# Plant and Animal Genome Sequencing and Bioinformatics Solutions at MedGenome

Overview

Over the last decade, rapid improvements in DNA sequencing technology have transformed the field of genomics and is now an essential tool in genetic and clinical research. The ability to generate chromosome-level genomes or sequence specific genomic regions of interest has tremendous implications in a variety of applications including development, disease, evolution, and crop breeding. At MedGenome, we have leveraged our experience in genomics technologies and developed a comprehensive, customizable service for high-quality genome assembly and annotation.

A. Assembly

Platform	Coverage
Long-reads	50-100x
Short-reads	50x
Hi-C	100x

Workflow
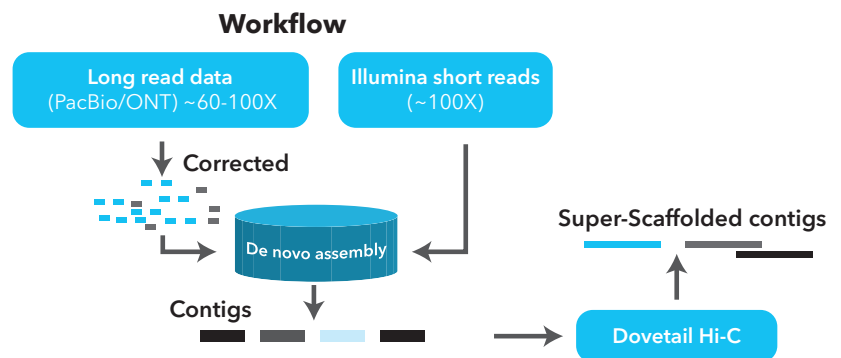
B. Structural Annotation
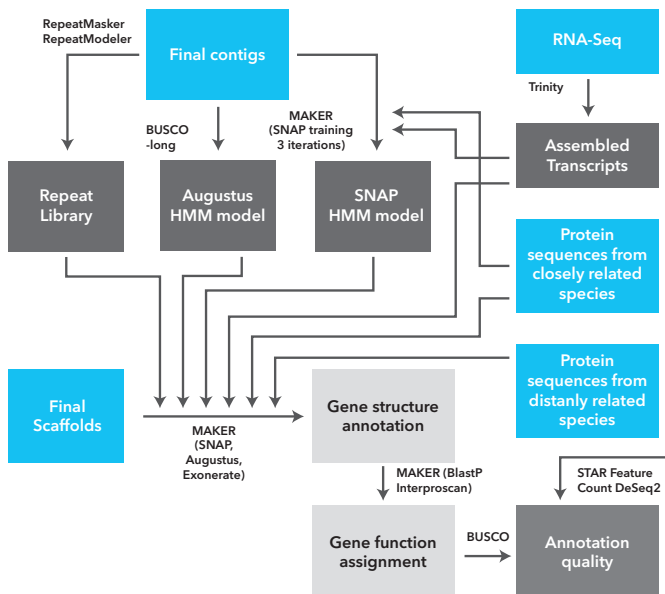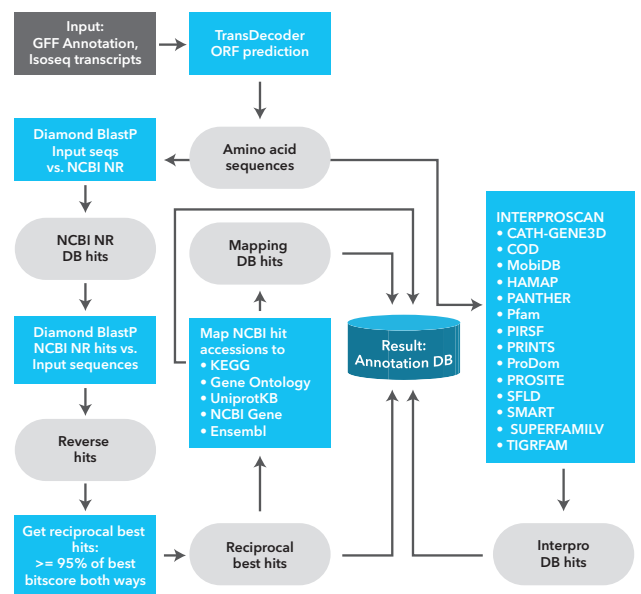

C. Functional Annotation

Figure 3. *A) Starting from a variety of tissues or cells, high-molecular weight DNA is extracted for long- and short-read library preparation and sequenced at the recommended depth. Tissues or cells will also be processed for Hi-C library preparation and sequenced to about 100x depth of a given genome. Long and short reads will be used for generating an initial draft genome assembly that will be assessed for accuracy and contiguity. This draft genome will then be scaffolded using Hi-C data. B) Gene models will be predicted in the final genome using various types of evidence such as transcriptome and protein sequence data along with ab initio gene prediction tools. The predicted gene models will be assessed for quality and refined over multiple iterations. C) The final gene model set will then be annotated using publicly available databases for homology-based function assignment.*

Case Study: The Indian cobra reference genome and transcriptome enables comprehensive identification of venom toxins

Suryamohan and colleagues report a *de novo* near-chromosomal genome assembly of *Naja naja*, the Indian cobra, a highly venomous, medically important snake. With a scaffold N50 of 223.35 Mb, this is the most contiguous reptilian genome published till date. A comprehensive annotation pipeline identified 23,248 predicted protein-coding genes, of which 12,346 were venom-gland-expressed genes and constitute the 'venom-ome'. Venom gene-specific gene models were generated to identify 139 genes from 33 toxin families. This will aid in synthetic venom production through recombinant toxin expression and will aid in the rapid development of safe and effective synthetic antivenom. Additionally, this genome will serve as a reference for snake genomes, support evolutionary studies and enable venom-driven drug discovery.
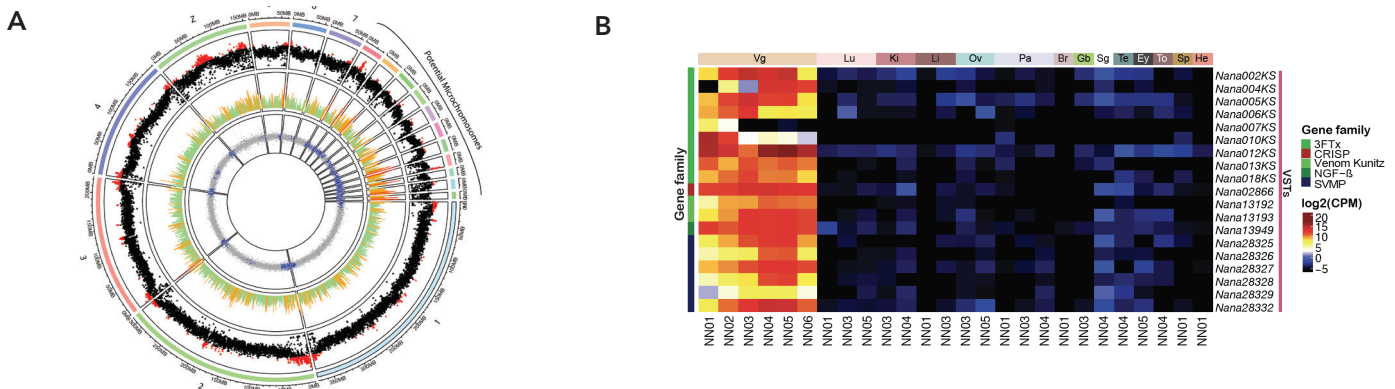
Link to Nature Article: https://www.nature.com/articles/s41588-019-0559-8

A

B

Figure 2. *A) Circos plot of the chromosomal Indian cobra genome assembly depicting the chromosomes, repetitive content, GC content and other characteristics. B) Transcriptome data overlaid on the high-quality genome allowed identification of genes primarily expressed in the Indian cobra venom gland*

Highlights of our Services and Bioinformatics Capabilities

- Support with experimental design and selection of appropriate genome assembly workflow
- High-molecular weight DNA extraction (blood, flash frozen tissue, plants)
- End-to-end solution in library preparation, sequencing, genome assembly and annotation

Analysis deliverables

- Raw data (FASTQ for genome sequencing, BAM files from Hi-C)
- Chromosomal-level scaffolded genome assembly and assembly statistics
- Repetitive element annotation and statistics
- Genome annotation and quality assessment statistics

Advanced Deliverables

- Custom gene family annotation
- Differential gene expression analysis
- Whole genome comparative analysis

Our Partner

agrigenome

Genome and transcriptome sequencing	Genome and transcriptome assembly	Annotation	Review data	Delivere data to customer
8-10 weeks	8-10 weeks	8-10 weeks	2 weeks	