CTT DUPLICATED GENE ANNOTATION

TEAM NAME: PerlSquad
perlsquadou@gmail.com

- Alex Mayle am218112@ohio.edu

- Brian Reynolds br058215@ohio.edu

- Eric Keep ek349112@ohio.edu

- Favour Ogundare fo893113@ohio.edu

- Robert Smith rs659612@ohio.edu

CLIENT: Zhihua Hua hua@ohio.edu

```
Availability: 6:30 - 7:00 pm Tuesday/Thursday

Preferred Method of Communication: E-mail
```

INITIAL PROJECT DESCRIPTION

> *What is the maximum number of genes encoded in each genome?*

Students work to answer this question by 'finalizing' a PERL-based annotation code, >that implements a closing Target Trimming algorithm, to >re-annotate the missing genes in their preferred genomes.

INTENDED USER BASE

Academia, Bioinformatics Personnel.

FIRST ADOPTER USER GROUP

Academia, Bioinformatics Personnel.

EQUIPMENT NEEDS

```
Machine running Perl with required dependencies:
    Bioperl, BioPro, GeneWIse, Blast, HMMER


Current Requirements:
    - Server to compile resulting executables,
    - Provided by Dr. Hua.


    - Create account on server so students can access fil
es on server.


    - Tools/Server for Web-Site for developed stand alone
 program.


    - Eventually, student research assistants along with
other
       professors for alpha/beta testing
```

FEATURES

> *Executable File:- specific to eukaryotic systems :- application is not*

web based

- Application takes input from user in the firm of annotation file, sequence of proteins and family name.

- Application finds genome, then finds where previous annotation file missed genes and or potential new genes, translate new signs, using dependency GeneWise to use previous annotated sequences.

- Application outputs how many genes belong to input's family. Genomic DNA Sequence, Annotation Transcript sequence, Peptide Sequence

INITIAL USE CASES / USAGE SCENARIOS

To annotate duplicated genes that may have been missed in the initial annotation

NON-FUNCTIONAL REQUIREMENTS

Security is not a priority, all code used will have open source propiety rights Dependencies and requirements for the program will be listed on the eventual site and the executables FAQ.