Specialty Journal of Pharmacognosy, Phytochemistry, and Biotechnology
ISSN: 3062-441X
2026, Volume 6, Issue 1, Page No: 110-120

Available online at: www.galaxypub.co/page/journals

Galaxy Publication

Shared Metabolites Do Not Prove Shared Biosynthesis: A Host–Endophyte Attribution Model Using Isotope Logic, Genomic Capacity, Compartment Evidence, and Cultivation Stability

Lina Hassan[1]*, Omar Khalaf[2], Reem Jaber[1], Rania Mostafa[3]

[1] Department of Host–Endophyte Biosynthetic Attribution, Faculty of Pharmacy, University of Jordan, Amman, Jordan.
[2] Department of Isotope Logic and Genomic Capacity, Faculty of Pharmacy, Jordan University of Science and Technology, Irbid, Jordan.
[3] Department of Compartment Evidence and Cultivation Stability, Faculty of Pharmacy, Al-Balqa Applied University, Salt, Jordan.

***E-mail** ✉ lina.hassan@ju.edu.jo

Received: 19 January 2026; **Revised:** 29 April 2026; **Accepted:** 02 May 2026

ABSTRACT

The recurrent detection of the same specialized metabolite in a plant and one of its endophytes is frequently interpreted as evidence that both partners possess the capacity to synthesize that compound. This inference is stronger than the observation itself. Shared chemistry can arise through autonomous biosynthesis, host metabolic induction, microbial pathway activation, metabolite or precursor exchange, biotransformation, cultivation-dependent expression, or analytical misassignment. This article develops a proposed host–endophyte biosynthetic-attribution model that separates chemical identity from biosynthetic origin and integrates four principal evidence domains: isotope logic, genomic capacity, compartment evidence, and cultivation stability. The model treats these domains as complementary rather than interchangeable. Stable-isotope incorporation can constrain precursor routing but may remain ambiguous under cross-feeding; compatible biosynthetic genes indicate capacity but not expression; spatial localization constrains where metabolites occur but not necessarily where they are synthesized; and persistent axenic production strengthens microbial attribution without independently establishing pathway sufficiency. Higher confidence therefore arises from convergence among orthogonal observations and, where feasible, functional perturbation or pathway reconstruction. The proposed framework also accommodates competing explanations and evidence that decreases confidence, including metabolite disappearance during serial cultivation, absence of required enzymatic capacity, inconsistent spatial distributions, or failure of predictions under controlled testing. Its purpose is not to assign universal numerical scores or claim validated predictive performance, but to make biosynthetic claims more explicit, falsifiable, and experimentally prioritized. The framework is intended to improve interpretation of host–endophyte chemistry and reduce premature producer attribution in natural-product discovery.

Keywords: Endophytes, Biosynthetic attribution, Stable-isotope tracing, Biosynthetic gene clusters, Spatial metabolomics, Natural products

How to Cite This Article: Hassan L, Khalaf O, Jaber R, Mostafa R. Shared Metabolites Do Not Prove Shared Biosynthesis: A Host–Endophyte Attribution Model Using Isotope Logic, Genomic Capacity, Compartment Evidence, and Cultivation Stability. Spec J Pharmacogn Phytochem Biotechnol. 2026;6(1):110-20. https://doi.org/10.51847/LNaA54bOOI

Introduction

Endophytic microorganisms are important sources of chemically diverse natural products, but some of the most consequential claims in this field concern compounds already known from their host plants. Detection of an apparently identical metabolite in a plant and an isolated endophyte creates an attractive biological narrative: the microorganism may possess a parallel or transferred biosynthetic pathway and could therefore represent an alternative production platform. Yet the history of endophytic taxol research illustrates how chemical detection,

strain-associated production, and reproducible pathway attribution can become conflated. Critical re-evaluation of this literature has emphasized that repeated claims of fungal taxol production do not, by themselves, establish a stable and independently demonstrated fungal biosynthetic system [1].

The interpretive problem is broader than any one compound. Plant–endophyte associations are metabolically interactive systems in which microbial colonization can alter host specialized metabolism, while plant-derived chemical environments can influence microbial physiology and secondary metabolism [2]. Consequently, finding a molecule in both partners is compatible with several mechanistic histories. The endophyte may synthesize it independently; the plant may synthesize more of it after colonization; one partner may transform a precursor supplied by the other; or the measured signal may reflect transport, persistence, or analytical ambiguity rather than parallel biosynthesis.

Experimental plant–endophyte systems reinforce this distinction between metabolic influence and producer identity. In poplar, fungal endophyte colonization can alter leaf chemistry and ecological phenotype, demonstrating that the presence of a microorganism may change host metabolite abundance without requiring the microorganism itself to manufacture each affected molecule [3]. Spatial metabolomic work likewise shows that endophyte colonization can reorganize host metabolic patterns across tissues and cellular regions, making chemical proximity biologically informative while leaving biosynthetic origin unresolved unless localization is coupled to stronger mechanistic evidence [4].

The central problem is therefore evidential rather than terminological. “Present in the endophyte,” “associated with colonization,” “encoded by the endophyte,” and “biosynthesized autonomously by the endophyte” are different claims. This article proposes a host–endophyte attribution model that preserves these distinctions and organizes evidence according to what it can and cannot establish. The framework gives particular weight to isotope incorporation, compatible genomic capacity, organism- or compartment-resolved localization, and persistence during controlled cultivation, while allowing functional reconstruction and contradictory evidence to revise an initial attribution. Its purpose is to replace binary producer labels with a more disciplined account of how biosynthetic origin is supported, challenged, and ultimately tested.

Why shared chemistry does not reveal biosynthetic origin

Chemical similarity becomes especially misleading when biological interaction itself changes production. Plant metabolites can act as signals or environmental modifiers of microbial secondary metabolism; in a root-associated *Streptomyces* endophyte, plant coumarins altered natural-product biosynthesis, illustrating that host chemistry can regulate microbial output without implying that host and microbe use the same biosynthetic pathway [5]. Conversely, persistent detection of a host-associated compound after isolation may strengthen the case for microbial production but remains conceptually separate from complete pathway demonstration. Reported taxol production by an endophytic *Aspergillus niger* under repeated cultivation is therefore evidence relevant to persistence, not an independent demonstration that every enzymatic step of a fully autonomous fungal pathway has been established [6].

Genomic information changes the quality of the attribution problem because it asks whether the proposed producer contains biochemically plausible machinery. Gene mining linked to cannabidiol production in *Fusarium solani*, for example, illustrates how chemical observations can be strengthened when candidate biosynthetic functions are investigated alongside production [7]. Even then, genomic compatibility does not make every pathway step certain, nor does it exclude metabolic contributions from another organism.

The reverse causal direction must also remain visible. Endophytes and endophyte-derived elicitors can stimulate biosynthetic and regulatory responses in medicinal plants, increasing host specialized-metabolite accumulation [8]. Thus, an increase in a “shared” compound after inoculation may actually provide evidence for host induction rather than microbial synthesis. The proposed classification in **Table 1** separates the principal interpretive states that can produce the same superficial observation—chemical sharing—while deliberately avoiding assignment of a producer before discriminating evidence is available.

Table 1 organizes the principal proposed interpretations of plant–endophyte metabolite sharing and the evidential distinctions that prevent chemical co-detection from being treated as biosynthetic proof.

Table 1. Proposed classification of mechanistically distinct explanations for an apparently shared host–endophyte metabolite

Interpretive state	What the observation may represent	Evidence that would strengthen the interpretation	Principal attribution boundary
Autonomous microbial production	Endophyte synthesizes the metabolite independently	Persistent axenic production plus compatible pathway evidence	Culture detection alone does not establish pathway sufficiency
Host induction	Colonization stimulates host biosynthesis	Host-resolved pathway activation with compartment-specific metabolite increase	Increased abundance does not identify the microorganism as producer
Interaction-dependent microbial production	Host compounds or conditions activate microbial metabolism	Reproducible induction in defined microbial cultures	Dependence on host signals differs from shared biosynthetic machinery
Independent dual production	Both partners possess routes to the same product	Organism-resolved precursor incorporation and compatible biosynthetic capacity in both	Structural identity does not imply pathway identity
Exchange or biotransformation	One partner supplies precursor, intermediate, or substrate used by the other	Tracer-resolved transfer and product formation in controlled compartments	Label transfer can mimic autonomous synthesis
Apparent sharing	Transport, carryover, contamination, or analytical misassignment	Orthogonal chemical identification and stringent separation controls	Detection without origin-resolving evidence remains non-attributive

Note: The categories are a proposed conceptual classification derived from the article's synthesis; they are not validated quantitative attribution classes.

Competing explanations for plant–endophyte metabolite sharing

A rigorous attribution model must accommodate not only multiple producers but multiple states of the same producer. Fungal secondary metabolism is extensively controlled by transcriptional, chromatin, developmental, and environmental regulation, so a biosynthetic gene cluster may be present while its associated chemistry is undetectable under routine laboratory conditions [9]. Consequently, loss of a metabolite after isolation cannot be interpreted automatically as evidence that the endophyte never possessed the relevant biosynthetic capacity. It may instead indicate that the ecological cue, regulatory state, nutrient environment, or interacting partner required for expression has been removed.

Current understanding of fungal specialized metabolism reinforces this conditionality: biosynthetic capacity and observed production occupy different evidential levels because pathway expression depends on regulatory and environmental context [10]. This distinction creates two opposing attribution errors. Persistent culture production may be overinterpreted as complete autonomous biosynthesis, whereas failure to detect a compound in standard culture may be overinterpreted as evidence against microbial capacity.

Epigenetic regulation creates a particularly important difficult case. Endophytic fungal pathways that are cryptic or weakly expressed under baseline conditions may become chemically visible after chromatin-modifying or related interventions [11]. Therefore, cultivation instability is meaningful evidence, but its interpretation depends on whether disappearance reflects true loss, regulatory silencing, strain change, or absence of an inducing interaction.

Functional manipulation provides a way to separate some of these possibilities. Systematic overexpression of secondary-metabolism transcription factors in *Aspergillus nidulans* has demonstrated that latent biosynthetic output can be exposed experimentally when regulatory constraints are altered [12]. Such interventions do not prove that every silent endophytic pathway behaves similarly, but they establish that "not produced under the tested condition" is weaker than "biosynthetic capacity absent." The attribution problem should therefore be framed as competing hypotheses subjected to progressively discriminating tests, rather than as a binary choice based on whether a metabolite appears in a chromatogram.

Evidence capable of discriminating between origins

Stable-isotope experiments provide one of the clearest routes from chemical occurrence toward biosynthetic inference because they can test whether defined precursors are incorporated into a product and how labelled

carbon or other atoms move through metabolism. Context-specific isotope tracing, metabolic-flux logic, and stable-isotope-assisted metabolomics collectively show that informative interpretation depends on tracer selection, temporal design, isotopologue structure, and analytical discrimination from naturally occurring isotope patterns [13–15]. In host–endophyte systems, the additional problem is cross-feeding: a label introduced into one organism may enter exchanged metabolites and subsequently appear in a product formed by the other.

Tracer evidence therefore requires analytical controls rather than simple detection of enrichment. High-resolution mass-spectrometric approaches can identify tracer incorporation across untargeted metabolomic data, but reliable attribution depends on distinguishing true labelled features from background and processing artefacts [16]. Chemical identity itself is another independent layer. Best-practice metabolomics guidance emphasizes that annotation and structural identification carry different levels of confidence [17]. A biosynthetic-origin claim cannot be stronger than the evidence that the supposedly shared metabolite is actually the same chemical entity in both partners.

Compartment evidence addresses a different question: where does the compound or its intermediates occur? Mass-spectrometry imaging can resolve plant metabolites spatially across tissues and cellular regions [18], and increasingly contributes to reconstruction of the spatial organization of natural-product biosynthesis [19]. Yet localization must not be collapsed into synthesis. Transport, secretion, uptake, ion suppression, spatial resolution, and matrix effects can all separate a measured location from the true site of production; recent plant MSI analyses therefore support spatial constraint of biosynthetic hypotheses rather than automatic producer assignment [20].

Genomics tests whether the proposed organism possesses plausible biosynthetic machinery. Contemporary BGC-detection tools, curated experimentally characterized cluster resources, and large fungal genome surveys make it possible to compare candidate pathways with known or predicted biosynthetic capacity [21–23]. This evidence is powerful when required enzymatic functions are compatible with the chemistry, but genome content is still not synonymous with expression, flux, or product formation.

Table 2 compares the principal evidence domains according to the attribution question each can answer, the inference it cannot support alone, and the controls needed before evidence is integrated in the proposed model.

Table 2. Comparative analytical matrix for evidence used to discriminate host and endophyte biosynthetic origin

Evidence domain	Primary attribution question	Strongest defensible inference	Critical control or complement	What the evidence cannot establish alone
Stable-isotope logic	Which biological route incorporates a defined precursor?	Direct metabolic participation in labelled precursor conversion	Organism-resolved exposure, time course, cross-feeding controls	Autonomous synthesis when label can move between partners
Chemical identity	Is the same molecular entity actually being compared?	Confidence that host and endophyte signals represent the same compound	Orthogonal spectra, standards where available, calibrated annotation	Which organism synthesized the compound
Compartment evidence	Where are product and intermediates localized?	Spatial compatibility with a proposed production or transfer route	Adequate spatial resolution and transport-aware interpretation	Site of synthesis from colocalization alone
Genomic capacity	Does the candidate producer encode compatible biosynthetic functions?	Plausibility of pathway capacity	Assembly quality, functional annotation, expression or biochemical testing	Active pathway or product formation
Cultivation stability	Does production persist outside the original association?	Reproducible production under defined axenic conditions	Serial passage, contamination controls, identity confirmation	Complete pathway autonomy
Evidence convergence	Do independent modalities support the same origin hypothesis?	Higher qualitative confidence in a working attribution	Explicit competing hypotheses and falsification tests	Validated probability, universal score, or causal certainty

Note: Evidence domains are complementary. The proposed framework does not assign numerical weights or treat any single domain as universally sufficient.

A host–endophyte attribution model

The proposed Host–Endophyte Attribution Model treats biosynthetic origin as a hypothesis that must survive comparison with plausible alternatives, not as a label attached when two samples contain the same compound. Metabologenomics demonstrates why this matters: correlations across fungal metabolomes and genomes can connect metabolites with candidate biosynthetic gene clusters, but the resulting pairs remain assignments that gain strength through additional evidence rather than direct demonstrations of causality [24]. Bioactivity-driven metabologenomics further shows how orthogonal information can narrow candidate metabolite–cluster relationships [25]. The model therefore begins only after chemical identity has been established sufficiently for the attribution question being asked.

Four evidence domains are then considered in parallel. Isotope logic asks whether precursor incorporation is compatible with synthesis by the proposed organism and whether cross-feeding has been excluded. Genomic capacity asks whether required enzymatic chemistry is plausible from the organism's genome. Compartment evidence asks whether product, intermediates, and producing biomass occupy spatial relationships compatible with the hypothesis. Cultivation stability asks whether production persists when host material and uncontrolled exchange are removed. Integrated genomic–metabolomic workflows support treating such data streams as complementary rather than substitutable [26]. Biosynthetic gene-cluster family analysis adds another constraint by asking whether candidate genomic machinery is related to pathways capable of comparable chemistry, without treating sequence relatedness as product identity [27].

A further distinction concerns evidence dependence. Cultivation and genomic observations can appear mutually supportive while still sharing the same hidden assumption that the cultured isolate is the relevant biological producer. The model therefore asks whether at least one line of evidence directly interrogates metabolic participation or functional pathway behavior. This is why a plausible BGC plus repeated product detection is stronger than either observation alone but remains vulnerable to contamination, uptake, or incompletely resolved pathway chemistry. Conversely, isotope evidence without genomic plausibility may indicate transformation of an imported precursor rather than autonomous construction of the complete scaffold. The evidential value lies in how modalities constrain one another, not in simply accumulating positive findings.

The model also distinguishes supportive absence from noninformative absence. Failure to find a compatible BGC is potentially contradictory only when genome completeness and pathway expectations are adequate. Failure to observe an intermediate spatially may be uninformative when sensitivity or transport is limiting. Similarly, failure of axenic production becomes more informative after several biologically appropriate culture conditions have been tested, yet remains different from evidence that the organism lacks the relevant capacity. These asymmetries are necessary because negative findings differ markedly in what they can exclude.

The proposed decision logic is asymmetric. Concordant evidence can raise confidence, but a serious contradiction can redirect the attribution even when several supportive observations remain. For example, persistent axenic production is weakened if the proposed producer lacks credible enzymatic capacity; genomic compatibility is weakened if controlled isotope tracing repeatedly assigns precursor incorporation to the host; and spatial colocalization is weakened if the compound persists where microbial biomass is absent. Conversely, negative culture results should be interpreted in light of possible pathway silencing rather than treated as automatic falsification.

Table 3 translates this reasoning into a proposed workflow that preserves the independence of the evidence domains and makes explicit when an attribution should remain unresolved rather than being forced toward either partner.

Table 3. Proposed decision workflow for host–endophyte biosynthetic attribution

Decision point	Question to resolve	Evidence that advances attribution	Evidence that redirects or limits it	Proposed decision state
Identity gate	Is the same compound sufficiently established in both sources?	Orthogonal structural agreement	Ambiguous annotation or isomerism	Resolve identity first
Origin hypotheses	Which host, microbial, dual, or exchange routes remain plausible?	Explicit competing explanations	Premature single-producer assumption	Maintain alternatives

Orthogonal testing	Do isotope, genome, compartment, and culture evidence converge?	Independent concordant observations	Cross-feeding, absent capacity, transport, instability	Increase or decrease confidence
Conflict resolution	Can discordant evidence be mechanistically explained?	Reproducible context-dependent explanation	Unresolved contradiction	Keep attribution provisional
Functional escalation	Can the candidate pathway be perturbed or reconstructed?	Product-linked functional response	Noninformative or host-dependent failure	Strengthen, revise, or retain uncertainty

Note: This workflow is proposed as a qualitative attribution architecture, not a validated scoring algorithm.

The mature synthesis is shown in **Figure 1**, in which competing origin hypotheses remain separate from evidence domains and converge only through a proposed confidence field bounded by falsification and validation requirements.

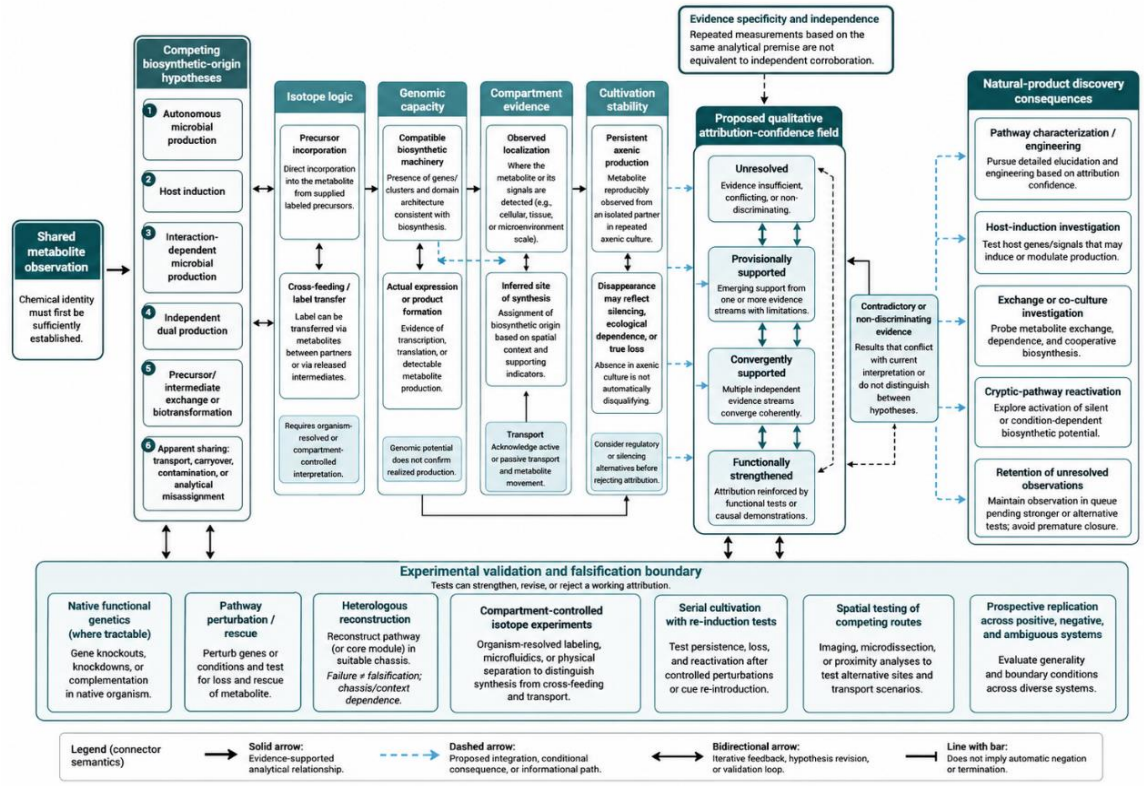

Figure 1. Proposed evidence-convergence architecture for host–endophyte biosynthetic attribution

Confidence in biosynthetic assignment

Confidence should reflect the convergence and independence of evidence rather than the number of observations accumulated. Genomics and metabolomics are complementary precisely because each constrains a different part of the inference: genomes delimit plausible biosynthetic machinery, whereas metabolomics records chemistry actually observed under specified conditions [28]. In the proposed model, multiple measurements derived from the same analytical premise do not automatically constitute independent corroboration. Repeated LC–MS detection, for example, may improve reproducibility of occurrence but does not substitute for genomic or tracer evidence.

The analytical identity of the metabolite remains a ceiling on downstream attribution. Mass-spectrometric workflows used for fungal endophyte natural products differ in their structural resolving power, and orthogonal characterization is particularly important when an origin claim depends on distinguishing closely related metabolites [29]. A confident pathway assignment to the wrong chemical structure is still a failed attribution.

Functional metabolomics adds another useful but bounded layer by connecting chemical features to biological phenotypes. Such associations can prioritize active metabolites and informative conditions, but phenotype linkage does not by itself reveal the pathway that produced the responsible molecule [30]. The same discipline applies to

computational linkage. Phylogeny-aware metabologenomics can improve natural-product–BGC assignments by accounting for organismal relatedness that can otherwise generate spurious correlations [31], yet an improved association is not equivalent to functional pathway proof.

Confidence also depends on specificity. A genomic observation is more discriminating when it explains chemically demanding transformations required for the candidate structure rather than merely identifying generic secondary-metabolism enzymes. An isotope experiment is more informative when alternative transfer routes predict distinguishable labelling patterns. Spatial evidence is stronger when the predicted producer and intermediates occupy a pattern that competing hypotheses would not readily generate. Thus, the proposed states are not a checklist in which four positive boxes automatically yield a high assignment; the decisive question is whether the combined evidence meaningfully separates one origin hypothesis from its strongest alternatives.

Evidence provenance should remain visible when communicating an assignment. A statement such as "microbial biosynthesis is convergently supported" should be accompanied in the surrounding analysis by the observations carrying that inference and the unresolved alternatives. This discourages confidence labels from becoming detached from the experiments that produced them and allows later contradictory data to revise the assignment without redefining the model.

Accordingly, the proposed confidence states are deliberately qualitative: unresolved, provisionally supported, convergently supported, and functionally strengthened. These terms describe evidential positions rather than numerical probabilities. A claim should move downward as well as upward when new evidence contradicts the assumed producer. The strongest useful assignment is therefore not the one with the largest volume of supportive data, but the one for which plausible alternatives have been actively tested and the remaining interpretation is consistent across structurally different forms of evidence.

Applications to natural-product discovery

Biosynthetic attribution changes what should be pursued after a host-associated metabolite is detected. If genomic and functional evidence supports an autonomous microbial pathway, the endophyte becomes a candidate source for pathway characterization, enzyme discovery, engineering, or controlled production. Enzyme-targeted genome mining coupled to experimental characterization has demonstrated how fungal genomic information can be converted into resolved natural-product pathways rather than remaining a catalogue of predicted capacity [32]. The proposed model uses this distinction to prevent producer designation from preceding the evidence required to justify producer-focused development.

An unresolved attribution can also be productive. Evidence for host induction may redirect discovery toward signaling molecules or interaction-dependent enhancement of plant biosynthesis. Evidence for precursor exchange may reveal division of metabolic labor and suggest co-culture rather than monoculture as the biologically relevant unit. Genomic compatibility with poor cultivation stability may justify experiments designed to reactivate silent microbial pathways rather than abandonment of the strain. Conversely, repeated chemical detection without credible biosynthetic capacity or organism-resolved incorporation should lower the priority of claims that depend specifically on autonomous microbial synthesis.

For dereplication and prioritization, the model can separate novelty of chemistry from novelty of biosynthetic ownership. Rediscovery of a known plant metabolite in an endophyte is not necessarily uninteresting, but its value depends on what is newly demonstrated: an independent pathway, unusual enzymology, metabolic exchange, or a host–microbe regulatory mechanism. This shifts emphasis from announcing chemical overlap to identifying the mechanistic feature that makes the overlap scientifically consequential.

These consequences are summarized in **Table 4** as boundaries between what an attribution state can reasonably motivate and what still requires validation.

Table 4. Proposed discovery actions and validation boundaries arising from biosynthetic attribution

Attribution situation	Reasonable discovery action	Key unresolved boundary	Validation priority
Convergent microbial support	Characterize pathway and production conditions	Native pathway sufficiency and regulation	Functional genetics or reconstruction
Host-induction support	Investigate signaling and host pathway response	Direct versus indirect elicitation	Controlled host perturbation

Exchange/biotransformation support	Resolve transferred precursor or intermediate	Direction and necessity of exchange	Compartment-controlled tracing
Genomic capacity without stable production	Test regulatory or ecological activation	Silence versus incorrect BGC assignment	Expression and induction experiments
Chemical sharing without origin evidence	Retain as unresolved observation	Producer identity	Orthogonal origin-discriminating tests

Note: The actions are proposed research priorities, not validated predictions of discovery success.

Experimental priorities suggested by the model

The most informative next experiment depends on which uncertainty currently dominates. When microbial genomic capacity is credible but pathway sufficiency remains uncertain, heterologous reconstruction can provide strong functional evidence. Development of engineered *Streptomyces* hosts for expression of natural-product biosynthetic gene clusters illustrates the value of a compatible chassis for testing encoded capacity [33]. Successful reconstruction that yields the expected product can strongly support sufficiency of the introduced genetic system, whereas failure is less decisive because missing partners, regulation, precursor supply, toxicity, or host incompatibility can produce false-negative outcomes.

Functional genetics in the native organism should be preferred when tractable because perturbing a candidate biosynthetic gene while monitoring the proposed product links genotype and chemistry within the original cellular context. Rescue or complementation can further distinguish pathway-specific effects from nonspecific loss of fitness. Where native manipulation is unavailable, heterologous reconstruction remains valuable but should be interpreted as a test of the introduced genetic system rather than a complete recreation of the natural host–endophyte interaction.

For isotope experiments, experimental separation should be designed around exchange rather than assuming it away. Labelled precursor supplied to the host, the endophyte, and defined co-cultures can be compared across time, with physical separation or compartment-specific sampling where feasible. Tracer incorporation should then be interpreted alongside chemical identity and biomass localization. A label appearing in microbial material after host feeding may indicate transfer rather than microbial de novo synthesis; reciprocal designs are therefore more informative than a single labelled condition.

Cultivation stability should likewise become a structured perturbation rather than a one-time observation. Serial passage, defined media, re-exposure to candidate host cues, and parallel monitoring of candidate pathway expression can distinguish persistent production from reversible silencing. Spatial evidence is most useful when it tests competing routes—for example, whether intermediates accumulate with microbial biomass, in host cells, or across an interface—rather than merely producing visually compelling colocalization maps.

Finally, prospective evaluation of the proposed model should include known positive, negative, and deliberately ambiguous systems. Attribution criteria should be specified before outcome interpretation, and investigators should document evidence that weakens as well as strengthens the favored hypothesis. Cross-laboratory replication would be particularly valuable for culture-stability and isotope findings that are sensitive to growth conditions and analytical handling. The model should be revised if contradictory evidence repeatedly exposes missing explanations, and rejected as a decision aid if its distinctions fail to improve reproducibility or experimental discrimination. Its immediate value is therefore methodological: making uncertainty testable and directing experiments toward the specific evidential gap that prevents a stronger biosynthetic claim.

Conclusion

Shared metabolites are observations about chemistry, not automatic statements about biosynthetic origin. In host–endophyte systems, the same chemical pattern can arise through autonomous production, host induction, interaction-dependent pathway activation, independent dual biosynthesis, precursor exchange, biotransformation, transport, or analytical error. Treating these possibilities as competing explanations changes the evidential question from whether a compound can be detected in both partners to which mechanistic account best survives discriminating tests.

The proposed Host–Endophyte Attribution Model organizes that evaluation around isotope logic, genomic capacity, compartment evidence, and cultivation stability, with functional perturbation or pathway reconstruction providing opportunities for stronger validation. No domain is universally sufficient. Tracer incorporation can be confounded by exchange, genome content by silent or incorrectly assigned pathways, spatial localization by

transport, and culture behavior by regulatory loss or ecological dependence. Confidence should therefore depend on convergence among genuinely orthogonal observations and should decrease when important contradictions remain unresolved.

This framework is intentionally qualitative and evidence-bounded. It does not provide validated probabilities, universal thresholds, or prospective performance claims. Its practical contribution is to separate chemical identity, biosynthetic plausibility, metabolic participation, spatial compatibility, persistent production, and functional sufficiency into claims that can be tested rather than conflated. Applied this way, shared host–endophyte chemistry becomes not proof of shared biosynthesis, but a starting hypothesis whose origin can be progressively constrained, challenged, and, where evidence permits, more securely attributed.

Acknowledgments: None

Conflict of Interest: None

Financial Support: None

Ethics Statement: None

References

1. Stadler M, Kolarik M. Taxol is NOT produced sustainably by endophytic fungi! – A case study for the damage that scientific papermills can cause for the scientific communities. Fungal Biol Rev. 2024;49:100367. doi:10.1016/j.fbr.2024.100367
2. Kumari P, Deepa N, Trivedi PK, Singh BK, Srivastava V, Singh A. Plants and endophytes interaction: A "secret wedlock" for sustainable biosynthesis of pharmaceutically important secondary metabolites. Microb Cell Fact. 2023;22(1):226. doi:10.1186/s12934-023-02234-8
3. Walther C, Vallet M, Reichelt M, Giri P, Rothe B, Negwer EJ, et al. A fungal endophyte alters poplar leaf chemistry, deters insect feeding and shapes insect community assembly. Ecol Lett. 2025;28(2). doi:10.1111/ele.70007
4. Aufrecht J, Veličković D, Tournay R, Couvillion S, Balasubramanian V, Winkler T, et al. Endophyte-induced systemic spatial reprogramming of metabolism in *Populus trichocarpa* roots under drought. New Phytol. 2026;251(1):438-54. doi:10.1111/nph.71227
5. Diab E, Du C, Tigani W, Elsayed SS, van Wezel GP. Plant coumarins modulate natural product biosynthesis in a *Streptomyces* root endophyte. J Nat Prod. 2026;89(3):980-91. doi:10.1021/acs.jnatprod.5c01626
6. Gamal A, Fikry E, Tawfeek N, El-Shafae AM, El-Sayed ASA, El-Domiaty MM. Production and bioprocessing of Taxol from *Aspergillus niger*, an endophyte of *Encephalartos whitelockii*, with a plausible biosynthetic stability: Antiproliferative activity and cell cycle analysis. Microb Cell Fact. 2024;23(1):78. doi:10.1186/s12934-024-02356-7
7. Cui J, Wu W, Zhang S, Li S, Xiao C, Cui S, et al. Gene mining and optimization of the biosynthesis of cannabidiol in *Fusarium solani* XG4-2. J Nat Prod. 2025;88(7):1591-604. doi:10.1021/acs.jnatprod.5c00298
8. Wang Y, Chen N, Deng K, Zhong X, Li Z, Li L, et al. Effects and molecular mechanism of endophytic elicitors on the accumulation of secondary metabolites in medicinal plants. Front Microbiol. 2025;16:1558567. doi:10.3389/fmicb.2025.1558567
9. Keller NP. Fungal secondary metabolism: Regulation, function and drug discovery. Nat Rev Microbiol. 2019;17(3):167-80. doi:10.1038/s41579-018-0121-1
10. Eagan JL, Keller NP. Fungal secondary metabolism. Curr Biol. 2025;35(11). doi:10.1016/j.cub.2025.02.029
11. Liu R, Peng XP, Newman DJ, Purchase D, Li G, Kusari S. Unlocking the metabolic potential of endophytic fungi through epigenetics: A paradigm shift for natural product discovery and plant–microbe interactions. Nat Prod Rep. 2025;42(10):1690-716. doi:10.1039/D5NP00028A
12. Guo S, Pardeshi L, Qin L, Cheung CY, Liu X, Fan L, et al. Systematic over-expression of secondary metabolism transcription factors to reveal the pharmaceutical potential of *Aspergillus nidulans*. Commun Biol. 2025;8(1):1444. doi:10.1038/s42003-025-08840-z

13. Kambhampati S, Hubbard AH, Koley S, Gomez JD, Marsolais F, Evans BS, et al. SIMPEL: Using stable isotopes to elucidate dynamics of context specific metabolism. Commun Biol. 2024;7(1):172. doi:10.1038/s42003-024-05844-z
14. Allen DK, Young JD. Tracing metabolic flux through time and space with isotope labeling experiments. Curr Opin Biotechnol. 2020;64:92-100. doi:10.1016/j.copbio.2019.11.003
15. Bueschl C, Kluger B, Neumann NKN, Doppler M, Maschietto V, Thallinger GG, et al. MetExtract II: A software suite for stable isotope-assisted untargeted metabolomics. Anal Chem. 2017;89(17):9518-26. doi:10.1021/acs.analchem.7b02518
16. Hoffmann F, Jaeger C, Bhattacharya A, Schmitt CA, Lisec J. Nontargeted identification of tracer incorporation in high-resolution mass spectrometry. Anal Chem. 2018;90(12):7253-60. doi:10.1021/acs.analchem.8b00356
17. Alseekh S, Aharoni A, Brotman Y, Contrepois K, D'Auria J, Ewald J, et al. Mass spectrometry-based metabolomics: A guide for annotation, quantification and best reporting practices. Nat Methods. 2021;18(7):747-56. doi:10.1038/s41592-021-01197-1
18. Yin Z, Huang W, Li K, Fernie AR, Yan S. Advances in mass spectrometry imaging for plant metabolomics—expanding the analytical toolbox. Plant J. 2024;119(5):2168-80. doi:10.1111/tpj.16924
19. Zou Y, Tang W, Li B. Exploring natural product biosynthesis in plants with mass spectrometry imaging. Trends Plant Sci. 2025;30(1):69-84. doi:10.1016/j.tplants.2024.08.002
20. Liu Z, Qin A, Zhang Y, Zhao Q, Li M, Liu H, et al. Mass spectrometry imaging: Principles and applications in plant research. New Phytol. 2026;250(1):102-14. doi:10.1111/nph.70834
21. Blin K, Shaw S, Vader L, Szenei J, Reitz ZL, Augustijn HE, et al. antiSMASH 8.0: Extended gene cluster detection capabilities and analyses of chemistry, enzymology, and regulation. Nucleic Acids Res. 2025;53(W1). doi:10.1093/nar/gkaf334
22. Zdouc MM, Blin K, Louwen NLL, Navarro J, Loureiro C, Bader CD, et al. MIBiG 4.0: Advancing biosynthetic gene cluster curation through global collaboration. Nucleic Acids Res. 2025;53(D1). doi:10.1093/nar/gkae1115
23. Robey MT, Caesar LK, Drott MT, Keller NP, Kelleher NL. An interpreted atlas of biosynthetic gene clusters from 1,000 fungal genomes. Proc Natl Acad Sci U S A. 2021;118(19). doi:10.1073/pnas.2020230118
24. Caesar LK, Butun FA, Robey MT, Ayon NJ, Gupta R, Dainko D, et al. Correlative metabologenomics of 110 fungi reveals metabolite–gene cluster pairs. Nat Chem Biol. 2023;19(7):846-54. doi:10.1038/s41589-023-01276-8
25. Ayon NJ, Earp CE, Gupta R, Butun FA, Clements AE, Lee AG, et al. Bioactivity-driven fungal metabologenomics identifies antiproliferative stemphone analogs and their biosynthetic gene cluster. Metabolomics. 2024;20(5):90. doi:10.1007/s11306-024-02153-8
26. Avalon NE, Murray AE, Baker BJ. Integrated metabolomic–genomic workflows accelerate microbial natural product discovery. Anal Chem. 2022;94(35):11959-66. doi:10.1021/acs.analchem.2c02245
27. Navarro-Muñoz JC, Selem-Mojica N, Mullowney MW, Kautsar SA, Tryon JH, Parkinson EI, et al. A computational framework to explore large-scale biosynthetic diversity. Nat Chem Biol. 2020;16(1):60-8. doi:10.1038/s41589-019-0400-9
28. Caesar LK, Montaser R, Keller NP, Kelleher NL. Metabolomics and genomics in natural products research: Complementary tools for targeting new chemical entities. Nat Prod Rep. 2021;38(11):2041-65. doi:10.1039/D1NP00036E
29. Gupta P, Verma A, Rai N, Singh AK, Singh SK, Kumar B, et al. Mass spectrometry-based technology and workflows for studying the chemistry of fungal endophyte derived bioactive compounds. ACS Chem Biol. 2021;16(11):2068-86. doi:10.1021/acschembio.1c00581
30. Vitale GA, Geibel C, Minda V, Wang M, Aron AT, Petras D. Connecting metabolome and phenotype: Recent advances in functional metabolomics tools for the identification of bioactive natural products. Nat Prod Rep. 2024;41(6):885-904. doi:10.1039/D3NP00050H
31. Boldt J, Porten C, Haeckl FPJ, Hug JJ, Panter F, Steglich M, et al. Phylogeny-aware metabologenomics accurately assigns natural products to biosynthetic gene clusters. Microb Biotechnol. 2026;19(1). doi:10.1111/1751-7915.70298
32. Tang J, Matsuda Y. Discovery of fungal onoceroid triterpenoids through domainless enzyme-targeted global genome mining. Nat Commun. 2024;15(1):4312. doi:10.1038/s41467-024-48771-7

33. Klumbys E, Xu W, Koduru L, Heng E, Wei Y, Wong FT, et al. Discovery, characterization, and engineering of an advantageous *Streptomyces* host for heterologous expression of natural product biosynthetic gene clusters. Microb Cell Fact. 2024;23(1):149. doi:10.1186/s12934-024-02416-y