

caArray 060 - Exporting experiment annotation and data file into MAGE-TAB and GEO SOFT formats

Question: How do I export experiment annotation and data file into MAGE-TAB and GEO SOFT formats?

Contents

- [Answer](#)
 - [caArray Download Options](#)
 - [Exporting to MAGE-TAB](#)
 - [Exporting to GEO SOFT](#)
 - [Option 1. Annotation file only](#)
 - [Option 2. GEO SOFT submission package](#)
 - [Requirement for Exporting to GEO Soft](#)
 - [Samples of MAGE-TAB Template](#)
- [Have a comment?](#)

Topic: caArray Usage

Release: caArray 2.3.1 and later

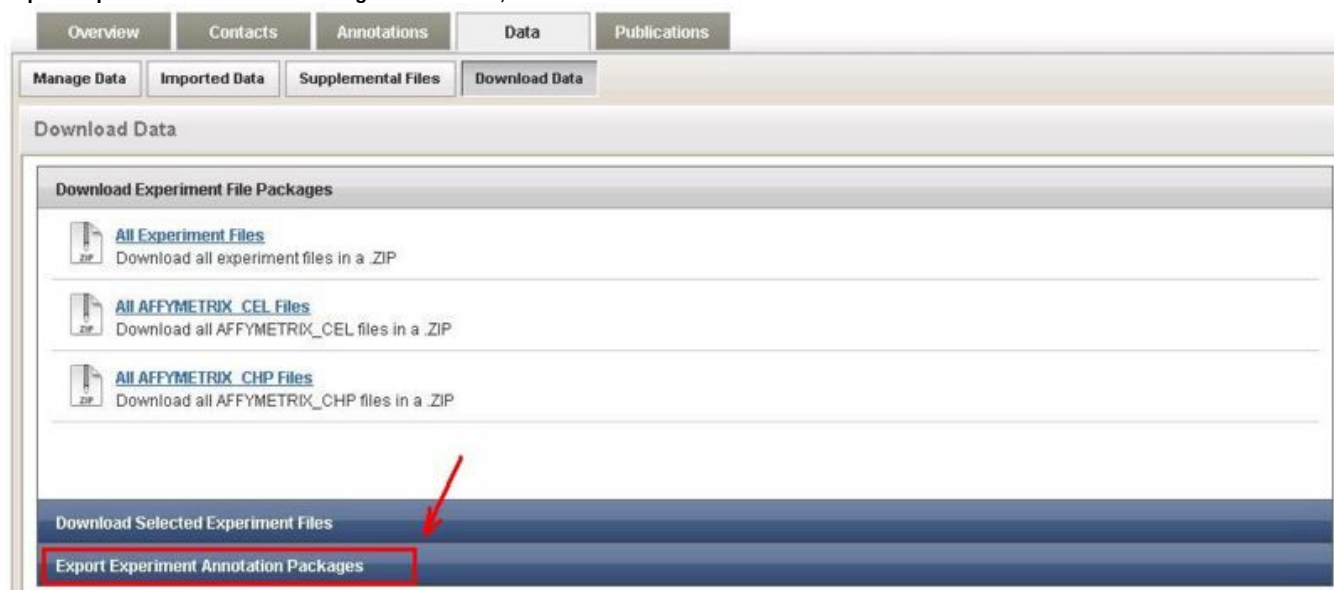
Date entered: 5/20/2010

Answer

caArray Download Options

Starting with caArray 2.3, the **Export Experiment Annotation Packages** option in caArray allows users to export the experiment annotations (and optionally data) into MAGE-TAB format and GEO SOFT format for GEO submission.

Export Experiment Annotation Packages under Data, Download Data



Exporting to MAGE-TAB

The user can generate MAGE-TAB files that consolidate the related annotation information about the experiment in the caArray repository.

To perform this task, click the MAGE-TAB (Annotations Only) link.

Links to Select an Option to Export Experiment Annotation Packages

Download Data

Download Experiment File Packages

Download Selected Experiment Files

Export Experiment Annotation Packages

You may choose to download experiment data in one of two GEO SOFT export types, or in MAGE-TAB format.

Microarray and Gene Expression Tabular Format (MAGE-TAB)

 [MAGE-TAB \(Annotations Only\)](#) ← 1

Export experiment annotations in MAGE-TAB IDF and SDRF format.

Gene Expression Omnibus (GEO) Simple Omnibus Format in Text (SOFT)

 [GEO SOFT \(File Only\)](#) ← 2

Export experiment annotations in GEO SOFT format.

 [GEO SOFT \(Submissions Package\)](#) ← 3

Export experiment annotations in GEO SOFT format, packaged with experiment data files.

The MAGE-TAB generated is a live export of the current state of the experiment's annotations, capturing all previous changes made via MAGE-TAB import and/or the Annotations GUI.

This process will generate an IDF and a SDRF file containing information about the experiment including biomaterial-hybridization-data chains, biomaterial characteristics and material type, source providers, labeled extract labels, experiment title and description, and term source (vocabulary) information. Export of other attributes like experimental factors, protocols, publications, persons, etc. is deferred.

Exporting to GEO SOFT

The export to GEO SOFT option allows the user to generate a data package that will meet the requirement for [GEO SOFT submission](#). This option is currently available only for Affymetrix experiments in caArray v.2.3.1. There are two options to export to GEO SOFT:

Option 1. Annotation file only

caArray validates that all the mandatory annotations are present, translates the experiment annotations into GEO SOFT format and prompts users to save the generated .soft.txt file (see **Arrow 2 in Figure 2**).

Option 2. GEO SOFT submission package

This option allows the user to export both annotation files and data files. caArray first validates that all the mandatory annotations are present, then translates the experiment annotations into GEO SOFT format and prompts users to save an archive containing the .soft.txt file, and all raw, derived and supplementary files (see **Arrow 3 in Figure 2**). If the archive is less than 2 GB, the package is provided in the .zip format; otherwise, it is provided in tar.gz format.

This export option is **available** only for the experiments whose annotation and data files were uploaded with MAGE-TAB format.

Requirement for Exporting to GEO Soft

Before launching the GEO SOFT Export, caArray will validate the experiments and make sure the following requirements are met:

1. The experiment must have at least one array design.
2. The provider for the array designs must be Affymetrix.
3. All array designs associated with the experiment must be ones for which caArray has the GEO accession.
4. Every hybridization must have at least one raw data file.
5. Every hybridization must have a derived data file of type AFFYMETRIX_CHP.
6. The following protocols must be present: extract, label, hybridization, scan and data processing. (These are entered via MAGE-TAB.)
7. There must be at least one characteristic or factor value that is present for every hybridization somewhere in its biomaterial chain.
8. For every biomaterial chain, the Material Type of the extract or labeled extract must be present.
9. For every biomaterial chain, the label of the labeled extract must have a non-empty, non-null value.

10. caArray informs the user of the result of the validation, and provides the list of errors if the validation did not succeed.

Samples of MAGE-TAB Template

Only the experiments entered into caArray via MAGE-TAB file upload are eligible for GEO SOFT Export. Various knowledge base entries (refer to [caArray 037 - MAGE-TAB Files](#)) have been created explaining MAGE-TAB format files and how to upload MAGE-TAB them. A sample package, including datafiles and MAGE-TAB annotation files (.idf and .sdrf files), can be downloaded below to help the user get a quick start: templates for [MAGE-TAB format and related datafiles](#).

An adventurous user may also download MAGE-TAB annotation files directly from caArray as explained in this [article](#).

Have a comment?

Please leave your comment in the [caArray End User Forum](#).