Hereditary Genetics: Current Research

ISSN: 2161-1041

Bana et al., Hereditary Genet Curr Res 2018, 7:1
DOI: 10.4172/2161-1041.1000191

Opinion Article Open Access

The Red Deer *Cervus elaphus* Genome CerEla1.0: Sequencing, Annotating, Genes, Chromosomes

Nóra Á Bana[1,2], Anna Nyiri[1], János Nagy[2], Krisztián Frank[1,2], Tibor Nagy[1], Viktor Stéger[1], Mátyás Schiller[1], Péter Lakatos[4], László Sugár[2], Péter Horn[2], Endre Barta[1,5] and László Orosz[1,3*]

[1]*Agricultural Biotechnology Institute, National Agricultural Research and Innovation Center, Szent-Györgyi Albert str. 4, H-2100 Gödöllő, Hungary*

[2]*Department of Animal Breeding Technology and Management Faculty of Agricultural and Environmental Sciences, Kaposvár University, Guba Sándor str. 40, H-7400 Kaposvár, Hungary*

[3]*Faculty of Sciences, Department of Genetics, Eötvös Loránd University, Pázmány Péter ave. 1/c, H-1117 Budapest, Hungary*

[4]*Department of Internal Medicine, Semmelweis University, Korányi Sándor str. 2/a, 1083 Budapest, Hungary*

[5]*Department of Biochemistry and Molecular Biology, University of Debrecen, Nagyerdei Ave 98, 4032 Debrecen, Hungary*

[*]**Corresponding author:** László Orosz, Faculty of Sciences, Department of Genetics, Eötvös Loránd University, Budapest, Hungary, Tel: +36 1 411 6500; E-mail: orosz.laszlo@abc.naik.hu

Rec date: Jan 15, 2018; **Acc date:** Jan 19, 2018; **Pub date:** Jan 22, 2018

Citation: Bana NA, Nyiri A, Nagy J, Frank K, Nagy T, et al. (2018) The Red Deer *Cervus elaphus* Genome CerEla1.0: Sequencing, Annotating, Genes, Chromosomes. Hereditary Genet Curr Res 7: 191. doi:10.4172/2161-1041.1000191

Opinion

The first genome assembly CerEla1.0 (NCBI, MKHE00000000) for red deer *C. elaphus* has been presented in this paper. The CerEla1.0 could serve manifold Genome Wide Association Study (GWAS). Red deer is an emblematic member of the natural megafauna of the Northern Hemisphere and has been present in human culture since the Neolithic. Humans introduced and spread the species in the Southern Hemisphere. Red deer, the mythological Wonder Deer, also the Royal Game of the Middle Ages in Europe is surrounded by respect and amazement in many cultures. Today red deer is not only one of the most desired game, but also a farm breed for venison, velvet antler products and tonic. Red deer is getting recognized as a model animal for bone, osteoporosis and regeneration research, as well as for population and evolutionary studies.

Red deer stag DNA was sequenced by Illumina Technology at 74x coverage. The ALLPATHS-LG assembly of the reads resulted in 34.7 × 10^3 scaffolds. For building the red deer pseudochromosomes a pre-established genetic (recombination) map was used as the main anchor points. A nearly complete co-linearity appeared between the array of the deer map points/map marker scaffold sequences and the order and orientation of the orthologous sequences in the syntonic bovine regions. Syntenies were also conserved at the in-scaffold level. The final CerEla1.0 assembly contains 26108 scaffolds and contigs and spans 3.4 Gbp including the NNN-s inserted between contigs during the scaffolding and between scaffolds. In nearly all genomic segments the cM distances corresponded uniformly to 1.34 Mbp, due to the too many "NNN-s" inserted by the Allpaths-LG, 1.25-fold uniformly more than in the bovine homologous regions. Chromosomal rearrangements between deer vs. cattle were demonstrated. In the resulting red deer pseudochromosome sequences 2.8 million heterozygous SNPs, 365 thousand indels and 19368 protein coding genes were identified along with positions for centromerons. This *de novo* assembly demonstrates the utilization of an approach of dual references, i.e., when a target genome (here *C. elaphus*) has already a pre-established genetic map and is combined with the well-established whole genome sequence of a closely related species (here *B. taurus*). The reference genome CerEla1.0 of (*Cervus elaphus hippelaphus*) and its annotation, in accordance with fresh data from other programmes, are under continuous monitoring and updating. If the sequence data of an SNP-based map marker will be available, the updating CerEla1.0 will be possible using the approach described in this work. The sequence and the pseudochromosome complement of CerEla1.0 may provide a basis and a rich source for broader interests, including, among others, conservation genetics, refined evolution and population studies within the family Cervidae (e.g. fallow deer *Dama dama* or roe deer *Capreolus capreolus*) as well as in a wider neighbourhood of ruminants and Pecora. CerEla1.0 also provides a source for chromosome-specific microsatellite sets. A large number of SNP/heterozygotic sites were identified (2.8 × 10^6 SNVs, 3.6 × 10^5 indels) and aligned to the deer pseudochromosomes. Chromosome specific microsatellite sets may shed light on inbreeding/outbreeding, help in the identification of gene introgressions, of descents for autosomal, maternal and paternal lineages, forensic identification, or defining allelic compositions behind of phenotypes important, for example, in game management. The exploration of the genetic secret of record antlers becomes possible by Genome Wide Association Studies. The applications and utilizations in several fields of medical research (e.g. bone and osteoporosis research, organ development and regeneration, robust tissue proliferation/tumour biology) are also feasible [1].

Data availability

The raw reads have been deposited into the SRA database (SRR4013902). The reference genome sequence has been submitted to the NCBI database and can be accessed at the following accession number (MKHE00000000). The gene annotation and the variation tracks are available for browsing and downloading from the JBrowse web page http://emboss.abc.hu/wonderdeer/JBrowse

Reference

1. Bana NA, Nyiri A, Nagy J, Frank K, Nagy T, et al. (2018) The Red Deer Cervus elaphus Genome CerEla1.0: Sequencing, Annotating, Genes, Chromosomes. Mol Genet Genomics 293: 1-22.