

## Sequencing of reindeer (*Rangifer tarandus*) genomes: insights into evolution, domestication and adaptation

M.W. Betsha<sup>1,2</sup>, K. Pokharel<sup>1</sup>, M. Yao<sup>3</sup>, M. Honkatukia<sup>1</sup>, N. Mazzullo<sup>4</sup>, J. Peippo<sup>1</sup>, T. Reilas<sup>1</sup>, P. Soppela<sup>4</sup>, A. Stammer-Grossman<sup>4</sup>, F. Stammer<sup>4</sup>, K.H. Røed<sup>5</sup>, J. Kantanen<sup>1\*</sup>

<sup>1</sup> *Natural Resources Institute Finland, Jokioinen, Finland*

<sup>2</sup> *Department of Environmental and Biological Sciences, University of Eastern Finland, Kuopio, Finland*

<sup>3</sup> *BGI-Tech, Guangdong, China*

<sup>4</sup> *Arctic Center, University of Lapland, Rovaniemi, Finland*

<sup>5</sup> *Norwegian University of Life Sciences, Oslo, Norway*

\* *e-mail: juha.kantanen@luke.fi*

**Key words:** *de novo* sequencing, mitochondrial DNA, resequencing, subspecies, taxonomy

**Motivation and Aim:** Semi-domesticated reindeer (*Rangifer tarandus*) have pivotal economic, societal and cultural value for indigenous people and pastoralists in northern and subarctic regions in Eurasian. Currently, there exist several semi-domesticated and wild *Rangifer*-populations, the taxonomic status of which has been actively debated. To examine genetic diversity, domestication history, taxonomy and adaptation, we deep-sequenced and *de novo* assembled one reindeer genome, resequenced 23 other *Rangifer* sp. samples and analysed whole mitochondrial DNA of these individuals.

**Materials and Methods:** Genomic DNA of a 1-year-old Finnish male reindeer was sequenced at high coverage (100X) on the Illumina HiSeq2500 and 4000 platforms. Seven paired-end DNA libraries with insert sizes ranging from 170bp to 20kb were constructed. The genome assembly, annotation, and orthology analysis were conducted using a robust bioinformatics pipeline. In the resequencing approach, 23 semi-domesticated and wild reindeer and caribou were sequenced (10X) using Illumina HiSeq2500 platform. The data were subjected e.g. for the principal component analysis. The complete mitochondrial genome for reindeer was assembled using the MITObim v.1.9 software.

**Results:** A total of 300.5 Gb of clean data was assembled using SOAP denovo resulting into 256,454 scaffolds (N50 = 502 Kb) with cumulative scaffold length of 2.66 Gb and spanning 90 % of the estimated (2.9 Gb) genome size of reindeer. Using a homology based approach, the reindeer genome was predicted to harbour 27,332 protein coding genes, 98 % of which were functionally annotated. Reindeer displayed a heterozygosity level, which is 2.3 and 1.5 times higher than that found in taurine cattle or yak. The resequenced animals grouped into two main clusters: northern European and northern Russian/northern American. Our study provides new information on the architecture of mitochondrial genome of reindeer (16,451 bp).

**Conclusion:** The draft quality of the reference genome along with the annotations will provide important insights into the evolution and demographic history of the reindeer and taxonomy of *Rangifer* sp. Our findings suggest that there have been at least two domestication events in the history of reindeer.

**Acknowledgements:** The present study is supported by the Academy of Finland (Grant No. 286040).