caArray Road Map

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

Development Focus
High Priority Items: To be implemented after 2.4.0 release.
The following are high priority features that will be incorporated in future caArray releases: • Fix collaborator group access to unimported files • Integrate caArray with biospecimen systems (caTissue) • Simplify the import of large data sets • Support Grid API access to protected data • Enable development and distribution of plug-in parsers
Intermediate Priority Items
The following are additional features that will be incorporated in future caArray releases. These features were voted by participants of the monthly caArray User meeting in June, 2009: • Provide a writable API to allow programmatic loading of data into caArray as part of existing pipelines (this is implemented in caArray Importer, a community contribution by Dr. Matt Eldridge (https://cabig-kc.nci.nih.gov/Molecular/forums/viewtopic.php?f=6&t=497). [vote: 11] • Enable easy upload of large data sets (>2GB). [vote: 5] • Controlled vocabulary (Suggest alternative terms already available, Allow curator to merge duplicate terms). [vote: 5] • Collaboration group ownership should not be tied to experiment ownership. [vote: 3] • Allow user to reset the status of files stuck in the Importing state. [vote: 3] • Allow adding a new organism through the UI. [vote: 1] • Support upload of files in the tar.gz format. [vote: 0] • Add ability to move Supplementary files back to the Manage Data tab. [vote: 0]

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

Version	**Development Focus**	**Release Date**
2.4.0	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: • Agilent: ♦GEML/xml array designs (aCGH, gene expression and miRNA) ♦Raw TXT data files (aCGH, gene expression and miRNA) • Nimblegen ? Community Code Contribution: ♦NDF array designs ♦Pair Report (raw and normalized) data files • Illumina: ♦BGX/TXT array designs	August 2010

	♦Gene expression: Sample Probe Profile TXT files with unique Probe_Id ♦Genotyping: Processed matrix TXT files with unique IlmnID values • Affymetrix: ♦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	
2.3.1	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: • 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	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: • 15578: Redesign Grid API to support backwards compatibililty via external data model • 18908: Keep an audit trail of permissions changes made to an experiment.	October 2009
2.2.1	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: • 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	**80 features/issues were addressed** • Highlights: ♦Sample search: ◊ 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: ◊ The user can generate and download a MAGE-TAB file set	January 2009

	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: ◊ 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: ◊ 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: ◊ 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.	
2.1.1	**Become part of the Life Science Distribution**	October 2008