



Agenda

- ▶ Overview of caMOD and its goals
- ▶ Salient Features
- ▶ Technical approach
- ▶ caBIG compatibility
- ▶ Q & A

Goals

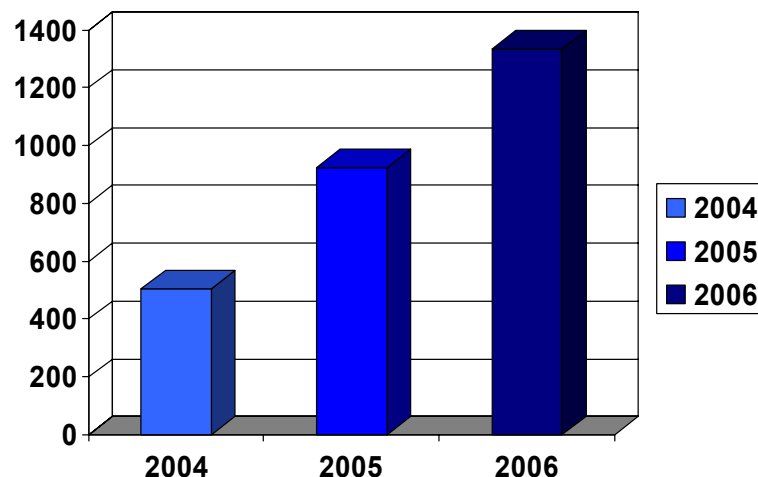
1. Cancer models that recapitulate many aspects of the genesis, progression, and clinical course of human cancers are valuable resources to cancer researchers engaged in a variety of basic, translational, clinical, and epidemiological investigations.

To provide an application and infrastructure supporting cancer models of diverse species to help researchers understand the genesis, progression, and clinical course of human cancer.

2. The NCI Mouse Models of Human Cancers Consortium (MMHCC) is a collaborative program designed to derive and characterize animal models, and to generate resources, information, and innovative approaches to the application of animal models in cancer research.

Provide a collaborative repository of internal and publicly available cancer models data and images supporting translational research.

3. Strive for a comprehensive knowledge source providing a wealth of information on cancer models, some pertinent aspects of which may never be available in a publication.



Avg # of Unique Visitors / Month

The cancer models database (caMOD) is a web-based resource that provides information about animal models for human cancer to the public research community

History

- ▶ January 2000
 - Prototype is presented during the Mouse Models of Human Cancers (MMHCC) Steering Committee Meeting
 - MMHCC adopts the Cancer Models Database (caMOD) as one of their initiatives
- ▶ July 2000
 - NCICB assumes responsibility for caMOD
- ▶ Spring 2001
 - caMOD 1.0 released (2-tier application)
- ▶ 2001 – 2005
 - Based on user feedback the application is constantly updated and improved.
- ▶ December 2005
 - caMOD 2.0 released (n-tier application, based on caBIG compliance guidelines)

Cancer Models Database v2.0

- **Submission**--Data in caMOD are extracted from the public scientific literature by curators and verified by the scientists who generated or worked with the models, or they are directly submitted by scientists.
- **Search**--Users can retrieve information about the making of models, their genetic descriptions, histopathology, derived cell lines, associated images, carcinogenic interventions, microarray data, and therapeutic trials in which the models were used. caMOD provides links to PubMed for associated publications and other resources such as mouse repositories, detailed information about altered gene, pathway affected, and information about human clinical trials that utilize the same compounds as the pre-clinical trials in the animal models.
- **System Function Administration**--The Admin function provides services for user registration, review of submitted models and database management.

The screenshot displays the caMOD website interface. At the top, the National Cancer Institute logo is visible. The main header features the caMOD logo and the text "Cancer Models Database". A navigation bar includes links for "QUICK LINKS", "HOME", "REGISTER", "SEARCH MODELS", "SUBMIT MODELS", and "HELP". A sidebar on the left lists "NCI HOME", "NCICB HOME", "EMICE WEBSITE", "CANCER IMAGES DATABASE", "NCICB SUPPORT", and "HELP". The main content area is titled "Welcome to the Cancer Models Database" and contains sections for "Search Models" (describing a query tool for model data) and "Submit Models" (describing a submission tool for human cancer models). A "LOGIN TO caMOD" section includes input fields for "Login ID" and "Password", along with "Login", "Register", and "Help" buttons. Below this, there are sections for "WHAT'S NEW" (with a "View What's New" link) and "DID YOU KNOW?" (with a "Release Notes" link and a note that caMOD 2.0 was released on Dec 21st).

Navigating the submission pages

Most parts contain multiple pages

SUBMIT & EDIT MODELS
Editing Model: **MMTV-ret (transgenic mice) (184)**
[View Model Status \(Edited-approved \)](#)

- MODEL CHARACTERISTICS
- GENETIC DESCRIPTION
 - Enter Engineered Transgene
 - ret
 - Enter Assoc Expression
 - Enter Genomic Segment
 - Enter Targeted Modification
 - Enter Induced Mutation
 - Enter Spontaneous Mutation
- CARCINOGENIC INTERVENTIONS
- PUBLICATIONS
- HISTOPATHOLOGY
- THERAPEUTIC APPROACHES
- CELL LINES
- IMAGES
- MICROARRAYS
- MODEL AVAILABILITY
- XENOGRAFT

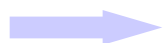
Multiple entries per category are possible

SUBMIT & EDIT MODELS
Editing Model: **MMTV-ret (transgenic mice) (184)**
[View Model Status \(Edited-approved \)](#)

- MODEL CHARACTERISTICS
- GENETIC DESCRIPTION
- CARCINOGENIC INTERVENTIONS
- PUBLICATIONS
- HISTOPATHOLOGY
 - Enter Histopathology
 - Salivary Gland (MMHCC)
 - Enter Assoc Metastasis
 - Enter Clinical Marker
 - Mammary Gland (MMHCC)
 - Enter Assoc Metastasis
 - Enter Clinical Marker
- THERAPEUTIC APPROACHES
- CELL LINES
- IMAGES
- MICROARRAYS
- MODEL AVAILABILITY
- XENOGRAFT

Search Result Detail Pages

MODEL DETAILS	HOME	REGISTER	SEARCH MODELS	SUBMIT MODELS	HELP																								
Viewing Model: C3(1)/SV40 Tag	Model Characteristics - Model: C3(1)/SV40 Tag																												
<ul style="list-style-type: none"> ➤ MODEL CHARACTERISTICS ➤ GENETIC DESCRIPTION ➤ CARCINOGENIC INTERVENTIONS ➤ PUBLICATIONS ➤ HISTOPATHOLOGY ➤ THERAPEUTIC APPROACHES ➤ CELL LINES ➤ IMAGES ➤ MICROARRAYS ➤ TRANSPLANT/XENOGRAFT ➤ BACK TO SEARCH RESULTS 	<table border="1"> <tr> <td>Model Descriptor</td> <td>C3(1)/SV40 Tag</td> </tr> <tr> <td>Official Nomenclature</td> <td> <ul style="list-style-type: none"> • FVB/N-TgN(C3-1-TAg)cJeg </td> </tr> <tr> <td>Genotype</td> <td> <ul style="list-style-type: none"> • C3(1)/SV40 Tag </td> </tr> <tr> <td>Species</td> <td>Mus musculus</td> </tr> <tr> <td>Strain</td> <td>FVB/N</td> </tr> <tr> <td>Experimental Design</td> <td>A 4.5 kb fragment of the 5' flanking region of the rat C3(1) gene (highly expressed in the rat ventral prostate) was used to drive the expression of the early region of the SV40 Tag.</td> </tr> <tr> <td>Phenotype</td> <td>SV40 Tag is expressed in several tissues including the mammary epithelium of females and the prostate epithelium of male mice. 100% of female mice develop multifocal low grade mammary intraepithelial neoplasia (MIN) at about 8 weeks which progresses to high grade MIN at about 12 weeks of age which resemble human DCIS. Invasive adenocarcinomas develop at about 16 weeks of age leading to the formation of 1-3 grossly palpable lesions. Females are generally euthanized by 6 months of age due to tumor burden. Approximately 15% of mice in the FVB/N background develop pulmonary metastases, whereas about 50% of mice in a mixed FVB/N x SV129 background develop lung mets. Male mice develop low grade intraepithelial prostate neoplasia (PIN) at about 3 months of age which progress to high grade PIN at about 5 months of age. Invasive carcinomas appear after about 7 months of age, primarily in the ventral prostate. Approximately 40% of males surviving to 9 months of age develop invasive carcinomas. Metastases are rare. Other proliferative lesions in association with Tag expression are seen in male and female mice including mixed tumor formation associated with heterotopic bone formation in the sweat glands of the feet, lesions in the nasovomerous glands, salivary glands, bulbourethral and urethral glands.</td> </tr> <tr> <td>Website for add. info</td> <td></td> </tr> <tr> <td>Breeding Notes</td> <td>Due to reduced life-span of females, breeding of heterozygotes is best achieved using males. Females should be bred prior to 4 months of age. Homozygous animals can be bred together and are available.</td> </tr> <tr> <td>Sex Distribution of the Phenotype</td> <td>Both Sexes</td> </tr> <tr> <td>Submitted by</td> <td>Green, Jeff</td> </tr> <tr> <td>Principal Investigator / Lab</td> <td>Green, Jeff</td> </tr> </table>	Model Descriptor	C3(1)/SV40 Tag	Official Nomenclature	<ul style="list-style-type: none"> • FVB/N-TgN(C3-1-TAg)cJeg 	Genotype	<ul style="list-style-type: none"> • C3(1)/SV40 Tag 	Species	Mus musculus	Strain	FVB/N	Experimental Design	A 4.5 kb fragment of the 5' flanking region of the rat C3(1) gene (highly expressed in the rat ventral prostate) was used to drive the expression of the early region of the SV40 Tag.	Phenotype	SV40 Tag is expressed in several tissues including the mammary epithelium of females and the prostate epithelium of male mice. 100% of female mice develop multifocal low grade mammary intraepithelial neoplasia (MIN) at about 8 weeks which progresses to high grade MIN at about 12 weeks of age which resemble human DCIS. Invasive adenocarcinomas develop at about 16 weeks of age leading to the formation of 1-3 grossly palpable lesions. Females are generally euthanized by 6 months of age due to tumor burden. Approximately 15% of mice in the FVB/N background develop pulmonary metastases, whereas about 50% of mice in a mixed FVB/N x SV129 background develop lung mets. Male mice develop low grade intraepithelial prostate neoplasia (PIN) at about 3 months of age which progress to high grade PIN at about 5 months of age. Invasive carcinomas appear after about 7 months of age, primarily in the ventral prostate. Approximately 40% of males surviving to 9 months of age develop invasive carcinomas. Metastases are rare. Other proliferative lesions in association with Tag expression are seen in male and female mice including mixed tumor formation associated with heterotopic bone formation in the sweat glands of the feet, lesions in the nasovomerous glands, salivary glands, bulbourethral and urethral glands.	Website for add. info		Breeding Notes	Due to reduced life-span of females, breeding of heterozygotes is best achieved using males. Females should be bred prior to 4 months of age. Homozygous animals can be bred together and are available.	Sex Distribution of the Phenotype	Both Sexes	Submitted by	Green, Jeff	Principal Investigator / Lab	Green, Jeff				
Model Descriptor	C3(1)/SV40 Tag																												
Official Nomenclature	<ul style="list-style-type: none"> • FVB/N-TgN(C3-1-TAg)cJeg 																												
Genotype	<ul style="list-style-type: none"> • C3(1)/SV40 Tag 																												
Species	Mus musculus																												
Strain	FVB/N																												
Experimental Design	A 4.5 kb fragment of the 5' flanking region of the rat C3(1) gene (highly expressed in the rat ventral prostate) was used to drive the expression of the early region of the SV40 Tag.																												
Phenotype	SV40 Tag is expressed in several tissues including the mammary epithelium of females and the prostate epithelium of male mice. 100% of female mice develop multifocal low grade mammary intraepithelial neoplasia (MIN) at about 8 weeks which progresses to high grade MIN at about 12 weeks of age which resemble human DCIS. Invasive adenocarcinomas develop at about 16 weeks of age leading to the formation of 1-3 grossly palpable lesions. Females are generally euthanized by 6 months of age due to tumor burden. Approximately 15% of mice in the FVB/N background develop pulmonary metastases, whereas about 50% of mice in a mixed FVB/N x SV129 background develop lung mets. Male mice develop low grade intraepithelial prostate neoplasia (PIN) at about 3 months of age which progress to high grade PIN at about 5 months of age. Invasive carcinomas appear after about 7 months of age, primarily in the ventral prostate. Approximately 40% of males surviving to 9 months of age develop invasive carcinomas. Metastases are rare. Other proliferative lesions in association with Tag expression are seen in male and female mice including mixed tumor formation associated with heterotopic bone formation in the sweat glands of the feet, lesions in the nasovomerous glands, salivary glands, bulbourethral and urethral glands.																												
Website for add. info																													
Breeding Notes	Due to reduced life-span of females, breeding of heterozygotes is best achieved using males. Females should be bred prior to 4 months of age. Homozygous animals can be bred together and are available.																												
Sex Distribution of the Phenotype	Both Sexes																												
Submitted by	Green, Jeff																												
Principal Investigator / Lab	Green, Jeff																												
QUICK LINKS																													
NCI HOME																													
NCICB HOME																													
EMICE WEBSITE																													
CANCER IMAGES DATABASE																													
NCICB SUPPORT																													
HELP																													



caGRID Demo

▶ <http://cagrid-browser.nci.nih.gov/>

```
<caBIGXMLQuery name="my-query">  
  <Target name="gov.nih.nci.camod.domain.AnimalModel">  
    <Objects name="gov.nih.nci.camod.domain.AnimalModel">  
      <Property name="id" value="3" />  
    </Objects>  
  </Target>  
</caBIGXMLQuery>
```

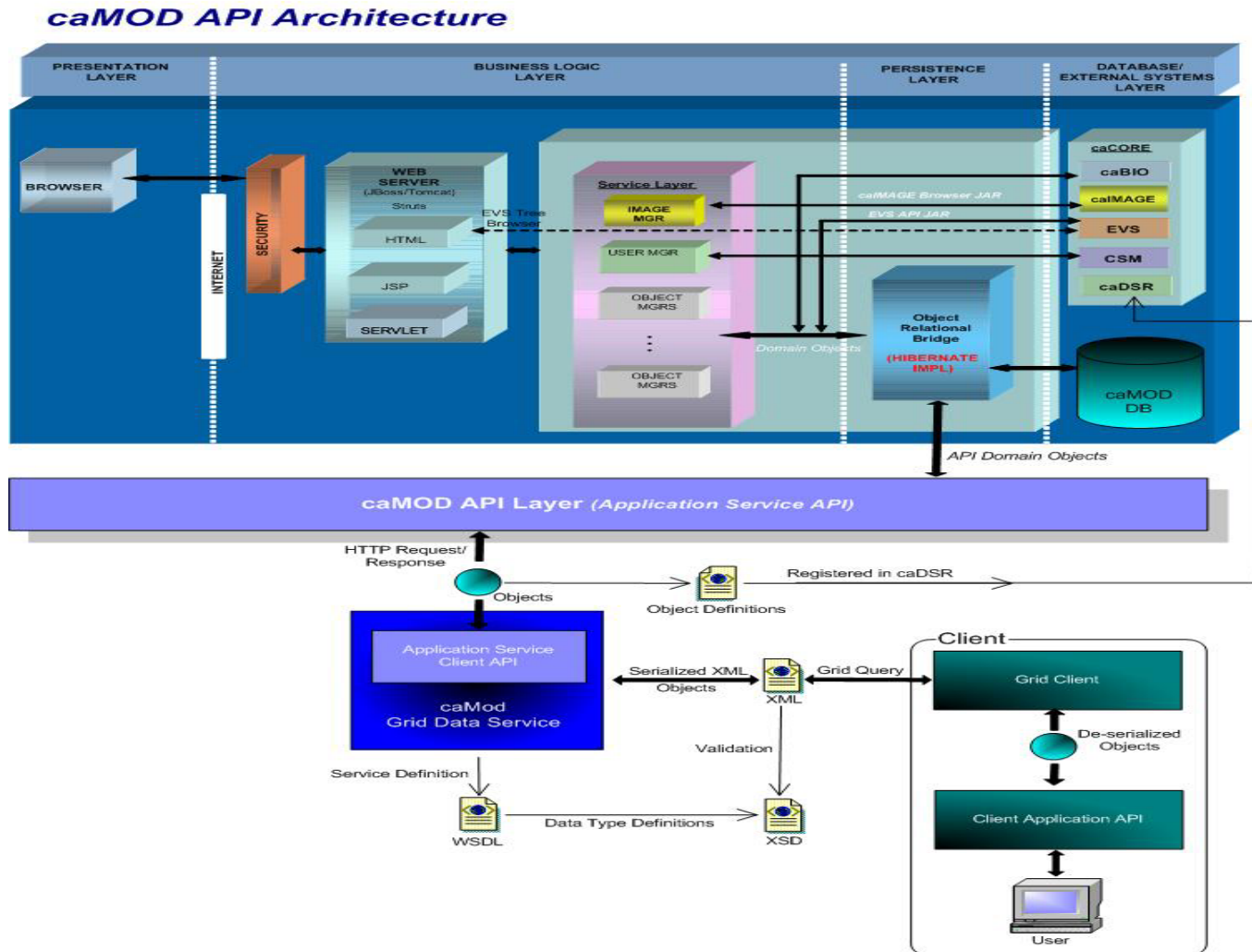

Interoperability

- ▶ caMOD has been designed, architected and constructed to facilitate interoperability with other systems, following caBIG guidelines.
- ▶ Information Providers to caMOD:
 - caBIO to retrieve gene info and clinical trials info through remote API
 - EVS to provide concept codes and preferred descriptions for concepts through caBIO EVS API
 - PubMed
 - Jackson Laboratory Resources
 - NCI's Developmental Therapeutics Program
 - caArray to store microarray data
 - caIMAGE server to store images
- ▶ Information Consumers: caMOD provides information to other systems
 - CMAP
 - BioGopher
 - Websites such as eMice references specific models in caMOD
 - caELMIR (future)

Evolution to caBIG Compliance

Maturity Model	Legacy	Bronze	Silver	Gold
Programming and Messaging Interfaces	<ul style="list-style-type: none"> - No programmatic interfaces to the system are available. Only local data files in a custom format can be read - Data transfer mechanisms implemented only on an ad hoc basis 	<ul style="list-style-type: none"> - Programmatic access to data from an external resource is possible 	<ul style="list-style-type: none"> - Well-described API's approved by the caBIG Architecture workspace provide access to data in the form of data objects that are instances of classes represented by a domain model - Electronic data formats reviewed and approved by the caBIG Architecture Workspace are supported for both input to and output from the system - Messaging protocols approved by the caBIG Architecture Workspace are supported wherever messaging is indicated by the use cases 	<ul style="list-style-type: none"> - All features of Silver, plus: - Service-oriented components produce or use resources in the form of grid services that use XML as the primary interchange format. - Interoperable with caGrid data grid architecture being developed by caBIG Architecture Workspace- Other features to be determined by caBIG Architecture workspace
Vocabularies / Terminologies & Ontologies	<ul style="list-style-type: none"> - Free text used throughout for data collection 	<ul style="list-style-type: none"> - Use of publicly accessible controlled vocabularies as well as local terminologies - Terminologies must include definitions of terms that meet caBIG VCDE workspace guidelines 	<ul style="list-style-type: none"> - Terminologies reviewed and validated by the caBIG VCDE Workspace used for all appropriate data collection fields and attributes of data objects - Term definitions must meet VCDE Workspace guidelines 	<ul style="list-style-type: none"> - All features of Silver, plus: - Full adoption of caBIG terminology standards as approved by the VCDE workspace. Terminologies must be available through a caGrid service
Data Elements	<ul style="list-style-type: none"> - No Structured metadata is recorded 	<ul style="list-style-type: none"> - Data element descriptions are maintained with sufficient definitional depth to enable a subject matter expert to unambiguously interpret the contents of the resource without contacting the original investigator - Data elements are built using controlled terminology - Metadata is stored and publicized in an electronic format that is separate from the resource that is being described 	<ul style="list-style-type: none"> - Common Data Elements (CDEs) built from controlled terminologies and according to practices validated by the VCDE workspace are used throughout - CDEs are registered as ISO/IEC 11179 metadata components in the caBIG Context of the cancer Data Standards Repository (caDSR) 	<ul style="list-style-type: none"> - All features of Silver, plus: - CDEs designated as caBIG Standards by the VCDE workspace are used - Metadata is advertised and discoverable via the caGrid services registry
Information Models	<ul style="list-style-type: none"> - No model describing the system is available in electronic format 	<ul style="list-style-type: none"> - Diagrammatic representation of the information model is available in electronic format 	<ul style="list-style-type: none"> - Object-oriented domain information models are expressed in UML as class diagrams and as XMI files, and are reviewed and validated by the VCDE Workspace 	<ul style="list-style-type: none"> - All features of Silver, plus: - Information models are harmonized across the caBIG Domain Workspaces

Architecture



caMOD APIs

- ▶ A read-only non-remote API exists for the retrieval of cancer models called “eQBE” from the common-persistence package. Extending the foundation of the concept of basic QBE in Hibernate, eQBE provides a mechanism for object searching based on example object graphs using the objects’ property values as criteria for a query. Used internally in caMOD.
- ▶ Using the caCore SDK Toolkit, the external caMOD API Layer is built upon the generated framework and domain objects (API Domain Objects) along with the caCore Application Service remote interface infrastructure.
- ▶ Integration of the caMOD application into the caGrid infrastructure includes the deployment of the caMOD Grid Data Service node and registration to the caBIG Index Service.

Example of internal caMOD Java API

For the AnimalModel class defined in caMOD with eQBE we can query properties of member objects as follows.

```
Person thePI = new Person();
thePI.setFirstName("Jane");
thePI.setLastName("Doe");
AnimalModel theQBEModel = new AnimalModel();
theQBEModel.setPrincipalInvestigator(thePI);
List theMatchingModels = Search.query(theQBEModel);
```

The result List of animal models will include all AnimalModel objects whose associated Person object (through the PrincipalInvestigator link) has a firstName property value set to "Jane" and a lastName property set to "Doe".

By default, Search.query will apply an equals evaluation to the property values.

Example of external caMOD remote Java API

For the AnimalModel class defined in caMOD we can query based on properties of the object as follows.

```
ApplicationService appService = ApplicationServiceProvider.getApplicationService();
try {
    AnimalModel theAnimalModel = new AnimalModelImpl();
    theAnimalModel.setModelDescriptor("ARF+/-Emu-myc");

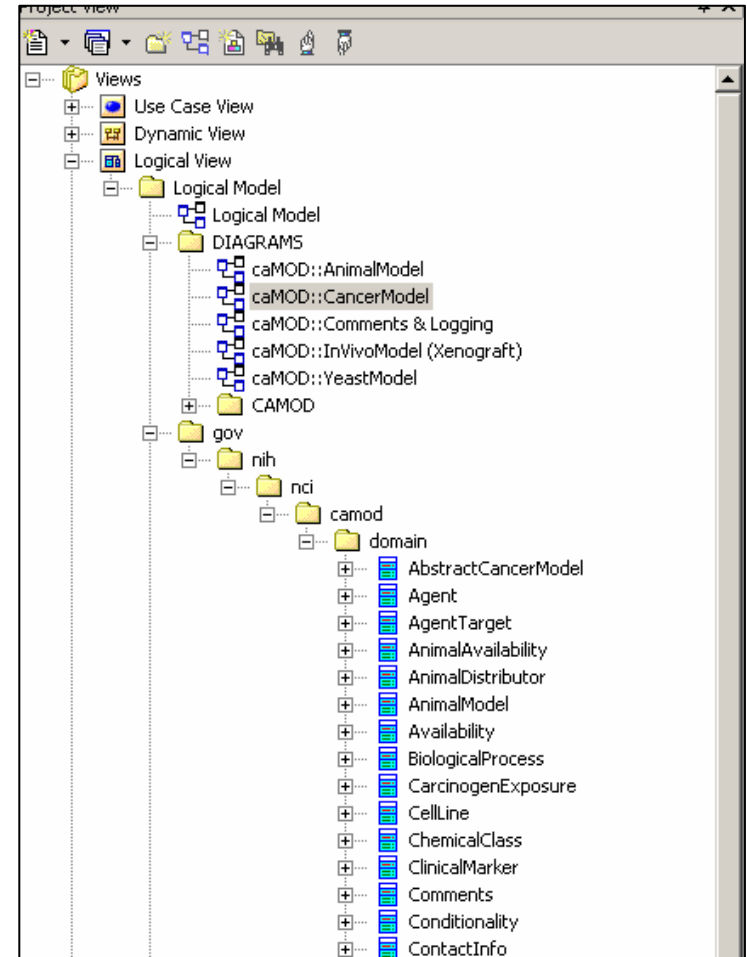
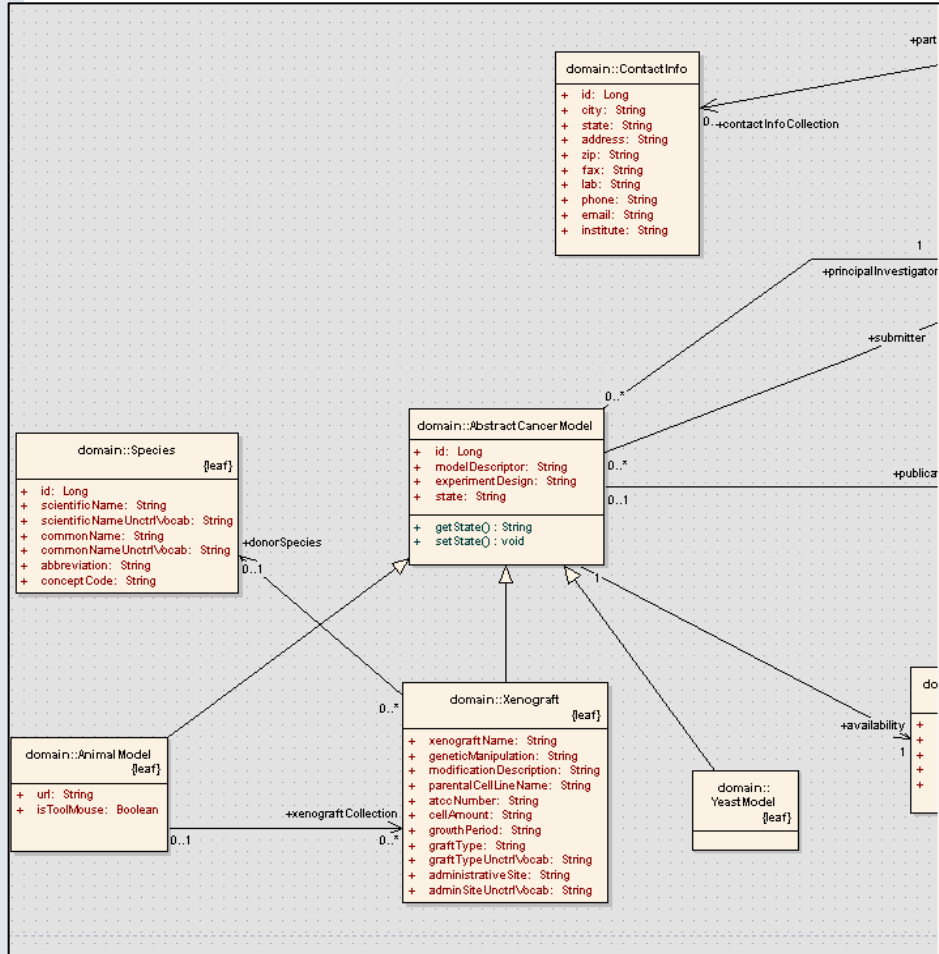
    List theAnimalModels = appService.search(AnimalModel.class, theAnimalModel);
    for (Iterator theIterator = theAnimalModels.iterator(); theAnimalModels.hasNext();)
    {
        AnimalModel theAnimalModel = (AnimalModel) theIterator.next();
        System.out.println("Id: " + theAnimalModel.getId() + "\n" + "\tType:" +
            theAnimalModel.getModelDescriptor() + "\n");
    }
}
catch (Exception e) {
    e.printStackTrace();
}
```

The result List of animal models will include all AnimalModel objects whose modelDescriptor property is equal to “ARF+/-Emi-mvc”.

caMOD Grid Node

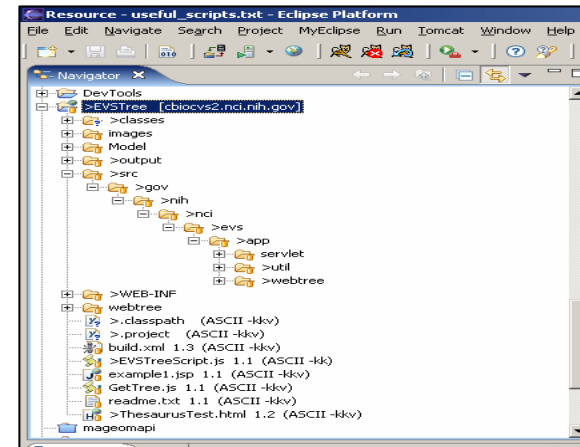
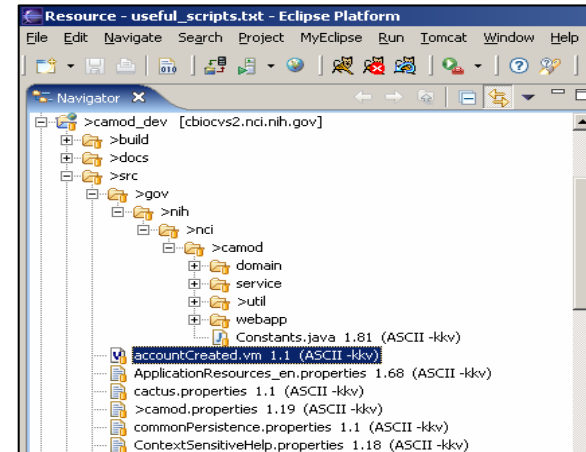
- ▶ The caMOD Grid Data Service node is generated using the caGrid Toolkit.
- ▶ The Data Type Definitions extracted from the caDSR registered caMOD model are used to build an XML schema or XSD that is used to define the input and output data types of the Grid services.
- ▶ The XSD is registered in the Global Model Exchange (GME) and used to describe the contract of the grid service and to validate the XML serialization of the API Domain Objects during data requests.
- ▶ Applications communicate with the caMOD Grid Data Service node following the caGrid services protocol.
- ▶ The caMOD API domain objects are serialized/deserialized by the caMOD Grid Data Service to and from XML adhering to the XML Schema (XSD) that is registered in the Global Model Exchange (GME)

UML Model



UML Model

- ▶ 76 Domain objects categorized under
 - caMOD::AnimalModel
 - caMOD::CancerModel
 - caMOD::InVivoModel (Xenograft)
 - caMOD::YeastModel
- ▶ Application specific classes under
 - caMOD::Comments & Logging
 - caMOD::Admin
 - caMOD::Preferences
- ▶ EVSTree utilizes caBIO 3.1 API



Vocabulary Usage

- ▶ NCI Thesaurus for murine organ and diagnosis terms
- ▶ NCI Thesaurus for human anatomical terms (release 2.1)
- ▶ NCI Thesaurus for staining methods (release 2.1)
- ▶ NCI Thesaurus for mouse and rat strains (release 2.1)
 - EVSTree – Reusable component for rendering EVS concepts

NATIONAL CANCER INSTITUTE **EVS** ENTERPRISE VOCABULARY SERVICES

TO LOCATE A TERM: 1 Enter a word or phrase in SEARCH 0/7 2 Navigate to the term in the DATA TREE

SEARCH: Entry
Enter Search Word or Phrase Below

Search: **SEARCH**

Match Entire Search Terms Only

SEARCH: Results
Results of Your Search Will Appear Below

Entering a search term above will perform a search for the term entered as well as alternate meanings.

For example, entering 'throat' will search for 'pharynx' as well as 'throat'.

DATA TREE
Browse DataTree Directly, Or Use Search for Prompting

- Murine Tissue Type
 - Body Fluids and Substances (MMHCC)
 - Cardiovascular System (MMHCC)
 - Digestive System (MMHCC)
 - Embryonic Tissue (MMHCC)
 - Endocrine Gland System (MMHCC)
 - Integument System (MMHCC)
 - Lymphohematopoietic System (MMHCC)
 - Musculoskeletal System (MMHCC)
 - Nervous System (MMHCC)
 - Reproductive System (MMHCC)

Common Data Elements

The screenshot shows the CDE Browser interface in Microsoft Internet Explorer. The browser address bar displays <http://cdebrowser.nci.nih.gov/CDEBrowser/>. The page header includes the National Cancer Institute logo and the text "U.S. National Institutes of Health | www.cancer.gov".

The main content area is divided into two sections:

- Left Panel (Tree View):** A "Refresh tree" section with a tree structure:
 - caDSR Contexts
 - caBIG (NCI cancer Biomedical Informatics Grid)
 - Classifications
 - C3PR
 - Cancer Models Database 2.0
 - Cancer Molecular Pages
 - CAP Cancer Checklists
 - caTIES
 - Clinical Trial Management Systems
 - Commercial Partners
 - Data Standards
 - Demonstration Applications
 - Genomic Identifiers
 - Grid-enablement of Protein Information Resource
 - Integrative Cancer Research
 - caArray 1.0
 - Dir Chall Lung Study

- Right Panel (Table):** A table of common data elements. The table has columns: Long Name, Document Text, Owned By, Used By Context, Registration Status, and Workflow. The table is sorted by Registration Status (Ascending).

<input type="checkbox"/>	Long Name	Document Text	Owned By	Used By Context	Registration Status	Workflow
<input type="checkbox"/>	Molecular Genetic Abnormality Identifier <i>java.lang.Long</i>		caCORE	caBIG		RELEASED
<input type="checkbox"/>	Administration Administration Age <i>java.lang.String</i>		caBIG			DRAFT
<input type="checkbox"/>	Administration Dose <i>java.lang.String</i>		caBIG			DRAFT
<input type="checkbox"/>	Administration Identifier <i>java.lang.Long</i>		caBIG			DRAFT
<input type="checkbox"/>	Administration Route of Administration <i>java.lang.String</i>		caBIG			DRAFT
<input type="checkbox"/>	Administration Treatment Regimen <i>java.lang.String</i>		caBIG			DRAFT
<input type="checkbox"/>	Administration Uncontrolled Text <i>Route of Administration</i> <i>java.lang.String</i>		caBIG			DRAFT
<input type="checkbox"/>	Animal Cancer Model Approval Status		caBIG			DRAFT

At the top right of the table area, there is a "Sort order" dropdown set to "(Default) Registration Status>>Workflow Status>>Long Name [Ascending]". Below it are three buttons: "Add to CDE Cart", "Add to CDE compare list", and "Compare CDEs".