

The first goat genome sets a good example for facilitating de novo assembly of large genomes

December 23 2012

In a collaborative study published online today in *Nature Biotechnology*, researchers from Kunming Institute of Zoology, Chinese Academy of Sciences, BGI, and other institutes, have completed the first genome sequence of domestic goat by a robust approach integrated with next-generation sequencing (NGS) and whole-genome mapping (WGM) technologies. The goat genome is the first reference genome for small ruminant animals and may help to advance the understanding of distinct ruminants' genomic features from non-ruminant species. This work also yields a valuable experience for facilitating the de novo assemblies of large, complex genomes in the future.

Goats are recognized as an important member of the world [livestock industry](#), and with many unique biological features. They are an important economic resource in many developing countries around the world, especially in China and India. However, despite their agricultural and biological importance, breeding and genetic studies of goats have been hampered by the lack of a high quality reference [genome sequence](#). The goat [genome](#) sequence will be useful for facilitating the identification of SNP markers for marker-assisted breeding, and improving the utility of the goat as a biomedical model and bioreactor.

With the availability of next-generation sequencing (NGS), draft assemblies are easy to generate nowadays. However, to finish a sequence to the chromosome level remains a hard nut to crack. In this study, the

results show that a single NGS platform, when combined with whole-[genome mapping](#) technology, could produce a finished assembly much faster and with high quality than other currently available mapping strategies such as BACs or FISH. Through this integrated approach, researchers obtained the ~2.66 Gb goat reference genome from a female Yunnan black goat.

Transposable elements (TEs) are major components of mammalian genomes and contribute to gene and/or genome evolution. The TEs in goat genome are similar to those of cattle, and contain large numbers of ruminant-specific repeats, such as SINE-tRNA and SINE-BovA. It is reported that SINE-BovA repeat expanded primarily in the cattle genome. However, in this study, researchers found the SINE-tRNA repeat expanded specifically in the goat genome.

Through constructing a phylogenetic tree among goats, cattle, horses, dogs, opossums and humans, researchers found the goat shared a common ancestor with cattle about 23 million years ago. Further comparison analysis revealed 44 rapidly evolving [genes](#) under positive selection, seven of which are immune-related genes and three are pituitary hormone or related genes. The immune-related genes identified also exist in cattle. The findings suggest that the rapid evolution of pituitary hormones may be related to the different features between goat and cattle in milk production, development rates of the fetus and/or hair variation.

The major histocompatibility complex (MHC) plays an important role in the immune system. In this study, the goat MHC was found to be located on chromosome 23 and contains two regions with length of 2.25 Mb and 360 kb, respectively. With the high quality genome assembly, further understanding of the goat MHC will be useful for immunological studies and vaccine development.

One of the distinguishing characteristics of mammals is the protective growth known as hair. It is produced by hair follicles within the skin, which could provide either protection (guard hairs) or insulation (underfur). The two major hair follicles include the primary hair follicle that produces only coat hair in all mammals, and the secondary hair follicle that can produce the cashmere or "fine hair" in certain mammals, including goats and antelopes. Despite a 2,500-year history and the extent of raw cashmere production, people are lack of understanding of the molecular mechanisms of cashmere formation and development.

Researchers conducted transcriptomic analysis on the primary and secondary follicles of a cashmere goat, revealing 51 genes that are differentially expressed between the two types of hair follicles. Keratin and keratin-associated proteins are the main structural proteins of hair fibers, determining the quality of fiber together. In the study, 29 keratin genes and 30 keratin-associated protein genes were detected in both types of follicles. Interestingly, they found two keratin genes and ten keratin-associated protein genes were consistently differentially expressed between primary and secondary hair follicles, suggesting that the keratin-associated protein genes may be more important in determining the structure of cashmere fibers. In addition to the keratin genes and keratin-associated protein genes, researchers also found several enzymes of amino acid biosynthesis, with implications in regulating primary hair growth and hair cycle.

Xun Xu, Deputy Director of BGI, said, "The [goat](#) reference genome is an important stepping stone in the molecular breeding of cashmere goats, and will help to advance the comparative studies on ruminants. The transcriptomic analysis on the primary and secondary follicles will open a new way for better improving the quality cashmere wool."

Provided by BGI Shenzhen

Citation: The first goat genome sets a good example for facilitating de novo assembly of large genomes (2012, December 23) retrieved 9 April 2024 from <https://phys.org/news/2012-12-goat-genome-good-de-novo.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.