Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 January 2022 doi:10.20944/preprints202201.0254.v1

Formalizing insect morphological data: a model-based, extensible insect anatomy ontology and its potential applications in biodiversity research and informatics

Jennifer C. Girón*,
Department of Entomology, Purdue University, West Lafayette, Indiana, USA
Natural Science Research Laboratory, Museum of Texas Tech University, Lubbock, Texas, USA.
entiminae@gmail.com; jgirondu@purdue.edu; jennifer.giron@ttu.edu
https://orcid.org/0000-0002-0851-6883

Sergei Tarasov
Finnish Museum of Natural History, Pohjoinen Rautatiekatu 13, FI-00014 Helsinki, Finland.
sergei.tarasov@helsinki.fi
https://orcid.org/0000-0001-5237-2330

Luis Antonio González Montaña
Universidad de los Llanos, Facultad de Ciencias Básicas e Ingeniería, Villavicencio, Meta, Colombia.
luis.gonzalez.montana@unillanos.edu.co; lagonzalezmo@unal.edu.co
https://orcid.org/0000-0002-9136-9932

Nicolas Matentzoglu
Semanticly Ltd, London, UK.
nicolas.matentzoglu@gmail.com
https://orcid.org/0000-0002-7356-1779

Aaron D. Smith
Department of Entomology, Purdue University, West Lafayette, Indiana, USA.
pimeliinae@gmail.com; smit3866@purdue.edu
https://orcid.org/0000-0002-1286-950X

Markus Koch
University of Bonn, Institute of Evolutionary Biology and Ecology, An der Immenburg 1, 53121 Bonn, Germany.
mkoch@evolution.uni-bonn.de
https://orcid.org/0000-0003-3247-2988

Brendon E. Boudinot
University of California, Davis, One Shields Ave, CA, USA.
Friedrich-Schiller-Universität Jena, Institut für Zoologie und Evolutionsforschung, Erbertstraße 1, 07743 Jena, Germany.
boudinotb@gmail.com
https://orcid.org/0000-0002-4588-0430

Patrice Bouchard
Canadian National Collection of Insects, Arachnids and Nematodes, Agriculture and Agri-Food Canada, 960 Carling Avenue, Ottawa, Ontario, K1A 0C6, Canada.
patrice.bouchard@canada.ca
https://orcid.org/0000-0001-7443-4547

Roger Burks

University of California, Riverside, Entomology Department, 900 University Ave. Riverside, CA, USA.
burks.roger@gmail.com; roger.burks@ucr.edu
https://orcid.org/0000-0003-3032-7939

Lars Vogt
TIB Leibniz Information Centre for Science and Technology, Welfengarten 1B, 30167 Hannover, Germany.
lars.m.vogt@googlemail.com
https://orcid.org/0000-0002-8280-0487

Matthew Yoder
University of Illinois, Illinois Natural History Survey, Champaign, IL, USA.
diapriid@gmail.com
https://orcid.org/0000-0002-5640-5491

David Osumi-Sutherland
European Bioinformatics institute (EMBL-EBI) WellcomeTrust Genome Campus, Cambridge UK.
davidos@ebi.ac.uk
https://orcid.org/0000-0002-7073-9172

Frank Friedrich
Universität Hamburg, institut fur Spezielle Zoologie, Hamburg, Germany.
frank.friedrich@uni-hamburg.de

Rolf Beutel
Friedrich-Schiller-Universität Jena, Institut für Zoologie und Evolutionsforschung, Erbertstraße 1, 07743 Jena, Germany.
rolf.beutel@uni-jena.de
https://orcid.org/0000-0002-0433-7626

István Mikó
University of New Hampshire, Durham, NH, USA.
istvan.miko@gmail.com; istvan.miko@unh.edu
https://orcid.org/0000-0003-2938-9075

*Corresponding author: Jennifer C. Girón, entiminae@gmail.com
Short title: Insect Anatomy Ontology

 doi:10.20944/preprints202201.0254.v1

ABSTRACT

The spectacular radiation of insects has produced a stunning diversity of phenotypes. During the last 250 years, research on insect systematics has generated hundreds of terms for naming and comparing those phenotypes. In its current form, this terminological diversity is presented in natural language and lacks formalization, which prohibits computer-assisted comparison using semantic web technologies. Here we propose a Model for Describing Insect Anatomical Structures (MoDIAS) which incorporates structural properties and positional relationships for standardized, consistent, and reproducible descriptions of insect phenotypes. We applied the MoDIAS framework in creating the ontology for the Anatomy of the Insect Skeleto-Muscular system (AISM). The AISM is the first general insect ontology that aims to cover all taxa by providing generalized, fully logical, and queryable, definitions for each term. It was built using the Ontology Development Kit (ODK), which maximizes interoperability with Uberon (Uberon multi-species anatomy ontology) and other basic ontologies, enhancing the integration of insect anatomy into the broader biological sciences. A template system for adding new terms, extending and linking the AISM to additional anatomical, phenotypic, genetic, and chemical ontologies is also introduced. The AISM is proposed as the backbone for taxon-specific insect ontologies and has potential applications spanning systematic biology and biodiversity informatics, allowing users to (1) use controlled vocabularies and create semi-automated computer-parsable insect morphological descriptions; (2) integrate insect morphology into broader fields of research, including ontology-informed phylogenetic methods, logical homology hypothesis testing, evo-devo studies, and genotype to phenotype mapping; and (3) automate the extraction of morphological data from the literature, enabling the generation of large-scale phenomic data, by facilitating the production and testing of informatic tools able to extract, link, annotate, and process morphological data. This system will allow for clear and semantically interoperable integration of insect phenotypes in biodiversity studies.

Keywords: Morphology, insects, biodiversity research, ontology development.

The ubiquitous distribution and stunning species richness of insects has generated a great diversity of phenotypes that fuel research in biodiversity, systematics, and various other biological fields. Roughly 90% of studies describing insect anatomy deal with structures related to the skeleto-muscular system (Deans et al. 2012a; Iyer et al. 2016; Adachi et al. 2020; Sommer 2020; Gotoh et al. 2021); the remaining 10%, in general, deal with the nervous system (e.g., Loesel et al. 2013), the midgut (e.g., Monteiro et al. 2014), the endocrine system (e.g., Page Jr and Amdam 2007), fat bodies (e.g., de Oliveira and Cruz-Landim, 2003), etc. Thousands of morphological terms referring to the insect skeleto-muscular system have historically emerged due to several general processes: 1) most basic terms (e.g., head, wings, legs, etc.) have been borrowed from vertebrate anatomy due to functional or positional similarity; 2) some terms have been created *de novo* to name exclusive insect (or arthropod) structures (i.e., sclerite, tergite); 3) many terms have been repeatedly adopted across distant insect lineages to name similar structures located in similar areas of the body (e.g., cercus in Diplura vs. cercus in Hymenoptera; Snodgrass 1935); 4) the continuous reassessment of insect morphology in light of new comparative or phylogenetic data, constantly changes terms and their definitions; and 5) often, the definition of a term in subsequent studies, as in the "telephone game", suffers from interpretational deviations, thereby, producing a significantly different meaning that may eventually become widely adopted.

The interplay of these term-generating processes brings two major persisting problems. First, numerous terms in the corpus seriously suffer from semantic ambiguities such as, homonymy (the same term is used for unrelated structures), polysemy (the same term is used for different but related -similar- structures) and synonymy (different terms with the same meaning) (Bolshoy and Lacková 2021). Second, many terms and definitions reflect the history of their usage rather than accurate anatomical concepts. Moreover, some terms refer to common spatio-structural properties, others refer to a common function or a common developmental or presumed common evolutionary origin, and some terms even refer to a mixture of these categories (Vogt et al. 2010). Consequently, interpreting and analyzing phenotypic data becomes unnecessarily difficult for non-experts and integrating phenotype data with other sources of data in the life sciences is very difficult and time-consuming.

These problems are compounded by our trend to see and characterize elements and developmental/evolutionary processes of the insect exoskeleton similarly to that of vertebrates (Snodgrass 1963), which resulted not only in misunderstanding of insect evolution and development, but also in an over-complicated system worsening the above-mentioned issues of insect morphological terminology. Bones, the main elements of the vertebrate skeleton, develop from well separated cell clusters into a complex scaffold mirroring the intricate 3D connections of cells with different function and origin (Bitsch and Bitsch 2002; Wang et al. 2017; Blumer 2021). Bones are connected to each other by different types of joints, whose accurate functioning requires the interplay of unrelated elements, including ligaments, articular cartilages, and synovial fluid (Blumer 2021).

In comparison, the insect exoskeleton, the cuticle, is a simple, acellular product of the single-layered outer epithelium, the epidermis (Hall 1975; Adler 2017; Denk-Lobnig and Martin 2020) and its movable elements, the sclerites, are only stiffer regions of the cuticle that are surrounded by more flexible ones of the same origin (conjunctivae), granting mobility (Fig. 1). Therefore, unlike the vertebrate system, the insect skeleto-muscular system should be simply modeled using clearly identifiable and consistently organized building blocks of the continuous cuticle (Fig. 2). These building blocks —sclerites, conjunctiva, and formative elements (Klass 2008; Klass and Matushkina 2012); Table 1)— can also be used as anatomical landmark entities (i.e., disjointed intrinsically identifiable anatomical entities; Young 1993) for identifying units of comparison across different species (i.e., non-evolutionary comparative homology assessment; Vogt 2017). It is worth pointing out that the distinction between building blocks is not always clear, since semi-membranous areas also occur (e.g., often parts of the epipharynx, the wing articulation), so everything is more or less a continuum (Fig. 2), with (gradually) different degrees of sclerotization. The concept of "building blocks" (i.e., unambiguously defined sclerites), may be justified for pragmatic reasons in most cases, but it is still a simplification.

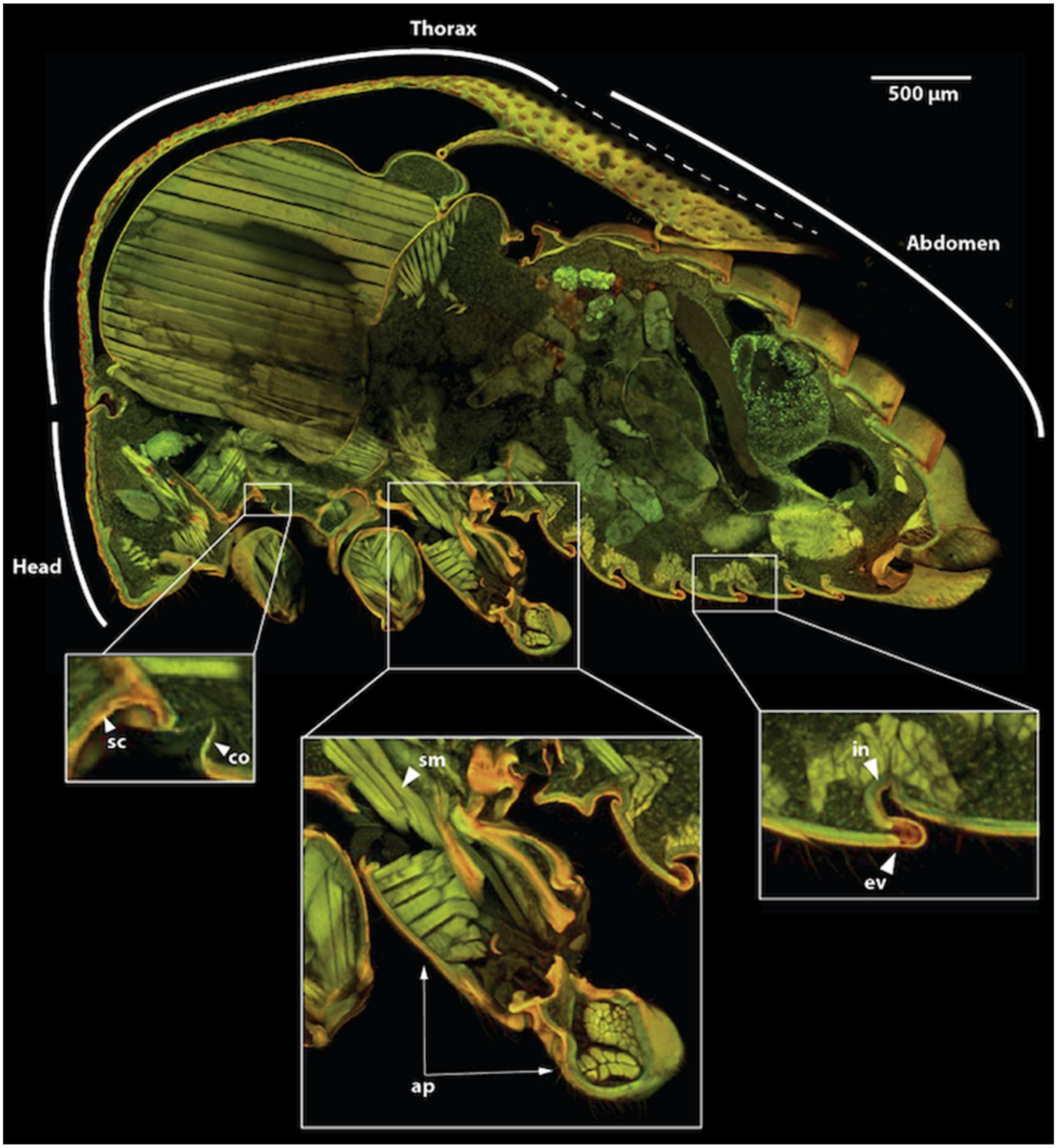

Figure 1. Autofluorescence-based CLSM micrograph showing the general structure of a sagittal section of the insect integument in an adult treehopper, genus *Ceresa* sp. (Membracidae). Excitation wavelength: 488, emission wavelengths: 500–580 pseudocolor green for conjunctivae, muscles, and other soft structures and 580–700 pseudocolor red for sclerotized components.

Abbreviations: sc: sclerite; co: conjunctiva; sm: skeletal muscle; ap: appendage; in: invagination; ev: evagination.

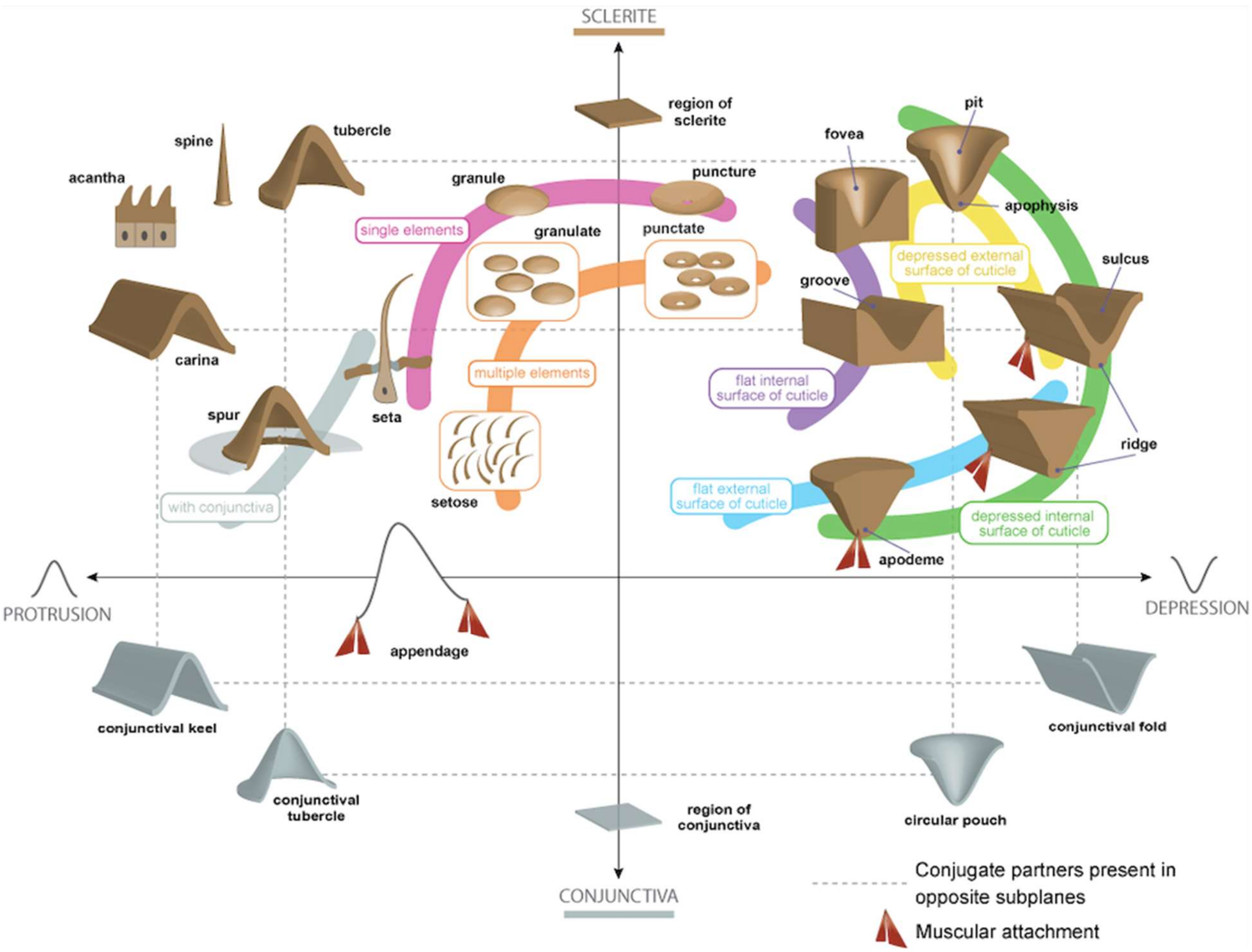

Figure 2. Schematic representation of some of the structural components of MoDIAS: a conceptual Model for Describing Insect Anatomical Structures. The two principal structural properties that characterize the insect cuticle are: 1) degree of flexibility (y axis), ranging from sclerite (stiff, at the top) to conjunctiva (flexible, at the bottom), and 2) degree of curvature (x axis; protrusion -left- to flat -center- to depression -right-); the degree of curvature of the external and internal surfaces of cuticular protrusions and depressions can be different: when both run in parallel, they form hollow protrusions (top left subplane, e.g., carina, tubercle) or hollow

depressions (top right subplane, yellow band e.g., pit, sulcus); the external surface can be depressed with the internal surface flat (top right subplane, purple band, e.g., fovea, groove); the external surface can be flat with the internal surface depressed (top right subplane, blue band, e.g., ridge, apodeme). Additional properties: quantity (single vs. multiple elements; pink and orange bands, respectively); shape can be observed throughout each subplane (e.g., sclerotized protrusions can range from elongated –carina– to rounded –tubercle–); same for depressions (elongate -groove- vs. rounded -fovea-).

Table 1. Cuticular elements and structural properties used in the AISM.

Subclass	Definition	URI
region of cuticle	The region of the insect integument (UBERON:6007284) that is part of chitin-based cuticle (UBERON:0001001)	AISM:0000174
sclerite	The region of the cuticle (AISM:0000174) that is less flexible than the neighboring conjunctiva(e) (conjunctiva(e) (AISM:0000004) that the sclerite is continuous with)	AISM:0000003
conjunctiva	The region of the cuticle (AISM:0000174) that is more flexible than the neighboring sclerite(s) (AISM:0000003) (sclerite(s) that the conjunctiva is continuous with)	AISM:0000004
cuticular depression	The region of the cuticle that corresponds to a concave surface	AISM:0000005

cuticular invagination	The region of cuticle (AISM:0000174) that corresponds with an invagination of the single layer epidermis (epithelial fold; UBERON:0005157). The cuticular invagination sometimes corresponds to a cuticular depression (concavity on the surface of the cuticle; AISM:0000005)	AISM:0000006
cuticular protrusion	The region of the cuticle that corresponds to a convex surface	AISM:0000008
cuticular evagination	The region of cuticle (AISM:0000174) that corresponds with an evagination of the cuticle and the single layer epidermis (epidermal fold; UBERON:0005157). The cuticular evagination usually corresponds to a cuticular protrusion (convexity on the surface of the cuticle; AISM:0000008)	AISM:0000027
anatomical region	A 3D region in space without well-defined compartmental boundaries; for example, the dorsal region of an ectoderm. [e.g., anterior region (BSPO:0000071); lateral region (BSPO:0000082); ventral margin (BSPO:0000684)]	BSPO:0000070

 doi:10.20944/preprints202201.0254.v1

somatic muscle	A muscle structure (UBERON:0005090) of invertebrates whose origin and insertion sites are in basal side of the epidermis or structures derived from it. The simplest somatic muscles consist of a single cell and associated extracellular structures.	UBERON:0014895

Ontologies have become a fundamental technology for semantic management and inference with biological knowledge (Smith et al. 2007; Balhoff et al. 2010; Deans et al. 2015; Dahdul et al. 2018; Tarasov 2019). An ontology is a logic-based representation of concepts and their relationships across a domain for modeling complex interactions in data (Deans et al., 2012a; Balhoff et al. 2013; Deans et al. 2015). In biology, ontologies serve two major purposes: they can be used as controlled vocabularies for stabilizing terminology and facilitate communication between scientists (Deans et al., 2012b), and as engines for inferring new complementary knowledge out of the encoded data. Therefore, ontology is a suitable technology for addressing the problems of understanding and interpreting terminology in insect morphology.

To date, there are seven ontologies dealing with the anatomy of different arthropod lineages, four of them dedicated to insects: Hymenoptera Anatomy Ontology (Yoder et al. 2010), *Drosophila* Anatomy Ontology (Osumi-Sutherland et al. 2013); *Tribolium* Ontology (Dönitz et al. 2013); and Mosquito Ontology (Topalis et al. 2008). However, given the narrow scope of each, none of them can be generally applied to insects as a whole. For the most part, these existing ontologies do not consider the interconnectedness of the whole cuticular system in their definitions, and those definitions tend to be idiosyncratic in the sense that they are taxon-specific and provide only textual/natural language definitions without much of a logical description, which prevents ontology-wide reasoning and inference.

One of the existing arthropod anatomy ontologies (Collembola Anatomy Ontology - CLAO; González-Montaña 2021a; https://github.com/luis-gonzalez-m/Collembola) has been successfully used in the production of semantic-based morphological descriptions (González-Montaña 2021b), demonstrating the potential for using anatomy ontologies in taxonomic research.

As a starting point towards a more stable, understandable, and interoperable terminology in insect morphology, in this study we provide a conceptual Model for Describing Insect Anatomical Structures (MoDIAS) at any developmental stage, in both formal and natural languages, and solely based on their structural properties and spatial relationships. We also provide the first universally applicable anatomy ontology for insects, the Anatomy Ontology of Skeleto-Muscular system (AISM), which is a formalized representation of MoDIAS that incorporates general terms for insect anatomy, including generalized definitions, while integrating them with other relevant ontologies. We provide ontology reasoning examples using the AISM and demonstrate its robustness and extensibility using the Ontology Development Kit (ODK). The AISM provides a computer-parsable controlled vocabulary for the insect skeleto-muscular system with a broad range of applications, including serving as a backbone for taxon-specific ontologies, providing opportunities to data mine the existing literature, as well as producing semantically enhanced descriptions; it also has the potential for integration in evo-devo research, phenotype to genotype mapping, and logical homology assessment analyses.

METHODS

Model for Describing Insect Anatomical Structures (MoDIAS)

The Model for Describing Insect Anatomical Structures (MoDIAS) incorporates structural properties and positional relationships to characterize anatomical structures used in morphological descriptions involving the insect skeleto-muscular system. We adopted anatomical concepts from Richards and Richards (1979), Snodgrass (1963), and Klass (2008) regarding the structural properties of the insect cuticle to define elementary building blocks (Fig. 2). For the positional relationships, we use terms referring to relative position of a given block along the body axes (dorsal, lateral, distal, etc.) and its connectedness to other structures (e.g., continuity, attachment). Using this model, each insect anatomical structure can be described and defined as one or more building blocks that are specifically related to other building blocks.

Creating and editing the AISM

The ontology for the Anatomy of the Insect Skeleto-Muscular system (AISM.owl) and accompanying file system were generated using the Ontology Development Kit (ODK, Matentzoglu et al. 2021; https://github.com/INCATools/ontology-development-kit) and edited with Protégé version 5.5.0 (Musen 2015). All the files are available on GitHub at https://github.com/insect-morphology/aism (Girón et al. 2021a). The ODK uses ROBOT-based workflows (Jackson et al. 2019; http://robot.obolibrary.org/) to automatically generate imports from related external ontologies including Uberon (Mungall et al. 2012) and the OBO relations ontology (Smith et al. 2005), and to drive quality control tests under continuous integration. It also provides a semi-automated release process supporting the generation of release products enhanced by the results of OWL reasoning.

Throughout this text we use **bold lettering** to indicate ontology classes, *italics* when referring to object properties, and use ID numbers to specify each. ID numbers are composed of the ontology prefix followed by colon and a number e.g., AISM:0000003. This ID represents a link (http://purl.obolibrary.org/obo/AISM_0000003) to an online version of the encoded information.

Following the principles proposed by MoDIAS, we created terms referring to the elementary building blocks of the insect skeleto-muscular system, as well as generalized terms from the glossary presented by Beutel et al. (2014). Each term has a *label* and a series of specific annotation properties including *sensu* (AISM:0000171), *definition*, *contributor*, and *has exact synonym* (Table. 2).

Table 2. Main annotation properties used in the AISM.

Annotation	AISM usage	URI
label	A term indicated by a word or set of words to unambiguously name an insect anatomical structure.	https://www.w3.org/2000/01/rdf-schema#label

definition	A natural language statement to describe an insect anatomical structure, constructed by articulating the appropriate subclass of descriptors.	http://purl.obolibrary.org/obo/IAO_0000115
has exact synonym	Alternative labels applied to the defined insect anatomical structure. Should be accompanied by a *sensu* annotation.	http://www.geneontology.org/formats/oboInOwl#hasExactSynonym
sensu	Bibliographic reference with its corresponding DOI (or other link to it), and the textual definition of the term according to that reference.	http://purl.obolibrary.org/obo/AISM_0000171
contributor	The person who composed the definition or added the subclass of descriptor	http://purl.obolibrary.org/dc/elements/1.1/contributor
creation date	The date when the definition was composed in year-month-day format.	http://geneontology.org/formats/oboInOwl#creation_date
date_modified	Date on which the resource was changed.	http://purl.org/dc/terms/modified
curator note	& Additional comments to clarify or expand on the presented definition. Should be accompanied by *contributor* and *creation_date*.	http://purl.obolibrary.org/obo/IAO_0000232

 doi:10.20944/preprints202201.0254.v1

foaf:depiction	Associated image or images illustrating the structure being defined, linked by a DOI or URL.	http://xmlns.com/foaf/0.1/depiction

Each AISM term is unambiguously labeled and formally represented by as many subclass of descriptors as necessary to clearly characterize the term using object properties and associated classes. Terms from existing general anatomy ontologies [e.g., Uberon multi-species anatomy ontology (Mungall et al. 2012); BFO: Basic Formal Ontology (Spear et al. 2016)], and supporting ontologies [e.g., BSPO: Biological Spatial Ontology (Dahdul et al. 2014); CARO: Common Anatomy Reference Ontology (Haendel et al. 2008); PATO: Phenotype And Trait Ontology (Gkoutos et al. 2005using; RO: Relation Ontology (Mungall et al. 2021)] were imported using the ODK.

An effort was made to maximize the inclusion of existing terms, avoiding duplication of existing object properties and general higher classes; we made sure that the class definitions offered in existing ontologies were compatible with the intended usage in the AISM before importing a class. Each term of the AISM is accompanied by a verbatim logical definition that translates each set of subclass of descriptors into natural language. Each definition in the AISM is intended to be broad enough to be applicable across Insecta, in a similar way as Uberon provides generalized definitions for animals (Mungall et al. 2012). As a convention, labels for muscles (which are given in English) and conjunctivae are given using their sites of attachment from proximal to distal, anterior to posterior, or dorsal to ventral.

We used *continuous with* (RO:0002150) for sclerite-conjunctiva attachments, whereas *adjacent to* (RO:0002220) for sclerite-sclerite articulations. We also propose the object properties *encircles* (AISM:0000078) and *encircled by* (AISM:0000079) to annotate the relationship between ring sclerites and their corresponding conjunctivae (e.g., femur, antennomere).

Reasoning

We used a template (https://github.com/insect-morphology/aism/blob/master/AISM_template_examples.tsv) to create AISM-based instances and definitions to demonstrate how the terms and generalized definitions provided in the AISM can be used to fit insect taxon-specific definitions more closely. In this template we represented different paired cuticular structures of the abdominal tergites as individuals (instances), for the orders Archaeognatha, Zygentoma, Dermaptera, Ephemeroptera (Baetidae), Hemiptera (Aphididae), Psocodea, and Coleoptera (Carabidae larva) (Table 3; see Fig. 3 for a schematic representation of these definitions). Using ROBOT (Jackson et al. 2019; http://robot.obolibrary.org/) we generated an OWL file from this template, which included terms from the AISM and other ontologies. This template-based OWL file was then merged with the AISM. We ran a series of DL queries in Protégé, using ELK 0.5 as a reasoner on this merged ontology to verify the fit of the provided taxon-specific definitions with the terms and definitions available in the AISM (Table 4). The expectation was that the queries would return the appropriate instances, depending on the properties indicated in the template.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 January 2022 doi:10.20944/preprints202201.0254.v1

Table 3. Example of template to specify new terms to include in AISM or AISM-derived ontologies. This template can be extended to add more descriptors. See https://github.com/insect-morphology/aism/blob/master/AISM_template_examples.tsv.

Term suggestion	definition	type of cuticular element	location	bilaterality	anterior to	lateral to	has part
cercus_ Archaeognatha	The paired protrusion of the dorsal region of the postabdomen that is anterior to the anus and composed of cercomeres.	'cuticular protrusion'	'dorsal postabdomen'	'bilaterally paired'	'insect anus'		'cercomeres'
cercus_ Zygentoma	The paired appendage of the dorsal region of the postabdomen that is anterior to the anus	appendage	'dorsal postabdomen'	'bilaterally paired'	'insect anus'		'cercomeres'

	and composed of cercomeres.						
cercus_ Dermaptera	The paired appendage of the dorsal region of the postabdomen that is anterior to the anus and composed of cercomeres.	appendage	'dorsal postabdomen'	'bilaterally paired'	'insect anus'		'cercomeres'
tergalius_ Ephemeroptera	The paired appendage of the dorsal region of the preabdomen that is lateral to the abdominal tergite.	appendage	preabdomen	'bilaterally paired'		'abdominal tergite'	
cercus_ Ephemeroptera	The paired appendage of the dorsal region of the postabdomen that	appendage	'dorsal postabdomen'	'bilaterally paired'	'insect anus'		'cercomeres'

 doi:10.20944/preprints202201.0254.v1

	is anterior to the anus and composed of cercomeres.						
cornicle_Aphididae	The paired cuticular protrusion of the dorsal region of the preabdomen.	'cuticular protrusion'	preabdomen	'bilaterally paired'			
cercus_Psocodea	The paired region of the dorsal region of the postabdomen that is anterior to the anus and composed of a collection of setae.	'region of cuticle'	'dorsal postabdomen'	'bilaterally paired'	'insect anus'		'collection of setae'
urogomphus_ Carabidae	The paired cuticular protrusion of the dorsal region of the	'cuticular protrusion'	'dorsal postabdomen'	'bilaterally paired'	'insect anus'		

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 18 January 2022 doi:10.20944/preprints202201.0254.v1

	postabdomen that is anterior to the anus.

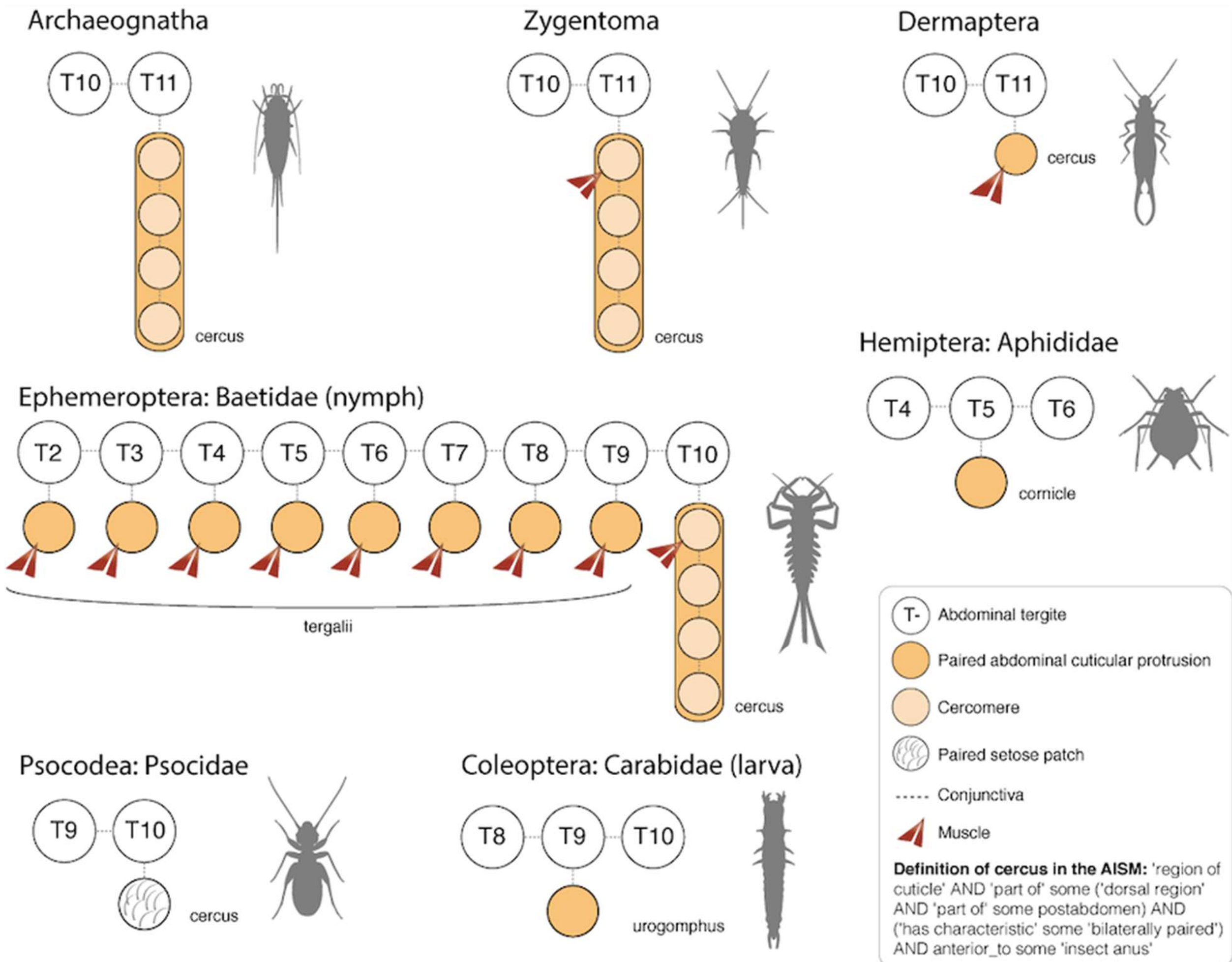

Figure 3. Schematic representation of taxon-specific definitions for paired cuticular structures of the abdominal tergites (not including structures of the genitalia) across different insect orders. Textual definitions for each structure are provided in Table 3.

Table 4. Example DL queries and their results using the AISM.

DL Query	**Resulting subclasses and individuals**
Internal queries	
'cuticular protrusion' and ('part of' some 'cuticle of insect head')	**Classes:** 'cuticle of insect mandible', 'cuticle of insect maxilla', 'cuticle of labial palpus'

	'cuticle of maxillary palpus', cuticle of antenna, antennifer, cuticle of galea, cuticle of glossa, cuticle of labium, cuticle of labrum, cuticle of lacinia, cuticle of ligula, cuticle of mouthpart, cuticle of paraglossa
'cuticle of insect appendage' and ('part of' some 'cuticle of insect thorax')	**Classes:** 'cuticle of fore leg', 'cuticle of fore wing', 'cuticle of hind leg', 'cuticle of hind wing', 'cuticle of insect leg', 'cuticle of insect wing', 'cuticle of mid leg', 'cuticle of mesopretarsus', 'cuticle of metapretarsus', 'cuticle of pretarsus', 'cuticle of propretarsus'
'part of' some 'cuticle of fore leg'	**Classes:** 'procoxal-protrochanteral conjunctiva', 'profemoro-protibial conjunctiva', 'protibio-protarsal conjunctiva', 'protrochantero-profemoral conjunctiva', 'cuticle of procoxa', 'cuticle of profemur', 'cuticle of propretarsus', 'cuticle of protarsus', 'cuticle of protibia', 'cuticle of protrochanter'
'cuticle of appendage segment' and 'part of' some 'cuticle of antenna' and 'adjacent to' some 'head capsule'	'cuticle of scapus'
Queries for taxon-specific definitions	

'part of' some 'cuticle of insect abdomen' and 'region of cuticle' and 'has characteristic' some 'bilaterally paired'	**Classes:** 'cuticle of gonocoxa IX', 'cuticle of gonocoxa VIII', 'cuticle of gonostylus IX', 'cuticle of gonostylus VIII', cuticle of cercus, cuticle of paramere **Instances:** cercus_Archaeognatha, cercus_Dermaptera, cercus_Ephemeroptera, cercus_Psocodea, cercus_Zygentoma, cornicle_Aphididae, tergalius_Ephemeroptera, urogomphus_Carabidae
'part of' some preabdomen and 'region of cuticle' and 'has characteristic' some 'bilaterally paired'	**Instances:** cornicle_Aphididae, tergalius_Ephemeroptera
'part of' some postabdomen and 'region of cuticle' and 'has characteristic' some 'bilaterally paired'	**Classes:** 'cuticle of gonocoxa IX', 'cuticle of gonocoxa VIII', 'cuticle of gonostylus IX', 'cuticle of gonostylus VIII', cuticle of cercus, cuticle of paramere **Instances:** cercus_Archaeognatha, cercus_Dermaptera, cercus_Ephemeroptera, cercus_Psocodea, cercus_Zygentoma, urogomphus_Carabidae

'part of' some postabdomen and 'cuticular protrusion' and 'has characteristic' some 'bilaterally paired'	**Classes:** 'cuticle of paramere'
	Instances: cercus_Archaeognatha, cercus_Dermaptera, cercus_Ephemeroptera, cercus_Zygentoma, urogomphus_Carabidae
'part of' some postabdomen and 'cuticular protrusion' and 'has characteristic' some 'bilaterally paired' and 'anterior to' some 'insect anus'	**Classes:** none
	Instances: cercus_Archaeognatha, cercus_Dermaptera, cercus_Ephemeroptera, cercus_Zygentoma, urogomphus_Carabidae
'part of' some postabdomen and 'cuticle of insect appendage' and 'has characteristic' some 'bilaterally paired' and anterior_to some 'insect anus'	**Classes:** none
	Instances: cercus_Dermaptera, cercus_Ephemeroptera, cercus_Zygentoma

In addition, to demonstrate the interoperability of the AISM with existing ontologies, we provide an example on how to describe a particular insect species phenotype. The ability to relate structures across additional ontologies was also illustrated by linking structures of the AISM with the circulatory system using relationships from the Relation Ontology (RO) and terms from the Ontology of Arthropod Circulatory Systems (OArCS).

RESULTS AND DISCUSSION

MoDIAS: Model for Describing Insect Anatomical Structures

MoDIAS is a descriptive model: a set of principles based on traditional approaches to classify and characterize cuticular elements based on their structural properties and relative position. It is proposed as a baseline to generate consistent and reproducible descriptions of insect skeleto-muscular structures.

The insect endo- and exoskeleton is a continuous entity that can be considered as a single anatomical structure (Klass 2008; similar to the skin of a vertebrate). Cuticular elements can be defined and distinguished from each other by variations along five key properties: 1) degree of flexibility (i.e., stiffness or resistance to deflection: sclerite vs. conjunctiva), 2) degree of surface curvature (i.e., deviations from a flat surface: depression vs. protrusion), 3) presence of muscular attachments, 4) quantity (single vs. multiple repeated elements), and 5) shape (circular, elongate). The interplay of these properties determines the features of each elementary building block that, together with its spatial relations and connectedness, allows for modelling of the entire structural diversity of the insect skeleto-muscular system. Similar categorical sets of properties have been employed in other semantic descriptive models of anatomical systems (e.g., OArCS, Ontology of Arthropod Circulatory System; Wirkner et al. 2017) to allow for better data structuring.

Based on their degree of flexibility there are two main regions of the cuticle: 1) sclerites, which are relatively stiff, and 2) conjunctivae which are relatively flexible and provide mobility (Klass and Matushkina 2012; Fig. 2, y axis). Even though these different regions of the cuticle are often characterized by their histological properties (sclerites with thick exocuticle with sclerotin vs. conjunctiva with thin exocuticle without sclerotin; Beutel et al. 2014), these are not discernible without histological sections.

Regions of the cuticle are also classified by the degree of curvature of their internal and external surfaces (Fig. 2, x axis and color bands on top right subplane), which not only provide information about external components but also allows for the linking of structural variations of the underlying epidermis. The external surface can be flat, convex (cuticular protrusions) or concave (cuticular depressions). Cuticular protrusions, if they correspond to evaginations of the cuticle (i.e., cuticular protrusion corresponding to an internal cuticular depression; Fig. 2, top left subplane), can correspond to either the evagination of a single cell membrane (e.g., seta) or to the evagination of a region of the single-layered outer epithelium (e.g., spurs, lobes) (Richards and Richards 1979). Appendages differ from other cuticular protrusions (e.g., spurs or lobes) in that they are connected to the rest of the body by somatic muscles (Fig. 2). When individual elements like a carina or a seta are repeated across a region of the cuticle they generate texture on that particular surface, forming sculpture or pilosity, respectively. Ring sclerites often represent repetitive subdivisions of appendages that can be either musculated (appendage segments) or non-musculated (meres).

Cuticular depressions (Fig. 2, top right subplane) vary in constitution depending on the orientation of the external and the internal surfaces of the cuticle: when both run in parallel, they form hollow depressions (e.g., pit, sulcus; Fig. 2, top of green band); the external surface can be depressed, with the internal surface flat (e.g., fovea, groove; Fig. 2, purple band); or the external surface can be flat, with the internal surface depressed (e.g., ridge, apodeme; Fig. 2, blue band, bottom of green band); this particular kind of cuticular depression forms strengthened areas across the body, providing mechanical stability, and frequently constitute sites for muscle attachment (Klass and Matushkina,2012; Beutel et al. 2014).

Each of these elementary building blocks with their particular features can be specifically characterized by their connections and spatial relations to other elementary building blocks, including positional relationships (e.g., dorsal, ventral, distal, proximal, medial, lateral), connectedness (e.g., continuous with, encircled by, adjacent to) and further phenotypic descriptors (color, relative size). This specific characterization results in accurate, consistent, and reproducible descriptions of insect anatomical structures. If employed correctly, MoDIAS-based natural language definitions should be easily translated into logical definitions and instance-based semantic phenotype descriptions of individual specimens using ontologies, so that information of the insect skeleto-muscular system can be accessible for machine processing.

AISM: Ontology for the Anatomy of the Insect Skeleto-Muscular system

In its version 2021-12-13, the ontology for the Anatomy of the Insect Skeleto-Muscular system (AISM) contains 1647 classes, where 532 are AISM terms, from which 384 are subclasses of **insect region of integument** (UBERON:6007284); it uses 24 object properties and

20 annotation properties. All other terms have been imported from existing ontologies as part of the basic imports using the Ontology Development Kit (ODK; see methods), which not only brings the specified terms, but also all their hierarchically associated terms to preserve the logical integrity of the ontology and maximize interoperability.

The insect integument, as a continuous structure, is composed of **chitin-based cuticle** (UBERON:0001001); therefore, every component of this continuous structure is designated as a **region of cuticle** (AISM:0000174), which is the parent class for all skeletal anatomical structures in the AISM (Fig. 4). Interpreting the skeleto-muscular system of insects as a set of consistently organized components and following the framework proposed by MoDIAS, each class included in the AISM is defined logically in OWL by some combination of: (1) kind of cuticular element [e.g., **sclerite** (AISM:0000003), **conjunctiva** (AISM:0000004), **cuticular depression** (AISM:0000005), **cuticular protrusion** (AISM:0000008), **skeletal muscle tissue** (UBERON:0001134), among others]; (2) location of structure in the body [e.g., *part of* the **insect thorax** (AISM:0000108), *anterior to* the **abdominal tergite I** (AISM:0000021)]; (3) connected structures indicated by specific relations and spatial descriptors [e.g., *adjacent to* **posterior margin** (BSPO:0000672) of **abdominal sternite III** (AISM:0004105)]. In this way the continuous nature of the insect integument is considered, making explicit statements about connectivity between parts and providing positional/spatial localization for each structure.

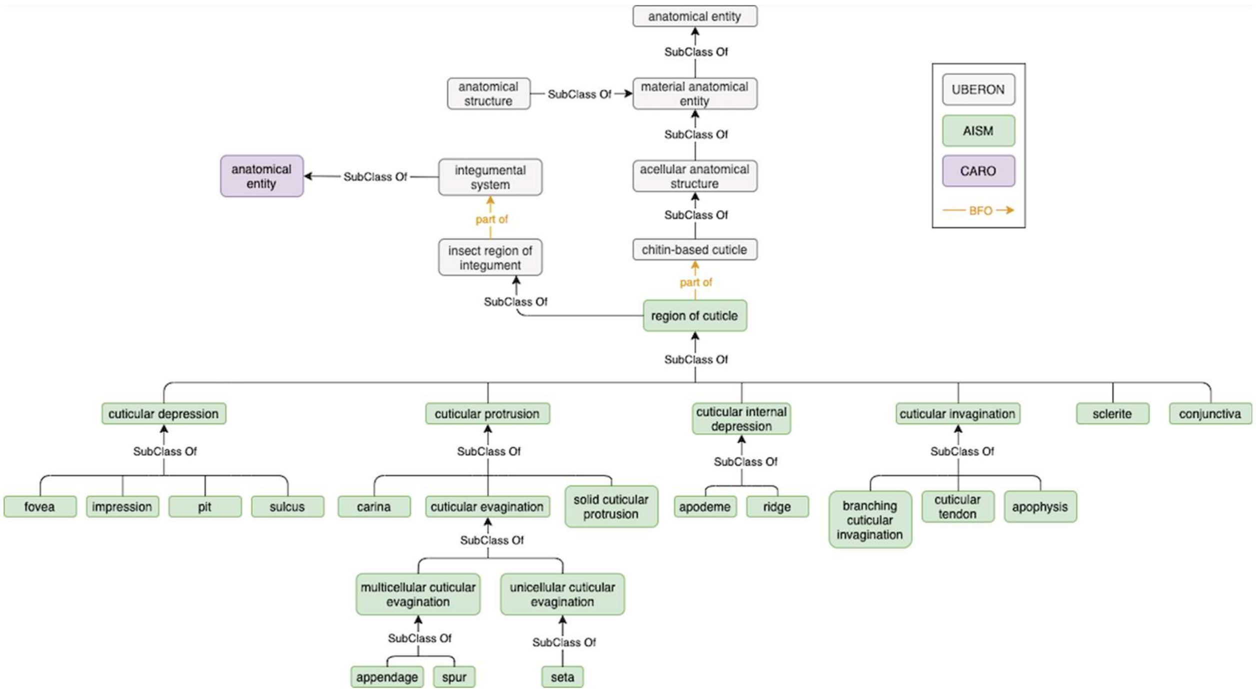

Figure 4. Graphic representation of high-level classes in AISM and some of their children, including hierarchy and elements from other ontologies.

In addition, each class (i.e., label and descriptors) is accompanied by a set of annotation properties including a natural language definition that has been created from the annotated descriptors (or vice versa). These definitions include the contributor who constructed the definition and the date of creation and date of modification in format year-month-day. When available, references for textual definitions from the literature have been annotated on each label using the annotation property *sensu* (AISM:0000171), which includes the full citation of the reference, a DOI or link when the reference is available online, and the verbatim definition provided in the text, in quotation marks (see Yoder et al. 2010). When explicitly mentioned in the literature, synonyms are added using the has_exact_synonym property indicating the *sensu* where this synonymy is proposed.

Using templates to curate and extend the AISM

In order to make the AISM maximally accessible and re-usable, the AISM aims for MoDIAS-based definitions that follow consistent, simple patterns. To ensure that users of the AISM can also easily generate MoDIAS/AISM-compliant descriptions of anatomical structures, we provide a template system for composing definitions (i.e., https://github.com/insect-morphology/aism/blob/master/AISM_template_examples.tsv). This template system can be used to provide formal descriptions of insect anatomical structures, or for extending the AISM with taxon-specific terms (subclasses). Even if the templates are not directly used, they provide guidance for the types of definitions that are compatible with the MoDIAS/AISM approach to defining terms.

The aim of the templates is to ensure that users provide the specific type of cuticular element (Fig. 2) and its appropriate location within the insect body. Users may further refine the location by specifying the structure's relative location via multiple statements using relations such as *adjacent to*, *posterior to*, and *dorsal to*. The template also includes a free text comment column allowing additional information to be provided in a less formal manner. Once this detailed, MoDIAS-compliant description is provided, users may also propose a commonly used term for the described structure, such as 'cercus'. The advantage of this approach is that it forces users to provide an accurate description of the structure's properties and location not captured by the generally used term.

Even without additional processing, the filled-out template constitutes a shareable and accessible controlled description of anatomical structures. Because the template corresponds to a standard OWL template system, it can also be used to generate new terms or instances, either for the AISM, or for extending the AISM with taxon-specific terms.

In table 3 we present a few examples of terms to refer to different paired structures of the abdomen of different insect taxa. For instance, the cercus of Archaeognatha, defined as the paired protrusion of the dorsal region of the postabdomen that is anterior to the anus and composed of cercomeres is indicated in the template specifying its type of cuticular element (cuticular protrusion, AISM:0000008), its location (dorsal postabdomen, AISM:0000523), its laterality (bilaterally paired, PATO:0040024), its position regarding other structures (anterior to [BSPO:0000096] insect anus [AISM:0004197]), and its composition (cuticle of cercomeres, AISM:0004199). As examples are specified, OWL reasoning can be used to provide a list of candidate terms in AISM that conform to the definition.

The current implementation of the templates relies on users following the specification. In the future, we plan to make use of the CEDAR template system (https://more.metadatacenter.org/tools-training/cedar-template-tools/#design-template) to provide auto-completion and constraints on column content, guiding and constraining users to ensure the correct types of terms are added in each column. We also plan to integrate a term suggestion option to avoid replication and detect potential synonyms.

Reasoning with the AISM

Each of the queries tested returned the expected outcomes in terms of subclasses and individuals (Table 4). Across Insecta, abdominal protrusions are highly variable in position, shape, and components, and in many cases the morphological interpretations of these structures and their features have been problematic over time. The terms and broad definitions presented in the AISM have the capability to incorporate the broad variation presented in our example taxa. By adding subclasses and relationships to AISM terms it is possible to characterize taxon-specific structures. For instance, the cornicles of Aphididae (Hemiptera) are paired cuticular protrusions located on the dorsal surface of the abdominal tergite 5 (sometimes abdominal tergite 6); the existing terms and definitions incorporated in the AISM allow for accommodating all these details into a definition for cornicle in a potential Hemiptera-specific ontology. Similarly, the tergalii of Ephemeroptera, paired appendages of the preabdomen located on the lateral region of the abdominal tergites (Kluge 2004) can be easily defined, and the particular abdominal tergites where the tergalii are present could be specified.

The different kinds of cerci present in our example taxa were also easily characterized, as they follow the generalized definition proposed in the AISM (the bilaterally paired region of the cuticle of the dorsal region of the postabdomen, that is anterior to the anus): in Dermaptera the cercus was characterized as an appendage (with muscular attachment) and composed of a single cercomere (Fig. 3). In Psocodea it was defined as a region of the cuticle that is anterior to the anus and bears a collection of setae (Fig. 3); we followed the definition presented by Yoshizawa (2005), even though the definition of this particular surface (anterior to the anus) in Psocodea as cercus has been contentious. In Zygentoma and Ephemeroptera the cercus was characterized as a paired appendage (with muscular attachment) of the dorsal region of the postabdomen, composed of cercomeres, whereas in Archaeognatha, the cercus is a cuticular protrusion (without muscular attachment) of the dorsal region of the postabdomen, composed of cercomeres. Our queries also returned the Coleoptera urogomphus as a cercus, as this undivided cuticular protrusion satisfies the requirements of the AISM definition for cercus. This demonstrates the power of the AISM's homology free approach, as these urogomphi are structurally equivalent to cercus, but not homologous, a similarity that would be obscured if we only relied in homology-biased terminology. On the other hand, the same query did not recover the cornicles, as these are paired projections of the abdomen, but located in a different abdominal region. The amount of detail incorporated into each definition will depend on the intended use of the ontology. Indeed, the number and sequence of cercomeres can be specified, along with the presence of setae, scales, or other relevant features.

Describing phenotypes with the AISM

The AISM has been conceived as the backbone ontology for taxon-specific ontologies. The broadly applicable terms it contains are intended as superclasses for terms defining more taxon-specific anatomical structures. For instance, in an AISM-based Coleoptera-specific ontology, the class **elytron** would be created as a subclass of the class **fore wing** (AISM:0000037), adding descriptors related to specific properties of elytron, such as subclass of the class **sclerite** (AISM:0000003). Similarly, in an AISM-based Diptera-specific ontology the class **haltere** would be a subclass of the class **hind wing** (AISM:0000038).

In general, for describing specific insect phenotypes with the AISM, a series of Entity-Quality statements can be used (e.g., Washington et al. 2009), taking advantage of the high interoperability of the AISM and the broad range of available existing ontologies including those for phenotype (PATO: Phenotype And Trait Ontology), taxonomy (NCBITaxon: National Center for Biotechnology Information organismal classification), spatial relationships (BSPO: Biological Spatial Ontology), among others. The template system proposed here can accommodate additional descriptors and relationships to better define structures within the AISM.

It is possible to represent phenotypes like a yellow profemur on a chalcid wasp: [(Chalcididae [NCBITaxon:92425] AND *has part* [BFO:0000051] some **profemur** [AISM:0000070]) AND (*has characteristic* [RO:0000053] some **yellow** [PATO:0000324])]. Using the template system, an additional column for color would be needed.

In our exercise linking AISM to OArCS, no new terms were required, just additional linkages between existing terms and relationships, for example, the **alary muscle** (OARCS:0000151) is *attached to* (RO:0002371) both the **heart** (OARCS:0000253) and the **abdominal tergite** (AISM:0004057). The term **alary muscle** would be imported using the ODK, bringing the necessary hierarchically linked terms and properties to be able to construct the logically appropriate axioms.

Taxon-specific ontologies can be linked to specialized taxonomic ontologies if those were available (e.g., see Stucky 2019). An example of taxonomic ontologies is the Vertebrate Taxonomy Ontology (VTO; Midford et al. 2013), which provides a comprehensive taxonomic hierarchy for vertebrates. It incorporates classes from the Taxonomic rank vocabulary (http://www.obofoundry.org/ontology/ncbitaxon.html) and the NCBI organismal classification

(National Center for Biotechnology Information; http://www.obofoundry.org/ontology/taxrank.html).

All these approaches to phenotypic descriptions can be implemented using and extending the proposed template system. Furthermore, integration between the AISM and existing ontologies like Uberon, the *Drosophila* Anatomy Ontology (FBBT; Costa et al. 2013), and the Hymenoptera Anatomy Ontology (HAO; Yoder et al. 2010) can be improved over time by adding cross-reference annotations to each shared term.

Taxonomy, morphology, and evo-devo: AISM on different granularity levels

Similar surface modifications of the insect skeleto-muscular system can correspond with cardinally different epithelial modifications: multicellular invaginations/evaginations of the epidermal cell layer (e.g., cuticular depressions, spurs, pits, and appendages), invaginations/evaginations of a single cell membrane (e.g., cuticular component of sensilla), changes in the thickness of the cuticle (i.e., modifications that do not correspond to any epithelial fold, e.g., impression, acantha or carina), and in some cases the combination of these categories.

Changes in the geometry of the epithelial sheet resulting in invaginations and evaginations are governed by genes that define changes in the shape of epidermal cells or regulate cell proliferation (Zartman and Shvartsman 2010; Hannezo et al. 2014; Gotoh et al. 2021) while genes involved in the reorganization of the cytoskeleton are governing similar geometrical changes of the membrane of a single cell (Lees and Waddington 1942; Bitan et al. 2012; Djokic et al. 2020). A third set of genes are involved in surface characteristics that are related to cuticle thickness, which are related to processes regulating cuticle deposition (Adler 2017; Jan et al. 2017; Tajiri, 2017; Zhao et al. 2017). These processes are also separated in time and space; evaginations and invaginations happen during the last larval and early pupal stages, while cuticle deposition starts in the late pupal stage (Andersen 2012).

It is evident that differentiating these superficially similar structures would be key to accurately understanding phenotypic diversity and morphological evolution. However, the differences between practical approaches to anatomy across different knowledge domains represent a huge communication gap that hinders progress towards a more integrative view of anatomy (Richards and Richards 1979): (1) morphology aims to interpret the structural identity and connectivity of anatomical structures (Snodgrass 1951); it uses dissections and section-based methods ranging from histology to μ-CT, and usually focuses on a handful of specimens in each study; (2) taxonomy focuses for the most part on externally visible structures with diagnostic value; each study can involve thousands of specimens in a comparative framework; (3) evo-devo studies gene expression on developing structures; the taxonomic breadth is usually limited to model organisms that are reared under laboratory conditions.

These knowledge domains refer to anatomy at different granularity levels and from different frames of reference (Vogt 2019) across different shared themes (i.e., taxonomic, individual count, developmental stage), which causes interoperability problems and misunderstanding among disciplines, due to concept shifting of anatomical entities. The AISM provides a bridge, in the form of a controlled vocabulary, to facilitate communication, by using an interconnected hierarchy of superficial cuticular elements (anatomical surfaces) and the hierarchy of deeper structures that reveal developmental and structural properties of the single-layered outer epithelium. The AISM provides an opportunity to link insect phenotypes to genotypes across developmental stages and taxonomic groups via the Gene Ontology (Ashburner 2000), and to metabolic processes via the Protein Ontology (PRO; Natale et al. 2017) and the ontology for Chemical Entities of Biological Interest (CHEBI; Hastings et al. 2015).

Homology and the AISM

Evolutionary homology is a central concept in biology, whereby structural similarity has evolved through shared ancestry in different taxa (Minelli and Fusco 2013; Wagner 2014). The definitions of the classes included in the AISM are merely descriptive in anatomical terms and do not constitute homology statements, although they may serve to assess the primary criteria of position and similarity (see comparative homology, units of comparison; Vogt 2017), so that instead of asking whether the cercus in Archaeognatha is homologous to the cercus in Psocodea, we can ask if a multisegmented appendage on the 11th tergite in one group is homologous with a setose patch on the 10th tergite of the other (see our examples for abdominal cuticular protrusions, Table 3). There are data models such as the one proposed by Mabee et al. (2020), where homology relationships can be logically formalized between anatomical structures of different taxonomic units. This approach requires elements from anatomy ontologies, taxonomic ontologies, and the Evidence and Conclusion Ontology (ECO, Chibucos et al. 2014). Under this scenario, the AISM would serve as one of the components required to assess homology statements across different taxonomic groups of insects.

Bringing insects into the Phenomic Era

Nomenclatural rules require that the establishment of an animal taxon new to science "be accompanied by a description or definition that states in words characters that are purported to differentiate the taxon" or "by a bibliographic reference to such a published statement" (Article 13.1 in ICZN 1999). Taxonomists have described over one million species of insects worldwide (Stork 2018). These descriptions constitute vast amounts of information that could be efficiently mined, compared, interpreted, and analyzed, just like any large molecular dataset nowadays. However, these phenotypic descriptions are presented in non-standard natural language form, and are therefore, inaccessible for machine interpretation (Balhoff et al. 2010). Ontologies and knowledge graphs offer a system to represent entire knowledge domains in an organized, standardized, consistent, and logical manner, so that information can be processed and quality-checked by computers (Arp et al. 2015).

There are informatic tools that allow data extraction from the literature, based on XML markup, which have been used primarily for extracting taxonomic information from PDF files (Penev et al. 2011). For instance, GoldenGATE-Imagine (Sautter et al. 2007; https://github.com/plazi/GoldenGATE-Imagine), which is used by Plazi (http://plazi.org). There are also tools that use ontologies for annotating anatomical, phenotypic, and taxonomic data (Phenex; Balhoff et al. 2010). Lücking et al. (2021) provide an overview of methods for semantic annotation of bibliographic records and introduce a system to use multiple annotations for terms; the authors also introduce the BIOfid-portal (https://www.biofid.de/en/search/), which is an on-line tool for accessing the semantics of biodiversity texts in German. The annotation method proposed by Lücking et al. (2021) is partly based on the MATTER conceptual framework for annotations (Model, Annotate, Train and Test, Evaluate, and Revise; Pustejovsky & Stubbs 2012).

The AISM provides the key to annotate phenotypic information for insects, extracted from the literature. Combining or expanding these and similar informatic tools can generate large-scale phenotypic datasets, unlocking multiple avenues of research including, among others, genotype to phenotype associations, evo-devo studies, and the use of Artificial Intelligence and ontological inference (Jackson et al. 2018) to analyze morphological evolution across insects. Phenotypic data generated with the aid of the AISM would greatly contribute to increase links in the Biodiversity Knowledge Graph (Page 2013).

Availability

The AISM is available on GitHub at https://github.com/insect-morphology/aism (Girón et al. 2021a) as well as on the OBO Foundry at http://www.obofoundry.org/ontology/aism.html. The GitHub repository is open for collaborative editing. We provide a manual on how to edit the AISM (https://github.com/insect-morphology/Manual; (Girón et al. 2021b), including the use of templates, and how to use the AISM as the starting point for developing taxon-specific ontologies.

CONCLUSION

Here we provided a Model for Describing Insect Anatomical Structures (MoDIAS) that incorporates structural properties and positional relationships to define anatomical structures of insects, independent of developmental stage, homology assumptions, or taxonomic group. Following the set of principles established by MoDIAS, we created the first universally applicable anatomy ontology for insects, the Ontology for the Anatomy of the Insect Skeleto-Muscular system (AISM). The AISM provides a basic backbone of generalized and unambiguously labeled terms for the anatomy of the skeleto-muscular system of insects. Each term is accompanied by natural language definitions translated into sets of subclass of descriptors to provide logical definitions in the ontology. Built using the Ontology Development Kit, which is a free, open source, and OBO Foundry-supported system, the AISM is interoperable with existing ontologies in the biological sciences, open for editing and refinement, and extensible to tackle taxon-specific ontologies.

The AISM opens new opportunities for phenomic-scale research in biology by providing computer-parsable formalization and a controlled vocabulary for insect anatomy. The potential application of AISM spans all biological domains, including phenotype comparison and description, and ontology-informed phylogenetic methods (Tarasov 2019).

FUNDING

Funding was provided by the NSF ARTS Program (DEB #1523605 and DEB #2009247 to Aaron Smith). Sergei Tarasov was supported by the Academy of Finland grant: 339576, and three-year grant from the University of Helsinki. Markus Koch was supported by grants of the German Research Foundation (DFG-project MO 412/54-1) for the Specialized Information Service Biodiversity Research (BIOfid). Brendon Boudinot was supported by the Alexander von Humboldt Stiftung (Germany).

REFERENCES

Adachi, H., Matsuda, K., Nishida, K., Hanson, P., Kondo, S. & Gotoh, H. (2020) Structure and development of the complex helmet of treehoppers (Insecta: Hemiptera: Membracidae). *Zoological Letters* 6, 3. https://doi.org/10.1186/s40851-020-00155-7

Adler, P.N. (2017) Gene expression and morphogenesis during the deposition of Drosophila wing cuticle. *Fly* 11, 194–199. https://doi.org/10.1080/19336934.2017.1295188

Andersen, S.O. (2012) Cuticular sclerotization and tanning. In: *Insect molecular biology and biochemistry*. Elsevier, pp. 167–192.

Arp, R., Smith, B. & Spear, A.D. (2015) *Building Ontologies with Basic Formal Ontology*. MIT Press, Cambridge, MA, USA, 248 pp.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M. & Sherlock, G. (2000) Gene Ontology: tool for the unification of biology. *Nature Genetics* 25, 25–29. https://doi.org/10.1038/75556

Balhoff, J.P., Dahdul, W.M., Kothari, C.R., Lapp, H., Lundberg, J.G., Mabee, P., Midford, P.E., Westerfield, M. & Vision, T.J. (2010) Phenex: Ontological Annotation of Phenotypic Diversity. *PLOS ONE* 5, e10500. https://doi.org/10.1371/journal.pone.0010500

Balhoff, J.P., Mikó, I., Yoder, M.J., Mullins, P.L. & Deans, A.R. (2013) A Semantic Model for Species Description Applied to the Ensign Wasps (Hymenoptera: Evaniidae) of New Caledonia. Systematic Biology 62, 639–659. https://doi.org/10.1093/sysbio/syt028

Beutel, R.G., Friedrich, F., Yang, X.-K. & Ge, S.-Q. (2014) Insect Morphology and Phylogeny *Insect morphology and phylogeny: a textbook for students of entomology*. De Gruyter. https://doi.org/10.1515/9783110264043

Bitan, A., Rosenbaum, I. & Abdu, U. (2012) Stable and dynamic microtubules coordinately determine and maintain *Drosophila* bristle shape. *Development* 139, 1987–1996. https://doi.org/10.1242/dev.076893

Bitsch, C. & Bitsch, J. (2002) The endoskeletal structures in arthropods: cytology, morphology and evolution. *Arthropod Structure & Development* 30, 159–177. https://doi.org/10.1016/S1467-8039(01)00032-9

Blumer, M.J.F. (2021) Bone tissue and histological and molecular events during development of the long bones. *Annals of Anatomy - Anatomischer Anzeiger* 235, 151704. https://doi.org/10.1016/j.aanat.2021.151704

Bolshoy, A. & Lacková, Ľ. (2021) Illusions of Linguistics and Illusions of Modern Synthesis: Two Parallel Stories. *Biosemiotics* 14, 115–119. https://doi.org/10.1007/s12304-021-09429-9

Chibucos, M.C., Mungall, C.J., Balakrishnan, R., Christie, K.R., Huntley, R.P., White, O., Blake, J.A., Lewis, S.E. & Giglio, M. (2014) Standardized description of scientific evidence using the Evidence Ontology (ECO). *Database* 2014. https://doi.org/10.1093/database/bau075

Costa, M., Reeve, S., Grumbling, G. & Osumi-Sutherland, D. (2013) The Drosophila anatomy ontology. *Journal of Biomedical Semantics* 4, 32. https://doi.org/10.1186/2041-1480-4-32

Dahdul, W.M., Cui, H., Mabee, P.M., Mungall, C.J., Osumi-Sutherland, D., Walls, R.L. & Haendel, M.A. (2014) Nose to tail, roots to shoots: spatial descriptors for phenotypic diversity in the Biological Spatial Ontology. *Journal of Biomedical Semantics* 5, 34. https://doi.org/10.1186/2041-1480-5-34

Dahdul W., Manda P., Cui H., Balhoff J.P., Dececchi T.A., Ibrahim N., Lapp H., Vision T., Mabee P.M. (2018) Annotation of phenotypes using ontologies: a gold standard for the training and evaluation of natural language processing systems. *Database*, bay110: 17 pp. https://doi.org/10.1093/database/bay110

de Oliveira, V.P. & Cruz-Landim, C. (2003) Morphology and function of insect fat body cells: A review. *Biociências* 11, 195–205.

Deans, A.R., Mikó, I., Wipfler, B. & Friedrich, F. (2012a) Evolutionary phenomics and the emerging enlightenment of arthropod systematics. *Invertebrate Systematics* 26, 323–330. https://doi.org/10.1071/IS12063

Deans, A.R., Yoder, M.J., Balhoff, J.P. (2012) Time to change how we describe biodiversity. *Trends in Ecology & Evolution* 27, 78–84. https://doi.org/10.1016/j.tree.2011.11.007

Deans AR, Lewis SE, Huala E, Anzaldo SS, Ashburner M, Balhoff JP, Blackburn DC, Blake JA, Burleigh JG, Chanet B, Cooper LD, Courtot M, Csösz S, Cui H, Dahdul W, Das S, Dececchi TA, Dettai A, Diogo R, Druzinsky RE, Dumontier M, Franz NM, Friedrich F, Gkoutos GV, Haendel M, Harmon LJ, Hayamizu TF, He Y, Hines HM, Ibrahim N, Jackson LM, Jaiswal P, James-Zorn C, Köhler S, Lecointre G, Lapp H, Lawrence CJ, Novère NL, Lundberg JG, Macklin J, Mast AR, Midford PE, Mikó I, Mungall CJ, Oellrich A, Osumi-Sutherland D, Parkinson H, Ramírez MJ, Richter S, Robinson PN, Ruttenberg A, Schulz KS, Segerdell E, Seltmann KC, Sharkey MJ, Smith AD, Smith B, Specht CD, Squires RB, Thacker RW, Thessen A, Fernandez-Triana J, Vihinen M, Vize PD, Vogt L, Wall CE, Walls RL, Westerfeld M, Wharton RA, Wirkner CS, Woolley JB, Yoder MJ, Zorn AM, Mabee P (2015) Finding Our Way through Phenotypes. PLOS Biology 13: e1002033. https://doi.org/10.1371/journal.pbio.1002033

Denk-Lobnig, M. & Martin, A.C. (2020) Divergent and combinatorial mechanical strategies that promote epithelial folding during morphogenesis. *Current Opinion in Genetics & Development* 63, 24–29. https://doi.org/10.1016/j.gde.2020.02.014

Djokic, S., Bakhrat, A., Tsurim, I., Urakova, N., Rasgon, J.L. & Abdu, U. (2020) Actin bundles play a different role in shaping scales compared to bristles in the mosquito Aedes aegypti. *Scientific Reports* 10, 14885. https://doi.org/10.1038/s41598-020-71911-0

Dönitz, J., Grossmann, D., Schild, I., Schmitt-Engel, C., Bradler, S., Prpic, N.-M. & Bucher, G. (2013) TrOn: An Anatomical Ontology for the Beetle Tribolium castaneum. *PLOS ONE* 8, e70695. https://doi.org/10.1371/journal.pone.0070695

Girón, J.C., Mikó, I., González Montaña, L.A., Tarasov, S. & Matentzoglu, N. (2021) Ontology for the Anatomy of the Insect Skeleto-Muscular System (AISM), v2021-09-07. https://doi.org/10.5281/zenodo.4660322

Girón, J.C., González Montaña, L.A. & Mikó, I. (2021b) How to create AISM-based insect anatomy ontologies using the Ontology Development Kit (ODK) and edit them in Protégé. https://doi.org/10.5281/zenodo.4642556

Gkoutos, G.V., Green, E.C., Mallon, A.-M., Hancock, J.M. & Davidson, D. (2004) Using ontologies to describe mouse phenotypes. *Genome Biology* 6, R8. https://doi.org/10.1186/gb-2004-6-1-r8

González-Montaña, L. (2021a). Collembola Anatomy Ontology. GitHub repository. https://github.com/luis-gonzalez-m/Collembola. https://doi.org/10.5281/zenodo.4660386

González-Montaña, L. (2021b) Semantic-based methods for morphological descriptions: An applied example for Neotropical species of genus *Lepidocyrtus* Bourlet, 1839 (Collembola: Entomobryidae). *Biosystematics and Ecology* 1, e71620: 16 pp. https://doi.org/10.1553/biosystecol.1.e71620

Gotoh, H., Adachi, H., Matsuda, K. & Lavine, L.C. (2021) Epithelial folding determines the final shape of beetle horns. *Current Opinion in Genetics & Development* 69, 122–128. https://doi.org/10.1016/j.gde.2021.03.003

Haendel, M.A., Neuhaus, F., Osumi-Sutherland, D., Mabee, P.M., Mejino, J.L.V., Mungall, C.J. & Smith, B. (2008) CARO – The Common Anatomy Reference Ontology. In: A. Burger, D. Davidson, and R. Baldock (Eds), Anatomy Ontologies for Bioinformatics: Principles and Practice. Computational Biology. Springer, London, pp. 327–349. https://doi.org/10.1007/978-1-84628-885-2_16

Hall, B.K. (1975) Evolutionary Consequences of Skeletal Differentiation. *American Zoologist* 15, 329–350. https://doi.org/10.1093/icb/15.2.329

Hannezo, E., Prost, J. & Joanny, J.-F. (2014) Theory of epithelial sheet morphology in three dimensions. *Proceedings of the National Academy of Sciences* 111, 27–32. https://doi.org/10.1073/pnas.1312076111

Hastings, J., Owen, G, Dekker, A., Ennis, M., Kale, N., Muthukrishnan, V., Turner, S., Swainston, N., Mendes, P., Steinbeck, C. (2016) ChEBI in 2016: Improved services and an expanding collection of metabolites. *Nucleic Acids Research*, 44: D1214–D1219. https://doi.org/10.1093/nar/gkv1031

ICZN (1999) *International Code of Zoological Nomenclature. Fourth edition, adopted by the International Union of Biological Sciences.* International Trust for Zoological Nomenclature, London.

Iyer, J., Wang, Q., Le, T., Pizzo, L., Grönke, S., Ambegaokar, S.S., Imai, Y., Srivastava, A., Troisí, B.L., Mardon, G., Artero, R., Jackson, G.R., Isaacs, A.M., Partridge, L., Lu, B., Kumar, J.P. & Girirajan, S. (2016) Quantitative Assessment of Eye Phenotypes for Functional Genetic Studies Using *Drosophila melanogaster*. *G3 Genes|Genomes|Genetics* 6, 1427–1437. https://doi.org/10.1534/g3.116.027060

Jackson L.M., Fernando P.C., Hanscom J.S., Balhoff J.P., Mabee P.M. (2018) Automated Integration of Trees and Traits: A Case Study Using Paired Fin Loss Across Teleost Fishes. *Systematic Biology*, 67: 559–575. https://doi.org/10.1093/sysbio/syx098

Jackson, R.C., Balhoff, J.P., Douglass, E., Harris, N.L., Mungall, C.J. & Overton, J.A. (2019) ROBOT: A Tool for Automating Ontology Workflows. *BMC Bioinformatics* 20, 407. https://doi.org/10.1186/s12859-019-3002-3

Jan, S., Liu, S., Hafeez, M., Zhang, X., Dawar, F.U., Guo, J., Gao, C. & Wang, M. (2017) Isolation and functional identification of three cuticle protein genes during metamorphosis of the beet armyworm, Spodoptera exigua. *Scientific Reports* 7, 16061. https://doi.org/10.1038/s41598-017-16435-w

Klass, K.-D. (2008) The female abdomen of ovipositor-bearing odonata (insecta: Pterygota). *Arthropod Systematics & Phylogeny* 66, 45–142. https://www.zobodat.at/pdf/Arthropod-Systematics-Phylogeny_66_0045-0142.pdf

Klass, K.-D. & Matushkina, N.A. (2012) The exoskeleton of the female genitalic region in Petrobiellus takunagae (Insecta: Archaeognatha): Insect-wide terminology, homologies, and functional interpretations. *Arthropod Structure & Development* 41, 575–591. https://doi.org/10.1016/j.asd.2012.06.003

Kluge, N. (2004) *The phylogenetic system of Ephemeroptera*. Springer Science & Business Media, 442 pp. https://link.springer.com/book/10.1007/978-94-007-0872-3

Lees, A.D. & Waddington, C. (1942) The development of the bristles in normal and some mutant types of *Drosophila melanogaster*. *Proceedings of the Royal Society of London. Series B - Biological Sciences*. https://doi.org/10.1098/rspb.1942.0019

Loesel, R., Wolf, H., Kenning, M., Harzsch, S. & Sombke, A. (2013) Architectural Principles and Evolution of the Arthropod Central Nervous System. In: A. Minelli, G. Boxshall, and G. Fusco (Eds), *Arthropod Biology and Evolution: Molecules, Development, Morphology*. Springer, Berlin, Heidelberg, pp. 299–342.

Lücking, A., Driller, C., Stoeckel, M., Abrami, G., Pachzelt, A. & Mehler, A. (2021) Multiple annotation for biodiversity: developing an annotation framework among biology, linguistics and text technology. Language Resources and Evaluation. https://doi.org/10.1007/s10579-021-09553-5

Mabee, P.M., Balhoff, J.P., Dahdul, W.M., Lapp, H., Mungall, C.J. & Vision, T.J. (2020) A Logical Model of Homology for Comparative Biology. *Systematic Biology* 69, 345–362. https://doi.org/10.1093/sysbio/syz067

Matentzoglu, N., Mungall, C., Osumi-Sutherland, D., Balhoff, J., Douglass, E., Vasilevsky, N., Larralde, M., Buttigieg, P.L., bjonnh, Duncan, B., Laporte, M.-A., white, marijane, Overton, J.A., Seager, J., Calder, J., Blumberg, K., Lee, R. & Robb, S. (2021) *INCATools/ontology-development-kit: 2021-04-04 release*. Zenodo. https://doi.org/10.5281/zenodo.4662067

Midford, P.E., Dececchi, T.A., Balhoff, J.P., Dahdul, W.M., Ibrahim, N., Lapp, H., Lundberg, J.G., Mabee, P.M., Sereno, P.C., Westerfield, M., Vision, T.J. & Blackburn, D.C. (2013) The vertebrate taxonomy ontology: a framework for reasoning across model organism and species phenotypes. *Journal of Biomedical Semantics* 4, 34. https://doi.org/10.1186/2041-1480-4-34

Minelli, A. & Fusco, G. (2013) Homology. In: K. Kampourakis (Ed), *The Philosophy of Biology: A Companion for Educators*. History, Philosophy and Theory of the Life Sciences. Springer Netherlands, Dordrecht, pp. 289–322. https://doi.org/10.1007/978-94-007-6537-5_15

Monteiro, E.C., Tamaki, F.K., Terra, W.R. & Ribeiro, A.F. (2014) The digestive system of the "stick bug" Cladomorphus phyllinus (Phasmida, Phasmatidae): A morphological, physiological and biochemical analysis. *Arthropod Structure & Development* 43, 123–134. https://doi.org/10.1016/j.asd.2013.11.005

Mungall, C., Osumi-Sutherland, D., Overton, J.A., Matentzoglu, N., Balhoff, J., Harris, N., Brush, M., Touré, V., Sinclair, M., Poelen, J., Bretaudeau, A., Cain, S., Haendel, M., Vasilevsky, N., Hammock, J., Laporte, M.-A., Jensen, M. & Larralde, M. (2021) *oborel/obo-relations: 2021-03-08 Release*. Zenodo. https://doi.org/10.5281/zenodo.4589530

Mungall, C.J., Torniai, C., Gkoutos, G.V., Lewis, S.E. & Haendel, M.A. (2012) Uberon, an integrative multi-species anatomy ontology. *Genome Biology* 13, R5. https://doi.org/10.1186/gb-2012-13-1-r5

Musen, M.A. (2015) The Protégé project: a look back and a look forward. *AI Matters* 1, 4–12. https://doi.org/10.1145/2757001.2757003

Natale, D.A, Arighi, C.N., Blake, J.A., Bona, J., Chen, C., Chen, S.-C., Christie, K.R., Cowart, J., D'Eustachio, P., Diehl, A.D., Drabkin, H.J., Duncan, W.D., Huang, H., Ren, J., Ross, K., Ruttenberg, A., Shamovsky, V., Smith, B., Wang, Q., Zhang, J., El-Sayed, A., Wu, C.H. (2017) Protein Ontology (PRO): enhancing and scaling up the representation of protein entities. *Nucleic Acids Research* 45: D339–D346. https://doi.org/10.1093/nar/gkw1075

Osumi-Sutherland, D., Marygold, S.J., Millburn, G.H., McQuilton, P.A., Ponting, L., Stefancsik, R., Falls, K., Brown, N.H. & Gkoutos, G.V. (2013) The Drosophila phenotype ontology. *Journal of Biomedical Semantics* 4, 30. https://doi.org/10.1186/2041-1480-4-30

Page R.D.M. (2013) BioNames: linking taxonomy, texts, and trees. *PeerJ*, 1: e190. https://doi.org/10.7717/peerj.190

Page, R.E. & Amdam, G.V. (2007) The making of a social insect: developmental architectures of social design. *BioEssays* 29, 334–343. https://doi.org/10.1002/bies.20549

Penev, L., Lyal, C., Weitzman, A., Morse, D., King, D., Sautter, G., Georgiev, T., Morris, R., Catapano, T. & Agosti, D. (2011) XML schemas and mark-up practices of taxonomic literature. ZooKeys 150: 89–116. https://doi.org/10.3897/zookeys.150.2213

Pustejovsky. J. & Stubbs, A. (2012) Natural Language Annotation for Machine Learning: A Guide to Corpus-Building for Applications. O'Reilly Media, Inc., 342 pp. https://www.google.com/books/edition/Natural_Language_Annotation_for_Machine/A57TS7fs8MUC?hl=en&gbpv=0

Richards, A.G. & Richards, P.A. (1979) The cuticular protuberances of insects. *International Journal of Insect Morphology and Embryology* 8, 143–157. https://doi.org/10.1016/0020-7322(79)90013-8

Sautter, G., Böhm, K., Agosti, D. (2007) Semi-automated xml markup of biosystematic legacy literature with the GoldenGATE editor. Pacific Symposium on Biocomputing, 12: 391–402. https://doi.org/10.1142/9789812772435_0037

Smith, B., Ceusters, W., Klagges, B., Köhler, J., Kumar, A., Lomax, J., Mungall, C., Neuhaus, F., Rector, A.L. & Rosse, C. (2005) Relations in biomedical ontologies. *Genome Biology* 6, R46. https://doi.org/10.1186/gb-2005-6-5-r46

Smith B., Ashburner M., Rosse C., Bard J., Bug W., Ceusters W., Goldberg L.J., Eilbeck K., Ireland A., Mungall C.J., Leontis N., Rocca-Serra P., Ruttenberg A., Sansone S.-A., Scheuermann R.H., Shah N., Whetzel P.L., Lewis S. (2007) The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. *Nature Biotechnology*, 25: 1251–1255. https://doi.org/10.1038/nbt1346

Snodgrass, R.E. (1935) *Principles of Insect Morphology*. Cornell University Press, 768 pp. https://www.jstor.org/stable/10.7591/j.ctv1nhm1j

Snodgrass, R.E. (1951) Anatomy and Morphology. *Journal of the New York Entomological Society* 59, 71–73. https://www.jstor.org/stable/25005398

Snodgrass, R.E. (1963) A contribution toward an encyclopedia of insect anatomy. *Smithsonian Miscellaneous Collections* 146, 1–48. https://www.biodiversitylibrary.org/page/30259679

Sommer, R.J. (2020) Phenotypic plasticity: from theory and genetics to current and future challenges. *Genetics* 215, 1–13. https://doi.org/10.1534/genetics.120.303163

Spear, A.D., Ceusters, W. & Smith, B. (2016) Functions in Basic Formal Ontology. *Applied Ontology* 11, 103–128. https://doi.org/10.3233/AO-160164

Stork N.E. (2018) How many species of insects and other terrestrial arthropods are there on Earth? *Annual Review of Entomology*, 63: 31–45. https://doi.org/10.1146/annurev-ento-020117-043348

Stucky, B., Balhoff, J., Barve, N., Barve, V., Brenskelle, L., Brush, M., Dahlem, G., Gilbert, J., Kawahara, A., Keller, O., Lucky, A., Mayhew, P., Plotkin, D., Seltmann, K., Talamas, E., Vaidya, G., Walls, R., Yoder, M., Zhang, G. & Guralnick, R. (2019) Developing a vocabulary and ontology for modeling insect natural history data: example data, use cases, and competency questions. *Biodiversity Data Journal* 7, e33303. https://doi.org/10.3897/BDJ.7.e33303

Tajiri, R. (2017) Cuticle itself as a central and dynamic player in shaping cuticle. *Current Opinion in Insect Science* 19, 30–35. https://doi.org/10.1016/j.cois.2016.10.009

Tarasov, S. (2019) Integration of Anatomy Ontologies and Evo-Devo Using Structured Markov Models Suggests a New Framework for Modeling Discrete Phenotypic Traits. *Systematic Biology* 68, 698–716. https://doi.org/10.1093/sysbio/syz005

Topalis, P., Tzavlaki, C., Vestaki, K., Dialynas, E., Sonenshine, D.E., Butler, R., Bruggner, R.V., Stinson, E.O., Collins, F.H. & Louis, C. (2008) Anatomical ontologies of mosquitoes and ticks, and their web browsers in VectorBase. *Insect Molecular Biology* 17, 87–89. https://doi.org/10.1111/j.1365-2583.2008.00781.x

Vogt, L. (2017) Assessing similarity: on homology, characters and the need for a semantic approach to non-evolutionary comparative homology. *Cladistics* 33, 513–539. https://doi.org/10.1111/cla.12179

Vogt, L. (2019) Levels and building blocks—toward a domain granularity framework for the life sciences. *Journal of Biomedical Semantics* 10, 4. https://doi.org/10.1186/s13326-019-0196-2

Vogt, L., Bartolomaeus, T. & Giribet, G. (2010) The linguistic problem of morphology: structure versus homology and the standardization of morphological data. *Cladistics* 26, 301–325. https://doi.org/10.1111/j.1096-0031.2009.00286.x

Wagner, G.P. (2014) *Homology, Genes, and Evolutionary Innovation*. Princeton, New Jersey, United States, 496 pp. https://press.princeton.edu/books/hardcover/9780691156460/homology-genes-and-evolutionary-innovation

Wang, M., Li, Z. & Zhou, Z. (2017) Insight into the growth pattern and bone fusion of basal birds from an Early Cretaceous enantiornithine bird. *Proceedings of the National Academy of Sciences* 114, 11470–11475. https://doi.org/10.1073/pnas.1707237114

Washington, N.L., Haendel, M.A., Mungall, C.J., Ashburner, M., Westerfield, M. & Lewis, S.E. (2009) Linking human diseases to animal models using ontology-based phenotype annotation. *PLOS Biology* 7, e1000247. https://doi.org/10.1371/journal.pbio.1000247

Wirkner, C.S., Göpel, T., Runge, J., Keiler, J., Klussmann-Fricke, B.-J., Huckstorf, K., Scholz, S., Mikó, I., J. Yoder, M. & Richter, S. (2017) The first organ-based ontology for arthropods (Ontology of Arthropod Circulatory Systems - OArCS) and its integration into a novel formalization scheme for morphological descriptions. *Systematic Biology* 66, 754–768. https://doi.org/10.1093/sysbio/syw108

Yoder, M.J., Mikó, I., Seltmann, K.C., Bertone, M.A. & Deans, A.R. (2010) A Gross Anatomy Ontology for Hymenoptera. *PLOS ONE* 5, e15991. https://doi.org/10.1371/journal.pone.0015991

Yoshizawa, K. (2005) Morphology of Psocomorpha (Psocodea: 'Psocoptera'). *Insecta matsumurana. New series: journal of the Faculty of Agriculture Hokkaido University, series entomology* 62, 1–44. http://hdl.handle.net/2115/10524

Young, B.A. (1993) On the necessity of an archetypal concept in morphology: With special reference to the concepts of "structure" and "homology." Biology and Philosophy 8, 225–248. https://doi.org/10.1007/BF00850483

Zartman, J.J. & Shvartsman, S.Y. (2010) Unit Operations of Tissue Development: Epithelial Folding. *Annual Review of Chemical and Biomolecular Engineering* 1, 231–246. https://doi.org/10.1146/annurev-chembioeng-073009-100919

Zhao, X., Gou, X., Qin, Z., Li, D., Wang, Y., Ma, E., Li, S. & Zhang, J. (2017) Identification and expression of cuticular protein genes based on Locusta migratoria transcriptome. *Scientific Reports* 7, 45462. https://doi.org/10.1038/srep45462

Figure captions

 doi:10.20944/preprints202201.0254.v1

Figure 1. Autofluorescence-based CLSM micrograph showing the general structure of a sagittal section of the insect integument in an adult treehopper, genus *Ceresa* sp. (Membracidae). Excitation wavelength: 488, emission wavelengths: 500–580 pseudocolor green for conjunctivae, muscles, and other soft structures and 580–700 pseudocolor red for sclerotized components. Abbreviations: sc: sclerite; co: conjunctiva; sm: skeletal muscle; ap: appendage; in: invagination; ev: evagination.

Figure 2. Schematic representation of some of the structural components of MoDIAS: a conceptual Model for Describing Insect Anatomical Structures. The two principal structural properties that characterize the insect cuticle are: 1) degree of flexibility (y axis), ranging from sclerite (stiff, at the top) to conjunctiva (flexible, at the bottom), and 2) degree of curvature (x axis; protrusion -left- to flat -center- to depression -right-); the degree of curvature of the external and internal surfaces of cuticular protrusions and depressions can be different: when both run in parallel, they form hollow protrusions (top left subplane, e.g., carina, tubercle) or hollow depressions (top right subplane, yellow band e.g., pit, sulcus); the external surface can be depressed with the internal surface flat (top right subplane, purple band, e.g., fovea, groove); the external surface can be flat with the internal surface depressed (top right subplane, blue band, e.g., ridge, apodeme). Additional properties: quantity (single vs. multiple elements; pink and orange bands, respectively); shape can be observed throughout each subplane (e.g., sclerotized protrusions can range from elongated –carina– to rounded –tubercle–); same for depressions (elongate -groove- vs. rounded -fovea-).

 doi:10.20944/preprints202201.0254.v1

Figure 3. Schematic representation of taxon-specific definitions for paired cuticular structures of the abdominal tergites (not including structures of the genitalia) across different insect orders. Textual definitions for each structure are provided in Table 3.

Figure 4. Graphic representation of high-level classes in AISM and some of their children, including hierarchy and elements from other ontologies.