

Mapping Load

Author: Lori Corbani

Created: May 23, 2003

Last Modified: June 23, 2003 14:14

1 Purpose of Document

To describe the Mapping Load product.

2 Overview

The purpose of the Mapping Load is to:

1. To load new or update existing experiment records (one per chromosome) into MGI Mapping tables and set these attributes:
 - Mapping Reference (J:)
 - Experiment Type
 - Gene
 - Assay Type
2. To assign MGI Accession IDs to the new experiments.
3. To update a Gene's Chromosome and/or Band.

3 Inputs

1. Reference (J:)
2. Experiment Type
3. A tab-delimited file in the following format (no column headings):

Table 1: Input File

Field	Name	Example	Type	Max Length
1*	MGI Accession ID	MGI:12345	string	30
2*	Chromosome	1	string	2
3*	Update Marker Chromosome?	yes no	string	3
4	Band	A1	string	20

Table 1: Input File

Field	Name	Example	Type	Max Length
5*	Mapping Assay	assembly	string	80

Fields with "*" require a non-null value. Other fields must be present but can be blank.

Field 2:

Any mouse chromosome

4 Outputs

1. 5 BCP files (inputfile name is prepended to each bcp file name):
 - MLD_Expts.bcp
 - MLD_Marker.bcp
 - MLD_Expt_Marker.bcp
 - MLD_Notes.bcp
 - ACC_Accession.bcp
2. Diagnostics file (inputfile.MMDDYYYY.diagnostics) that includes all SQL commands issued by the load.
3. Error file (inputfile.MMDDYYYY.error) that includes any data validation errors.

5 Processing

1. Verify Mode.
 - if mode = full: delete existing experiment records; create new experiments
 - if mode = incremental: append data to existing experiments
 - if mode = preview: set "DEBUG" to True
2. Verify the J:is valid.

If the verification fails, write the error to the error file and skip the record.
3. Create the master Experiment records and Mapping Accession IDs (MLD_Expts, ACC_Accession).
4. For each line in the input file:
 1. Verify the MGI Accession ID is valid (not null).

If the verification fails, write the error to the error file and skip the record.
 2. Verify the Mapping Assay is valid (not null).

If the verification fails, write the error to the error file and skip the record.

3. Verify the Chromosome is valid (not null).
If the verification fails, write the error to the error file and skip the record.
4. Create the MLD bcp records (MLD_Marker, MLD_Expt_Marker).
5. Load the bcp files into the database.

6 Usage

To execute:

Python Program: mappingload.py

Parameters:

- -S server
- -D mgd database
- -U user
- -P password file
- -M mode (*full, incremental or preview*)
- -I input file
- -R reference (J: in format #####)
- -E experiment type ("TEXT-Physical Mapping")