Eurasia Specialized Veterinary Publication

Entomological Research Letters

ISSN: 3108-639X

2026, Volume 6, Issue 1, Page No: 71-81
Copyright CC BY-NC-SA 4.0
Available online at: www.esvpub.com/

Probabilistic Inference Offers a Stronger Basis for Integrative Insect Species Boundaries Than Binary Taxonomic Decisions

Pieter Botha[1]*, Anel Van Wyk[1], Johan Marais[2]

[1]Department of Probabilistic Inference and Species Boundaries, Faculty of Biosciences, Stellenbosch University, Stellenbosch, South Africa.
[2]Department of Integrative Taxonomy and Binary Decisions, Faculty of Life Sciences, University of Pretoria, Pretoria, South Africa.

***E-mail** ✉ pieter.botha@sun.ac.za

ABSTRACT

Integrative insect systematics increasingly combines morphology, molecular and genomic data, ecology, geography, and behaviour, yet the transition from heterogeneous evidence to a categorical species decision remains conceptually uneven. Different evidence classes measure different biological properties, vary in diagnostic resolution, and can be statistically or biologically dependent. We develop a proposed Bayesian conceptual architecture in which species boundaries are treated as competing hypotheses supported to differing degrees rather than as immediate binary outcomes. Prior information is separated from likelihood evidence; morphological, molecular-genomic, and ecological-behavioural contributions remain analytically distinct; evidence dependence and model adequacy are assessed before synthesis; and posterior boundary support is separated from the downstream decision to recognize or reject a species hypothesis. The framework also distinguishes lack of detected differentiation from affirmative counterevidence and maintains sampling, model, geographic, and taxonomic uncertainty after evidence integration. Insect examples demonstrate why mitochondrial divergence, genomic structure, morphology, ecological association, and behavioural differentiation cannot be treated as interchangeable votes. The framework is not an empirically validated delimitation algorithm and does not prescribe universal probability thresholds. Instead, it provides a testable systems-level structure for making evidence provenance, conditionality, dependence, and uncertainty explicit in integrative species-boundary inference.

Keywords: Bayesian species delimitation, Integrative taxonomy, Insect systematics, Genomic discordance, Evidence dependence, Taxonomic uncertainty.

Received: 02 April 2026
Revised: 25 May 2026
Accepted: 28 May 2026

How to Cite This Article: Botha P, Van Wyk A, Marais J. Probabilistic Inference Offers a Stronger Basis for Integrative Insect Species Boundaries Than Binary Taxonomic Decisions. Entomol Appl Sci Lett. 2026;6(1):71-81. https://doi.org/10.51847/gTTu0RTyZn

Introduction

Species delimitation increasingly operates at the intersection of multiple data types, but integrating those data remains more difficult than accumulating them. Contemporary integrative-taxonomy scholarship identifies heterogeneous species concepts, non-universal markers, and partly subjective evidence fusion as continuing challenges [1]. In insect systematics specifically, phylogenomics is increasingly interpreted alongside morphology, behaviour, ecology, development, spatial information, and natural history rather than as an isolated source of taxonomic authority [2]. The unresolved problem is therefore not whether more evidence should be gathered, but how biologically different observations should alter confidence in competing species-boundary hypotheses.

© 2026 Entomological Research Letters

Empirical insect studies expose this inferential problem. In Polynesian Pseudoloxops, multilocus relationships were considered together with morphometrics, genital morphology, host association, and geography [3]. Phylogenomics, morphology, population-genetic evidence, and biogeography substantially reorganized the historical interpretation of the Acromyrmex octospinosus complex [4]. Conversely, pronounced morphological variability in figitid wasps illustrates how character states conventionally used for diagnosis can overlap or vary within presumed species [5]. These examples demonstrate neither that integration always splits nor that it always lumps taxa; they demonstrate that different observations can change the credibility of a boundary hypothesis in different directions.

This article therefore develops a proposed probabilistic evidence architecture for integrative insect species boundaries. Its central distinction is between observations, evidence-class-specific support, dependence among evidence, conditional posterior support, and the eventual taxonomic decision. The objective is not to replace formal taxonomy with a probability value, but to prevent categorical nomenclatural outputs from concealing uncertainty already present in the evidence used to reach them.

Materials and Methods

The framework is a synthesis constructed from methodological species-delimitation research and insect case studies rather than from a newly assembled dataset. Its Bayesian language is conceptual: no numerical priors, likelihood functions, posterior probabilities, Bayes factors, or decision thresholds are estimated here. Probabilistic species delimitation nevertheless provides an established methodological basis for retaining multiple possible species identities rather than forcing immediate categorical assignment [6]. The present model extends that logic to the provenance and interpretation of heterogeneous integrative evidence.

Three operations are kept separate. First, prior information represents knowledge available before evaluating the focal evidence, including geography, previously proposed taxonomy, sampling design, and biologically plausible demographic histories. Second, evidence is partitioned into morphology, molecular-genomic information, and ecological-behavioural information rather than collapsed into a common vote. Third, the dependence and adequacy of those evidence sources are examined before any proposed posterior synthesis.

This separation is necessary because species-boundary models themselves differ biologically. Migration-aware multispecies-coalescent approaches show that inference changes when post-divergence gene flow is admitted rather than implicitly excluded [7]. Evaluations of genomic delimitation strategies likewise show that alternative methods have different sensitivities to population structure and divergence histories [8]. Most importantly, multispecies-coalescent analyses can delimit genetic structure that does not intrinsically identify whether the inferred units are independently evolving species [9]. Models that incorporate gene flow therefore address a demographic history rather than turning migration itself into a binary species criterion [10].

Accordingly, the proposed architecture treats every inferential output as conditional on the data, model, sampling design, and biological interpretation that generated it. A highly resolved molecular partition is thus evidence about lineage structure; it becomes evidence about a species boundary only through an explicit interpretive step.

Results and Discussion

Prior information can improve biological coherence while also introducing circularity. Geographic subdivisions or historically named entities may define candidate units before molecular analysis, yet geographically widespread taxa can contain continuous or structured populations that model-based delimitation interprets as separate entities. Reanalysis of such conditions demonstrates that isolation by distance, geographic structure, and a priori grouping can contribute to oversplitting under some multispecies-coalescent applications [11]. Existing names therefore should not be treated automatically as high-confidence priors.

Sampling completeness is a second prior-model interface. Recent analysis of gene-flow inference shows that omitted lineages can alter the inferred identity or magnitude of introgression events, with effects depending on demographic assumptions and analytical method [12]. In the proposed architecture, incomplete taxon sampling consequently remains an uncertainty attribute attached to the model rather than disappearing after a posterior value is produced.

Prior information is therefore used to define plausible hypotheses and expose assumptions, not to predetermine their answer. Geography, previous taxonomy, known dispersal constraints, and sampling gaps can shape what models deserve consideration, but none is converted here into an invented numerical prior.

Morphological likelihood

Morphology contributes to species-boundary inference only in relation to the expected variation of the character system being measured. Three-dimensional analysis of male genital morphology in Nylanderia showed that some phylogenomic distinctions had corresponding morphological separation whereas others did not [13]. The relevant signal is therefore not simply "different" versus "same," but whether observed variation is more compatible with differentiation among candidate lineages than with variation expected within them.

Integrative work on Nearctic Phymata similarly illustrates that mitochondrial, nuclear, morphometric, and ecological evidence can produce partially discordant partitions, while morphometrics can help adjudicate particular molecular ambiguities [14]. Conversely, phylogenomic investigation of morphologically cryptic Nylanderia complexes demonstrates that weak external diagnosis can coexist with substantial genomic lineage structure [15].

The proposed model therefore distinguishes absence of detected morphological differentiation from morphological evidence against a boundary. These are not equivalent states. Failure to recover a difference may reflect character conservatism, sex restriction, developmental stage, measurement resolution, or inadequate sampling. Strong overlap across well-sampled, biologically relevant characters may provide counterevidence, but only under conditions that justify that interpretation. Morphological likelihood must consequently retain the taxon, character system, life stage, sex, and sampling design under which the observation was obtained.

Molecular and genomic likelihood

Molecular evidence often provides the highest-dimensional description of lineage structure, but dimensionality is not equivalent to taxonomic certainty. In the Bactrocera dorsalis complex, genomic evidence exposed limits of mitochondrial barcode diagnosis in a globally important pest complex, demonstrating that mitochondrial similarity or divergence need not map cleanly onto genome-wide lineage history [16]. The proposed architecture therefore keeps mitochondrial, reduced-representation, multilocus, and whole-genome evidence distinguishable rather than treating them collectively as a single "genetic" likelihood.

Whole-genome data can sharpen separation among levels of divergence while simultaneously revealing substantial structure within currently recognized species. Analyses of the Papilio machaon complex illustrate this dual role: genome-wide information can resolve evolutionary partitions yet still require interpretation of which partitions represent within-species differentiation and which support species-level hypotheses [17]. More data therefore reduce some forms of sampling uncertainty without eliminating the biological problem of assigning structure to hierarchical level.

Analytical uncertainty also persists within genomic data. Comparisons of gene-flow inference in Drosophila demonstrate that different methods can infer different cross-species migration histories and exhibit different performance under tested simulations [18]. Thus, disagreement among demographic models is not an inconvenience that should be averaged away; it is evidence about model sensitivity. At a different scale, broad analysis of insect COI data shows extensive heterogeneity in intraspecific divergence and undermines the assumption that a single sequence-distance threshold can universally partition insect species [19].

For this framework, molecular-genomic likelihood therefore has at least four separable dimensions: marker provenance, genomic breadth, demographic-model dependence, and the biological level represented by the recovered structure. None supplies a universally valid species threshold. Instead, each constrains how strongly a molecular pattern can support a particular boundary hypothesis.

Because these dimensions are easy to collapse under the label "molecular support," their interpretive differences and principal gaps are more efficiently separated in a coverage-and-gap map than repeated narratively.

Table 1. Molecular evidence states that alter the inferential meaning of genomic differentiation in insect species-boundary analysis

Boundary interpretation and decisive check	Principal uncertainty or gap	Molecular-genomic state with evidential standing	Biological setting with observed basis

Supports marker-specific differentiation, not automatically a species boundary **Test against independent nuclear or genome-wide data where available**	Introgression, lineage sorting, sex-biased history, marker genealogy	Mitochondrial partition without equivalent genome-wide resolution Evidence-derived synthesis	Closely related or introgressed insect lineages Barcode or mitochondrial clustering contrasted with nuclear/genomic evidence
Demonstrates population or lineage structure whose taxonomic level remains to be interpreted **Compare structure with morphology, geography, demography, and other independent evidence**	Sampling density, geographic continuity, unsampled populations	Genome-wide structure nested within a named species Evidence-derived synthesis	Geographically structured populations or recent divergence Broad genomic differentiation among sampled populations
Strengthens robustness of the genomic component of a boundary hypothesis **Evaluate model adequacy rather than count agreeing analyses as separate votes**	Shared assumptions or common data can make model agreement non-independent	Genomic separation concordant across several analytical models Proposed inferential classification	Candidate lineages with sufficient genomic sampling Similar partitions recovered across distinct model formulations
Indicates model-sensitive rather than settled demographic evidence **Conduct sensitivity analysis across biologically plausible models**	Identifiability, taxon sampling, population-size assumptions	Gene-flow history changes with analytical method Evidence-derived synthesis	Recently diverged or introgressing lineages Competing demographic methods infer different migration events or directions
Provides an operational classification rule but weak justification for a universal biological boundary **Replace universal cutoff logic with taxon- and evidence-conditioned interpretation**	Strong variation in within-species divergence among taxa	Fixed genetic-distance cutoff applied across heterogeneous taxa Evidence-derived synthesis	DNA-barcode delimitation across broad insect diversity One distance threshold separates candidate units
Improves description of structure but does not independently establish species status **Couple genomic resolution to evidence from other biological dimensions and explicit species hypotheses**	Structure–species distinction persists	Additional loci increase resolution without resolving biological level Proposed inferential classification	Genome-rich species complexes Narrower confidence in lineage structure as genomic information increases

Ecological and behavioral evidence

Ecological and behavioural observations occupy a different inferential position from sequence or morphology because they may describe interactions relevant to lineage persistence without directly measuring ancestry. In the beet-leafhopper species complex, courtship vibration and morphometric evidence corroborated molecularly delimited operational units and helped resolve distinctions that were ambiguous from other characters [20]. Such behavioural evidence can be especially informative when signals participate in mate recognition, but signal divergence alone does not establish complete reproductive isolation or quantify realized gene flow.

Ecological information is typically more indirect. In Ulopsina leafhoppers, molecular delimitation and morphology were combined with ecological-niche modelling and biogeographic reconstruction, allowing environmental and geographic context to be related to inferred lineages [21]. Modelled habitat suitability, however, is not itself an isolation mechanism. Geographic separation, host association, or niche differentiation may explain why a boundary hypothesis is plausible, may generate an alternative population-structure explanation, or may identify conditions under which further sampling is needed.

The proposed architecture therefore separates contextual ecological evidence from process-proximal behavioural evidence. Neither category receives an automatic evidential rank. Behavioural characters may be biologically close to premating isolation but restricted to one sex, stage, season, or experimental context. Ecological differences may be broadly observable yet partly inherited from geography rather than independently generated.

Their contribution depends on whether they supply genuinely new information or reproduce structure already represented by sampling location, morphology, or molecular ancestry.

This distinction is sufficiently multidimensional that a state-space representation is needed to preserve both inferential relevance and dependence rather than placing ecology and behaviour on a single evidence ladder.

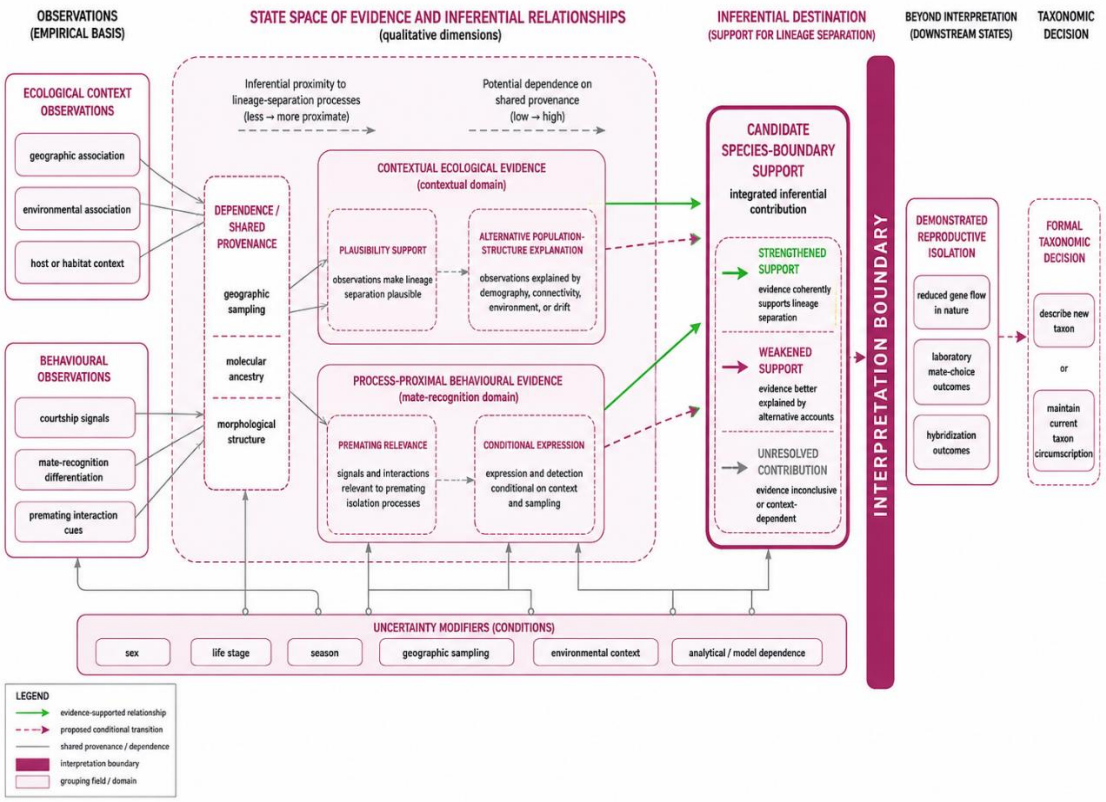


Figure 1. Where ecological context and mate-recognition signals change species-boundary support without becoming taxonomic verdicts

Evidence dependence

Integrative evidence is most informative when apparent corroboration does not merely reproduce the same underlying structure. Genetic delimitation and reproductive isolation provide a direct example. Comparative work in Drosophila found substantial but incomplete concordance between coalescent species hypotheses and reproductive-isolation criteria [22]. Agreement therefore can strengthen interpretation, but neither criterion can simply be substituted for the other.

Dependence also enters through models. Simulation studies show that violations of assumptions used by multispecies-coalescent analyses can alter inferred boundaries [23]. Existing taxonomy may then influence how molecular partitions are interpreted: butterfly analyses comparing classical taxonomy with multispecies-coalescent delimitation show that historical classifications and molecular hypotheses can be reconciled without treating either as automatically decisive [24]. Marker scale creates another dependency. In Chanohirata leafhoppers, mitochondrial and nuclear loci, delimitation procedures, morphology, and biogeography contributed overlapping but nonidentical information [25].

The proposed dependence audit therefore asks whether two observations have separate biological provenance, whether they share specimens or geography, whether one was used to define the other, and whether they depend on common model assumptions. Concordance surviving this audit is more informative than simple agreement among method outputs. Conversely, discordance can identify an informative boundary condition rather than a failed analysis.

Dependence is also asymmetric. Two concordant molecular analyses may share most of their information, whereas a narrowly sampled behavioural assay can contribute genuinely different evidence despite its smaller apparent scope. The framework therefore rejects any assumption that data volume, methodological sophistication, or dimensionality determines evidential independence. What matters is whether an observation changes the relative

plausibility of competing boundary hypotheses after information already represented elsewhere has been accounted for.

These distinctions require adjudication across the source of dependence rather than a count of agreeing datasets.

Table 2. Adjudicating when concordant evidence strengthens an insect species boundary and when it duplicates the same signal

Uncertainty or boundary and decisive check	Observable basis and warranted reading	Analytical setting
Independent biological corroboration		
Proposed adjudication state		
Independence must be justified, not assumed **Test shared sampling, coding, and model inputs before integration**	Concordant lineage pattern in data with separate provenance Can strengthen support for the same boundary hypothesis	Different character systems sampling distinct biological processes
Shared geographic provenance		
Evidence-derived synthesis		
Isolation by distance or unsampled intermediates can mimic discrete units **Increase spatial sampling and evaluate continuous structure**	Several evidence classes reproduce the same spatial partition Agreement may partly reflect one geographic structure	Molecular, morphological, or ecological differences covary with collection locality
Shared specimens and ancestry		
Proposed adjudication state		
Correlated genealogies or organismal covariance **Model covariance or retain evidence as a linked block**	Repeated concordant clusters Replication increases description but not necessarily independent support	Multiple molecular markers or traits measured on the same individuals
Model-dependent concordance		
Evidence-derived synthesis		
Common misspecification can propagate across methods **Compare biologically different models and conduct sensitivity tests**	Similar species partitions across methods Supports robustness only within the shared assumption space	Several analyses use related coalescent assumptions
Criterion concordance with exceptions		
Evidence-derived synthesis		
Taxon and assay scope constrain transferability **Preserve exceptions and avoid treating either criterion as universal ground truth**	Related criteria converge but retain discordant cases Supports a relationship between criteria without identity	Genetic boundary support and reproductive-isolation evidence mostly agree
Cross-class discordance		
Proposed adjudication state		
Could reflect introgression, plasticity, conserved morphology, sampling, or error **Investigate the mechanism of discordance before categorical adjudication**	Competing lineage configurations Narrows plausible explanations and exposes scale or process differences	Morphology, molecules, ecology, or behaviour support different partitions

Posterior species-boundary support

Posterior species-boundary support is the point at which the proposed architecture combines evidence while retaining its provenance. Integrative analysis of Chinese Plautia demonstrates that morphology, mitochondrial information, genome-wide SNP data, and multiple delimitation approaches can converge on coherent lineage hypotheses [26]. Such convergence provides a biological rationale for increased confidence, but the number of agreeing methods is not itself a posterior probability.

Accordingly, proposed posterior support is conditional on four preceding questions: whether the prior is defensible, whether each evidence class measures what its interpretation assumes, whether dependence has been addressed, and whether relevant alternative models remain plausible. A posterior is therefore not a taxonomic truth state. Probabilistic delimitation itself is model-conditional [6], while assumption violations can modify model-based boundaries [23].

Posterior support can also decrease. Morphological overlap, evidence for continuous geographic structure, or a demographic model accommodating introgression may make a previously simple split less plausible without proving conspecificity. The architecture therefore permits evidence to strengthen, weaken, or leave essentially unchanged the relative support for a candidate boundary. This directional flexibility is essential because integrative taxonomy must be capable of both splitting and lumping historical classifications rather than embedding an implicit preference for discovery of additional species.

The figure below compresses this distinction by making evidence accumulation, dependence correction, alternative explanations, and the taxonomic interpretation boundary simultaneously visible.

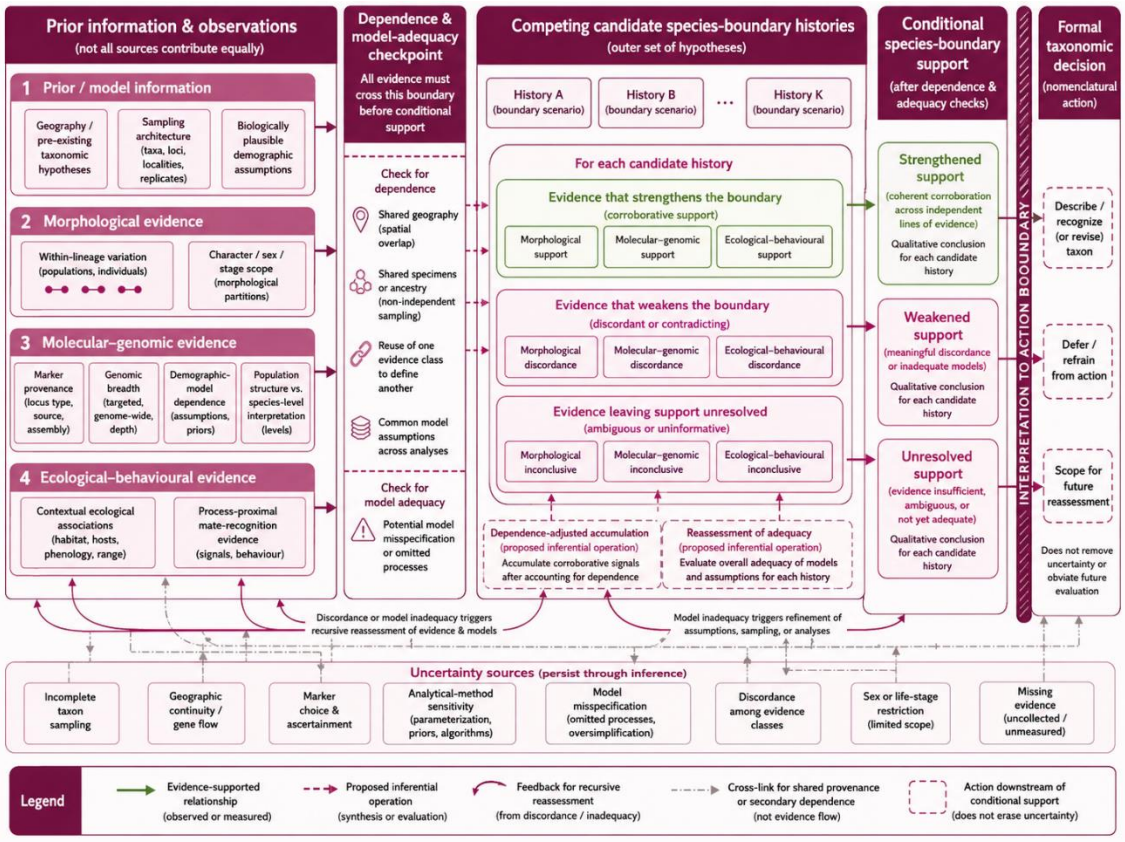


Figure 2. Conditional support survives evidence integration only when provenance, dependence, and alternative species histories remain visible

Decision thresholds versus continuous support

Taxonomic practice ultimately requires decisions, but an operational decision need not imply that biological support was naturally binary. Comparative genomic work on allopatric Lepidoptera found that barcode and genomic perspectives were not uniformly concordant, illustrating how geographic separation can generate difficult cases in which divergence is continuous while direct tests of reproductive interaction are unavailable [27]. A universal threshold would suppress precisely the uncertainty that makes those cases informative.

The distinction is reinforced by multimodal machine-learning work combining molecular and morphological information for species identification [28]. Improved classification of specimens assigned to accepted labels addresses identification performance; it does not independently validate the historical species boundaries underlying those labels. Likewise, broad variation in insect COI distances makes a universal barcode cutoff difficult to defend [19].

The proposed framework therefore separates a continuous inferential state from a downstream decision rule. A decision can be necessary for naming, identification, regulation, or communication while remaining explicitly conditional on current evidence.

The resulting states are best distinguished by whether the evidence changes support for a boundary and whether a categorical action is warranted, rather than by assigning an unsupported score.

Table 3. Keeping continuous species-boundary support distinct from the categorical actions that taxonomy eventually requires

Decision consequence, boundary or uncertainty, and research response	Inference with evidential standing	Evidential basis and warranted reading
Defer categorical confidence where possible **Unresolved discordance is informative** **Target the evidence class capable of discriminating alternatives**	Weak or conflicting support Proposed state	Major evidence classes disagree or key sampling is absent Several biological histories remain plausible
Avoid treating method count as accumulated certainty **Dependence limits information gain** **Reassess provenance and shared assumptions**	Concordant but dependent support Proposed state	Multiple outputs agree but share data, geography, or assumptions Evidence is coherent but partly redundant
Taxonomic action may be more defensible **Model and sampling uncertainty remain** **Seek replication across geography, stage, sex, or data type where material**	Broad heterogeneous corroboration Proposed state	Distinct evidence classes converge after dependence checks Boundary hypothesis gains conditional support
Any categorical threshold is assumption-sensitive **Reproductive interaction may be unobservable** **Expand genomic, morphological, ecological, and geographic comparisons**	Allopatric continuous divergence Evidence-derived synthesis	Geographic separation accompanies molecular differentiation Species-level versus population-level interpretation remains difficult
Does not by itself validate those classes as species **Training labels inherit prior taxonomy** **Validate delimitation independently of classifier performance**	High identification accuracy Evidence-derived synthesis	Combined traits classify specimens into accepted labels Demonstrates recognition of existing classes
Produces a categorical operational outcome **Decision does not erase inferential uncertainty** **Report rationale, alternatives, and residual uncertainty**	Formal taxonomic decision Proposed state	Evidence has been interpreted under an explicit criterion Converts conditional inference into nomenclatural action

Uncertainty communication

Uncertainty should be carried forward rather than appended as a generic limitation. Next-generation delimitation scholarship emphasizes separating lineage discovery, taxonomic interpretation, and broader geographic or evolutionary context [29]. The proposed reporting scheme therefore distinguishes evidence uncertainty, model uncertainty, sampling uncertainty, and decision uncertainty.

These components can move independently. Missing taxa may alter gene-flow inference [12], whereas analytical method can change the inferred history even when genomic observations are held broadly comparable [18]. A taxonomic account should consequently state what evidence supports the boundary, what important alternatives remain, and which uncertainty attaches to the data versus the interpretive model. This approach does not prevent a taxonomic decision; it makes the inferential conditions of that decision recoverable.

A useful uncertainty statement should also identify scale. Uncertainty about a barcode genealogy is not equivalent to uncertainty about whether two populations exchange genes, and neither is equivalent to uncertainty about whether the available evidence justifies a nomenclatural act. Similarly, a result obtained from adult males cannot automatically resolve larval diagnosis, female morphology, or seasonal behavioural variation. Making those scales explicit prevents a local observation from silently becoming a species-wide assertion.

The proposed architecture reframes integrative delimitation as conditional evidence synthesis rather than evidence voting. Insect studies combining molecular, morphological, ecological, and behavioural information demonstrate

that different evidence classes can contribute complementary information to the same boundary problem [3, 13, 20]. Their value lies not in the number of channels agreeing but in what biological uncertainty each channel resolves.

Molecular evidence requires the same discipline. Barcode discordance, method-sensitive gene-flow inference, and heterogeneous within-species COI divergence demonstrate that the molecular domain itself cannot be treated as a homogeneous source of certainty [16, 18, 19]. Probabilistic delimitation provides a formal precedent for retaining uncertainty among hypotheses [6], but the present architecture extends that logic conceptually rather than claiming a validated new estimator.

The architecture also clarifies what stronger evidence would mean. Stronger support is not synonymous with more characters, more loci, or more algorithms. It arises when additional observations discriminate among plausible alternative explanations while surviving sensitivity to sampling and model choice. This formulation turns discordance into a research guide: instead of asking which data type should overrule the others, investigators can ask which competing explanation generates the discordance and what observation would most effectively distinguish those explanations.

Three limitations are central. First, the evidence base spans several insect orders but does not establish universality across Insecta, and immature stages are less well represented than adults. Second, geographic and taxonomic sampling can alter inference; missing lineages are therefore not merely a logistical weakness [12]. Third, reproductive-isolation and coalescent evidence show substantial but incomplete correspondence [22], so no evidence class can be elevated to a universal ground truth.

Validation should test whether explicitly modelling evidence provenance and dependence changes reproducibility, sensitivity to sampling, or stability of taxonomic conclusions relative to simpler integration strategies. Such tests should span recent and deep divergences, sympatric and allopatric systems, sexually dimorphic and cryptic taxa, and datasets with known discordance. Until such validation exists, the architecture is best regarded as a proposed framework for transparent reasoning rather than a demonstrated improvement in delimitation accuracy. Its practical value is the discipline of keeping observation, model, support, and decision separable, consistent with broader arguments that lineage inference and taxonomic interpretation should not be collapsed [29].

Conclusion

Integrative insect species delimitation becomes more transparent when heterogeneous observations are treated as conditional evidence rather than interchangeable votes. The proposed architecture separates prior assumptions, evidence-class-specific likelihood, dependence, model adequacy, conditional species-boundary support, and the downstream taxonomic decision. It preserves discordance, sampling limitations, and alternative interpretations instead of forcing them into a binary endpoint. This model does not define a universal species criterion or probability threshold, and it is not yet empirically validated as superior to existing workflows. Its contribution is a testable structure for asking what each observation supports, what information it shares with other evidence, which assumptions condition the inference, and what uncertainty remains when a formal taxonomic decision is ultimately made.

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. Ribeiro TMA, Espíndola A. Integrated phylogenomic approaches in insect systematics. Curr Opin Insect Sci. 2024;61:101150. doi:10.1016/j.cois.2023.101150
3. 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
4. Mera-Rodríguez D, Fernández-Marín H, Rabeling C. Phylogenomic approach to integrative taxonomy resolves a century-old taxonomic puzzle and the evolutionary history of the Acromyrmex octospinosus species complex. Syst Entomol. 2025;50(3):469-94. doi:10.1111/syen.12665
5. Ferrer-Suay M, Heimpel GE, Rakhshani E, Selfa J. Deciphering morphological variability: Addressing taxonomic ambiguities in contemporary species delimitation (Hymenoptera, Figitidae). Insects. 2026;17(1):54. doi:10.3390/insects17010054
6. Hua X, Moritz C. A phylogenetic approach to delimitate species in a probabilistic way. Syst Biol. 2025;74(5):723-39. doi:10.1093/sysbio/syaf004
7. Kornai D, Jiao X, Ji J, Flouri T, Yang Z. Hierarchical heuristic species delimitation under the multispecies coalescent model with migration. Syst Biol. 2024;73(6):1015-37. doi:10.1093/sysbio/syae050
8. Hausdorf B. Species delimitation using genomic data: Options and limitations. Mol Ecol. 2025;34(8):e17717. doi:10.1111/mec.17717
9. 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
10. 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
11. 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
12. Cheng S, Flouri T, Zhu T, Yang Z. The impact of incomplete taxon sampling on inference of gene flow by Bayesian and summary methods using genomic sequence data. Syst Biol. 2026:syag023. doi:10.1093/sysbio/syag023
13. 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
14. Masonick P, Weirauch C. Integrative species delimitation in Nearctic ambush bugs (Heteroptera: Reduviidae: Phymatinae): Insights from molecules, geometric morphometrics and ecological associations. Syst Entomol. 2020;45(1):205-23. doi:10.1111/syen.12388
15. 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
16. Doorenweerd C, San Jose M, Geib S, Barr N, Rubinoff D. Genomic data reveal new species and the limits of mtDNA barcode diagnostics to contain a global pest species complex (Diptera: Tephritidae: Dacinae). Syst Entomol. 2024;49(2):279-93. doi:10.1111/syen.12616
17. Cassar LF, Nabholz B, Reboud EL, Chevalier E, Lafon BJ, Cotton AM, et al. Whole-genome data shed light on speciation and within-species differentiation of the Papilio machaon complex around the Mediterranean Basin. Syst Entomol. 2025;50(4):692-712. doi:10.1111/syen.12675
18. Ji J, Roberts T, Flouri T, Yang Z. Inference of cross-species gene flow using genomic data depends on the methods: Case study of gene flow in Drosophila. Syst Biol. 2025;74(4):566-82. doi:10.1093/sysbio/syaf019
19. 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
20. Sinaiko G, Dietrich CH. Untangling the beet leafhopper (Hemiptera: Cicadellidae) species complex: A case study in integrative taxonomy. Insect Syst Divers. 2025;9(3):4. doi:10.1093/isd/ixaf017
21. Wang D, Dietrich CH, Xue Q, Cao Y, Zhang Y. Phylogeny, molecular species delimitation, and biogeography of the cryptic subgenus Ulopsina (Hemiptera: Cicadellidae: Ulopinae). Insect Syst Divers. 2025;9(3):10. doi:10.1093/isd/ixaf020
22. 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

23. 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
24. Matos-Maraví P, Wahlberg N, Antonelli A, Penz CM. Species limits in butterflies (Lepidoptera: Nymphalidae): Reconciling classical taxonomy with the multispecies coalescent. Syst Entomol. 2019;44(4):745-56. doi:10.1111/syen.12352
25. Wang D, Dietrich CH, Zhang Y. Phylogenetics, historical biogeography and molecular species delimitation of Chanohirata (Hemiptera: Cicadellidae: Deltocephalinae: Penthimiini). Syst Entomol. 2024;49(1):173-92. doi:10.1111/syen.12612
26. Zhang Y, Li Y, Liang J, Liu H, Xue H, Wang S, et al. Integrative taxonomy for species delimitation in genus Plautia (Hemiptera: Heteroptera: Pentatomidae) from China. Insect Syst Divers. 2026;10(3):ixag023. doi:10.1093/isd/ixag023
27. 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
28. Yang B, Zhang Z, Yang CQ, Wang Y, Orr MC, Wang H, et al. Identification of species by combining molecular and morphological data using convolutional neural networks. Syst Biol. 2022;71(3):690-705. doi:10.1093/sysbio/syab076
29. 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