Nucleic Acids Research, 2024, 1–9
https://doi.org/10.1093/nar/gkae1093
Database issue

The Natural Products Atlas 3.0: extending the database of microbially derived natural products

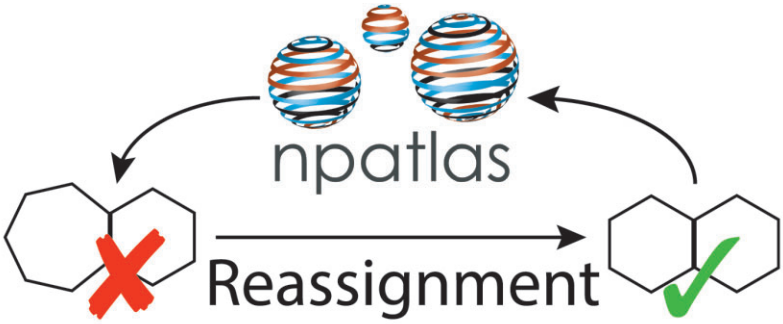
Ella F. Poynton[1], Jeffrey A. van Santen[1], Matthew Pin[1], Marla Macias Contreras[2], Emily McMann[1], Jonathan Parra[3,4], Brandon Showalter[1], Liana Zaroubi[1], Katherine R. Duncan[5] and Roger G. Linington[1,*]

[1]Department of Chemistry, Simon Fraser University, 8888 University Drive, Burnaby, British Columbia, V5A 1S6, Canada
[2]Institute of Pharmacy and Biomedical Sciences, University of Strathclyde, 161 Cathedral St, Glasgow G4 0RE, UK
[3]Centro de Investigaciones en Productos Naturales (CIPRONA), Universidad de Costa Rica, San José 11501-2060, Costa Rica
[4]Centro Nacional de Innovaciones Biotecnológicas (CENIBiot), CeNAT-CONARE, San José 1174-1200, Costa Rica
[5]Biosciences Institute, Catherine Cookson Building, Newcastle University, Newcastle upon Tyne NE2 4HH, UK
[*]To whom correspondence should be addressed. Tel: +1 778 7823517; Fax: +1 778 7823765; Email: rliningt@sfu.ca

Abstract

The Natural Products Atlas is a database of microbially derived natural products that contains structures, producing organism taxonomy, biosynthetic and chemical ontology classifications, grouping by compound classes and cross-links to a suite of other natural product-related data resources. The database is supported by a web server that includes functionality to browse the collection, search the database using both chemical structures and text/numerical terms and visualize the chemical diversity it contains using interactive dashboards. In the current database release, we have curated 1347 papers, increasing the number of compounds to 36 545. In addition, we have initiated a large-scale effort to incorporate data from papers reporting structural reassignments and revisions to previously published structures. This effort led to the incorporation of 590 corrections to existing entries, significantly improving the accuracy of the dataset. The Natural Products Atlas may be accessed at www.npatlas.org.

Graphical abstract



Introduction

The chemistry of the microbial world has a profound effect on many areas of our lives. From the influence of the human microbiome on health and wellbeing, to the protection or destruction of crops and food stocks by environmental strains, microbial chemistry contributes substantially to processes that shape society. Notwithstanding the importance of this topic, until recently there was no open access repository containing the structures of all known microbial metabolites. The Natural Products Atlas was developed to address this gap, with the goal of collecting, curating and distributing information on all known microbial secondary metabolites.

The Natural Products Atlas was originally published in late 2019 (1), with a second paper describing extensions and improvements to the platform published in 2022 (2). These two papers have now been cited 352 times (Web of Science, September 2024). Among these, approximately a quarter of papers focus on the discovery of individual classes of natural products (NPs; NP discovery), a quarter are review papers on different elements of NP discovery or databases and half

Received: September 13, 2024. **Revised:** October 18, 2024. **Editorial Decision:** October 18, 2024. **Accepted:** October 24, 2024

describe large-scale examination of NPs' chemical space (Figure 1). These topics range from new database construction to omics data integration and encompass a wide range of applications that would not have been possible without the availability of the Natural Products Atlas as a central repository of known microbial NPs' structures.

In the previous Atlas release, we completed a thorough review of the original publications for all structures in the database missing assignments of one or more chiral centers. We performed this reanalysis to ensure that the structure representations in the database contained all information from the original papers, rather than simplified structural depictions. During this effort, it became clear that an appreciable number of NPs are published without full configurational assignment. These structures are sometimes later updated through total synthesis or reisolation, defining the configurations at these unknown positions. In addition, a smaller number of cases involve the correction of structures that were originally incorrectly assigned, changing either configurations at one or more chiral centers, or the gross connectivity of the carbon skeleton.

Because the Natural Products Atlas reports the original isolation paper for each compound, reassignments to structures were not tracked in previous versions of the database. To address this shortcoming, we have initiated a program to find and curate all papers reporting corrections or updates to existing entries. This paper describes the methodology and outcome from this effort, which has led to the correction of hundreds of structures and significantly improved the accuracy of the database. In addition, we describe the curation of the most recent set of papers reporting new microbial NPs and discuss current discovery trends revealed from these data.

Materials and methods

Curation of newly reported NPs

Curation of new papers (defined below) reporting microbial NP discovery followed our established curation pipeline. Briefly, titles and abstracts from a set of 50 journals known to contain NP papers were classified using an in-house support vector machine model to identify papers describing new NP isolation. A named entity recognition tool was then used to extract compound names and source organism taxonomies from the title/abstract text, and relationships between organisms and compounds were defined using a relationship model. Compound names were searched against PubChem to acquire chemical structures where available. Finally, article data were formatted for our data curation web server and assigned to curators for manual review and database insertion.

Curation of reassignment data

Preparation of literature training set

To create a training set for classifier development for articles reporting reassignment of NP structures we collected and manually inspected a large body of article titles and abstracts (~5000) from the organic chemistry literature. Because reports of reassignment are rare, we enriched the set by including results from targeted keyword searches and articles containing the names of microbial NPs. We then manually reviewed the title and abstract of each article, classifying them as either about reassignment (1) or not (0).

Creation of random forest classifier

Using the literature training set described above (1025 articles about reassignment and a balanced number of 'negative' articles that did not report reassignments) we created a random forest classifier to identify articles reporting structure reassignments for NPs. Following hyperparameter tuning and parameter optimization, the model returned an accuracy of 85.4%, recall of 80.6% and precision of 87.7% using an 80:20 train test split that resulted in a holdout set of 410 articles (204 = 1, 206 = 0). An attempt was made to introduce class imbalance in favor of positive results to improve the recall of the model. However, this led to large increases in the false positive rate, which was not practical for downstream manual curation, so no biased weighting was used in the final model.

Classification of literature articles

Titles and abstracts for 268 824 articles derived from 10 chemistry journals from the period 2000–2023 were collected from the NCBI E-utilities resource and classified using the reassignment classifier, returning 8468 candidate articles. Because the model was not trained to differentiate between microbially derived and plant/invertebrate-derived compounds, these articles were scanned for compound names and related terms from the Natural Products Atlas database. Articles related to microbial NPs were retained, yielding 4418 candidate articles for manual review.

Curation of reassignment data

Candidate reassignment articles were combined with lists of articles derived from targeted keyword searches, reformatted into a flat comma-separated values (csv) format that included the Natural Product Atlas Identification (NPAID) and Simplified Molecular Input Line Entry System (SMILES) string where available, and divided into datasets for manual review. Curators examined each row of the dataset to decide whether the article was about NP reassignment and whether the compound being reassigned was of microbial origin. Compounds meeting these criteria were then updated in the table by adding the corrected structure and the type of reassignment. Completed datasets were compiled into a single file, reviewed for accuracy and removal of duplicates and sent to the developer team for database insertion. In some cases, reassigned compounds were missing from the Atlas, or compounds were reassigned to previously reported structures that were not in the Atlas database. In both cases, the original isolation article for the missing compound was found, curated and inserted into the Atlas using our standard article curation pipeline to ensure the inclusion of all corrected structures.

Incorporation of additions and corrections

User deposited articles were exported from the backend server, standardized using in-house literature standardization tools and reformatted for insertion into the Natural Products Atlas curation web server. Each article was then manually reviewed in the curation platform using our standard curation workflow and inserted into the database using a comprehensive set of checks for standardization and uniqueness against the existing database entries.

User-deposited corrections were exported from the backend server, reformatted to csv format and manually reviewed for accuracy. Manual review involved reexamination of the original paper(s) reporting new compound discovery, and com-

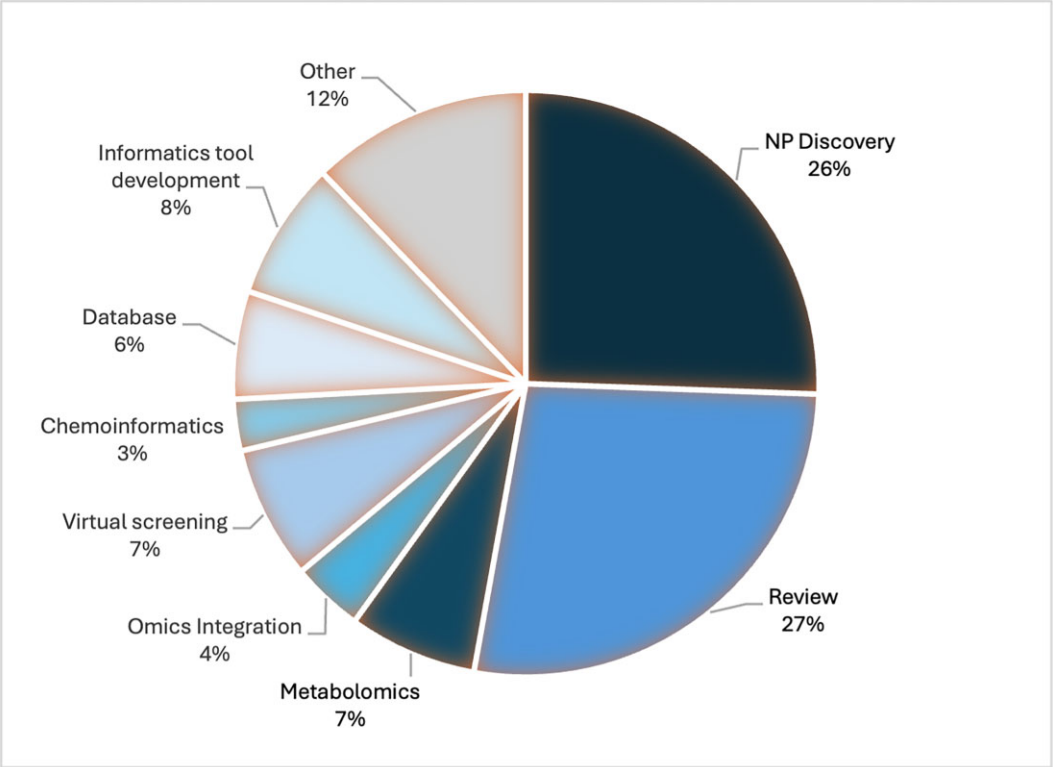


Figure 1. Distribution of subject areas covered by papers citing the Natural Products Atlas (2019–2024).

parison with the proposed correction to confirm the accuracy of the proposed change. In some cases (e.g. addition of synonyms), the correction involved insertion of new articles into the database and the creation of a new entry in the synonym table, linked to the compound page. In others (e.g. correction to source organism taxonomy), the correction included insertion of new taxonomic terms (if not already present in the database) and updating of links between compound and taxonomy tables. Corrections that passed manual review were subjected to checks for standardization and uniqueness and updated directly in the database using the administrator Application Programming Interface as described above.

Results

Curation of newly reported NPs

The previous publication describing the Natural Products Atlas included articles published up to the end of 2020. In this new release, we have extended the coverage up to the end of 2022, reviewed additional papers from the period prior to 2021 and added selected coverage of papers from 2023/24. In total, this required the inspection and manual curation of 1347 research articles using the Natural Products Atlas curation platform. Of these, 1308 articles were added to the database, with an average of 3.05 new compounds per article. This curation effort increased the size of the database by 3993 compounds to a total of 36 545. The ratio of fungal to bacterial compounds in 2021/22 was 4.4:1. This is significantly higher than the previous ratio for the Atlas database (2.4:1 for 2019/20) and follows a recent trend toward new compound discovery from fungi.

Curation of user-submitted articles and corrections

Since the inception of the Natural Products Atlas project, the website has included functionality for the deposition of articles missing from the database (i.e. the addition of compounds not present in the database; https://www.npatlas.org/deposit), and the reporting of additions or corrections to the existing data (i.e. updating of information for compounds currently held in the database; https://www.npatlas.org/corrections). To date, community use of these functions has been low. Since 2019, we have received depositions of 237 articles, more than half of which were from the Atlas development team. Similarly, we have received 155 additions/corrections, many of which were also from the development team. Nevertheless, these community contributions are very important for improving the accuracy of the database as they are often submitted by domain-specific experts who find subtle errors that would be very difficult for the Atlas team to detect using automated tools. Additional submissions from the community are therefore strongly encouraged. For this release, we reviewed and incorporated all new data from both sources as described in the 'Materials and methods' section. This included updates to incorrect structures (27), addition of synonyms (46) and various corrections to taxonomy terms or isolation references.

Addition of structural reassignment information

Identification of articles reporting structure reassignments

The inclusion of reassignment data is a new feature for this release of the Natural Products Atlas. In this context, reassignment articles are defined as articles that updated or correct the structures of previously published NPs. To identify articles covering these topics, we pursued a two-pronged approach. First, we performed classical literature searches using keyword

terms including 'structure reassignment', 'structural revision', etc. Second, we manually classified ∼2000 articles from the synthetic products/NPs literature into two groups (true/false) based on whether they included information about structure reassignments and used this training set to create a random forest classifier to identify reassignment articles from the primary literature. Classifying ∼270 000 articles from PubMed (see 'Materials and methods' section) returned 8468 candidate articles about structure reassignments. After the removal of duplicates, the title and abstract from each article were searched against lists of compound names and genera from the Atlas database to identify reassignment articles relevant to microbial NPs. Finally, article lists (4418 articles total) were provided to curators for review.

Curation of reassignment data

Curators manually reviewed each article to identify the compound(s) being reassigned, identified those that were microbial NPs and manually extracted the NPAID (if available) or name and structure (if no NPAID available) of the original structure, the SMILES representation of the reassigned structure and the type of reassignment. In total, curators reviewed 4418 articles, of which 411 contained reassignments of microbial NPs. Together these results identified 157 compounds not previously included in the Atlas database, 494 compounds that required structure reassignment and 81 compounds previously included in the Atlas that required removal (see 'Discussion' section).

Data management for structure reassignments

Addition of reassignment data to the Atlas necessitated changes to the underlying database schema. The central identifier in the database is the NPAID (e.g. NPA012345). This is defined in a 'compound' table that is at the heart of the schema. Until now, this table has also included information about the structure of the compound associated with this NPAID. However, with the inclusion of reassignment data, the structural representation is no longer fixed for the lifetime of each entry. Therefore, compound structure was removed to a separate table, allowing both the currently accepted structure and any previous (incorrect) structures to be associated with the same NPAID. These two tables are connected by a linker table that lists the substance ID (NPAID), the compound structure ID and the reference ID for the reference reporting that structure. In this way, information is preserved about the provenance of each version of the structure, in line with our goal of associating every data point with a primary reference from the scientific literature.

Incorporating reassignment information into data display and search functions

Addition of reassignment data also created new considerations for search options and information display. For example, should incorrect structures be included in standard search results? On the one hand, the answer should be no, since there is no currently available data supporting their presence in nature. On the other hand, the answer should be yes, since these structures are present in the scientific literature but are now known to be incorrect; information that could be of significant value to researchers studying these compounds. To resolve this issue, reassigned structures are excluded from search results but are available in the JSON version of the download options.

A second consideration concerns substances whose structures have been disproven. In rare cases, published structures are reassigned to existing NPs (i.e. their structures were initially incorrectly assigned as new compounds, and subsequently reassigned to a previously published structure). In these cases, the NPAID of the misassigned compound becomes deprecated, since structures must be unique in the database. This raises the question of how to display the information on the incorrect structure. Removing these data may lead users to conclude that the molecule is legitimate but that the entry is missing from the Natural Products Atlas. However, retaining it suggests that the structure is legitimate. To resolve this issue, we added a banner to the compound page for each deprecated structure, stating that the structure had been reassigned to a different entry in the compound pages and providing a hyperlink to the currently accepted compound page (see 'Discussion' section). Deprecated compounds are excluded from search results unless the Advanced Search options are used. To access deprecated structures, include the search term 'has_exclusions' == TRUE in the Advanced Search page. Deprecated structures are excluded from the download files, since these files are often used as training sets for Artificial Intelligence (AI)/Machine Learning (ML) development (3–5) as well as classical dereplication (6–8). However, the deprecated structures are available as a separate file in the Zenodo repository (see 'Data availability' section).

Addition of reassignment data has created a new layer of information that must be displayed to users. Compounds may have one or more reassignment articles associated with them, and consequently several different structure representations over their discovery lifetimes. To present this to users, compounds with reassignments now include a reassignment table on the compound page that lists the original isolation structure and reference, as well as a chronological list of reassignments with their respective structures. The structure displayed as the main image on the compound page (and listed in the download file) is the currently accepted version of the structure. As with all data in the Natural Products Atlas, this provides a direct link between each data point and the primary reference from which it was derived.

Because reassignments sometimes lead to compounds being corrected to previously published structures we have also implemented synonym handling. Unlike most chemical databases, these synonyms are also all tied to references from the primary literature. Where available, synonyms are now displayed as a separate table below the main data table on each compound page. Both accepted compound names and synonyms are included in name searches. All synonyms are also included in the download files.

Interoperability and cross-linking to other NPs resources

One useful attribute of the Natural Products Atlas is that it maintains cross-links to other domain-specific resources in NPs research. These include the MIBiG database of biosynthetic gene clusters (https://mibig.secondarymetabolites.org) (9), the GNPS repository of mass spectrometry (MS) data (https://gnps.ucsd.edu) (10) and the NP-MRD database of nuclear magnetic resonance (NMR) data (https://np-mrd.org) (11,12). In the latest release, we have updated the cross-links to all three of these repositories, capturing both new data in those repositories that links to existing Natural Products At-

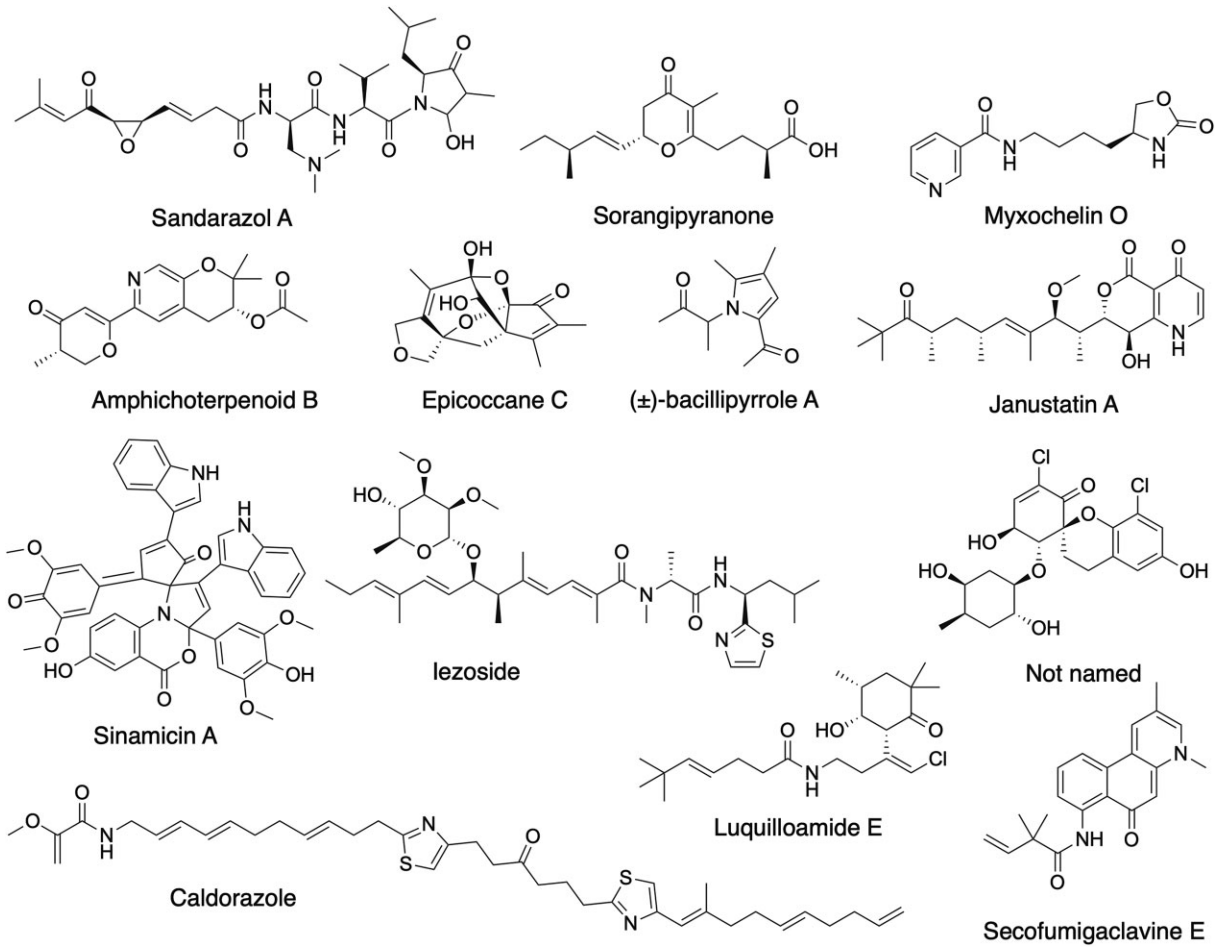


Figure 2. Examples of microbial NP structures (2021–2022) with low structure similarity scores to existing entries in the Natural Products Atlas (13–24).

las entries, and links to newly added compounds in the Atlas database. In addition, we have added cross-links to the Collaborative Microbial Metabolite Center (https://cmmc.gnps2.org) knowledgebase, which is a new data ecosystem for microbial metabolites in human microbiome datasets.

A complication of adding reassignment data is that this will reduce the number of cross-links to other resources unless those resources also update their representations of corrected structures. However, given that experimental data such as MS and NMR spectra contain spectroscopic properties of chemical matter that are not impacted by the appended structure representation it is still useful to create these cross-links, even if the data repository holds an outdated representation of the structure. To resolve this issue, we search all structural variations for each compound and include the cross-link to the most recent structure for which a match is found.

Discussion

New curations

The rate of new microbial NP discovery remains high, with 1437 new compounds reported in 2021 and 1516 in 2022. Among the compounds reported in 2021/22, 71 were from genera not previously included in the Natural Products Atlas, suggesting that exploration of new areas of taxonomic space remains an active area of investigation. In terms of chemical novelty, performing structure similarity comparison for each compound reported in 2021/22 against all compounds in the previous release of the database (Morgan fingerprinting radius = 2, Dice similarity scoring) revealed 47 compounds with maximum similarity scores ≤ 0.5. This demonstrates that unusual scaffold discovery continues to be a regular element of NPs science, and that the natural world remains an important resource for architecturally complex and unusual molecules (Figure 2). Although the ratio of new compounds reported from fungal and bacterial sources in 2021/22 was 4.4:1, the ratio for compounds with low similarity scores was ~1:1 (24 bacterial and 23 fungal). Care must be taken with inferring the relative 'novelty' of compounds from both sources because the overall number of compounds is low, and includes several cases where multiple compounds were reported from a single family, influencing this ratio.

Reassignments

Considerations for incorporating reassignment data. A major advancement for this release of the Atlas is the addition of reassignment data. Reassignment data are defined as data published after the initial report of the NP that update or correct the original structure. These range from the assignment of one or more chiral centers whose configurations were not originally determined, to revision of the gross carbon skeleton of the structure. This is distinct from the correction of database errors introduced during curation/ingestion (e.g.

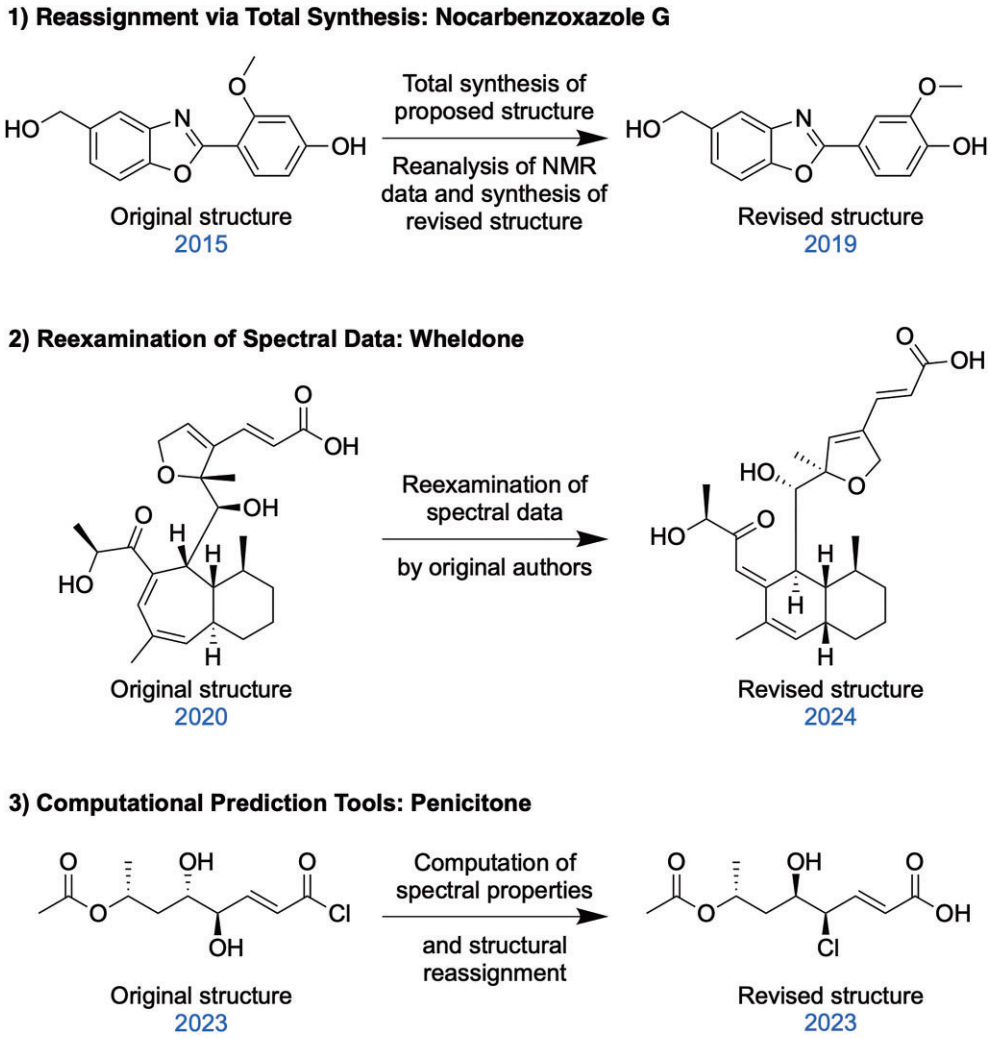


Figure 3. Examples of sources of structure reassignment (28–33).

misrepresentation of structures that were correctly reported in the literature or misspelling of compound names or taxonomic terms).

Maintaining an accurate and comprehensive record of scientific knowledge is the central objective of all scientific databases. Yet the pressures of daily publication of new information and periodic updates and corrections to existing knowledge make this difficult to sustain. Many academic database projects are never updated after their initial release (25). Others engage in periodic curation efforts (26) or adopt a 'Wiki' style curation model that federates curation to the scientific community (27). However, most of these efforts prioritize the addition of new information, rather than updating or correcting previous entries.

Because reassignments are typically reported as part of larger studies, it can be difficult to find this information unless it is explicitly stated in the abstract text. This is in part because most journal copyright rules preclude the large-scale downloading and automatic data extraction from full article text. In addition, discussions of structure reassignments can often be complicated, with a wide range of outcomes, making these data slow to curate, and complicated to manage. Together these issues help to explain why rates of curation of reassignment data are currently low.

Reassignments are typically identified via one of three routes: (i) total synthesis that reveals inconsistencies between the spectral data for the original molecule and the synthetic material of the proposed structure; (ii) re-isolation of the compound and re-examination of the spectral data or (iii) re-examination of the original data using computational spectral prediction tools (Figure 3). In all cases, the outcome is that new information is made available that updates our knowledge about the structure of the reported molecule.

In general, reassignments fall into one of three categories: (i) The compound is unique, but the originally reported structure requires revision. (2) The compound was originally reported as unique but is the rediscovery of a known compound whose spectral data were misinterpreted. (3) The compound does not have the reported structure, but the correct structure cannot be determined.

Each of these scenarios requires a different action in the Natural Products Atlas database. For scenario 1, a new representation of the structure is added to the database, and this is assigned as the current structure for that compound and displayed in the compound page. The original structure remains associated with the original isolation paper and is available in the download JSON file (see 'Results' section). For scenario 2, the NPAID for the incorrect structure is deprecated (since

this molecule never existed) and the assigned name is added as a synonym to the NPAID for the correct structure. The reassignment paper is used as the provenance reference for this synonym so that users can easily determine why the incorrect name is associated with the correct structure. In scenario 3, the NPAID is also deprecated, and the structure removed from default searches and download files. In cases where NPAIDs are deprecated, banners are added to compound pages stating that the compound has been removed and explaining why this deprecation occurred.

In total, curation of 4418 articles led to the inclusion of 590 structure reassignments including the removal of 81 incorrect compounds. Interestingly, this effort identified a small number of instances where multiple research groups reported synthesis of the same structure but reached different conclusions about the correct structure representation. For example, stevastelin C3 was originally reported from a *Penicillium* fungus in 1994 (34). A total synthesis was reported in 2002 but unusually, a footnote in this paper indicated that the original spectral data were not available and therefore the synthetic material had not been compared with the original NP (35). A subsequent total synthesis in 2004 (36) was able to acquire spectral data for the original compound, revealing a mismatch in the structures. Reanalysis of the original data followed by total synthesis of the proposed structure led to a revision of the structure of the NP.

Separately, we encountered one case where a NP (medermycin/lactoquinomycin A) (37,38) was reassigned on the basis of chemical and spectral data (39), but the original structure was reinstated by a rebuttal paper from a third research group six months later (40). These complex cases highlight the intricacies of curating reassignment data. With the move toward AI/ML-enabled extraction of information from the scientific literature, it is tempting to suggest that data curation could become a largely automated process. However, cases like these highlight the challenges associated with such an approach and suggest that manual review will remain an essential element of high-quality data curation for the foreseeable future.

It is important to note that the effort to curate all structure reassignments in microbial NP chemistry is not yet complete. It is not always straightforward to identify papers reporting structure reassignments, particularly if this information is only discussed in passing in the main body of the text. It is also possible that the database contains compounds that require revision, but that these structures have yet to be subjected to rigorous reanalysis by the community. Advances in computational prediction of spectral data are expected to simplify the challenge of identifying structures that require such reanalysis. However, this would be made easier if the training sets used to build these models were accurate, which is not possible if data from the primary literature contains high error rates, creating a 'Catch-22′ scenario. Progress toward highly accurate spectral prediction for NPs is unlikely to make significant headway until reassignment data are more consistently incorporated into existing databases and training sets.

License terms

Conclusion

The latest release of the Natural Products Atlas adds new compounds published since the previous release, updates cross-links and user-submitted corrections and additions and now includes information on structural reassignments, improving the accuracy of information on compounds that have been subjected to structural reevaluation.

Data availability

The data underlying this article are available in the Natural Products Atlas at www.npatlas.org. Documentation for the API is available at https://www.npatlas.org/api/v1/docs. Previous releases of the Natural Products Atlas database are available in Zenodo at https://dx.doi.org/10.5281/zenodo.3530792.

Acknowledgements

We thank Darren Scobie for contributions to initial data curation.

Author contributions: E.F.P. (Conceptualization, Investigation, Methodology, Data curation, Formal analysis); J.A.vS. (Conceptualization, Investigation, Methodology, Data curation, Software, Validation, Visualization); M.P. (Investigation, Resources); M.M.C. (Data curation); E.M. (Data curation); J.P. (Data curation); B.S. (Data curation); L.Z. (Data curation); K.R.D. (Data curation); R.G.L. (Conceptualization, Data curation, Formal analysis, Funding acquisition, Project administration, Supervision, Writing – Original draft, Writing – review & editing).

Funding

Natural Sciences and Engineering Research Council of Canada [to R.G.L.]; Department for Environment, Food and Rural Affairs, UK Government (DEFRA) [to K.R.D.]; UK Research and Innovation [EP/X03142X/1 to K.R.D.]; HORIZON EUROPE Marie Skłodowska-Curie Actions [101072485 to K.R.D.]. Funding for open access charge: Natural Sciences and Engineering Research Council of Canada.

Conflict of interest statement

J.A.vS. is an employee of Unnatural Products. The other authors declare no conflicts of interest.

References

1. van Santen,J.A., Jacob,G., Singh,A.L., Aniebok,V., Balunas,M.J., Bunsko,D., Neto,F.C., Castaño-Espriu,L., Chang,C., Clark,T.N., *et al.* (2019) The Natural Products Atlas: an open access knowledge base for microbial natural products discovery. *ACS Cent. Sci.*, **5**, 1824–1833.
2. van Santen,J.A., Poynton,E.F., Iskakova,D., McMann,E., Alsup,T.A., Clark,T.N., Fergusson,C.H., Fewer,D.P., Hughes,A.H., McCadden,C.A., *et al.* (2022) The Natural Products Atlas 2.0: a database of microbially-derived natural products. *Nucleic Acids Res.*, **50**, D1317–D1323.
3. Reher,R., Kim,H.W., Zhang,C., Mao,H.H., Wang,M., Nothias,L.F., Caraballo-Rodriguez,A.M., Glukhov,E., Teke,B., Leao,T., *et al.* (2020) A convolutional neural network-based approach for the rapid annotation of molecularly diverse natural products. *J. Am. Chem. Soc.*, **142**, 4114–4120.
4. Ricart,E., Pupin,M., Müller,M. and Lisacek,F. (2020) Automatic annotation and dereplication of tandem mass spectra of peptidic natural products. *Anal. Chem.*, **92**, 15862–15871.
5. Margulis,E., Dagan-Wiener,A., Ives,R.S., Jaffari,S., Siems,K. and Niv,M.Y. (2021) Intense bitterness of molecules: machine learning for expediting drug discovery. *Comput. Struct. Biotechnol. J.*, **19**, 568–576.
6. Maw,Z.A., Grunwald,A.L., Haltli,B.A., Cartmell,C. and Kerr,R.G. (2024) Discovery of the lipopeptides albubactins A–H from *Streptomyces albidoflavus* RKJM0023 via chemical elicitation with rhamnolipids and synthesis of albubactin A. *J. Nat. Prod.*, **87**, 1682–1693.
7. Kyei,L., Piedl,K., Menegatti,C., Miller,E.M. and Mevers,E. (2024) Discovery of biofilm inhibitors from the microbiota of marine egg masses. *J. Nat. Prod.*, **87**, 1635–1642.
8. Morehouse,N.J., Flewelling,A.J., Liu,D.Y., Cavanagh,H., Linington,R.G., Johnson,J.A. and Gray,C.A. (2023) Tolypocaibols: antibacterial lipopeptaibols from a *Tolypocladium* sp. Endophyte of the marine Macroalga *Spongomorpha arcta*. *J. Nat. Prod.*, **86**, 1529–1535.
9. Terlouw,B.R., Blin,K., Navarro-Muñoz,J.C., Avalon,N.E., Chevrette,M.G., Egbert,S., Lee,S., Meijer,D., Recchia,M.J.J., Reitz,Z.L., *et al.* (2023) MIBiG 3.0: a community-driven effort to annotate experimentally validated biosynthetic gene clusters. *Nucleic Acids Res.*, **51**, D603–D610.
10. Wang,M., Carver,J.J., Phelan,V.V., Sanchez,L.M., Garg,N., Peng,Y., Nguyen,D.D., Watrous,J., Kapono,C.A., Luzzatto-Knaan,T., *et al.* (2016) Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.*, **34**, 828–837.
11. Wishart,D.S., Sayeeda,Z., Budinski,Z., Guo,A., Lee,B.L., Berjanskii,M., Rout,M., Peters,H., Dizon,R., Mah,R., *et al.* (2022) NP-MRD: the Natural Products Magnetic Resonance Database. *Nucleic Acids Res.*, **50**, D665–D677.
12. Pin,M., Poynton,E.F., Jordan,T., Kim,J., Ledingham,B., van Santen,J.A., Yang,V., Maras,A., Tavangar,P., Gautam,V., *et al.* (2023) A data deposition platform for sharing nuclear magnetic resonance data. *J. Nat. Prod.*, **86**, 2554–2561.
13. Panter,F., Bader,C.D. and Müller,R. (2021) The sandarazols are cryptic and structurally unique Plasmid-encoded toxins from a rare myxobacterium. *Angew. Chem. Int. Ed. Engl.*, **60**, 8081–8088.
14. Okoth,D.A., Hug,J.J., Mándi,A., Kurtán,T., Garcia,R. and Müller,R. (2021) Structure and biosynthesis of sorangipyranone—a new γ-dihydropyrone from the myxobacterial strain MSr12020. *J. Ind. Microbiol. Biotechnol.*, **48**, kuab029.
15. Wang,D.-G., Niu,L., Lin,Z.-M., Wang,J.-J., Gao,D.-F., Sui,H.-Y., Li,Y.-Z. and Wu,C. (2021) The discovery and biosynthesis of nicotinic myxochelins from an *Archangium* sp. SDU34. *J. Nat. Prod.*, **84**, 2744–2748.
16. Jiang,M., Wu,Z., Wu,Q., Yin,H., Guo,H., Yuan,S., Liu,Z., Chen,S. and Liu,L. (2021) Amphichoterpenoids A–C, unprecedented picoline-derived meroterpenoids from the ascidian-derived fungus *Amphichorda felina* SYSU-MS7908. *Chin. Chem. Lett.*, **32**, 1893–1896.
17. Cao,P.-R., Li,M., Zhang,J.-S., Zheng,Y.-L., Chen,J., Zhao,Y.-Q., Qi,X.-D., Zhu,P.-H., Gu,Y.-C., Kong,L.-Y., *et al.* (2022) Epicoccanes A–D, four oxidative dimers of pyrogallol analogues from *Epicoccum nigrum*. *Org. Lett.*, **24**, 6789–6793.
18. Qin,L., Yong,K., Lian,X.-Y. and Zhang,Z. (2022) New metabolites (±)-bacillipyrrole A and bacillipyrazine A from the Mariana Trench-associated bacterium *Bacillus subtilis* SY2101. *Phytochem. Lett.*, **49**, 79–82.
19. Park,J.-D., Li,Y., Moon,K., Han,E.J., Lee,S.R. and Seyedsayamdost,M.R. (2022) Structural elucidation of cryptic algaecides in marine algal-bacterial symbioses by NMR spectroscopy and MicroED. *Angew. Chem. Int. Ed. Engl.*, **61**, e202114022.
20. Ge,Y., Ma,Y., Zhao,M., Wei,J., Wu,X., Zhang,Z., Yang,H., Lei,H. and Wu,B. (2022) Exploring gabosine and chlorogentisyl alcohol derivatives from a marine-derived fungus as EcGUS inhibitors with informatic assisted approaches. *Eur. J. Med. Chem.*, **242**, 114699.
21. Ohno,O., Iwasaki,A., Same,K., Kudo,C., Aida,E., Sugiura,K., Sumimoto,S., Teruya,T., Tashiro,E., Simizu,S., *et al.* (2022) Isolation of caldorazole, a thiazole-containing polyketide with selective cytotoxicity under glucose-restricted conditions. *Org. Lett.*, **24**, 4547–4551.
22. Zhang,J., Zhao,W.-Y., Wang,C., Yi,J., Yu,Z.-L., Deng,S., Zhang,H.-L., Huo,X.-K., Sun,C.-P. and Ma,X.-C. (2021) Identification, semisynthesis, and anti-inflammatory evaluation of 2,3-seco-clavine-type ergot alkaloids from human intestinal fungus *Aspergillus fumigatus* CY018. *Eur. J. Med. Chem.*, **224**, 113731.
23. Kurisawa,N., Iwasaki,A., Teranuma,K., Dan,S., Toyoshima,C., Hashimoto,M. and Suenaga,K. (2022) Structural determination, total synthesis, and biological activity of iezoside, a highly potent Ca^{2+}-ATPase inhibitor from the marine cyanobacterium *leptochromothrix valpauliae*. *J. Am. Chem. Soc.*, **144**, 11019–11032.
24. Kim,G.J., Mascuch,S.J., Mevers,E., Boudreau,P.D., Gerwick,W.H. and Choi,H. (2022) Luquilloamides, cytotoxic lipopeptides from a Puerto rican collection of the filamentous marine cyanobacterium *Oscillatoria* sp. *J. Org. Chem.*, **87**, 1043–1055.
25. Nainala,V.C., Rajan,K., Kanakam,S.R.S., Sharma,N., Weißenborn,V., Schaub,J. and Steinbeck,C. (2024) COCONUT 2.0: a comprehensive overhaul and curation of the collection of open natural products database. ChemRxiv doi: https://chemrxiv.org/engage/chemrxiv/article-details/66dee5e212ff75c3a1dbfd54, 10 September 2024, preprint: not peer reviewed.
26. Moumbock,A.F.A., Gao,M., Qaseem,A., Li,J., Kirchner,P.A., Ndingkokhar,B., Bekono,B.D., Simoben,C.V., Babiaka,S.B., Malange,Y.I., *et al.* (2021) StreptomeDB 3.0: an updated compendium of streptomycetes natural products. *Nucleic Acids Res.*, **49**, D600–D604.
27. Rutz,A., Sorokina,M., Galgonek,J., Mietchen,D., Willighagen,E., Gaudry,A., Graham,J.G., Stephan,R., Page,R., Vondrášek,J., *et al.* (2022) The LOTUS initiative for open knowledge management in natural products research. *eLife*, **11**, e70780.
28. Sun,M., Zhang,X., Hao,H., Li,W. and Lu,C. (2015) Nocarbenzoxazoles A–G, benzoxazoles produced by halophilic *Nocardiopsis lucentensis* DSM 44048. *J. Nat. Prod.*, **78**, 2123–2127.
29. Kim,T., Lee,S.-A., Noh,T., Choi,P., Choi,S.-J., Song,B.G., Kim,Y., Park,Y.-T., Huh,G., Kim,Y.-J., *et al.* (2019) Synthesis, structure revision, and cytotoxicity of nocarbenzoxazole G. *J. Nat. Prod.*, **82**, 1325–1330.
30. Knowles,S.L., Raja,H.A., Isawi,I.H., Flores-Bocanegra,L., Reggio,P.H., Pearce,C.J., Burdette,J.E., Rokas,A. and Oberlies,N.H. (2020) Wheldone: characterization of a unique scaffold from the coculture of *Aspergillus fischeri* and *Xylaria flabelliformis*. *Org. Lett.*, **22**, 1878–1882.
31. Rangel-Grimaldo,M., Earp,C.E., Raja,H.A., Wood,J.S., Mardiana,L., Ho,K.L., Longcake,A., Williamson,R.T., Palatinus,L.,

Hall,M.J., *et al.* (2024) Wheldone revisited: structure revision via DFT-GIAO chemical shift calculations, 1,1-HD-ADEQUATE NMR spectroscopy, and X-ray crystallography studies. *J. Nat. Prod.*, **87**, 2095–2100.
32. Xie,F., Xia,D.-D., Duan,H.-J., Sun,Y., Zi,Z.-F., Wan,D.-Y., Zhou,H. and Ding,Z.-T. (2023) Two decarestrictine analogs from the soil-derived fungus *Penicillium* sp. YUD18003 associated with *Gastrodia elata*. *Chem. Biodivers.*, **20**, e202300566.
33. Novitskiy,I.M., Elyashberg,M., Bates,R.W., Kutateladze,A.G. and Williams,C.M. (2023) Penicitone: structural reassignment of a proposed natural product acid chloride. *Org. Lett.*, **25**, 7796–7799.
34. Morino,T., Masuda,A., Yamada,M., Nishimoto,M., Nishikiori,T., Saito,S. and Shimada,N. (1994) Stevastelins, novel immunosuppressants produced by *Penicillium*. *J. Antibiot. (Tokyo)*, **47**, 1341–1343.
35. Sarabia,F., Chammaa,S. and López-Herrera,F.J. (2002) A macrolactonization approach to the stevastelins. *Tetrahedron Lett.*, **43**, 2961–2965.
36. Kurosawa,K., Matsuura,K. and Chida,N. (2005) Total synthesis of stevastelins B3 and C3: structure confirmation of stevastelin B3 and revision of stevastelin C3. *Tetrahedron Lett.*, **46**, 389–392.
37. Okabe,T., Nomoto,K., Funabashi,H., Okuda,S., Suzuki,H. and Tanaka,N. (1985) Lactoquinomycin, a novel anticancer antibiotic. II. Physico-chemical properties and structure assignment. *J. Antibiot. (Tokyo)*, **38**, 1333–1336.
38. Tatsuta,K., Ozeki,H., Yamaguchi,M., Tanaka,M. and Okui,T. (1990) Enantioselective total synthesis of medermycin (lactoquinomycin). *Tetrahedron Lett.*, **31**, 5495–5498.
39. Léo,P.-M., Morin,C. and Philouze,C. (2002) Structure revision of medermycin/lactoquinomycin A and of related C-8 glycosylated naphthoquinones. *Org. Lett.*, **4**, 2711–2714.
40. Williamson,R.T., McDonald,L.A., Barbieri,L.R. and Carter,G.T. (2002) In support of the original medermycin/lactoquinomycin A structure. *Org. Lett.*, **4**, 4659–4662.

Received: September 13, 2024. **Revised:** October 18, 2024. **Editorial Decision:** October 18, 2024. **Accepted:** October 24, 2024