

Camelid genomics: Anticipating the future

Warren E. Johnson^{1*}, Polina Perelman^{1, 2}, and Stephen J. O'Brien¹

¹Laboratory of Genomic Diversity, National Cancer Institute, Frederick, MD 21702, USA

²Institute of Cytology and Genetics, Russian Academy of Sciences, Novosibirsk, 630090, Russia

*To whom correspondence should be addressed. E-mail: johnsonw@ncifcrf.gov

Technological and computational advances are rapidly propelling us into the genomics era. Through the study of the genomes of living species we shall develop a radically different perspective of biology as constraints that limited our approaches to biological enterprises will be lifted in unexpected and disorienting ways. One of the implications of this genomic restructuring is that we will more rapidly decipher and then modify biological processes. Where the genomics era will lead us is for the novelist or futurist to predict and describe. But as researchers, managers, policy makers, and world citizens it is important to prepare and be ready to take full advantage of the opportunities, because they are many. As a community interested in camelid species, it is our responsibility to embark on that process now.

There are numerous ways that we, as a community can influence the process, at the individual, institutional, and national level. Certainly the improvement and adaptation of new technologies, tools, and reagents will catalyze research activity. The most important recent development in camelid genomics was the selection of alpaca among the mammalian species chosen for whole genome sequencing with funding by the National Human Genome Research Institute. From this effort there is currently a draft 3.0-fold-coverage sequence map of the alpaca genome consisting of 11 million sequence-reads assembled into 504,327 contigs (segments), which cover 1.73 Gb (gigabase-pairs) or about 72% of the alpaca genome. Efforts to increase the assembly to a “finished” (>10-fold sequence) are ongoing and are likely to be completed by 2011. Additional genetic tools are under development for the alpaca, including an alpaca radiation hybrid (RH) panel for producing a framework physical map of genetic markers, a critical tool for annotation of the genome sequence, plus a SNP (single nucleotide polymorphism) array suitable for genome association studies.

The alpaca is the first camelid for which comprehensive genome map and whole genome sequence information is being developed, and thus will serve as a reference for all other domestic and wild camelids. The situation is similar to several other domestic species like cattle, pig, horse, dog, cat or chicken, where one representative of a genus provides essential map information, genome analysis tools and resources for the study of disease or population genetics of other related species. Soon genomic focus on old-world Bactrian and Dromedary camels will promote the integration of genetic approaches in the management and husbandry of domesticated camels, particularly in Eurasia and North Africa.

Although the development of molecular tools for the camelids is a crucial step in characterizing the genetic basis of phenotypic differences of economic interest, such as productivity traits, an equally necessary and perhaps more rate-limiting step is more traditional. Genetic tools are useless without biological questions and biological samples, so there is an

urgent need for coordinated efforts to collect suitable biological samples for genetic analyses that are matched with phenotypic data from our animals. The time is now for the gathering of blood samples with detailed biomedical and phenotypic feature annotation suitable for identifying the gene bases of these fascinating hereditary differences.

Fortuitously, the diversity among these species and more specifically within the camelid family provides numerous opportunities for comparative genomic assessment and for studying biologically and economically important traits. For example, camelids differ from other ruminants in several ways, including a three-chambered instead of a four-chambered foregut, an upper lip split into two separately mobile parts, and uniquely among mammals they have elliptical red blood cells and in addition to the normal antibodies a type of antibody lacking the light chain. These species have also developed unique ecological adaptations including extensive adaptations that permit them to live in harsh, arid environments or at high altitudes. Camelids also offer numerous opportunities for important discoveries as research animals in the genomics era, including as models for human congenital defects. One of our major goals will be to identify expressed genes and genetic markers for these and other phenotypic variation (e.g. parameters related with production-important traits, disease susceptibility, etc.).

Relative to other important agricultural species, camelid species are somewhat behind in utilizing genomic tools. However, the potential for the development and use of these tools are huge, and the recently initiated alpaca genome project is providing a strong launching pad for camelid genomic research. We now need to join as a community to provide the natural context in a way that bridges traditional political, cultural, and economic boundaries, and includes multidisciplinary approaches by a diverse range of specialists with the common goal of increasing our understanding and improving the health, welfare, and management of domestic and wild camelids.