

caArray Roadmap

The list below outlines the priority features ordered and grouped by priorities.

Development Focus - Features planned for 2.5.0

The following features are planned for caArray 2.5.0:

- Large data support
 - Upon upload, store files (as-is) on the file system instead of in the database.
 - Provide an Uploader client that allows the User to select a large number of files from a local directory and upload to caArray, with support for "pause" and "resume" and failure recovery. (E.g., Flickr Uploadr)
 - Break import into smaller transactions to avoid long-running transactions and to avoid MySQL transaction import limit on parsed data values.
- Other
 - Plugin architecture that supports easy addition of new parsers as well as different file storage options
 - Update the technology stack (Jboss 5.1, Java 6)

Development Focus - Features planned beyond 2.5.0

The following features are next in line after caArray 2.5.0:

- Grid security and Single sign-on to support TRANSCEND
- Audit trail to support TRANSCEND
- File-level security (if necessary for TRANSCEND)
- Maintenance and Bug fixes including the Collaborator group defect which prevents collaborators from seeing files that have been uploaded but not yet imported.

Major Features in Scheduled and Past Releases

The table below outlines the major features in **scheduled and past release(s)**.

Version	Development Focus	Release Date
2.4.1	<p>This release represents a feature release of the caArray 2 software. Highlights of caArray 2.4.1 are:</p> <ul style="list-style-type: none">• Visibility into the queue of import and validation jobs, and where your job is in relation to other jobs in the system. Ability to cancel a queued job. For details about this feature, refer to the Submitting Data to an Experiment chapter in the caArray User's Guide.• Search for experiments by publication author or PubMed ID.• Organisms are now restricted to terms from the NCBI Taxonomy.• Bug fix to allow proper import of Agilent array design files where the chromosome name is not in a strictly numeric or X/Y format.• The installer no longer depends on an internet connection.• Fix to bug where experiments with multiple array designs may cause copy number data to get imported as "Imported (Not Parsed)".	September 2011
2.4.0	<p>This release represents a feature release of the caArray 2 software: Focus is around data parser. The following data parsers will be included in caArray 2.4.0:</p> <ul style="list-style-type: none">• Agilent:<ul style="list-style-type: none">◦ GEML/xml array designs (aCGH, gene expression and miRNA)◦ Raw TXT data files (aCGH, gene expression and miRNA)• Nimblegen - Community Code Contribution:<ul style="list-style-type: none">◦ NDF array designs◦ Pair Report (raw and normalized) data files• Illumina:<ul style="list-style-type: none">◦ BGX/TXT array designs◦ Gene expression: Sample Probe Profile TXT files with unique Probe_Id◦ Genotyping: Processed matrix TXT files with unique IlmnID values• Affymetrix:<ul style="list-style-type: none">◦ AGCC/Command Console formats for CDF, CEL and CHP files.◦ CNCHP files with copy number and LOH data.• Copy number data in MAGE-TAB Data Matrix format	August 2010
2.3.1	<p>This release represents a feature release of the caArray 2 software: Focus is around data export to GEO. Highlights of caArray 2.3.1 include:</p> <ul style="list-style-type: none">• caArray will allow the user to export an experiment into GEO SOFT format. This SOFT package can then be submitted to GEO.• The initial release will support single-channel experiments that have Affymetrix CHP data files, and that are based on array designs (platforms) already in GEO.• Focussing on CHP files allows us to package the native data files in the SOFT package as opposed to having to derive data tables to include in the SOFT submission.	February 2010
2.3.0	<p>This release represents a feature release of the caArray 2 software: Focus is around new programmatic interface to caArray application. Highlights of caArray 2.3.0 include:</p> <ul style="list-style-type: none">• 15578: Redesign Grid API to support backwards compatibility via external data model• 18908: Keep an audit trail of permissions changes made to an experiment.	October 2009

2.2.1	<p>This release represents a minor release of the caArray 2.x software: Focus is around access control and ownership of experiments. Highlights of caArray 2.2.1 include:</p> <ul style="list-style-type: none"> • 19324: Genepix data files must require an accompanying SDRF and sample names must not be mandated. • 15443: Ability to change experiment ownership • 19532: System administrator should be able to change collaborator group ownership. 	April/May 2009
2.2.0	<p>80 features/issues were addressed</p> <ul style="list-style-type: none"> • Highlights: <ul style="list-style-type: none"> ◦ Sample search: <ul style="list-style-type: none"> ■ The user can now search for biomaterials (Samples and Sources) in the system. ■ The keyword search can be applied to one of a list of standard or predefined categories, or to arbitrary categories previously imported into the system via Characteristic[] columns in a MAGE-TAB SDRF document. ◦ Export experiment annotations in MAGE-TAB format: <ul style="list-style-type: none"> ■ The user can generate and download a MAGE-TAB file set describing the annotations of an experiment. ■ The generated MAGE-TAB includes the relationship between biomaterials, hybridizations and data files, and also characteristics of the biomaterials. ■ The generated MAGE-TAB does not contain information about experimental factors, protocols, publications or people. ◦ Bulk-update of experiment annotations using MAGE-TAB import: <ul style="list-style-type: none"> ■ When a MAGE-TAB file set is being imported, the System recognizes references to existing biomaterials and hybridizations by name. ■ Biomaterial linkages can be modified in an additive way via MAGE-TAB import. Existing linkages and biomaterials cannot be deleted via MAGE-TAB import. ■ Biomaterial characteristics can be modified or new characteristics added via MAGE-TAB import. ◦ Support for GEO and ScanArray data: <ul style="list-style-type: none"> ■ Data files can be designated as GEO SOFT or GEO GSM and then imported (without parsing) and associated with samples. ■ ScanArray CSV data files can be imported (without parsing) and associated with samples. ◦ Enhanced usability of permissions management workflow: <ul style="list-style-type: none"> ■ While setting sample-selective access for an experiment, the user can do a keyword search for Samples against a set of categories, select the Samples of interest, and grant access to the selected Samples as a whole. 	January 2009
2.1.1	Become part of the Life Science Distribution	October 2008