

Integrating Genomic Knowledge Sources Through an Anatomy Ontology

J.H. Gennari, A. Silberfein, and J.C. Wiley

Pacific Symposium on Biocomputing 10:115-126(2005)

INTEGRATING GENOMIC KNOWLEDGE SOURCES THROUGH AN ANATOMY ONTOLOGY

JOHN H. GENNARI, ADAM SILBERFEIN, & JESSE C. WILEY

*Biomedical & Health Informatics, The Information School, & Comparative Medicine,
University of Washington, Seattle, WA, 98195, USA*

Modern genomic research has access to a plethora of knowledge sources. Often, it is imperative that researchers combine and integrate knowledge from multiple perspectives. Although some technology exists for connecting data and knowledge bases, these methods are only just beginning to be successfully applied to research in modern cell biology. In this paper, we argue that one way to integrate multiple knowledge sources is through anatomy—both generic cellular anatomy, as well as anatomic knowledge about the tissues and organs that may be studied via microarray gene expression experiments. We present two examples where we have combined a large ontology of human anatomy (the FMA) with other genomic knowledge sources: the gene ontology (GO) and the mouse genomic databases (MGD) of the Jackson Labs. These two initial examples of knowledge integration provide a proof of concept that anatomy can act as a hub through which we can usefully combine a variety of genomic knowledge and data.

1 The Problem: Overwhelming, Distributed Genomic Knowledge

Modern biology researchers are hampered by the need to integrate information from rapidly developing and diverse knowledge sources. As a general problem, researchers in computer science and informatics have developed methods for combining and integrating data and knowledge bases. However, these methods are only beginning to be applied to practical problems in biology research. In this paper, we present an example of knowledge integration where we use an anatomy ontology as a hub through which we have connected data sources for two disparate views of cell biology.

One can view modern cell biology research as having two branches. In one core branch of research, molecular biologists and biochemists use model systems such as cell culture and yeast based assays to examine general principles of cell biology. These researchers describe the structure and function of the abstract cell, irrespective of how that cell participates in larger systems. Alternatively, researchers also study anatomy-specific aspects of biology, such as developmental biology and disease pathology. Necessarily, the latter approach is tissue-specific; however, their work must also understand and be consistent with the more generic approach to cell biology. Specifically, if the generic model of the cell specifies all possible genetic interactions and functions, then a tissue specific model must account only for those genes that are expressed in the tissue.

An ideal informatics solution would allow researchers who focus on either approach to see and understand results from both these views. Since anatomy underlies all of biological research, we propose an anatomy-based platform for integrating data sources. The long-term goal of this platform is to allow researchers to associate genes with cellular function from both viewpoints—both at the level of an ‘abstract cell’ and in tissue-specific fashion.

Our motivation for this work is linked to the goals and informatics needs of a current genomic research effort. One of the co-authors (JCW) is also a member of the Comparative Mouse Genomics Centers Consortium (CMGCC). The primary goal of the CMGCC is to identify, and produce genetic mouse models for, human genetic variants of genes believed to be ‘environmentally sensitive’ and linked to human pathological conditions (www.niehs.nih.gov/cmgcc/). The goal is to use the mouse models developed to explore genetically conferred diseases prevalence. The mouse consortium is building genetic mouse models of low frequency variants of genes involved in cell cycle control and DNA repair mechanisms—two target biological processes widely believed to be of significance in determining disease prevalence. In this research context, linking together functional genetic information with anatomy is of particular significance. Although a given mouse model may be developed with a specific pathological condition in mind, the genetic manipulation will affect biological processes distributed across the entire organism and may have different manifestations across different tissue types.

In this paper, we provide two proof-of-concept demonstrations where anatomy is the central hub for both tissue-specific and for abstract cellular genetic data. For anatomic concepts, we use the Foundational Model of Anatomy (FMA), a comprehensive, well-structured ontology of anatomy [1,2]. In our first demonstration, we show how the FMA can support the investigation of the abstract cell, by connecting the FMA ontology of cellular anatomy (the structure and sub-components of the cell) with the Gene Ontology (GO) hierarchy of cell component terms [3]. Given this connection, our tool can browse GO-annotated databases from within the FMA viewer.

As a second demonstration, we have also connected the FMA to a database of tissue-specific gene expression results. Specifically, we used the Mouse Genomic Database from Jackson Labs that includes anatomic information about gene expression results in the mouse [4]. To scope our effort, we looked at tissues of the brain, using the FMA organization and description of brain regions and components. In Section 3, we provide details and screenshots from both of these demonstration projects. However, before giving these details, we first describe the relevant bioinformatics resources and efforts in knowledge sharing.

2 Standards and Knowledge Bases for Bioinformatics Sharing

Our work builds on a number of important bioinformatics resources. In general, there are many groups working to standardize bioinformatics knowledge and data. For example, the Microarray Gene Expression Data Society aims to facilitate the sharing of microarray data through the development of standards for experimental design descriptions and data descriptions (www.mged.org/). Our work is designed to include other bioinformatics resources as they become available.

2.1 Gene Annotation and The Gene Ontology

The Gene Ontology (GO) is a structured, controlled vocabulary to allow molecular biologist to better share data and knowledge about the roles of gene products [3]. The usual use of the GO is to provide researchers with a standard language for annotating a gene—providing information about a gene product, such as its molecular function. It is organized into three hierarchies: (1) a set of terms for the molecular functions of a gene product, (2) terms that describe the larger-scale biological processes that may involve the gene product, and (3) terms that include the cellular components that may be important or relevant for the gene product. The GO views all cell types equally, and does not include any tissue-specific information. The GO is developed collaboratively and is an evolving vocabulary, with new versions released monthly.

An important aspect of the GO is that a set of databases exist that contain information about particular gene products, in particular research species, that have been studied and annotated with the GO controlled vocabulary. For example, the GO was developed by the groups developing *Flybase*, the *Saccharomyces* Genome Database, and the Mouse Genome Database (MGD), so the genes in all of these sources are annotated with appropriate GO terms. Currently the GO web pages list more than 30 such annotated databases contributed by about 15 groups worldwide. GO provides a uniform search capability: researchers can use GO terms to retrieve related gene products across these multiple databases.

Some researchers have pointed out that the GO has some problems and weaknesses, if viewed as a formal ontology rather than a controlled vocabulary [5,6]. However, in our current work, we focus on the GO as a simple entry point into the annotation databases such as the MGD. Our assumption is that the GO will improve over time, allowing for better, higher quality inferences and search capabilities.

2.2 Standards and Ontologies for Anatomy

For our work, we are interested in efforts to standardize gene expression annotations about anatomy—information about the source of a tissue sample. One group that is

focused on anatomy for gene expression is the Standards and Ontologies for Functional Genomics (SOFG) group (www.sofg.org/). This group recognizes a problem faced by a genetic researcher working with animal models. One would like to annotate results with anatomic information that is consistent across both the animal model of disease (e.g., mouse) and human, the eventual target for therapies. Even within a single species, there is variability and inconsistency in anatomic labeling. As with other forms of meta-data, the tendency is for scientists to use informal, natural language terms, rather than terms from a controlled vocabulary or ontology.

As steps toward ameliorating this situation, the SOFG group has identified a number of on-line resources for anatomy ontologies, and developed a short “SOFG anatomy entry list” (SAEL). This list of about 100 anatomic terms represents the most commonly used anatomic terms for annotating gene expression results. The resource list includes anatomy ontologies developed by groups devoted to functional genomics in particular species (e.g., an anatomy ontology for the mouse), as well as long-standing efforts to describe human anatomy—either for medicine and pathology (OpenGalen, www.opengalen.org/), or in purely structural terms, by the Foundational Model of Anatomy (see below). The expectation is that each of these anatomy resources would match or map their anatomic terms to the entry list, and then this entry list could be used as a way to align or link together different ontologies of anatomy [7]. As we show in Section 3, we share this aim of aligning ontologies via anatomy.

2.3 *The Foundational Model of Anatomy(FMA)*

The Foundational Model of Anatomy is a long-standing project to describe all of human anatomy as a symbolic ontology of concepts and relationships [1,2]. The FMA was not designed especially for the genomics domain, nor indeed for any particular biomedical viewpoint. Rather, it was designed with the idea that if one could represent in a principled, formal manner the truth about biological structure, then the resulting knowledge base could become a reference ontology for all of biomedical informatics [1]. The FMA is already fairly comprehensive (with more than 70,000 concepts), and it continues to grow and evolve.¹

In contrast to anatomy ontologies that are designed specifically for annotating tissue-specific genetic data, the FMA is designed for and by anatomists, and this perspective has both advantages and problems. On the one hand, the FMA represents a

¹ The FMA is related to, but different from the UW Digital Anatomist (UWDA) terminology of the UMLS. The UWDA was the predecessor of the FMA, and within the UMLS this terminology was kept in sync with the FMA only through the Summer of 2002. Since then, the FMA team has added thousands of new classes, including many of the sub-cellular anatomy terms that are our focus.

much more comprehensive view of anatomy—both broader in scope and with greater detail than is necessary for current research in cell biology. However, this wealth of detail and information can overwhelm and obscure a simpler view of anatomy—as exemplified by the SAEL list of only 100 terms that covers ~80% of current anatomic annotations [7]. Additionally, the FMA contains no explicit information about the function or physiology of anatomy. Therefore, it also contains no information about dysfunction, or pathology, which is often a focus for the genomic researcher.

The FMA uses the Protégé environment for knowledge-based systems to store, retrieve, and manage its anatomic concepts and relationships [8]. Protégé is an open-source, extensible environment with a large user base. These characteristics make it easy to adapt the system for special needs. As we present in the next section, we were able to modify the Protégé user interface to connect directly to specific bioinformatics databases.

3 Results: an Anatomy Ontology as an Information Portal

Our claim is that a comprehensive anatomy ontology such as the FMA can be fruitfully used as a hub to integrate a variety of genomic and cell-biology knowledge sources. In this section, we present two examples: (1) We connected the FMA concepts of cellular anatomy to the GO annotation databases by linking to the GO terms in its cellular components tree, and (2) We connected gene expression results from MGD to the concepts in the FMA that describe brain anatomy. Figure 1 shows our anatomy ontology as a hub with tissue-specific knowledge on the left, and generic GO knowledge on the right. As we show, the user interacts with a single Protégé user interface that integrates the multiple knowledge sources into a single view.

For both systems, we leveraged the extensibility of the Protégé environment. For each linking system, we built a Java plug-in component that modified the behavior of the default user interface. In response to user browsing actions, our plug-ins access the relevant mapping tables, and then use this information to construct JDBC calls to locally stored copies of the relevant DBs. Our plug-ins then display the data returned from these calls directly within the FMA viewer.

3.1 Mapping Cellular Anatomy: the FMA and GO

As we mentioned earlier, the GO describes function, process, and cell structure independent of anatomy, tissue, or cell type. In contrast, the FMA looks at anatomy structurally, independent of function or physiology. Thus, the only way to directly link the FMA and GO knowledge sources is through cell structure.

Of course, the anatomic names and concept organization for cellular structure in the FMA do not always match the terms and organization used in GO's cell compo-

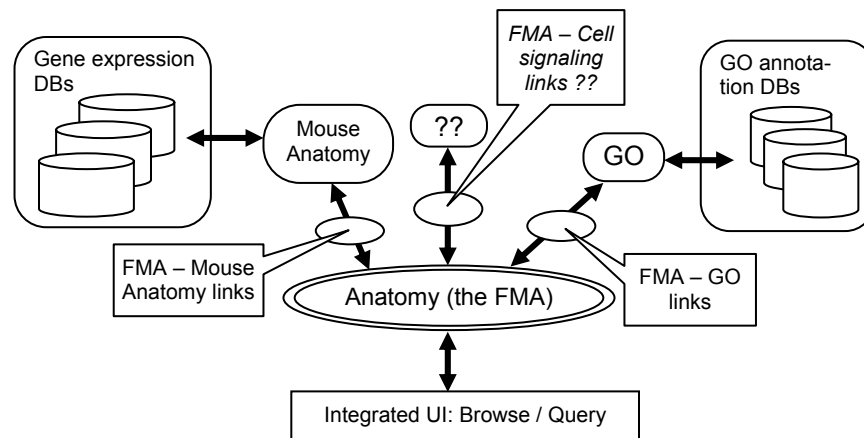


Figure 1. A diagram where our anatomy ontology acts as a hub for integrating knowledge. Mappings to GO and to mouse anatomy are described in Section 3.1 and 3.2; mappings to cell-signaling knowledge sources are future work.

ment hierarchy. If our goal is to view the GO knowledge sources from within the broad FMA anatomy ontology, we must connect terms in the FMA to terms in GO. As an initial proof-of-concept, we hand-built a database table that connects about 150 terms in the FMA’s ontology of cell parts to the corresponding terms in the GO. With this table, we then implemented a viewer within Protégé that allows direct access to the GO annotation databases from within the FMA Protégé viewer. Figure 2 shows an example of this connection. Within the FMA, we have browsed to the concept “wall of lysosome”, and we can then see both the corresponding GO term (“lysosomal membrane”) and information from the GO databases (Flybase & MGI, in this case) that show information about the gene products marked as associated with the lysosomal membrane. Currently, our interface also provides a simple table with the details about each of these GO annotations. (Users can view this additional information by selecting a particular GO annotation and hitting the “V” button.)

By itself, this system provides a new view onto the GO databases—it organizes the information according to the FMA’s formal definitions of cellular anatomy. However, the real strength of our approach is that this viewer can be combined with other anatomy-centric viewers of genetic information, as we describe below.

3.2 Mapping Gene Expression Data: MGD, FMA, & GO

In contrast to studying cellular behavior across all cell types, one may want to study a specialized set of cells—ones associated with a particular organ or tissue type. Gene expression data is one way of understanding which genes are active in which anat-

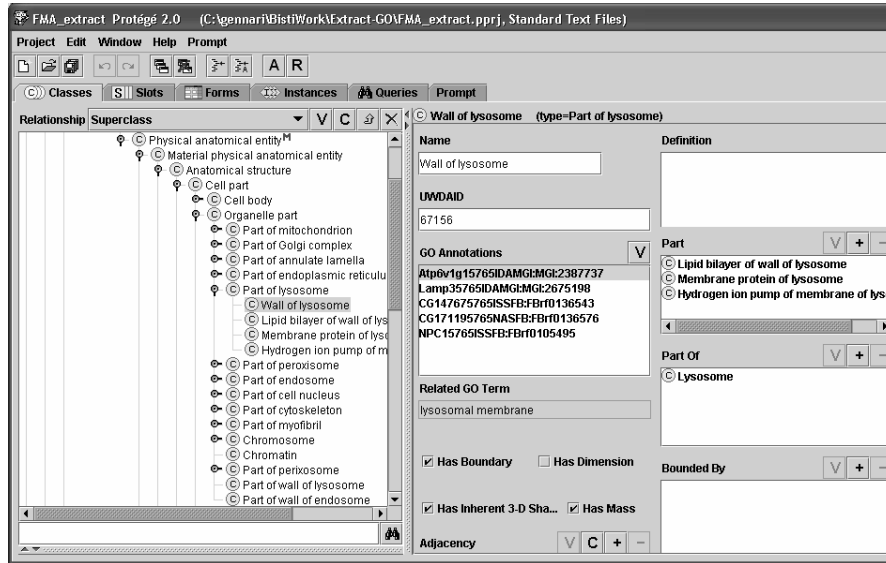


Figure 2. A screen from Protégé showing the FMA information about the wall of the lysosome and highlighting the gene products that are annotated with the GO term “lysosomal membrane”. In this case, information from MGI and Flybase are displayed. Additional details about each annotated object can be retrieved via the “V” (view) button.

omic parts. Therefore, scientists need an understanding of anatomy to effectively use and organize this data. To date, there is not a standard source for anatomic knowledge for annotation of gene expression results—as we described earlier, this is the concern and work of the SOFG group. As a concrete example, we have connected the FMA with one source of gene expression data for the mouse: the Mouse Genome Database (MGD) [4]. As with the GO—FMA mappings described in section 4.1, we created a table connecting anatomy terms as used by the MGD to concepts defined in the FMA.² To scope our work, we choose to focus on the Brain regions. (Fortunately, there are few anatomic differences between mouse brain regions and human ones.) With our mapping tables, we can connect FMA brain regions with data in the MGD that lists relevant gene expression results for that region. Figure 3 shows an example interface, with the hippocampus selected, and indicating that one can retrieve 615 gene expression results for that region.

The connection between the FMA, the GO, and the MGD data provides interesting capabilities for the cell biology researcher. For example, the ability to view gene

² We understand (personal communication from M. Ringwald) that the MGD is in the process of updating/changing the anatomic terms used to be more consistent with the SOFG efforts

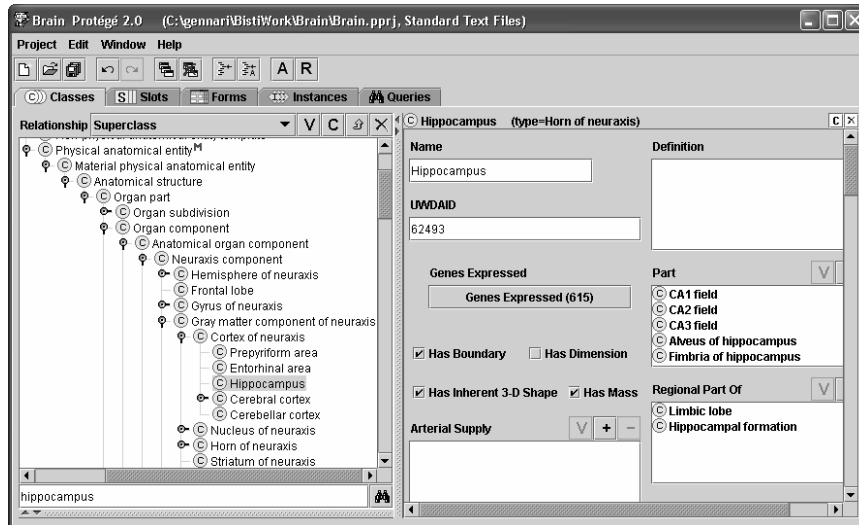


Figure 3. A screen showing FMA brain regions and the link to MGI data about gene expression results for that region. Selecting the “Genes Expressed” button produces a table of 615 results (in this case), and for each of these, its associated GO terms. As a next step, for one of these genes, one can select a specific GO term to further explore, as shown in Fig 4.

expression data by both GO term and tissue type will be of utility for the CMGCC. As we described in the introduction, this consortium studies specific pathologies via animal models, whose genetic alteration may extend beyond the intended target pathology. As an example of how genetic modifications can have different effects in different tissues, consider one target of interest for the CMGCC: the cell cycle regulatory gene cyclin D1. It is well known that cyclin D1 expression is related to colorectal cancer [9], centrocytic lymphomas [10], and mammary adenocarcinoma development [11]. However, cyclin D1 expression has also been linked with the induction of apoptosis within the post-mitotic neurons in the brain [12]. Consequently, research groups focused on cancer biology would not necessarily notice the impact of genetic manipulations of cyclin D1 in the brain as the anticipated phenotype would be the exact opposite observed in other tissues—namely, cyclin D1 is associated with cell proliferation in the context of cancer research and neuronal degeneration and death within the brain. Thus, to fully explore the impact of human genetic variants within mouse models, it is essential to possess information about both the tissue specificity of expression and the biological process within which that gene is involved.

From the interface in Figure 3, one can retrieve a table (not shown) of all the gene expression results for the selected region. If the user chooses to look at cyclin

Gene expression results for regulation of cell cycle

325 unique gene results for regulation of cell cycle

Symbol	Name	Structure	Detected
Mas1	MAS1 oncogene	hindbrain	yes
Mas1	MAS1 oncogene	hindbrain	no
Akt1	thymoma viral proto-oncogene 1	hippocampus	yes
Bcr	breakpoint cluster region homolog	hippocampus	yes
Ccna2	cyclin A2	hippocampus	yes
Ccnb1-rs1	cyclin B1, related sequence 1	hippocampus	yes
Ccnb1-rs1	cyclin B1, related sequence 1	hippocampus	no
Ccnd1	cyclin D1	hippocampus	no
Ccnd1	cyclin D1	hippocampus	yes
Ccnd2	cyclin D2	hippocampus	yes
Ccng1	cyclin G1	hippocampus	yes
Igf2	insulin-like growth factor 2	hippocampus	yes
Igfb1	integrin beta 1 (fibronectin receptor beta)	hippocampus	yes
Jun	Jun oncogene	hippocampus	yes

Gene Ontology (GO) Classifications

Category	GO Term	Evidence	Reference
Component	cyclin-dependent	IPI	MGI:2677082
Component	nucleus	IDA	MGI:2448203
Component	cytosol	IDA	MGI:2448203
Function	cyclin-dependent	IDA	MGI:2448203
Function	protein kinase activity	IDA	MGI:2652284
Function	protein binding	IPI	MGI:2652284
Function	protein binding	IPI	MGI:2677082
Function	protein binding	IPI	MGI:67412

Evidence Codes

Figure 4 A screen shot showing the set of genes that (1) have results associated with the hippocampus, and (2) are annotated with the GO process term “regulation of cell cycle”. The lower half of the Figure shows the full set of GO terms associated with cyclin D1.

D1, then one also sees all the GO annotations for that gene. Since the consortium is interested in the GO process labeled “regulation of cell-cycle”, we choose that GO term to further browse. Figure 4 then shows all genes expressed within the hippocampus that are also involved in that particular GO process. One can then continue to iteratively explore the GO annotations of Cyclin D1 by selecting other terms (see bottom half of Figure 4), and then seeing the other expressed genes organized by brain structure.

4 Related Work in Knowledge Base & Data Base Integration

The systems we have described are small prototypes—they are limited by the size of the mapping tables we built between the FMA concepts and the other knowledge sources. However, the general task of linking or integrating knowledge bases and databases has been well-studied in computer science. The challenge we face in linking knowledge sources into a central anatomy hub is a specific example of the more general task of knowledge integration in bioinformatics. For example, the Bio-

Mediator project has developed methods for answering queries about genetic tests across multiple sources [13,14]. Their approach does not rely on static tables of maps between terminologies, and instead builds more dynamic rules that describe how terms match across sources.

More generally still, in knowledge-base research, there are methods for semi-automatically determining the set of matching concepts across two ontologies. For example, within the Protégé environment, there is a plug-in tool known as Prompt, that allows users to link any two related ontologies [15]. Outside of Protégé, researchers have used similar methods to align a small portion of the FMA with the OpenGalen medical and anatomic terminology [16]. Eventually, we may want to augment our mapping work by using these sorts of methods and tools. However, compared to anatomy as a whole, cellular anatomy may not need that large of an ontology—e.g., the entire GO cell structure tree consists of less than 1300 terms. These more automated approaches will be important in the longer term, as the number and size of new genomic knowledge sources increases.

5 Future Work and Discussion

We have argued that an anatomy ontology (the FMA) can function as a unifying hub for integrating knowledge spanning from subcellular components to macro-anatomical features. We recognize that our demonstration of two examples does not suffice to show the broad applicability of our ideas. However, the ability to browse clustered data by GO terms and tissue expression distribution (shown in Figure 4), rather than hunting and searching gene by gene, is of immediate utility to at least the CMGCC community. We also acknowledge that we have only demonstrated our approach with the subset of data for brain regions. With other anatomic concepts there may be more significant differences across species. Future work will look at a more complete mapping across species, and as we described in the previous section, we may leverage tools like Prompt for helping to create these mappings.

We plan to connect other sources into our anatomy hub. For example, we know of other work that provides tissue-specific gene expression data. We may be able to link knowledge from the Gene Expression Atlas (symatlas.gnf.org/SymAtlas/), or to information reported about tissue-specific mouse transcriptions on RIKEN cDNA microarrays [17]. Of course, crossing species boundaries to the mouse raises a separate set of issues, especially when the anatomy is significantly different [18].

In the longer term, we believe that an important resource to link to our anatomy hub is knowledge about cell-signaling pathways (see middle spoke of Figure 1). The distinction we made between general-principle driven cellular biology research and tissue-specific biology research, has the same significance in the context of modeling

cellular signaling. That is, signaling pathways should be understood both from the view of the abstract cell, and in a tissue-specific manner, accounting for genetic conditions present in specific cell types. Thus, our next research step is to combine anatomy (as in the FMA) with a comprehensive ontology for cell signaling. We have already begun work designing a cell-signaling ontology and are examining related work in this area, such as that by the BioPax group (www.biopax.org). The combination of anatomy with signaling knowledge should provide significant benefits to the biology researcher who is looking for a unified view of genomic and proteomic knowledge.

The ultimate success of our work depends on several assumptions. The primary assumption is that it is possible to align diverse ontologies about cell biology, and that it will be useful to do so. The second assumption is anatomy is a useful organizing hub to connect the various sorts of functional genetic information that currently resides in distributed databases. Our results so far are an encouraging proof of principle exploration of these key assumptions. Specifically, our success in mapping parts of the GO cellular component information and the Jackson lab's mouse anatomy ontology onto the FMA suggests that commonly utilized biological structure ontologies can be linked together and viewed within an anatomy ontology.

Integrating and connecting bioinformatics knowledge resources is a long-term research task. However, given the flood of information and the need to combine data and knowledge, we believe this goal is an important one. Unfortunately, knowledge about anatomy, cellular physiology, and pathology is unlikely to be codified into a single ontology that all stake-holders can agree on. For the foreseeable future there will be different ontologies with different perspectives, and researchers must be able to appropriately combine knowledge from these different views and ontologies.

Acknowledgments

We thank Cornelius Rosse and Jose Mejino for explanations and insights into the FMA. Partial funding for this work was provided by the BISTI planning grant (#P20 LM007714) from the National Library of Medicine.

References

- [1] Rosse C, Mejino JLV. A Reference Ontology for Bioinformatics: The Foundational Model of Anatomy. *Journal of Biomedical Informatics* 2003; **36**:478-500.
- [2] Rosse C, Mejino JLV, Modayur BR, Jakobovits RM, Hinshaw KP, Brinkley JF. Motivation and Organizational Principles for Anatomical Knowledge Representation: The Digital Anatomist Symbolic Knowledge Base. *Journal of the American Medical Informatics Association* 1998; **5** (1):17-40.
- [3] GeneOntologyConsortium. Creating the gene ontology resource: design and implementation. *Genome Res* 2001; **11** (8):1425-1433.

- [4] Blake J, Richardson J, Bult C, Kadin J, Eppig J. MGD: The Mouse Genome Database. *Nucleic Acids Research* 2003; **31**:193-195.
- [5] Smith B, Williams J, Schulze-Kremer S. The ontology of the gene ontology. *Proceedings, AMIA Annual Symposium*, Washington, D.C., 2003. 609-613.
- [6] Wroe CJ, Stevens R, Goble CA, Ashburner M. A Methodology to Migrate the Gene Ontology to a Description Logic Environment Using DAML+OIL. *Pacific Symposium on Biocomputing*, 2003. 624-235.
- [7] Aitken S, Baldock R, Bard J, et al. SAEL -- The SOFG anatomy entry list. *ISMB 2004*, Glasgow, Scotland, 2004.
- [8] Gennari JH, Musen MA, Fergerson RW, Grosso WE, Crubezy M, Eriksson H, Noy NF, Tu SW. The evolution of Protégé: An environment for knowledge-based systems development. *International Journal of Human-Computer Studies* 2003; **58** (1):89-123.
- [9] Bartkova J, Lukas J, Strauss M, Bartek J. The PRAD-1/cyclin D1 oncogene product accumulates aberrantly in a subset of colorectal carcinomas. *Int J Cancer* 1994; **58** (4):568-573.
- [10] Lovec H, Grzeschiczek A, Kowalski MB, Moroy T. Cyclin D1/bcl-1 cooperates with myc genes in the generation of B-cell lymphoma in transgenic mice. *Embo J* 1994; **13** (15):3487-3495.
- [11] Wang TC, Cardiff RD, Zukerberg L, Lees E, Arnold A, Schmidt EV. Mammary hyperplasia and carcinoma in MMTV-cyclin D1 transgenic mice. *Nature* 1994; **369** (6482):669-671.
- [12] Kranenburg O, van der Eb AJ, Zantema A. Cyclin D1 is an essential mediator of apoptotic neuronal cell death. *Embo J* 1996; **15** (1):46-54.
- [13] Mork P, Halevy A, Tarczy-Hornoch P. A Model for Data Integration Systems of Biomedical Data Applied to Online Genetic Databases. *Proceedings of the Annual AMIA Fall Symposium*, Washington, D.C., 2001. 473-477.
- [14] Shaker R, Mork P, Barclay M, Tarczy-Hornoch P. A Rule Driven Bi-Directional Translation System for Remapping Queries and Result Sets Between a Mediated Schema and Heterogeneous Data Sources. *Proceedings, AMIA Fall Symposium*, San Antonio, TX, 2002. 692-696.
- [15] Noy NF, Musen MA. The PROMPT suite: Interactive tools for ontology merging and mapping. *Int J of Human-Computer Studies* 2003; **59** (6):983-1024.
- [16] Zhang S, Bodenreider O. Aligning representations of anatomy using lexical and structural methods. *Proceedings, AMIA Fall Symposium*, Washington, D.C., 2003. 753-757.
- [17] Bono H, Yagi K, Kasukawa T, et al. Systematic expression profiling of the mouse transcriptome using RIKEN cDNA microarrays. *Genome Res* 2003; **13** (6B):1318-1323.
- [18] Travillian RS, Rosse C, Shapiro LG. An Approach to the Anatomical Correlation of Species through the Foundational Model of Anatomy. *Proceedings, AMIA Fall Symposium*, Washington, D.C., 2003. 669-673.