Eurasia Specialized Veterinary Publication

Entomological Research Letters

ISSN: 3108-639X

2024, Volume 4, Issue 2, Page No: 130-140
Copyright CC BY-NC-SA 4.0
Available online at: www.esvpub.com/

Species Concepts, Data Conflict, and Evidential Standards in Integrative Insect Taxonomy Under Increasing Genomic Complexity

Omar Khalid[1]*, Sara Nadeem[1], Bilal Farooq[2], Hina Saeed[1]

[1]Department of Species Concepts and Data Conflict, Faculty of Life Sciences, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia.
[2]Department of Evidential Standards and Genomic Complexity, Faculty of Biosciences, Qatar University, Doha, Qatar.

***E-mail** ✉ omar.khalid@kfupm.edu.sa

ABSTRACT

Genome-scale data have transformed insect systematics by exposing population structure, cryptic lineages, introgression, discordant genealogies, and phenotypically inconspicuous divergence at unprecedented resolution. Greater resolution, however, does not make species inference self-validating. Species delimitation still requires decisions about what biological process produced an observed partition, whether apparently corroborating evidence is genuinely independent, and whether evidence supporting lineage separation is sufficient for reproducible diagnosis and naming. This critical review evaluates these problems across coalescent delimitation, phylogenomics, gene flow, reticulation, morphological and ecological discordance, and integrative taxonomic practice. The evidence shows that population structure can mimic species boundaries, gene flow can coexist with lineage integrity, mitochondrial and nuclear histories may diverge, diagnostic morphology can depend on sex or developmental condition, and genomic scale does not eliminate model or sampling error. The review therefore develops a proposed distinction among evidence detection, process interpretation, lineage delimitation, diagnostic corroboration, and nomenclatural assignment. Under this synthesis, conflict is not treated as failed integration but as information about competing biological or analytical explanations. Evidential strength depends not simply on the number of datasets agreeing, but on whether they interrogate partly independent biological dimensions and survive explicit alternative-process tests. This approach is intended as a critical framework for more transparent insect taxonomy under increasing genomic complexity rather than as a universal species criterion.

Keywords: Species delimitation, Integrative taxonomy, Phylogenomics, Gene flow, Evidential independence, Taxonomic diagnosis

Received: 22 October 2024
Revised: 10 December 2024
Accepted: 11 December 2024

How to Cite This Article: Khalid O, Nadeem S, Farooq B, Saeed H. Species Concepts, Data Conflict, and Evidential Standards in Integrative Insect Taxonomy Under Increasing Genomic Complexity. Entomol Res Lett. 2024;4(2):130-40.
https://doi.org/10.51847/QMdZokaVJm

Introduction

Genome-enabled taxonomy has changed the scale at which insect diversity can be detected, but it has not eliminated the inferential decisions that connect variation to species status. Contemporary delimitation increasingly combines genomic, morphological, ecological, geographic and computational evidence, yet current syntheses still identify heterogeneous species concepts, complex evolutionary histories and difficult data integration as persistent sources of uncertainty [1–3]. Computational approaches can organize multidimensional evidence and reveal latent clustering, but the resulting clusters remain taxonomic hypotheses rather than self-validating species assignments [4]. The central problem is therefore no longer simply whether enough characters

© 2024 Entomological Research Letters

have been measured. It is whether the measured differences identify the same biological boundary, whether they arise from independent processes, and whether the inference remains defensible under plausible alternatives.

This problem is especially acute in insects. Large effective population sizes, incomplete lineage sorting, heterogeneous dispersal, recurrent geographic subdivision, sex-biased movement, host association and occasional hybridization can generate sharply different signatures among loci and phenotypes. Conversely, recently diverged or morphologically conservative species may remain difficult to diagnose despite strong genomic differentiation. Increasing the number of loci can make particular partitions highly precise while leaving unresolved whether those partitions represent population structure, lineage divergence, reticulation, model misspecification or a taxonomically meaningful species boundary.

This review therefore distinguishes five inferential operations that are often compressed into a single act of "species delimitation": detecting structure, identifying the evolutionary or demographic process producing that structure, judging whether the resulting lineage hypothesis warrants species-level interpretation, establishing a reproducible diagnosis, and linking that diagnosis to nomenclatural identity. The article does not propose another universal species concept. Instead, it develops a proposed evidential-independence and inference-boundary synthesis in which agreement among data classes is strongest when their biological bases and failure modes are sufficiently distinct, and discordance triggers explanation rather than automatic rejection of one dataset.

Species concepts in operational taxonomy

Species concepts matter operationally because delimitation methods do not observe species directly; they estimate patterns expected under particular models of lineage history. The multispecies coalescent is especially useful for describing genealogical separation, but it can identify population structure as if it were species-level divergence when structured populations are treated as candidate species [5]. Empirical comparison with reproductive isolation reinforces the same caution from another direction: coalescent partitions and reproductive isolation can correspond substantially without being equivalent criteria [6]. A genomically distinct population is therefore not transformed into a species merely because one model assigns it a high posterior probability.

Geographic sampling further complicates interpretation. Isolation by distance and long-standing population structure can produce apparently discrete partitions when sampling is spatially discontinuous, allowing widespread taxa to be over-split under some multispecies-coalescent analyses [7]. At the opposite extreme, explicit demographic modeling shows that gene flow can persist between diverging lineages and should therefore be modeled rather than treated as an automatic contradiction of species status [8]. These findings undermine two symmetrical shortcuts: "structure means species" and "gene flow means one species."

The relationship between genetic similarity and taxonomic status is also non-monotonic. Under asymmetric migration or introgression, sequences sampled from different species can sometimes be more closely related than sequences sampled within the same nominal species [9]. Such outcomes are not evidence that molecular data lack taxonomic value. They show instead that similarity is an observation requiring process interpretation.

Operational taxonomy should therefore separate a species criterion from an evidence-generating model. Reproductive isolation, lineage independence, ecological differentiation and diagnosability answer different biological questions, while coalescent, clustering or distance-based methods estimate different manifestations of those questions. A defensible insect delimitation should state which criterion is being operationalized, which biological processes could imitate its expected signal, and which evidence would discriminate among those alternatives. This does not eliminate conceptual pluralism; it makes the consequences of that pluralism explicit.

Genome-scale variation

Phylogenomics reduces stochastic uncertainty by sampling many loci, but systematic error can become more rather than less consequential when model misspecification, paralogy, taxon sampling or data-processing assumptions affect large datasets consistently [10]. Genome-scale confidence must therefore be distinguished from taxonomic validity. A highly supported partition may be precise under the selected analytical model while remaining biologically ambiguous.

Allopatric Lepidoptera illustrate why genomic depth alone cannot supply a universal divergence threshold. Nuclear target-enrichment data reveal markedly different degrees of differentiation among geographically separated populations that might appear similarly distinct under a single mitochondrial marker [11]. Conversely, phylogenomic analyses of morphologically cryptic *Nylanderia* demonstrate the capacity of multilocus datasets to resolve complexes in which external morphology provides limited initial separation [12]. These cases are

complementary rather than contradictory: genomic data are powerful for detecting structure, but the taxonomic interpretation of that structure remains conditional on geography, sampling and independent corroboration.

Genome-scale evidence also changes existing classifications unevenly. In Polleniidae, mitogenomic and morphological evidence recover both congruent and conflicting relationships, showing that historical classifications may contain a mixture of robust characters and misleading similarities [13]. Large-scale butterfly systematics demonstrates that target-enrichment datasets can be combined with legacy Sanger sequences to increase taxonomic coverage, although heterogeneous missing data and uneven sampling remain relevant [14]. Anchored phylogenomics of Miltogramminae similarly supported revision of higher classification while retaining explicit limitations imposed by incomplete genus and species coverage [15].

These studies collectively argue against a simple "genomics replaces morphology" narrative. Genome-scale variation is most informative when it clarifies which partitions are stable across analytical treatments, which are sensitive to sampling or marker choice, and which conflict with phenotypic or geographic expectations. The important transition is therefore from more loci to more explicit inference. Genomic datasets can reveal cryptic lineages and reorganize classification, but they cannot by themselves determine whether discordance reflects introgression, incomplete lineage sorting, geographic structure, convergent morphology or inadequate sampling.

Gene flow and reticulation

Gene-tree discordance is expected under several evolutionary processes, including incomplete lineage sorting, migration and introgression, and the same topological discordance can therefore have more than one historical explanation [16]. This non-uniqueness is central to species delimitation because a taxonomist may otherwise interpret reticulation as evidence against lineage separation, or interpret discordant loci as direct evidence of hybridization, without first excluding alternative sources of conflict.

Even formal introgression statistics have identifiable failure modes. Among-species substitution-rate variation can generate signals resembling introgression in some analytical settings, meaning that rate heterogeneity must be considered before discordant site patterns are converted into reticulation histories [17]. Unsampled or extinct donor lineages create a different problem: "ghost" introgression can produce patterns that simpler topological summaries assign to the wrong donor or recipient, whereas fuller use of topology and branch-length information can improve discrimination among histories [18]. Neither result invalidates introgression analysis; both demonstrate that reticulation is an inferred process rather than a directly observed property of a discordant tree.

Insects provide equally important empirical counterexamples to the idea that gene exchange necessarily destroys species boundaries. In a western Canadian *Phyciodes* contact region, genome-wide SNP data supported the integrity of four species despite limited gene flow, while mitochondrial COI haplotypes were extensively shared among taxa [19]. Mito-nuclear discordance can also arise from sex-biased dispersal, because mitochondrial and nuclear genomes experience different inheritance and movement histories [20]. Thus mitochondrial sharing may be compatible with introgression, sex-biased movement, incomplete lineage sorting, or combinations of these processes; its meaning is system dependent.

For integrative taxonomy, the relevant question is not simply whether gene flow exists but whether the magnitude, timing and genomic distribution of exchange are compatible with persistent lineage differentiation. Reticulation should therefore be treated as a process layer superimposed on lineage history rather than as a binary veto on species recognition. A strong analysis asks whether the same observations remain plausible under incomplete lineage sorting, geographic structure, rate heterogeneity, unsampled donors or sex-specific movement.

Because these scenarios can generate superficially similar conflict while requiring different taxonomic interpretations, their distinctions are more efficiently compared as a mechanism–boundary matrix than as a sequential list.

Table 1. Gene-flow and reticulation patterns that alter the taxonomic meaning of genomic discordance

Biological setting with observed basis	Interpretation for speci within its limit	Discriminating check
Persistent lineage integrity with limited exchange		
Evidence-derived		
Recently diverged or contact-zone taxa	Gene flow does not automatically erase a species boundary Strength depends on sampling, genomic coverage and demographic history	Test whether differentiation is genome-wide and stable across localities

Genome-wide clustering persists although some loci or individuals indicate admixture		
Mitochondrial sharing across genomically distinct taxa		
Evidence-derived		
Contact zones or recently separated lineages **COI haplotypes cross nominal species while nuclear genomic structure remains distinct**	Mitochondrial identity cannot by itself establish conspecificity Introgression and incomplete lineage sorting may produce similar patterns	Compare mitochondrial history with multilocus nuclear structure
Sex-biased movement generating marker conflict		
Evidence-derived		
Taxa in which dispersal differs between sexes **Mitochondrial and nuclear datasets imply different geographic boundaries**	Marker discordance can arise from different inheritance and dispersal pathways Mechanism is taxon specific and requires biological support	Test sex-specific dispersal and marker-specific structure
Rate-variation mimicry		
Evidence-derived		
Phylogenomic datasets with heterogeneous substitution rates **Introgression statistic is significant despite no independently established exchange**	Apparent introgression may reflect model violation False-positive risk depends on rate regime and statistic	Evaluate rate heterogeneity and alternative tests
Ghost-reticulation ambiguity		
Evidence-derived		
Incompletely sampled radiations **Observed topology fits several donor–recipient histories**	The apparent source of introgression may be misassigned Unsampled or extinct lineages are inherently difficult to exclude	Use branch-length and likelihood information and expand sampling
Reticulation-uncertainty state		
Proposed synthesis		
Any lineage set with unresolved discordance **Multiple processes remain compatible with the same conflict pattern**	Taxonomic interpretation should remain conditional rather than forced No single process has been discriminated	Preserve unresolved status until alternatives are tested

Morphological and ecological discordance

Discordance becomes analytically useful when different datasets are allowed to disagree for biologically interpretable reasons. In a tiger beetle complex, mitochondrial, genomic and morphological evidence did not recover identical units, demonstrating that even extensive integrative sampling can leave genuine conflict rather than a single consensual signal [21]. In *Nylanderia*, three-dimensional male genital morphometrics provided diagnostic structure that could be compared with phylogenomic lineages, while some mitochondrial groupings lacked corresponding genomic or morphological independence [22]. The implication is not that genital morphology is intrinsically superior to COI, but that character systems can differ in lineage specificity.

Ecological and physiological divergence introduces another inferential level. Alpine diving beetles showed genomic, morphological and physiological differentiation consistent with rapid ecotypic divergence and incipient speciation [23]. Such concordance can strengthen an evolutionary interpretation of population differentiation, but "incipient" itself marks an unresolved endpoint: ecological differentiation does not automatically supply a taxonomic diagnosis or demonstrate completion of speciation.

Morphology is also conditional on developmental state. In overlapping *Psylliodes* populations, some diagnostic structures varied in reliability before full sclerotization, showing that a character may be informative in mature adults yet unreliable in recently emerged individuals [24]. Geographic structure creates a different limitation. Within the *Poecilimon bosphoricus* group, geography predicted genetic divergence better than existing morphology-based taxonomy, while gene flow and incomplete lineage sorting remained plausible contributors to

discordance [25]. Genetic conflict with current names therefore identifies a taxonomic problem without automatically proving which names should change.

Finally, external adult morphology is not the only phenotypic evidence available. Reproductive anatomy and immature-stage characters helped distinguish externally similar eupelmid genera, illustrating that taxonomically informative character systems may be hidden when analyses are restricted to adult habitus [26]. Although this is a genus-level result rather than a species-delimitation test, it supports a broader methodological principle: the apparent absence of morphological differentiation can reflect incomplete character sampling.

Taken together, these cases argue for replacing "morphology versus molecules" with a more specific question: which phenotype, measured at which life stage and spatial scale, is expected to track the lineage process under investigation? Morphological, ecological and physiological discordance should therefore be interpreted relative to development, sex, geography, inheritance and evolutionary timescale rather than treated as a unitary category of contradictory evidence.

Evidential independence

Integrative taxonomy is often described as stronger when multiple data classes agree, but simple agreement does not establish independence. Molecular and morphological characters can reinforce the same phylogenetic hypothesis, as demonstrated by the combined analysis of extensive morphological and molecular characters in Epyrinae [27]. Yet two datasets can also share an underlying cause: geographic isolation may generate both genomic structure and correlated phenotypic differentiation, while several mitochondrial loci largely represent one inherited genomic history. Counting such datasets as multiple independent confirmations can exaggerate evidential strength.

The strongest insect studies instead use new evidence to challenge an initial hypothesis. In *Pseudoloxops*, molecularly generated groupings were evaluated against morphology, genital structure, host association and geography before formal species recognition [28]. In the *Temnothorax salvini* group, genomic hypotheses were compared with morphometric and other taxonomic evidence, and the resulting species hypotheses were subsequently diagnosable rather than resting on genomic clustering alone [29]. These examples do not establish one mandatory workflow, but they illustrate why integrative support is more informative when later evidence could realistically falsify the initial partition.

This review therefore proposes three distinctions. Redundant corroboration occurs when nominally different measurements are strongly coupled to the same process or sampling structure. Conditional independence occurs when evidence classes interrogate partly different biological dimensions but still share important confounders. Orthogonal corroboration occurs when a lineage hypothesis survives evidence with substantially different biological bases and failure modes, for example genomic differentiation together with stage-appropriate diagnostic morphology whose distribution is not merely a restatement of geography. These categories are proposed analytical terms, not validated certainty grades.

A second distinction concerns scale. Evidence may be individually reliable yet mismatched to the taxonomic question: population-genomic structure can be robust at a regional scale, ecological differentiation may be habitat-specific, and a morphological diagnosis may function only in one sex or developmental state. Independence therefore cannot be separated from portability. Evidence supporting species recognition should be evaluated not only for agreement but also for the spatial, developmental and analytical domain across which that agreement remains interpretable.

The cross-scale problem is summarized below because the distinction among redundancy, partial independence and genuinely challenging corroboration is difficult to recover from a simple count of agreeing datasets.

Table 2. Cross-scale evidential relationships that determine whether agreement constitutes independent corroboration

Evidential-independence with evidential standing	Typical taxonomic contex with observed basis	Consequence for interpretation, principal uncertainty or boundary, and research or testing implication
Shared-process redundancy	Several measurements respond to the same spatial or inheritance structure	Agreement adds less independent support than the number of datasets suggests
Proposed synthesis	Multiple datasets agree but are biologically coupled	Common cause may be unrecognized

		Identify shared inheritance, geography or developmental origin
Marker-family redundancy	Multiple loci belong to the same largely linked or uniparentally inherited history	Repetition increases precision more than inferential breadth
Proposed synthesis	Several markers recover the same partition	Linkage and shared genealogy limit independence Contrast with evidence from a different inheritance system
Conditional independence	Genomic, ecological or morphological evidence interrogates different dimensions but shares one major confounder	Support is meaningful but remains conditional Geography, environment or sampling may jointly structure datasets
Proposed synthesis	Partial concordance across evidence classes	Test the shared confounder explicitly
Orthogonal corroboration	Candidate lineages are tested with evidence having different biological bases and failure modes	Stronger support for a lineage interpretation Complete independence is rarely demonstrable
Proposed synthesis	Initial partition survives independent diagnostic or biological challenge	Prefer evidence capable of falsifying the initial hypothesis
Scale-mismatched agreement	Evidence is reliable but collected at different spatial, developmental or analytical scales	Apparent conflict may reflect non-comparable scales Transferability of diagnosis or structure is limited
Evidence-derived synthesis	Concordance changes with locality, sex, caste or life stage	Match sampling scale to the taxonomic claim
Diagnostic independence	A genomic lineage also possesses reproducible characters useful for identification	Delimitation and diagnosis reinforce different inferential stages Diagnostic trait may not transfer across all stages or ranges
Proposed synthesis	Lineage evidence and practical diagnosis converge	Test diagnostic portability beyond discovery samples

The proposed standard is therefore not "more datasets are better." It is that evidence should be weighted qualitatively by what it measures, what could generate a false concordance, what alternatives it can exclude, and whether its interpretation survives changes in biological and analytical scale. This reframes evidential conflict from an inconvenience to be averaged away into a source of information about the processes producing taxonomic disagreement.

Standards for delimitation

A delimitation standard should evaluate the conditions under which a result was produced, not merely whether a method returned a supported partition. Simulation work shows that multispecies-coalescent inference can change when model assumptions are violated [30], while comparisons among delimitation methods show that relative performance depends on population size, divergence history and gene flow [31]. Method agreement is therefore not a substitute for model adequacy.

Distance thresholds require the same restraint. Broad analysis of insect COI variation found substantial intraspecific divergence in many taxa, undermining the assumption that one fixed cutoff reliably separates all species [32]. Within Entiminae, estimated thresholds varied among genera even within a single beetle subfamily [33]. Barcode distances remain useful for identification, screening and hypothesis generation, but their taxonomic meaning is conditional on sampling and taxonomic context.

Sampling architecture is part of this standard. A genome-wide partition recovered from sparse, geographically separated samples may be highly repeatable yet still confound discontinuous sampling with lineage discontinuity. Conversely, extensive locus sampling cannot compensate for missing populations at predicted contact zones. The relevant robustness question is therefore whether the lineage hypothesis persists when the spatial and analytical design is capable of exposing clines, admixture or alternative topology. This follows the broader phylogenomic warning that systematic biases can survive very large datasets [10].

This review therefore proposes that delimitation claims pass three qualitative boundaries: process plausibility, model robustness and evidential challenge. A candidate lineage should remain credible after plausible alternatives such as geographic structure, incomplete lineage sorting and reticulation are considered; its principal partition

should not depend solely on one vulnerable analytical assumption; and at least one informative evidence class should be capable of challenging rather than simply restating the initial genomic result. These are not numerical criteria and do not assign certainty grades.

The resulting logic is compressed in **Figure 1** because standards for delimitation operate across observation, process interpretation and taxonomic inference rather than as a single analytical step.

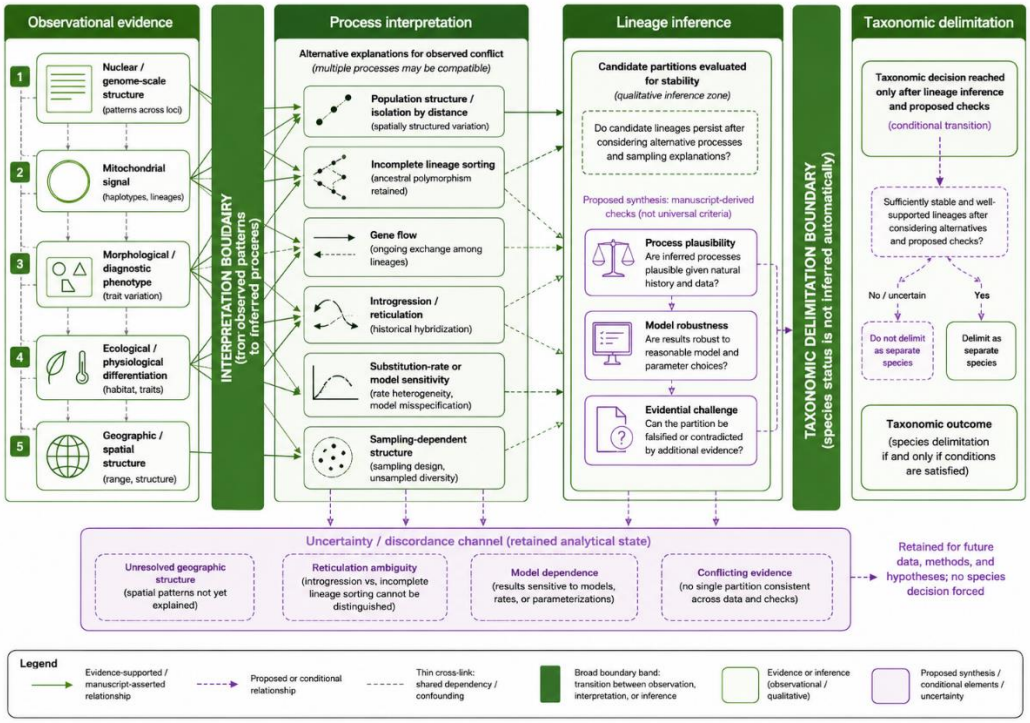


Figure 1. Where genomic partitions become taxonomic claims: layered boundaries for delimitation under conflict

Standards for diagnosis and naming

Delimitation and diagnosis should remain separate. A lineage may be strongly supported genomically yet lack a portable character for identification, while a diagnostic difference may exist without establishing evolutionary independence. Sequence provenance adds another boundary: type-derived DNA records can strengthen reproducibility by linking sequence data to name-bearing material, but such linkage does not itself demonstrate species distinctness [34].

Formal revision illustrates the additional work required after lineage discovery. In the *Cerurina–Afrocerura* revision, phylogenomic, barcode, morphological and type-associated evidence were integrated within a nomenclatural treatment [35]. This does not create a universal prescription, but it shows why naming is an evidence-traceability operation as well as an inference about lineage status. Diagnosis must also remain portable across relevant sex, life stage, geography and material condition; otherwise the apparent species boundary may be difficult to reproduce outside the discovery dataset.

Diagnosis need not mean one conspicuous external character. It may rest on reproducible combinations of morphology, genital structure, developmental characters or molecular identifiers, provided their scope is stated. The *Nylanderia* evidence shows that sex-specific structures can add information not recovered by mitochondrial clustering alone [22], while *Psylliodes* demonstrates that character reliability can vary with maturity [24]. A diagnosis should therefore report where it works, not imply portability that has not been tested.

The principal failure modes are therefore better represented conditionally than as a checklist.

Table 3. Failure modes separating lineage delimitation from reproducible diagnosis and naming

Boundary condition and decisive check	State or failure mode with evidential standing	Evidential situation and warranted reading

Crypsis may be genuine **Seek independent diagnostic or operational identification evidence**	Genomic-only lineage Proposed synthesis	Stable genomic partition without portable diagnosis Supports a lineage hypothesis but not automatically a reproducible identification route
Developmental transformation alters character expression **Validate across relevant stages**	Stage-limited diagnosis Evidence-derived synthesis	Character works only in a particular maturity or life stage Diagnosis may fail in other material
Other forms may remain cryptic **State diagnostic scope explicitly**	Sex- or caste-limited diagnosis Proposed synthesis	Character is restricted to one sex or caste Identification depends on specimen composition
Unsampled clines or populations may overlap **Test range-wide portability**	Geographically restricted diagnosis Proposed synthesis	Character combination holds only within discovery range Transfer beyond sampled region is uncertain
Database metadata may be incomplete **Strengthen voucher and type linkage**	Name–sequence provenance gap Evidence-derived	Sequence cannot be securely connected to name-bearing material Molecular identification can inherit naming error
Formal diagnosis and type context are missing **Keep lineage inference and nomenclatural act separate**	Delimitation–naming collapse Proposed synthesis	Analytical cluster is treated directly as a named species Inference and nomenclature become conflated

Critical synthesis

The central difficulty in genomic insect taxonomy is not data scarcity but inferential compression. Population structure can be mistaken for species separation [5], discordant datasets can reflect genuine biological conflict [21], and sequence identity still requires nomenclatural anchoring [34]. Treating these as one decision obscures where uncertainty enters.

This separation also changes the role of disagreement. If genomic and morphological evidence conflict, the task is not to select the dataset with the greatest resolution, but to identify which causal or analytical histories could generate the observed asymmetry. The same principle applies when several genomic methods agree: concordance among methods sharing assumptions may be less informative than a disagreement that reveals model sensitivity. Model-violation studies make this distinction especially important [30]. Accordingly, unresolved status is an informative taxonomic outcome rather than a procedural failure.

The proposed synthesis therefore distinguishes detection, process interpretation, lineage delimitation, diagnostic corroboration and naming. Discordance is carried forward until competing explanations are tested; agreement is weighted qualitatively by evidential independence rather than dataset count. Rate heterogeneity can mimic introgression [17], geography can organize genetic divergence more strongly than existing taxonomic boundaries [25], and genomic hypotheses can nevertheless gain independent morphological support [29]. The framework consequently permits three legitimate outcomes: supported species-level interpretation, supported lineage structure with taxonomic status unresolved, or unresolved conflict requiring additional evidence. It does not prescribe a universal species concept, threshold or weighting formula.

Future taxonomic practice

Future practice should couple genomic throughput with explicit validation. Large-scale integrative workflows can accelerate hypothesis discovery [3], and multidimensional computational methods can help organize heterogeneous evidence [4], but neither function substitutes for taxonomic interpretation. Geographic sampling should be expanded where allopatry or spatial structure could generate false discreteness [11], while type-linked sequence infrastructure can improve the reproducibility of names attached to molecular records [34].

Reporting should consequently make assumptions and negative evidence visible. Future revisions would benefit from recording which competing histories were evaluated, which evidence classes were genuinely capable of contradicting the preferred hypothesis, and which sex, stage or geographic samples remain absent. Such transparency would allow later genomic or museum-based evidence to update a taxonomic hypothesis without retrospectively treating the earlier decision as more certain than it was.

The proposed standards require prospective comparison across insect orders, demographic histories, genomic architectures and diagnostic systems. Because delimitation methods perform differently across simulated scenarios [31], validation should ask not merely whether workflows recover expected taxa, but when they preserve uncertainty appropriately

Conclusion

Increasing genomic resolution changes what insect taxonomy can detect more rapidly than it changes what species status means. Strong delimitation therefore requires separation of observed structure from its biological explanation, taxonomic interpretation, diagnosis and naming. The proposed synthesis treats conflict as evidence to be explained and corroboration as strongest when it crosses genuinely different biological and analytical failure modes. Its value is procedural rather than universal: it identifies where claims should remain conditional, where further evidence is required, and where genomic confidence must not be mistaken for taxonomic certainty.

Acknowledgments: None

Conflict of Interest: None

Financial Support: None

Ethics Statement: None

References

1. 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
2. Vences M, Miralles A, Dufresnes C. Next-generation species delimitation and taxonomy: Implications for biogeography. J Biogeogr. 2024;51(9):1709-22. doi:10.1111/jbi.14807
3. Hartop E, Srivathsan A, Ronquist F, Meier R. Towards large-scale integrative taxonomy (LIT): Resolving the data conundrum for dark taxa. Syst Biol. 2022;71(6):1404-22. doi:10.1093/sysbio/syac033
4. Pyron RA. Unsupervised machine learning for species delimitation, integrative taxonomy, and biodiversity conservation. Mol Phylogenet Evol. 2023;189:107939. doi:10.1016/j.ympev.2023.107939
5. Sukumaran J, Knowles LL. Multispecies coalescent delimits structure, not species. Proc Natl Acad Sci U S A. 2017;114(7):1607-12. doi:10.1073/pnas.1607921114
6. Campillo LC, Barley AJ, Thomson RC. Model-based species delimitation: Are coalescent species reproductively isolated? Syst Biol. 2020;69(4):708-21. doi:10.1093/sysbio/syz072
7. Chambers EA, Hillis DM. The multispecies coalescent over-splits species in the case of geographically widespread taxa. Syst Biol. 2020;69(1):184-93. doi:10.1093/sysbio/syz042
8. Jackson ND, Carstens BC, Morales AE, O'Meara BC. Species delimitation with gene flow. Syst Biol. 2017;66(5):799-812. doi:10.1093/sysbio/syw117
9. Jiao X, Yang Z. Defining species when there is gene flow. Syst Biol. 2021;70(1):108-19. doi:10.1093/sysbio/syaa052
10. Young AD, Gillung JP. Phylogenomics—principles, opportunities and pitfalls of big-data phylogenetics. Syst Entomol. 2020;45(2):225-47. doi:10.1111/syen.12406
11. Joshi M, Espeland M, Huemer P, deWaard J, Mutanen M. Species delimitation under allopatry: Genomic insights within and across continents in Lepidoptera. Insect Syst Divers. 2024;8(5):7. doi:10.1093/isd/ixae027
12. Williams JL, Zhang YM, LaPolla JS, Schultz TR, Lucky A. Phylogenomic delimitation of morphologically cryptic species in globetrotting Nylanderia (Hymenoptera: Formicidae) species complexes. Insect Syst Divers. 2022;6(1):10. doi:10.1093/isd/ixab027
13. Johnston NP, Piwczyński M, Trzeciak P, Walczak K, Szpila K. Integration of mitogenomic and morphological data disentangles the systematics of Pollenia and establishes a revised phylogenetic hypothesis for the Polleniidae. Syst Entomol. 2023;48(2):296-315. doi:10.1111/syen.12576

14. Espeland M, Nakahara S, Zacca T, Barbosa EP, Huertas B, Marín MA, et al. Combining target enrichment and Sanger sequencing data to clarify the systematics of the diverse Neotropical butterfly subtribe Euptychiina (Nymphalidae, Satyrinae). Syst Entomol. 2023;48(4):498-570. doi:10.1111/syen.12590
15. Johnston NP, Pape T, Piwczyński 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
16. Jiao X, Flouri T, Yang Z. Multispecies coalescent and its applications to infer species phylogenies and cross-species gene flow. Natl Sci Rev. 2021;8(12):nwab127. doi:10.1093/nsr/nwab127
17. Koppetsch T, Malinsky M, Matschiner M. Towards reliable detection of introgression in the presence of among-species rate variation. Syst Biol. 2024;73(5):769-88. doi:10.1093/sysbio/syae028
18. Pang XX, Zhang DY. Detection of ghost introgression requires exploiting topological and branch length information. Syst Biol. 2024;73(1):207-22. doi:10.1093/sysbio/syad077
19. Wingert BD, Campbell EO, Acorn JH, Sperling FAH. Genomic integrity of Phyciodes butterfly species in a region of contact (Lepidoptera: Nymphalidae). Insect Syst Divers. 2024;8(2):4. doi:10.1093/isd/ixae006
20. Eberle J, Bazzato E, Fabrizi S, Rossini M, Colomba M, Cillo D, et al. Sex-biased dispersal obscures species boundaries in integrative species delimitation approaches. Syst Biol. 2019;68(3):441-59. doi:10.1093/sysbio/syy072
21. Duran DP, Laroche RA, Roman SJ, Godwin W, Herrmann DP, Bull E, et al. Species delimitation, discovery and conservation in a tiger beetle species complex despite discordant genetic data. Sci Rep. 2024;14:6617. doi:10.1038/s41598-024-56875-9
22. Williams JL, Punnath A, Fernández MB, Calcaterra LA, LaPolla JS, Lucky A. “Picking up signals” in male genital morphospace and integrating phylogenomics to delimit Neotropical Nylanderia Emery species (Hymenoptera: Formicidae). Insect Syst Divers. 2024;8(5):6. doi:10.1093/isd/ixae028
23. Pallarés S, Ortego J, Carbonell JA, Franco-Fuentes E, Bilton DT, Millán A, et al. Genomic, morphological and physiological data support fast ecotypic differentiation and incipient speciation in an alpine diving beetle. Mol Ecol. 2024;33(17):e17487. doi:10.1111/mec.17487
24. Berrilli E, Biondi M, D'Alessandro P, Salvi D. Cryptic, sibling or neither of the two? Integrative species delimitation of Psylliodes flea beetles with overlapping ranges. Zool Scr. 2023;52(3):235-48. doi:10.1111/zsc.12574
25. Çıplak B, Yahyaoğlu Ö, Uluar O, Knowles LL. Understudied regions and messy taxonomy: Geography, not taxonomy, is the best predictor for genetic divergence of the Poecilimon bosphoricus species group. Syst Entomol. 2024;49(2):221-36. doi:10.1111/syen.12615
26. Chen YM, Qu XR, Shang CJ, Peng LF, Zang LS. Identification of morphologically similar genera based on reproductive systems and immature stages: A case study of Anastatus and Mesocomys (Hymenoptera: Eupelmidae). Insect Syst Divers. 2024;8(5):2. doi:10.1093/isd/ixae031
27. Colombo WD, Tribull CM, Waichert C, Azevedo CO. Integrative taxonomy solves taxonomic impasses: A case study from Epyrinae (Hymenoptera, Bethylidae). Syst Entomol. 2022;47(3):504-29. doi:10.1111/syen.12544
28. Balukjian B, Van Dam MH. Using integrative taxonomy to delimit species in the plant bug genus Pseudoloxops (Hemiptera: Miridae) in French Polynesia. Insect Syst Divers. 2024;8(5):1. doi:10.1093/isd/ixae012
29. Prebus MM. Phylogenomic species delimitation in the ants of the Temnothorax salvini group (Hymenoptera: Formicidae): An integrative approach. Syst Entomol. 2021;46(2):307-26. doi:10.1111/syen.12463
30. Barley AJ, Brown JM, Thomson RC. Impact of model violations on the inference of species boundaries under the multispecies coalescent. Syst Biol. 2018;67(2):269-84. doi:10.1093/sysbio/syx073
31. Luo A, Ling C, Ho SYW, Zhu CD. Comparison of methods for molecular species delimitation across a range of speciation scenarios. Syst Biol. 2018;67(5):830-46. doi:10.1093/sysbio/syy011
32. Zhang H, Bu W. Exploring large-scale patterns of genetic variation in the COI gene among Insecta: Implications for DNA barcoding and threshold-based species delimitation studies. Insects. 2022;13(5):425. doi:10.3390/insects13050425
33. Ma Z, Ren J, Zhang R. Identifying the genetic distance threshold for Entiminae (Coleoptera: Curculionidae) species delimitation via COI barcodes. Insects. 2022;13(3):261. doi:10.3390/insects13030261

34. Renner SS, Scherz MD, Schoch CL, Gottschling M, Vences M. Improving the gold standard in NCBI GenBank and related databases: DNA sequences from type specimens and type strains. Syst Biol. 2024;73(2):486-94. doi:10.1093/sysbio/syad068
35. Mulvaney LRJ, St Laurent RA, László GM, Schintlmeister A. Integrative taxonomic revision of the Cerurina–Afrocerura genus-pair (Lepidoptera, Noctuoidea, Notodontidae, Cerurinae). Insect Syst Divers. 2024;8(5):5. doi:10.1093/isd/ixae029