Abstract.—In order to maximize the utility of future studies of trilobite ontogeny, we propose a set of standard practices that relate to the collection, nomenclature, description, depiction, and interpretation of ontogenetic series inferred from articulated specimens belonging to individual species. In some cases, these suggestions may also apply to ontogenetic studies of other fossilized taxa.

Introduction

Recent years have seen progress in describing ancient ontogenies in ways that can be compared with those of living taxa, even at the level of patterns and mechanisms of developmental control. By conducting morphometric analysis of appropriate data sets derived from fossils it is now possible to move beyond simply describing sequences of ontogenetic stages, and to address questions of high interest for evolutionary developmental biology. This provides insight into how developmental processes evolve and how such processes affect the evolution of organismal body patterning. As such studies progress in number and taxonomic scope, it is becoming possible to assess variation in developmental trajectories among and across clades, and thus to move beyond the typological approach dictated by rare exemplars. It is increasingly apparent that even the ancient fossil record can reveal subtle patterns of microevolutionary-scale variation that has potential for insights into how aspects of body patterning evolved (Sánchez, 2012). As a result, interest in describing and interpreting ancient ontogenetic series has burgeoned.

Investigations of fossilized ontogenetic series can be aided by defining standard analytical practices, including evaluation of the limitations that fossilization places on our ability to interpret them. Here, we review concepts and procedures relating to the description and interpretation of articulated trilobite ontogeny. Although the focus of this contribution is on articulated trilobites (a general overview of trilobite ontogeny can be found in the legend of Figure 1), many of these issues apply generally in studies of ancient ontogeny, in particular in other arthropod or arthropod-related taxa. Our aim is to highlight methodological standards that may increase the comparative value of future studies.
Ontogeny offers several potential benefits. Firstly, it may clarify the methodology of investigation in order to ensure that a common descriptive framework is applied among different studies and taxa. This, in turn, facilitates the extraction of comparative information from individual case studies, and thus may enhance understanding of the evolution of development. Secondly, a standard protocol helps highlight the limitations of information and interpretation that fossilization imposes. This aids authors, reviewers, and editors in insuring that the various strengths and weaknesses of further studies are immediately evident.

Outline and caveats of a standard study

The nature of the sample.—Assessment of the sample’s characteristics is an essential first step in evaluating and interpreting patterns of variation seen among specimens belonging to a single species.

Any study must report repository information and comment on the curation state of the material considered. Preferably, an official registration/catalogue number should be available for each specimen, and details of these provided as supplementary material to published studies. Ideally, each specimen analyzed is both identified and stored individually, so that it can be recovered easily when next needed. If that is not the case, readers should be informed. In addition, the systematics of the species, including a clear diagnosis, must be included in the study if not already published. The rationale for uniting presumptive juveniles and adult specimens in the same ontogenetic series should also be made explicit. Photographs of specimens should clearly illustrate relevant features and best results are commonly achieved by applying standard paleontological photographic techniques, such as coating with ammonium chloride sublimate or magnesium oxide (e.g., Feldman, 1989). Arrows indicating the thoracic/pygidial boundary can be helpful, especially for species with a homonomous trunk condition (Hughes, 2003).

Attention should be paid to the geological context of the material analyzed. This involves documentation of the site information and number of specimens collected, and should also include discussion of key geological indicators concerning the sample. For example, analysis of the stratigraphic interval (including section thickness) from which the specimens have been derived is of critical importance for inferring the span of time over which specimens in the sample accumulated. It is necessary to note whether a sample comes from a single bed (sensu Patzkowsky and Holland, 2012), from a series of similar bedsets, or from units representing different depositional conditions. Blending of data from specimens occurring in different beds need not invalidate an analysis, but blending of data does place important constraints on the interpretation of patterns revealed. Conversely, where information on precise stratigraphic occurrence is available, it offers valuable opportunity to examine how variation is partitioned among collections with modest environmental or temporal differences. Information on lithology bears on the degree of compaction witnessed in the sample and should be noted (e.g., calcareous mudstones are commonly less compacted than claystones). In addition, the cuticular condition of the specimen (testate, internal mold, external mold, etc.) may also bear on measurements obtained.

For certain studies, the frequency of occurrence of exuviae attributed to any given stage or specific morph is relevant for testing alternative hypotheses regarding the ontogeny and/or demographics of a species. In such cases, consideration of whether specimens can be determined to be exuviae or carcasses is of importance because this strongly affects the expected frequency distribution under each hypothesis (Sheldon, 1988; Hartnoll and Bryant, 1990). However, making such a determination, even in cases in which the exoskeleton remains articulated, is commonly challenging. Only in exceptional cases can the majority of a sample’s specimens be unambiguously assigned to either category.

Coming to the dataset itself, the vagaries of fossil preservation and recovery mean that often not all specimens containing valuable information are complete in all characters of interest. Accordingly, sample size for each kind of measurement acquired should be specified, and will likely vary within the dataset. For example, it is necessary to state that out of N specimens available, the number of thoracic segments could be counted confidently on X and the number of pygidial segments (often harder to determine) on only Y out of N (or X), or that size measurements, such as body length and cephalic width, were
available for W specimens, whereas only for Z specimens was it possible to obtain more comprehensive landmark-based morphometric data.

Studies that contain morphometric data should also include estimates of measurement error in specimen size and, where relevant, shape. The extent to which taphonomy affects morphology depends on the preservational quality of the specimens considered (as noted above). As with blending data from different beds, the taphonomic modification of form need not invari-

ably exclude biologically informative studies, but the effects of taphonomic modification should be carefully gauged. For example, the shapes of articulated but compressed complete spec-

imens (e.g., Hong et al., 2014; Holmes et al., 2020a) are more strongly influenced by taphonomic factors than those of exquisitely preserved silicified sclerites (e.g., Webster, 2015), and so the biologically meaningful questions that can be asked of such materials are necessarily different. It is key to ensure that a biologically informative question can be assessed realistically given the material available. Estimates of measurement error assist in this process by identifying patterns that stand out from noise, but do not, in themselves, discriminate biological from taphonomically induced patterns. That task requires con-

sideration of whether patterns observed mimic the expectations of taphonomically induced variance.

Description and interpretation.—In approaching the study of the ontogeny of fossil species, it is important, as far as possible, to separate the description of fossils from their
developmental interpretation.

Intraspecific data may show variation in size and shape, including in discrete characters, such as segment numbers. Infering the ontogenetic process that produced the observed pattern of variation is an exercise in probabilistic inference because different processes can commonly produce the same pattern. With respect to trilobite development, terms such as morph, degree, segment, and tagma are morphological terms used to describe the phenotypic condition of a specimen. In contrast, terms such as meraspid, holaspis, anamorphic, epimorphic, stage, and instar are developmental terms and, as such, can be employed only after a given ontogenetic interpretation of the data has been made explicit and justified. In early descriptions of trilobite ontogeny, such categorical differences had limited importance because the principal aim was to reveal the broad outline of how trilobites developed. However, as more subtle aspects of ancient developmental control are dissected, categorical differences become more important.

For arthropod fossil species, which grew in a stepwise fash-

ion, ontogenetic reconstruction often starts from seeking pat-

terns in the variation of form that allow us to partition the study sample in a number of distinct morphological categories: any such categories of form are referred to as morphs. Morphs are established on a strictly descriptive basis, which considers the state of a discrete character, or of a combination of several discrete characters (e.g., the number of thoracic and pygidial segments, the latter counted, conventionally, to exclude the termi-

nal piece). Conversely, the inference that one or a subset of those morphs represents one sequential ‘step’ in development, known in arthropods as a stage or instar, is an interpretative undertaking that must not be confused with the prior, descriptive work of morph recognition. If this critical distinction between morphs and stages is overlooked, an ambiguous ontogenetic series reconstruction can result (e.g., Dai et al., 2017, fig. 7, in which all morphs were presented as sequential stages).

The distinction between description and interpretation may be reflected in the terminology adopted. In trilobite ontogeny, degree constitutes a morph that is defined by the number of thoracic segments. Degrees are generally referred to only during the meraspid period of development (e.g., a “degree X meraspid”). Because the meraspid period is a developmental phase, the terms degree and meraspid are categorically distinct, and not necessary coupled. For instance, in Aulacopleura koninckii (Bar-

rande, 1846), specimens with 18 thoracic segments (i.e., degree 18 specimens) include both meraspid individuals that would sub-

sequently attain 19, 20, 21, or 22 thoracic segments, and the holaspids for which 18 was the mature thoracic segment number (Fusco et al., 2004)—these could logically be referred to as “degree 18 meraspids” and “degree 18 holaspids,” respectively.

In some arthropods, it may be difficult to consistently and correctly identify the morphological criteria used to group specimens into morphs or degrees. For example, precise counting of thoracic segment numbers can be difficult, especially during the meraspid period. Comparison with isolated (meraspid) pygidia of approximately the same size range may be necessary for rec-

ognizing and describing the articulation separating these regions at different stages. Any criteria used to identify articulations, and thus the partition of the trunk into thoracic and pygidial seg-

ments, should be described.

A final note about the term “segment.” The description of trilobite segmentation (as either pattern or process) is typically limited to a dorsal view of the exoskeleton: what is actually observed is the subdivision of the dorsal exoskeleton into scler-

ties, or tergites, including both the articulated tergites of the thorax and their non-articulated serial homologues within the cephalon (where discernible) and in the pygidium. Pragmatically, in most papers as well as herein, the term segment is applied to all these (articulated or not) serially homologous exoskeletal units (Hughes et al., 2006). However, because in some arthro-
pods there is a mismatch between dorsal and ventral segmental patterns (Fusco and Minelli, 2013), and the developmental pro-
cesses forming ventral and dorsal serially homologous structures can operate independently (Janssen et al., 2004), the term segment should not be interpreted as referring to either a modular morphological unit of the whole body or to a develop-

mental unit of the main body axis (Fusco, 2008).

Stage assignment.—In arthropods, postembryonic growth occurs mainly in a stepwise manner, in pace with the occurrence of ecdysis (Minelli and Fusco, 2013), and it seems natural to describe the ontogeny based on successive stages. However, it should be noted that the assignment of specimens to developmental stages is a kind of inference that is not always feasible, because stages may lack unique size range or morphological markers that distinguish them.

Formally, two main types of ontogenetic morphometric data may be obtained from fossil series of molting animals, cross-sectional and mixed cross-sectional data (Cock, 1966). Cross-sectional data are those for which assignment of a given specimen to a certain developmental stage can be done with
confidence on the basis of some morphological criterion (e.g., the number of thoracic segments in immature trilobites, when appropriate, or membership in a distinct size class). Mixed cross-sectional data are those for which a criterion of stage assignment is not available, which is often the case in trilobite specimens with the mature number of segments, or when distinct sizes classes are not evident. Both types of data can be used in studies of fossil ontogeny (e.g., Fusco et al., 2016; Hopkins, 2020), but each requires different processing. When a criterion for stage assignment is available, both relative (allometric) and absolute (stage-based) growth analyses are possible. When this is not the case, only size-related shape changes can be investigated (e.g., Holmes et al., 2020b), with no possibility to separate static (i.e., within-stage) and ontogenetic (i.e., between-stages) allometry (Klingenberg, 2016).

The interpretation of stages is a critical step in any ontogenetic analysis of absolute growth, and care should be paid to justifying any particular staging hypotheses. Justification must be based on some kind of evidence. For example, if using segment numbers (either in the thorax, or in the trunk), the resulting per-stage growth rates or intra-stage size variation should exhibit some properties such as (proportional) regularity among stages. Alternatively, a given staging hypothesis can be supported on the basis of specific morphological features (e.g., exoskeletal ornament) that are seemingly added sequentially from stage to stage. Stage assignments also can be made using a criterion of size and/or shape clustering (e.g., with respect to the cranidium), although reliable assignment of individuals to particular stages is often only possible for the earliest stages. Methods for identifying instars based on size data were reviewed by Webster (2015).

Whether based on qualitative, quantitative-discrete, or quantitative-continuous characters, it is always possible (and often the case) that reliable stage assignment is only feasible for a subset of specimens and/or stages. Beyond that subset, investigation must switch from cross-sectional data analysis to analysis of mixed cross-sectional data. As a result, certain analyses commonly apply only to particular portions of the ontogeny.

Choice of the staging criteria should also consider the study’s objectives. For example, if the focus of a particular study is to resolve the ontogenetic dynamics of segment release, then stage assignment should be made using a criterion independent of the number of segments (e.g., a size-clustering based on cranial diameter size), otherwise the possibility of identifying within-stage variation in thoracic segment number is precluded. Conversely, if the focus is to determine per-stage growth rate and its variation, a criterion of stage assignment independent of any assumptions about growth patterns should be adopted. Whatever the case, the effects of potential confounding factors in stage assignment on the results should be discussed.

Illustration of the inferred ontogeny.—Because illustrations effectively present and transmit interpretations, particular care should be taken in the preparation of any diagrammatic representation of the inferred ontogeny and the legend that accompanies it.

Following McNamara et al. (2003) and Minelli et al. (2003), segmentation schedule diagrams have become common in studies of trilobite ontogeny. Their purpose is to illustrate the ontogenetic pathway by which a representative individual developed. Where more than a single developmental pathway apparently existed within a taxon, alternative segmentation schedules can be presented (e.g., Hughes et al., 2006, fig. 5A). Such diagrams may also illustrate competing developmental hypotheses (Hou et al., 2015, fig. 7).

Data analysis might produce an ontogenetic hypothesis for a taxon under study that a researcher could decide to represent through some kind of graphical schematization. There can be several reasons for this: to be consistent with the fact that not all the different details of the hypothesis may have the same evidentiary support, to suitably highlight a specific aspect of development, or simply to provide a sketch of the inferred ontogeny. Figure 1 contrasts two different schematizations (Fig. 1.2, 1.3) of the same ontogenetic hypothesis (Fig. 1.1) for the segmentation of an imaginary trilobite species, based on observed or conjectured morphs. Observations of this species suggest that there were multiple alternative developmental pathways, of which the pathways drawn in Figure 1.1 are plausible candidates. The scheme depicted in Figure 1.2 is made by simply arranging the observed morphs firstly by degree and then by the number of pygidial segments. This generates a confusion between morphs (segmental condition) and stages, and conflates the development of individual trilobites with the pattern of variation in the sample as a whole. If read as the ontogeny of an individual trilobite, as is likely, the scheme clearly does not correspond to the ontogenetic hypothesis one aims to depict (Fig. 1.1). On the contrary, the scheme depicted in Figure 1.3, while showing only one (presumably the most common) of the six segmentation pathways that individual trilobites could have followed in this case, is consistent with the hypothesized ontogeny in Figure 1.1. Information about the existence of variation in the pattern of segment addition (and thus in the number of pygidial segments at each stage), and the existence of more than one pattern of segment addition, can be conveyed with the legend or with other illustrations.

In the theoretical example given above (also in Dai et al., 2017, fig. 7), directly equating the morphological pattern observed with the developmental sequence shown in Figure 1.2 entails an unlikely process of intermittent loss of trunk segments between instars. Such a pattern is not observed in extant arthropods, and is most unlikely to have occurred in trilobites. The same potential problem exists with some other published segmentation schedules for polymerid trilobites in which two instars per meraspid degree are shown (e.g., Dai et al., 2014, fig. 6; Lei, 2016, fig. 10; Du et al., 2020, fig. 9). Although multiple instars evidently did occur within the earliest part of the meraspid period in many trilobites, before the release of any freely articulating segments in the thorax (e.g., Zhang and Clarkson, 1993), it should be reiterated that published accounts of multiple meraspid instars in those trilobites with functional thoracic segments are putative developmental scenarios, and remain hypotheses to be tested using independent evidence (also see Hou et al., 2015, p. 508–511).

Use of nomenclature.—In describing the ontogeny of an extinct species, the developmental nomenclature of the extant group to which the fossil taxon belongs, or is closely related, should be applied as far as possible.

Careful application of the developmental nomenclature of extant organisms provides the best comparative framework
available from which to highlight both the similarities and differences between the ontogenies of living and extinct forms (see Hughes et al., 2006, p. 621, on the misinterpretation of trilobite development in a major text book in invertebrate zoology resulting from trilobite-specific terminology use). To this aim, reference to a standard manual of the group can be of help (e.g., for arthropods, see Minelli et al., 2013).

For instance, in arthropods the term “segmentation” is used to describe both a morphological feature (translational body symmetry) and the developmental process that generates it (Fusco and Minelli, 2013; Dai et al., 2017; Du et al., 2020). The term “somitogenesis,” which is sometimes used in trilobite literature to indicate the appearance of new segments in the trunk prior to the onset of maturity (McNamara et al., 2006; Lei, 2016), is not normally used for the segmentation process in extant arthropods, especially in the case of post-embryonic segmentation (anamorphic development). What is observed during post-embryonic development in anamorphic arthropods is the appearance of new exoskeletal segmental units in the posterior of the trunk, which may not coincide with other aspects of segment generation. This is the reason why in fossils phrases such as “segment appearance” or “segment morphological expression” should be preferred to phrases such as “segment generation” or “segment proliferation.”

Similarly, “tagmosis” is used to indicate a morphological characteristic (a form of body organization) as well as the developmental processes that generate it (Fusco and Minelli, 2013). Although there is little consensus on how tagmata should be defined in modern arthropods, the process of tagmosis (also called tagmatization) in some way describes the ontogenetic subdivision of the main body axis into major morpho-functional units. In trilobites and their close relatives, the peculiar process of their development known as release, which involved the formation of a new functional segment articulation at the posterior end of the anterior-most segment of the pygidium, has been described as tagmosis (e.g., McNamara et al., 2006). Segment release is not observed in extant arthropods, but in many anamorphic taxa segment “appearance” followed by “maturation” at a later stage (e.g., with regard to the formation of the appendages) occurs and this pattern is not generally regarded as a change in tagmosis. As the segmental boundary between thorax and pygidium shifted posteriorly during part of ontogeny (defining the meraspis period), the trilobite thorax and pygidium have been suggested to be parts of one tagma, the trunk (Minelli et al., 2003).

Conclusions

Here we do not present a comprehensive review of the methodology of ontogenetic analysis of articulated trilobites and their relatives, but rather address specific topics that may help the growing number of case studies to be of best comparative value. It is our hope that application of this methodology can further advance understanding of the developmental basis of ancient evolution.

Acknowledgments

We thank all editors, L. Amati, and an anonymous reviewer for their helpful comments. NCH’s contribution was supported by the US National Science Foundation grant EAR-1849963 and by the Fulbright Academic and Professional Excellence Award 2019 APE-R/107 kindly hosted at the Indian Statistical Institute, Kolkata. XZ is funded by Natural Science Foundation of China (Grant Nos 41621003, 41890840 and 41930319) and the 111 Project (D17013). The paper is a contribution to IGCP 686 project “The stratigraphic and magmatic history of Early Paleozoic equatorial Gondwana and its associated evolutionary dynamics.”

References
