

Notice de synthèse

en vue de l’Habilitation à Diriger des Recherches

Discipline : Chimie

Titre provisoire du mémoire

*« Augmenter les sciences moléculaires par l’intégration et l’intelligence artificielle : des spectres
aux holobiontes, de l’expérience à la découverte »*

Louis-Félix Nothias

Chaire de Professeur Junior (CPJ), CNRS

Responsable de l’HolobiomicsLab — équipe eDIAM

Institut de Chimie de Nice (ICN), UMR 7272 CNRS, Université Côte d’Azur

Garante : Élise Dumont (Professeure, ICN ; membre de l’IUF)

Résumé

Mes recherches développent des méthodes de chimie analytique/computationnelle et d’intelligence artificielle (IA) pour les sciences moléculaires. Cette démarche, résolument interdisciplinaire et transdisciplinaire, s’inscrit dans une évolution profonde du métier de chercheur. Depuis ma thèse en chimie des produits naturels, ma trajectoire suit un fil unique : **lever les verrous successifs qui limitent l’exploration du métabolome — sa nature et ses fonctions, des organismes aux écosystèmes**. J’ai d’abord contribué à des méthodes devenues des références pour l’*annotation* des spectres de fragmentation — réseaux moléculaires, annotation structurale computationnelle (algorithmique et par apprentissage automatique) —, ainsi qu’à des bibliothèques spectrales publiques en métabolomique et produits naturels [28–30], qui alimentent les modèles prédisant les propriétés structurales des molécules à partir de leurs spectres ; puis à l’*acquisition* expérimentale des spectres de masse et à l’*intégration* multi-omique, des écosystèmes microbiens (*Earth Microbiome Project*) aux *holobiontes* (hôte et microbiome associé) ; avant d’aborder le *raisonnement* assisté par graphes de connaissances et systèmes d’IA agentique opérant avec de grands modèles de langage. Ces avancées convergent vers une conviction : l’annotation n’est plus le verrou principal ; les prochains leviers sont l’intégration et l’IA. D’une part, pour rapprocher le calcul des instruments de laboratoire, interconnecter les omiques dès l’expérience et augmenter la performance analytique en temps réel ; d’autre part, pour intégrer données moléculaires et littérature scientifique, et accroître ainsi notre capacité à produire des connaissances à partir des données massives générées en continu par l’expérimentation. C’est le programme que je structure aujourd’hui, dans le cadre d’une Chaire de Professeur Junior CNRS, au sein de l’équipe eDIAM, autour de trois piliers — « des spectres aux molécules », « des atomes aux écosystèmes », « de l’expérience à la découverte ». Cette dynamique multi-omique s’étend du microbiome aux écosystèmes (cnidaires, projet METASTRESS) jusqu’à la médecine de précision (4P) avec le LabCom Preciomics-AI en préparation. Le programme s’appuie sur des projets financés (ANR-MetaboLinkAI ; IDEX UniCA — Starting package, Académie 4, RAPID-NP ; ANR-POPS), structuré par la candidature ERC HOLOBIOMICS et prolongé par des demandes en cours (ANR AI-FORGE, LabCom ANR Preciomics-AI), incarnés dans des briques logicielles ouvertes et réutilisables comme dans des méthodes transposables à d’autres laboratoires (amélioration de la performance analytique par modernisation et mise à niveau d’instruments existants), plaçant la science ouverte et l’humain au centre de la co-production augmentée par l’IA des moyens et des connaissances scientifiques.

1. Contexte et fil conducteur

La spectrométrie de masse est devenue la méthode de choix pour l’analyse moléculaire détaillée du vivant et de son environnement [15] ; sa sensibilité et sa large gamme de détection par électronébulisation en font l’instrument central de la métabolomique. Mais une fraction majeure des signaux acquis reste non interprétée : le métabolome demeure largement sous-caractérisé — typiquement, moins de 5 % des signaux détectés sont annotés par les bibliothèques spectrales [14]. Formé à la chimie des produits naturels durant ma thèse (Université de Corse, 2015 ; diterpènes antiviraux d’*Euphorbia*) [1], j’ai consacré mes travaux suivants à rendre ce métabolome plus accessible, lisible, puis exploitable, en combinant chimie analytique, informatique et, aujourd’hui, intelligence artificielle.

Ma trajectoire se lit comme la levée de verrous successifs : *annoter* les spectres, *acquérir* plus profondément les données, *intégrer* les omiques (initialement à l’échelle des écosystèmes, désormais en modélisant les holobiontes), puis *raisonner* sur ces données à l’aide de l’IA. Cette section retrace ce bilan (§2), explicite pourquoi l’intégration et l’IA constituent les prochains leviers — catalyseurs à plusieurs niveaux, de l’expérimentation à la méta-interprétation (§3) —, et présente le programme de recherche auquel je me consacre désormais (§4).

2. Bilan : quatre verrous levés

Annoter — donner un sens chimique aux spectres

Ma thèse (Université de Corse, co-réalisée à l’Institut de Chimie des Substances Naturelles, CNRS, Gif-sur-Yvette) m’a confronté à la difficulté d’isoler et d’élucider la structure de produits naturels complexes (esters de diterpènes d’*Euphorbia*) [1] ; mes premières applications des réseaux moléculaires m’en ont révélé le potentiel comme les limites pour décrypter leurs voies de fragmentation, et la valeur d’une annotation — même partielle — obtenue à grand débit. Cette expérience a fondé mon fil conducteur. Lors de mon post-doctorat (Université de Californie San Diego, avec le Prof. Pieter Dorrestein), j’ai co-développé des cadres méthodologiques de référence pour la génération de *réseaux moléculaires* à partir des spectres de fragmentation collectés lors d’expériences de métabolomique par spectrométrie de masse — le *Feature-Based Molecular Networking* (*Nature Methods*, 2020) [2], aujourd’hui largement adopté ; sa combinaison avec le fractionnement bioguidé (*bioactivity-based molecular networking*) [31] ; et son extension reliant les espèces ioniques d’un même composé (*ion identity molecular networking*) [6]. J’ai ensuite contribué directement au projet de recherche mené par le Prof. S. Böcker (Université de Iéna, Allemagne), à des méthodes d’annotation structurale computationnelle — algorithmiques et par apprentissage automatique — capables d’annoter des métabolites dépourvus de spectre de référence, connus comme inconnus (*Nature Biotechnology*, 2021 et 2022) [3,4]. J’ai par ailleurs contribué à des bibliothèques spectrales publiques en métabolomique et produits naturels — la *NIH Natural Products Library* [29] et des bibliothèques de fragmentation multi-niveaux (MSn) ouvertes (*MSnLib*, *Nature Methods*, 2025) [28], ainsi qu’une bibliothèque de suspects à l’échelle des dépôts publics [30], diffusées via le répertoire GNPS — ressources communautaires qui alimentent les modèles prédisant les propriétés structurales des molécules à partir de leurs spectres. Ces méthodes ont conduit l’équipe participante, que j’ai coordonnée pour l’occasion, aux meilleures performances de la compétition internationale CASMI 2022 — un jalon, preuve de concept d’un pipeline d’annotation *de bout en bout* couplant en cascade adduits ioniques, formule moléculaire, classe chimique, appariement à l’échelle des dépôts, information taxonomique et annotation structurale —, même si des méthodes alternatives remarquables pour certaines tâches d’annotation ont depuis émergé. Bilan : l’annotation, associée à un traitement complet du signal, longtemps le principal goulet d’étranglement, n’est plus le verrou dominant pour les spectres effectivement acquis.

Acquérir — atteindre la profondeur du métabolome

Des travaux liés à ma thèse m’avaient montré que le fractionnement, couplé à l’analyse LC-MS/MS, révèle une profondeur spectrale bien supérieure à celle d’un extrait brut [31] : convaincu que le verrou

se déplaçait vers l’acquisition, j’ai expérimenté, avec Hosein Mohimani (Carnegie Mellon University), un concept d’acquisition itérative et algorithmique en spectrométrie de masse (*MS2Planner*, *Bioinformatics*, 2021) [5], démontrant qu’un pilotage intelligent de l’instrument augmente à la fois la couverture et la qualité des spectres de fragmentation, rejoignant les efforts d’autres groupes pionniers de l’acquisition pilotée par ordinateur (acquisition itérative *data-set-dependent* [33] ; cadre *ViMMS* [36]). Ce travail a fondé ma conviction qu’il faut interconnecter l’instrument et le calcul, jusqu’à des acquisitions auto-optimisées par IA en temps réel. Cette ligne se poursuit dans mon équipe : les travaux de thèse de M. Bekbergenova (prédiction de l’identifiabilité des spectres pour mieux cibler l’acquisition) et la contribution de L. Pradi au projet *DeepMeta* visent à mettre en place une acquisition de spectrométrie de masse pilotée par l’IA.

Intégrer — du microbe à l’écosystème

Dès ma thèse, l’analyse de la diversité chimique du genre *Euphorbia* — l’interpréter à l’échelle de dizaines d’espèces et en formuler des hypothèses écologiques et évolutives [32] — m’avait confronté au défi de l’intégration à l’échelle des organismes et des écosystèmes. Au sein du *Center for Microbiome Innovation*, j’ai conçu et co-implémenté le volet métabolomique et contribué à l’intégration multi-omique de l’*Earth Microbiome Project* (*Nature Microbiology*, 2022) [7], l’un des plus vastes inventaires moléculaires du microbiome à l’échelle globale. J’ai également participé à des méthodes de *genome mining* reliant clusters de gènes biosynthétiques et métabolites microbiens (MetaMiner, *Cell Systems*, 2019 [8] ; Nerpa, 2021 [9]). Ces expériences ont forgé ma vision : les métabolites spécialisés du microbiome échappent majoritairement aux stratégies usuelles et exigent d’**interconnecter les omiques dès la conception expérimentale**, et non seulement au stade de l’analyse. Ce cheminement vers l’intégration s’est étendu à des holobiontes variés — éponges [10], coraux et macroalgues [11,12], avec des ressources reliant génomes et métabolomes [13] — et se poursuit dans le projet *ANR-POPS* [26] (polyamines microbiennes) — une dynamique appelée à s’accélérer avec la maturation des méthodes.

Raisonner — amorcer l’assistance par l’IA

Une fois l’annotation rendue largement systématique, l’enjeu se déplace vers le *raisonnement* : donner du sens à cette masse de données pour en construire des connaissances. Comme chercheur associé à l’Université de Genève (avec le Prof. Jean-Luc Wolfender), j’ai gagné des compétences en graphes de connaissances pour les métabolites bioactifs et prototypé, avec l’équipe Wimmics (Centre Inria d’Université Côte d’Azur, F. Gandon), un système multi-agent fondé sur des grands modèles de langage (*Large Language Models*, LLMs) interrogeant ces graphes en langage naturel (préfigurant MetaboT [18] et Mimosa-AI). Cette étape marque le passage de l’annotation au *raisonnement* assisté, et la transition vers mon programme actuel.

3. Charnière : l’intelligence artificielle, prochain levier

Ces quatre étapes dessinent moins une conclusion qu’une trajectoire : l’annotation demeure en partie un goulot d’étranglement, mais cesse progressivement d’être le verrou dominant — un horizon proche vers lequel la communauté s’avance et auquel mon programme contribue. Les défis qui demeurent — détecter des métabolites rares *in situ*, relier génomes et chimie, naviguer une littérature et des données massives — relèvent désormais de l’*intégration* et du *raisonnement*, là où l’essor d’une science augmentée par l’IA ouvre une voie nouvelle [34,35]. Ma conviction, héritée de ce parcours, est double : (i) les omiques doivent être interconnectées expérimentalement dès le départ, et non assemblées *a posteriori* ; (ii) l’IA, notamment couplée aux graphes de connaissances, peut, en temps réel, augmenter la performance analytique de la spectrométrie de masse, et — à la faveur de l’émergence des modèles de fondation, en mobilisant la littérature scientifique dès l’expérimentation et en enrichissant l’interprétation — accroître la production de connaissances [18]. C’est l’objet du programme de recherche que je structure et déploie.

4. Vision : le programme eDIAM

Ma vision de recherche s'intègre dans le programme collectif de l'équipe eDIAM, que j'ai co-conçu avec Élise Dumont et que je co-structure à travers les différents projets de mon groupe de recherche, que je déploie et orchestre. Au titre de ma Chaire de Professeur Junior CNRS, je coordonne le groupe HolobiomicsLab au sein de l'équipe eDIAM (*Équipe d'Intelligence Artificielle pour les Sciences Moléculaires*, ICN), dont les recherches, co-dirigées avec la Prof. Élise Dumont, développent des méthodes d'IA pour la chimie en combinant approches expérimentales, informatique, science des données et modélisation moléculaire. Le programme s'organise en trois **axes** — « des spectres aux molécules », « des atomes aux écosystèmes », « de l'expérience à la découverte » — au sein desquels les quatre verrous du bilan (annoter, acquérir, intégrer, raisonner) sont re-déployés et augmentés par l'IA : aucun axe n'épuise un verrou, et plusieurs projets sont délibérément transversaux.

« Des spectres aux molécules » — caractérisation et découverte moléculaires augmentées

Cet axe va du signal aux caractéristiques structurales, puis à la molécule. Il mobilise d'abord l'*acquisition* et l'*annotation* : orchestrer caractérisation expérimentale (chromatographie, spectrométrie de masse, résonance magnétique nucléaire), élucidation structurale et modélisation prédictive, avec auto-optimisation continue des instruments — vers des spectromètres de nouvelle génération qui adaptent leurs acquisitions en temps réel par des boucles de rétroaction IA-instrument (projets DeepMeta et apprentissage par renforcement). Il porte aussi la **découverte de produits naturels** : la candidature AI-FORGE (ANR-NRF, demande en cours) [22] s'attaque au « fossé génome-métabolome » — la majorité des clusters biosynthétiques microbiens demeurant orphelins, faute de lien établi entre génomes et métabolites [13] —, qu'elle traite comme un enjeu explicitement **multimodal** : faire converger spectrométrie de masse multiniveaux (du tandem à la mobilité ionique, IMS-CCS), résonance magnétique nucléaire, biologie de synthèse et graphes de connaissances dans des espaces de représentation partagés — l'*intégration* y étant un levier au service de l'élucidation moléculaire. Ce volet s'inscrit dans la continuité de ma thèse en chimie des produits naturels, dont il représente une évolution transdisciplinaire, mise en œuvre de manière collaborative : le projet *RAPID-NP* [25] (financement post-doctorat IdEx de Thaiz Rodrigues Teixeira, 2025–2027 ; collaboration INSERM, V. Libis) développe l'expression hétérologue de clusters biosynthétiques et la purification à haut débit de peptides microbiens grâce à l'isolement par affinité, fournissant le pendant expérimental « produits naturels » d'AI-FORGE — concrétisant un pipeline transdisciplinaire de nouvelle génération pour les produits naturels, d'abord appliqué aux microorganismes et ayant, à l'avenir, vocation à s'étendre à d'autres sources avec les progrès de la biologie de synthèse.

« Des atomes aux écosystèmes » — intégration multi-omique et interopérabilité

Cet axe déploie le verrou de l'*intégration* à l'échelle des systèmes vivants, du microbiome — fort de l'expérience de l'*Earth Microbiome Project* [7] — aux holobiontes. Il s'agit de construire des espaces de représentation unifiés et des modèles de fondation où convergent données multimodales, propriétés physico-chimiques et biologiques, modélisation et littérature scientifique : un continuum porté par le consortium MetaboLinkAI (ANR-FNS, financé) que je co-coordonne [21]. Une première démonstration à grande échelle est engagée avec l'*Earth Metabolome Initiative* : en collaboration avec P.-M. Allard (Université de Fribourg), l'analyse en MSn de plusieurs milliers d'extraits de plantes — et une collection d'extraits de champignons avec J.-L. Wolfender (Université de Genève) — éprouve le cadre MetaboLinkAI à l'échelle d'inventaires chimiques massifs et pan-écosystémiques. Il prolonge la vision formalisée dans la candidature ERC HOLOBIOMICS (non financée, mais structurante) [23] — observer le « holométabolome » des holobiontes (hôte et microbiome associé).

En collaboration avec É. Röttinger (IRCAN), le projet *METASTRESS* [24] en donne une incarnation : l'analyse couplée **transcriptomique-métabolomique** de la réponse au stress thermique et de la longévité chez le cnidaire *Nematostella vectensis*, en étendant le cadre MetaboLinkAI à l'intégration de deux omiques sur graphes de connaissances (thèse CNRS MITI de M. Féraud ; stage de M2 de D. Boukhajou, binôme transdisciplinaire financé par l'Académie 4, IdEx UniCA ; au sein

de l’Institut Fédératif de Recherche Ressources Marines, UniCA). Cette dynamique multi-omique se prolonge vers la santé avec le **LabCom Preciomics-AI** [27] — laboratoire commun CNRS / techincare (deeptech soutenue par BPI France) **en cours de préparation** (demande de financement LabCom ANR en cours ; pas encore effectif), sous direction scientifique conjointe (L.-F. Nothias, côté académique ; M. Benahmed, techincare) — dédié à l’intégration des approches multi-omiques (épigénomique, métabolomique, exposomique...) et de données hétérogènes (analyse de métaux lourds, anamnèse, facteurs environnementaux et de mode de vie, littérature scientifique), par l’IA, au service d’une médecine prédictive, préventive, personnalisée et participative (4P), avec l’Hôpital Lenval et le CHU de Nice comme partenaires cliniques.

« De l’expérience à la découverte » — des systèmes d’IA agentique ouverts pour les sciences

Concevoir des architectures d’IA multi-agents pour la co-production humain-IA des expériences et des connaissances, diffusées comme briques logicielles ouvertes. L’enjeu est de bâtir une infrastructure agentique **souveraine et flexible** — fondée sur des modèles et des outils ouverts, auditables et reproductibles — plutôt que de dépendre de boîtes noires propriétaires. À terme, ces agents doivent fermer la boucle « hypothèse → expérience → interprétation → nouvelle hypothèse », le chercheur demeurant décisionnaire à chaque étape — chaque raisonnement et chaque action du système restant transparents et auditables (*IA explicable*). Ce verrou conjugue *intégration* et *raisonnement* : graphes de connaissances, standardisation et ontologies de référence pour naviguer le continuum des données massives — spectres, littérature — produites en continu, et en extraire des hypothèses vérifiables. Dans le cadre de la vision de l’équipe eDIAM, mon groupe co-développe à cette fin un écosystème d’outils ouverts : *Perspicacité* [16] (exploration agentique et tracée de la littérature), *Mimosa-AI* (système auto-évolutif d’agents d’IA capables de réaliser des tâches scientifiques réelles) [17], *MetaboT* [18] / *MetaboKG* [19] (graphes de connaissances pour la métabolomique), et, en préparation, leurs déclinaisons — *AgenticScienceBuilder* (substrat à provenance vérifiable transformant des publications en tâches scientifiques reproductibles pour l’évaluation d’agents, et outils et compétences pour les systèmes agentiques) [17], Consilium (génération d’hypothèses et identification d’opportunités) et Studium [20] (enseignement assisté par l’IA). Une exigence transverse les relie : chaque décision reste **traçable, explicable et soumise au jugement humain** — l’IA augmente l’expert, elle ne le remplace pas.

Structuration, encadrement et coopération

Ce programme s’inscrit dans la structuration d’une nouvelle équipe (eDIAM), qui se déploie à la fois à l’Institut de Chimie de Nice et au Centre Inria d’Université Côte d’Azur (interface chimie × intelligence artificielle), et d’une politique d’encadrement qui encourage ses membres à l’autonomie, à l’auto-apprentissage et à l’usage d’outils d’IA générative — une pratique installée depuis 2023, dont nous avons développé un savoir-faire désormais bien établi. J’encadre actuellement deux doctorants — deux autres étant attendus en validation des acquis de l’expérience —, deux post-doctorants et un assistant ingénieur, et j’ai assuré le montage et la coordination de projets — financés (chaire CPJ CNRS, ANR-MetaboLinkAI, IDEX UniCA — Starting package, Académie 4, RAPID-NP —, ANR-POPS) comme candidatures structurantes (ERC HOLOBIOMICS ; ANR AI-FORGE et LabCom ANR Preciomics-AI, en cours). Il s’appuie sur des collaborations internationales (Suisse, États-Unis, Inde, Corée du Sud, Allemagne, Maroc) et sur l’environnement local (notamment la chaire affiliée 3iA Côte d’Azur). Mes valeurs — **science ouverte, interdisciplinarité et transdisciplinarité**, autonomie technologique (logiciels ouverts) et **humain au centre** — guident autant la recherche que la formation des jeunes chercheurs à la double compétence chimie / informatique appliquée. Dans un paysage technologique en évolution rapide, marqué par la compétition et la coopération internationales — y compris avec les grands acteurs du numérique, qui ne cachent pas leur ambition de dépasser les acteurs institutionnels —, la vision scientifique que je m’efforce d’appliquer et de faire adopter à mes collaborateurs est la nécessité d’une capacité d’adaptation continue : dans l’apprentissage comme dans l’agilité des choix stratégiques et le pragmatisme. Même si nous mobili-

sons aujourd’hui surtout des LLMs généralistes pour des prototypes et preuves de concept rapides, notre objectif est d’évaluer et de déployer des modèles plus frugaux (architectures simplifiées ou petits LLMs affinés) dès que possible — notamment pour la mise en production et l’adoption par la communauté.

5. Synthèse et perspectives

De la chimie des produits naturels à l’IA pour les sciences moléculaires, mon parcours a systématiquement cherché à lever le verrou suivant, sans attendre que le précédent soit pleinement résolu — certains se levant d’ailleurs plus efficacement lorsqu’on les aborde ensemble, simultanément ou en séquence, autour d’un horizon partagé. Ces expériences — scientifiques, mais aussi de montage de projets, de co-animation d’équipe et d’accompagnement de jeunes chercheurs vers une culture de laboratoire fondée sur l’entraide, l’autonomisation, l’auto-formation et l’interdisciplinarité — me conduisent aujourd’hui à solliciter l’Habilitation à Diriger des Recherches, afin d’encadrer pleinement la relève scientifique autour de cette vision : une métabolomique de nouvelle génération — intégrée aux autres omiques et connectée à l’instrument dès l’expérimentation —, augmentée par une IA ouverte et centrée sur l’humain, « de l’expérience à la découverte ».

Ce programme accompagne aussi une évolution du métier de chercheur que j’aborde comme à la fois une expérimentation personnelle et collective : apprendre à travailler avec des systèmes d’IA — jusque dans le pilotage et la coopération de l’équipe comme dans l’accélération de la réalisation des projets — tout en gardant l’humain et le jugement scientifique au centre. L’IA y ouvre aussi une opportunité singulière pour l’**interdisciplinarité**, en abaissant les barrières techniques entre domaines et en facilitant le dialogue entre chimie, biologie et science des données. J’y vois enfin le rôle de la coopération et des échanges internationaux, non seulement comme catalyseurs d’innovation, mais comme ferment d’un socle commun et durable, à l’heure où nos sociétés affrontent une incertitude croissante et des défis sans frontières.

Références

Travaux de l’auteur (sélection) :

- [1] Nothias-Scaglia L.-F. *et al.* Antiviral activity of diterpene esters on chikungunya virus and HIV replication. *J. Nat. Prod.* 78, 1277–1283 (2015).
- [2] Nothias L.-F. *et al.* Feature-based molecular networking in the GNPS analysis environment. *Nat. Methods* 17, 905–908 (2020).
- [3] Dührkop K., Nothias L.-F. *et al.* Systematic classification of unknown metabolites using high-resolution fragmentation mass spectra. *Nat. Biotechnol.* 39, 462–471 (2021).
- [4] Hoffmann M.A., Nothias L.-F. *et al.* High-confidence structural annotation of metabolites absent from spectral libraries. *Nat. Biotechnol.* 40, 411–421 (2022).
- [5] Zuo Z., Cao L., Nothias L.-F.*, Mohimani H.* MS2Planner : improved fragmentation spectra coverage in untargeted mass spectrometry by iterative optimized data acquisition. *Bioinformatics* 37, i231–i236 (2021).
- [6] Schmid R. *et al.* (incl. Nothias L.-F.) Ion identity molecular networking for mass spectrometry-based metabolomics. *Nat. Commun.* 12, 3832 (2021).
- [7] Shaffer J.P., Nothias L.-F. *et al.* Standardized multi-omics of Earth’s microbiomes reveals microbial and metabolite diversity. *Nat. Microbiol.* 7, 2128–2150 (2022).
- [8] Cao L. *et al.* (incl. Nothias L.-F.) MetaMiner : a scalable peptidogenomics approach for discovery of ribosomal peptide products with blind modifications. *Cell Syst.* 9, 600–608 (2019).
- [9] Kunyavskaya O. *et al.* (incl. Nothias L.-F.) Nerpa : a tool for discovering biosynthetic gene clusters of bacterial nonribosomal peptides. *Metabolites* 11, 693 (2021).
- [10] Zhang S. *et al.* (incl. Nothias L.-F.) Comparative metabolomic analysis reveals shared and unique chemical interactions in sponge holobionts. *Microbiome* 10, 22 (2022).
- [11] Wegley Kelly L. *et al.* (incl. Nothias L.-F.) Distinguishing the molecular diversity, nutrient content and energetic potential of exometabolomes produced by macroalgae and reef-building corals. *PNAS* 119, e2110283119 (2022).
- [12] Quinlan Z.A. *et al.* (incl. Nothias L.-F.) Microbial community metabolism of coral reef exometabolomes. *Environ. Microbiol.* 27, e70064 (2025).

- [13] Schorn M.A. *et al.* (incl. Nothias L.-F.) A community resource for paired genomic and metabolomic data mining. *Nat. Chem. Biol.* 17, 363–368 (2021).
- [14] Koester I., Quinlan Z.A., Nothias L.-F. *et al.* Illuminating the dark metabolome of *Pseudo-nitzschia*–microbiome associations. *Environ. Microbiol.* 24, 5408–5424 (2022).
- [15] Wolfender J.-L. *et al.* (incl. Nothias L.-F.) Metabolomics in ecology and bioactive natural products discovery : challenges and prospects for a comprehensive study of the specialised metabolome. *Chimia* 76, 954–963 (2022).

Projets et production en préparation :

- [16] Perspicacité-AI — cadre ouvert d’exploration agentique et tracée de la littérature (manuscrit en préparation).
- [17] AgenticScienceBuilder / Mimosa-AI — substrat à provenance vérifiable pour l’évaluation d’agents sur des tâches scientifiques (Mimosa-AI : Legrand *et al.*, *arXiv* 2603.28986, 2026, préprint ; AgenticScienceBuilder : manuscrit en préparation).
- [18] MetaboT — agents conversationnels (LLM multi-agents) pour l’interrogation en langage naturel de graphes de connaissances métabolomiques (Bekbergenova *et al.*, *Research Square*, 2025, préprint ; doi :10.21203/rs.3.rs-6591884/v1).
- [19] MetaboKG — cadre de graphe de connaissances analyse-centré pour la métabolomique non ciblée (Feraud *et al.*, *arXiv* 2605.24706, 2026).
- [20] Studium — cadre d’apprentissage agentique (« Learning Workflow Graph ») (manuscrit en préparation).
- [21] MetaboLinkAI — projet ANR–FNS bilatéral (financé), www.metabolinkai.net.
- [22] AI-FORGE — candidature ANR–NRF (en cours) : élucidation par IA des clusters biosynthétiques orphelins.
- [23] HOLOBIOMICS — candidature ERC Starting Grant 2023 (non financée).
- [24] METASTRESS — projet IdEx UniCA (É. Röttinger – L.-F. Nothias) : multi-omique du stress chez les cnidaires.
- [25] RAPID-NP — post-doctorat IdEx (T. Rodrigues Teixeira ; collab. INSERM) : expression hétérologue de BGC, peptides marins.
- [26] ANR-POPS — projet ANR (partenaire) : métabolites microbiens.
- [27] Preciomics-AI — LabCom CNRS / techincare (en préparation) : multi-omique & IA pour la médecine de précision (4P).

Bibliothèques spectrales (contributions et ressources) :

- [28] Brungs C. *et al.* (incl. Nothias L.-F.) MSnLib : efficient generation of open multi-stage fragmentation mass spectral libraries. *Nat. Methods* (2025).
- [29] NIH Natural Products Library — bibliothèque spectrale publique, diffusée via la plateforme GNPS. <https://gnps.ucsd.edu>.
- [30] Bittremieux W. *et al.* (incl. Nothias L.-F.) Open access repository-scale propagated nearest neighbor suspect spectral library for untargeted metabolomics. *Nat. Commun.* (2023).

Références complémentaires :

- [31] Nothias L.-F. *et al.* Bioactivity-based molecular networking for the discovery of drug leads in natural product bioassay-guided fractionation. *J. Nat. Prod.* (2018).
- [32] Ernst M., Nothias L.-F. *et al.* Assessing specialized metabolite diversity in the cosmopolitan plant genus *Euphorbia* L. *Front. Plant Sci.* 10, 846 (2019).
- [33] Broeckling C.D. *et al.* Comprehensive tandem-mass-spectrometry coverage of complex samples enabled by data-set-dependent acquisition. *Anal. Chem.* 90, 8020–8027 (2018).

Contexte (littérature) :

- [34] Gil Y., Greaves M., Hendler J., Hirsh H. Amplify scientific discovery with artificial intelligence. *Science* 346, 171–172 (2014).
- [35] Wei J. *et al.* From AI for science to agentic science : a survey on autonomous scientific discovery. *arXiv* 2508.14111 (2025, préprint).
- [36] Wandy J. *et al.* ViMMS 2.0 : a framework for fragmentation strategies in LC-MS metabolomics. *J. Open Source Softw.* 7(71), 3990 (2022).

Fil scientifique — Louis-Félix Nothias

« Lever les verrous du métabolome » → augmenter les sciences moléculaires par l'intégration et l'IA | des spectres aux holobiontes (équipe eDIAM)

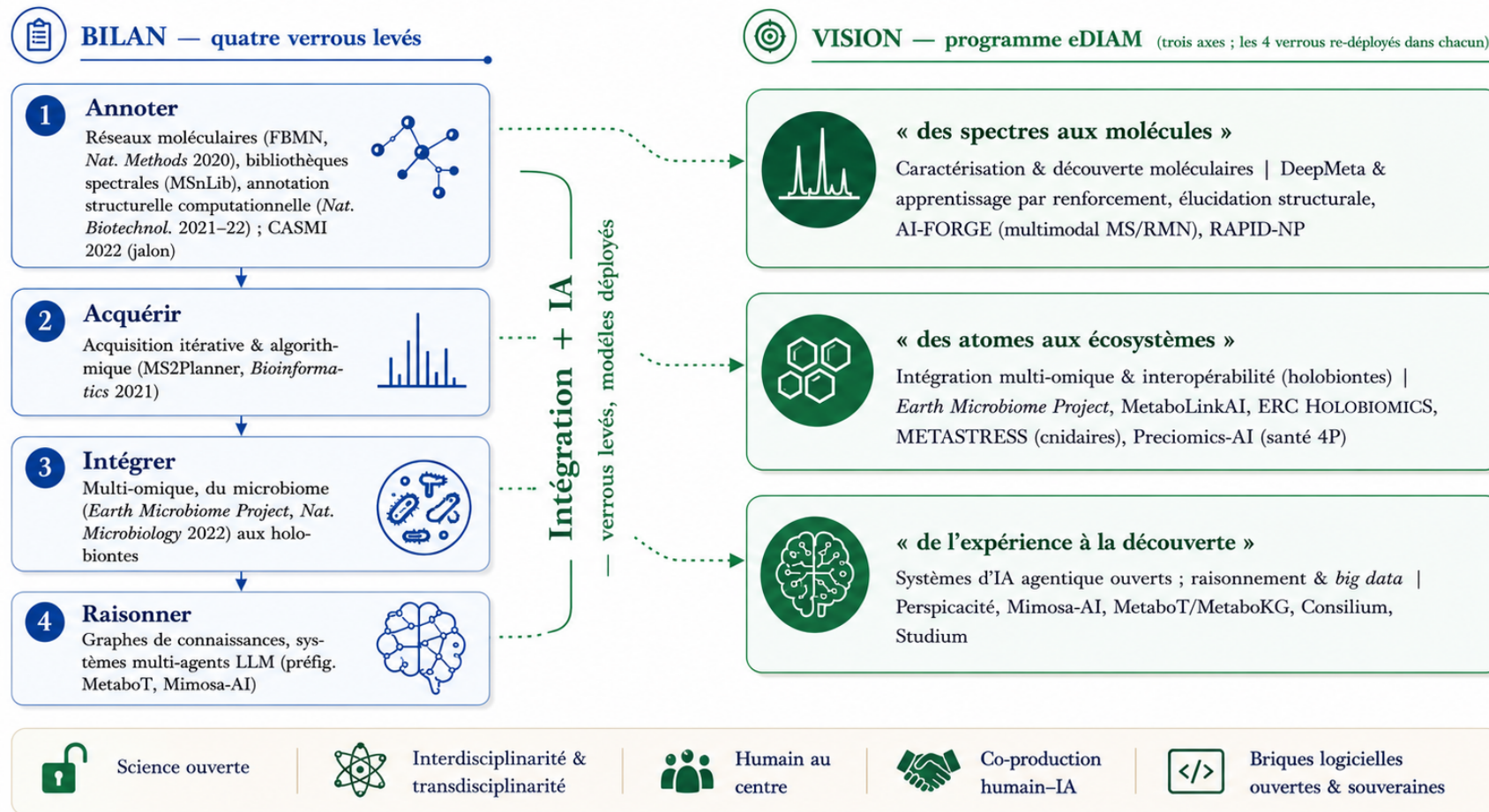


Figure — Synthèse graphique du fil scientifique : du bilan (« lever les verrous du métabolome ») au programme eDIAM (trois axes), les quatre verrous re-déployés et augmentés par l'intégration et l'IA.