STKE Connections Map Metadata Schema:

Translating the Connections Maps Database into SBML-Compatible XML

Shai Sachs,[1] Cal Collins,[1] Tina Underwood,[2] and Nancy R. Gough[3]*

(2004 October 7)

[1]Akaza Research, One Kendall Square, Building 400, 4th Floor, Cambridge, MA 02139, USA. [2]HighWire Press, 1454 Page Mill Road, Palo Alto, CA 94304, USA. [3]Science's STKE, AAAS, 1200 New York Avenue NW, Washington, DC 20005 USA.

*Corresponding author. E-mail, ngough@aaas.org

1.0 Introduction

This document describes a method for expressing information from the STKE Connections Map database in a Systems Biology Markup Language (SBML)-compatible XML format. SBML is an XML-based language that represents information about biochemical reaction, processes, networks, and pathways. The version of SBML for which this document is written is SBML Level 2, Version 1.

This document describes the XML schema for the Connections Map, therefore a basic working knowledge of XML is assumed and XML examples and notation will be used freely in this document. Also, a solid conceptual understanding of SBML (http://www.sbml.org/documents/.) and the STKE Connections Map database are required. The Connections Maps and SBML are described briefly.

This document and the associated schemas provide a set of rules for representing the Connections Map data in SBML-compliant XML. Recommendations on how to interpret, analyze, or manipulate the data are not supplied or intended. STKE makes no warranty of any kind with respect to the subject matter included herein. AAAS specifically disclaims all warranties, expressed, implied or otherwise, including without limitation, all warranties of merchantability, fitness for a particular purpose, or non-infringement.

1.1 The Connections Maps Cell Signaling Database

The Connections Maps at Science's STKE (http://stke.sciencemag.org/cm/) are presented through a dynamically generated graphical interface to the database of information on the components of cellular signaling pathways and their relations to one another. Information is provided by scientists with expertise in a given field and these "Pathway Authorities" are recruited by the editorial staff of the STKE. Each pathway is peer reviewed in association with publication of a Viewpoint article in a particular issue of *Science*.

Information in the database is divided into "canonical" and "specific", where canonical information represents knowledge abstracted from multiple organisms or cells and experimental paradigms. Specific information is limited to that information documented in a particular organism or cell type (such as *Saccharomyces cerevisae*), or related set of organisms (such as myocytes in mammals). Pathway

Authorities enter information about canonical and specific components into the database and then incorporate these components into pathways.

Each component has a minimum set of information that is independent of any pathway in which it appears. This is the "pathway-independent" component information displayed at the STKE site and is stored in the component record in the database. Once components are placed in a pathway, they acquire additional "pathway-dependent" component information, which is stored in the node record in the database. Only specific components can participate in specific pathways and only canonical components can participate in canonical pathways. Each specific component is derived from a canonical component "parent" and any components that share a canonical component parent are listed as homologs in the database (this can be functional or sequence homology).

Components are divided into three classes: simple, complex, and token. Simple components represent a molecule or group of molecules that always work together to perform a particular function. Complex components are comprised of more than one simple component. The actions of the components when in the complex are different from the actions when the simple components are not part of a complex or when they are part of a different complex. Token components represent a group of components and serve to make the pathways more manageable to visualize by grouping components of similar relevance within a pathway. Tokens are comprised of one or more simple or complex components.

Relationships among the components are described in the database as relation records. Components can be related to components within the same pathway (relation) or in different pathways (interpath relation).

The information in various fields in the database for pathways, components, nodes, and relations is a combination of data chosen from controlled vocabularies, data from other sources (such as PubMed or Flybase), and free text.

1.2 Introduction to SBML

SBML is a general-purpose formalism for representing models of biochemical pathways. The documentation describes SBML as language-neutral, but SBML is oriented toward expression in XML. In SBML, all information is stored inside within "elements". (In the SBML documentation, these are called "components", but the term "components" has a specific meaning in the STKE Connections Maps database and thus, we have chosen to refer to these SBML entities as elements.) Each element can contain plain text (as a specified data type, a controlled vocabulary, or unrestricted character string), a set of more elements, or no text. Furthermore, each element can have attributes. The SBML schema specifies a set of elements, and the concepts that the elements are meant to represent; it also specifies how these elements are to be structured.

The most relevant elements of the SBML format are described briefly and the order of presentation is based on relevancy of the elements to expressing the Connections Maps data. A valid SBML file has the following structure:

- An <SBML> element that contains:
    - A single <model> element, representing a system of biochemical pathways. Each element contains
        - A <listOfFunctionDefinitions>, representing a set of mathematical functions.
        - A <listOfUnitDefinitions>, representing a set of units for measuring quantities like volume and concentration.

- A <listOfCompartments>, representing a set of physical compartments in which species are located. This element contains:
    - One or more <compartment> elements, each representing a single physical compartment (a compartment is defined as a well-stirred container of chemical entities in which the species are, in mathematical terms, exactly evenly distributed across space).
- A <listOfSpecies> that contains
    - Zero or more <species> elements, each representing a single entity that participates in a pathway.
- A <listOfParameters>, representing a set of variables for use in mathematical functions.
    - A <listOfRules>, representing a set of constraints on parameters.
- A <listOfReactions>, representing a set of interactions among species. This element contains
    - Zero or more <reaction> elements, each representing a single interaction. This element contains
        - A <listOfReactants>, specifying the species that react. This element contains
            - Zero or more <speciesReference> elements, each representing a single reactant species.
        - A <listOfProducts>, specifying the species that are produced. This element contains:
            - Zero or more <speciesReference> elements, each representing a single product species.
        - A <listOfModifiers>, specifying the species that modify the reaction. This element contains:
            - Zero or more <modifierSpeciesReference> elements, each representing a single modifier species.
        - A <kineticLaw>, specifying the mathematical function that defines the rate at which a reaction takes place.
- A <listOfEvents>, describing a set of biological events.

Any SBML element may also contain the following elements, the contents of which are specified by domain-specific XML schemas as needed:

- <notes>, for human-generated free-text descriptions of the parent element
- <annotation>, for computer-generated, structured descriptions of the parent element.

2.0 Defining the Connections Maps Data in SBML Format

There is not a one-to-one correspondence between the information specified in the SBML Level 2, Version 1 language and the data in the Connections Maps database. For example, SBML offers no way to distinguish between complex, token, and simple components. The STKE Connections Maps XML schema is as an extension of the SBML format and will define which SBML elements are best suited for storing various information from the Connections Maps database and will specify the structure of the <annotation> element within various SBML elements relevant to the Connections Map.

2.1 Resource ID numbers

Several tables in the Connections Map database use a resource_id as a primary field. A resource id has the format "stkecm;XX_YY", where XX indicates the table for which the resource_id is the primary key, and YY is a number which makes the resource_id unique.

Almost all elements in SBML contain id attributes, which are meant to distinguish one element from another. SBML ids must start with a letter or an underscore, which may be followed by any number of letters, underscores or digits.

It is important to note that Connections Map resource_ids are not valid as id attributes for SBML elements, because resource_ids contain semicolons. When resource_ids are used as id attributes, the semicolon is replaced with an underscore, so that the id attribute would appear as "stkecm_XX_YY".

2.2 Auditing Information

Several tables in the Connections Map database contain fields used to track data modification and the completeness or approval status of the data. SBML does not explicitly have an explicit element to store this information, therefore, this information is stored in the <annotation> element of the element corresponding to the table. The following fields may be expressed in an <auditing> element of the <annotation> element, in the following way:

- authority_id maps to SBML attribute auditing.authorityID
- created_by maps to SBML attribute auditing.createdBy
- creation_date maps to SBML attribute auditing.creationDate
- last_modified_by maps to SBML attribute auditing.lastModifiedBy
- last_modified_date maps to SBML attribute auditing.lastModifiedDate

Note that the reference_count field common in many tables is not mapped to an <auditing> attribute, because this information can be inferred from the number of <citation> elements within same <listOfCitations> element within the same <annotation> element.

3.0 STKE Pathway Authorities and Database Users Schema

Data in the Connections Maps database may be entered or modified by Pathway Authorities, STKE editors, technical staff, and database curators. The users are defined by userIDs, which are values used for the following attributes:

- annotation.auditing.authorityID
- annotation.auditing.createdBy
- homolog.lastModifiedBy
- synonym.lastModifiedBy
- relationGroup.lastModifiedBy

The database users' information is exported to a single XML file conforming to the STKE Database Users schema. The root element of this schema is a SBML element <userLibrary>, which has no attributes and one nonrepeatable child element, a <listOfUsers> element.

The <listOfUsers> element also has no attributes and only one type of child element, <user>, which is repeatable. Information on the Pathway Authorities is stored in the cm_user table, with one row storing information about a single Authority. Therefore, each row of the cm_user table maps to a

single <userLibrary>:<listOfUsers>:<user> element. The link to the external userLibrary is made through a <userReference> element within the SBML file.

The fields in the cm_user table are mapped to attributes in the SBML element <user> as follows:

- cm_user_id maps to SBML attribute user.id
- cm_fname maps to SBML attribute user.firstName
- cm_lname maps to SBML attribute user.lastName
- cm_finitials to SBML attribute user.initials
- cm_email maps to SBML attribute user.email
- cm_au_affiliation maps to SBML attribute user.affiliation
- cm_department maps to SBML attribute user.department
- cm_address maps to SBML attribute user.address
- cm_city maps to SBML attribute user.city
- cm_state maps to SBML attribute user.state
- cm_country maps to SBML attribute user.country
- cm_zip maps to SBML attribute user.zip
- The standard auditing information is mapped to SBML element <user>:<auditing>.

Here is an example of a Connections Map User Library:

```
<?xml version="1.0" standalone="yes"?>
<sbml xmlns="http://www.sbml.org/sbml/level2" level="2" version="1">

<userLibrary>
  <listOfUsers>
      <user id="amgbell" firstName="Allan" lastName="Bell" affiliation="HighWire
Press"></user>
      <user id="mwhite" firstName="Morris" lastName="White" initials="F."
email="morris.white@joslin.harvard.edu" department="Howard Hughes Medical Institute, Harvard
Medical School" address="One Joslin Place" city="Boston" state="MA" country="USA"
zip="02215" affiliation="Joslin Diabetes Center"></user>
      <user id="ngough" firstName="Nancy" lastName="Gough" initials="R."
affiliation="Science's STKE"></user>
  </listOfUsers>
        </userLibrary>
```

4.0 The Component Library

Component information is independent of pathway information, and, thus, is expressed in a separate XML file, the "Component Library." This file contains information about all of the components in the STKE Connections Maps database.

The root element of the Component Library is a SBML <componentLibrary> element that has no attributes, a repeatable <controlledVocabularyReference> element, and one nonrepeatable child element, a <listOfComponents> element.

The <controlledVocabularyReference> has a type attribute and a location attribute, similar to the <controlledVocabularyReference> element in the <model> : <annotation> element. Every

componentLibrary must define one <controlledVocabularyReference> element with type set to "organism," and one <controlledVocabularyReference> element with type set to "species_type." These elements are necessary to interpret the controlled vocabulary terms referenced by component.organism and component.type.

The <listOfComponents> element also has no attributes and only one type of child element, <component>, which is repeatable. Information about components is stored in the cm_components table, with one row storing information about a single component. Therefore, each row of the cm_components table maps to a single <component> element in the <componentLibrary> : <listOfComponents> element.

The fields in the cm_components table map to attributes and child elements in the SBML <component> element, as follows:

- cm_components.resource_id maps to SBML attribute component.id.
- cm_components.name maps to SBML attribute component.name.
- cm_components.abbr maps to SBML attribute component.abbreviation
- cm_components.data_type maps to SBML attribute component.dataType, and must be "simple", "complex", or "token".
- cm_components.type maps to SBML attribute component.type, but should be omitted if data_type is "complex" or "token".
    - Note that type is a foreign key into the cm_controlled_vocabulary table.
- cm_components.organism maps to attribute component.organism
    - Note that organism is a foreign key into the cm_controlled_vocabulary table.
- cm_components.parent_id maps to SBML attribute component., if parent_id is not *NULL*.
- cm_components.description maps to SBML element <component> : <description>.
- The standard auditing information maps to SBML element <component> : <auditing>.
- The standard citation information maps to SBML element <component> : <listOfCitations>.

4.1 Component Authorities

All components have at least one authority. If a component X, with resource_id RX has two authorities, then there are two rows in the cm_authorities table with cm_authorities.resource_id = RX. Each of these rows is stored in a single SBML <authority> element within a <listOfAuthorities> element within the component that corresponds to X (and has component.id = RX).

The <authority> element has a userLibrary attribute, which is a URI describing the location of the user library that describes the users defined in the listOfAuthorities. The userLibrary attribute is required.

The fields in cm_authorities are mapped to attributes of the SBML <authority> element as follows:

- cm_authorities.authority_id maps to SBML attribute authority.id
- cm_authorities.corresponding maps to SBML attribute authority.corresponding.

The resource_id field is implicitly stored in the id attribute of the ancestor <component> element. The following are inferred: (i) component_family can be inferred from class_type, parent_id, and resource_id, (ii) class_type can be inferred from parent_id.

<listOfAuthorities>

```
        <authority id="tinau" corresponding ="1">
        <authority id="amgbell" corresponding ="0">
</listOfAuthorities>
```

4.2 Homologs

Specific components derived from the same canonical component are considered homologs. For example, two components, X and Y, are both homologs if they are both specific and the canonical component to which X belongs is the same as the canonical component to which Y belongs (i.e., X.parent_id = Y.parent_id).

If a component, X, with resource_id r_x has ten homologs, then there are ten rows in cm_component_homolog with cm_component_homolog.resource_id = r_x. Each of these rows is stored in a single SBML <homolog> element within a <listOfHomologs> element within the <component> element that corresponds to X (and has component.id=r_x). There is only one field in cm_component_homolog that is mapped to a single attributes of the <homolog> element as follows:

- cm_component_homolog.homolog_id maps to SBML attribute homolog.id

The cm_component_homolog.resource_id field is not mapped, because it is implicitly stored in the id attribute of the ancestor <component> element.

4.3 Synonyms

Alternate names for components are stored as synonyms. If a component, X, with resource_id r_x has ten synonyms, then there are ten rows in cm_component_synonym with cm_component_synonym.resource_id = r_x. Each of these rows is stored in a single SBML <synonym> element within a <listOfSynonyms> element within the <component> element that corresponds to X (and has component.id=r_x). The fields in cm_component_synonym are mapped to attributes of the <synonym> element as follows:

- cm_component_synonym.synonym maps to SBML attribute synonym.value
- cm_component_synonym.last_modified_date maps to SBML attribute synonym.lastModifiedDate.
- cm_component_synonym.last_modeified_by maps to SBML attribute synonym.lastModifiedBy.

The resource_id field is not mapped, because it is implicitly stored in the id attribute of the ancestor <species> element.

4.4 Complex and Token Components

Complex and token components are comprised of several simple constituent components. If a component, X, with resource_id r_x has ten constituents, then there are ten rows in cm_component_complex with cm_component_complex.resource_id = r_x. Each of these rows is stored in a single SBML <constituent> element within the <listOfConstituents> element of the <component> element that corresponds to X (and has component.id=r_x). The fields in cm_component_complex are mapped to attributes of the <constituent> element as follows:

- cm_component_complex.constituent_id maps to SBML attribute constituent.id

- cm_component_complex.position maps to SBML attribute constituent.position
- The standard auditing information is mapped to SBML <constituent> : <auditing>.

5.0 STKE Connections Maps Controlled Vocabulary Schema

The Connections Map database uses controlled vocabularies to express many different values, such as a node's subcellular localization or a reaction's type. Although all controlled vocabulary terms are stored in the same table, there are several distinct controlled vocabularies, grouped together by the cm_controlled_vocabulary.table_attribute field. For example, all subcellular localization vocabulary terms have the same value in the table_attribute field.

Each controlled vocabulary will be exported in its entirety to a single XML file conforming to an STKE Connections Map Controlled Vocabulary schema. The root element of this schema is a SBML <controlledVocabulary> element. The link to the external controlled vocabulary is made through a <controlledVocabularyReference> element within the SBML file.

The schema for the external Controlled Vocabulary is defined as follows:

- The root element is a single SBML <controlledVocabulary> element. The <controlledVocabulary> element has one attribute and one element:
    - SBML attribute controlledVocabulary.type is set to the cm_controlled_vocabulary.table_attribute field that defines the controlled vocabulary.
    - A single SBML <listOfTerms> element with no attributes.
        - For each term in the controlled vocabulary, a single <term> element is included in the SBML <listOfTerms> element. The values from that term's row in the cm_controlled_vocabulary table are mapped to the following attributes:
            - cm_controlled_vocabulary.resource_id maps to the SBML attribute term.id
            - cm_controlled_vocabulary.cv_term maps to the SBML attribute term.value
            - cm_controlled_vocabulary.parent_id maps to the SBML attribute term.parentID, if parent_id is not *NULL*.
            - The standard auditing information is included in an SBML <auditing> element within the <term> element.

The controlledVocabulary.type, term.id, and term.value attributes are required. The parentID is required, if parent_id field is not *NULL*. The <listOfTerms> element is required.

For example, the following fragment of XML displays some of the terms for the component type vocabulary:

```
<controlledVocabulary type="component_type">
        <listOfTerms>
                <term id="rc" value="diglyceride" parentID="rp" />
                <term id="rp" value= "lipid" />
        </listOfTerms>
</controlledVocabulary>
```

The files containing controlled vocabularies are linked to the main SBML file through <controlledVocabularyReference> elements in the <model> : <annotation> : <listOfControlledVocabularies> element. The <controlledVocabularyReference> element has the following structure:

- The SBML type attribute is equal to the controlledVocabularyReference.type attribute in the file containing the controlled vocabulary.
- The SBML location attribute specifies a URI for the XML file containing the controlled vocabulary.

The file that exists at the URI specified by a controlledVocabularyReference.location attribute must have a controlledVocabulary.type attribute equal to the controlledVocabularyReference.type attribute of the corresponding <controlledVocabularyReference> attribute in the main Connections Maps SBML file (Table 1). Those attributes and attributes whose values are resource_ids for controlled vocabulary terms will be noted explicitly.

Each Connections Map SBML file is required to define <controlledVocabularyReference> elements with the types shown Table 1. All pathway files, which are mapped to <model> elements, must include all five of the <controlledVocabularyReference> elements.

The Component Library also must define <controlledVocabularyReference> elements for two of the types shown in Table 1. The <controlledVocabularyReference> elements must contain two attributes, controlledVocabularyReference.type and controlledVocabularyReference.location. Location is the URI for the controlled vocabulary and type represents either organism or species_type.

Table 1. Controlled vocabularies and their usage in the Connections Maps database.

Controlled Vocabulary type	**Required for <model> element**	**Required for Component Libary**	**Attributes in the STKE Connections Maps Schema that refer to this Controlled Vocabulary**
organism	X	X	Pathway.organism and component.organism
tissue_cell	X		Pathway.tissue
species_type	X	X	Component.type
subcellular_localization	X		Species.localization
relation_method	X		Relation.method

For example, part of the SBML <model> : <annotation> element might appear as follows:

```
<listOfControlledVocabularies>
      <controlledVocabularyReference type="organism"
location="http://stke.sciencemag.org/cm/vocabularies/organism.xml"/>
      <controlledVocabularyReference type="tissue_cell"
location="http://stke.sciencemag.org/cm/vocabularies/tissue_cell.xml"/>
      <controlledVocabularyReference type="relation_method"
location="http://stke.sciencemag.org/cm/vocabularies/relation_method.xml"/>
```

```
      <controlledVocabularyReference type="species_type"
location="http://stke.sciencemag.org/cm/vocabularies/species_type.xml"/>
      <controlledVocabularyReference type="subcellular_localization"
location="http://stke.sciencemag.org/cm/vocabularies/subcellular_localization.xml"/>
</listOfControlledVocabularies>
```

An example of the <controlledVocabularyReference> elements for the Component Library might appear as follows:

```
<listOfControlledVocabularies>
   <controlledVocabularyReference type="species_type"
location="http://stke.sciencemag.org/cm/vocabularies/component_type.xml"/>
   <controlledVocabularyReference type="organism"
location="http://stke.sciencemag.org/cm/vocabularies/organism.xml"/>
         </listOfControlledVocabularies>
```

6.0 Citations

All entities in the Connections Map database may be annotated with citations to published literature. The relationship between an entity and a citation is established in the resource_related table, with the resource_id field holding the resource_id of the entity, and the related_resource_id field holding the resource_id of the citation.

For each citation listed in the resource_related table, a SBML <citation> element must be included in a <listOfCitations> element within the element corresponding to that resource related to the citation. Each citation listed also has a row in separate citation table: cm_book_citation, cm_entrez_citation, cm_database_citation, cm_freeform_citation or cm_journal_citation.

In addition to the fields unique to the citation table, each citation will have an attribute called "type" that contains one of the following values: book, protein, nucleotide, database, freeform, or journal.

The fields of the cm_book_citation are mapped to attributes of the SBML <citation> element as follows:

- cm_book_citation.book_title maps to SBML attribute title
- cm_book_citation.authors maps to SBML attribute authors
- cm_book_citation.publisher maps to SBML attribute publisher
- cm_book_citation.editor maps to SBML attribute editor
- cm_book_citation.edition maps to SBML attribute edition
- cm_book_citation.publication_location maps to SBML attribute publicationLocation
- cm_book_citation.pages maps to SBML attribute pages
- cm_book_citation.pub_date maps to SBML attribute pubDate

The fields of the cm_freeform_citation are mapped to attributes of the SBML <citation> element as follows:

- cm_freeform_citation.doc_type maps to SBML attribute docType
- cm_freeform_citation.title maps to SBML attribute title
- cm_freeform_citation.journal maps to SBML attribute journal

- cm_freeform_citation.author maps to SBML attribute author
- cm_freeform_citation.year maps to SBML attribute year
- cm_freeform_citation.volume maps to SBML attribute volume
- cm_freeform_citation.pages maps to SBML attribute pages

The fields of the cm_entrez_citation are mapped to attributes of the SBML <citation> element as follows:

- (type = "nucleotide" or "protein")
- cm_entrez_citation.definition maps to SBML attribute definition
- cm_entrez_citation.,source maps to SBML attribute source
- cm_entrez_citation.abstract_url maps to SBML attribute abstractUrl
- cm_entrez_citation.resource_id (without prefix) maps to SBML attribute accession

The fields of the cm_database_citation are mapped to attributes of the SBML <citation> element as follows:

- cm_database_citation.dbname maps to SBML attribute dbname
- cm_database_citation.accession maps to SBML attribute accession

The fields of the cm_journal_citation are mapped to attributes of the SBML <citation> element as follows:

- cm_journal_citation.resource_id (minus the prefix) maps to SBML attribute pmid
- cm_journal_citation.doc_type maps to SBML attribute docType
- cm_journal_citation.title maps to SBML attribute title
- cm_journal_citation.authors maps to SBML attribute authors
- cm_journal_citation.journal maps to SBML attribute journal
- cm_journal_citation.volume maps to SBML attribute volume
- cm_journal_citation.issue maps to SBML attribute issue
- cm_journal_citation.pages maps to SBML attribute pages
- cm_journal_citation.pretty_date maps to SBML attribute date

An example with each type of citation follows:

<listOfCitations>
<citation type="freeform" docType="Research Report" title="The Caenorhabditis elegans Antennapedia homologue in migration cells controls their direction of migration" authors="Salser, S. J., Kenyon, C." year="1992" journal="Nature" volume="355" pages="255-258"/>
<citation type="database" dbname="FlyBase" accession="FBgn0036045"/>
<citation type="book" title="Cell Culture in the Neurosciences" authors="Guroff, G." publisher="Plenum" editor="Bottenstein & Sato" edition=" " publicationLocation=" " pages="245-272" pubDate="1985-01-01 00:00:00.0"/>
<citation type="nucleotide" abstractUrl="http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=Nucleotide&list_uids=10880136&dopt=GenBank" accession="NP_004032"/>
<citation type="protein" definition="G-protein alpha-t2 subunit" source="house mouse" abstractUrl="http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=Protein&list_uids=193432&dopt=GenPept" accession="AAC37650"/>

<citation type="journal" pmid="11560894" authors="Natarajan, L., Witwer, N.E., Eisenmann, D.M." docType="Research Report" title="The divergent Caenorhabditis elegans beta-catenin proteins BAR-1, WRM-1 and HMP-2 make distinct protein interactions but retain functional redundancy in vivo." journal="Genetics" volume="159" pages="159-72" issue="1" date="2001 Sep"/>
</listOfCitations>

7.0 Pathways

Information about pathways is stored in the cm_pathway table, with one row holding information about one pathway. Because a pathway is conceptually a collection of interactions among several biochemical entities, it is represented by the SBML <model> element.

Each row in cm_pathway exported to a single SBML file and corresponds to a <model> element in that file. There is a separate SBML file for each pathway. The fields of cm_pathway are mapped to attributes of <model>, or attributes and elements of a <model> : <annotation> : <pathway> element, as follows:

- cm_pathway.resource_id maps to SBML attribute model.id
- cm_pathway.path_name maps to SBML attribute model.name
- cm_pathway.parent_id maps to SBML attribute pathway.parentID, if parent_id is not *NULL*.
- cm_pathway.organism maps to SBML attribute pathway.organism.
    - Organism is a foreign key into the cm_controlled_vocabulary table.
- cm_pathway.tissue_cell maps to attribute pathway.tissue.
    - Tissue_cell is a foreign key into the cm_controlled_vocabulary table.
- cm_pathway.path_description maps to SBML element <pathway> : <description>
- The standard auditing information maps to SBML element <pathway> : <auditing>.
- The standard citations information maps to SBML element <pathway> : <listOfCitations>

The following can be inferred: cm_pathway.class_type can be inferred from cm_pathway.parent_id and cm_pathway.family_name can be inferred from cm_pathway.class_type and cm_pathway.parent_id.

If either cm_pathway.organism or cm_pathway.tissue_cell is *NULL*, then the corresponding attribute in <pathway> may be omitted.

7.1 Defining Pathway Authorities

All pathways have at least one authority. If a pathway X, with resource_id RX has two authorities, then there are two rows in the cm_authorities table with cm_authorities.resource_id = RX. Each of these rows is stored in a single <authority> element within a <listOfAuthorities> element within the pathway.

The model element has a userLibrary attribute, which is a URI describing the location of the User Library that describes the users defined in the listOfAuthorities. The userLibrary attribute is required.

The fields in cm_authorities are mapped to attributes of the <authority> element as follows:

- cm_authorities.authority_id maps to SBML attribute authority.id
- cm_authorities.corresponding maps to SBML attribute authority.corresponding.

The cm_authorities.resource_id field is not mapped, because it is implicitly stored in the id attribute of the ancestor <pathway> element.

Example of corresponding Authorities:

<listOfAuthorities>
<authority id="tinau" corresponding ="1">
<authority id="amgbell" corresponding ="0">
</listOfAuthorities>

7.2 Components Associated with Pathways

The <model> : <annotation> element must contain an element named componentLibraryReference, which two attributes, one named location and one named species.type. The value of componentLibraryReference.location is the URI for the Component Library. The Component Library describes the components from which the nodes in the pathway are derived.

8.0 Compartments

SBML requires that all models that define species also specify the compartments in which those species are located. The Connections Maps includes localization information for nodes, however, the compartments specified by these localization attributes do not meet the criteria of a well-stirred container of chemical entities in which the species are, in mathematical terms, exactly evenly distributed across space. However, Connections Map nodes are expressed as SBML species, which requires a component element. Therefore, a single "dummy" compartment, in which all species are said to exist, is used. This compartment has id "stkecm_CT_0" and has name "Connections Map Compartment." Consequently, every STKE Connections Map model has the following <listOfCompartments> element:

<listOfCompartments>
<compartment id="stkecm_CT_0" name="Connections Map Compartment" />
</listOfCompartments>

Furthermore, each <species> element must have species.compartment set to "stkecm_CT_0".

9.0 Nodes

A node is a component that has been incorporated in a pathway. For each node in Connections Map database, there is a row in the cm_node table.

A node is mapped to an SBML <species> element. Node information is stored in the <node> element of the <annotation> element of the <species> element that corresponds to the component to which that node belongs. The mapping of cm_node fields to <node> attributes is as follows:

- cm_node.resource_id maps to SBML attribute species.id
- cm_node.path_id maps to SBML attribute node.pathID, if path_id is different from the cm_pathway.resource_id of the pathway being exported.
    - Note that if a node with a pathID attribute is involved in a relation, then that relation must be an interpath relation.
- cm_node.component_id maps to the SBML attribute node.componentID, and refers to a component in the Component Library.
- cm_node.y_pos maps to SBML attribute node.yPos

- cm_node.x_pos maps to SBML attribute node.xPos
- cm_node.type maps to SBML attribute node.type
    - Note that type is a foreign key into the cm_controlled_vocabulary table.
- cm_node.localization_id maps to SBML attribute node.localization
    - Note that localization_id is a foreign key into the cm_controlled_vocabulary table.
- cm_node.abbr maps to SGML attribute node.name
- cm_node.description maps to SBML element <node> : <description>
- The standard auditing information maps to SBML element <node> : <auditing>.
- The standard citations information maps to SBML element <node> : <listOfCitations>.

The following can be inferred, cm_node.organism can be inferred from the value of the pathway.organism attribute within the ancestor <model> element.

10.0 Relations

A relation expresses an interaction between two nodes: The effect of one node on another is positive, negative, neutral, or unknown. Each relation is stored as a single row in the cm_relation table. A relation is mapped to an SBML <reaction> element. For each node in a pathway that is being exported, every relation in which that node is involved (either as a source or as a target) must also be exported in a single <reaction> element. The mapping from fields in cm_relation to attributes in the SBML <reaction> element, or to attributes and elements in the <reaction> : <annotation> : <relation> element are as follows:

- cm_relation.resource_id maps to SBML attribute reaction.id
- cm_relation.source_node_id maps to SBML attribute modifierSpeciesReference.species in the element <reaction> : <listOfModifiers> : <modifierSpeciesReference>
    - There must be a <species> element in the <listOfSpecies> with species.id = source_node_id.
- cm_relation.target_node_id maps to SBML attribute speciesReference.species in the element <reaction> : <listOfProducts> : <speciesReference>.
    - There must be a <species> element in the <listOfSpecies> with species.id = target_node_id.
- cm_relation.sign maps to SBML attribute relation.sign, and must be "+", "-", "0", or "?".
- cm_relation.method maps to SBML attribute relation.method
    - Note that method is a foreign key into the cm_controlled_vocabulary table.
- cm_relation.reliability maps to SBML attribute relation.reliability, and must be "Demonstrated", "Strongly Implied", "Implied", or "Speculative".
- cm_relation.group_id maps to SBML attribute relation.groupID, if it is not *NULL*.
- cm_relation.description maps to SBML element <relation> : <description>.
- The standard auditing information maps to SBML element <relation> : <auditing>.
- The standard citations information maps to SBML element <relation> : <listOfCitations>.
- The SBML attribute reaction.reversible is always set to false.

The following attributes can be inferred, cm_relation.class_type can be inferred from the id of the <compartment> element containing the species involved in the reaction.

10.1 Relation Groups

Some relations are grouped to disambiguate their interaction with other relations. Each such relation has a non-NULL value in the group_id field of its row in the cm_relation table, and this value is stored in the <reaction> : <annotation> : <group> element for the relevant relation. Suppose that a relation with group_id r_g is exported to the SBML file. The row in the cm_map_group table with

resource_id r_g is mapped to a <listOfReactions> : <reaction> : <annotation> : <relationGroup> element, as follows:

- cm_map_group.resource_id maps to SBML attribute relationGroup.id
- cm_map_group.group_type maps to SBML attribute relationGroup.type, and must be "alternative reactions", "competing reactions", "coupled reactions", "negative feedback", "negative feedforward", "positive feedback", "positive feedforward", or "simultaneous reactions"
- cm_map_group.lastModifiedBy maps to SBML attribute relationGroup.lastModifiedBy
- cm_map_group.lastModifiedDate maps to SBML attribute relationGroup.lastModifiedDate

The relationGroup.id and relationGroup.type attributes are required.

For example, suppose that the pathway being expressed has two groups of relations, with resource_id and group_type r_a and "coupled reactions", and r_b and "negative feedback", respectively. An example of the partial XML fragment for the <listOfReactions> element includes the following code (not all relation attributes are shown):

```
<listOfReactions>
        <reaction>
                <annotation>
                        <relationGroup id="ra" type="coupled reactions"
                        lastModifiedBy="Bowerman" lastModifiedDate="2000-01-10
                        13:05:46.566"/>
                </annotation>
        </reaction>
        <reaction>
        <annotation>
                        <relationGroup id=" rb" type="negative feedback"
                        lastModifiedBy="Bowerman" lastModifiedDate="2000-01-10
                        13:06:48.522"/>
                </annotation>
    </reaction>
```

11.0 Formal STKE Connections Map XML Schema

Four XML schemas that formally define the STKE Connections Map Metadata Schema accompany this document. The four schemas correspond to the following elements an SBML file:

1. componentLibrary.xsd defines the <componentLibrary> element in the Component Library.
2. controlledVocabulary.xsd defines the <controlledVocabulary> element in a Controlled Vocabulary file.
3. userLibrary.xsd defines the <userLibrary> element in the Pathway Authorities Library
4. model.xsd defines the <model> element in pathway file

APPENDIX A. SAMPLE EXPORTED SBML FILE

The following is a part of the SBML file that might be generated from an export of the "C. elegans Endoderm Induction Wnt Pathway". Its purpose is to clarify the concepts in the STKE Connections Map Metadata Schema. Only a few nodes and relations have been exported.


```
<?xml version="1.0" standalone="yes"?>
<sbml xmlns="http://www.sbml.org/sbml/level2" xmlns:xs="http://stke.sciencemag.org/schemas/model.xsd" level="2" version="1">
  <model id="stkecm_CMP_9993" name="Fas Signaling Pathway in Cardiomyocytes">
    <annotation>
      <listOfControlledVocabularies>
        <controlledVocabularyReference type="organism" location="http://stke.sciencemag.org/cm/vocabularies/organism.xml"/>
        <controlledVocabularyReference type="tissue_cell" location="http://stke.sciencemag.org/cm/vocabularies/tissue_cell.xml"/>
        <controlledVocabularyReference type="relation_method" location="http://stke.sciencemag.org/cm/vocabularies/relation_method.xml"/>
        <controlledVocabularyReference type="species_type" location="http://stke.sciencemag.org/cm/vocabularies/species_type.xml"/>
        <controlledVocabularyReference type="subcellular_localization"
location="http://stke.sciencemag.org/cm/vocabularies/subcellular_localization.xml"/></listOfControlledVocabularies>
      <componentLibraryReference location="http://stke.sciencemag.org/component_library.xml"/>
      <userLibraryReference location="http://stke.sciencemag.org/user_library.xml"/>
      <pathway parentID="stkecm_CMP_7966" organism="stkecm_CV_2663" tissue="stkecm_CV_1577">
        <description>&lt;FONT FACE="verdana, arial, helvetica, sans serif" size="2"&gt;Fas is a member of the death domain-containing subgroup of
the tumor necrosis factor (TNF) receptor superfamily that induces apoptosis or necrosis in a variety of cells. Fas is investigated mainly with respect to
its death-inducing capabilities, but it can also act as a co-stimulator in T cell activation. Defects in the Fas pathway are associated with autoimmunity
and lymphoproliferation in humans [autoimmune lymphoproliferative syndrome (ALPS)] and mice [lymphoproliferative disease (lpr) and gld mice]
(Nagata, 1999; Siegel &lt;i&gt;et al&lt;/i&gt;., 2000; Krammer, 2000). Fas has also a role in cardiac hypertrophy. Cardiac hypertrophy is a reaction to
biomechanical stress of pressure overload, for example, that caused by arterial hypertension. This adaptive growth is primarily compensatory, but
sustained hypertrophy leads to heart failure (Hunter &lt;i&gt;et al&lt;/i&gt;., 1999). FasL increases cell size and protein synthesis of neonatal rat
cardiomyocytes. In particular, FasL upregulates mRNA for atrial natriuretic factor (ANF), a hallmark of the hyperthrophic response of cardiomyocytes
(Badorff &lt;i&gt;et al&lt;/i&gt;., 2002). In vivo, lpr mice showed a rapid onset of left ventricular dilatation and failure, absence of compensatory
hypertrophy, and significantly increased mortality after pressure overload induction (Badorff &lt;i&gt;et al&lt;/i&gt;., 2002). Inactivation of glycogen
synthase kinase 3-beta (GSK3&lt;FONT FACE="SYMBOL"&gt;b&lt;/FONT&gt;) by phosphorylation is necessary and sufficient for cardiomyocyte
hypertrophy in vitro (Cross &lt;i&gt;et al&lt;/i&gt;., 1995; Haq &lt;i&gt;et al&lt;/i&gt;., 2000). FasL transiently increased Akt activity and
```

GSK3<FONT FACE="SYMBOL">b</FONT> phosphorylation in cardiomyocytes of wild-type mice, but not in cardiomyocytes of lpr mice (Badorff <i>et al</i>., 2002).</description>
<listOfAuthorities>
<authority id="haraldwajant" corresponding="1"/></listOfAuthorities>
<auditing authorityID="haraldwajant" createdBy="haraldwajant" creationDate="2002-03-14 03:34:22.096" lastModifiedBy="haraldwajant" lastModifiedDate="2003-05-07 12:31:44.116"/>
<listOfCitations>
<citation type="journal" pmid="8524413" authors="Cross, D.A., Alessi, D.R., Cohen, P., Andjelkovich, M., Hemmings, B.A." docType="Journal Article" title="Inhibition of glycogen synthase kinase-3 by insulin mediated by protein kinase B." journal="Nature" volume="378" pages="785-9" issue="6559" date="1995 Dec 21-28"/>
<citation type="journal" pmid="10690403" authors="Nagata, S." docType="Review, Academic" title="Fas ligand-induced apoptosis." journal="Annu Rev Genet" volume="33" pages="29-55" date="1999"/>
<citation type="journal" pmid="10528039" authors="Hunter, J.J., Chien, K.R." docType="Review, Tutorial" title="Signaling pathways for cardiac hypertrophy and failure." journal="N Engl J Med" volume="341" pages="1276-83" issue="17" date="1999 Oct 21"/>
<citation type="journal" pmid="11018058" authors="Haq, S., Choukroun, G., Kang, Z.B., Ranu, H., Matsui, T., Rosenzweig, A., Molkentin, J.D., Alessandrini, A., Woodgett, J., Hajjar, R., Michael, A., Force, T." docType="Journal Article" title="Glycogen synthase kinase-3beta is a negative regulator of cardiomyocyte hypertrophy." journal="J Cell Biol" volume="151" pages="117-30" issue="1" date="2000 Oct 2"/>
<citation type="journal" pmid="11048730" authors="Krammer, P.H." docType="Review, Tutorial" title="CD95's deadly mission in the immune system." journal="Nature" volume="407" pages="789-95" issue="6805" date="2000 Oct 12"/>
<citation type="journal" pmid="11101867" authors="Siegel, R.M., Chan, F.K., Chun, H.J., Lenardo, M.J." docType="Review, Academic" title="The multifaceted role of Fas signaling in immune cell homeostasis and autoimmunity." journal="Nat Immunol" volume="1" pages="469-74" issue="6" date="2000 Dec"/>
<citation type="journal" pmid="11827997" authors="Badorff, C., Ruetten, H., Mueller, S., Stahmer, M., Gehring, D., Jung, F., Ihling, C., Zeiher, A.M., Dimmeler, S." docType="Journal Article" title="Fas receptor signaling inhibits glycogen synthase kinase 3 beta and induces cardiac hypertrophy following pressure overload." journal="J Clin Invest" volume="109" pages="373-81" issue="3" date="2002 Feb"/></listOfCitations>
</pathway>
</annotation>
<listOfCompartments>
<compartment id="stkecm_CT_0" name="Connections Map Compartment"/></listOfCompartments>
<listOfSpecies>
<species id="stkecm_CMN_10000" compartment="stkecm_CT_0">
<annotation>
<node pathID="stkecm_CMP_9993" componentID="stkecm_CMC_9995" xPos="392.0" yPos="258.0" type="stkecm_CV_1400" localization="stkecm_CV_1003" abbreviation="AKT/PKB">

<auditing authorityID="haraldwajant" createdBy="haraldwajant" creationDate="2002-03-14 04:43:04.676" lastModifiedBy="haraldwajant" lastModifiedDate="2002-05-30 17:36:49.79"/></node>
</annotation>
</species>
<species id="stkecm_CMN_10002" compartment="stkecm_CT_0">
<annotation>
<node pathID="stkecm_CMP_9993" componentID="stkecm_CMC_9996" xPos="398.0" yPos="378.0" type="stkecm_CV_1311" localization="stkecm_CV_1003" abbreviation="GSK-3">
<auditing authorityID="haraldwajant" createdBy="haraldwajant" creationDate="2002-03-14 04:43:53.973" lastModifiedBy="haraldwajant" lastModifiedDate="2002-05-30 17:43:01.61"/></node>
</annotation>
</species>
<species id="stkecm_CMN_9999" compartment="stkecm_CT_0">
<annotation>
<node pathID="stkecm_CMP_9993" componentID="stkecm_CMC_9994" xPos="382.0" yPos="132.0" type="stkecm_CV_1310" localization="stkecm_CV_7504" abbreviation="Fas">
<auditing authorityID="haraldwajant" createdBy="haraldwajant" creationDate="2002-03-14 04:42:06.086" lastModifiedBy="haraldwajant" lastModifiedDate="2002-05-30 17:33:34.506"/></node>
</annotation>
</species>
<species id="stkecm_CMN_10001" compartment="stkecm_CT_0">
<annotation>
<node pathID="stkecm_CMP_9993" componentID="stkecm_CMC_9998" xPos="394.0" yPos="512.0" type="stkecm_CV_1303" localization="stkecm_CV_1003" abbreviation="ANF">
<description><FONT FACE="verdana, arial, helvetica, sans serif" size="2">ANF is a hallmark of the cardiac hypertrophy response. Inhibition of glycogen synthase kinase 3 beta (GSK3<FONT FACE="SYMBOL">b</FONT>) results in the translocation of its downstream effector nuclear factor of activated T cells (NF-AT) into the nucleus and initiation of hypertrophic gene expression (Haq <i>et al</i>., 2000).</description>
<auditing authorityID="haraldwajant" createdBy="haraldwajant" creationDate="2002-03-14 04:43:15.236" lastModifiedBy="haraldwajant" lastModifiedDate="2002-05-30 17:39:57.17"/>
<listOfCitations>
<citation type="journal" pmid="11018058" authors="Haq, S., Choukroun, G., Kang, Z.B., Ranu, H., Matsui, T., Rosenzweig, A., Molkentin, J.D., Alessandrini, A., Woodgett, J., Hajjar, R., Michael, A., Force, T." docType="Journal Article" title="Glycogen synthase kinase-3beta is a negative regulator of cardiomyocyte hypertrophy." journal="J Cell Biol" volume="151" pages="117-30" issue="1" date="2000 Oct 2"/></listOfCitations>

```
        </node>
       </annotation>
      </species>
    </listOfSpecies>
    <listOfReactions>
     <reaction id="stkecm_CMR_10003" reversible="false">
      <annotation>
       <relation sign="+" method="stkecm_CV_7696" reliability="Demonstrated">
        <description>FasL transiently increased Akt activity and GSK3&amp;#223; phosphorylation in cardiomyocytes of wt mice, but not in cardiomyocytes of lpr mice (Badorff et al., 2002). Akt-mediated phosphorylation of GSK3&amp;#223; inhibits the activity of the latter (Cross et al., 1995).</description>
        <auditing authorityID="haraldwajant" createdBy="haraldwajant" creationDate="2002-03-14 04:44:35.67" lastModifiedBy="haraldwajant" lastModifiedDate="2002-03-14 04:47:40.233"/>
        <listOfCitations>
         <citation type="journal" pmid="8524413" authors="Cross, D.A., Alessi, D.R., Cohen, P., Andjelkovich, M., Hemmings, B.A." docType="Journal Article" title="Inhibition of glycogen synthase kinase-3 by insulin mediated by protein kinase B." journal="Nature" volume="378" pages="785-9" issue="6559" date="1995 Dec 21-28"/>
         <citation type="journal" pmid="11827997" authors="Badorff, C., Ruetten, H., Mueller, S., Stahmer, M., Gehring, D., Jung, F., Ihling, C., Zeiher, A.M., Dimmeler, S." docType="Journal Article" title="Fas receptor signaling inhibits glycogen synthase kinase 3 beta and induces cardiac hypertrophy following pressure overload." journal="J Clin Invest" volume="109" pages="373-81" issue="3" date="2002 Feb"/></listOfCitations>
       </relation>
      </annotation>
      <listOfProducts>
       <speciesReference species="stkecm_CMN_10000"/></listOfProducts>
      <listOfModifiers>
       <modifierSpeciesReference species="stkecm_CMN_9999"/></listOfModifiers>
     </reaction>
     <reaction id="stkecm_CMR_10004" reversible="false">
      <annotation>
       <relation sign="-" method="stkecm_CV_1266" reliability="Demonstrated">
        <description>FasL transiently increased Akt activity and GSK3&amp;#223; phosphorylation in cardiomyocytes of wt mice, but not in cardiomyocytes of lpr mice (Badorff et al., 2002). Akt-mediated phosphorylation of GSK3&amp;#223; inhibits the activity of the latter (Cross et al., 1995).</description>
```

<auditing authorityID="haraldwajant" createdBy="haraldwajant" creationDate="2002-03-14 04:44:39.34" lastModifiedBy="haraldwajant" lastModifiedDate="2002-03-14 04:49:35.826"/>
<listOfCitations>
<citation type="journal" pmid="8524413" authors="Cross, D.A., Alessi, D.R., Cohen, P., Andjelkovich, M., Hemmings, B.A." docType="Journal Article" title="Inhibition of glycogen synthase kinase-3 by insulin mediated by protein kinase B." journal="Nature" volume="378" pages="785-9" issue="6559" date="1995 Dec 21-28"/>
<citation type="journal" pmid="11827997" authors="Badorff, C., Ruetten, H., Mueller, S., Stahmer, M., Gehring, D., Jung, F., Ihling, C., Zeiher, A.M., Dimmeler, S." docType="Journal Article" title="Fas receptor signaling inhibits glycogen synthase kinase 3 beta and induces cardiac hypertrophy following pressure overload." journal="J Clin Invest" volume="109" pages="373-81" issue="3" date="2002 Feb"/></listOfCitations>
</relation>
</annotation>
<listOfProducts>
<speciesReference species="stkecm_CMN_10002"/></listOfProducts>
<listOfModifiers>
<modifierSpeciesReference species="stkecm_CMN_10000"/></listOfModifiers>
</reaction>
<reaction id="stkecm_CMR_10005" reversible="false">
<annotation>
<relation sign="-" method="stkecm_CV_7696" reliability="Strongly Implied">
<description>Inhibition of GSK3beta results in the translocation of its downstream effector NF-AT into the nucleus and initiation of the hypertrophic gene expression (Haq et al., 2000)</description>
<auditing authorityID="haraldwajant" createdBy="haraldwajant" creationDate="2002-03-14 04:44:43.086" lastModifiedBy="haraldwajant" lastModifiedDate="2002-03-14 04:54:26.526"/>
<listOfCitations>
<citation type="journal" pmid="11018058" authors="Haq, S., Choukroun, G., Kang, Z.B., Ranu, H., Matsui, T., Rosenzweig, A., Molkentin, J.D., Alessandrini, A., Woodgett, J., Hajjar, R., Michael, A., Force, T." docType="Journal Article" title="Glycogen synthase kinase-3beta is a negative regulator of cardiomyocyte hypertrophy." journal="J Cell Biol" volume="151" pages="117-30" issue="1" date="2000 Oct 2"/></listOfCitations>
</relation>
</annotation>
<listOfProducts>
<speciesReference species="stkecm_CMN_10001"/></listOfProducts>
<listOfModifiers>
<modifierSpeciesReference species="stkecm_CMN_10002"/></listOfModifiers>
</reaction>

```
    </listOfReactions>
  </model>
</sbml>
```

APPENDIX B – SAMPLE COMPONENT LIBRARY

The following is sample portion of the componentLibrary.

```
<?xml version="1.0" encodeing="UTF-8" standalone="yes"?>
<sbml xmlns="http://www.sbml.org/sbml/level3" level="2" version="1">
<componentLibrary>
  <listOfControlledVocabularies>
    <controlledVocabularyReference type="species_type" location="http://stke.sciencemag.org/cm/vocabularies
/component_type.xml"/>
    <controlledVocabularyReference type="organism" location="http://stke.sciencemag.org/cm/vocabularies/org
anism.xml"/></listOfControlledVocabularies>
  <listOfComponents>
    <component id="stkecm_CMC_10031" name="Histidine protein kinase" abbreviation="CKI2" dataType="simple"
type="stkecm_CV_1303" organism="stkecm_CV_2525" parentID="stkecm_CMC_9726">
      <description>CK12 is a histidine protein kinase that can activate cytokinin signaling.</description>
      <auditing authorityID="jsheen" createdBy="jsheen" creationDate="2002-03-19 00:05:15.513" lastModified
By="jsheen" lastModifiedDate="2004-08-24 16:07:52.226"/>
      <listOfAuthorities>
        <authority id="jsheen" corresponding="1"/></listOfAuthorities>
      <listOfHomologs>
        <homolog id="stkecm_CMC_10027"/>
        <homolog id="stkecm_CMC_10028"/>
        <homolog id="stkecm_CMC_14245"/>
        <homolog id="stkecm_CMC_14252"/></listOfHomologs>
      <listOfSynonyms>
        <synonym value="AHK5" lastModifiedBy="ngough" lastModifiedDate="2004-08-24 16:07:52.533"/></listOfS
ynonyms>
    </component>
    <component id="stkecm_CMC_9727" name="Histidine phosphotransfer protein" abbreviation="HPT" dataType="s
imple" type="stkecm_CV_1303">
      <description>Histidine phosphotransfer proteins (HPTs) mediate histidine-aspartate phosphorelay.</des
```

cription>
<auditing authorityID="jsheen" createdBy="jsheen" creationDate="2002-02-27 11:42:24.403" lastModified
By="jsheen" lastModifiedDate="2003-10-04 11:06:05.486"/>
<listOfAuthorities>
<authority id="jsheen" corresponding="1"/></listOfAuthorities>
<listOfCitations>
<citation type="journal" pmid="10664616" authors="Urao, T., Yamaguchi-Shinozaki, K., Shinozaki, K."
docType="Review, Tutorial" title="Two-component systems in plant signal transduction." journal="Trends Pla
nt Sci" volume="5" pages="67-74" issue="2" date="2000 Feb"/>
<citation type="journal" pmid="11158442" authors="Suzuki, T., Sakurai, K., Ueguchi, C., Mizuno, T."
docType="Journal Article" title="Two types of putative nuclear factors that physically interact with histi
dine-containing phosphotransfer (Hpt) domains, signaling mediators in His-to-Asp phosphorelay, in Arabidops
is thaliana." journal="Plant Cell Physiol" volume="42" pages="37-45" issue="1" date="2001 Jan"/>
</listOfCitations>
<listOfSynonyms>
<synonym value="AHP1" lastModifiedBy="magpie1" lastModifiedDate="2003-10-04 11:06:05.883"/>
<synonym value="AHP2" lastModifiedBy="magpie1" lastModifiedDate="2003-10-04 11:06:05.883"/>
<synonym value="AHP3" lastModifiedBy="magpie1" lastModifiedDate="2003-10-04 11:06:05.883"/>
<synonym value="AHP4" lastModifiedBy="magpie1" lastModifiedDate="2003-10-04 11:06:05.883"/>
<synonym value="AHP5" lastModifiedBy="magpie1" lastModifiedDate="2003-10-04 11:06:05.883"/></listOf
Synonyms>
</component>
<component id="stkecm_CMC_8218" name="B cell antigen receptor" abbreviation="BCR" dataType="complex">
<description><FONT FACE="verdana, arial, helvetica, sans serif" size="2">The B cell receptor (B
CR) is composed of a membrane-bound immunoglobulin molecule (mIg) noncovalently associated with disulfide-l
inked heterodimers of Ig-alpha and Ig-beta (Ig-<FONT FACE="SYMBOL">a</FONT>/<FONT FACE="SYMB
OL">b</FONT>). </description>
<auditing authorityID="jcambier" createdBy="jcambier" creationDate="2002-01-30 14:23:26.113" lastModi
fiedBy="jcambier" lastModifiedDate="2003-12-19 18:15:40.11"/>
<listOfAuthorities>
<authority id="jcambier" corresponding="1"/></listOfAuthorities>
<listOfSynonyms>
<synonym value="Big Jim & the twins" lastModifiedBy="magpie1" lastModifiedDate="2003-12-19 18:1
5:41.71"/></listOfSynonyms>

```
    <listOfConstituents>
      <constituent id="stkecm_CMC_6912" position="1">
        <auditing authorityID="jcambier" createdBy="jcambier" creationDate="2002-01-30 14:27:14.27" lastM
odifiedBy="jcambier" lastModifiedDate="2002-01-30 14:27:14.27"/></constituent>
      <constituent id="stkecm_CMC_8219" position="2">
        <auditing authorityID="jcambier" createdBy="jcambier" creationDate="2002-01-30 14:27:14.27" lastM
odifiedBy="jcambier" lastModifiedDate="2002-01-30 14:27:14.27"/></constituent>
    </listOfConstituents>
  </component>
  <component id="stkecm_CMC_8377" name="STAT target genes" abbreviation="Genes" dataType="token">
    <description>Target genes that are activated by dimeric STAT complexes typically contain a "GAS" elem
ent similar to the consensus TTC(N2-4)GAA.</description>
    <auditing authorityID="chorvath" createdBy="chorvath" creationDate="2002-02-02 09:22:14.653" lastModi
fiedBy="chorvath" lastModifiedDate="2003-08-26 09:50:41.91" completeness="incomplete"/>
    <listOfAuthorities>
      <authority id="chorvath" corresponding="1"/></listOfAuthorities>
    <listOfConstituents>
      <constituent id="stkecm_CMC_6711" position="1">
        <auditing authorityID="chorvath" createdBy="chorvath" creationDate="2002-03-03 11:36:19.483" last
ModifiedBy="chorvath" lastModifiedDate="2002-03-03 11:36:19.483"/></constituent>
    </listOfConstituents>
  </component>
 </listOfComponents>
</componentLibrary>
```

APPENDIX C. STKE CONNECTIONS MAP METADATA

The following table includes the complete list of the SBML elements and attributes with sample values, whether the element or attribute is required (R is for required and O is for optional), whether the element is repeatable (Y is for yes, N is for no), the source field in the database, and additional notes. The root <SBML> element is excluded.

An element or attribute is required if it must be present within its immediate ancestor element for the document to conform to both the SBML specifications and the STKE Connections Map Metadata Schema. However, if an element or attribute is optional and a value is available from the source database field, that value should be provided in the SBML file. For example, the <listOfCitations> element is optional; therefore, documents lacking this element are compliant with the specification. However, the <citation> element is required; therefore, documents having a <listOfCitations> element, but lacking a <citation> element within the <listOfCitations> element, are invalid.

An element is repeatable if several elements–each pertaining to different data, but conforming to the same part of the specification–may be present. The element is only repeatable with respect to its immediate ancestor element. For example, <user> may be repeated within <listOfUsers>, but <listOfUsers> may not be repeated within <model>. Attributes are never repeatable in XML documents.

XML Outline number	**SBML Element or Attribute Elements are enclosed in <>.**	**R or O**	**Repeatable**	**Database Field**	**Notes**
1.0	<model>				
1.1	model.id	R	N	cm_pathway.resource_id	
1.2	model.name	R	N	cm_pathway.path_name	
1.3	<annotation>	R	N		
1.3.1	<listOfControlledVocabularies>	R	N		Must include all controlled vocabularies, type="organism", "tissue_cell", relation_method", "species_type", "subcellular_localization".
1.3.1.1	<controlledVocabularyReference>	R	Y		Used to refer to an external controlled vocabulary file.
1.3.1.1.1	controlledVocabularyReference.type	R	N	cm_controlled_vocabulary.table_attribute	

1.3.1.1.2	controlledVocabularyReference.location	R	N		Refers to the URI of the XML file containing the controlled vocabulary.
1.3.2	<componentLibraryReference>	R	N		
1.3.2.1	componentLibraryReference.location	R	N		Refers to the URI of the Component Library, which contains information about the components from which the nodes in pathways are derived.
1.3.3	<userLibraryReference>	R	N		
1.3.3.1	userLibraryReference.location	R	N		Refers to the URI of the User Library, which contains information about the components from which the nodes in pathways are derived.
1.3.4	<pathway>	R	N		
1.3.4.1	pathway.parentID	O	N	cm_pathway.parent_id	Should be omitted if parent_id is *NULL*.
1.3.4.2	pathway.organism	O	N	cm_pathway.organism	This is a foreign key into the cm_controlled_vocabulary table (see 1.3.1).
1.3.4.3	pathway.tissue	O	N	cm_pathway.tissue_cell	This is a foreign key into the cm_controlled_vocabulary table (see 1.3.1).
1.3.4.4	<description>	R	N	cm_pathway.path_description	Should be omitted if path_description is *NULL*.
1.3.4.5	<listOfAuthorities>	R	N		
1.3.4.5.1	<authority>	R	Y		Must be at least one authority listed. Corresponds to a row from the cm_authorities table with cm_authorities.resource_id equal to attribute pathway.id. If only one authority, then

					corresponding="1", if more than one authority, then only one will be designated with corresponding="1" and all others will be set to corresponding="0".
1.3.4.5.1.1	authority.id	R	N	cm_authorities.authority_id	
1.3.4.5.1.2	authority.corresponding	R	N	cm_authorities.corrresponding	
1.3.4.6	<auditing>	O	N		
1.3.4.6.1	auditing.authority.id	R	N	cm_pathway.authority_id	
1.3.4.6.2	auditing.createdBy	R	N	cm_pathway.created_by	
1.3.4.6.3	auditing.creationDate	R	N	cm_pathway.creation_date	
1.3.4.6.4	auditing.lastModifiedBy	R	N	cm_pathway.last_modified_by	
1.3.4.6.5	auditing.lastModifiedDate	R	N	cm_pathway.last_modified_date	
1.3.4.7	<listOfCitations>	O	N		
1.3.4.7.1	<citation>	O	Y		The citations come from different tables depending on the type: cm_journal_citation, cm_entrez_citation, cm_book_citation, cm_database_citation, or cm_freeform_citation.
1.3.4.7.1.1	type	R	N		
1.3.4.7.1.2	title	O	N		
1.3.4.7.1.3	authors	O	N		
1.3.4.7.1.4	publisher	O	N		
1.3.4.7.1.5	editor	O	N		
1.3.4.7.1.6	edition	O	N		
1.3.4.7.1.7	publicationLocation	O	N		
1.3.4.7.1.8	pages	O	N		
1.3.4.7.1.9	pubDate	O	N		
1.3.4.7.1.10	docType	O	N		

1.3.4.7.1.11	journal	O	N		
1.3.4.7.1.12	year	O	N		
1.3.4.7.1.13	volume	O	N		
1.3.4.7.1.14	issue	O	N		
1.3.4.7.1.15	source	O	N		
1.3.4.7.1.16	abstractUrl	O	N		
1.3.4.7.1.17	accession	O	N		
1.3.4.7.1.18	dbname	O	N		
1.3.4.7.1.19	definition	O	N		
1.3.4.7.1.20	date	O	N		
1.3.4.7.1.21	pmid	O	N		
1.4	<listOfCompartments>	R	N		
1.4.1	<compartment>	R	N		
1.4.1.1	compartment.id	R	N		Must always be "stkecm_CT_0."
1.4.1.2	compartment.name	R	N		Must always be "Connections Map Compartment."
1.5	<listOfSpecies>	R	N		
1.5.1	<species>	R	Y		
1.5.1.1	species.id	R	N	cm_node.resource_id	
1.5.1.2	species.compartment	R	N		Must always be "stkecm_CT_0."
1.5.1.3	<annotation>	R	N		
1.5.1.3.1	<node>	R	N		Corresponds to a row from the cm_node table with cm_node.resource_id equal to attribute species.id (1.5.1.1)
1.5.1.3.1.1	node.pathID	R	N	cm_node.path_id	
1.5.1.3.1.2	node.componentID	R	N	cm_node.component_id	
1.5.1.3.1.3	node.yPos	R	N	cm_node.y_pos	
1.5.1.3.1.4	node.xPos	R	N	cm_node.x_pos	
1.5.1.3.1.5	node.type	R	N	cm_node.type	This is a foreign key into the cm_controlled_vocabulary table

					"species_type" (see 1.3.1).
1.5.1.3.1.6	node.localization	R	N	cm_node.localization_id	This is a foreign key into the cm_controlled_vocabulary table "subcellular_localization" (see 1.3.1).
1.5.1.3.1.7	node.abbreviation	R	N	cm_node.abbr	
1.5.1.3.1.8	<description>	O	N	cm_node.description	
1.5.1.3.1.9	<auditing>	O	N		
1.5.1.3.1.9.1	auditing.authority.id	R	N	cm_node.authority_id	
1.5.1.3.1.9.2	auditing.createdBy	R	N	cm_node.created_by	
1.5.1.3.1.9.3	auditing.creationDate	R	N	cm_node.creation_date	
1.5.1.3.1.9.4	auditing.lastModifiedBy	R	N	cm_node.last_modified_by	
1.5.1.3.1.9.5	auditing.lastModifiedDate	R	N	cm_node.last_modified_date	
1.5.1.3.1.10	<listOfCitations>	O	N		
1.5.1.3.1.10.1	<citation>	R	Y		The citations come from different tables depending on the type: cm_journal_citation, cm_entrez_citation, cm_book_citation, cm_database_citation, or cm_freeform_citation.
1.5.1.3.1.10.1.1	type	R	N		
1.5.1.3.1.10.1.2	title	O	N		
1.5.1.3.1.10.1.3	authors	O	N		
1.5.1.3.1.10.1.4	publisher	O	N		
1.5.1.3.1.10.1.5	editor	O	N		
1.5.1.3.1.10.1.6	edition	O	N		
1.5.1.3.1.10.1.7	publicationLocation	O	N		
15.1.3.1.10.1.8	pages	O	N		
1.5.1.3.1.10.1.9	pubDate	O	N		
1.5.1.3.1.10.1.10	docType	O	N		
1.5.1.3.1.10.1.1	journal	O	N		

1					
1.5.1.3.1.10.1.12	year	O	N		
1.5.1.3.1.10.1.13	volume	O	N		
1.5.1.3.1.10.1.14	issue	O	N		
1.5.1.3.1.10.1.15	source	O	N		
1.5.1.3.1.10.1.16	abstractUrl	O	N		
1.5.1.3.1.10.1.17	accession	O	N		
1.5.1.3.1.10.1.18	dbname	O	N		
1.5.1.3.1.10.1.19	definition	O	N		
1.5.1.3.1.10.1.20	date	O	N		
1.5.1.3.1.10.1.21	pmid	O	N		
1.6	<listOfReactions>	O	N		
1.6.1	<reaction>	R	Y		
1.6.1.1	Reaction.id	R	N	cm_relation.resource_id	
1.6.1.2	Reaction.reversible	R	N		Always set to “false”.
1.6.1.3	<annotation>	R	N		
1.6.1.3.1	<relation>	R	N		
1.6.1.3.1.1	relation.sign	R	N	cm_relation.sign	Must be “+”, “-”, “0”, or “?”.
1.6.1.3.1.2	relation.method	R	N	cm_relation.method	This is a foreign key into the cm_controlled_vocabulary table "relation_method" (see 1.3.1).

1.6.1.3.1.3	relation.reliability	R	N	cm_relation.reliability	Must be “Demonstrated”, “Strongly Implied”, “Implied”, or “Speculative”.
1.6.1.3.1.4	Relation.groupID	O	N	cm_relation.group_id	Should be omitted if group_id is NULL.
1.6.1.3.1.5	<description>	O	N	cm_relation.description	
1.6.1.3.1.8	<auditing>	O	N		
1.6.1.3.1.8.1	auditing.authority.id	R	N	cm_relation.authority_id	
1.6.1.3.1.8.2	auditing.createdBy	R	N	cm_relation.created_by	
1.6.1.3.1.8.3	auditing.creationDate	R	N	cm_relation.creation_date	
1.6.1.3.1.8.4	auditing.lastModifiedBy	R	N	cm_relation.last_modified_by	
1.6.1.3.1.8.5	auditing.lastModifiedDate	R	N	cm_relation.last_modified_date	
1.6.1.3.1.9	<listOfCitations>	O	N		
1.6.1.3.1.9.1	<citation>	R	Y		The citations come from different tables depending on the type: cm_journal_citation, cm_entrez_citation, cm_book_citation, cm_database_citation, or cm_freeform_citation.
1.6.1.3.1.9.1.1	type	R	N		
1.6.1.3.1.9.1.2	title	O	N		
1.6.1.3.1.9.1.3	authors	O	N		
1.6.1.3.1.9.1.4	publisher	O	N		
1.6.1.3.1.9.1.5	editor	O	N		
1.6.1.3.1.9.1.6	edition	O	N		
1.6.1.3.1.9.1.7	publicationLocation	O	N		
1.6.1.3.1.9.1.8	pages	O	N		
1.6.1.3.1.9.1.9	pubDate	O	N		
1.6.1.3.1.9.1.10	docType	O	N		
1.6.1.3.1.9.1.11	journal	O	N		
1.6.1.3.1.9.1.12	year	O	N		
1.6.1.3.1.9.1.13	volume	O	N		

1.6.1.3.1.9.1.14	issue	O	N		
1.6.1.3.1.9.1.15	source	O	N		
1.6.1.3.1.9.1.16	abstractUrl	O	N		
1.6.1.3.1.9.1.17	accession	O	N		
1.6.1.3.1.9.1.18	dbname	O	N		
1.6.1.3.1.9.1.19	definition	O	N		
1.6.1.3.1.9.1.20	date	O	N		
1.6.1.3.1.9.1.21	pmid	O	N		
1.6.2	<listOfRelationGroups>	O	N		
1.6.2.1	<relationGroup>	R	Y		Corresponds to a row in the cm_map_groups table
1.6.2.1.1	relationGroup.id	R	N	cm_map_group.resource_id	
1.6.2.1.2	relationGroup.type	R	N	cm_map_group.group_type	Must be “alternative reactions”, “competing reactions”, “coupled reactions”, “negative feedback”, “negative feedforward”, “positive feedback”, “positive feedforward”, or “simultaneous reactions”
1.6.2.1.3	relationGroup.lastModifiedBy	O	N	cm_map_group.last_modified_by	
1.6.2.1.4	relationGroup.lastModifiedDate	O	N	cm_map_group.last_modified_date	
1.6.3	<listOfProducts>	R	N		
1.6.3.1	<speciesReference>	R	N		This element is repeatable in SBML Level 2, but not in the STKE Connections Map Metadata Schema, because every relation in the Connections Map has exactly one modifier (upstream node) and exactly one product (downstream node).
1.6.3.2	speciesReference.species	R	N	cm_relation.target_node_id	A <species> element with

					species.id equal to speciesReference.species must exist in the <listOfSpecies> element (1.5).
1.6.4	<listOfModifiers>	R	N		
1.6.4.1	<modifierSpeciesReference>	R	N		This element is repeatable in SBML Level 2, but not in the STKE Connections Map Metadata Schema, because every relation in the Connections Map has exactly one modifier (upstream node) and exactly one product (downstream node).
1.6.4.1.1	modifierSpeciesReference.species	R	N	cm_relation.source_node_id	A <species> element with species.id equal to modifierSpeciesReference.species must exist in the <listOfSpecies> element (1.5).

APPENDIX D. STKE CONNECTIONS MAP COMPONENT LIBRARY METADATA

The following table includes the complete list of the XML elements and attributes that define the Component Library Metadata file. The table also describes whether the element or attribute is required (R is for required and O is for optional), whether the element is repeatable (Y is for "yes", N is for "no"), the source field in the database, and some notes.

An element or attribute is required if it must be present within its immediate ancestor element for the document to conform to the STKE Connections Map Metadata Schema. However, if an element or attribute is optional and a value is available from the source database field, that value should be provided in the SBML file.

An element is repeatable if it several elements – each pertaining to different data, but conforming to the same part of the specification – may be present. The element is only repeatable with respect to its immediate ancestor element. Attributes are never repeatable in XML documents.

XML Outline Number	**SBML Element or Attribute**	**R or O**	**Repeatable**	**Database Field**	**Notes**

1	<componentLibrary>	R	N		
1.1	<controlledVocabularyReference>	R	Y		
1.1.1	controlledVocabularyReference.type	R	N	cm_controlled_vocabulary.table_attribute	Must include controlled vocabularies type="organism" and "species_type".
1.1.2	controlledVocabularyReference.location	R	N		
1.2	<listOfComponents>	R	N		
1.2.1	<component>	R	Y		
1.2.1.1	component.id	R	N	cm_component.resource_id	
1.2.1.2	component.name	R	N	cm_component.name	
1.2.1.3	component.abbreviation	R	N	cm_component.abbr	Should be omitted if cm_component.abbreviation is *NULL*.
1.2.1.4	component.dataType	R	N	cm_component.data_type	Must be "simple", "complex" or "token".
1.2.1.5	component.type	R	N	cm_component.type	This field is a foreign key into the cm_controlled_vocabulary table. Should be omitted if cm_component.data_type is "complex" or "token". Otherwise, this attribute is required.
1.2.1.6	component.organism	O	N	cm_component.organism	This field is a foreign key into the cm_controlled_vocabulary table. Should be omitted if cm_component.organism is *NULL*.
1.2.1.7	component.parentID	O	N	cm_component.parent_id	Should be omitted if cm_component.parent_id is *NULL*
1.2.1.8	<description>	R	N	cm_component.description	
1.2.1.9	<listOfAuthorities>	R	N		

1.2.1.9.1	<authority>	R	Y		Must be at least one authority listed. Corresponds to a row from the cm_authorities table with cm_authorities.resource_id equal to attribute component.id. If only one authority, then corresponding="1", if more than one authority, then only one will be designated with corresponding="1" and all others will be set to corresponding="0".
1.2.1.9.1.1	authority.id	R	N	cm_authorities.authority_id	
1.2.1.9.1.2	authority.corresponding	R	N	cm_authorities.corrresponding	
1.2.1.10	<auditing>	O	N		
1.2.1.10.1	auditing.authority.id	R	N	cm_component.authority_id	
1.2.1.10.2	auditing.createdBy	R	N	cm_component.created_by	
1.2.1.10.3	auditing.creationDate	R	N	cm_component.creation_date	
1.2.1.10.4	auditing.lastModifiedBy	R	N	cm_component.last_modified_by	
1.2.1.10.5	auditing.lastModifiedDate	R	N	cm_component.last_modified_date	
1.2.1.11	<listOfHomologs>	O	N		
1.2.1.11.1	<homolog>	R	Y		Corresponds to a row from the cm_component_homolog table with cm_component_homolog.resource_id equal to attribute component.id (1.2.1.1)
1.2.1.11.1.1	homolog.id	R	N	cm_component_homolog.homolog_id	
1.2.1.12	<listOfSynonyms>	O	N		

1.2.1.12.1	<synonym>	R	Y		Corresponds to a row from the cm_component_synonym table with cm_component_synonym.resource_id equal to attribute component.id (1.2.1.1)
1.2.1.12.1.1	synonym.value	R	N	cm_component_synonym.value	
1.2.1.12.1.2	synonym.lastModifiedDate	O	N	cm_component_synonym.last_modified_date	
1.2.1.12.1.3	synonym.lastModifiedBy	O	N	cm_component_synonym.last_modifiedy_by	
1.2.1.13	<listOfConstituents>	O	N		Should be omitted if component.dataType (1.2.1.4) is "simple". Required otherwise.
1.2.1.13.1	<constituent>	R	Y		Corresponds to a row from the cm_component_complex table with cm_component_complex.resource_id equal to attribute component.id (1.2.1.1)
1.2.1.13.1.1	constituent.id	R	N	cm_component_complex.resource_id	
1.2.1.13.1.2	constituent.position	R	N	cm_component_complex.position	
1.2.1.13.1.3	<auditing>	O	N		
1.2.1.13.1.3.1	auditing.authority.id	R	N	cm_component_complex.authority_id	
1.2.1.13.1.3.2	auditing.createdBy	R	N	cm_component_complex.created_by	
1.2.1.13.1.3.3	auditing.creationDate	R	N	cm_component_complex.creation_date	
1.2.1.13.1.3.4	auditing.lastModifiedBy	R	N	cm_component_complex.last_modified_by	

1.2.1.13.1.3.5	auditing.lastModifiedDate	R	N	cm_component_complex.last_modified_date	
1.2.1.14	<listOfCitations>	O	N		
1.2.1.14.1	<citation>	R	Y		The citations come from different tables depending on the type: cm_journal_citation, cm_entrez_citation, cm_book_citation, cm_database_citation, or cm_freeform_citation.
1.2.1.14.1.1	type	R	N		
1.2.1.14.1.2	title	O	N		
1.2.1.14.1.3	authors	O	N		
1.2.1.14.1.4	publisher	O	N		
1.2.1.14.1.5	editor	O	N		
1.2.1.14.1.6	edition	O	N		
1.2.1.14.1.7	publicationLocation	O	N		
1.2.1.14.1.8	pages	O	N		
1.2.1.14.1.9	pubDate	O	N		
1.2.1.14.1.10	docType	O	N		
1.2.1.14.1.11	journal	O	N		
1.2.1.14.1.12	year	O	N		
1.2.1.14.1.13	volume	O	N		
1.2.1.14.1.14	issue	O	N		
1.2.1.14.1.15	source	O	N		
1.2.1.14.1.16	abstractUrl	O	N		
1.2.1.14.1.17	accession	O	N		
1.2.1.14.1.18	dbname	O	N		
1.2.1.14.1.19	definition	O	N		
1.2.1.14.1.20	date	O	N		
1.2.1.14.1.21	pmid	O	N		

APPENDIX E. STKE CONNECTIONS MAP CONTROLLED VOCABULARY METADATA

The following table includes the complete list of the XML elements and attributes that define the Controlled Vocabulary file. The table also describes whether the element or attribute is required (R is for required and O is for optional), whether the element is repeatable (Y is for "yes", N is for "no"), the source field in the database, and some notes.

An element or attribute is required if it must be present within its immediate ancestor element for the document to conform to the STKE Connections Map Metadata Schema. However, if an element or attribute is optional and a value is available from the source database field, that value should be provided in the SBML file.

An element is repeatable if several elements–each pertaining to different data, but conforming to the same part of the specification–may be present. The element is only repeatable with respect to its immediate ancestor element. Attributes are never repeatable in XML documents.

Outline Number	**SBML Element or Attribute**	**R or O**	**Repeatable**	**Database Field**	**Notes**
1	<controlledVocabulary>	R	N		
1.1	<listOfTerms>	R	N		
1.1.1	<term>	R	Y		
1.1.1.1	term.id	R	N	cm_controlled_vocabulary.resource_id	
1.1.3	term.value	R	N	cm_controlled_vocabulary.cv_term	
1.1.4	term.parentID	O	N	cm_controlled_vocabulary.parent_id	Should be omitted if cm_controlled_vocabulary.parent_id is *NULL*.

APPENDIX F. STKE CONNECTIONS MAP USER LIBRARY METADATA

The following table includes the complete list of the XML elements and attributes that define the userLibrary file. The table also describes whether the element or attribute is required (R is for required and O is for optional), whether the element is repeatable (Y is for "yes", N is for "no"), the source field in the database, and some notes.

An element or attribute is required if it must be present within its immediate ancestor element for the document to conform to the STKE Connections Map Metadata Schema. However, if an element or attribute is optional and a value is available from the source database field, that value should be provided in the SBML file.

An element is repeatable if several elements–each pertaining to different data, but conforming to the same part of the specification–may be present. The element is only repeatable with respect to its immediate ancestor element. Attributes are never repeatable in XML documents.

Outline Number	**SBML Element or Attribute**	**Required**	**Repeatable**	**Database Field**	**Notes**
1	<listOfUsers>	R	N		
1.1	<user>	R	Y		
1.1.1	user.id	R	N	cm_user.user_id	
1.1.2	user.firstName	R	N	cm_user.fname	
1.1.3	user.lastName	R	N	cm_user.lname	
1.1.4	user.initials	O	N	cm.user.finitials	Should be omitted if cm_user.initials is *NULL*.
1.1.5	user.email	O	N	cm_user.email	Should be omitted if cm_user.email is *NULL*.
1.1.6	user.department	O	N	cm_user.department	Should be omitted if cm_user.department is *NULL*
1.1.7	user.address	O	N	cm_user.address	Should be omitted if cm_user.address is *NULL*
1.1.8	user.city	O	N	cm.user.city	Should be omitted if cm_user.city is *NULL*.
1.1.9	user.state	O	N	cm_user.state	Should be omitted if cm_user.state is *NULL*.
1.1.10	user.country	O	N	cm_user.country	Should be omitted if cm_user.country is *NULL*

1.1.11	user.zip	O	N	cm_user.zip	Should be omitted if cm_user.zip is *NULL*
1.1.12	user.telephone	O	N	cm.user.telephone	Should be omitted if cm_user.telephone is *NULL*.
1.1.15	user.affiliation	O	N	cm_user.au_affiliation	Should be omitted if cm_user.au_affiliation is *NULL*