**Uberon: Unification of multi-species vertebrate anatomy ontologies for comparative biology**

M.A. Haendel1([haendel@ohsu.edu](mailto:haendel@ohsu.edu)), J.P. Balhoff 2([balhoff@nescent.org](mailto:balhoff@nescent.org)), F. Bastian3 ([frederic.bastian@unil.ch](mailto:frederic.bastian@unil.ch)), D.C. Blackburn4 ([dblackburn@calacademy.org](mailto:dblackburn@calacademy.org)), J.A. Blake5, ([Judith.Blake@jax.org](mailto:Judith.Blake@jax.org)), A. Comte3 ([Aurelie.Comte@unil.ch](mailto:Aurelie.Comte@unil.ch)), W.M. Dahdul2,6 ([wasila.dahdul@usd.edu](mailto:wasila.dahdul@usd.edu)), T.A. Dececchi2,6 ([alex.dececchi@usd.edu](mailto:alex.dececchi@usd.edu)), R.E. Druzinsky7 ([druzinsk@uic.edu](mailto:druzinsk@uic.edu)), T.F. Hayamizu5 ([Terry.Hayamizu@jax.org](mailto:Terry.Hayamizu@jax.org)), N. Ibrahim8 ([nibrahim@uchicago.edu](mailto:nibrahim@uchicago.edu)), S.E. Lewis9 ([suzi@berkeleybop.org](mailto:suzi@berkeleybop.org)), P.M. Mabee2,6 ([pmabee@usd.edu](mailto:pmabee@usd.edu)), A. Niknejad3 ([Anne.Niknejad@unil.ch](mailto:Anne.Niknejad@unil.ch)), M. Robinson-Rechavi3 ([Marc.Robinson-Rechavi@unil.ch](mailto:Marc.Robinson-Rechavi@unil.ch)), P.C. Sereno8 ([dinosaur@uchicago.edu](mailto:dinosaur@uchicago.edu)), C.J. Mungall9([cjmungall@lbl.gov](mailto:cjmungall@lbl.gov))

1 Dept Med Info & Epid, Oregon Health & Sci. Univ.; 2Natl Evol Synth Center; 3Dept. Eco and Evol, U of Lausanne; 4Dept Vert Zool & Anthro, Cal Acad Sci; 5The Jackson Laboratory; 6Dept Biol, U So Dakota; 7Dept of Oral Biol, U of IL, 8Dept Org Biol & Anat, U of Chicago; 9Lawrence Berkeley Natl Lab

**KEYWORDS: evolutionary biology, vertebrates, anatomy ontology, morphological variation**

Elucidating evolutionary processes, disease and developmental dysfunction requires understanding variation in phenotype. For ‘natural’ phenotypic variation across species, multi-species AOs have been developed. The Phenoscape project ([www.phenoscape.org](http://www.phenoscape.org/)) has been working to make vertebrate multi-species (Teleost, Vertebrate Skeletal, and Amphibian AOs; TAO, VSAO, AAO) and ssAOs interoperable for the purposes of connecting morphological variation with candidate genes. The Bgee project ([http://bgee.unil.ch](http://bgee.unil.ch/)) vertebrate Homologous Organs Group (vHOG) ontology specifies homology among high-level organ groups for the purposes of annotating and querying gene expression across species. Previously, multi-species AOs contained a mixture of unique and overlapping content. This hampered integration and coordination due to the need to maintain cross-references or inter-ontology equivalence axioms to the ssAOs, or perform massive obsolescence and modular import. To rectify this, stakeholders decided to unify multi-species AOs with Uberon because of its metazoan scope and existing integration and synchronization methods.

Here we present the unification of five multi-species anatomy ontologies (Uberon, TAO, VSAO, AAO and vHOG) resulting in interoperability among disparate data and research groups (Phenoscape, Bgee, MODs). In essence, these ontologies were merged into Uberon. Where there was overlapping content the redundant and/or conflicted classes were reviewed and text and axioms improved before consolidating. Where there was no overlapping content, the proper axioms were added to include the new classes. Separate ID-spaces were retained for each source ontology. To enable an efficient workflow including division of labor and editorial rights, Uberon separates axioms into a core ontology that is maintained in OBO and an OWL extension ontology (ext) that imports the core file. Unique classes from the multi-species AOs, now in the ext, maintain relationships to Uberon core classes. The classes in the ext are maintained by Phenoscape curators in Protégé. We also developed an attribution strategy whereby we create citation records to promote attribution of stakeholder contributions (see<http://purl.obolibrary.org/obo/uberon/references/index.html>).

The newly broadened Uberon ontology is a single expert resource for vertebrate anatomy terminology that permits reasoning across anatomical variation in extinct and extant taxa. It permits queries for candidate genes using annotations for phenotypes from the systematics, medical, and model organism communities as well as for logical definitions in Cell and Gene Ontologies. This new extended Uberon is available at:<http://purl.obolibrary.org/obo/uberon/ext.owl>. We invite contributions to Uberon via the tracker, monthly calls, and direct editing.

**References**

[1] Mungall, C. J., C. Torniai, G. V. Gkoutos, S.E. Lewis, and , M. A. Haendel. 2012. Uberon, an integrative multi-species anatomy ontology. *Genome Biology 13, R5.* <http://genomebiology.com/2012/13/1/R5>

**Outline**

**Introduction**

1. Background (1/2 page)
   1. species specific AOs
   2. Uberon
   3. multi-species AOs
2. Issues
   1. common design patterns
      1. Fig 1 - different modeling as per below
   2. synchronization
   3. query across ontologies

**Methods**

1. Merger
   1. Roll up methods
      1. Fig 2 - overview as per below
      2. Fig 3 - specific example
   2. taxon specificity inside uberon
   3. implications and repercussions
2. Ontology maintenance
   1. Repositories/Uberon.org/documentation
   2. Distributed editing

**Results and Discussion**

1. Ontology usage
   1. example queries/projects
2. Conclusions

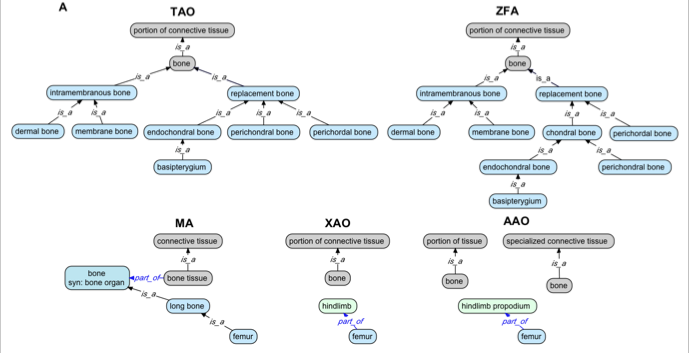


Fig 1. will show only bottom half, comparing femurs

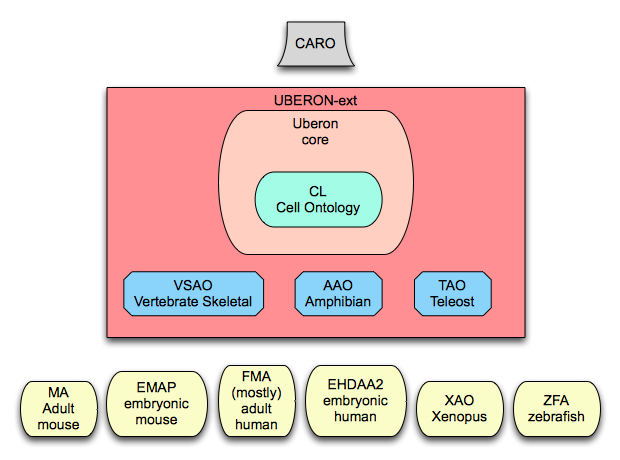


Fig 2 - will rename big box to just uberon

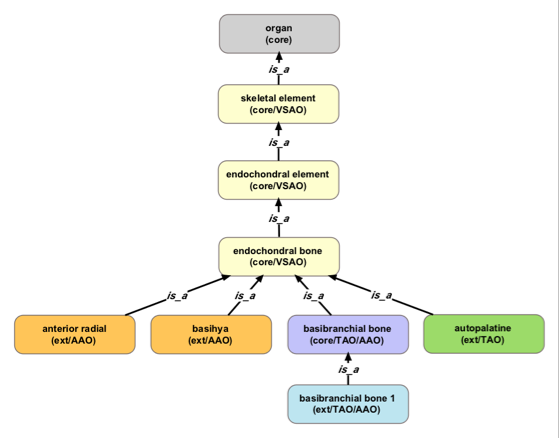


Fig 3- re-evaluate choice of example

**Introduction**

**Background**

B)

Previously, multi-species AOs contained a mixture of unique and overlapping content. This hampered integration and coordination due to the need to maintain cross-references or inter-ontology equivalence axioms to the ssAOs, or perform massive obsolescence and modular import. To rectify this, stakeholders decided to unify multi-species AOs with Uberon because of its metazoan scope and existing integration and synchronization methods. The Uberon multi-species anatomy ontology was constructed to provide a generalization of anatomical structures across metazoans, with an emphasis on chordates [Mungall 2012]. The original purpose for Uberon was to enable query across model organisms based on different types of similarity, including asserted homology, function, development, etc, for the purposes of identifying candidate models for human disease based on similarity of anatomical phenotypes [Washington 2009].

C)

For ‘natural’ phenotypic variation across species, multi-species AOs have been developed. The Teleost Anatomy Ontology (TAO) covers the teleosts (bony fishes), with an emphasis on the skeletal system [Dahdul 2010] while the Amphibian Anatomy Ontology (AAO) covers amphibians [Maglia 2006]. The vertebrate skeletal anatomy ontology (VSAO) was constructed with similar goals, focusing on the skeletal system, but extended its reach to all vertebrates [Dahdul 2012]. Previously all of these ontologies contained some mixture of unique specific contributions together with overlapping content. This resulted in both duplication of effort, as each ontology defined classes such as ‘vertebra’ and ‘nervous system’. In addition, from a users’ perspective, integration was hampered by the need to maintain cross-references or inter-ontology equivalence axioms.The Phenoscape project ([www.phenoscape.org](http://www.phenoscape.org/)) has been working to make vertebrate multi-species (TAO, VSAO, AAO) and ssAOs (MGI, Xenbase and Zfin) interoperable for the purposes of connecting morphological variation with candidate genes. Semi-independently the Bgee project ([http://bgee.unil.ch](http://bgee.unil.ch/)) has focused more on ontogentic and expression level patterns in a more restricted set of model systems. Also the vertebrate Homologous Organs Group (vHOG) ontology has worked to assert homology among high-level organ groups for the purposes of annotating and querying gene expression across species. To reduce the duplication of effort caused by multiple groups are working in parallel, we therefore decided to combine efforts into a single multi-species ontology. We opted to use Uberon as the target ontology, as it had the most extensive coverage across metazoans, and was well-integrated with existing single-species anatomy ontologies

**Issues**

### Merge process

Our goal was to merge three source ontologies (TAO, AAO and VSAO) into a target ontology (Uberon). These ontologies were previously maintained in OBO format using OBO-Edit but were translated into OWL Format. Owl was chosen as it allows for greater expressivity and use of OWL based tools, modularization for easier reuse and interoperability, and use of better annotation properties for attribution and for UI and community-specific needs. This shift in formats required training Phenoscape curators in Protege. At the start of the work described here, most structures shared between two or more existing ontologies were already present in the core Uberon, and mapped back to the source ontology. We therefore took all classes that were unique to each source ontologies, and not shared with Uberon, and moved them into an ontology that extends the target ontology (hereafter called ‘ext’). For class moved, we generated a new primary UBERON IDs which contain cross-references to the original term ID, but we preserved all original text definitions in each source ontology using an ‘external definition’ annotation property. Expert vetted relationships from the VSAO were protected in Uberon. VSAO definitions were added as primary definitions and VSAO labels were also chosen as primary in Uberon. We rewrote axioms (relationships) such superclasses and assertions were using Uberon IDed terms, not IDs from source ontologies. For example in the AAO caput glenoidale (AAO\_0000801) is part\_of some centra, while in Uberon the same term (Uberon ID 3000801) is asserted to be an ‘organ part’ and part\_of some ‘vertebral centrum’.

### 

**Ontology maintenance**

### Division and coordination of labor

The separation of axioms into a core ontology and an extension allows a division of labor and editorial rights. Original Uberon terms were distinguished as ‘core’, even when shared by terms coming in from other ontologies (in particular because they were shared). Unique terms that from the other ontologies were considered ‘ext’ terms. Core ontology terms are displayed in black and ext terms in gray in Protege for ease of distinguishing editing rights. Maintenance and editing of the core ontology requires specialized tooling to keep the ontology in sync with other anatomy ontologies. However, the ext ontology contains subclasses of core classes and can be maintained by Phenoscape curators in Protege without this specialized tooling. The distinction between core and ext terms is largely an editorial artefact and disappears with the released artifact. As of 2013-MM-DD, the import chain are combined into the main Uberon release.

Allocation of responsibilities and coordination of editing rights were critical parts of the merging process among these different communities. Lead developers of the uberon (CM and MAH) have editing rights for the core uberon. Proposed changes to the core terms are submitted through the Uberon tracker and vetted by larger community. Phenoscape curators have made major contributions to the core ontology via the Uberon tracker, design documents, meetings and workshops. Phenoscape curators (NI, AD, WD) have editing rights for the ext. Changes to ext are often done through calls between editors and/or using the provisional term tacker to list proposed terms and definitions.

Contributions to the ontology are marked as metadata in the ontology, at the level of classes, axioms and the ontology itself. Authorship of definitions for TAO terms, new terms defined by Alex and Nizar ORCID IDs. Ontology development tools

We used Protege 4, plus the Obo-Edit graph component [Dietze et al, in prep] from Obo-Edit [Day-Richter et al.], plus the depictions plugin [Balhoff, in prep?], and additional plugins that support an OBO style of development.

### Centralized reasoning using Jenkins

As described [Mungall 2012], Uberon includes a large number of axioms to ensure both internal consistency and consistency with other ontologies. We required that the ontology always satisfy these conditions. In addition, we required checks that the ontology did not violate certain syntactic conditions (e.g. every class should have exactly one label and no more than one text definition). Some of these checks are difficult to run locally on the ontology developer’s local machine, so we made use of the OBO library Jenkins Continuous Integration server[REF]. Every commit performed by an editor triggers a validation and build pipeline, which also generated an obo-format file which is used in Phenoscape character annotation in Phenex and entity matching through CharaParser [Cui et al?].

**Consistency checks**

We have implemented a process by which the xrefs and equivalence axioms are checked between the single-species anatomy ontologies and the ontologies now integrated into uberon. For example, if a zebrafish class has an xref to a TAO class that has not been obsoleted in Uberon without a replaced by, these will be unsatisifable and will be fixed in the source zebrafish file. (Placeholder text...)

Different methods of coordination were considered, and we decided to combine these ontologies, eliminating the need to maintain cross-references or inter-ontology equivalence axioms. Governance, editing rights (core vs. ext), etc. - describe

**Add: why merger was needed, what was driving it, effort needed:**

**General goals:**

1) Promote interoperability between different data sources, including Phenoscape and MOD data. This means development of a quality data set annotated with ontologies that allows

inference across species.

2) Promote attribution of contributions to ontology development

3) Promote an efficient workflow that reduces manual labor and allows contribution.

4) Promote best practices in ontology development, resulting in logically consistent and orthogonal ontologies. This will allow inference across data sets and sources.

**Phenoscape requirements/needs:**

1. Need an application ontology that enables annotation, reasoning, and browsing of comparative vertebrate knowledge.

2. Need to be able to add to and edit this application ontology in a timely fashion, in particular the skeletal components.

3. Need to be able to selectively see taxon/community specific labels and definitions (Note this is jointly coordinated by changes in ontology and UI).

4. We need processed versions of all the MOD ontologies that put their classes into the appropriate subclass relationships to classes in the application (multi-species) ontology. ZFA, XAO, MGI and other AOs are currently synchronized with Uberon, where the goal is a common representation and use reasoning to ensure consistency is maintained.

**Advantages of using a single application ontology that uses skeletal terms from VAO, AAO, UBERON and TAO**

1. Eliminate modular import process of VAO into TAO, AAO, and AmAO

2. Enable curators to add terms to a single ontology without concern over taxonomic scope.

3. Have a single resource in the immediate future (by end of June) that we can use to address the above requirements

Need based due to annotations for phenoscape. The lack of specificity for terms (i.e. podiumal terms). Disambiguating how different communities used terms.

Need to reflect in this paper the lots of discussion that was required by the merger; research with citations by Paul & Nizar initially, Alex also, discussion with Terry, Wasila, Paula, David, etc. for months. Give examples of major types of problems encountered (maybe just one?). Different expertise and backgrounds of scientists sharing and using this resource, including MODs

Significantly different -- more than addition of 8K terms and defs, but vetting, editing, referencing, consolidating etc. across different communities.

...

An understanding of biological diversity necessarily requires an understanding the variation in morphological form, and how such variation has evolved. One way in which to examine this variation is to compare anatomical features based on structural, functional, and developmental similarity.

-----

### Previous work

The Uberon multi-species anatomy ontology was constructed to provide a generalization of anatomical structures across metazoans, with an emphasis on chordates [Mungall 2012]. The original purpose for Uberon was to enable query across model organisms based on different types of similarity, including asserted homology, function, development, etc, for the purposes of identifying candidate models for human disease based on similarity of anatomical phenotypes [Washington 2009]. Here we describe recent enhancements to Uberon

The vertebrate skeletal anatomy ontology (VSAO) was constructed with similar goals, focusing on the skeletal system of vertebrates [Dahdul 2012]. The Teleost Anatomy Ontology (TAO) covers the teleosts (bony fishes), with an emphasis on the skeletal system [Dahdul 2010]. The Amphibian Anatomy Ontology (AAO) covers amphibians [Maglia 2006]. Previously all of these ontologies contained some mixture of unique specific contributions together with overlapping content. This resulted in both duplication of effort, as each ontology defined classes such as ‘vertebra’ and ‘nervous system’. In addition, from a users’ perspective, integration was hampered by the need to maintain cross-references or inter-ontology equivalence axioms.

We therefore decided to combine efforts into a single multi-species ontology. We opted to use Uberon as the target ontology, as it had the most extensive coverage, and was well-integrated with existing single-species anatomy ontologies

Add summary of Phenoscape import effort (VSAO, TAO). TAO was updated with import of VSAO. TAO, VSAO, and AAO no longer updated or being maintained independently.

As of 2013-MM-DD, as a result of our combined efforts, the public release of Uberon included additional coverage across non-model organism anatomy, particularly in the realm of skeletal structures. SOME STATS HERE. In addition, multiple additional classes have been added and fixes made to existing parts of the ontology, with a major contribution coming from the requirements and efforts of the Phenoscape project. In this update, we provide a summary of these updates.

Similarly, a few workshops had taken place with a focus on specific vertegrate anatomical systems, and this work needed a home. Neural crest, FEED, homology/digits

# **Methods**

Many of the ontology extension and maintenance methods are described in the original Uberon publication. We briefly summarize these here, and describe some recent changes, as well as describing specifically the method for combining the multiple multi-species ontologies.

### Merge process

Our goal was to merge three source ontologies (TAO, AAO and VSAO) into a target ontology. These ontologies were previously maintained in OBO format using OBO-Edit. Because the target ontology was an OWL ontology called ‘ext’ which owl-imports the core Uberon ontology, this required training Phenoscape curators in Protege. This entails move to OWL for expressivity and use of OWL based tools, modularization for easier reuse and interoperability, and use of better annotation properties for attribution and for UI and community-specific needs. At the start of the work described here, most structures shared between two or more existing ontologies were already present in the core Uberon, and mapped back to the source ontology. We therefore took all classes that were unique to each of the other source ontologies and moved them into an ontology that extends the target ontology (hereafter called ‘ext’). For each such class moved, we generated a new UBERON identifier. We preserved all original text definitions in each source ontology using an ‘external definition’ annotation property. We rewrote axioms (relationships) such that source IDs were replaced by Uberon

IDs. For example …Terms merged from the three ontologies were given primary UBERON IDs which contain cross-references to the original term IDs.

Expert vetted relationships from the VSAO were protected in Uberon. VSAO definitions were added as primary definitions and VSAO labels were also chosen as primary in Uberon.

The resulting combined ontology required additional cleanup...

2. What is the ext? ‘core’ and ‘ext’ created in merger. What is ‘core’? No concept of this before this merger. How defined? Created so that Phenoscape could edit uberon.

Need to discuss with Melissa/Chris how big our “slice” was, how do you get “slices”, can someone take multiple different “slices” to make a new ext?

### We contributed terms to core (how many?) plus terms to ext (how many?) contributed text and logical definitions (quantify). Synonyms.

### Division and coordination of labor

The separation of axioms into a core ontology and an extension allows a division of labor and editorial rights. Original Uberon terms were distinguished as ‘core’, even when shared by terms coming in from other ontologies (in particular because they were shared). Unique terms that from the other ontologies were considered ‘ext’ terms. Core ontology terms are displayed in black and ext terms in gray in Protege for ease of distinguishing editing rights. Maintenance and editing of the core ontology requires specialized tooling to keep the ontology in sync with other anatomy ontologies. However, the ext ontology contains subclasses of core classes and can be maintained by Phenoscape curators in Protege without this specialized tooling. The distinction between core and ext terms is largely an editorial artefact and disappears with the released artifact. As of 2013-MM-DD, the import chain are combined into the main Uberon release.

Allocation of responsibilities and coordination of editing rights were critical parts of the merging process among these different communities. Lead developers of the uberon (CM and MAH) have editing rights for the core uberon. Proposed changes to the core terms are submitted through the Uberon tracker and vetted by larger community. Phenoscape curators have made major contributions to the core ontology via the Uberon tracker, design documents, meetings and workshops. Phenoscape curators (NI, AD, WD) have editing rights for the ext. Changes to ext are often done through calls between editors and/or using the provisional term tacker to list proposed terms and definitions.

Contributions to the ontology are marked as metadata in the ontology, at the level of classes, axioms and the ontology itself. Authorship of definitions for TAO terms, new terms defined by Alex and Nizar ORCID IDs. Ontology development tools

We used Protege 4, plus the Obo-Edit graph component [Dietze et al, in prep] from Obo-Edit [Day-Richter et al.], plus the depictions plugin [Balhoff, in prep?], and additional plugins that support an OBO style of development.

### Centralized reasoning using Jenkins

As described [Mungall 2012], Uberon includes a large number of axioms to ensure both internal consistency and consistency with other ontologies. We required that the ontology always satisfy these conditions. In addition, we required checks that the ontology did not violate certain syntactic conditions (e.g. every class should have exactly one label and no more than one text definition). Some of these checks are difficult to run locally on the ontology developer’s local machine, so we made use of the OBO library Jenkins Continuous Integration server[REF]. Every commit performed by an editor triggers a validation and build pipeline, which also generated an obo-format file which is used in Phenoscape character annotation in Phenex and entity matching through CharaParser [Cui et al?].

**Consistency checks**

We have implemented a process by which the xrefs and equivalence axioms are checked between the single-species anatomy ontologies and the ontologies now integrated into uberon. For example, if a zebrafish class has an xref to a TAO class that has not been obsoleted in Uberon without a replaced by, these will be unsatisifable and will be fixed in the source zebrafish file. (Placeholder text...)

# **Results**

## The process of integrating different anatomical ontologies highlighted several inconsistencies in terminology both between and within anatomical and/or medical nomenclatures. A good illustration of this process is the issue of major appendicular divisions (see supplemental information), which have been used to refer to different sections (ref) and tissues of the limb (ref) by different research communities (ref).

## In addition, some terms have been used to describe fundamentally different anatomical parts, a rare but significant problem that surfaced during the process of term creation as well as annotation. One example is the use of the term acropodium, which should only refer to the phalanges (), but is used by some authors to refer to the entire manus skeleton () minus the mesopodium (see Figure xx). We suggest keeping acropodium (or acropodial skeleton) for the phalanges and introduce a new term for the metapodium + acropodium, digitipodium (Figure xx).

## Merging of multiple multi-species ontologies

The most significant advance since the initial publication of papers describing Uberon, TAO, VSAO and AAO has been the combination of these ontologies into a combined ontology. This required a considerable amount of coordination between multiple groups.

### Merging of TAO

Prior to this work, a number of TAO classes had been obsoleted and replaced by VSAO classes, and these were incorporated as part of the VSAO merge, below. Other classes represented largely teleost-specific structures. These were moved into Uberon, and tagged as belonging to TAO. Additional manual reconciliation took place. As TAO was originally created by cloning ZFA, the ontology contained a large number of classes that were not truly generalized beyond Danio rerio, including a large number of cell types. These were obsoleted.

### Merging of AAO

..

### Merging of VSAO

As VSAO consistent largely of mid-high level classes, many of these were already present in Uberon at the start of this work. The VSAO editors had worked closely with the Uberon editors to reconcile differences and coordinate changes.

Merging of vHOG??

### - mention stages?

We also created “homology notes” fields for nnn classes derived from vHOG organ association table.

### Incorporation of FEED

[awating Rob D’s response] We created relationships for muscle attachment and innervation. In future we may add relationships to muscle action to NBO [Gkoutos2012].

...

### Generation of new classes to support Phenoscape annotation

In addition to merging these ontologies, we augmented Uberon with a large number of classes to support description of such things as A, B and C. These will be described in full details in a separate publication.

## Inline documentation

...

## Integration of depictions

This update also incorporates a new effort to include images depicting instances of ontology classes in a variety of species. These were generated in part by mining Wikipedia/DBpedia, and also by manual curation using image depictions plugin for Protege 4 [Balhoff, todo?].

## Neuroanatomy

TODO - discuss this with Maryann

## Applications

The original vHOG was intended primarily for integrating gene expression data. The original Uberon was intended primarily for phenotype data integration (between model organisms and human) and structuring GO. The original TAO and AAO were intended for evolutionary character annotation within their respective taxa. The original VSAO was intended as a mid-upper level reference ontology for skeletal structures. The combined ontology can serve all these purposes.

In addition, new applications have been developed. FANTOM5. LAMHDI. eagle i. Others...

# **Discussion**

## Info here about CARO, previous thoughts about starting with ss-AOs and a top level one, and finding the middle. In the end, it didn’t make sense to maintain the multi-species AOs at lower taxonomic levels. (some should go in discussion)

## Benefits of integration

The combination of multiple multi-species ontologies required considerable coordinated effort and compromise, but the end result benefits both the maintainers of these ontologies and their under users.

Prior to the integration,

--Multispecies AOs (TAO, AAO, VSAO) not xrefed to MODs.

--Uberon had relatively sparse coverage of vertebrate anatomical structures that were not present in at least two core model organisms or human (i.e. the classes generalized classes in FMA, MA, XAO and ZFA). For example: LIST SOME EXAMPLES HERE...

We also improved many of the sections of Uberon that were already described in the original paper by applying expert anatomical knowledge from an evolutionary perspective. EXAMPLES: limb sibdivisions, digitopodium examples, etc.

The users of AAO and TAO gained in other ways. For the subset that were present in Uberon (e.g. classes such as heart, brain) the gain involved integration with other ontologies, including single-species anatomy ontologies, as this integration had already been performed as described in the original Uberon paper.

The integration also involved compromise. The resulting ontology is more complex than an ontology that was developed specifically to satisfy the annotation requirements of Phenoscape. For example, many appendicular subdivisions are in some sense represented multiple times in parallel hierarchy. Uberon contains not only digit, manus, pectoral limb, but also digit skeleton, manus skeleton, pectoral limb skeleton, as well as digit bone, manus bone, pectoral limb bone. From the point of view of evolutionary systematics annotation, only the skeleton terms might be required; an ontology that was restricted to these terms would be more usable for this user group, and might perform better in automated entity matching. We can recapitulate this simplicity to some extent by creating ontology views [ref: modularization paper], but this pushes the complexity to the release process.

Another compromise is in the choice of primary labels for classes. We decided to use evolutionary terminology where possible, as such terminology is the most neutral with respect to taxonomy. For example: manus and manual digit for hand and finger; ectepicondyle instead of lateral epicondyle of humerus [Alex to provide additional info]. We carried this naming convention all the way to substructures, e.g. distal phalanx of manual digit 1. This nomenculature comes at some expense of users who are more familiar with human-centric terminology. We ameliorate this to some extent by the use of special purpose domain-centric synonyms, but recognize that at this time tool support is such that each class effectively has a single primary label.

## Attribution

The ontology includes an ontology header with multiple annotations stating the creators and contributors of the ontology. In addition, the design documents are also annotated with authorship and contribution information. Relationships are not attributed currently in uberon -- but can we do this? Can we attribute VSAO relationships to Dahdul et al., 2012?

We will need to add to IAO a new annotation property for community specific definitions (which may be concatenated in the current application ontology for now in definition or comments field in obo file, parsed in separate annotations properties in Owl file).

We will use annotation properties to ensure that there is adequate attribution and documentation of decisions. These are: Definition Editor: GROUP: Phenoscape and/or PERSON:Nizar etc. The curator notes property will be used to record design decisions, point to tracker items, etc. The comments field can be used for any additional comments. We will include taxon-specific notes where relevant, we will use a new annotation property for this.

Chris has created an attribution schema here:

<https://github.com/cmungall/uberon/blob/master/uberon.references>

in which different working groups at various times and places may be attributed. We will continue to work with the OBO Foundry and NESCent to define best practices in ontology attribution using persistent and resolvable URIs.

## Ontology development strategy

...

## Future integration with non-vertebrate multi-species ontologies

Our goal is to create a collection of federated anatomical ontologies covering multi-cellular organisms. Some groups of taxa will be included within Uberon, others will be developed as separate ontologies that will work together with Uberon as part of a multi-species OWL import chain. In particular, the new Poriferan (sponge) ontology [Thacker, this issue] and Mollusc ontology [Gonzales, this issue] are being developed as separate ontologies that reuse a set of CARO and Uberon classes (with the mollusc ontology making use of a larger set of Uberon terms that Porifierans). As the Arthropod Anatomy Ontology ARTHRO [??, this issue] is developed further, a number of classes from Uberon will be obsoleted and ceded to this ontology. We are also seeking domain experts to either flesh out Uberon or bud off new ontologies covering organisms such as cnindarians, annelids and echinoderms, all of which have some superficial level of coverage at the moment.

[Diagram showing new ontology landscape].

..

# **Conclusions**

..

# **Author contributions**

AD and NI were the primary ontology editors for the merged ontology, generated new terms, resolved merge issues, evaluated and fixed existing Uberon classes, contributed high level design, providing expert evolutionary knowledge. CJM provided software support, devised the merge plan, ensured consistency with existing ontologies and extended the ontology in non-skeletal areas. MAH executed the merge plan, contributed expert knowledge and oversaw the ontology development. PM, PS and SEL oversaw the project and provided the environment and resources required. WD provided expert knowledge and assisted with the merge of TAO, VSAO and Uberon. DB and ES provided experty knowledge and assisteed with the merge of AAO. AN, AC, FB and MRR contributed vHOG. RD contributed FEED. JB provided technical support and developed plugins for the ontology curators to use.

# **References**

..

# **Figures**

..

Figure 1: Venn diagram showing provenance of classes in the new combined Uberon. For each ontology that was used as a source, a subset of classes had equivalent classes in Uberon, with the remaining classes representing new subclasses of existing Uberon classes. In addition, a number of new classes were added either to support annotation in the phenoscape project, or other projects.

…

Figure 2: A visual depiction of a subset of the ontology showing the class XXXX and descendant classes. Colors indicate the provenance of the classes, with zzz coming from AAO/TAO, yyy coming from VSAO and the rest in the core Uberon. Classes marked foo were generated by Phenoscape curators to support annotation. Graphs from the graph viewer from OboEdit.

...

Possible figure:<http://douroucouli.files.wordpress.com/2012/07/mouth.jpg?w=630&h=615>

Figure 3: Example of visualization of depiction axioms from uberon/depictions.owl

Figure 4: Major divisions of the manus skeleton.

Figure 5: possible shape of future ontology landscape. Dedicated single-species ontologies remain, and become more tightly integrated with Uberon. Separate federated multi-species anatomy ontologies ‘bud off’, importing some portion of the core ontology. Examples (existing): Profiera, Mollusc, Arthropod. Future: echinoderm, annelid. The plant ontology is also show side-by-side.

**SUPPLEMENTARY DATA**

**I. General introduction to “-Podium” terms**

Haeckel (1895) introduced many “-podium” terms in his treatise, *Phylogenie*, to refer to skeletal elements and their developmental anlagen. These terms, clearly, were not created to refer to composite bone/flesh limb segments but rather to skeletal elements and their developmental precursors. The use of “-podium” terms in reference to the vertebrate appendicular skeleton remains the most common usage in the recent literature, as exemplified in recent discussions of the adult appendage morphology of transitional basal sarcopterygians (Shubin et al. 2004) or the fin-limb transition (Hall 2007).

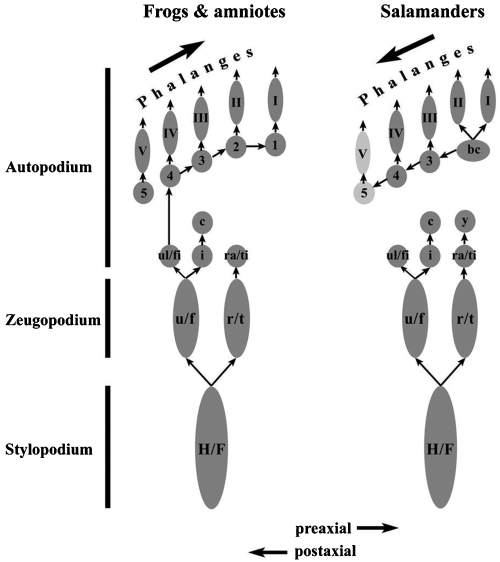
Nevertheless, some authors in the field of developmental biology have sometimes used these terms (often referring to the ‘autopod’ or ‘zygopod’) to refer to ‘limb segments’ (e.g. Chiu and Hamrick, 2002; Tamura et al., 2008), usually in the context of the *limb buds* (Tamura et al., 2008; Shou et al., 2005). In developmental biology the boundary between the so called ‘autopodium’ and ‘zeugopodium’ is well documented (Wagner and Chiu, 2001) based on the expression domains of *Hoxa 11* and *Hoxa 13* but the definitions of the terms themselves are often surprisingly vague and the usage of ‘-podium’ terms in this field makes it difficult to establish how much the meaning of these terms diverges from the original meaning. Even high profile articles are prone to this problem (Wagner and Chiu, 2001), as exemplified by the example below taken verbatim from Wagner and Chiu (2001) - our comments are in brackets and italics:

**II. “Limbs” and “fins”—basic taxonomy of terms**

“The archetypal limb [*The term ‘limb’ suggests that the author refers to skeletal and soft tissue structures*] of a tetrapod consists of three major segments: the upper limb or the stylopodium, the lower limb or the zeugopodium, and the hand/foot [*Again, the terms ‘hand’ and ‘foot’ refer to multi tissue segments*] or the autopodium (Fig. 1A). The stylopodium consists of one long bone [*here the wording clearly (correctly) suggests that the stylopodium* ***consists*** *of one long bone in the forelimb and another long bone in the hindlimb – but no other tissues*], the humerus in the forelimb, attached to the shoulder girdle and the femur in the hind limb, attached to the pelvic girdle. The zeugopodium is primarily [*in the case of the Zeugopodium the wording is different and suggests that it is ‘primarily’ composed of two long bones – even though no other tissues are mentioned*] composed of two long bones, the radius and ulna, and the tibia and fibula in the fore and hind limb, respectively. The autopodium consists of two segments, a proximal mesopodium and a distal acropodium. The mesopodium is a complex of nodular elements in most tetrapods, and is called the carpus[*saying that the mesopodium IS the carpus (i.e. mesopodium=carpus, or, in ontology terms, carpus is\_a mesopodium) is significant, since the carpus is a complex of bones*] in the hand and tarsus in the foot. The acropodium is a series of small long bones [*again, the acropodium is clearly described a s a series of bones*], the metacarpals and metatarsals as well as the digits [*‘digits’ is a more ambiguous term, because it can refer to skeletal and multi tissue segments*].”

As the above example shows, “-podium” terms are often loosely defined and used interchangeably to refer to skeletal segments or entire limb segments.

**In the case of the manus and pes skeletons, the following segments can be identified:**



**Figure modified from Frobisch (year?). Left manus/pes; preaxial = anterior; postaxial = posterior.**

The mesopodium/basipodium consists of the carpal/tarsal bones, the metapodium consists of the metacarpals / metatarsals, the acropodium of the phalanges.

The terms “preaxial” and “postaxial” refer to the anterior and posterior portions of the limb/ fin skeleton with the dividing line considered the metapterygial axis. In Tetrapoda this refers to the long axis of the stylopodial skeletal element (humerus in the forelimb, femur in the hindlimb).

**Manus terminology:** We suggest keeping acropodium (or acropodial skeleton) for the phalanges, as you suggested, and introducing a new term for the metapodium + acropodium, **digitipodium**. The term ‘digitipodium’ has been added on 10/11/2012, see:<https://github.com/cmungall/uberon/issues/131>.

--How to handle reports of taxa ‘lacking digits’? Metacarpals are NOT part of digits; acropodium = digits.

Digit = phalanges + sesamoids (including those at metaphalangeal - prehallux and prepollex should not be part of the digits -- although a minority view does see them as digits (argument has been made)

How best to say ‘digit 1 skeleton is missing’? ‘digit one skeleton, absent’ or ‘phalanx part\_of digit 1, absent’

**Major divisions**

**Skeleton**

1. manus/pes skeleton (autopodium)
   1. carpus/tarsus (mesopodium or basipodium)

i. carpal/tarsal bones (terms already defined)

* 1. metacarpals/metatarsals (metapodium)

i. metacarpal/metatarsal bones (terms already defined)

* 1. manual/pedal phalanges (acropodium)

i. phalanx terms

**Segment terms**

**1.** Hand/manus and foot/pes

a. wrist/ankle

b. metacarpal segment/metatarsal segment (metapodium)

c. Manual/pedal Digits

(hand proper=digitipodium)

Another major issue with digits is the problem of homology, both across taxa and between the fore and hindlimb. These include: the relationship between digits and radials, digit loss, sesamoid and “false digits” as well as hyperphalangy and polydactyly. We will avoid discussions of homology issues here as this aspect goes beyond the scope of this paper.

**III. Other issues relating to developmental biology**

The CVAO should refer to terms used or applicable to the adult skeleton. Commonly used terms can be defined concisely without trying to accommodate the sometimes less precise usage in developmental biology. This is well illustrated in the case of “-podium” terms.

The mesopodium is a skeletal term that refers to the carpus/tarsus in the vertebrate skeleton, as exemplified in literature on the fin/limb transition (see Shubin and Davis 2004). This skeletal segment is clearly distinct from the remainder of the autopodium (i.e. the metacarpals and phalanges). In developmental biology, however, the term “mesopodium” sometimes is regarded as including the skeletal parts of the first digit, or thumb (e.g., Woltering and Duboule 2010). While expression of the *Hoxa 13* gene appears to suggest that, in developmental terms, such a close association might exist, a definition of ‘mesopodium’ along these lines would not make any sense for morphologists. Numbering digits, which is a widespread in comparative anatomy, is considered problematic by some developmental biologists (Woltering and Duboule 2010).

Similar problems are likely to arise when terms such as “axial skeleton” and “cranial skeleton” are defined. Many useful and important musculoskeletal divisions disappear or become blurred in early development. For example, there is no clear boundary between the hand plate and the forearm, as digit ligaments are found in the forearm at early stages of development (Bard 2008). In adult animals, however, such a distinction is useful. Similarly, the traditional distinction between the axial and appendicular skeleton runs counter to recent developmental evidence, which suggests that the caudal series may well be part of the appendicularskeleton (Minelli, 2003). For example, acid treatment of some amphibian taxa can lead tail blastema to generate limbs, rather than a new tail (Maden, 1993; Brokes, 1997; Minelli, 2003).

Since a human developmental ontology, as well as *C. elegans* and *Drosophila* developmental ontologies have been created (available in BioPortal) to complement adult anatomy ontologies for these taxa, it is advisable to follow the same approach for the Comparative Vertebrate Anatomy Ontology, rather than trying to incorporate both adult and developmental terms in the same ontology (as has been attempted in the *Xenopus* Anatomy Ontology).

=====