



Eurasia Specialized Veterinary Publication

Entomological Research Letters

ISSN: 3108-639X

2026, Volume 6, Issue 2, Page No: 67-77

Copyright CC BY-NC-SA 4.0

Available online at: www.esvpub.com/

Genomes, Imaging, AI-Assisted Morphology, and Reproducibility at the Emerging Frontier of Integrative Insect Systematics

Lina Hassan^{1*}, Omar Khalaf², Reem Jaber¹, Rania Mostafa³

¹Department of Genomic and Imaging Systematics, Faculty of Biosciences, University of Jordan, Amman, Jordan.

²Department of AI-Assisted Morphology and Reproducibility, Faculty of Life Sciences, Jordan University of Science and Technology, Irbid, Jordan.

³Department of Integrative Systematics and Emerging Frontiers, Faculty of Sciences, Al-Balqa Applied University, Salt, Jordan.

*E-mail ✉ lina.hassan@ju.edu.jo

ABSTRACT

Integrative insect systematics is entering a phase in which genome-scale sequencing, high-resolution imaging, artificial intelligence and increasingly reproducible computational infrastructure can be combined within the same evidential workflow. The opportunity is substantial, but the principal difficulty is inferential rather than technological: genomic resolution is not equivalent to an unambiguous species history, visible three-dimensional morphology is not automatically homologous or diagnostic, and machine recognition is not taxonomic adjudication. This horizon-scanning review examines these emerging interfaces and develops a proposed evidence-accountability perspective for keeping observation, measurement, computational prediction and systematic inference analytically distinct. Recent insect phylogenomics demonstrates both increased resolution and persistent conflict among loci, data types and histories. Micro-computed tomography extends morphological accessibility to historical and fossil material, while deep-learning systems increasingly extract useful signals from digitized and field imagery. Their transferability nevertheless depends on taxon representation, life stage, imaging domain, annotation and treatment of unknown taxa. The emerging frontier is therefore best understood not as progressive automation of taxonomy, but as tighter integration of complementary evidence with explicit validation boundaries. Reproducibility, provenance and specimen-linked multimodality become essential when computationally generated observations are expected to support durable systematic conclusions.

Keywords: Integrative taxonomy, Insect phylogenomics, Micro-computed tomography, Computer vision, Automated morphology, Reproducible systematics.

Received: 28 August 2026

Revised: 08 October 2026

Accepted: 10 October 2026

How to Cite This Article: Hassan L, Khalaf O, Jaber R, Mostafa R. Genomes, Imaging, AI-Assisted Morphology, and Reproducibility at the Emerging Frontier of Integrative Insect Systematics. *Entomol Appl Sci Lett.* 2026;6(2):67-77.
<https://doi.org/10.51847/Ro5gKk7uBN>

Introduction

Insect systematics is moving beyond workflows in which morphology, molecular evidence and natural-history information are accumulated as largely separate streams. Contemporary phylogenomics increasingly operates within integrative analyses in which genomic evidence is interpreted alongside morphology, ecology, behaviour and developmental information [1]. The resulting opportunity is not simply to acquire more data, but to connect evidence types without treating them as interchangeable measures of the same biological quantity.

Artificial intelligence intensifies this problem. Machine learning can assist the integration and interpretation of heterogeneous evidence used in species delimitation, but an algorithmic output remains an analytical result rather

than a species diagnosis in itself [2]. The distinction matters because taxonomic inference includes decisions about homology, lineage independence, evidence conflict, nomenclature and biological scope that need not be encoded in a recognition model.

At the same time, high-throughput insect monitoring is linking species discovery, identification, description and trait measurement into more continuous workflows [3]. This convergence creates a productive interface between biodiversity monitoring and systematics, yet operational identification in a monitoring pipeline should not be assumed to have the evidentiary status of specimen-based revision.

Two structural asymmetries further constrain the frontier. Arthropod systematics remains affected by life-stage-centric sampling, particularly when adult morphology and adult-linked records dominate evidence available for linking populations or names [4]. Published insect phylogenies are also distributed unevenly across families, indicating substantial gaps in the accessible phylogenetic knowledge base [5]. These deficiencies are sampling and knowledge-coverage problems, not evidence that poorly represented lineages are biologically less diverse or systematically less important.

The central argument developed here is therefore that technological convergence must be accompanied by evidential separation. Genomic sequences, reconstructed morphology, learned image features, classifier outputs and taxonomic conclusions may enter one integrated workflow, but they occupy different inferential positions. This review examines how that distinction changes the interpretation of genome-scale taxonomy, three-dimensional imaging and AI-assisted morphology before considering automated character extraction and, subsequently, provenance and computational reproducibility.

Current baseline in integrative systematics

Genome-scale sequencing transformed the amount of information available to insect phylogenetics, but it did not abolish the analytical problems that separate sequence acquisition from historical inference. Orthology assessment, gene-tree heterogeneity, data filtering, model choice and the relationship between individual loci and species histories remain consequential when datasets increase in size [6]. More sequence therefore expands inferential opportunity while simultaneously increasing the number of decisions that must be documented and tested.

The evidence base is also biologically uneven. Life stages remain unequally represented, so adult-centred morphological or molecular resources cannot automatically be extrapolated to immature forms [4]. Phylogenetic coverage is uneven for a different reason: available machine-readable trees represent published research effort and taxonomic harmonisation rather than a census of insect diversity [5]. These two missingness processes should remain separate. A lineage can be well represented by adult specimens yet poorly represented phylogenetically, or appear frequently in molecular studies while its immature stages remain weakly linked to named adults.

The current baseline can therefore be characterised as integrative but not evidentially interchangeable. Phylogenomic synthesis encourages combination of data types [1], AI-assisted delimitation makes explicit the need to validate machine-supported integration [2], and holistic monitoring demonstrates the value of connecting identification, specimens and traits [3]. Building on these developments, this review proposes an evidence-accountability principle: every new modality should be evaluated according to what it directly observes or computes, the inference required to turn that output into a systematic character or hypothesis, and the validation necessary before the result influences a taxonomic decision. This is an organisational synthesis proposed here, rather than an established universal standard.

Genome-scale taxonomy

Genome-scale species delimitation illustrates why dataset size and inferential certainty should not be conflated. In *Stylops*, whole-genome analyses showed that delimitation outcomes could change with locus sampling and method configuration; increasing the number of loci under one analytical approach produced additional inferred splits [7]. The appropriate conclusion is not that greater sampling is undesirable, but that delimitation sensitivity itself is evidence that should be reported. A more densely sampled partition is not automatically the biologically correct partition.

A related distinction separates topological resolution from concordance. Anchored-hybrid-enrichment analyses of *Kikihia* and *Maoricicada* cicadas increased resolution while gene- and site-level conflict and marked mito-nuclear discordance remained detectable [8]. A well-supported genomic topology can therefore coexist with heterogeneous histories. Such discordance may be compatible with biological processes including incomplete

lineage sorting or introgression, but analytical error and model inadequacy can generate overlapping patterns; the topology alone does not identify a unique cause.

Genome-enabled systematic decisions also occur at different inferential scales. Anchored phylogenomics materially informed revision of higher classification within Miltogramminae [9]. In Chanohirata, molecular species-delimitation analyses were integrated with phylogenetic and historical-biogeographic evidence rather than treated as an autonomous naming procedure [10]. Museomic analysis of Eurepini showed another pathway: historical collection material was incorporated into phylogenetic reconstruction and revisionary taxonomy, widening the specimen base available for systematic reassessment [11]. These cases should not be pooled as evidence for a common performance level. They demonstrate distinct uses of molecular information in higher classification, species-level assessment and collection-based revision.

The breadth of genome-scale applications is nevertheless expanding. Recent phylogenomic studies have reframed higher-level relationships in planthoppers, diving beetles and eurytomid wasps [12–14]. Their shared implication is breadth of application, not evidence for a universal marker strategy. In the eurytomid case, for example, character-history interpretation depended on a reconstructed phylogenetic framework, making evolutionary conclusions conditional on topology, coding and model assumptions rather than direct observations of historical transitions.

Large trees can also expose disagreement with existing taxonomy. Dense phylogenomics of the *Platydracus*-group revealed relationships with substantial implications for classification, but a phylogeny does not itself supply the morphological diagnoses or nomenclatural argument required for every formal change [15]. More directly, comparative work in Neuropterida found inconsistent behaviour among genomic data classes and analytical alternatives, demonstrating that genome-scale sources are not interchangeable [16]. The persistence of such conflict is especially important as increasingly large datasets are interpreted as if scale itself were a robustness criterion.

Conversely, recent work on Systelognatha shows that genome-scale analysis can clarify difficult higher-level backbones [17]. The lesson is therefore not scepticism toward phylogenomics, but a change in what should count as strong genomic evidence. Resolution is most informative when accompanied by explicit records of marker construction, missingness, orthology, competing analyses and unresolved conflict. This review consequently treats genome-scale taxonomy as a sensitivity-aware inference process, in which tree estimation, species delimitation, evolutionary interpretation and formal taxonomic action remain separate downstream operations.

High-resolution and 3D imaging

High-resolution imaging extends integrative systematics by changing which morphological structures can be observed without requiring those structures to become taxonomically decisive. Micro-computed tomography applied to historical geometrid material recovered detailed three-dimensional morphology that contributed to an integrative treatment of a long-overlooked moth [18]. Its systematic value arose from expanded access to characters within a broader evidential context, not from imaging acting as an independent species-delimitation mechanism.

The same observation–interpretation distinction becomes sharper in palaeontology. X-ray microtomography exposed otherwise inaccessible morphology in an enigmatic Mesozoic insect larva and allowed those structures to enter comparative evolutionary analysis [19]. Yet preservation, segmentation and character homology remain interpretive filters between a reconstructed structure and its phylogenetic meaning. A visible state can be uncertain, convergent or preservation-dependent.

Recent micro-CT-based palaeosystematics of Crato Formation beetles further illustrates the expansion of imaging into material for which conventional preparation can reveal only part of the relevant anatomy [20]. This frontier increases morphological accessibility, particularly for rare, fragile or embedded specimens, without establishing that every reconstructed feature has diagnostic value or that micro-CT has uniform throughput across taxonomic groups.

The key analytical issue is therefore where increasing visibility changes inference and where it does not. The stage-specific distinctions required to preserve that boundary are condensed below.

Table 1. Where three-dimensional visibility changes—and does not change—systematic inference

Imaging state and distinction	Specimen context and observable basis	Interpretive consequence and status	Uncertainty and required test
-------------------------------	---------------------------------------	-------------------------------------	-------------------------------

Image-enabled access to systematic evidence			
Imaging stage/state Non-destructive internal character access Focal systematic distinction Previously inaccessible versus previously unmeasured morphology	Biological or specimen context Historical adult museum insects Observable/evidential basis Reconstructed internal and external 3D structure	Interpretive consequence Adds candidate morphological evidence to an integrative treatment Status Evidence-derived	Boundary or uncertainty Resolution, specimen condition and segmentation can constrain visibility Research/test implication Compare reconstructed states with independently interpretable morphology where possible
Imaging stage/state Fossil character recovery Focal systematic distinction Visibility versus preservation-dependent state interpretation	Biological or specimen context Larval fossil material Observable/evidential basis Tomographically reconstructed hidden structures	Interpretive consequence Permits explicit coding and comparative assessment Status Evidence-derived	Boundary or uncertainty Preservation and homology remain separate from imaging success Research/test implication Test alternative character interpretations rather than treating reconstruction as self-validating
Imaging stage/state Embedded-specimen reconstruction Focal systematic distinction Physical accessibility versus systematic informativeness	Biological or specimen context Fossil Coleoptera in geological matrix Observable/evidential basis 3D anatomical reconstruction without conventional dissection	Interpretive consequence Extends the range of specimens entering systematic comparison Status Evidence-derived	Boundary or uncertainty Small specimen series and preservation constrain transfer Research/test implication Evaluate which reconstructed structures remain repeatable across specimens and observers
Character promotion checkpoints			
Imaging stage/state Character observability checkpoint Focal systematic distinction Image-derived structure versus systematic character	Biological or specimen context Historical, fossil or difficult-to-dissect material Observable/evidential basis Replicable visualization of a structure	Interpretive consequence Determines whether a feature can enter subsequent coding Status Proposed synthesis	Boundary or uncertainty Visibility alone does not establish homology or diagnosticity Research/test implication Require explicit transition from observed structure to coded character
Inference checkpoint			
Imaging stage/state Interpretation checkpoint Focal systematic distinction Character state versus phylogenetic/taxonomic conclusion	Biological or specimen context Any 3D-imaged systematic dataset Observable/evidential basis Character comparison embedded in broader evidence	Interpretive consequence Allows morphology to inform, rather than dictate, inference Status Proposed synthesis	Boundary or uncertainty Convergence, preservation and alternative homology hypotheses remain possible Research/test implication Report competing interpretations when image evidence does not uniquely resolve them

Ai-assisted morphological recognition

AI-assisted morphology changes systematics most substantially when learned features are treated as testable evidence rather than invisible intermediates. Deep-learning-derived morphological information from digitized insect collections has been shown to carry complementary phylogenetic signal when combined with molecular data [21]. This result expands the potential evidential content of collection images, but it does not establish that every learned feature is a homologous morphological character. Feature usefulness and character identity are different validation problems.

The broader insect-recognition literature is correspondingly difficult to reduce to a single measure of progress. Machine-learning applications span different taxa, image conditions, class distributions, tasks and evaluation procedures, producing substantial heterogeneity in what reported model performance actually means [22]. A horizon scan should therefore resist a pooled “AI accuracy” narrative. Recognition of curated specimen photographs, detection of insects against natural backgrounds and classification of closely related taxa constitute different inference environments.

Field systems demonstrate that the distinction has practical consequences. Deep-learning pipelines can automate defined components of time-lapse observation of insects and floral environments [23]. Such systems offer scalable ecological observation, but camera detections remain contingent on visibility, behaviour, season, position and

training composition. They also need not produce a preserved specimen that can be re-examined when an identification is questioned. Operational recognition and specimen-backed taxonomy should consequently remain separate outputs.

Unknown taxa represent an even more direct test of model boundaries. Open-set and hierarchical approaches explicitly attempt to handle insects absent from training data, yet unseen taxa can still be assigned to visually similar known classes [24, 25]. An “unknown” or anomalous prediction is therefore best interpreted as an uncertainty signal. It is not evidence, by itself, that the observation represents an undescribed species.

Entomological computer vision has long been recognised as potentially transformative, while representative training material, annotation quality, validation and deployment conditions remain central constraints [26]. The emerging systematic opportunity is thus not autonomous machine taxonomy. It is an iterative recognition environment in which image-derived evidence can be compared with specimen identity, morphology, molecular evidence and taxonomic hierarchy, with failure states retained rather than discarded.

The logic of this boundary is compressed in **Figure 1**. The figure treats recognition as a sequence of changing evidence states and makes the points at which uncertainty should return to specimen or expert review visually explicit.

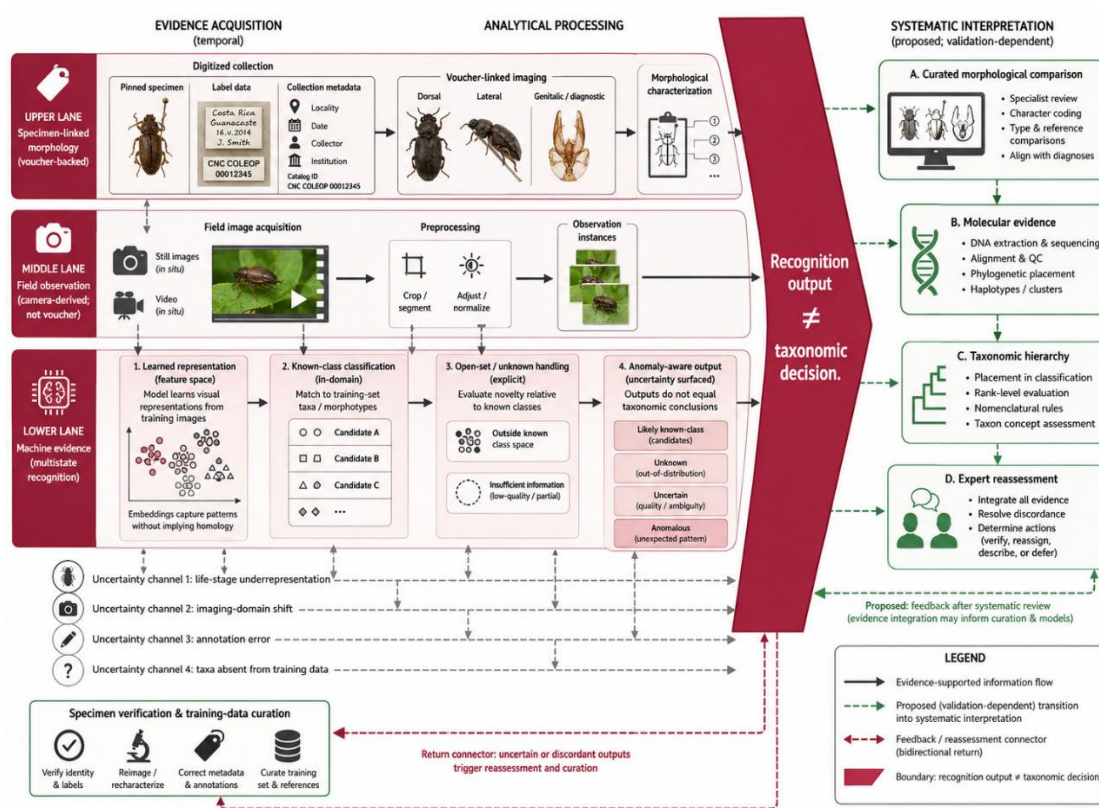


Figure 1. Recognition is not adjudication: signal states and uncertainty channels in AI-assisted insect morphology

Automated character extraction

Recognition and character extraction should not be treated as synonymous. A classifier can use image structure that improves prediction without those internal representations corresponding to repeatable, biologically interpretable character states. The demonstration that deep-learning-derived morphology can contribute complementary phylogenetic information provides a basis for testing such representations [21], but it also makes a measurement question unavoidable: what exactly has been extracted, and under what conditions can another analyst recover or interpret it?

Recent taxonomy-focused AI scholarship extends the horizon beyond identification toward automated feature extraction and multimodal analytical workflows while continuing to identify explainability, data quality and generalisation as unresolved constraints [27]. For insect systematics, this frontier is particularly consequential

because morphological characters are expected to remain interpretable across specimens, studies and often developmental or geographic variation. A computational feature useful within one training domain should not become a reusable systematic character merely because it predicts labels successfully.

This review therefore proposes a measurement-validity checkpoint between automated extraction and character use. Image-derived features should first be evaluated for repeatability, dependence on imaging conditions, taxonomic and life-stage coverage, and concordance with an independently interpretable biological signal. Evidence that learned features can contribute information does not remove this requirement [21]. Likewise, heterogeneous performance across insect image-identification tasks shows why domain-specific testing must precede transfer to new collections, field conditions or taxonomic groups [22].

The practical question is not whether character extraction can be automated, but what evidence would justify promoting an extracted feature from computational signal to reusable systematic evidence. **Table 2** converts that question into explicit predictions and tests.

Table 2. Tests required before machine-extracted morphology can function as reusable systematic evidence

Promotion state and context	Observable test	Meaning if supported	Boundary	Required next step	Evidential state
Technical repeatability and transfer					
Extraction state or prediction Feature repeatability Biological/taxonomic context Repeated images or reconstructions of the same specimen/state	Stability of extracted representation across technically reasonable repeats	Feature is less likely to be an imaging accident	Repeatability does not establish homology	Separate technical repeatability from biological interpretation	Proposed test
Extraction state or prediction Imaging-domain transfer Biological/taxonomic context Collection images, alternative imaging setups or field-derived imagery	Performance or feature stability under changed image domains	Evidence for portability beyond the development dataset	Lighting, orientation, background and resolution can alter representation	Validate transfer before reusing extracted traits across collections or monitoring systems	Proposed test
Extraction state or prediction Taxon and life-stage transfer Biological/taxonomic context Closely related, rare, immature or previously unseen insects	Behaviour on held-out taxa and developmental states	Identifies whether representation generalises beyond dominant training classes	Absence from training data can mimic novelty or failure	Report unknown/unsupported states rather than forcing assignment	Proposed test
Biological interpretability and cross-modal evidence					
Extraction state or prediction Morphological interpretability Biological/taxonomic context Candidate feature linked to observable anatomy	Concordance with independently inspectable structures or validated morphometrics	Supports use as interpretable morphological evidence	Predictive utility may arise from non-homologous image cues	Do not equate discriminative power with character homology	Proposed synthesis
Extraction state or prediction Complementary phylogenetic signal Biological/taxonomic context Digitized specimens analysed with molecular or other independent evidence	Comparative contribution of extracted morphology within joint analyses	Supports a role as an additional evidence stream	Complementarity does not imply replacement of molecular or curated morphology	Test sensitivity with and without machine-derived evidence	Evidence-derived
Extraction state or prediction Multimodal consistency Biological/taxonomic context Image-derived features compared with molecular, specimen and expert evidence	Concordance and explicit discordance among evidence streams	Makes cross-modal agreement and conflict auditable	Agreement can arise despite shared sampling or labelling bias	Preserve discordant cases for reassessment rather than collapsing them into one label	Proposed synthesis
Reusable-character promotion					
Extraction state or prediction Reusable character promotion Biological/taxonomic context	Documented extraction procedure, domain limits and	Provides a defensible transition from	No single test establishes universal diagnosticity	Require explicit validation provenance	Proposed synthesis

Feature proposed for downstream taxonomic or phylogenetic reuse	independent validation	computational feature to reusable evidence	before character reuse
---	------------------------	--	------------------------

Digital vouchers and provenance

Multimodal systematics becomes auditable when derived data remain linked to biological material. A 2026 insect dataset connects imagery, DNA, specimen identity and annotation at individual and trap scales [28], although its sampling design and predominance of trap-captured material limit transfer beyond the represented fauna. Digital extended specimens generalise this principle by linking physical vouchers to images, sequences and other derivatives through persistent digital objects [29]. Earlier extended-specimen infrastructure likewise treated collections as networks connecting genetic, phenotypic and ecological information [30]. This review therefore proposes derivative-to-voucher persistence: consequential genomic, imaging or AI outputs should retain a resolvable link to their specimen or explicitly documented observation. Connectivity improves traceability; it does not verify identification.

Reproducible computational workflows

Provenance must extend through computation. Nextflow illustrates how explicit workflow logic can make complex analyses portable and repeatable [31], while Bioconda provides a distinct layer of versioned software and dependency reconstruction [32]. Neither determines whether the biological assumptions encoded in an analysis are appropriate.

Inference provenance also requires explicit models and preprocessing. IQ-TREE 2 makes model choice, tree estimation and support procedures recoverable analytical decisions [33]. BUSCO formalises genome-completeness assessment [34]. Orthology inference is another upstream decision, and current OrthoFinder methods make that step explicit and scalable [35]. These tools motivate a proposed end-to-end accountability chain from source data through QC, orthology, environment and model to archived output; reproducible execution remains necessary but insufficient for taxonomic validity.

Because these controls operate at different stages, the table below separates reconstruction of the analysis from validation of its systematic interpretation.

Table 3. Checkpoints separating computational reproducibility from systematic validity

Recoverability dimension	Voucher linkage	Workflow logic	Environment	Model/QC/orthology	Evidence-to-inference handoff
Recoverable evidence	Persistent source identifiers	Machine-readable steps	Versioned dependencies	Explicit analytical settings	Complete provenance chain
What it permits	Trace derivatives	Re-execution	Environment reconstruction	Audit of upstream choices	Independent reassessment
Boundary	Linkage $\neq$ correct identification	Re-execution $\neq$ validity	Tooling $\neq$ appropriate analysis	Support/QC $\neq$ species truth	Computational agreement $\neq$ taxonomic adjudication
Required record	Voucher/event metadata	Workflow and inputs	Software versions	Models, QC and orthology outputs	biological validation record

The same distinction is represented visually as an evidence-to-inference boundary rather than a linear automation pathway.

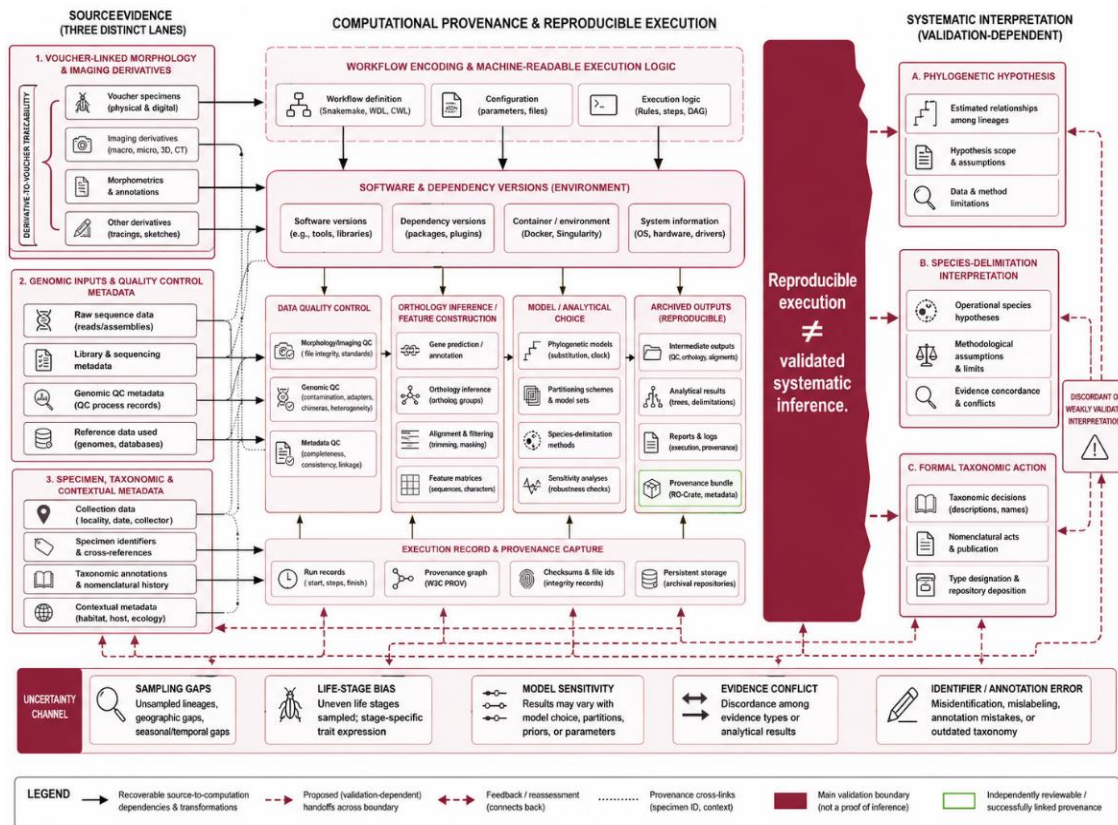


Figure 2. Where reproducibility ends and systematic validation begins

Emerging risks and validation requirements

Genome-scale evidence can remain sensitive to locus sampling [7], retain mito-nuclear discordance despite improved resolution [8], or differ among genomic data classes [16]. Such conflict should be reported as unresolved among biological history, model sensitivity and analytical error unless evidence discriminates among them.

AI requires a different stress test. Learned morphology can contribute useful signal [21], but taxa absent from training data can still be assigned to familiar classes [24], including under anomaly-aware hierarchical approaches [25]. Domain transfer, unknown-class handling and specimen-level reassessment should therefore precede taxonomic use. Likewise, multimodal provenance improves traceability [28] and digital extended specimens preserve resolvable links [29], but persistent identifiers do not guarantee correct labels. Workflow orchestration reduces computational opacity [31]. Explicit model records expose inferential choices [33]. Genome QC documents input condition [34], while orthology documentation makes another consequential preprocessing step inspectable [35]. None substitutes for biological validation.

Five-to-ten-year research horizon

The most defensible near-term direction is specimen-linked multimodality rather than replacement of one evidence class. Integrative phylogenomics supports complementary evidence [1], deep-learning morphology can add signal [21], and linked image-DNA resources demonstrate operational multimodality [28]. Adoption will remain uneven across taxa, life stages, collections and regions.

AI should become more uncertainty-aware and taxonomically structured. Integrative species-delimitation scholarship already frames AI as assistance requiring validation [2]. Open-set methods represent one route for accommodating unknown taxa [24], while anomaly-aware systems provide another [25]. Newer taxonomy-focused work extends the agenda toward feature extraction and multimodal workflows [27]. Novelty flags should therefore trigger examination, not automated species discovery.

High-value imaging can extend accessible morphology in historical [18] and fossil material [20] if derivative-to-voucher provenance persists [29]. In parallel, machine-readable workflows [31], versioned environments [32], explicit QC [34] and documented orthology [35] should make future analyses reconstructable without prescribing a single software stack. The priority is auditable inference, not technological uniformity.

Conclusion

The emerging frontier of insect systematics is not automation alone but accountable integration. Genomes may increase resolution while retaining conflict; 3D imaging reveals morphology without guaranteeing homology; AI recognises patterns without adjudicating species boundaries; and reproducible computation reconstructs analyses without validating biological conclusions. The proposed evidence-accountability perspective therefore separates observation, computational measurement, uncertainty, provenance and taxonomic inference while keeping them linked enough for independent reassessment.

Acknowledgments: None

Conflict of Interest: None

Financial Support: None

Ethics Statement: None

References

1. Ribeiro TM, Espíndola A. Integrated phylogenomic approaches in insect systematics. *Curr Opin Insect Sci.* 2024;61:101150. doi:10.1016/j.cois.2023.101150
2. Karbstein K, Kösters L, Hodač L, Hofmann M, Hörandl E, Tomasello S, et al. Species delimitation 4.0: Integrative taxonomy meets artificial intelligence. *Trends Ecol Evol.* 2024;39(8):771-84. doi:10.1016/j.tree.2023.11.002
3. Meier R, Hartop E, Pylatiuk C, Srivathsan A. Towards holistic insect monitoring: Species discovery, description, identification and traits for all insects. *Philos Trans R Soc Lond B Biol Sci.* 2024;379(1904):20230120. doi:10.1098/rstb.2023.0120
4. Caterino MS, Recuero E. Overcoming life stage-centric biases illuminates arthropod diversity, systematics and biology. *Syst Entomol.* 2024;49(3):345-54. doi:10.1111/syen.12624
5. Chesters D. Where are the biggest gaps in phylogenetic coverage of insect diversity? *Syst Entomol.* 2025;50(1):221-36. doi:10.1111/syen.12652
6. Johnson KP. Putting the genome in insect phylogenomics. *Curr Opin Insect Sci.* 2019;36:111-7. doi:10.1016/j.cois.2019.08.002
7. Lähteenaro M, Straka J, Forshage M, Hovmöller R, Nakase Y, Nilsson AL, et al. Phylogenomic species delimitation of the twisted-winged parasite genus *Stylops* (Strepsiptera). *Syst Entomol.* 2024;49(2):294-313. doi:10.1111/syen.12618
8. Stukel M, Porczak AE, Gordon ERL, Vailionis J, Haji D, Buckley TR, et al. Phylogenomics improves the phylogenetic resolution and provides strong evidence of mito-nuclear discordance in two genera of a New Zealand cicada (Hemiptera: Cicadidae) species radiation. *Syst Entomol.* 2024;49(2):237-57. doi:10.1111/syen.12613
9. Johnston NP, Pape T, Piwczynski M, Wallman JF, Wiegmann BM, Cassel BK, et al. Anchored phylogenomics and revised classification of the Miltogramminae (Diptera: Sarcophagidae). *Syst Entomol.* 2024;49(1):138-55. doi:10.1111/syen.12609
10. Wang D, Dietrich CH, Zhang Y. Phylogenetics, historical biogeography and molecular species delimitation of *Chanohirata* (Hemiptera: Cicadellidae: Deltocephalinae: Pentimiini). *Syst Entomol.* 2024;49(1):173-92. doi:10.1111/syen.12612
11. He S, Su YN, Tan MK, Zwick A, Warren BH, Robillard T, et al. Museomics, molecular phylogeny and systematic revision of the Eurepini crickets (Orthoptera: Gryllidae: Eneopterinae), with description of two new genera. *Syst Entomol.* 2024;49(3):389-411. doi:10.1111/syen.12622
12. Deng J, Stroiński A, Szwedo J, Ghanavi HR, Yapar E, Franco DC, et al. Phylogenomic insights into the relationship and the evolutionary history of planthoppers (Insecta: Hemiptera: Fulgoromorpha). *Syst Entomol.* 2025;50(3):495-518. doi:10.1111/syen.12666

13. Bergsten J, Nylander JAA, Ospina OE, Lemmon AR, Miller KB. Whole genome shotgun phylogenomics resolve the diving beetle tree of life. *Syst Entomol.* 2025;50(4):940-74. doi:10.1111/syen.12685
14. Zhang YM, Delvare G, Blaimer BB, Cruaud A, Rasplus JY, Brady SG, et al. Phasing in and out of phytophagy: Phylogeny and evolution of the family Eurytomidae (Hymenoptera: Chalcidoidea) based on ultraconserved elements. *Syst Entomol.* 2025;50(4):780-93. doi:10.1111/syen.12682
15. Zhao QH, Brunke A, Reyes-Hernández JL, Hansen AK, Solodovnikov A. A big phylogenomic tree with bigger taxonomic implications for the world's biggest rove beetles (Coleoptera, Staphylinidae, Staphylinini, Platydracus-group). *Syst Entomol.* 2025;50(4):1041-66. doi:10.1111/syen.12691
16. Zhang R, Wang L, Tian S, Liu Y, Jiang Y, Zhou X, et al. Inconsistent performance of multi-type genomic data in phylogenomics of neuropteridan insects, with solutions toward conflicting results. *Syst Entomol.* 2025;50(4):855-75. doi:10.1111/syen.12684
17. Yang X, Wang Y, Huo QB, Rehman A, Grubbs SA, Cai C, et al. Phylogenomics reveals the backbone phylogeny of Systellognatha (Insecta: Plecoptera). *Mol Phylogenet Evol.* 2026;217:108547. doi:10.1016/j.ympev.2026.108547
18. Englund M, Lee KM, Staude H, Duplouy A, Hausmann A, Laiho E, et al. 130 years from discovery to description: Micro-CT scanning applied to construct integrative taxonomy of forgotten moth from Southern Africa (Lepidoptera: Geometridae). *Syst Entomol.* 2024;49(3):507-25. doi:10.1111/syen.12627
19. Badano D, Fratini M, Maugeri L, Palermo F, Pieroni N, Cedola A, et al. X-ray microtomography and phylogenomics provide insights into the morphology and evolution of an enigmatic Mesozoic insect larva. *Syst Entomol.* 2021;46(3):672-84. doi:10.1111/syen.12482
20. Biffi G, Lepéco A, Beutel RG. Unearthing Mesozoic beetles through micro-computed tomography: A new family of Archostemata and the first representatives of Ommatidae and †Notocupes from the Crato Formation. *Syst Entomol.* 2026;51(3):e70061. doi:10.1111/syen.70061
21. Hunt R, Reyes-Hernández JL, Shaw JJ, Solodovnikov A, Pedersen KS. Integrating deep learning derived morphological traits and molecular data for total-evidence phylogenetics: Lessons from digitized collections. *Syst Biol.* 2025;74(3):453-68. doi:10.1093/sysbio/syae072
22. Gao Y, Xue X, Qin G, Li K, Liu J, Zhang Y, et al. Application of machine learning in automatic image identification of insects - a review. *Ecol Inform.* 2024;80:102539. doi:10.1016/j.ecoinf.2024.102539
23. Bjerger K, Karstoft H, Mann HMR, Høye TT. A deep learning pipeline for time-lapse camera monitoring of insects and their floral environments. *Ecol Inform.* 2024;84:102861. doi:10.1016/j.ecoinf.2024.102861
24. Badirli S, Picard CJ, Mohler G, Richert F, Akata Z, Dundar M. Classifying the unknown: Insect identification with deep hierarchical Bayesian learning. *Methods Ecol Evol.* 2023;14(6):1515-30. doi:10.1111/2041-210X.14104
25. Bjerger K, Geissmann Q, Alison J, Mann HMR, Høye TT, Dyrmann M, et al. Hierarchical classification of insects with multitask learning and anomaly detection. *Ecol Inform.* 2023;77:102278. doi:10.1016/j.ecoinf.2023.102278
26. Høye TT, Ärje J, Bjerger K, Hansen OLP, Iosifidis A, Leese F, et al. Deep learning and computer vision will transform entomology. *Proc Natl Acad Sci U S A.* 2021;118(2):e2002545117. doi:10.1073/pnas.2002545117
27. Lu S, Fang C, Zeng H, Hu R, Wei C, Miao R, et al. Advancing biological taxonomy in the AI era: Deep learning applications, challenges, and future directions. *Sci China Life Sci.* 2026;69(1):37-50. doi:10.1007/s11427-025-3074-8
28. Orsholm J, Quinto J, Autto H, Banelyte G, Chazot N, deWaard J, et al. A multi-modal dataset for insect biodiversity with imagery and DNA at the trap and individual level. *Sci Data.* 2026;13(1):630. doi:10.1038/s41597-026-07251-x
29. Hardisty AR, Ellwood ER, Nelson G, Zimkus B, Buschbom J, Addink W, et al. Digital Extended Specimens: Enabling an extensible network of biodiversity data records as integrated digital objects on the Internet. *Bioscience.* 2022;72(10):978-87. doi:10.1093/biosci/biac060
30. Lendemer J, Thiers B, Monfils AK, Zaspel J, Ellwood ER, Bentley A, et al. The Extended Specimen Network: A strategy to enhance US biodiversity collections, promote research and education. *Bioscience.* 2020;70(1):23-30. doi:10.1093/biosci/biz140
31. Di Tommaso P, Chatzou M, Floden EW, Barja PP, Palumbo E, Notredame C. Nextflow enables reproducible computational workflows. *Nat Biotechnol.* 2017;35(4):316-9. doi:10.1038/nbt.3820

32. Gruning B, Dale R, Sjödin A, Chapman BA, Rowe J, Tomkins-Tinch CH, et al. Bioconda: Sustainable and comprehensive software distribution for the life sciences. *Nat Methods*. 2018;15(7):475-6. doi:10.1038/s41592-018-0046-7
33. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol*. 2020;37(5):1530-4. doi:10.1093/molbev/msaa015
34. Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. BUSCO update: Novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol Biol Evol*. 2021;38(10):4647-54. doi:10.1093/molbev/msab199
35. Emms DM, Liu Y, Belcher L, Holmes J, Kelly S. OrthoFinder: Improved phylogenetic orthology inference with enhanced accuracy and scalability. *Nat Methods*. 2026;23(7):1327-33. doi:10.1038/s41592-026-03126-6