caArray 009 - Highlights for Building caArray's MAGE-TAB Annotation Files

Highlights for Building caArray's MAGE-TAB Annotation Files

Topic: caArray Usage

Release: Up to caArray 2.2

Date entered: 02/17/2009

Contents of this Page

Introduction

caArray has been designed to handle MAGE-TAB data using certain conventions. The caArray User Guide devotes a section in Chapter 7 – Submitting Data to Importing MAGE-TAB and an entire appendix to working with MAGE-TAB files in caArray. The appendix specifically includes much specific detail from recognized data fields to validation rules essential for importing MAGE-TAB files. Highlights are outlined below to offer you a quick start.

Recognized Fields and Validation Rules for MAGE-TAB Annotation Files

- You can describe your experiment in great detail with the convenience of managing MAGE-TAB files in spreadsheet format. For example, a biomaterial (source, sample, extract, or labeled extract) named in one column can be followed by any number of characteristics columns containing annotation data about that biomaterial.
  It should be noted that caArray has requirements about which fields should be used and the order in which they must be listed. For example, in an SDRF file, the biomaterials columns must follow the order of Source Name - Sample Name - Extract Name - Labeled Extract Name. See Biomaterial Column Order in an SDRF.
- Controlled vocabularies should be used in IDF and SDRF files. The commonly used vocabularies can be found in MGED Ontology (MO) or the NCI Thesaurus. Each TERM SOURCE REF should have an entry in the IDF file. If term source is unknown, use "caArray" in the TERM SOURCE REF column. See Term Source REF.
- All column types are not mandatory in the SDRF. caArray can auto-generate missing biomaterials and associate protocols intelligently.

For a detailed list of recognized fields and validation rules for IDF and SDRF files, refer to Appendix A in the User's Guide.

MAGE-TAB Files: Upload and Import

See Importing MAGE-TAB Data in the caArray User's Guide for more information about all points noted in this section.

1. MAGE-TAB Files must be imported into experiments. That is, an experiment needs to be created before MAGE-TAB files are imported.
2. Although only one IDF is allowed per import session, multiple IDF files are allowed per experiment.
3. IDF, SDRF(s), data file(s) referenced in SDRF files must be validated and imported together.
4. Data files that are not referred to by SDRF files can also be imported into the same experiment, but they must be uploaded, validated and imported in a separate session from IDF/DSRF files.
5. MAGE-TAB ADF and Data Matrix files can be uploaded, but caArray does not parse these files. the exception is copy number data matrix files, which caArray does parse. See About File Types in caArray.

SDRF File Has Higher Priority

Annotations can be configured via the Annotations tab in caArray or via an SDRF file (Annotations Tab vs. MAGE-TAB Annotation Files in caArray). The designation in the SDRF is authoritative. For example: The tissue site for one sample is set to be "Lung" from the annotation interface, but it is set to be "Brain" in the SDRF file. Upon data upload, the tissue site will be shown as "Brain", since SDRF file has higher priority than "Annotation" interface. See SDRF Decides Raw versus Derived Data File.

Known Bug List in caArray

Known Bug List for MAGE-TAB Files

Error Message	Known Bugs
"Term Source Ref is not preceded by valid data type"	Factor Value can not have "Term Source REF" followed

Further Readings on MAGE-TAB Files

- For more information on MAGE-TAB Files, refer to caArray 007 - MAGE-TAB Files.
- For more information on when to use MAGE-TAB Annotation Files, refer to caArray 008 - Using the Annotations tab or MAGE-TAB annotation files to annotate an experiment.
- For more information on How to upload MAGE-TAB files, refer to caArray 002 - Uploading MicroArray Gene Expression Data into caArray.

Have a comment?

Please leave your comment in the caArray End User Forum.