

Code Documentation/User Guide

Chromosome 8

Jennifer J. Stiens
j.j.stiens@gmail.com
Bioinformatics, Birkbeck College
28 April, 2018

Data Access Routines

Overall Schema:

These scripts are utilised by the middle-tier scripts to query the database for specific gene information. The scripts utilise a config_db.py file which must be set up by the database host and includes database login parameters.

Instructions for middle-tier (business logic layer):

Import the directory and module into the middle-tier script.

For example:

```
from data_access import seq_query
```

To call the functions within the script, use 'module_name.function_name(arg)'.

For example:

```
results = seq_query.seq_query(acc)
```

Module descriptions:

I. list_query

Description:

=====

This program will access database using pymysql to obtain list of all genes and gene identifiers from chromosome 8. Returns tuple listing the four identifiers (accession number, gene name, protein product, chromosome location) for each gene in database.

Usage:

=====

list_query

II. seq_query

Description:

=====

Pymysql query program used to return sequence information for specified gene using accession number.

Usage:

=====

seq_query ACC

Functions:

=====

seq_query(acc) Returns tuple with accession number and gene sequence.

III. coding_query

Description:

=====

Pymysql query program used to return coding information for specified gene using accession number.

Usage:

=====

coding_query ACC

Functions:

=====

coding_query(acc) Returns tuple: (accession number, codon start, exon positions)