

Unannotated genome

Alignment of a family of 3D7 VSA proteins

translated genome (6 frames)

HMMer hits (coordinates)

List of all putative protein sequences

Curated list of protein sequences

DNA sequences

- HMM profile built with the 3D7 alignment.
- Genome searched for homologues with the HMM profile.

Custom made Perl scripts extract the protein sequences using the coordinates.

- All sequences aligned and verified with ClustalW.
- Sequences labelled "gene", "pseudogene" or "truncated".

Control : The new protein sequences are used to build a HMM profile and the entire process is repeated.

Perl script converts the coordinates from the translated genome to extract the DNA sequence from the original genome

