptation

My first major contribution was the development of a method to use genetic loci identified from genome wide association studies (GWAS), along with population genetic datasets, in order to detect subtle shifts in allele frequency that are coordinated across a very large number of loci scattered across the genome. Such a pattern is expected when natural selection acts on a phenotype which is highly polygenic, but had in general remained beyond the limit of detection due initially to a lack of the necessary data, and subsequently of the appropriate statistical methods. While a previous paper had demonstrated how to perform such a test in the case of a single pair of populations, our paper described the appropriate way to carry out such tests for an arbitrary number of populations with an arbitrary demographic history, properly accounting for multiple testing and other factors. Our paper also included an extensive discussion of the sort of confounding factors that could generate false signals of adaptation in polygenic traits, and the methods we propose have served as the basis for studies of polygenic adaptation in human populations for a range of traits including height, BMI, telomere length and bone mineral density. Through various collaborations I am involved in or aware of, the methods I developed are also being applied to study polygenic adaptation in a range of non-human systems, including

Arabidopsis, maize, and various species of pine.

1. Berg JJ and Coop G. A population genetic signal of polygenic adaptation. PLOS Genetics 2014; 10:8. PMCID: PMC4125079

Sweeps from Standing Variation

The second major contribution from my graduate work included a thorough analysis of the statistical footprint of so called “sweeps from standing variation”. In this scenario, one supposes that a change in the environment has caused a previously neutral variant segregating in the population to become strongly adaptive and sweep to fixation. However, because much of the signature of a classic selective sweeps depends on events which occur early in the sweep when the allele is still at low frequency (and because this is the part of an adaptive allele’s history where the sweep from standing variation model differs from the classical sweep model), the statistical signatures left behind can differ substantially between the two. I undertook a thorough analysis of the impact on patterns of allelic and haplotypic diversity left behind after a sweep of this kind, with a particular eye toward identifying the kinds of signatures that will be most useful for future methods developers seeking to identify such events.

1. Berg JJ and Coop G. A coalescent model for a sweep of a unique standing variant. Genetics 2015; 201:2. PMCID: PMC4596678

D. Additional Information: Research Support and/or Scholastic Performance

YEAR	SCIENCE COURSE TITLE	GRADE
	University of Wisconsin - Madison	
2007	Advanced General Chemistry	AB
2007	Multivariable Calculus	A
2008	Intro to Organic Chemistry	AB
2008	Linear Algebra and Differential Equations	A
2008	Ecology, Evolution, and Genetics	A
2008	Intermediate Organic Chemistry	B
2009	Cellular Biology Lab	B
2009	General Ecology	B
2009	Ecology, Evolution, and Genetics Lab	B
2009	Organismal Biology	B
2009	Paleobiology	B
2009	Cellular Biology	AB
2010	Organic Chemistry Lab	B
2010	Biology of Microorganisms	B
2010	Biological Interaction	A
2010	Evolutionary Biology	A
2010	Evolutionary Genomics	A
2010	Major Evolutionary Transitions	A
2010	Population Genetics	A
2010	Modeling in Population Genetics and Evolution	A

YEAR	SCIENCE COURSE TITLE	GRADE
	University of California, Davis	
2011	Principles of Population Biology (A)	A
2011	Mathematical Methods in Population Biology	A
2012	Principles of Population Biology (B)	A
2012	Probability	A
2012	Statistical Rethinking	A
2012	Principles of Population Biology (C)	A
2012	Stochastic Processes	A-