CARO: The Common Anatomy Reference Ontology

Haendel, M.A., Osumi-Sutherland, D.S. Neuhaus, F.,
Mabee, P.M., Mungall, C.J., Ashburner, M., and Smith, B.

Abstract

The Canonical Anatomy Reference Ontology (CARO) is being developed to facilitate interoperability between existing anatomy ontologies for different species, and will provide a template for building new anatomy ontologies. CARO has a structural axis of classification based on the top-level nodes of the Foundational Model of Anatomy. CARO will complement the developmental process sub-ontology of the GO Biological Process ontology, using it to ensure the coherent treatment of developmental stages, and to provide a common framework for the model organism communities to classify developmental structures. Definitions for the types and relationships are being generated by a consortium of investigators from diverse backgrounds to ensure applicability to all organisms. CARO will support the coordination of cross-species ontologies at all levels of anatomical granularity by cross-referencing types within the cell type ontology (CL) and the Gene Ontology (GO) Cellular Component ontology. A complete cross-species CARO could be utilized in other ontologies for cross-product generation.

1. Necessity of a canonical anatomy reference ontology

Genomes are modified over evolutionary time to produce a diversity of anatomical forms. Understanding the relationship between a genome and its phenotypic outcome requires an integrative approach that synthesizes knowledge derived from the study of biological entities at various levels of granularity, encompassing gene structure and function, development, phylogenetic relationships, and ecology.

Many model organism databases (MODs) collect large amounts of data on the relationship between genetic/genomic variation and morphological phenotypes in databases, which standardize the description of morphological phenotypes and gene expression patterns using types from anatomy ontologies specific to their species of interest. These ontologies have allowed the MODs to group phenotypic and gene expression data pertaining to particular anatomical types.[1] Methods of phenotype curation

---

[1] In keeping with the nomenclature of Smith et al., 2005a, we prefer the term 'type' to 'class'. Ontologies contain terms that refer to types of things in the real world. A type should not be confused with its instances. For example, a human anatomy ontology might contain the term 'foot'. This refers to the type

are being extended and standardized as part of the work of the National Center for Biomedical Ontology (NCBO), which aims to provide data-mining tools which can be applied across all species, in particular in support of queries relating to anatomical structures and associated genes. However, there is currently no system for standardizing the representation of anatomy in ontologies.

Cross-species standardization among anatomy ontologies would bring a number of benefits. First, it would allow the development of standardized tools for grouping and querying anatomy-linked data. Second, it is a prerequisite for inference of anatomically based phenotypic and gene expression data within and across species. Third, if anatomy ontologies were standardized, then a method for representing homology between anatomical types in different anatomy ontologies could be devised. Fourth, standardization would allow better interoperability between anatomy ontologies and other ontologies.

In this chapter, we propose a canonical anatomy reference ontology (CARO), which is designed to serve as a standardized, generic structural classification system for anatomical entities. We also propose a standardized set of relations for use in building anatomy ontologies, extending the set of relations already defined as part of the OBO Relations Ontology (RO; Smith et al., 2005a). By necessity, this proposal also begins to address the key issue of representation of homology between anatomical types in the context of anatomy ontologies.

This chapter summarizes progress on creating CARO, drawing on conclusions reached during an anatomy ontology workshop held in Seattle, WA, in September of 2006[2] sponsored by the National Center for Biomedical Ontology.

2. What is CARO?

CARO is an ontology of canonical anatomy. At its core is a single, structural classification scheme based on that developed by the Foundational Model of Anatomy (FMA), a well established ontology of human anatomy (Rosse and Mejino, 2003). One reason the FMA was chosen as a model for CARO is because the FMA adheres to the principles laid out by the OBO Foundry. CARO has adopted the policy of single inheritance based principally on the empirical observation that ontologies that allow multiple inheritance, while easier to build, are marked by characteristic errors, which generally result from the use of multiple classification schemes within a single ontology, leading to what has been called '*is_a* overloading'. This can be avoided by utilizing *genus-differentia* definitions of the terms in ontologies, in which each type is specified as a refinement (via some *differentia*) of an existing more general type (the *genus*, i.e. the corresponding parent type, in the *is_a* hierarchy). Definitions of this form are typically written along the lines of "An S *is_a* G which D". This provides unambiguous definitions that can be applied consistently and leads to clean classification hierarchies in which all

_____

human foot, of which your left foot is an instance. The collection of all such instances is the *extension* of the corresponding type.

[2] http://www.bioontology.org/wiki/index.php/Anatomy_Ontology_Workshop

types have a single (*is_a*) parent, and all children of a given type are disjoint (so that nothing can be an instance of both a type and its sibling).

CARO provides relations and the definitions for high-level anatomical types for canonical anatomy. A canonical anatomy gives an account of the 'prototypical' composition of the members of a given species.[3] This simplifies the task of constructing anatomy ontologies, as information captured in them, for example pertaining to part and location relationships, can differ radically in non-canonical types. Scientific communities have different perspectives on what constitutes canonical anatomy. Model organism biologists generally have a standard strain or strains that are considered 'wild-type' for their chosen species. Within medicine, canonical anatomy is a generalization deduced from qualitative observations that are implicitly sanctioned by their accepted usage by anatomists (Rosse et al. 1998, Smith et al., 2005b). Defining canonical anatomy is even more problematic in the context of evolutionary biology, where natural variation within a species is often the object of study. Taxonomists therefore utilize voucher or 'type' specimens to define what is representative for a given species.[4] Extensions of CARO to enable integration with the disease ontology (DO) or other ontologies representing pathology or non-canonical anatomy can be accomplished in due course; but such integration will be unfeasible except on the basis of a foundation of canonical anatomy in relation to which relevant deviations can be defined.

CARO includes structural definitions of many generic anatomical types such as cell, tissue, organ and organ system (see Table 1 for a complete list), organized in an *is_a* hierarchy. *Part_of* and other relations between these types will also be represented. CARO thereby provides a standardized reference ontology on which to build species-specific or taxon-specific anatomy ontologies or to reorganize existing ontologies. This can be achieved by using a clone of CARO to create upper-level types for a species or taxon-specific ontology. As part of a species or taxon-specific ontology, the cloned types will refer to anatomical types in the species or taxon in question. Each of these types cloned from CARO will have an *is_a* relationship to the corresponding CARO type, and will inherit from the latter its definition.

The CARO types 'cell' and 'cellular component' are potential root nodes for two existing non-species-specific anatomy ontologies: GO Cell Component and OBO Cell Type. Work is already under way to coordinate definitions and type names that are common to CARO and the latter ontologies, and definitions in all three ontologies will cross-reference each other.

A structural classification alone is not sufficient for the complete representation of anatomy. Other classification systems required for this task include an ontology of functions applicable to anatomical structures and an ontology of qualities such as shape (see Figure 1). Types from ontologies of function and quality can be used in conjunction

---

[3] For a more detailed analysis see chapter 14.
[4] International Commission on Zoological Nomenclature, INTERNATIONAL CODE OF ZOOLOGICAL NOMENCLATURE online, chapter 13: The type concept in
nomenclature, Article 61. Principles of Typification. http://www.iczn.org/iczn/index.jsp

with CARO types to build combined anatomy ontologies for single species with multiple inheritance 'views'. For example, components of the immune system are grouped solely on a functional basis; they are not part of some single structure or group that can be defined in CARO. Some suitable ontologies of functions are already in existence or are planned (GO Molecular Function, Gene Ontology Consortium, 2006; FMP, Cook et al., 2004) however, it may be necessary to supplement these ontologies with others still to be created.

Anatomical types classified under CARO can also be linked to types representing biological processes in which they participate, such as those found in the Gene Ontology Biological Process Ontology (GO BP) or in developmental stage ontologies (see Representing Development, later in this chapter). The formalism for combining definitions of types from different parent ontologies in a definition follows the *genus and differentia* methodology described earlier.

RELATION TO TIME / GRANULARITY	CONTINUANT				OCCURRENT
	INDEPENDENT		DEPENDENT		
ORGAN AND ORGANISM	**Organism** (NCBI Taxonomy)	**Anatomical Entity** (CARO)	**Organ Function** (FMP)	**Phenotypic Quality** (PaTO)	**Biological Process** (GO)
CELL AND CELLULAR COMPONENT	**Cell** (CL)	**Cellular Component** (GO)	**Cellular Function** (GO)		
MOLECULE	**Molecule** (ChEBI, SO, RnaO, PrO)		**Molecular Function** (GO)		Molecular Process (GO)

Figure 1. Coverage of species-independent ontologies relevant to biology.

CARO is an ontology of independent anatomical continuants. Continuants have a continuous existence through time. Dependent continuant entities are things that inhere in independent continuant entities, such as qualities, shapes, roles, functions. Occurrents (processes) have temporal parts which unfold in time (every occurrent depends on one or more independent continuants as its participant or bearer). The ontology prefixes shown in parentheses are either under development (FMP, RnaO, PrO) or are available at OBO: http://obo.sourceforge.net/browse.html

3. CARO structure and definitions

The first version of CARO is under active development, and currently there are approximately 50 types. A CARO listserve and wiki track discussion of the ontology and related subjects. A CARO pre-version file can be accessed on the wiki in obo format. The wiki is at: http://bioontology.org/wiki/index.php/CARO:Main_Page

CARO is an ontology of independent anatomical continuants. Continuants have a continuous existence through time. Dependent continuant entities are things that inhere in independent continuant entities, such as qualities, shapes, roles, functions. Occurrents (processes) have temporal parts which unfold in time (every occurrent depends on one or more independent continuants as its participant or bearer). The ontology prefixes shown in parentheses are either under development (FMP, RnaO, PrO) or are available at OBO: http://obo.sourceforge.net/browse.html

The first version of CARO is under active development, and currently there are approximately 44 types. A CARO listserve and wiki are being used for discussion of the ontology and related subjects. A CARO pre-version file can be accessed on the wiki in obo format. The wiki is at: http://bioontology.org/wiki/index.php/CARO:Main_Page

The CARO types and definitions are based on the topmost nodes of the FMA (Rosse and Mejino, 2003). The top levels of the FMA provide a rich set of abstract structural classifications that take into account qualities such as dimensions and contiguity, and cover many levels of granularity – from whole organism down to cell parts. All of these characteristics have made the FMA an ideal starting point for CARO. However, many of the FMA type definitions are not applicable to all species; some are mammal-specific, some are human-specific, and some are specific to only adult human. The definitions of these types have been generalized in CARO to be inclusive of more species. Organismal domain specialists will be required to validate the CARO types, in much the same way that human anatomists were required to build and validate the FMA. In addition, the FMA is incomplete in its treatment of developmental structures and developmental relations. As the representation of developmental anatomy in ontologies is central to the functioning of multiple model organism databases, we have begun to extend the CARO classification scheme to fill this gap. Figure 2 shows the relations between types in CARO, and Table 1 the definitions for these types.
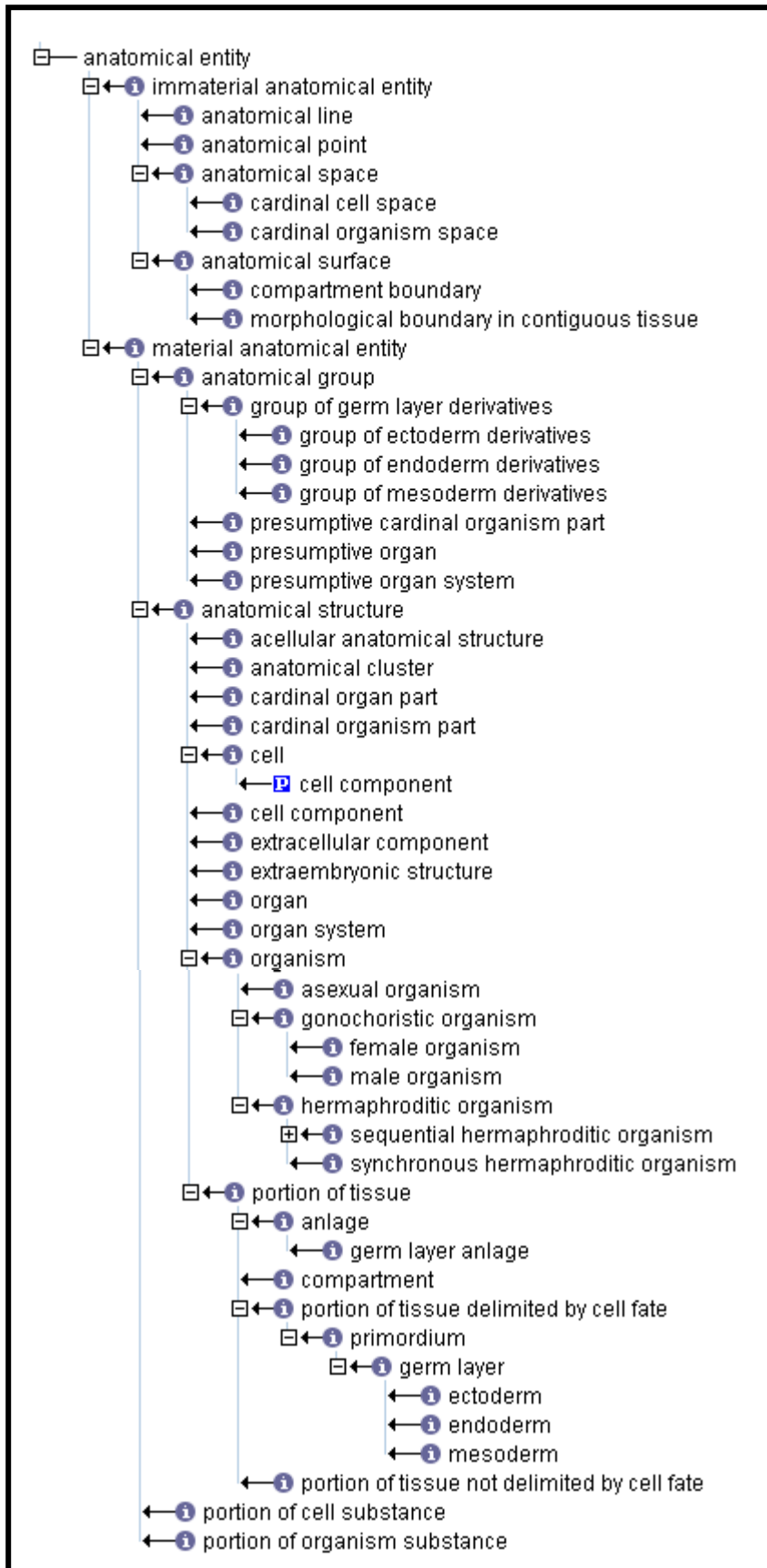
```
□── anatomical entity
    □←❶ immaterial anatomical entity
        ←❶ anatomical line
        ←❶ anatomical point
        □←❶ anatomical space
            ←❶ cardinal cell space
            ←❶ cardinal organism space
        □←❶ anatomical surface
            ←❶ compartment boundary
            ←❶ morphological boundary in contiguous tissue
    □←❶ material anatomical entity
        □←❶ anatomical group
            □←❶ group of germ layer derivatives
                ←❶ group of ectoderm derivatives
                ←❶ group of endoderm derivatives
                ←❶ group of mesoderm derivatives
            ←❶ presumptive cardinal organism part
            ←❶ presumptive organ
            ←❶ presumptive organ system
        □←❶ anatomical structure
            ←❶ acellular anatomical structure
            ←❶ anatomical cluster
            ←❶ cardinal organ part
            ←❶ cardinal organism part
            □←❶ cell
                ←Ⓟ cell component
            ←❶ cell component
            ←❶ extracellular component
            ←❶ extraembryonic structure
            ←❶ organ
            ←❶ organ system
            □←❶ organism
                ←❶ asexual organism
                □←❶ gonochoristic organism
                    ←❶ female organism
                    ←❶ male organism
                □←❶ hermaphroditic organism
                    ⊞←❶ sequential hermaphroditic organism
                    ←❶ synchronous hermaphroditic organism
            □←❶ portion of tissue
                □←❶ anlage
                    ←❶ germ layer anlage
                ←❶ compartment
                □←❶ portion of tissue delimited by cell fate
                    □←❶ primordium
                        □←❶ germ layer
                            ←❶ ectoderm
                            ←❶ endoderm
                            ←❶ mesoderm
                ←❶ portion of tissue not delimited by cell fate
        ←❶ portion of cell substance
        ←❶ portion of organism substance
```

Figure 2. Relations between CARO types. (i) represents *is_a*, and (p) *part_of* as defined in the RO.

Table 1. CARO Definitions

anatomical entity	Biological entity which constitutes the structural organization of a biological organism or is an attribute of that organization.
immaterial anatomical entity	Anatomical entity which is a three-dimensional space, surface, line or point associated with a material physical anatomical entity.
material anatomical entity	Anatomical entity which has mass.
anatomical structure	Material anatomical entity which has inherent 3D shape and is generated by coordinated expression of the organism's own genome.
anatomical group	Material anatomical entity which has as its members the maximal number of anatomical structures of the same type.
portion of organism substance	Material anatomical entity in a gaseous, liquid, semisolid or solid state, with or without the admixture of cells and biological macromolecules; produced by anatomical structures or derived from inhaled and ingested substances that have been modified by anatomical structures as they pass through the body.
portion of cell substance	Portion of organism substance located within a cell.
anatomical cluster	Anatomical structure which has as its parts a heterogeneous collection of organs, organ parts, cells, or organism part subdivisions that are adjacent to, or continuous with, one another; does not constitute an organ, organ system, organ system subdivision, organism part or organism part subdivision.
anatomical point	Immaterial anatomical entity of zero dimension, which forms a boundary of an anatomical line or surface.
anatomical line	Immaterial anatomical entity of one dimension, which forms a boundary of an anatomical surface or is a modulation of an anatomical surface.
anatomical surface	Immaterial anatomical entity of two dimensions, that is demarcated by anatomical lines or points
anatomical space	Immaterial anatomical entity of three dimensions, which is generated by morphogenetic or other physiologic processes; is surrounded by one or more anatomical structures; contains one or more organism substances or anatomical structures.
cell space	Anatomical space which is part of a cell.
organism	Anatomical structure which is an individual member of a species.
asexual organism	Organism that does not produce gametes.
gonochoristic organism	Organism that has male and female sexes.
female organism	Organism that can produce female gametes.
male organism	Organism that can produce male gametes.
hermaphroditic organism	Organism that can produce both male and female gametes.
sequential hermaphroditic organism	Organism that produces gametes first of one sex, and then later of the other sex.
synchronous hermaphroditic organism	Organism that produces both male and female gametes at the same time.
cell	Anatomical structure which has as its direct parts a maximally connected cell compartment surrounded by a plasma membrane.
epithelial cell	Cell which has as its part a cytoskeleton that allows for tight cell to cell contact and for apical-basal cell polarity where the basal part is directed towards a basal lamina.
cell component	Anatomical structure which is part of the cell.
acellular anatomical structure	Anatomical structure which consists of cell parts and cell substances and together does not constitute a cell or a tissue.
basal lamina	Acellular anatomical structure which consists of a thin sheet of fibrous proteins that underlie and support the cells of an epithelium. It separates the cells of an epithelium from any underlying tissue.
portion of tissue	Anatomical structure which has as its parts cells of one or more types spatially arranged in a characteristic pattern.
epithelium	Portion of tissue, which consists of one or more layers of epithelial cells connected to each other by cell junctions and which is underlain by a basal lamina.
unilaminar epithelium	Epithelium which consists of a single layer of epithelial cells.
multilaminar epithelium	Epithelium which consists of more than one layer of epithelial cells that may or may not be in contact with a basement membrane.
atypical epithelium	Epithelium the cells of which do not conform to the typical arrangement of unilaminar and multilaminar epithelium.
simple squamous epithelium	Unilaminar epithelium which consists of a single layer of squamous cells.

simple columnar epithlium	Unilaminar epithelium which consists of a single layer of columnar cells.
simple cuboidal epithelium	Unilaminar epithelium which consists of a single layer of cuboidal cells.
multi-tissue anatomical structure	Anatomical structure which has as its parts two or more portion of tissue types which form a distinct morphological aggregate in an organ; the boundary of which may be fiat or bona fide or both.
organ	Anatomical structure which has as its parts a maximal group of multi-tissue anatomical structures, which together constitute the entire organ.
solid organ	Organ the unshared parts of which do not surround macroscopic anatomical spaces; only its shared parts (subdivisions of hollow tree organs) contain anatomical spaces.
parenchymatous organ	Solid organ which consists of parenchyma and connective tissue stroma; the stroma subdivides the parenchyma into lobes, segments, lobules, acini, or cortex and medulla.
nonparenchymatous organ	Solid organ which consists of organ parts that are arranged as fascicles or sheets.
cavitated organ	Organ the unshared parts of which surround one or more macroscopic anatomical spaces.
organ with organ cavity	Cavitated organ in which its unshared parts surround one continuous anatomical space.
organ with cavitated organ parts	Cavitated organ in which its unshared parts surround two or more macroscopic anatomical spaces.
organ system	Anatomical structure which has as its direct parts instances of one or more organ types which are interconnected with one another by zones of continuity.
multi-organ anatomical structure	Anatomical structure which has as its parts a maximal group of organ and multi-tissue anatomical structures, portions of tissue, or cells of diverse types, which together constitute the entire organism. Forms a distinct morphological subdivision of the organism.
extraembryonic structure	Anatomical structure which is contiguous with the embryo and is comprised of portions of tissue or cells that will not contribute to the embryo.
morphological boundary in contiguous tissue	A discontinuity in a contiguous tissue due to a change in shape and/or distribution of its component cells.
<u>portion of tissue delimited by cell fate</u> delimited	Portion of tissue whose boundaries delimit the majority of the precursors of some specific later type or types or anatomical structure.
<u>portion of tissue not delimited by cell fate</u>	Portion of tissue whose boundaries do not delimit the majority of the precursors of some specific later type or types of anatomical structure. DO WE REALLY NEED THIS – NORMALLY USE OF 'NOT' IN THIS WAY IS BAD PRACTICE
<u>anlage </u>(field)	Portion of tissue, not delimited from contiguous tissue by a morphological or compartment boundary, whose boundaries delimit the majority of the precursors of some specific later structure or structures
primordium (primitive structure)	Portion of tissue whose boundaries delimit the majority of the precursors of some specific later structure or structures.
compartment	Portion of tissue which is delimited from contiguous portions of tissue by a compartment boundary.
compartment boundary	Anatomical surface which inhibits mixing of cells between portions of tissue with different lineages.
germ layer	Primordium formed as one of the earliest subdivisions of an embryo, and whose boundaries delimit the majority of the precursors of a high proportion of all later structures.
endoderm	Germ layer which is the innermost layer of the gastrulating embryo.
ectoderm	Germ layer which is the outermost layer of the gastrulating embryo.
mesoderm	Germ layer which is the middle layer of the gastrulating embryo.
group of germ-layer derivatives	Anatomical group which develops from a germ-layer.
group of ectoderm derivatives	Anatomical group which develops from the ectoderm.
group of mesoderm derivatives	Anatomical group which develops from the mesoderm.
group of endoderm derivatives	Anatomical group which develops from the endoderm.

Definitions which have been modified from those used by the FMA for use in CARO are discussed below.

Representing granularity
In order to represent different levels of granularity in CARO, the appropriate types must be specified in such a way as to be applicable across all taxa. The FMA has a well developed system for classifying structural types according to a hierarchy of granularity. Each level of the hierarchy defines the basic building blocks for the level above. Tissues, for example, are defined as aggregates of cells. However, because the FMA applies only to human anatomy, the FMA developers have used both this bottom up definition of structural types along with a vernacular top down naming system: an organ part is made up of multiple tissues; cardinal body part is made up of multiple organs. We have renamed some of these types to reflect their bottom up definition so that they can apply to a wide range of organisms. *Cardinal body part* is renamed, *multi-organ anatomical structure*, and *cardinal organ part* is renamed *multi-tissue anatomical structure*.

Portion of tissue: The term 'tissue' is used sometimes as a mass noun (compare: 'luggage', 'sugar') in such a way as to refer ambiguously to indeterminate amounts of cellular material. We prefer 'portion of tissue' (a count noun analogous to 'suitcase' or 'sugar-lump') to make it clear that the term refers unambiguously to a single discrete structure. In addition, we have altered the definition to make 'a characteristic pattern of specialized cells' one of the defining features of tissue, rather than 'similarly specialized cells' as we believe this to be more inclusive of different taxa and of developing structures. 'Characteristic' is used to signify that each type of portion of tissue is marked by a distinctive pattern of organization of cells of distinctive types.

Cardinal parts: Use of the word 'cardinal' implies a major subdivision of some anatomical structure. For example, all organs have mitochondria as parts, but this is not relevant when discussing the major regions of an organ. 'Cardinal' is used to mean *at the greatest level of granularity* as applied to entities of the pertinent type. We have used 'cardinal organism part', rather than 'cardinal body part', because 'body' means different things to different communities.

Cross-ontology coordination of CARO types
A number of types in CARO are present in other ontologies, such as the Gene Ontology Cellular Component (GO CC), and the Cell Type ontology (CL). Specifically, these types represent integration of different levels of anatomical granularity. Coordination of definitions between the GO CC, the CL, and CARO ontologies has begun, and these types will be linked via cross-references.

Organ: The definition of 'cardinal organ part' in the FMA specifies that there be portions of two or more types of tissue as its parts (Rosse and Mejino, 2003). However, this definition is not consistent with the use of the term 'organ' in other organisms' non-human anatomy. For example, the neuromast organ of the zebrafish consists of cells of a number of different cell types, but not tissues. Similarly, the fly fat body is thought of as an organ but consists of cells of only one tissue type. We have attempted to define organ purely structurally, on the basis of the types of boundaries it has. For this definition to work, we will almost certainly need to find a way to add maturity as a defining factor. The current definition should be considered provisional.

acellular anatomical structure: This type is synonymous with GO:0044421 *extracellular region part*.

Cell, intracellular component, and extracellular component: These types integrate different levels of anatomical granularity. The definitions are being coordinated between the GO CC, the CL, and CARO ontologies via cross-references.

cell: This type is synonymous with GO:0005623 *cell* and CL:0000000 *cell*.

The organism types: We include the whole organism as an anatomical structure to allow the formulation of part relations of sexually dimorphic anatomical structures. For example, humans have as parts gonads, but only male humans have testes. Different life strategies for reproduction have different corresponding anatomical structures, requiring that these organism types be defined in CARO.
epithelial cell: This type is synonymous with CL:0000066 *epithelial cell.*

4. Classification of developing structures

The process of development involves the gradual division of contiguous tissues into regions (portions) of tissue that follow different paths of differentiation. To reflect this, we define developing anatomical structures as subtypes of the type 'portion of tissue' (see Figure 3).[5] However, in doing so, we have classified these developmental types in such a way as to include non-structural differentia. This violates our directive to maintain a purely structural single inheritance. Because of the need to represent these entities and to begin the investigation as to how best to undertake this task, we have included them
cell component: This type is synonymous with GO:0044464 *cell part.*

basal lamina: This type is synonymous with GO:0005605 *basal lamina.*

In addition to these structural criteria for classification, we need to take into account the use of cell fate by developmental biologists as a criterion to define and name developing anatomical structures. Cell fate refers to the likely destination of particular groups of cells (or syncytial nuclei) under conditions of normal development. Developmental biologists traditionally define and name portions of tissue, at least in part, on the basis of restricted cell fate: lens placode, limb field, limb bud, fat-body primordium, and so on. The boundaries of these regions delimit the majority[6] of the precursors of some specific type or types of anatomical structure.

Developing portions of tissue

We distinguish two major subtypes of portion of tissue using a differentia based on cell fate: 'cell fate delimited portion of tissue' and 'portion of tissue not delimited by cell fate'. The former are portions of tissue whose boundaries delimit the majority of the

---

[5] The differentia we have chosen for subtypes of portion of tissue are based on suggestions from Volker Hartenstein (personal communication).
[6] Incomplete determination coupled with cell mixing at the boundaries of primordia often means that not all cells in a primordium will share the fate of the majority, and that cells from adjacent primordia do share that fate (Slack, 1991).

precursors of some specific later structure or group of structures. The latter either have either no restricted fate (as for some anatomical structures in some early vertebrate embryos; Slack, 1991) or have boundaries that cut across those delimiting the majority of the precursors of some specific later structure or group of structures. This type also encompasses adult portions of tissue.

The organism types

We include the whole organism as an anatomical structure to allow the formulation of part relations of sexually dimorphic anatomical structures. For example, humans have as parts gonads, but only male humans have testes. Different life strategies for reproduction have different corresponding anatomical structures, requiring that these organism types be defined in CARO.

We define 3 disjoint subtypes of 'cell fate delimited potion of tissue'. These are anlage, primordium and compartment. We define such a portion of tissue as a primordium if its boundaries with contiguous tissue are marked by a discontinuity in morphology – a difference in cell shape for example. If such a portion of tissue lacks morphologically distinct boundaries, we refer to it as an anlage (see Table 1 for complete definitions). This definition encompasses portions of tissue with bona-fide boundaries defined on the basis of non-morphlogical criteria such as gene expression. It also encompasses portions of tissue with fiat boundaries (Smith 2001) defined solely on the basis of cell fate.

5. Developing structure types

Although the favored terms for developing structures vary between different communities of developmental biologists ('rudiment', 'field', 'placode', 'bud', etc), we believe that our differentiation of primordia and anlagen on the basis of morphology reflects a very commonly made distinction. These high level nodes can be used to classify different embryonic structures while retaining terminology that is familiar to particular communities.

In CARO, analge and primordium are two of three disjoint sibling types. The third sibling of this group is 'compartment'. Within most contiguous portions of developing tissue, newly born cells can mix freely. But some are bisected by boundaries that newly born cells do not cross. Rather than being active barriers to cells, these boundaries are thought to be the result of differentiation of cells into groups that do not mix with each other (Dahmann and Basler, 1999). These divisions of contiguous developing tissues, known as compartments, are widely distributed throughout the animal kingdom, and are well documented for example in insects and vertebrates. The cells adjacent to compartment boundaries are often sites of expression of genes with key roles in pattern specification during development (Dahmann and Basler, 1999; Matsuoka et al., 2005). We define as a compartment, any portion of tissue which is delimited fully or in part by a compartment boundary and whose boundaries delimit the majority of the precursors of some specific later structure or group of structures.Prior to extensive morphogenesis and differentiation, most developing structures are sufficiently simple that they can be defined as a subtype of

the CARO term 'portion of tissue'. In some cases, types originally defined for adult structures are clearly applicable to developing structures. For example, the regions of the imaginal discs of Drosophila that will develop into adult appendages are consistent with the definition of columnar epithelium. One of the earliest divisions of all embryos of triploblastic species into portions of tissue, is the development of germ-layers. These divide the early embryo into portions of tissue, each with a distinct fate. The majority of cells that will form epidermis and neural tisssues, for example, can be found in the ectoderm. We treat germ-layers as we would any other portion of tissue with a cell fate delimited potion of tissue fate. We use the term germ-layer anlage (a child of anlage) to refer to germ-layers at stages prior to them becoming morphologically distinct. Once they have become morphologically distinct, we use the term germ-layer (a child of primordium).

Other developing tissues share many, although not all of the qualities of mature tissues. For example, many tissues of the early Drosophila embryo fit the definition of epithelia except that they lack a basal lamina. The number of generic structural types that are applicable to developing tissues will be expanded in future versions of CARO.

Grouping developing anatomical structures

In order to represent development at different levels of granularity, we group individual primordia and anlagen in CARO according to the type of anatomical entities they develop into: 'presumptive organ system', 'presumptive organ' and 'presumptive cardinal body part'. These types are not structurally contiguous and so cannot be classified as anatomical structures. As collections of distinct parts, they are subtypes of 'anatomical group'.

Our system also allows the gradual increases in granularity occurring during development to be captured in a consistent fashion. As development proceeds, developing structures of different granularity levels are formed. As they do, such structures can be reclassified from portions of tissue to organs to multi-tissue aggregates, etc. Terms for individual germ-layers are widely used within biology to refer both to the primary divisions of the embryo and to refer collectively to the group of anatomical structures in an embryo which are derived from a specific germ-layer. To reflect this usage, we have added terms to CARO referring to group of germ-layer derivatives, e.g. 'group of mesoderm derivatives'. This use of 'group' allows us to group the relevant anatomical structures using part relations, so avoiding multiple inheritance. As for 'developing organ system', these terms are subtypes of 'anatomical group'.

Use of structurally classified developmental types to curate gene expression and phenotypic data will make it possible to look for genes common to the development and maintenance of particular structural types and to the transitions from one structural type to another.

These generic structural types will provide a basic structural classification of developing structures. However, many important details of structural types specific to a single species or taxonomic group will need to be captured in the relevant leaf nodes of species-

specific anatomy ontologies. These details could be formalized by referencing structural qualities specified in the Phenotype Attribute and Trait Ontology (PATO)[7].

The limits of structural classification

Structural classification can only go so far in capturing the schemes defining and classifying portions of developing tissue division which developmental biologists consider important. CARO cannot provide these, but we think it important to specify how this might best be achieved in application ontologies built using CARO as a template.

Structures defined by shared cell fate:

Developmental biologists traditionally define and name portions of tissue, at least in part, on the basis of some shared fate: lens placode, limb field, limb bud, fat-body primordium, and so on. The boundaries of these regions delimit groups of cells that are precursors of some specific type or types of anatomical structure. For example, each of the pair of heart primordia in a Zebrafish embryo consists of all the members of a connected group of heart precursor cells.

Two examples:

  P = Heart Tube
  Q = Organ with organ cavity
  R = Heart (a subtype of 'organ with cavitated organ parts')

  P = Heart Anlage
  Q = Portion of Unilaminar Epithelium
  R = Heart

Other possible examples:

  P = wing pouch
  Q = portion of columnar epithelium
  R= wing

  P = fat body / gonad primordium
  Q = portion of tissue
  R = larval fat body and larval gonadal sheath muscle

This differentia distinguishes these tissues from other developing structures that do not correspond to all members of a connected group of cells sharing some fate. Hensen's node in the chicken embryo, for example, contains different precursors at different stages

---

[7] www.bioontology.org/wiki/index.php/PATO:Main_Page

of gastrulation, and does not delimit a connected group of cells sharing some particular fate (Selleck and Stern 1991).

Where applicable, this formalism can be used in leaf node definitions in species specific application ontologies. This is especially useful in cases where such domains are not yet mophologically distinct, but have been experimentally defined. It also provides a way to define germ-layers, mesoderm, ectoderm and endoderm - according to the classes of mature structure whose precursor cells they contain. Finally, as mature structures are named in these definitions, it is possible to use this information to group developing structures according to what they will develop into.

Compartments

Within most portions of developing tissue, newly born cells can mix freely. However, some such portions of tissue are split into two morphologically similar domains, each of which expresses some cell segregation mechanism that prevents cells from one domain from mixing with those in the other. The result is a smooth boundary between the two domains, like the interface between oil and water. These divisions of developing tissues, known as compartments (hereafter referred to as developmental compartment), occur throughout the animal kingdom (Dahmann and Basler, 1999). Cells adjacent to their boundaries often sites of expression of genes with key roles in pattern specification during development (Dahmann and Basler, 1999; Matsuoka et al., 2005).

Developmental compartments are clearly important divisions of developing tissues that need to be represented in anatomy ontologies, but as they are not structurally distinct domains they cannot be defined in CARO.   Instead they could be represented as a genus and differentia cross-product of an appropriate portion of tissue term with an appropriate function term.

5. Relations in CARO

An ontology is a controlled vocabulary that encapsulates the meanings of its terms in a computer parsable form. An anatomy ontology consists of statements composed of two kind of terms, denoting types and relations, respectively. Typically such statements involve two type terms A and B, so that they are of the form 'A *rel* B'. Relations commonly encountered in anatomical ontologies include the *is_a* relation, indicating that one type is a subtype of another, and the *part_of* relation. Examples of use include *Pancreas is_a Lobular organ* in the FMA and C*ell nucleus part_of Cell* in the GO Cellular Component ontology. However, anatomical ontologies are by no means limited to these two relations; the FMA employs a large number of spatial relations (Rosse and Mejino, 2003)[8] , and ontologies that encompass entities at various developmental stages typically link types using relations such as *develops_from*, for example the OBO Cell (CL) ontology and anatomical ontologies for model organisms such as fly and zebrafish.

---

[8] Also see 'spatial association relationship' at http://fme.biostr.washington.edu:8089/FME/index.html

Relations play an essential role in ontologies, since they are the primary bearer of semantic content (see chapter 14). To ensure a consistent use of terms that denote relationships within and across ontologies, it is important to agree on shared, unambiguous definitions of these terms. These definitions utilize the dependence of relationships between types (e.g. C*ell nucleus* and *Cell)* on the relationships between instances of these types (e.g. concrete cell nuclei and the cells which contain them), as is discussed in detail in the chapters 14 and 15 of this book. In this section, we will discuss the extension of the OBO Relations Ontology (RO; Smith et al., 2005) to provide relations that are necessary for CARO and species-specific anatomies. This extension comes in different flavors: (a) in some cases, we need to add new relations to capture important aspects of anatomical entities, (b) in other cases, we need to add new relations that further specify existing ones in order to better represent the dynamic changes within developing organisms, and (c) we need to consider relations that link anatomy ontologies to other ontologies.

Defining develops_from

The RO covers the most important relationships for anatomy ontologies, but lacks explicit definitions of many spatial relations that it would be desirable to include. Some of these are discussed in chapter 15 of this book. Further, for CARO to provide a representation of developmental anatomy, we need to define a relationship that represents the various ways that anatomical structures change through development. We lack a single, transitive relationship that can represent the transformation, fission and fusion of developing structures over time. Here we outline the relationship *develops_from*, which fulfills these criteria. In order to define *develops_from* we need to distinguish two cases. In the first case, some entity changes its properties but remains numerically identical; for example, if an adult develops from a child, then the adult will have different properties (e.g. a different weight and height) but it will be still the same individual. In contrast, if a zygote develops from a sperm cell and an ovum, then the zygote is not identical with either; but the zygote arises from the cell and the ovum. These two relations are used to define the type level relationships *transformation_of* and *derives_from*[9] in RO. Since it is often unknown during development whether one structure is a transformation of another or whether some portion of a structure arises from another one, we need a *develops_from* relation which covers both cases.

More formally, the develops_from relationship is defined as follows:
C *develops_from* D if and only if:
for any x and any time t, if x *instantiates* C *at* time t, then EITHER

> (i), for some time t', x *instantiates* D *at* t' and t' **precedes** t, and there is no time interval t'' such that x *instantiates* C *at* t'' and x *instantiates* D *at* t''; OR
> (ii)   for some time t', there is some y such that y *instantiates* D *at* t', and x **arises_from** y.

---

[9] To avoid confusion with the very different meaning of 'derives from' in an evolutionary context, we plan to rename this relationship '*arises_from'*. It is referred to as such in the following text.

While *develops_from* is a relationship between types, **precedes, succeeds, buds_from, and arises_from** hold between instances.

x **succeeds** y if and only if
>   (i) x and y are anatomical entities; and
>   (ii) x begins to exist at the same instant of time at which y ceases to exist; and
>   (iii) there is some anatomical structure z such that z is part of y when y ceases to exist and z is part of x when x begins to exist.

x **buds_from** y if and only if
>   (i) x and y are anatomical entities; and
>   (ii) there is some anatomical structure z such that z is part of y immediately before x begins to exist, and x **succeeds** z;
>   (iii) x continues to exist for some interval of time from the point when y begins to exist.

The relationship **arises_from** is the transitive closure of **arises_from:**

x **arises_from** y is defined recursively in the following way:
>   (i) if x **succeeds** y, then x **arises_from** y;
>   (ii) if x **buds_from** y, then x **arises_from** y;
>   (iii) if x **arises_from** y and y **arises_from** z, then x **arises_from** z;
>   (iv) x **arises_from** y holds only if (i), (ii) or (iii) holds.

The underlying relationship **arises_from** is also a relationship between instances; it can be defined in the following way:

x **arises_from** y if and only if
>   (i) x begins to exist at the same instant of time at which y ceases to exist;   and
>   (ii) there is some anatomical structure z such that z is part of right before y ceases to exist and z is part of x when x begins to exist.

Time-restricted part relationships

The parthood relations as defined in the RO do not adequately represent some dynamic aspects of developmental anatomy. In particular, the RO relationships *has_part* and *part_of*, both apply at all stages: X *has_part* Y means that every X, regardless of stage, have some Y as instance-level part. The Drosophila anatomy ontology, however, contains types of neuroblasts that are part of the ventral nerve cord primordium (VNC). As these neuroblasts divide, more types become identifiable – at stage 9 there are 10 types but by stage 11 there are 34 (Berger et al., 2001). We cannot capture the part relationship between these cell types and the VNC primordium using the *has_part* relation, because this would imply that all instances of the VNC have instances of each of these neuroblast types as a part at all stages. Similarly, the relation *part_of* also applies irrespective of stage. We can solve this dilemma by defining versions of *part_of* and

has_part applicable which are applicable only during the stages in which both partners in the relationship exist. The formal definitions of these relationships are:

C *time_restricted_part_of* D if and only if
for any x and any time t, if x *instantiates* C *at* time t, then there is a y such that (for some time t', y *instantiates* D *at* t' and x **part_of** y **at** t') AND (for all times t": if x **exists_at** t" and y **exists_at** t", then x is **part_of** y **at** t"))

C *time_restricted_has_part* D if and only if
for any x and any time t, if x *instantiates* C *at* time t, then there is a y such that (for some time t', y *instantiates* D *at* t' and y **part_of** x **at** t') AND (for all times t": if x **exists_at** t" and y **exists_at** t", then y is **part_of** x **at** t"))

Relationships linking separate ontologies

As mentioned above, the structural classification of anatomical entities in CARO is separate from the treatment of functional classification and of homology between anatomical entities across different species. In order to record function and homology information, the anatomical types within a species-specific anatomy ontology need to be linked to types in other ontologies, and the necessary relations – including *has_function* and *homologous_to* – will be added to the RO in due course. We discuss relations between developmental stage and anatomical types in the following section. Note that the spatial relations and the *develops_from* relation mentioned above are relations that are used within a given anatomical ontology; in contrast relations such as *has_function, homologous_to, starts_during* and *ends_during* are relationships that link types across different ontologies. *Is_a*, too, can link types across different ontologies, as for instance when we make an assertion to the effect that *Mouse organ is_a CARO:organ.*

6. Representing Stage

Development can be considered a process that *has_participant* (Smith et al., 2005a) whole organism. For any one species, events during development occur in a predictable order. However, the precise timing of these events is dependent on environmental conditions. Developmental biologists traditionally measure progress through (the process of) development relative to the occurrence of some standard series of events which can be easily and reliably scored (e.g. Campos Ortega and Hartenstein 1999; Nieuwkoop and Faber, 1994). A standard table of development divides development into stages, each delimited by a pair of events, and describes key events occurring within each stage.

For some organisms, not only is the order of events is consistent, but under standard laboratory conditions their timing relative to a reference event (e.g. conception) shows little variation. In these cases it is possible to define stages in terms of the period of time that elapsed since the reference event. This method of defining stages is particularly useful if no easily score-able morphological stage criteria are available. For example, in the zebrafish, early stages are often referred to either by morphological criteria or by time since fertilization, while the later stages are referred to exclusively by time since fertilization (Kimmel et al, 1995).

As stage series are necessarily species specific, ontologies representing individual stage series have to be constructed for each species. Minimally, a stage ontology will contain types for the stages that make up a standard table of development. The relative timing of these stages can be recorded using the relation *preceded_by* (Smith et al, 2005a). Stages can be grouped together into super-stages, or divided into sub-stages, with the latter having a *part_of* relationship to the stages themselves, which are in turn *part_of* super-stages. While stage series are species specific, many of the developmental processes described in standard tables of development are not. Information about the relative timing of developmental processes described in each standard table of development can be captured within species-specific stage ontologies. The relative timing of these processes to each other and to stage boundaries can be recorded using the relations *part_of*, *preceded_by* and an additional relationship *simultaneous_with*.[10] Linking these to relevant GO types such as cellularization (see Figure 3) will facilitate reasoning between species-specific stage ontologies.

We propose that these species-specific stage ontologies be used to record the periods of development during which anatomical structures exist by using the relationships *starts_during* and *ends_during* (a formalized version of the strategy used by ZFIN). These relationships link anatomy ontology types to appropriate types in the stage ontology. This will give a crude resolution to records of timing: the existence of X begins some time during stage N and ends some time during stage N′. The temporal resolution of these links could be improved, as data allows, in two ways. Where some standard system of substages has been defined, we can simply make *starts_during* and *ends_during* links to these substages. Alternatively, we can refine our record of the timing of the beginning or end of existence of an anatomical entity by instantiating these as events within the stage ontology and using *preceded_by* relations to processes beginning or ending within a stage (see Figure 3).
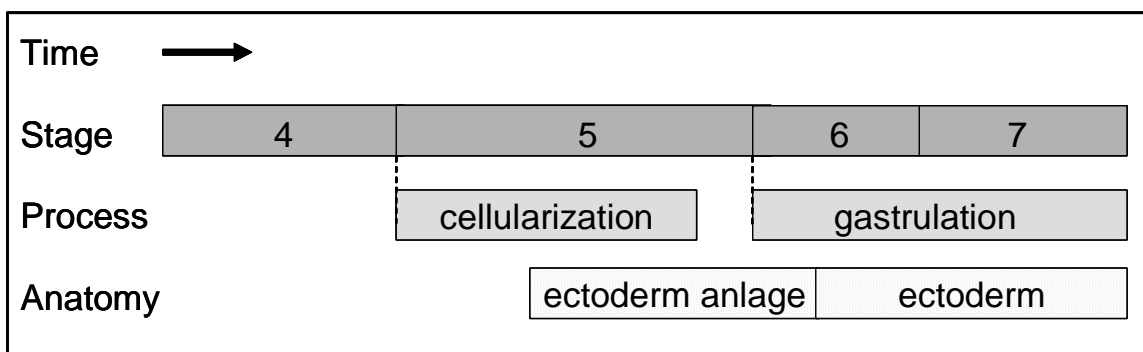

Figure 3. Relationship between anatomical entities, stage, and process.

Both stage and process   For each species, an ontology will be constructed containing types for stage and developmental process in a single ontology of occurents. Anatomical entities are contained in a separate ontology of continuants. The ends of each box represent events for which relative timing can be recorded using the relations *preceded_by* and *simultaneous_with*. These ordering relations will be used in conjunction

---

[10] To be defined in a future publication.

with *starts_during* and *ends_during* to define the period during which an anatomical entity exists. This example illustrates ectoderm development in the Drosophila embryo, wherein the ectoderm anlage *starts_during* stage 5, the ectoderm anlage *ends_during* stage 6, the ectoderm *starts_during* stage 6, the process gastrulation *preceded_by* cellularization, and gastrulation *simultaneous_with* stage 6 and stage 7.

7. CARO depth and application

The question of CARO depth is closely related to its utility in building new anatomy ontologies. The top-level types in CARO together with the relationships defined above can be used to structure application anatomy ontologies. However, the types in CARO are very generic relative to the types commonly defined within a species-specific anatomy ontology. This is because it is very difficult to further subtype CARO and remain within the bounds of disjoint structural definitions. For example, the compound eye of a Drosophila and the camera-lens eye of a human have little in common structurally, making it unlikely that the type 'eye' would be included in CARO (though these types might be grouped, outside of CARO, using the function 'to see'). However, it may be possible to achieve a disjoint set of structural definitions for particular monophyletic groups within multi-species anatomy ontologies.

A number of projects aim to generate anatomy ontologies of multiple taxa. In particular, the Cypriniformes Tree of Life (CToL)[11], the plant ontology[12], as well as the amphibian[13], and hymenoptera[14] anatomy ontologies. As in the case of species-specific anatomy ontologies, multi-species anatomy ontologies can also clone the CARO types for use as their topmost nodes. Within a multi-species anatomy ontology, a type that satisfies the definition of a CARO type will have an *is_a* relation to the CARO type with the *differentia* of a taxon rather than a species. For example, for the cypriniform fishes anatomy ontology, the cypriniform class 'organ' *is_a* CARO:organ, with the *differentia* being that it is an organ of a type found in Cypriniformes. CARO can in this way be used as a template for multi-species anatomy ontologies as well as for species-specific ones.

Currently, many ontology developers use an existing ontology when building a new one (as CARO itself is modeled on the FMA). For example, the zebrafish anatomy ontology has been used as a template for both fish and amphibian multi-species ontologies. This is because the zebrafish anatomy ontology refers to anatomical structures that evolved within chordates – a post-anal tail evolved at the level of Chordata, the lateral line system evolved at the level of Craniata, jaws evolved at the level of Gnathostomata, and bone at the level of Vertebrata (Figure 4).

---

[11] http://www.nescent.org/wg_fishevolution
[12] www.plantontology.org
[13] www.morphologynet.org
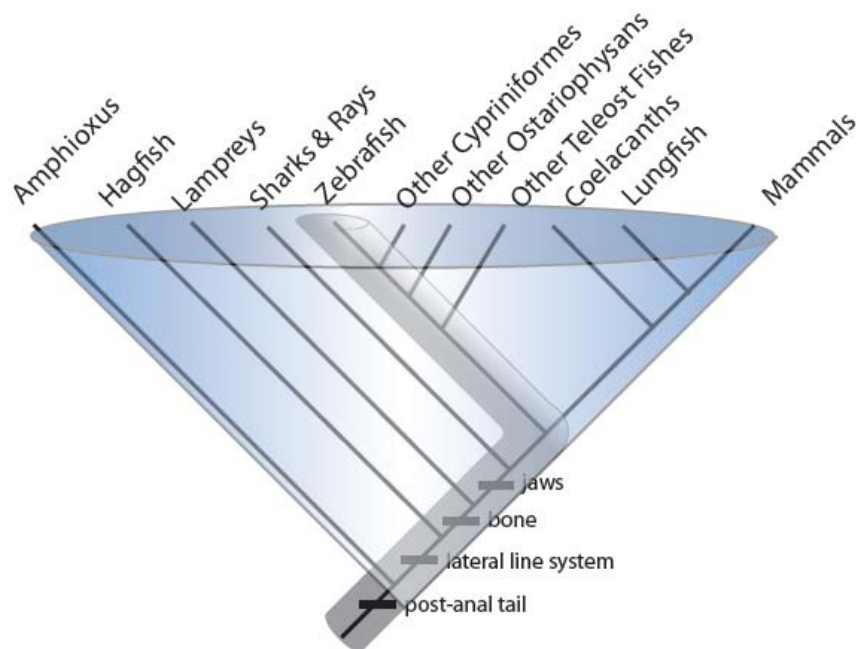[14] http://ceb.scs.fsu.edu/ronquistlab/ontology/wiki/index.php/Main_Page

Figure 4: Species-specific anatomy ontologies contain types applicable to more diverse taxa. The zebrafish anatomy ontology (inner darker cylinder) includes terms referring to features that evolved at various times in the chordate lineage. This ontology could be expanded to include anatomical structures found in all vertebrates (entire cone).

Within multi-species anatomy ontologies it is necessary to specify in which organisms the anatomical entities are applicable. This can be accomplished with the relation, *part_of_organism*, proposed by the CToL group to link anatomical entities to taxa within a taxonomy ontology.[15] Similarly, the types in CARO are not applicable to all organisms. For example, diploblastic animals such as cnidarians (a phylum that includes jellyfish and sea anemones) lack mesoderm (a proposed CARO term) while sponges may have no distinct germ-layers at all (Holland, 1998). CARO classes could also be linked to a taxonomy ontology to indicate which classes are applicable at various taxonomic levels. The purpose of cross-referencing multi-species anatomy ontologies and CARO to a taxonomic ontology would be to provide a user with choice of appropriate types. A similar method has been proposed to limit classes to specific taxa in other species-independent ontologies such as the GO or the CL (Waclaw Kusnierczyk, personal communication). It is important to note that cross-referencing anatomy and taxonomy ontologies in this manner does not specify homology.

8. Representing Homology

Methods for recording homology between types in anatomy ontologies are extremely important both to provide resources for evolutionary biologists and for the development of tools for inter-species inference regarding the molecular basis of morphological phenotypes or traits. Structures (including genes) are homologous if they evolved from

---

[15] CToL working group, National Evolutionary Synthesis Center.

some structure in a common ancestor, and homology implies genealogical descent as the vehicle of transfer of information. Homology must be addressed within the context of multi-species anatomy ontologies because of the very nature of how anatomical structures evolve. The reason anatomical types are structurally or functionally similar, and therefore classified together in some ontology, may be because they are evolutionarily related. However, many well documented counter examples exist. For example, both zebrafish and humans have a skull bone named the parietal bone, and another named the frontal bone. These could be grouped in an ontology on the basis of position within the skull and name. However, there is good evidence that the parietal bone in humans is homologous to the frontal bone in zebrafish (Schultze, 1985; Jollie, 1962). Thus, one cannot assume homology based on structural similarity or name.

We propose that homology information be captured independently of both structure and function information. Specifically, statements of homology are hypotheses and require evidence (codes) and attribution. This is particularly important to evolutionary biologists creating phylogenies, where different evidence is often used to generate different phylogenetic views. In light of this need to capture homology, a new relationship, *homologous_to*, is proposed to be included in the RO, but its definition is still under discussion[16]. The ontological implications for this new relationship are as yet untested. For instance, if two structures are deemed homologous, is this information transitive down *is_a* chains? Can two structures be homologous if none of their parts are homologous? Erwin and Davidson (2002) have suggested that the regulatory processes that underlie development may be homologous, whereas the creation of gross anatomical structures is specific to phyla or classes (and may not be homologous). In this respect, it is the processes or functions that are homologous whereas the structures are not.

To establish a homology relation between sister anatomical entities may require the determination of an evolutionary precursor in order to create sister subtypes within a multi-species anatomy ontology. It may prove difficult in some cases to define an evolutionary precursor purely on a structural basis and will require domain experts whose expertise spans large branches of the tree of life. However, it is possible that a function ontology used in combination with homology statements could overcome this difficulty. Multi-species anatomy ontologies will have to reconcile these homology issues with maintenance of disjoint definitions based on structure. It is important to note that even though one intended use for CARO is as a template for building multi-species anatomy ontologies, no homology between types is implied by common treatment within CARO, since CARO types are classified purely on the basis of structural criteria and not on evolutionary history.

9. Long term CARO goals

One of the long-term goals of CARO is to provide the source of standardized representations of anatomical types used in creating composite types of the kind found in ontologies such as the GO's Biological Process ontology. Like CARO, GO is cross-species, describing types of biological process that occur across a wide variety of species,

---

[16] CToL working group, National Evolutionary Synthesis Center.

encompassing types such as *heart development* and *neural tube closure*. Like CARO, GO is also canonical – it describes the features of typical, wild-type instances. At the present time, GO does not contain explicit references to types from an anatomical entity ontology. Instead, rough definitions of types such as *heart* and *neural tube* are 'embedded' inside the definitions of the corresponding GO types. This leads to redundancy, duplication of effort, inconsistency and a poor basis for cross-domain inference.

Once CARO is in use as a template for species-specific or multi-species anatomy ontologies, types from these ontologies along with their taxonomic reference can be referenced by the GO. GO will retain types such *as neural tube closure*, but the corresponding definitions can refer to definitions taken from CARO or from one of the multi-species or single-species anatomy ontologies created in a way which will allow the ontologies to be kept synchronized (Mungall, 2004).

The primary axis of classification in CARO is structural, not functional. This does not mean that CARO ignores function; rather it is the case that CARO insists that function be treated as a separate *orthogonal* ontology. Instead of stating that *verterbrate eye is_a sense organ* as we may do in a mixed classification, we instead state that *vertebrate eye has_function visual perception*, with the *is_a* parent of verterbrate eye being the appropriate structural supertype (i.e. cavitated organ). Separating structure from function in this way leads to cleaner ontology design, with each type having a single *is_a* parent. At the same time, this methodology still allows for cross-ontology queries, such as "find all genes active in sense organs." The organismal function ontology that will be used in conjunction with CARO or other anatomy ontologies is yet to be developed. Like CARO, this ontology will adhere to OBO Foundry principles and be itself placed in the OBO Foundry (http://obofoundry.org/). Many of these functions will be realized in biological processes of the kind found in the GO, so this ontology will be developed in coordination with the Gene Ontology Consortium.

One final consideration is that CARO compliance can be exploited to help build phylogenetic views of a given set of taxa. Since all species-specific and multi-species anatomy ontologies will have *is_a* links to CARO nodes, it will be possible to view an assembly of anatomical structures by limiting the taxonomic level. In combination with a set of homology statements, one could build different phylogenies based on different evidence. This is not unlike the current method of creating phylogenies, except that the anatomical structures are named and assigned to taxa in a standardized manner thereby providing links to other relevant data. For example, the development and function of homologous structures in two different species are likely to retain at least some of the molecular mechanisms present in the ancestral structure in their most recent common ancestor. CARO should in this way prove a useful organizational tool to facilitate the inference of molecular mechanisms underlying morphology.

References

Berger, C., Urban, J. and Technau, G.M. (2001) Stage-specific inductive signals in the Drosophila neuroectoderm control the temporal sequence of neuroblast specification *Development* 128:3243-3251.

Campos-Ortega, J. A. and Hartenstein, V. (1999) The Embryonic Development of Drosophila melanogaster. (Second Edition), Springer-Verlag, Berlin ISBN 3-540-57070-9

Cook, D.L., Mejino, J.L.V., and Rosse, C. (2004) Evolution of a Foundational Model of Physiology: Symbolic Representation for Functional Bioinformatics. *Medinfo* 11(1):336-40.

Dahmann, C. and Basler, K. (1999) Compartment boundaries: at the edge of development. *Trends Genet.* 15(8):320-6.

Erwin, D.H. and Davidson, E.H. (2002) The last common bilaterian ancestor. *Development* 129:3021-3032.

Farris, J.S. (1970) Methods for computing Wagner Trees. *Syst. Zool.* 19:83-92

Fitch, W.M. (1971) Toward defining the course of evolution: minimum change for a specific tree topology. *Syst. Zool.* 20, 406-416.

Gene Ontology Consortium (2006) The Gene Ontology (GO) project in 2006. 34:D322-6.

Holland, P.W.H. (1998) Major Transitions in Animal Evolution: A Developmental Genetic Perspective. *American Zoologist.* 38(6): 829-842.

Jollie, M. (1962). Chordate Morphology. New York, Reinhold Books.

Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B., and Schilling, T.F. (1995) Stages of embryonic development of the zebrafish. *Dev. Dyn.* 203:253-310.

Matsuoka, T. Ahlberg, P.E., Kessaris, N., Iannarelli, P., Dennehy, U., Richardson, W.D., McMahon, A.P., and Koentges, G. (2005) Neural crest origins of the neck and shoulder. *Nature* 436(7049):347-355

Mungall, C: Obol: integrating language and meaning in bio-ontologies. *Comparative and Functional Genomics*; Volume 5, Issue 6-7, 2004. Pages 509-520.

Nieuwkoop and Faber, 1994, Normal Table of Xenopus laevis. 3rd Ed.

Rosse, C. and Mejino, J.L.V. (2003) A Reference Ontology for Bioinformatics: The Foundational Model of Anatomy. *Journal of Biomedical Informatics* 36:478-500.

Rosse, C., Mejino J.L.V. , .Modayur B.R., Jakobovits R., Hinshaw K.P., Brinkley J.F. Motivation and organizational principles for anatomical knowledge representation: the Digital Anatomist symbolic knowledge base. *Journal of the American Medical Informatics Association*, 5(1):17-40, 1998.

Schultze, H. P. and M. Arsenault (1985). The panderichthyid fish Elpistostege: A close relative of tetrapods? *Paleontology* 28: 293-309.

Slack, J. (1991). From egg to Embryo (2nd Ed). Cambridge University Press.

Smith, B. (2001) Fiat Objects *Topoi,* 20(2):131–148.

Smith B, Ceusters W, Klagges B, Kohler J, Kumar A, Lomax J, Mungall CJ, Neuhaus F, Rector A, Rosse C (2005a) Relations in Biomedical Ontologies. *Genome Biology* 6:R46

Smith B, Kumar A, Ceusters W, Rosse C. (2005b) On carcinomas and other pathological entities. *Comparative and Functional Genomics*, vol 6, 7-8, 379-387, 2005.