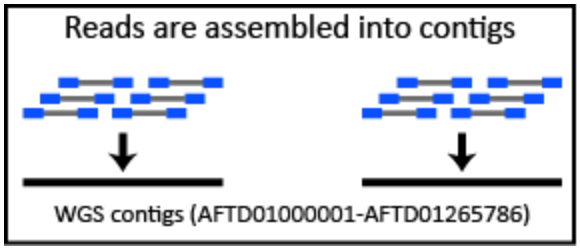
December 2011
Version 1.0

Tutorial 1 – Overview of CHO Genome Resources

CHO-K1 Genome Assembly and Annotation

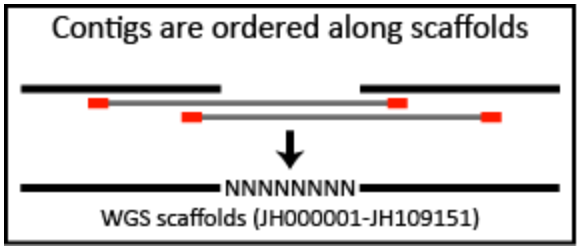
The CHO-K1 genome (CriGri_1.0) was assembled from Illumina sequencing data by a whole genome shotgun (**WGS**) approach in which the CHO genome is sheared into smaller pieces, sequenced, and then assembled into a draft genome sequence. The CHO-K1 genomic information is available at NCBI GenBank using the accession codes described below.

Overlapping reads from short-insert (< 500 bp) paired-end libraries are aligned and assembled into **contigs**. **Contigs** are continuous sequences containing no gaps.



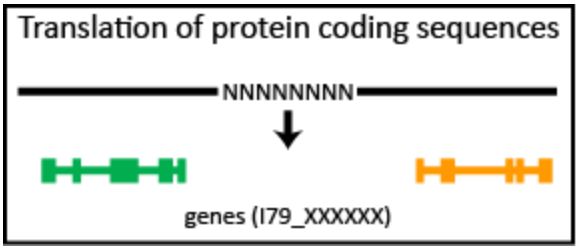
The CHO-K1 WGS contigs have accession numbers of AFTD01000001-AFTD01265786.

Reads from all paired-end libraries, in increasing order of insert size, are used to order contigs and assemble **scaffolds**. **Scaffolds** may contain gaps (size can be estimated from library insert size).



The CHO-K1 WGS scaffolds have accession numbers of JH000001-JH109151.

The WGS scaffolds are annotated by prediction of coding sequences (**CDS**) and protein sequences.



All genes derived from this genome sequencing project have been assigned the locus tag prefix I79.

Gene symbols were assigned to CHO protein-coding genes in the Chinese hamster genome database based on annotation of homologous proteins.

As of December 2011, The Chinese hamster genome database contains information for the CHO-K1 genome (Xu et al. 2011) and the Chinese hamster mitochondrial genome (Partridge et al. 2007).

December 2011
Version 1.0

Genome Sequence Publications

Xu X, Nagarajan H, Lewis NE, Pan S, Cai Z, Liu X, Chen W, Xie M, Wang W, Hammond S, Andersen MR, Neff N, Passarelli B, Koh W, Fan HC, Wang J, Gui Y, Lee KH, Betenbaugh MJ, Quake SR, Famili I, Palsson BO, Wang J (2011) The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line. *Nature Biotechnology* 29(8): 735-741.

Partridge MA, Davidson MM, Hei TK (2007) The complete nucleotide sequence of Chinese hamster (Cricetulus griseus) mitochondrial DNA. *DNA Sequence* 18(5): 341-346.