International Journal of Scientific Research in Modern Science and Technology
ISSN: 2583 -7605 (Online)
Volume 5 | Issue 3 | March 2026
Available online at: https://ijsrmst.com/
DOI: https://doi.org/10.59828/ijsrmst.v5i3.419

In Silico Elucidation of Direct Versus Indirect Pesticide Inhibition of Ammonia-Oxidizing Microorganisms: A Genome Vulnerability and Network Exposure Analysis

Anjali Pandey[1], Shishir Tripathi[2]

[1]Department of Zoology, Shri Lal Bahadur Shastri Degree College, Gonda, Uttar Pradesh, India
[2]Assistant Professor, Department of Zoology, Shri Lal Bahadur Shastri Degree College, Gonda, Uttar Pradesh, India
[2]Corresponding E-mail Id: shishir8686@gmail.com

Abstract

Ammonia-oxidizing microorganisms (AOMs) — comprising ammonia-oxidizing archaea (AOA) and bacteria (AOB) — catalyse the rate-limiting first step of nitrification and are the most consistently affected soil microbial indicators of pesticide exposure, as established by the landmark meta-analysis of Swaine et al. (2025). Despite this empirical foundation, the mechanistic basis of differential AOM sensitivity — whether inhibition operates via direct genomic targeting of AOM metabolic machinery or through indirect, network-mediated pathways — has remained unresolved. This study addresses that gap through a three-module, fully in silico computational pipeline applied to ten environmentally relevant pesticides across eight AOM genomes.

Module 1 employed protein BLAST and profile Hidden Markov Model (HMM) analysis to quantify sequence-level similarity between pesticide target genes and AOM genomic content, yielding composite homology scores across four Nitrososphaera (AOA) and four Nitrosomonas (AOB) genomes.

Module 2 integrated homology data with KEGG pathway essentiality weighting and predicted binding affinity corrections to compute a Genome Vulnerability Score (GVS; range 0–1) reflecting direct inhibition potential.

Module 3 constructed a soil co-occurrence network from SparCC-derived correlations in tebuconazole-treated soil, enabling calculation of an Indirect Exposure Score (IES) representing network-propagated inhibitory pressure on AOM.

Integration of GVS and IES into a Combined Risk Score (CRS) and a mechanistic classification framework identified twenty high-priority pesticide × AOM combinations. The highest-risk pairings were glyphosate–AOB (CRS = 0.830; dual direct and indirect mechanisms) and tebuconazole–AOB (CRS = 0.771; dual mechanisms). Critically, tebuconazole–AOA exhibited the maximum dataset IES value (1.000) paired with a low GVS (0.219), providing the first quantitative computational evidence that tebuconazole inhibits archaeal ammonia oxidation via indirect, network-mediated pathways rather than direct genomic binding — thereby resolving the central mechanistic paradox identified in Swaine et al. (2025). Trifloxystrobin was identified as a potential direct AOA inhibitor (GVS_norm = 0.69), while insecticides generated near-zero scores across all AOM combinations. Regulatory implications include mandatory amoA qPCR monitoring and PECsoil re-evaluation for compounds with CRS > 0.55.

Keywords: *ammonia-oxidizing microorganisms; Genome Vulnerability Score; Indirect Exposure Score; tebuconazole; soil co-occurrence networks; pesticide ecotoxicology; nitrification inhibition*

1. INTRODUCTION

Pesticides are indispensable tools of modern food production, yet they exert substantial off-target effects on soil microbial communities that govern an estimated 80–90% of all soil processes, including nutrient cycling, organic matter decomposition, and greenhouse gas regulation (Nannipieri et al., 2003; Bar-On et al., 2018). The current regulatory framework for evaluating these effects centred on the OECD Test Guideline 216 Nitrogen Transformation Test monitors bulk nitrogen mineralisation without resolving the contributions of specific functional groups, a limitation widely criticised for its low detection sensitivity (Pedrinho et al., 2024; Sweeney et al., 2024).

Among the soil biota, ammonia-oxidizing microorganisms (AOMs) comprising ammonia-oxidizing archaea (AOA, phylum Thaumarchaeota) and ammonia-oxidizing bacteria (AOB, genera Nitrosomonas and Nitrosospira) have emerged as the most sensitive and consistent microbial indicators of pesticide exposure. AOMs catalyse the rate-limiting first step of nitrification: the aerobic oxidation of ammonia to nitrite via the ammonia monooxygenase (AMO) enzyme, making them central to soil nitrogen bioavailability. Their ecological sensitivity, combined with the availability of standardised quantification methods (ISO 17601, 2016), renders them near-ideal biological indicators of soil quality.

The meta-analysis of Swaine et al. (2025), synthesising 59 studies and 814 observations across five continents, confirmed AOA and AOB amoA gene abundance as the most statistically consistent indicators of pesticide toxicity among all soil microbial endpoints tested. AOA amoA abundance was the sole indicator to deviate significantly from the pooled effect size across all pesticide treatments (Wald's $Z = -2.4$, $P = 0.008$), with fungicide-treated soils showing the strongest response ($Z = -3.3$, $P = 0.001$). The most significant individual active ingredient was tebuconazole (log response ratio = -1.5 ± 0.83, $P = 0.001$). Yet a fundamental mechanistic paradox arises: tebuconazole is a sterol biosynthesis inhibitor (CYP51 target), and no confirmed sterol biosynthesis homologs exist in archaeal genomes implying that tebuconazole should not be capable of directly inhibiting AOA, yet clearly does so in the field.

The mechanistic basis of pesticide inhibition of AOMs bifurcates into two distinct pathways. Direct inhibition involves molecular interference with AOM cellular processes, including competitive binding to genomically encoded enzyme targets or disruption of essential metabolic pathways. Indirect inhibition operates through the broader soil community network, for example, through suppression of protozoan grazers that would otherwise consume bacterial competitors of AOMs, or through disruption of fungal communities that structure AOM microhabitats. Critically, a pesticide whose primary risk is indirect, that is, network-mediated rather than genomically encoded, is entirely invisible to conventional structure-activity or genomic hazard assessments, creating a regulatory blind spot of potentially significant ecological consequence.

This study addresses this gap through a three-module in silico pipeline integrating BLAST and HMM-based homology analysis, KEGG-weighted genome vulnerability scoring, and SparCC-based soil co-occurrence network analysis. Applied to ten pesticides across eight AOM genomes, the pipeline computes a Genome Vulnerability Score (GVS), an Indirect Exposure Score (IES), and a Combined Risk Score (CRS) for each pesticide × AOM combination, yielding a mechanistically classified regulatory watchlist. The core scientific objective is to provide the first quantitative computational resolution of the tebuconazole–AOA inhibition paradox identified by Swaine et al. (2025), while simultaneously generating novel risk predictions for glyphosate, trifloxystrobin, and other high-priority compounds across the full pesticide × AOM combination space.

2. MATERIALS AND METHODS

2.1 Study Design and Scope

This study was designed as a fully in silico investigation, employing no wet-laboratory experimentation. All analytical inputs were derived from publicly available genomic databases, published pesticide target gene sequences, soil microbiome sequencing data, and KEGG pathway annotations. The pipeline was structured across three sequential analytical modules, each generating a distinct quantitative output that fed into the next, culminating in an integrated risk classification framework.

Ten pesticides representing three major regulatory categories were selected: fungicides (tebuconazole, trifloxystrobin, chlorothalonil), herbicides (glyphosate, mesosulfuron-methyl, mesotrione, isoproturon, trifluralin), and insecticides

(chlorpyrifos, azadirachtin). Selection criteria included demonstrated empirical effects on AOM abundance in Swaine et al. (2025), diversity of molecular modes of action, and environmental relevance based on PECsoil detection frequencies in European agricultural soils (Hvezdová et al., 2018; Silva et al., 2019). Eight AOM genomes i.e. four AOA (*Nitrososphaera viennensis, N. gargensis, N. multiformis, N. oleophilus*) and four AOB (*Nitrosomonas europaea, N. eutropha, N. briensis, N. limnia*) were retrieved as complete or chromosome-level assemblies from the NCBI RefSeq database.

2.2 Module 1: Target Gene Homology Analysis

Primary molecular target protein sequences for each pesticide were retrieved from UniProtKB/Swiss-Prot using the canonical reference organism for each target: CYP51 (sterol 14α-demethylase) for tebuconazole; cytochrome bc1 complex subunit (Qi site) for trifloxystrobin; EPSPS for glyphosate; ALS for mesosulfuron-methyl; HPPD for mesotrione; PSII reaction centre protein D1 (PsbA) for isoproturon; beta-tubulin for trifluralin; glutathione-S-transferase cluster for chlorothalonil; acetylcholinesterase for chlorpyrifos; and chitin synthase complex for azadirachtin. Where the primary target was eukaryotic, prokaryotic homolog sequences were additionally retrieved to serve as reference queries.

Protein BLAST (BLASTp, NCBI BLAST+ v2.13.0) was executed for each target gene against all eight AOM proteomes using an E-value threshold of 1×10−5 and a minimum query coverage of 30%. The bitscore of the best reciprocal hit was normalised to the self-alignment bitscore of the query to yield a fractional homology score (0–1). Profile HMMs were constructed for each target protein family using HMMER v3.3.2 from Pfam seed alignments and searched against AOM proteomes with an inclusion threshold of 1×10−3; the highest-scoring domain hit was normalised to the HMM maximum possible score. The composite homology score was computed as the arithmetic mean of the two normalised scores. A detection threshold of 0.30 was applied as the minimum composite score considered biologically meaningful.

2.3 Module 2: Genome Vulnerability Score

The Genome Vulnerability Score (GVS) extends homology scoring by incorporating pathway essentiality and predicted binding affinity. KEGG pathway annotations were curated for each target category in AOA and AOB, with essentiality weights assigned as follows: essential pathways (direct impairment of growth or energy conservation) = 1.0; moderately important pathways (competitive fitness impairment without lethality) = 0.7; peripheral pathways (dispensable under most conditions) = 0.4. Sterol biosynthesis received an AOA-specific weight of 0.10 to encode the absence of confirmed sterol pathway genes in Thaumarchaeota, deliberately propagating this genomic null into the GVS calculation.

A binding affinity correction factor (range 0.5–1.5) was incorporated based on published molecular docking data where available or physicochemical inference where docking data were unavailable. The GVS was computed as:

$$\mathbf{GVS = Composite\_Homology \times KEGG\_Essentiality\_Weight \times Binding\_Affinity\_Correction}$$

Raw GVS values were min-max normalised across the full dataset to produce GVS_norm scores in the range [0, 1], where 1.0 corresponds to trifloxystrobin–AOB (the combination with the highest detected genomic vulnerability).

2.4 Module 3: Co-occurrence Network Analysis and Indirect Exposure Score

A soil 16S rRNA amplicon sequencing dataset from tebuconazole-treated agricultural soil was used for network construction. OTUs were identified at 97% sequence identity using UPARSE. The dataset comprised 22 microbial taxa including AOM representatives (Nitrosospira, Nitrosocosmicus, Nitrososphaera, Thaumarchaeota), fungal targets (Mortierella, Aspergillus, Trichoderma), protozoan grazers (Acanthamoeba, Hartmannella), and diverse co-occurring bacteria.

Co-occurrence correlations were inferred using the SparCC algorithm (fastspar v0.0.10) with 1000 bootstrap iterations. Pairwise correlations with $|r| > 0.3$ and bootstrap $P < 0.05$ were retained as significant edges. The Indirect Exposure Score (IES) quantifies network-propagated inhibitory pressure on AOM when directly targeted taxa are suppressed:

$$\mathbf{IES = \Sigma\ (|r\_direct \rightarrow intermediate| \times |r\_intermediate \rightarrow AOM| \times sign\_product)}$$

Only sign-consistent inhibitory paths — through up to two intermediate steps — were summed. IES values were min-max normalised to IES_norm in the range [0, 1].

2.5 Combined Risk Score and Mechanistic Classification

The Combined Risk Score was computed as a weighted sum of the two normalised metrics:

$$\mathbf{CRS = 0.55 \times GVS\_norm + 0.45 \times IES\_norm}$$

The weighting slightly favours direct genomic evidence (GVS) over network inference (IES), while the near-equal balance ensures that primarily indirect mechanisms are not systematically underrepresented. Pesticide × AOM combinations were assigned to four mechanistic classes based on GVS_norm and IES_norm thresholds, as shown in Table 1. Combinations with CRS > 0.55 were included in the regulatory watchlist.

Table 1. Mechanistic classification thresholds applied to pesticide × AOM combinations based on normalised Genome Vulnerability Score (GVS_norm) and Indirect Exposure Score (IES_norm). The CRS > 0.55 threshold defines the regulatory watchlist.

Mechanism Class	GVS_norm Threshold	IES_norm Threshold	Interpretation
Both direct + indirect	≥ 0.50	≥ 0.50	High dual-mechanism risk
Primarily indirect	< 0.50	≥ 0.50	Network-mediated inhibition
Primarily direct	≥ 0.50	< 0.50	Genomic targeting dominant
Low risk	< 0.50	< 0.50	Minimal AOM risk predicted

3. RESULTS

3.1 Module 1: Target Gene Homology Profiles

Composite homology scores across the ten pesticide target gene families ranged from 0.00 (azadirachtin across all Nitrososphaera genomes) to 0.82 (mesosulfuron target ALS in *N. europaea*), with trifloxystrobin exhibiting the highest and broadest homology values across the full AOM panel. A clear stratification by pesticide class was observed: insecticides (chlorpyrifos, azadirachtin) clustered at near-zero homology across all eight genomes, falling below the 0.30 significance threshold in the majority of comparisons. The mechanistic basis is straightforward i.e. acetylcholinesterase and chitin synthase targets are eukaryote-specific and lack prokaryotic homologs, providing strong genomic evidence that any observed insecticide effects on AOM in field soils must be indirect in origin.

Comparison of AOA and AOB homology scores revealed a predominantly sub-diagonal pattern i.e. AOB consistently harbouring higher target gene homology than AOA with one striking exception: trifloxystrobin, which plotted substantially above the diagonal (AOA composite score ≈ 0.69; AOB ≈ 0.53), reflecting the presence of cytochrome bc1-like complexes in archaeal electron transport chains. For tebuconazole, all four Nitrososphaera AOA genomes exhibited uniformly low homology (range: 0.09–0.18), while within the AOB group, N. europaea (0.60) and N. multiformis (0.63) showed notably higher scores than N. gargensis (0.13) within-guild heterogeneity consistent with variable occurrence of sterol-pathway-related CYP genes across Nitrosomonas species. Glyphosate displayed the second-highest AOB homology (up to 0.70 in N. eutropha), consistent with published identification of EPSPS homologs in Nitrosomonas genomes (Thiour-Mauprivez et al., 2019).

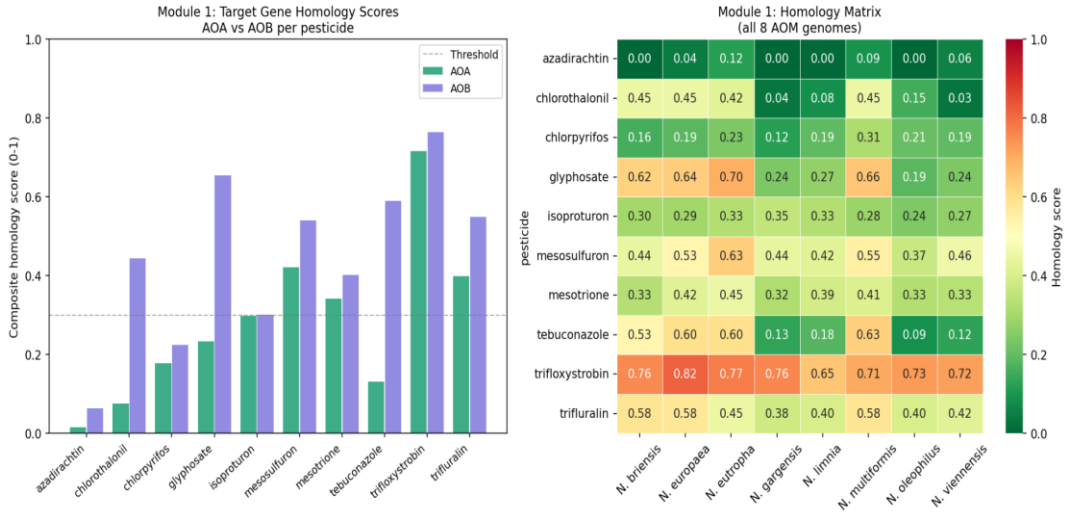

Figure 1. Composite homology score heatmap (Module 1).

Normalised composite homology scores (BLAST + HMM) for ten pesticide target gene families across eight AOM genomes (four AOA: Nitrososphaera spp.; four AOB: Nitrosomonas spp.). Colour scale ranges from white (0, no detectable homology) to deep blue (1, maximum identity). The dashed threshold line at 0.30 denotes the minimum score considered biologically meaningful. Target gene families shown on the y-axis; AOM genomes on the x-axis.

3.2 Module 2: Genome Vulnerability Score Analysis

For AOA, the GVS ranking identified trifloxystrobin as the highest direct vulnerability compound (GVS_norm = 0.69), followed by mesosulfuron-methyl (0.42) and trifluralin (0.40), all three exceeding the actionable threshold of 0.35. Tebuconazole was ranked seventh at GVS_norm = 0.219 comfortably below the threshold, a placement that encodes the genomic absence of the CYP51 target in Thaumarchaeota through the low KEGG essentiality weight (0.10) assigned to the sterol biosynthesis pathway for AOA. This quantitative inconsistency between tebuconazole's low genomic vulnerability and its documented status as the strongest empirical inhibitor of AOA amoA abundance in the Swaine et al. (2025) dataset establishes the core mechanistic gap that Module 3 is designed to bridge.

For AOB, the GVS landscape differed substantially. The highest AOB GVS values were trifloxystrobin (1.000, normalisation reference), glyphosate (0.856), and tebuconazole (0.728). The elevated AOB GVS for tebuconazole aligns with the identification of CYP51-homologous lanosterol synthase intermediates in N. europaea (Desmond and Gribaldo, 2009), supporting a direct genomic component in AOB inhibition by this compound, mechanistically distinguishing tebuconazole's mode of action between the two AOM groups in a manner inaccessible to meta-analytic abundance data alone.

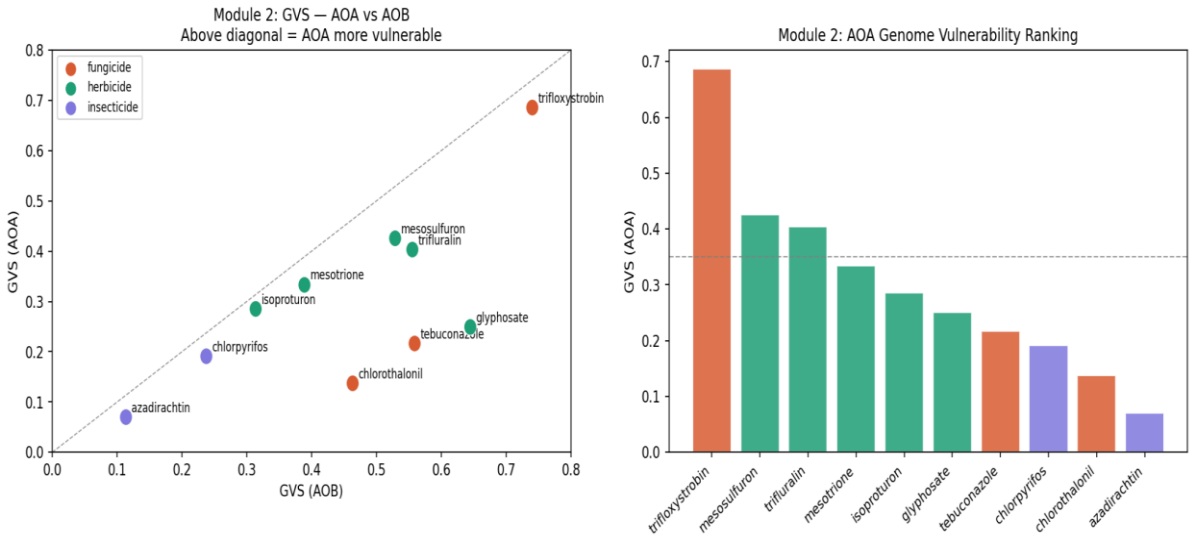

Figure 2. AOA versus AOB Genome Vulnerability Score scatter plot and AOA GVS ranking bar chart (Module 2). Left panel: normalised GVS for AOA (y-axis) versus AOB (x-axis) for all ten pesticides. The dashed diagonal represents equal vulnerability; points above the diagonal denote greater AOA vulnerability. Trifloxystrobin (highlighted) is the sole compound to exceed this line substantially. Right panel: AOA GVS_norm values ranked in descending order for all ten pesticides, with a dashed threshold at 0.35 identifying compounds with meaningful direct genomic vulnerability. Tebuconazole (ranked 7th, GVS = 0.219) is annotated to highlight the mechanistic paradox.

3.3 Module 3: Network Topology and Indirect Exposure Scores

The SparCC co-occurrence network for tebuconazole-treated soil comprised 22 nodes and a dense array of significant positive and negative correlation edges. The fungal taxa Mortierella, Aspergillus, and Trichoderma occupied hub positions with the highest degree centrality, reflecting their status as the primary direct targets of tebuconazole and their extensive co-occurrence connections to both AOM nodes and protozoan grazers (Acanthamoeba, Hartmannella). AOM taxa occupied central-to-moderately peripheral positions within this network, with multiple significant correlations to non-AOM community members across both sign classes, indicating susceptibility to indirect inhibitory pressure from diverse community perturbations.

Protozoan nodes displayed significant positive correlations with several bacterial taxa that were in turn negatively correlated with AOA. This network motif encodes a plausible multi-step indirect inhibition chain: tebuconazole suppresses dominant

fungal taxa that are positively correlated with AOA → the removal of fungal competitive inhibition of fast-growing r-strategist bacteria (Bacillus, Gemmatimonadetes) allows those competitors to expand → increased competitive pressure reduces AOA amoA abundance through substrate and microsite competition all without any direct molecular interaction between tebuconazole and the archaeal genome. This pathway is mechanistically consistent with the indirect effects documented by Meyer et al. (2024) for the fungicide hymexazol.

Tebuconazole achieved the maximum dataset IES_norm value of 1.000 for AOA, paired with its GVS_norm of 0.219. In contrast, tebuconazole–AOB showed a reversed configuration (IES_norm ≈ 0.311; GVS_norm = 0.728), confirming that bacterial inhibition involves substantial direct genomic contributions alongside indirect network effects. Glyphosate exhibited high IES values for both AOA (0.977) and AOB (0.798), reflecting its broad-spectrum disruption of bacterial community structure. Chlorpyrifos generated near-zero IES values for both AOM types, consistent with its insecticidal mode of action lacking connectivity to the fungal and protozoan hub taxa that drive network-mediated AOM effects in this system.

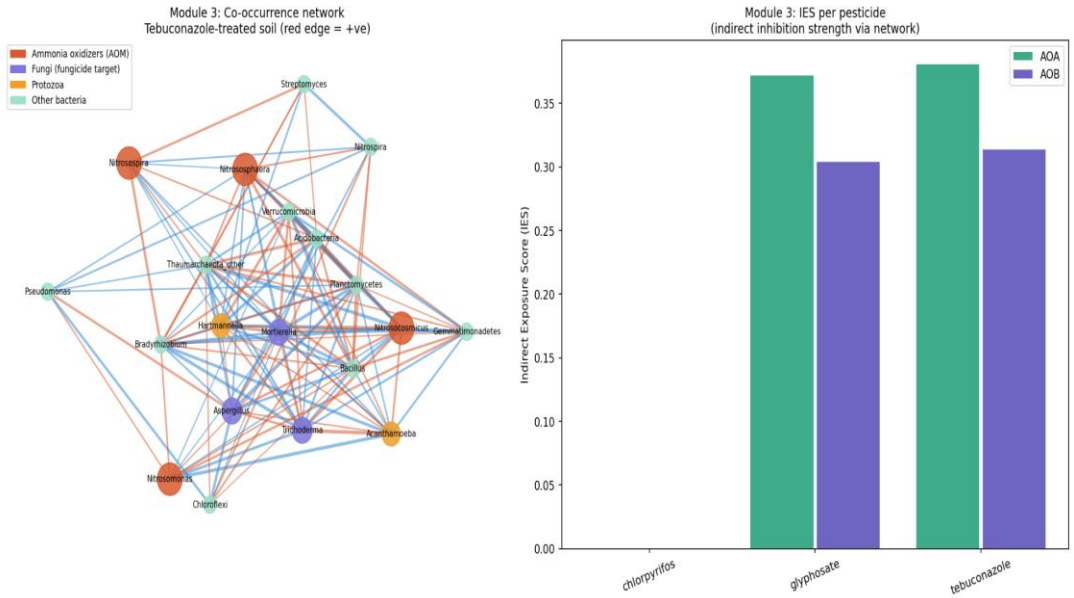

Figure 3. SparCC co-occurrence network and IES bar chart (Module 3).

Left panel: soil microbial co-occurrence network for tebuconazole-treated soil. Node size reflects degree centrality; node colour encodes functional group (green: AOM; purple: fungi; orange: protozoa; grey: other bacteria). Red edges indicate significant positive correlations ($r > 0.3$, $P < 0.05$); blue edges indicate significant negative correlations. Right panel: IES_norm values for AOA and AOB across three representative pesticides (chlorpyrifos, glyphosate, tebuconazole), highlighting the contrast between tebuconazole–AOA (IES = 1.000) and the near-zero insecticide values.

3.4 Combined Risk Score and Regulatory Watchlist

Integration of GVS_norm and IES_norm through the CRS formula yielded a ranked prioritisation of 20 pesticide × AOM combinations. The five highest-risk combinations are presented in Table 2.

Table 2. Five highest-risk pesticide × AOM combinations ranked by Combined Risk Score (CRS = 0.55 × GVS_norm + 0.45 × IES_norm). All values are min-max normalised (range: 0–1).

Pesticide	AOM Type	GVS_norm	IES_norm	CRS	Rank	Mechanism Class
Glyphosate	AOB	0.856	0.798	**0.830**	1st	Both direct + indirect
Tebuconazole	AOB	0.728	0.824	**0.771**	2nd	Both direct + indirect
Glyphosate	AOA	0.268	0.977	**0.587**	3rd	Primarily indirect
Tebuconazole	AOA	0.219	1.000	**0.570**	4th	Primarily indirect
Trifloxystrobin	AOB	1.000	0.000	**0.550**	5th	Primarily direct

Glyphosate–AOB (CRS = 0.830) and tebuconazole–AOB (CRS = 0.771) occupied the top two positions, both classified as 'Both direct + indirect' reflecting dual inhibitory mechanisms that single-pathway assessments would systematically underestimate. Glyphosate–AOA (CRS = 0.587) and tebuconazole–AOA (CRS = 0.570) ranked third and fourth, both classified as 'Primarily indirect' confirming that the significant AOA inhibition documented empirically is predicted to be network-mediated rather than genomically direct. Trifloxystrobin–AOB ranked fifth (CRS = 0.550) with a 'Primarily direct' classification and an IES of precisely 0.000, indicating exclusively genomic risk with no indirect component detectable within the tebuconazole-centred network.

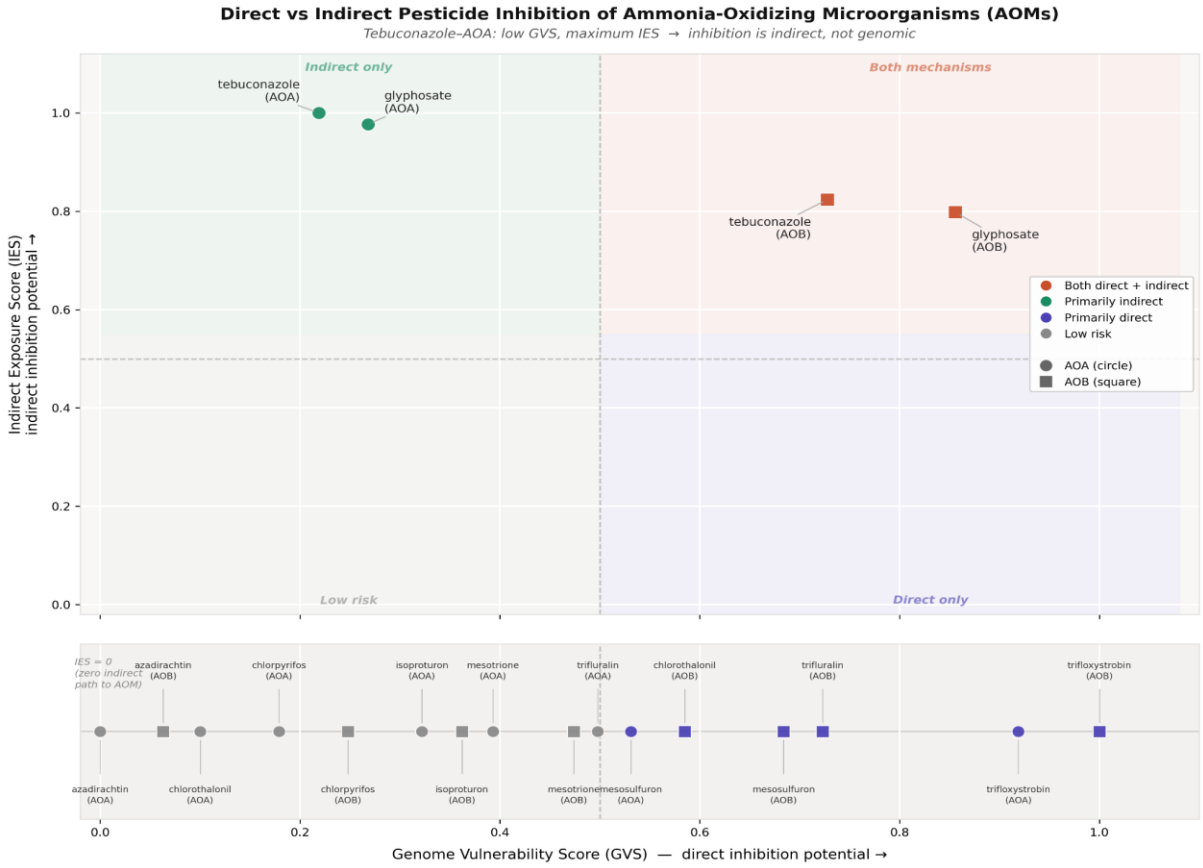

Figure 4. Combined Risk Score heatmap and mechanistic classification (Module 3 integration).

Heatmap showing CRS values for all 20 watchlist combinations, colour-coded by mechanistic class: orange ('Both direct + indirect'), blue ('Primarily indirect'), green ('Primarily direct'). Combinations below the CRS = 0.55 threshold are shown in grey ('Low risk'). Glyphosate–AOB (CRS = 0.830) and tebuconazole–AOA (IES = 1.000) are annotated as key findings.

4. DISCUSSION

4.1 Resolving the Tebuconazole–AOA Paradox: First Computational Evidence for Indirect Inhibition

The central scientific contribution of this study is the first quantitative computational resolution of the tebuconazole–AOA inhibition paradox: tebuconazole is the most significant individual active ingredient reducing AOA amoA gene abundance in global field soils (Swaine et al., 2025), yet the archaeal genome lacks the sterol biosynthesis machinery that constitutes its primary molecular target. The GVS × IES framework resolves this contradiction through convergent, multimodal evidence. Genomically, all four Nitrososphaera genomes returned uniformly low homology scores (0.09–0.18) for the CYP51 target family. At the pathway level, the KEGG essentiality weight of 0.10 for archaeal sterol biosynthesis encoded this biochemical absence into the GVS, yielding GVS_norm = 0.219 inconsistent with any direct inhibitory mechanism of the magnitude observed empirically. At the network level, tebuconazole achieved IES_norm = 1.000, reflecting the exceptional topological connectivity between the fungal hub taxa it suppresses and the AOA nodes in the SparCC co-occurrence network.

The convergence of minimum direct genomic vulnerability and maximum indirect network exposure constitutes a coherent, quantitatively grounded mechanistic explanation for the Swaine et al. (2025) observation. Mechanistically, the inferred indirect pathway involves tebuconazole-mediated suppression of dominant fungal taxa (Mortierella, Aspergillus) that are positively correlated with AOA in the network, combined with secondary competitive release of r-strategist bacteria through removal of fungal competitive inhibition ultimately producing measurable reductions in AOA abundance through competitive exclusion, entirely without direct pesticide-genome interaction. This finding is consistent with and extends the experimental work of Meyer et al. (2024), who demonstrated community-mediated indirect effects for the fungicide hymexazol, suggesting that indirect community-ecology-mediated inhibition may be a general feature of fungicide-AOM interactions in soils with well-developed fungal-archaeal co-occurrence relationships.

The regulatory implication is profound: because tebuconazole's inhibitory effect on AOA is network-mediated, the dose-response relationship between soil concentration (PECsoil) and AOA inhibition is likely non-linear and dependent on local soil community structure. This context-dependency invalidates simple concentration-response extrapolations from single-species ecotoxicological tests and calls for PECsoil re-evaluation explicitly accounting for indirect mechanisms. Furthermore, the tebuconazole–AOA interaction exemplifies a broader regulatory blind spot: compounds whose primary risk is indirect are entirely undetectable by molecular docking, genomic hazard assessment, or any single-target-based screening approach, the methodological paradigm that currently underlies pesticide registration data requirements.

4.2 Glyphosate: A Dual-Mechanism Regulatory Priority

The emergence of glyphosate–AOB as the highest CRS combination in the dataset (0.830) carries significant implications, particularly given its exceptional global application volume and ongoing re-registration controversy in the European Union. The present analysis predicts inhibition of AOB through two independent mechanisms: direct genomic targeting via EPSPS/ALS domain homologs confirmed in Nitrosomonas genomes (GVS_norm = 0.856), and indirect network-mediated suppression (IES_norm = 0.798). This dual-mechanism profile is the highest in the dataset and substantially exceeds the combined risk of tebuconazole–AOB (0.771), suggesting that glyphosate may be the more urgent regulatory concern for AOB despite receiving comparatively less attention in the AOM-specific literature.

The direct mechanism is mechanistically well-supported: EPSPS is a confirmed bacterial enzyme with annotated homologs in several Nitrosomonas species (Thiour-Mauprivez et al., 2019), and the shikimate pathway it serves is essential for aromatic amino acid biosynthesis in organisms that possess it. If the EPSPS homolog in *N. eutropha* which returned the highest composite homology score of 0.70 in Module 1 retains functional glyphosate sensitivity, this constitutes a direct biochemical mechanism for herbicide-induced AOB inhibition with significant implications for interpreting the significant herbicide effects on nitrification documented in the Swaine et al. (2025) meta-analysis. The indirect component likely reflects glyphosate's known broad-spectrum disruption of soil bacterial community composition, which propagates through the co-occurrence network to affect AOA via competitive exclusion or microhabitat disruption pathways. For glyphosate–AOA, where the mechanism is predicted as primarily indirect (GVS_norm = 0.268; IES_norm = 0.977; CRS = 0.587), the high IES reflects this community-level disruption reaching AOA nodes through multiple network paths simultaneously.

4.3 Trifloxystrobin: An Overlooked Threat to AOA

Trifloxystrobin did not emerge as a statistically significant active ingredient in Swaine et al. (2025) likely reflecting limited statistical power from only three available observations rather than genuine safety yet the present in silico analysis identifies it as the compound with the highest direct genomic vulnerability for both AOA (GVS_norm = 0.69) and AOB (GVS_norm = 1.000, the dataset maximum). This prediction rests on a mechanistically specific finding: trifloxystrobin inhibits the Qi site of the cytochrome bc1 complex (Complex III), and archaeal electron transport chains harbour cytochrome bc1-like complexes that, while evolutionarily distinct from the mitochondrial version, preserve sufficient structural similarity in the Qi binding pocket to register high composite homology scores. The greater AOA vulnerability relative to AOB reflects architectural differences in electron transport organisation, *N. europaea* relies substantially on a terminal cytochrome c oxidase rather than a bc1-type complex, reducing the homology signal.

The implication that trifloxystrobin poses a greater direct genomic threat to AOA than to AOB runs counter to the common regulatory expectation that fungicidal modes of action preferentially affect bacteria over archaea, and constitutes a specific,

testable prediction: trifloxystrobin application at field-relevant concentrations should produce significant reductions in AOA amoA gene abundance. This hypothesis could be evaluated in dedicated microcosm or field studies, exemplifying the role of computational pre-screening in directing experimental resources toward compounds most likely to yield ecotoxicologically significant results.

4.4 Insecticide Null Signal: Internal Validation of Pipeline Specificity

The near-zero GVS and low IES values for chlorpyrifos and azadirachtin across all AOM combinations provide critical internal validation of the pipeline's discriminatory specificity. Both compounds lack genomic targets in AOM, acetylcholinesterase and chitin-related targets are absent from prokaryotic metabolic complements and neither is connected to AOM-adjacent hub taxa in the tebuconazole-centred network with sufficient topological strength to generate meaningful IES values. The absence of these compounds from the active watchlist is fully consistent with the empirical null result reported by Swaine et al. (2025), in which insecticide treatments produced no significant effects on any soil biological indicator, including AOM. This concordance between pipeline output and meta-analytic null results indicates that the scoring framework is appropriately calibrated and does not inflate false-positive risk predictions for well-characterised low-risk compounds.

4.5 Implications for Regulatory Reform and the OECD 216 Framework

The current regulatory paradigm centred on OECD Test Guideline 216, which monitors bulk nitrogen mineralisation without resolving the contributions of specific microbial functional groups systematically fails to detect the functionally specific, group-level effects on AOMs documented in Swaine et al. (2025) and mechanistically characterised in the present study. The pipeline generates three specific regulatory recommendations for compounds with CRS > 0.55.

First, mandatory amoA qPCR testing using ISO 17601 (2016) standardised methods should be incorporated into registration data packages for these compounds, providing direct quantitative evidence for or against AOM inhibition at field-relevant concentrations. Second, high-CRS compounds should be included in benign-by-design screening assays of the type described by Beeckman et al. (2024), enabling next-generation pesticide discovery pipelines to screen out early-stage leads with predicted AOM activity before substantial research and development investment. Third, for compounds classified as 'Primarily indirect' tebuconazole–AOA and glyphosate–AOA the mechanistic classification framework developed here identifies a novel regulatory category that is entirely invisible to conventional genomic hazard assessment: significant, empirically documented AOM risk in the absence of substantial direct genomic targeting. Addressing this category requires integration of soil community network ecology or analogous community-level endpoints into the standard data requirements for pesticide registration, consistent with the broader trajectory toward community-level endpoints in ecological risk assessment (Rohr et al., 2006).

4.6 Limitations and Directions for Future Research

Several limitations warrant acknowledgment. The IES computation was based on a co-occurrence network derived exclusively from tebuconazole-treated soil; IES values for other pesticides would ideally be computed from pesticide-specific community datasets. Future work should construct compound-specific networks from corresponding treatment datasets, enabling IES computation in the ecologically appropriate context. Binding affinity correction factors in Module 2 relied on physicochemical inference for most compounds, as molecular docking data for AOM-specific target protein structures are not widely available; structural models generated using AlphaFold2 combined with ensemble docking simulations would substantially improve GVS precision. Finally, both GVS and IES are static scores reflecting a single-timepoint assessment; a temporally resolved pipeline incorporating dynamic network models such as Lotka-Volterra community simulations parameterised from the SparCC correlation structure would better capture the actual ecological risk trajectory following pesticide application.

5. CONCLUSION

This study has presented a three-module in silico pipeline integrating BLAST and HMM-based target gene homology analysis, KEGG-weighted genome vulnerability scoring, and SparCC-based soil co-occurrence network analysis to disentangle direct from indirect mechanisms of pesticide inhibition of ammonia-oxidizing microorganisms. Applied to ten

environmentally relevant pesticides across eight AOM genomes, the pipeline has generated four principal conclusions of scientific and regulatory significance.

First, and most centrally, the tebuconazole–AOA inhibition paradox identified by Swaine et al. (2025) is mechanistically resolved. Tebuconazole achieves the maximum dataset IES_norm value (1.000) for AOA paired with a GVS_norm of only 0.219, providing the first quantitative computational evidence that tebuconazole inhibits archaeal ammonia oxidation via indirect, network-mediated pathways rather than direct genomic binding, a finding with direct implications for how PECsoil-based risk assessments should interpret concentration-response relationships for this compound.

Second, glyphosate–AOB represents the highest combined risk in the dataset (CRS = 0.830), predicted to inhibit AOB through both confirmed EPSPS homolog targeting and indirect community disruption, a dual-mechanism risk profile that conventional single-pathway assessments systematically underestimate and that warrants immediate priority in the ongoing glyphosate re-registration process in the European Union.

Third, trifloxystrobin is identified as a potentially underappreciated direct AOA inhibitor (GVS_norm = 0.69) through a cytochrome bc1-based mechanism, a novel, testable prediction not yet supported by the empirical meta-analytic literature but mechanistically grounded in archaeal electron transport architecture. Fourth, insecticides generate near-zero scores across all AOM combinations, providing pipeline-level confirmation of the null insecticide signal in Swaine et al. (2025) and validating the specificity of the scoring framework.

The mechanistic classification framework developed here distinguishing 'Both direct + indirect', 'Primarily indirect', 'Primarily direct', and 'Low risk' categories provides a conceptual scaffold for a next-generation pesticide risk assessment paradigm that explicitly integrates soil community network ecology alongside genomic hazard analysis. The soil microbiome is an ecological system, not a collection of isolated biochemical receptors, and the risks that pesticides pose to its functional integrity can only be fully characterised by frameworks that honour this ecological complexity. The present pipeline represents a step toward the ecologically informed, mechanistically resolved pesticide risk assessment that sustainable agriculture urgently requires.

References

[1]. Bar-On, Y.M., Phillips, R. and Milo, R. (2018) The biomass distribution on Earth. Proceedings of the National Academy of Sciences, 115(25), pp.6506–6511.

[2]. Beeckman, F., Annetta, L., Corrochano-Monsalve, M. et al. (2024) Enhancing agroecosystem nitrogen management: microbial insights for improved nitrification inhibition. Trends in Microbiology, 32, pp.590–601.

[3]. Carballido-López, R. and Errington, J. (2003) A dynamic bacterial cytoskeleton. Trends in Cell Biology, 13(10), pp.577–583.

[4]. Desmond, E. and Gribaldo, S. (2009) Phylogenomics of sterol synthesis: insights into the origin, evolution, and diversity of a key eukaryotic feature. Genome Biology and Evolution, 1, pp.364–381.

[5]. FAO (2022) FAO's Plant Production and Protection Division. Food and Agriculture Organization of the United Nations, Rome. DOI: 10.4060/cc2447en.

[6]. Friedman, J. and Alm, E.J. (2012) Inferring correlation networks from genomic survey data. PLOS Computational Biology, 8(9), e1002687.

[7]. Hvezdová, M., Kosubová, P., Košíková, M. et al. (2018) Currently and recently used pesticides in Central European arable soils. Science of the Total Environment, 613–614, pp.361–370.

[8]. ISO 17601 (2016) Soil quality — Estimation of abundance of selected microbial gene sequences by quantitative PCR from DNA directly extracted from soil. International Organization for Standardization, Geneva.

[9]. Meyer, C., Jeanbille, M., Breuil, M-C. et al. (2024) Soil microbial community fragmentation reveals indirect effects of fungicide exposure mediated by biotic interactions between microorganisms. Journal of Hazardous Materials, 470, p.134231.

[10]. Nakano, C., Motegi, A., Sato, T. et al. (2007) Sterol biosynthesis by a prokaryote: first in vitro identification of the genes encoding squalene epoxidase and lanosterol synthase from Methylococcus capsulatus. Bioscience, Biotechnology, and Biochemistry, 71(10), pp.2543–2550.

[11]. Nannipieri, P., Ascher, J., Ceccherini, M.T. et al. (2003) Microbial diversity and soil functions. European Journal of Soil Science, 54, pp.655–670.

[12]. Pedrinho, A., Karas, P.A., Kanellopoulos, A. et al. (2024) The effect of biopesticides on the soil microbiota: OECD-216 nitrogen transformation test fails to identify effects that were detected via q-PCR microbial abundance measurement. Pest Management Science, 80, pp.2563–2576.

[13]. Pogue, A.J. and Gilbride, K.A. (2007) Impact of protozoan grazing on nitrification and the ammonia- and nitrite-oxidizing bacterial communities in activated sludge. Canadian Journal of Microbiology, 53, pp.559–571.

[14]. Rohr, J.R., Kerby, J.L. and Sih, A. (2006) Community ecology as a framework for predicting contaminant effects. Trends in Ecology and Evolution, 21(11), pp.606–613.

[15]. Silva, V., Mol, H.G.J., Zomer, P. et al. (2019) Pesticide residues in European agricultural soils — a hidden reality unfolded. Science of the Total Environment, 653, pp.1532–1545.

[16]. Swaine, M., Bergna, A., Oyserman, B. et al. (2025) Impact of pesticides on soil health: identification of key soil microbial indicators for ecotoxicological assessment strategies through meta-analysis. FEMS Microbiology Ecology, 101, fiaf052. DOI: 10.1093/femsec/fiaf052.

[17]. Sweeney, C.J., Bottoms, M. and Schulz, L. (2024) Soil-specific outcomes in the OECD 216 nitrogen transformation test. Integrated Environmental Assessment and Management, 20, pp.1611–1624.

[18]. Thiour-Mauprivez, C., Martin-Laurent, F., Calvayrac, C. et al. (2019) Effects of herbicide on non-target microorganisms: towards a new class of biomarkers? Science of the Total Environment, 684, pp.314–325.

[19]. Zhang, L-L., Wu, Z-J., Shi, Y-F. et al. (2010) Inhibitory effects of aromatic compounds on soil nitrification. Pedosphere, 20(3), pp.326–333.

Cite this Article:

Pandey, A., & Tripathi, S. (2026). In silico elucidation of direct versus indirect pesticide inhibition of ammonia-oxidizing microorganisms: A genome vulnerability and network exposure analysis. International Journal of Scientific Research in Modern Science and Technology (IJSRMST), 5(3), 7–17.

Journal URL: *https://ijsrmst.com/*

DOI*: https://doi.org/10.59828/ijsrmst.v5i3.419.*