

Using Equivalence Axioms from the Mammalian Phenotype Ontology to Facilitate Phenotype and Expression Comparisons

Susan M Bello, Terry F. Hayamizu, Cynthia L Smith, Martin Ringwald and The MGI Software Team

Mouse Genome Informatics
The Jackson Laboratory
Bar Harbor, USA

Abstract—Comparisons of expression and phenotypes associated with a gene can enhance the understanding of gene function. These comparisons can be difficult to make due to differences in the ontologies used to annotate the data. Using equivalence axioms in the Mammalian Phenotype (MP) ontology and mappings between Uber-anatomy (UBERON) and mouse developmental anatomy (EMAPA) terms MGI has implemented gene expression + phenotype comparison matrices for genes in MGI. These matrices used the shared anatomical concepts between expression and phenotype ontologies to facilitate comparisons between annotations in these two domains.

Keywords—phenotype; expression; mouse;

I. INTRODUCTION

Visualizing the correlation between gene expression patterns and phenotypes associated with mutations in the same gene enhances the ability of researchers to investigate gene function and identify potential candidate genes for disease. At MGI phenotype annotations use the Mammalian Phenotype ontology (MP) [1] and expression annotations use the developmental mouse anatomy ontology (EMAPA) [2] that incorporates both anatomical structures and the developmental stages when these structures are present. To facilitate comparisons between phenotype and expression, MGI (www.informatics.jax.org) has implemented comparison matrices (Fig. 1) displaying developmental and adult gene expression and genotype specific phenotype annotations. Using the equivalence axioms in the MP to determine the relationship between MP and EMAPA terms, these matrices display the expression and phenotype data for a given gene in the common anatomical framework of the hierarchically structured EMAPA ontology.

II. USE OF EQUIVALENCE AXIOMS

Over the years, the MP and EMAPA curators at MGI have worked with UBERON in establishing anatomical cross-references to enable the comparative analysis of expression and phenotype data across species. The MP to EMAPA mappings described here rely on the usage of the Uber-Anatomy ontology (UBERON) [3] terms in the equivalence axioms of MP terms and the mapping of UBERON terms to

EMAPA terms. For example the MP term ‘abnormal heart morphology’ has an equivalence axiom containing the UBERON term ‘heart’ (UBERON:0000948) which is mapped to EMAPA:16105. While this is a relatively simple example, these mappings also allow for MP terms to be associated with more than one anatomy term. For example the MP term ‘tracheoesophageal fistula’ (MP:0003321) maps to both esophagus (EMAPA:16833) and trachea (EMAPA:16853) as the equivalence axiom includes the UBERON terms for both anatomical structures. This approach also takes advantage of the hierarchical structures of the ontologies. Thus annotations to the MP term ‘abnormal semilunar valve morphology’ are linked to the anatomical term heart via the relation of the EMAPA term ‘semilunar valve’ (EMAPA:35760) to the EMAPA term ‘heart’ (EMAPA:16105).

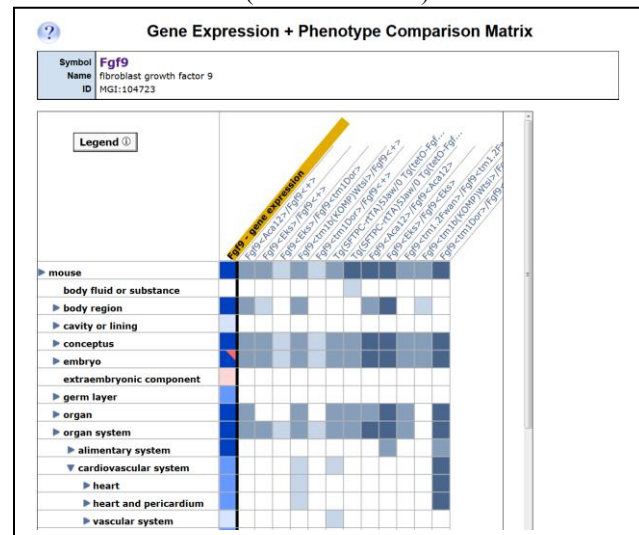


Fig. 1. Gene expression + phenotype comparison matrix for *Fgf9*

There are currently 6149 MP terms containing UBERON terms in the equivalence axiom, out of a total of 9086 MP terms with equivalence axioms (Table 1). Using this approach, annotations to 5067 MP terms can be mapped to the 1369 relevant mouse anatomical structure. There are two primary reasons why an MP term with a logical definition may not

MGI is supported by program project grant HG000330 from the National Human Genome Research Institute (NHGRI). GXD is supported by grant HD062499 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD).

map to a relevant anatomical term. Either (1) the UBERON file lacks the EMAPA mapping or (2) the MP equivalence axiom uses a term from a different ontology, typically the Gene Ontology (GO) or Cell Ontology (CL). Many of the GO terms used incorporate anatomy into the GO term. MP terms using GO terms are primarily anatomy specific development or physiology MP terms, such as ‘delayed caudal neuropore closure’ (MP:0012709) or ‘arrest of tooth development’ (MP:0000118). There are currently 1893 MP terms that use at least one GO term in their equivalence axiom and 1295 MP terms that use at least on CL term. Explicit incorporation of the anatomy terms related to these MP terms, when appropriate, would enhance the mapping. Not all MP terms refer to specific anatomical structures so there may always be some MP terms that do not map to EMAPA terms. For MP terms using GO in the axioms, this will require incorporating the GO equivalence axioms into the MGI mapping process to incorporate the GO axiom that already include a reference to a relevant anatomical term, or the addition of anatomy axioms to the appropriate MP terms when these don’t exist in the GO. The mapping of MP terms to EMAPA terms also supports links directly from MP terms in the Mammalian Phenotype Browser to the relevant EMAPA term in the MGI Anatomy Browser. This allows a user to easily traverse from a phenotype term to genes expressed in the relevant tissues and from an anatomy term to mice with phenotypes in those tissues. As mentioned above, most MP terms link to a single EMAPA term, with a few mapping to 2 or at most 3 UBERON terms. However, for EMAPA terms with at least one mapped MP term, the EMAPA term may map to multiple MP terms. The 6149 MP terms with UBERON in the logical definition use a total of just 1866 UBERON terms. Of these about half are used in in a single equivalence axiom, with a further 500+ used in 2 or 3 axioms. However, almost 400 UBERON terms are used in 4 or more MP term axiom with almost half of these being used in 6 or more axioms. For example the EMAPA term parathyroid gland (EMAPA:32812) is associated with 12 MP terms (Fig. 2).

This reflects the range and variability of phenotypes that many be displayed in a single tissue including both morphological and physiological phenotypes.

III. CONCLUSION

Overall the use of UBERON terms in equivalence axioms may be used to significantly enhance the ability of users to transverse between disparate types of data that share a common anatomical reference point. The new MGI gene expression + phenotype comparison matrices and tissue and phenotype cross-links in browsers represent a practical application of equivalence axioms in the MP and mappings from EMAPA to UBERON. These tools increase the ability of MGI’s users fully utilize the data in MGI.

ACKNOWLEDGMENT

We thank the MGI team and UBERON developers for all their hard work. The MGI Software Group includes; Richard M. Baldarelli, Jonathan S. Beal, Olin Blodgett, Jeffrey W. Campbell, Lori E. Corbani, Sharon C. Giannatto, Peter Frost, David B. Miers, James A. Kadin, Joel E. Richardson.

REFERENCES

- [1] C. L. Smith and J. T. Eppig, “The mammalian phenotype ontology: enabling robust annotation and comparative analysis,” *Wiley Interdiscip. Rev. Syst. Biol. Med.*, vol. 1, no. 3, pp. 390–399, Nov. 2009.
- [2] T. F. Hayamizu, R. A. Baldock, and M. Ringwald, “Mouse anatomy ontologies: enhancements and tools for exploring and integrating biomedical data,” *Mamm. Genome*, vol. 26, no. 9–10, pp. 422–430, Oct. 2015.
- [3] C. J. Mungall, C. Torniai, G. V Gkoutos, S. E. Lewis, and M. A. Haendel, “Uberon, an integrative multi-species anatomy ontology,” *Genome Biol.*, vol. 13, no. 1, p. R5, 2012.

TABLE I. SUMMARY OF MAPPING OUTCOMES

	Count of terms
MP terms with axioms	9086
MP terms with UBERON in axioms	6149
Total number of UBERON terms used in the MP	1866
Total number of MP terms that map to EMAPA	5067
MP terms with GO in axioms	1893
MP terms with CL in axioms	1295
Total number of EMAPA terms that map to MP terms	1369

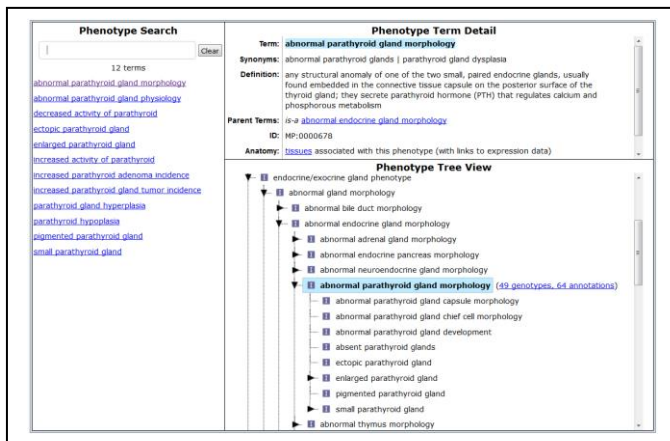


Fig. 2. Display of MP terms mapped to the EMAPA term parathyroid gland in the MP browser.