Eurasia Specialized Veterinary Publication

Entomological Research Letters

ISSN: 3108-639X

2023, Volume 3, Issue 1, Page No: 142-153
Copyright CC BY-NC-SA 4.0
Available online at: www.esvpub.com/

Distinguishing Incomplete Lineage Sorting, Introgression, and Analytical Error Is Essential for Interpreting Discordant Insect Phylogenies

Daniel Fischer[1]*, Laura Meier[1], Stefan Koch[2], Thomas Braun[3]

[1]Department of Phylogenomic Discordance and Lineage Sorting, Faculty of Life Sciences, University of Freiburg, Freiburg, Germany.
[2]Department of Introgression and Analytical Error, Faculty of Biosciences, Technical University of Munich, Munich, Germany.
[3]Department of Phylogenetic Interpretation and Discordance, Faculty of Sciences, University of Kiel, Kiel, Germany.

***E-mail** ✉ daniel.fischer@pharmazie.uni-freiburg.de

ABSTRACT

Phylogenomic discordance is increasingly an expected feature of insect systematics rather than an exceptional failure of inference. The difficulty is therefore not simply detecting conflict among loci, sites, gene trees, or analytical methods, but determining what that conflict can legitimately reveal about species history. Incomplete lineage sorting (ILS), introgression, gene-tree estimation error, model misspecification, orthology problems, rate heterogeneity, uneven taxon or locus sampling, missing data, and concentrated character support can generate partially overlapping signatures. Consequently, discordance itself is not a diagnosis. This Forum-compatible conceptual article develops a differential framework for separating these sources before phylogenetic conflict is translated into species-tree or classificatory conclusions. Evidence from alpine *Parnassius*, Aphididae, Odonata, *Heliconius*, *Anastrepha*, and the *Bactrocera dorsalis* complex illustrates biologically distinct discordance regimes, from predominantly ILS-associated rapid radiation to histories in which ancestral or ongoing gene flow contributes materially to conflicting relationships. Methodological evidence further shows that support, concordance, topology frequency, branch length, site pattern, and network structure answer different inferential questions and remain sensitive to distinct failure modes. We argue that interpretation should proceed through explicit separation of observed conflict, analytical adequacy, biological-model compatibility, and downstream species-history inference. The resulting hierarchy is proposed as an organizing architecture rather than an empirically validated diagnostic algorithm. Its central objective is to preserve mixed-cause and genuinely indeterminate outcomes instead of forcing discordant insect phylogenies into a premature binary explanation.

Keywords: Incomplete lineage sorting, Introgression, Insect phylogenomics, Gene-tree discordance, Species-tree inference, Analytical error

Received: 01 April 2023
Revised: 24 June 2023
Accepted: 29 June 2023

How to Cite This Article: Fischer D, Meier L, Koch S, Braun T. Distinguishing Incomplete Lineage Sorting, Introgression, and Analytical Error Is Essential for Interpreting Discordant Insect Phylogenies. Entomol Res Lett. 2023;3(1):142-53. https://doi.org/10.51847/VsTZEPRiZq

Introduction

Genome-scale datasets have not abolished phylogenetic disagreement; instead, they have revealed that discordance may originate from several biological and analytical processes whose observable signatures overlap. ILS, introgression, stochastic gene-tree variation, model misspecification, data-treatment decisions, and other sources of error can all separate inferred gene histories from a single species-tree representation, while methods for detecting introgression interrogate different combinations of topology frequencies, branch lengths, site

© 2023 Entomological Research Letters

patterns, or model likelihoods. Discordance should therefore be treated as an observation requiring differential diagnosis rather than as direct evidence for one historical mechanism [1, 2]. The central problem is inferential: an analysis may establish that loci disagree without establishing why they disagree, and a statistic capable of rejecting one simple history need not uniquely identify the alternative history responsible.

Insects provide particularly clear reasons not to assign cause from conflict alone. Rapidly radiating alpine *Parnassius* butterflies show extensive phylogenomic discordance for which ILS was the primary supported explanation and introgression was not prominent in the analysed dataset, whereas deep relationships within Aphididae were complicated by gene-tree discordance, introgression, and conditions associated with the gene-tree anomaly zone [3, 4]. These systems are contrasts, not estimates of how frequently particular causes occur across Insecta. They demonstrate that similar high-level observations—unstable relationships or disagreement among genealogies—can arise within materially different evolutionary and analytical contexts.

The contribution developed here is consequently narrower than a general review of phylogenomic conflict. The article asks how investigators should distinguish evidence compatible with ILS from evidence favoring introgression and from conflict plausibly amplified or generated by analytical and sampling processes. It separates evidence classes that are often conflated: branch support from concordance, coalescent compatibility from causal proof, introgression detection from reconstruction of its direction or timing, and inferred species trees from taxonomic decisions. The eventual discordance-triage hierarchy is explicitly proposed. It is intended to organize established evidence and expose unresolved alternatives, not to supply universal thresholds or claim prospective diagnostic performance.

Sources of phylogenetic discordance

Biological discordance encompasses histories at very different temporal and taxonomic scales. Phylogenomic analyses of Odonata provide evidence that ancestral introgression can contribute to deep evolutionary relationships; full-likelihood analysis of the *Heliconius erato–sara* group supports a history involving species divergence and multiple introgression events; and transcriptomic phylogenomics of *Anastrepha* indicates that ILS, ancestral introgression, and more recent gene flow can coexist during diversification [5-7]. These examples establish possibility and coexistence, not a common prevalence across insect orders. They also caution against treating ILS and introgression as mutually exclusive labels applied to an entire radiation. A lineage history can contain retention of ancestral polymorphism at one set of internodes and interspecific exchange at another, while individual genomic regions may retain different portions of that history.

The inferential problem becomes sharper when conventional indicators of certainty disagree with alternative historical models. In the *Bactrocera dorsalis* species complex, strongly supported concatenated relationships conflicted with coalescent and network interpretations in a system showing evidence of interspecific gene flow [8]. Importantly, the direction of that flow was not unambiguously resolved. The case illustrates why bootstrap or analogous support should be separated from historical adequacy: high support describes repeatability or signal under an inferred model and dataset, but it cannot guarantee that the model class captures reticulation. Conversely, the existence of discordance does not make a network automatically preferable. A reticulate interpretation requires positive evidence that distinguishes it from coalescent variation and from analytical distortion.

A useful conceptual distinction is therefore between biological multiplicity and analytical multiplicity. Biological multiplicity refers to the possibility that several historical processes contributed to the same set of gene-tree relationships. Analytical multiplicity refers to the possibility that several properties of data processing or model fitting alter how those histories are reconstructed. The two are not independent in practice: analytical error may obscure the expected pattern of a genuine biological process, whereas a biologically complex history may expose weaknesses in an otherwise adequate tree-based analysis. The diagnostic task is not to assign every discordant locus to a single cause, but to determine which causal classes remain compatible after progressively testing the assumptions on which each interpretation depends.

Expected signatures of incomplete lineage sorting

ILS is most informative diagnostically when it is formulated as a model expectation rather than used as a synonym for gene-tree disagreement. Under specified multispecies-coalescent conditions, discordant gene-tree or quartet frequencies have predictable relationships to the species-tree history, permitting formal assessment of whether an observed distribution is compatible with those expectations [9]. Such compatibility is evidence for an ILS-consistent explanation, but not proof that ILS is the sole process operating. Departure from the expected

distribution is likewise non-specific: gene-tree estimation error, introgression, model violation, or a combination of processes may move observed discordance away from the simple expectation.

This distinction is especially important in rapid radiations. Short intervals between divergences create limited time for ancestral alleles to coalesce before subsequent splitting, making extensive lineage sorting a plausible source of discordance. Yet recalcitrant phylogenomic histories need not decompose cleanly into a single mechanism. Work on an ancient rapid radiation outside insects demonstrated that gene-tree estimation error, ILS, and ancient gene flow can jointly account for persistent conflict [10]. The transferable point is methodological rather than taxon-specific: observing a signature expected under ILS does not require all residual discordance to share the same origin. The relevant question is whether the data remain compatible with ILS after uncertainty in gene-tree estimation and competing reticulate histories have also been examined.

Concordance factors add an important observation layer because they distinguish the amount of underlying agreement from conventional branch-support values. Gene concordance factors and site concordance factors quantify how much decisive gene or site information supports a focal relationship, thereby revealing conflict that can remain obscured behind strong summary support [11]. They do not, however, identify the process that produced low concordance. In an ILS-dominated radiation such as the analysed *Parnassius* system, widespread conflict can be interpreted in conjunction with coalescent expectations and independent tests [3]; low concordance by itself cannot be relabelled as ILS. This separation of discordance measurement from discordance explanation is fundamental to the proposed framework.

Expected ILS signatures should therefore be evaluated relationally. A short internode may increase biological plausibility; an appropriate distribution of alternative genealogies may increase model compatibility; and consistency across perturbations of data and analysis may reduce some competing analytical explanations. None of these observations independently establishes a unique cause. ILS becomes a stronger interpretation when multiple expectations are jointly satisfied and competing explanations fail to account for the same pattern, while mixed or off-model evidence should leave alternative causes open.

Expected signatures of introgression

Introgression changes genealogical histories in ways that may affect topology frequencies, branch-length distributions, genomic site patterns, and inferred historical parameters. Models that explicitly incorporate cross-species gene flow show that ignoring it can bias estimation of divergence times under tested conditions [12]. The direction and magnitude of such bias are scenario dependent, so altered historical parameters cannot be assumed to follow a universal pattern. More broadly, an introgression signal must be separated from its consequences: evidence that gene flow occurred does not automatically recover its timing, direction, number of events, donor identity, or full effect on the species tree.

Non-detection is equally difficult to interpret. Comparisons of Bayesian and heuristic tests have shown that methods can differ materially in their ability to identify cross-species introgression under particular data and model conditions [13]. The specific comparator is not an insect system, so its numerical performance should not be transferred to insect datasets. The general diagnostic implication is nevertheless important: failure of a lower-power procedure to detect gene flow cannot by itself establish historical isolation. Negative evidence becomes interpretable only in relation to the power, assumptions, lineage configuration, and sampling design of the test that generated it.

Emerging computational approaches expand rather than eliminate this dependency. Deep-learning methods such as ERICA have demonstrated the capacity to classify historical introgression patterns in simulated and empirical genomic applications, including *Heliconius* examples [14]. Such methods can contribute complementary evidence where their training and data assumptions are appropriate, but their output is not model-free historical proof. A machine-learning classification should therefore enter a discordance diagnosis as one evidence channel whose result requires convergence with other genealogical, branch-length, site-pattern, or explicit-model evidence.

The genomic distribution of introgression also matters. Comparative genomic work in *Heliconius* demonstrates that introgressed history can be heterogeneous across the genome and can contribute to obscured relationships within a radiation [15]. Consequently, a local excess of one genealogy need not represent the history of every genomic region, while a genome-wide summary can conceal localized reticulation. This mosaic property distinguishes the question “is there evidence of historical exchange?” from the stronger question “does one reticulate history adequately explain the species-level conflict?” It also blocks an unwarranted inference from

genomic heterogeneity to adaptive introgression unless independent evidence establishes the selective interpretation.

Taken together, expected introgression signatures are asymmetric in evidential strength. Positive concordant results across appropriately chosen methods may make a reticulate history more plausible, whereas a single significant statistic generally provides narrower information. Similarly, absence of a signal is informative only when the test had adequate opportunity to detect the relevant history. Introgression should therefore be inferred through convergent, assumption-aware evidence rather than by elevating any topology asymmetry, site statistic, or network edge into a complete reconstruction of evolutionary history.

Analytical and sampling error

Before discordance is attributed to biological history, the analytical path by which sequences become alignments, loci, gene trees, concordance summaries, and species trees must itself be examined. Methodological incongruence can enter through data assignment, orthology decisions, alignment and filtering, model choice, tree estimation, taxon representation, and downstream synthesis; no single correction removes all such artefacts [16]. The appropriate objective is therefore to test and minimize plausible error sources, not to assume that an "error-cleaned" dataset exists. In the proposed diagnostic logic, analytical adequacy is a prerequisite for strong biological attribution because every subsequent ILS or introgression test operates on inferred or processed data.

Orthology provides a useful example of why simple failure rules are inadequate. Gene duplication and loss can produce paralogous sequences that complicate species-tree reconstruction, yet simulations and empirical analyses indicate that some ILS-aware methods can remain robust to paralogs under tested conditions [17]. This does not make paralogy harmless. Instead, it converts orthology status into a sensitivity question: does the focal relationship persist when gene-family uncertainty, copy selection, or loci with ambiguous orthology are treated differently? A topology that changes sharply under plausible orthology alternatives carries a different inferential status from one that remains stable.

Rate heterogeneity can similarly distort the relationship between estimated gene histories and the species tree. Across analysed phylogenomic datasets, greater lineage-specific root-to-tip rate variation was associated with larger gene-tree/species-tree distances, whereas overall locus rate did not show the same association [18]. The evidence supports rate variation as an error-relevant diagnostic dimension, not as a unique explanation for discordance. Concordance measurement has its own sensitivities: site concordance factors can be affected by homoplasy and taxon sampling, and revised estimation procedures can reduce these effects without abolishing all bias [19]. Thus, apparent conflict may reflect both real historical heterogeneity and how that heterogeneity is measured.

Sampling incompleteness operates at several levels. High proportions of missing taxa across gene trees can degrade some summary approaches, while algorithms designed for severe missing-data conditions can improve recovery in benchmark settings [20]. Such benchmarks do not define a universal percentage at which an insect dataset becomes unreliable. Missingness is better treated structurally: investigators should ask whether absent loci or taxa are distributed in ways that selectively remove information from the relationships being diagnosed. More generally, genomic tree building contains multiple linked sources of stochastic and systematic error, so increasing sequence volume does not by itself guarantee an accurate history [21]. Species-tree approaches such as ASTRAL-III explicitly address gene-tree discordance expected under the multispecies coalescent and can use partially resolved gene trees [22], but an ILS-aware tree estimator does not by itself distinguish ILS from introgression.

The same caution applies one level earlier, because a gene tree is an estimate rather than a directly observed historical object. Maximum-likelihood programs such as IQ-TREE 2 provide extensive model and search capabilities for genomic inference [23], yet model choice and search sophistication cannot guarantee that the resulting topology equals the underlying genealogy. If alternative plausible models, partitions, alignments, or search settings change the relationship under investigation, that instability is evidence about inferential sensitivity and must precede a strong biological narrative.

Apparent certainty can also be highly concentrated. Comparative perturbation analyses have shown that a small number of genes or sites can disproportionately determine contentious phylogenomic relationships [24]. This phenomenon is not universal to every strongly supported node; rather, it motivates an influence audit. Relationships supported broadly across independent loci deserve a different interpretation from relationships whose support disappears when a very small subset of characters is removed. High bootstrap support and distributed phylogenetic evidence are therefore not interchangeable properties.

These problems are directly relevant to insect systematics. Entomological phylogenomics requires explicit attention to taxon and locus sampling, orthology, alignment, model selection, missing information, and inference strategy, with no single workflow guaranteed to be optimal across insect radiations [25]. The article therefore treats analytical and sampling error not as a residual explanation invoked after biological hypotheses fail, but as an active diagnostic class tested alongside them. This does not mean every discordant insect phylogeny is artefactual. It means that biological attribution is strongest when the relevant discordance survives plausible perturbations of the analytical and sampling decisions capable of generating or amplifying it.

The different error pathways cannot be compressed satisfactorily into a single "data quality" label because they act at different inferential stages and imply different corrective tests. The stage-specific distinctions below therefore separate how an apparent phylogenetic conflict is generated, observed, and stress-tested before it is carried forward as evidence about evolutionary history.

Table 1 establishes why the error branch of a discordance diagnosis must remain internally differentiated: orthology, rate structure, concordance estimation, missingness, gene-tree inference, concentrated character leverage, and sampling architecture do not produce equivalent evidential consequences. The same separation is needed visually because these mechanisms can either manufacture apparent conflict, magnify existing biological discordance, or obscure a real ILS or introgression signal.

Table 1. Stage-specific analytical failure states that can imitate, amplify, or conceal biological phylogenetic discordance

Limits and uncertainty	Evidence · Interpretive consequence	Research or test · Evidential standing	Biological setting
Orthology–paralogy instability			
Tested ILS-aware methods can tolerate paralogy under some conditions; robustness is not universal	Focal topology changes under alternative copy assignment, orthology filtering, or paralog inclusion Some apparent gene-tree conflict may depend on gene-family reconstruction rather than lineage history alone	Repeat species-tree inference under defensible orthology treatments and report relationship sensitivity Evidence-derived diagnostic state	Gene-family datasets used for insect species-tree reconstruction
Lineage-rate distortion			
Association does not establish rate variation as the unique cause of a conflicted node	Gene trees with greater root-to-tip rate variation show greater disagreement with species-tree estimates Rate heterogeneity can contribute to gene-tree estimation difficulty and apparent discordance	Compare focal relationships across rate profiles and models capable of representing relevant heterogeneity Evidence-derived diagnostic state	Loci with heterogeneous branch-specific evolutionary rates
Concordance-measurement distortion			
Improved estimation reduces rather than eliminates bias	Estimated concordance changes with treatment of homoplasy or taxon configuration Measured conflict may partly reflect the estimator and sampling configuration	Contrast support with gene/site concordance and test concordance sensitivity to taxon configuration Evidence-derived diagnostic state	Branches evaluated with site-level concordance statistics
Structured missingness			
Benchmark performance is dataset- and algorithm-dependent; no universal missing-data cutoff follows	Species-tree recovery changes with data completeness or alternative summary algorithms Missing information can destabilize species-tree reconstruction without implying a new biological process	Characterize where missing information occurs and perturb the analysis rather than applying a fixed percentage threshold Proposed diagnostic formulation grounded in method evidence	Gene-tree collections with uneven taxon or locus representation
Gene-tree model/search sensitivity			
Better-fitting or more sophisticated software does not make an inferred genealogy error-free	Alternative plausible models, partitions, alignments, or searches alter focal gene-tree relationships Downstream discordance can inherit uncertainty created during gene-tree estimation	Carry gene-tree uncertainty into species-tree and reticulation analyses instead of treating one topology as raw observation Proposed diagnostic formulation grounded in method evidence	Sequence-to-genealogy inference preceding any coalescent or introgression analysis
Character-leverage concentration			

Demonstrated for particular contentious relationships, not every supported node	Removal or down-weighting of a small subset of loci/sites changes topology or support substantially High aggregate support may represent concentrated rather than broadly distributed evidence	Perform locus/site influence or perturbation analyses before interpreting support as historical certainty Evidence-derived diagnostic state	Contentious internodes in large phylogenomic matrices
Workflow-coupled sampling error			
Optimal workflow is system-specific	Focal conclusions vary across defensible sampling or workflow choices Analytical stages can interact, so no isolated quality metric captures total inferential vulnerability	Audit taxon representation, locus composition, orthology, missingness and model sensitivity as linked stages Proposed integrative state	Insect phylogenomic datasets spanning specimen selection, loci, orthology and tree inference

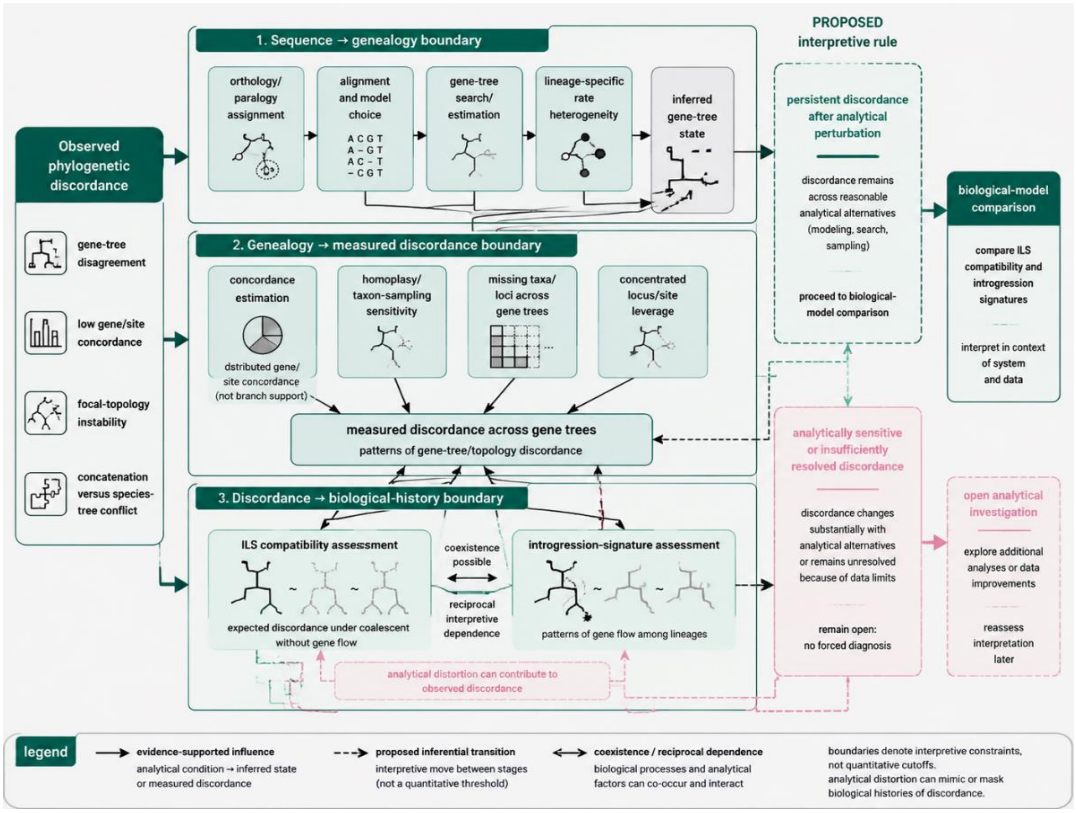


Figure 1. Analytical boundaries that determine when discordance can be carried forward as evidence of lineage history

Diagnostic comparisons

The central diagnostic difficulty is that no single observable property of a discordant phylogeny uniquely separates ILS, introgression, and analytical error. Quartet distributions can be evaluated against multispecies-coalescent expectations, but departure from those expectations does not identify the alternative process by itself [9]. Concordance factors quantify how widely a relationship is supported among genes or sites, but they remain descriptive rather than causal [11]. Introgression-oriented statistics and explicit models interrogate additional genealogical or sequence-pattern information, yet their interpretation depends on lineage configuration, assumptions, and statistical power [2]. Analytical perturbations answer a different question again: whether the apparent conflict survives plausible changes in data treatment or inference [16]. Consequently, diagnostic strength should come from agreement among evidence classes that fail for different reasons, rather than from maximizing one support statistic.

Sampling architecture belongs within this diagnostic comparison rather than being treated as a preliminary logistical matter. Analyses of Polyommatina showed that combining broad barcode sampling with a strategically

sampled multilocus backbone can recover higher-level phylogenetic structure more reliably than barcode-only inference under the studied design [26]. The precise sampling proportions are not general rules for insect phylogenomics. The broader implication is that representation determines which relationships contain enough independent information to discriminate among alternative histories. Sparse lineage sampling can hide intermediates, uneven locus coverage can change gene-tree composition, and strategically distributed multilocus information can alter the stability of systematic conclusions. Thus, when competing histories depend on different taxon or locus subsets, sampling sensitivity is itself evidence about the reliability of the inferred discordance.

The proposed diagnostic comparison therefore treats signatures as conditional combinations. ILS gains plausibility when discordance is compatible with coalescent expectations and remains stable under analytical perturbation. Introgression gains plausibility when reticulation-sensitive evidence converges across suitable methods and cannot be explained adequately by ILS or identified artefacts. Analytical error gains plausibility when focal conflict is strongly dependent on orthology, model, sampling, rate, missing-data, concordance-estimation, or character-leverage decisions. None of these categories is required to exclude the others. A mixed state is scientifically preferable when the available evidence indicates that biological and analytical processes jointly shape the observed pattern.

Hierarchical discordance-triage workflow

A useful hierarchy should begin with the inferential status of the observation rather than with a preferred biological explanation. The first proposed level is therefore discordance characterization: determine what actually disagrees—sites, genes, quartets, genomic regions, concatenated versus coalescent trees, or alternative analytical treatments. This level separates conventional support from distributed concordance because high branch support can coexist with extensive underlying conflict [11]. It also records whether discordance is localized to one contentious internode or distributed across a radiation.

The second proposed level is analytical adequacy and perturbation sensitivity. Orthology, rate variation, concordance estimation, missingness, taxon representation, model specification, and concentrated locus/site influence should be interrogated before a biological cause is privileged [16]. Rate-associated discordance warrants sensitivity analysis rather than automatic biological attribution [18]. Concordance estimates should be interpreted with awareness of homoplasy and taxon-sampling effects [19]. Missing-data patterns should be examined structurally rather than against a universal cutoff [20]. Gene trees should remain recognized as inferred objects whose model and search uncertainty propagates downstream [23]. When a relationship collapses repeatedly under reasonable perturbations, the appropriate result is analytically sensitive discordance, not a forced ILS or introgression diagnosis.

Only discordance that is sufficiently robust to those tests should move strongly into the third proposed level, coalescent compatibility. ILS-aware inference and quartet expectations provide a model-based basis for asking whether alternative gene histories have the distribution expected from ancestral polymorphism and successive lineage divergence [9]. Methods such as ASTRAL-III accommodate discordant gene trees under multispecies-coalescent assumptions [22], but neither an ILS-aware species tree nor an MSC-compatible quartet pattern demonstrates that introgression is absent. Evidence from rapid insect radiations shows why the appropriate language is “consistent with ILS” rather than “proved to be ILS” [3].

The fourth proposed level is reticulation challenge. Where discordance persists and ILS alone does not adequately account for the available evidence, analyses should examine topology asymmetry, branch-length information, sequence patterns, explicit gene-flow models, or complementary computational approaches appropriate to the lineage configuration [2]. Test power remains part of interpretation because failure to detect introgression under one method need not demonstrate its absence [13]. Furthermore, insect systems such as *Anastrepha* show that ILS and gene flow can coexist [7], while the *Bactrocera dorsalis* complex illustrates that alternative tree and network representations can remain difficult to resolve completely even when gene flow is implicated [8]. Reticulation testing should therefore permit three outcomes: evidence favoring introgression, evidence insufficient to distinguish alternatives, or evidence supporting a mixed history.

The final proposed level is species-history qualification, not immediate taxonomic revision. Evidence from the preceding layers is carried forward with its unresolved uncertainties. A strongly supported species-tree relationship that remains sensitive to gene flow, for example, should not be presented with the same historical confidence as a relationship stable across tree, coalescent, reticulation, and perturbation analyses. Conversely,

unresolved discordance is an acceptable endpoint. The hierarchy is designed to stop when the evidence cannot discriminate among alternatives rather than manufacture resolution.
Because these stages generate different testable expectations, a prediction-and-test ledger is more informative than assigning unsupported scores or confidence percentages.

Table 2. Prediction-and-test ledger for advancing, retaining, or stopping discordance diagnoses across the proposed triage hierarchy

Focal triage state	Limits and uncertainty	Evidence basis · Interpretation or consequence · Evidence · Interpretive consequence	Biological setting	Research or test · Status · Evidential standing
Discordance characterized but cause unassigned	Support and concordance are not equivalent; different conflict metrics measure different properties	Conflict is reproducible, but its distribution and analytical origin remain incompletely tested Establishes a phenomenon requiring explanation, not a mechanism	Any insect node with conflicting loci, sites, quartets, or analytical trees	Partition discordance by gene, site, quartet, genomic region, and analytical treatment Proposed triage state
Analytically sensitive discordance	Sensitivity does not prove that the underlying history lacks genuine biological discordance	Topology or concordance responds to orthology, rate, model, missingness, taxon, or character-leverage treatment Strong biological attribution should be deferred	Focal relationship changes under defensible data or model perturbations	Identify which perturbation controls the result and improve sampling/model adequacy Proposed stopping or return state
Persistent, MSC-compatible discordance	Compatibility does not exclude introgression or residual gene-tree error	Discordance survives analytical audit and fits relevant coalescent expectations ILS becomes a supported explanation	Rapid divergence or short-internode history where coalescent variation is plausible	Test whether reticulation-sensitive evidence provides additional explanatory information Proposed advancement state
Persistent discordance with reticulation signal	Timing, direction, donor identity, and number of events may remain unresolved	Independent introgression-sensitive evidence departs from a tree-only/ILS account Introgression becomes a plausible component of the history	Lineages with plausible historical contact or genomic exchange	Seek convergence among genealogy, branch-length, site-pattern, and explicit-model evidence Proposed advancement state
Mixed biological discordance	Relative contributions may be impossible to partition precisely	Some observations are coalescent-compatible while other independent evidence supports exchange ILS and introgression remain simultaneously plausible rather than mutually exclusive	Radiations in which lineage sorting and gene flow plausibly overlap	Report process coexistence and localize evidence by branch or genomic region where possible Proposed mixed state
Indeterminate discordance	More complex models do not guarantee identifiable history	Available evidence cannot discriminate among analytical error, ILS, introgression, or mixtures Unresolved causal status is retained explicitly	Sparse, low-power, model-sensitive, or topologically difficult datasets	Improve sampling or orthogonal evidence before stronger historical or classificatory claims Proposed legitimate endpoint

The workflow also requires a cross-scale view because observations at sites and loci become progressively stronger claims about genealogies, species histories, and ultimately classification. **Figure 2** therefore represents not a one-way decision tree but a layered architecture in which failed adequacy checks return the analysis to earlier stages and unresolved cases remain visible.

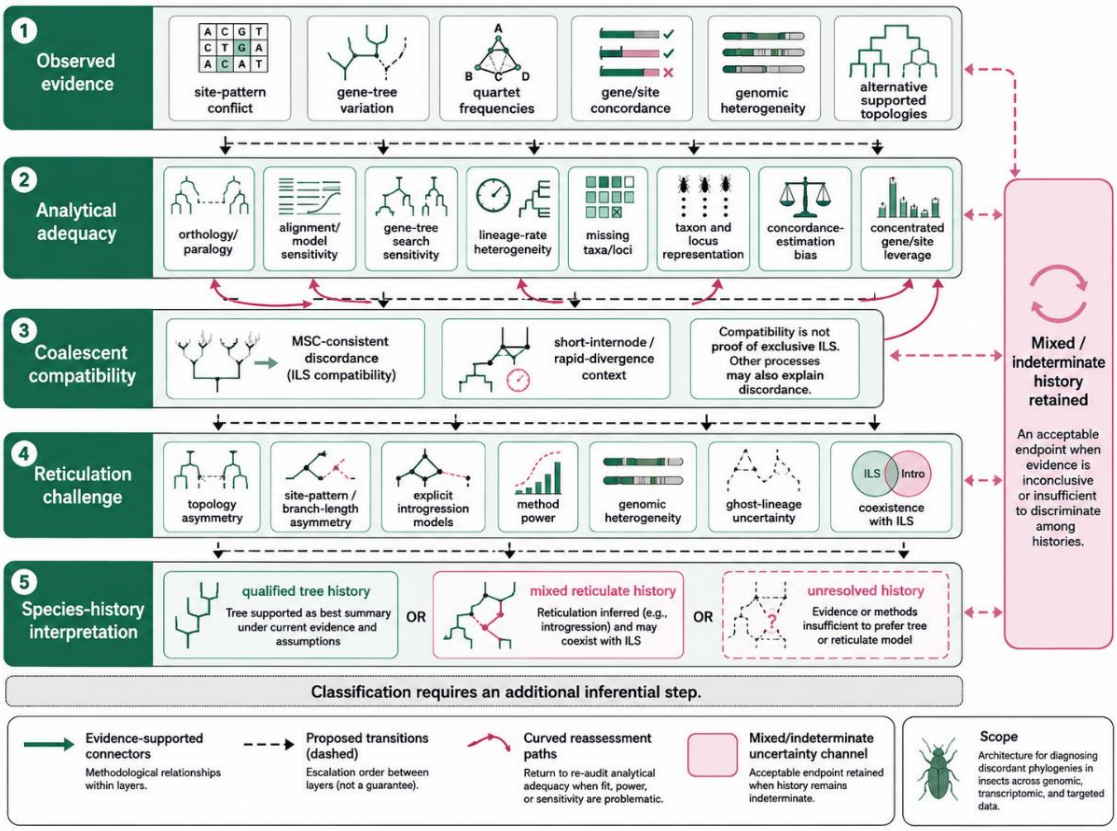


Figure 2. How discordance evidence changes inferential status across analytical, coalescent, reticulate, and species-history layers

Consequences for species trees and classification

Discordance triage becomes taxonomically important because species-tree inference and species delimitation answer different questions. Multispecies-coalescent delimitation may identify genetic structure that does not necessarily correspond to independently evolving species, particularly where population structure is interpreted as evidence of species boundaries [27]. This limitation does not make coalescent methods uninformative; rather, it means that recovered lineage structure must not be translated automatically into taxonomic rank.

Gene flow creates a complementary boundary. Species-delimitation models can explicitly consider migration, demonstrating that evidence of gene flow need not logically imply absence of species boundaries [28]. Conversely, a resolved bifurcating species tree does not erase evidence of reticulation. Classification therefore requires adjudicating whether the uncertainty relevant to a contentious node materially changes hypotheses of lineage independence, diagnosability, or broader systematic interpretation.

In insects, these distinctions are practically consequential. *Anastrepha* provides an example in which ILS and gene flow occur alongside morphospecies that largely retain separate evolutionary-lineage signals [7]. In the *Bactrocera dorsalis* complex, alternative genomic analyses demonstrate that strong tree support can coexist with gene-flow-associated ambiguity [8]. Sampling design can also affect higher-level systematic reliability [26]. None of these observations supplies a general rule for splitting or synonymizing taxa. They instead support a conservative principle: taxonomic change should be proportional to the convergence of evidence, while unresolved phylogenetic mechanism should remain visible in the classification argument.

The proposed framework therefore propagates uncertainty rather than deleting it at the species-tree stage. A node stable across analytical perturbations, coalescent evaluation, reticulation tests, and representative sampling carries a different evidential status from a node supported only under one analytical architecture. Where discordance remains mixed or indeterminate, nomenclatural stability may be preferable to treating an unresolved model comparison as a taxonomic fact.

Future methodological priorities

A first priority is to integrate evidence modalities without pretending that they are interchangeable. Topology frequencies, branch lengths, site patterns, gene/site concordance, explicit demographic models, network analyses, and newer computational classifiers answer different questions and have different failure modes [2]. Quartet-based tests provide model-based expectations for coalescent histories [9]. Concordance factors expose distributed versus concentrated disagreement [11]. Machine-learning approaches can add another evidence stream where training assumptions are appropriate [14]. Future comparative evaluations should ask when combinations of these modalities discriminate histories better than isolated statistics, while allowing mixed and unresolved outcomes.

A second priority is reproducible perturbation reporting. Methodological incongruence can arise throughout phylogenomic workflows [16], and insect-focused phylogenomics similarly emphasizes the importance of sampling, orthology, model choice, and inference strategy [25]. Studies of discordance should therefore report how focal conclusions respond to alternative orthology treatments, taxon/locus subsets, missing-data structures, rate conditions, concordance estimators, and plausible gene-tree models. The objective should not be a universal checklist threshold, but transparent characterization of which analytical choices control the historical conclusion.

A third priority is explicit propagation of uncertainty from gene trees into species trees and classification. Sampling strategies can materially affect systematic reconstruction [26], while coalescent structure and species status remain conceptually distinct [27]. Gene flow adds another dimension that cannot be reduced to either automatic species collapse or automatic network acceptance [28]. Empirical evaluation of the proposed hierarchy should therefore test whether staged, multimodal diagnosis improves interpretive reliability across insect groups with different divergence depths, genomic resources, sampling densities, and known histories. Such work would constitute validation of the framework; the synthesis presented here itself does not.

Conclusion

Discordant insect phylogenies should not be interpreted by assigning a biological cause directly from conflict, support, or a single diagnostic statistic. ILS, introgression, and analytical or sampling error generate overlapping signatures, may coexist, and operate across different inferential levels. The central contribution of this article is a proposed hierarchy that separates observed discordance from analytical adequacy, coalescent compatibility, reticulation evidence, species-history interpretation, and eventual classification.

Its principal safeguard is permission not to resolve every conflict. Persistent evidence may favor ILS, introgression, or a mixed history; analytical sensitivity may require renewed inference; and some cases should remain indeterminate until sampling or methodology improves. Species trees and taxonomic decisions should inherit these qualifications rather than conceal them. Treating discordance as a differential-diagnostic problem therefore offers a more defensible basis for interpreting complex insect phylogenomic histories without converting methodological confidence into unwarranted evolutionary certainty.

Acknowledgments: None

Conflict of Interest: None

Financial Support: None

Ethics Statement: None

References

1. Steenwyk JL, Li Y, Zhou X, Shen XX, Rokas A. Incongruence in the phylogenomics era. Nat Rev Genet. 2023;24(12):834-50. doi:10.1038/s41576-023-00620-x
2. Hibbins MS, Hahn MW. Phylogenomic approaches to detecting and characterizing introgression. Genetics. 2022;220(2):iyab173. doi:10.1093/genetics/iyab173
3. He B, Zhao Y, Su C, Lin G, Wang Y, Li L, et al. Phylogenomics reveal extensive phylogenetic discordance due to incomplete lineage sorting following the rapid radiation of alpine butterflies (Papilionidae: Parnassius). Syst Entomol. 2023;48(4):585-99. doi:10.1111/syen.12592

4. Owen CL, Miller GL. Phylogenomics of the Aphididae: Deep relationships between subfamilies clouded by gene tree discordance, introgression and the gene tree anomaly zone. Syst Entomol. 2022;47(3):470-86. doi:10.1111/syen.12542
5. Suvorov A, Scornavacca C, Fujimoto MS, Bodily P, Clement M, Crandall KA, et al. Deep ancestral introgression shapes evolutionary history of dragonflies and damselflies. Syst Biol. 2022;71(3):526-46. doi:10.1093/sysbio/syab063
6. Thawornwattana Y, Seixas FA, Yang Z, Mallet J. Full-likelihood genomic analysis clarifies a complex history of species divergence and introgression: The example of the erato-sara group of Heliconius butterflies. Syst Biol. 2022;71(5):1159-77. doi:10.1093/sysbio/syac009
7. Congrains C, Zucchi RA, de Brito RA. Phylogenomic approach reveals strong signatures of introgression in the rapid diversification of Neotropical true fruit flies (Anastrepha: Tephritidae). Mol Phylogenet Evol. 2021;162:107200. doi:10.1016/j.ympev.2021.107200
8. San Jose M, Doorenweerd C, Geib S, Barr N, Dupuis JR, Leblanc L, et al. Interspecific gene flow obscures phylogenetic relationships in an important insect pest species complex. Mol Phylogenet Evol. 2023;188:107892. doi:10.1016/j.ympev.2023.107892
9. Allman ES, Mitchell JD, Rhodes JA. Gene tree discord, simplex plots, and statistical tests under the coalescent. Syst Biol. 2022;71(4):929-42. doi:10.1093/sysbio/syab008
10. Cai L, Xi Z, Lemmon EM, Lemmon AR, Mast A, Buddenhagen CE, et al. The perfect storm: Gene tree estimation error, incomplete lineage sorting, and ancient gene flow explain the most recalcitrant ancient angiosperm clade, Malpighiales. Syst Biol. 2021;70(3):491-507. doi:10.1093/sysbio/syaa083
11. Minh BQ, Hahn MW, Lanfear R. New methods to calculate concordance factors for phylogenomic datasets. Mol Biol Evol. 2020;37(9):2727-33. doi:10.1093/molbev/msaa106
12. Tiley GP, Flouri T, Jiao X, Poelstra JW, Xu B, Zhu T, et al. Estimation of species divergence times in presence of cross-species gene flow. Syst Biol. 2023;72(4):820-36. doi:10.1093/sysbio/syad015
13. Ji J, Jackson DJ, Leaché AD, Yang Z. Power of Bayesian and heuristic tests to detect cross-species introgression with reference to gene flow in the Tamias quadrivittatus group of North American chipmunks. Syst Biol. 2023;72(2):446-65. doi:10.1093/sysbio/syac077
14. Zhang Y, Zhu Q, Shao Y, Jiang Y, Ouyang Y, Zhang L, et al. Inferring historical introgression with deep learning. Syst Biol. 2023;72(5):1013-38. doi:10.1093/sysbio/syad033
15. Edelman NB, Frandsen PB, Miyagi M, Clavijo B, Davey J, Dikow RB, et al. Genomic architecture and introgression shape a butterfly radiation. Science. 2019;366(6465):594-9. doi:10.1126/science.aaw2090
16. Fleming JF, Valero-Gracia A, Struck TH. Identifying and addressing methodological incongruence in phylogenomics: A review. Evol Appl. 2023;16(6):1087-104. doi:10.1111/eva.13565
17. Yan Z, Smith ML, Du P, Hahn MW, Nakhleh L. Species tree inference methods intended to deal with incomplete lineage sorting are robust to the presence of paralogs. Syst Biol. 2022;71(2):367-81. doi:10.1093/sysbio/syab056
18. Vankan M, Ho SYW, Duchêne DA. Evolutionary rate variation among lineages in gene trees has a negative impact on species-tree inference. Syst Biol. 2022;71(2):490-500. doi:10.1093/sysbio/syab051
19. Mo YK, Lanfear R, Hahn MW, Minh BQ. Updated site concordance factors minimize effects of homoplasy and taxon sampling. Bioinformatics. 2023;39(1):btac741. doi:10.1093/bioinformatics/btac741
20. Morel B, Williams TA, Stamatakis A. Asteroid: A new algorithm to infer species trees from gene trees under high proportions of missing data. Bioinformatics. 2023;39(1):btac832. doi:10.1093/bioinformatics/btac832
21. Kapli P, Yang Z, Telford MJ. Phylogenetic tree building in the genomic age. Nat Rev Genet. 2020;21(7):428-44. doi:10.1038/s41576-020-0233-0
22. Zhang C, Rabiee M, Sayyari E, Mirarab S. ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics. 2018;19(Suppl 6):153. doi:10.1186/s12859-018-2129-y
23. 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
24. Shen XX, Hittinger CT, Rokas A. Contentious relationships in phylogenomic studies can be driven by a handful of genes. Nat Ecol Evol. 2017;1(5):0126. doi:10.1038/s41559-017-0126
25. 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

26. Talavera G, Lukhtanov V, Pierce NE, Vila R. DNA barcodes combined with multilocus data of representative taxa can generate reliable higher-level phylogenies. Syst Biol. 2022;71(2):382-95. doi:10.1093/sysbio/syab038
27. 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
28. 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