A Graphical User Interface for a Comparative Anatomy Information System: Design, Implementation and Usage Scenarios

Ravensara S. Travillian, MS, MA[1]; Kremena Diatchka[2]; Tejinder K. Judge[3]; Katarzyna Wilamowska[4]; Linda G. Shapiro, PhD[4]

[1] Dept. of Biomedical and Health Informatics, University of Washington School of Medicine, Seattle, WA; [2] School of Electrical Engineering and Computer Science, Oregon State University, Corvallis, OR; [3]Dept. of Computer Science, University of Wisconsin; [4] Dept. of Computer Science and Engineering, University of Washington, Seattle, WA

ABSTRACT

Building on our previous design work in the development of the Structural Difference Method (SDM) for describing anatomical similarities and differences across species, we describe the design and implementation of the associated comparative anatomy information system (CAIS) interface and provide scenarios from the literature for its use by research scientists.

INTRODUCTION

In previous work, we described the development of the Structural Difference Method (SDM) formalism for representing the similarities and differences between homologous structures across different species [1]. Additionally, we proposed the design of a comparative anatomy information system (CAIS), based on the SDM, to support queries about those similarities and differences [2]. This paper reports on the development and implementation of a graphical user interface for that system, as well as on our experiments with the use of CAIS, including scenarios from rodent-human research that show how the system can be used for realistic studies.

THE CAIS SYSTEM

The CAIS system [2] was designed to allow a user to study the similarities and differences between anatomical entities in two species. As in the Emily query interface to the FMA [3], queries to the CAIS system have the basic form:

<anat. entity1> <query relation> <anat. entity2>

where <anat. entity 1> is an anatomical entity from the first species, <anat. entity 2> is an anatomical entity from the second species, and the query relation is one of the following operators: *similar-to, different, shared, not shared, union, is-homologous?,* and *is-different?*. Either <anat. entity1> or <anat. entity2> can be **Unknown**, in which case the system returns a mapping for the specified anatomical entity if one exists in the database. If there are two anatomical entities

specified, one in each species, or if the **Unknown** reference has been resolved, the system returns the information as requested by the chosen operator, as well as references to the scientific literature on which that information is based. The operators, defined based on graph matching [1,2], can be summarized as follows.

Result set operators:

similar-to: returns an anatomical iso-morphism (1-to-1 and onto correspondence) between the two homologous structures across species at the level of granularity (*e.g.*, `Organ`, `Organ part`, `Cell`) of the query if there is one, and returns `False` otherwise. For example, the `Left` and `Right atria` and `Left` and `Right ventricles` of the `Heart` are similar between the mouse and the human.

different: returns a non-null correspondence other than anatomical isomorphism (*e.g.*, a one-to-many relationship) between two homologous structures across species at the level of granularity of the query if there is one, and `False` if there is no mapping in the database. For example, the `Left lobes` of the mouse and human `Lungs` are different because they are in a 1:2 relationship.

shared: returns all the parts of the structure which occur in both species to the level of granularity specified. For example, the human and mouse brains both contain an `Amygdala`, so `Amygdala` would be one of the structures returned on a *shared* query on human and mouse `Brain`.

not shared: returns all the parts of the structure which occur in one species or the other, but not both, to the level of granularity specified; this is the set complement of the structures returned by *shared*. For example, the human brain *includes* `Gyri` and `Sulci` that mouse brains do not, so the *not shared* relation between human and mouse brains would contain those `Gyri` and `Sulci` (among other structures).

union: returns all the parts of the structure which occur either in one species or the other, or in both, to the level of granularity specified—in other words, the set union of the structures returned by the CAIS relationships *shared* and *not shared.*

Boolean operators:

is-homologous? returns `True` if the two entities selected for the query are homologous, and `False` if they are not.

is-different? is the opposite of *is-homologous?*—it returns `False` if the two entries selected for the query are homologous, and `True` if they are not.
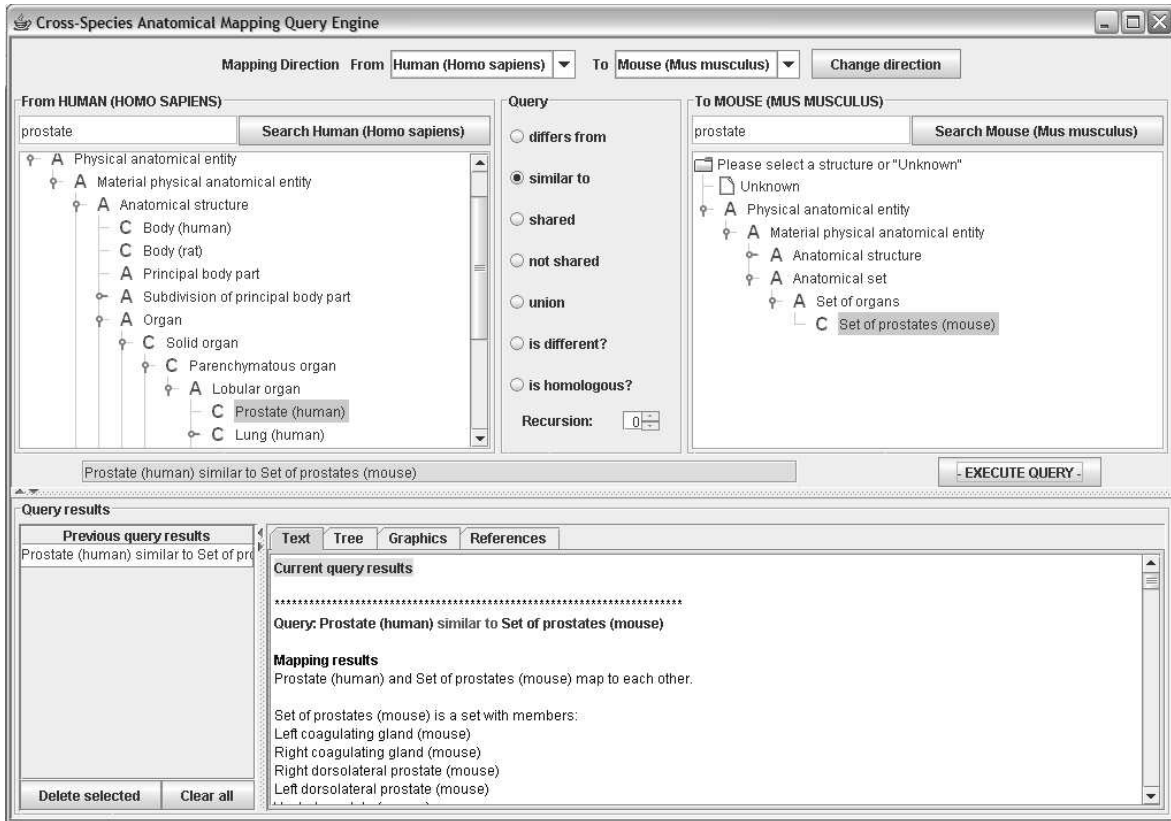

Figure 1 illustrates a screen shot of the CAIS graphical user interface that shows the results in text mode.

THE CAIS INTERFACE

To make the CAIS query functionality available to users, we have designed and implemented a graphical user interface (Figure 1). The CAIS interface is written in Java, and uses the Java API to access the Protégé-2000 database, in which rat, mouse, and human anatomical structures comprise a single hierarchy [2]. The CAIS interface provides the following functionalities.

1. choose the pair of species to compare from all species in the database,

2. select an anatomical entity from a hierarchy or search for one that the user

has entered and give him/her a choice if the entry is ambiguous,

3. inform the user if selected entities cannot be directly compared and indicate reasonable alternatives if they exist,

4. select the query operator from a list of choices,

5. show the user query in a string form as the user constructs it from the GUI,

6. compare the selected structures at multiple levels of the parts hierarchy as selected by the user (default is 1 level)

7. keep track of results from prior queries so the user can return to them, and

8. show the output in multiple forms including text, tree, graphics, and references.

Figure 1 shows a screen shot of the full user interface. The user has selected the species human on the left and mouse on the right. She has typed in "prostate" in the search area on the left, and the system has found the human prostate in the hierarchy and displayed it. She has also typed in "prostate" in the search area on the right, and the system has responded with a message, "Select from search results," and displayed four possibilities from which the user has selected "Set of prostates (mouse)". She has then selected the query operator "similar" and clicked on "Execute Query." The query has been executed, and the results displayed in text mode, since the text tab is the default display tab. As the text mode is very verbose, the user may wish next to look at the results in tree mode (Figure 2) or graphics format ((Figure 3). Tree results are returned as a structured hierarchy, down as many levels of the tree as was specified in the selected recursion level. In the graphics results, a representative graphic is included at each level of the hierarchy.
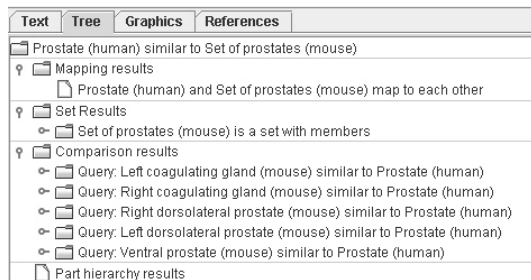

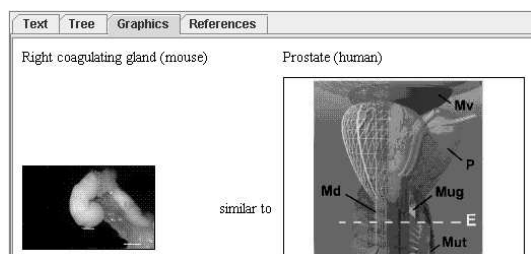Figure 2 illustrates the tree display mode.

Figure 3 illustrates the graphics display mode.

SCENARIOS

In order to illustrate the potential use of the CAIS system, we give several research scenarios from the literature. We motivate the need for such a tool in each scenario and give examples of CAIS queries (in simplified string form) that can be used by the researchers in these studies.

Scenario 1: Correlating prostatic lobes/organs

Dr. A is a pharmacological scientist who is studying the effect of candidate compounds for new prostate cancer drugs. Because different regions of the human prostate are subject to different diseases, those regions that develop benign prostatic hyperplasia do not develop cancer, and vice-versa. Dr. A wants to determine the rat-human homologies for the dorsolateral and ventral regions of the prostate, so that she can correlate the observed effects of the compounds in rat tissue with predictions for the effects in humans expected to be observed in later clinical trials. Specifically, her questions are: do the dorsolateral prostates of the rat correspond to the dorsolateral regions of the human prostate, and does the rat ventral prostate correspond to the ventral region of the human prostate (called anterior lobe in humans)?

The CAIS operator *similar-to* provides information for the researcher on what structures are homologous across species, what evidence exists that they are homologous (*e.g.*, traditional embryological studies, genetic assays), and the provenance or source of that information. Dr. A's queries will be:

<Dorsolateral prostate (rat)> <is-homologous?> <Dorsal lobe of prostate (human)>

<Ventral prostate(rat)> <is-homologous?> <Anterior lobe of prostate (human)>.

The attributed relationship returned by CAIS answers the researcher's query: documented in Dorothy Price's embryological work on "Comparative Aspects of Development and Structure in the Prostate" in the *National Cancer Institute Monograph* 1963 Oct. 12:1-27 [4]. The rat dorsolateral prostates are homologous to the dorsolateral lobes of the human, while the rat ventral prostate is *not* homologous to the anterior lobe of the human prostate. Based on this information, Dr. A. adjusts her expected correlations of the compound's effect accordingly.

Scenario 2: Correlating prostatic zones/organs

Dr. B is a pathologist who is formally developing new mouse models of human prostate cancer. Part of his evaluation is the application of analyses of previous results in mouse modeling of human prostate cancer and the determination of what those analyses imply for a mouse model that more soundly mirrors the initial development and the subsequent progression of prostatic tumors.

He has a candidate model in mind, pending confirmation of certain homologies. Given that the human peripheral zone is the region in which most prostate carcinomas originate, his question to establish the validity of that candidate rests on the results of the following correspondence: what is the mouse prostate region corresponding to the human peripheral zone of the prostate? His CAIS query based on that question will be in the form

<Unknown (mouse)> <similar-to> <Peripheral zone of prostate (human)>.

In this case, CAIS can be used to return not only the result set for the query, but also the references that back up the result, including, for example, the information that on the basis of an epidemiological study, Xue [5] reports that the mouse dorsolateral prostate corresponds to the peripheral zone of the human prostate, and that Roy-Burman [6] concurs on a preliminary basis, but cautions that Xue's assertion is based on descriptive data, and that the molecular studies that would confirm the correspondence remain to be carried out. Based on this information, Dr. B. determines that his mouse model is as yet insufficiently validated, and incorporates certain molecular assays on the dorsolateral prostate as part of the validation process for this model.

Scenario 3: Shared similarities and differences in prostate for tumor microenvironment

Dr. C wants to determine the best mouse tumor model for determining clinically relevant information on the response of tumors to a particular treatment effect. Bearing in mind the significant role the tumor host microenvironment (in this case, vasculature among other variables) can play in establishment of the tumor and its response to treatment, Dr. C. requires information on what aspects of the prostatic epithelium—the tumor microenvironment—are similar between the mouse and the human, and what aspects are different.

The queries *shared*, *not-shared*, and *union* provide information about the documented evolutionary possibilities for a given anatomical structure. For example, to confirm that the basic cellular structure of the mouse and human prostates are similar enough to support generalizing from the mouse tumor microenvironment to the human (a subset of Dr. C's eventual result set) the researcher may wish to verify that the prostates in both species consist of the same types of cells. This researcher's query would take the form

<Prostatic epithelium (mouse)> <shared> <Prostatic epithelium (human)>.

CAIS would return the result set that the prostatic epithelium in both species share the following cell types: {Secretory epithelial cell, Basal epithelial cell, Neuroendocrine epithelial cells}, citing Marker [7] and Garabedian [8], among others, as sources for this information, and verifying for this researcher that the species are histologically similar enough to validate a particular proposed study. The results of previous queries are accessible for use in building the compound query, which will return the totality of the shared features of the tumor microenvironment.

Scenario 4: Union of all normal stem cells as basis of a breast cancer tumor cell taxonomy

Dr. D is a tumor biologist who uses genome-wide expression analysis on normal luminal epithelial and myoepithelial/basal lineages of tumor cells for molecular classification of breast cancer, to the end of developing therapies that are less toxic than traditional radiation or chemotherapy treatment. As a first step in this research, he is working on a cross-species stem cell hierarchy, which he expects to reveal important aspects of the histogenesis of breast cancer evolution.

The CAIS operator *union* gives the range of all normal possibilities of these structures in the species under examination. CAIS will return all of the similarities and differences at all levels of granularity in the knowledge base in response to a *union* query. In order for Dr. D to obtain the desired information for his hierarchy, a detailed compound query on the relevant anatomical sites is necessary. One representative component of this compound query is

<Epithelial cell of mammary gland (mouse)> <union> <Epithelial cell of lactiferous duct tree (human)>.

The researcher builds up the query from components like this, and submits the query *in toto* to CAIS. Based on the information returned, Dr. D now has a result set from which he can derive his cell hierarchy, which will underlie his examination of breast cancer histogenesis.

DISCUSSION/CONCLUSIONS

Drawing on the Structural Difference Method [SDM], developed in previous work [1], we developed and implemented an application that extracts cross-species anatomical information from a Protégé-2000 database file, and allows the users to query the application about correspondences and differences in those anatomical structures. We implemented features to make the application more user-friendly, such as allowing the user to build a query by clicking, rather than being forced to remember the syntax, and by allowing the user to view and change the query as it is being constructed.

We provide a search feature, and control which classes can be searched and selected. All of the set operators developed for the SDM have been implemented, and permit different aspects of anatomical correlation to be queried. The tabs provide different views for users to choose among, including unstructured set results, a structured hierarchy of results, graphics for comparison, and attributed slots that describe the basis (embryological or genetic) of the anatomical correlation, and the provenance of the information it was based upon.

An important knowledge representation aspect remains for future work—although we hinted at the issue of conflicting sources in our first scenario, this version of our application does not deal systematically with sources that conflict with each other or with our model. Determining the appropriate way to resolve and represent such conflicting information will comprise a critical component of our future research.

Based on correspondence with domain experts, we are incorporating information on 5 different rodent organs (mammary gland, prostate, lung, ovary, and cervix), and preliminary feedback from users indicate a very welcome reception. In fact, the need for communicating these anatomical correspondences is becoming greater as the research into animal models of disease becomes more interdisciplinary and as researchers come from other backgrounds than traditional comparative anatomy. Our scenarios reflect the real need expressed by users for valid comparative anatomy information available in a user-friendly manner.

ACKNOWLEDGMENTS

REFERENCES

1. Travillian RS, Rosse C, Shapiro LG. An approach to the anatomical correlation of species through the Foundational Model of Anatomy. AMIA Annu Symp Proc. 2003;:669-73.

2. Travillian RS, Gennari JH, Shapiro LG. Of Mice and Men: Design of a Comparative Anatomy Information System. AMIA Annu Symp Proc. 2005.

3. Detwiler LT, Chung E, Li A, Mejino JL Jr, Agoncillo A, Brinkley J, Rosse C, Shapiro L. A relation-centric query engine for the Foundational Model of Anatomy. Medinfo. 2004;11(Pt 1):341-5.

4. Price D. Comparative aspects of development and structure in the prostate. Natl Cancer Inst Monogr. 1963 Oct;12:1-27.

5. Xue L, Yang K, Newmark H, Lipkin M. Induced hyperproliferation in epithelial cells of mouse prostate by a Western-style diet. Carcinogenesis. 1997 May;18(5):995-9.

6. Roy-Burman P, Wu H, Powell WC, Hagenkord J, Cohen MB. Genetically defined mouse models that mimic natural aspects of human prostate cancer development. Endocr Relat Cancer. 2004 Jun;11(2):225-54.

7. Marker PC, Donjacour AA, Dahiya R, Cunha GR. Hormonal, cellular, and molecular control of prostatic development. Dev Biol. 2003 Jan 15;253(2):165-74. Review.

8. Garabedian EM, Humphrey PA, Gordon JI. A transgenic mouse model of metastatic prostate cancer originating from neuroendocrine cells. Proc Natl Acad Sci U S A. 1998 Dec 22;95(26):15382-7.