



Expanding and exploring the chemical space of food chemicals and natural products

José L. Medina-Franco

DIFACQUIM research group

School of Chemistry

National Autonomous University of Mexico
(UNAM)



XXX Symposium on Bioinformatics and Computer-Aided Drug Discovery (BCADD-2024)
September 18th, 2024

Content

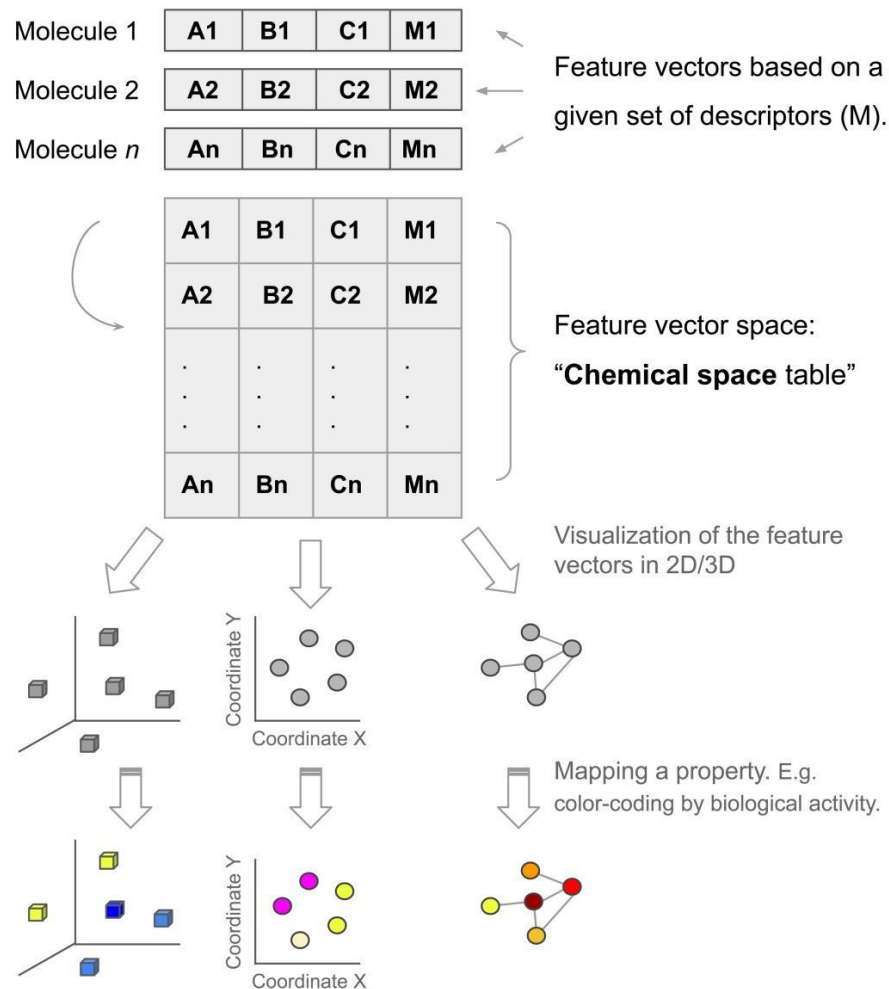
- Chemical space.
- Natural products in drug development.
- Molecular databases: food chemicals and natural products.
- **LANaPDB: Latin American Natural Product Database.**
- Summary.
- Outlook.

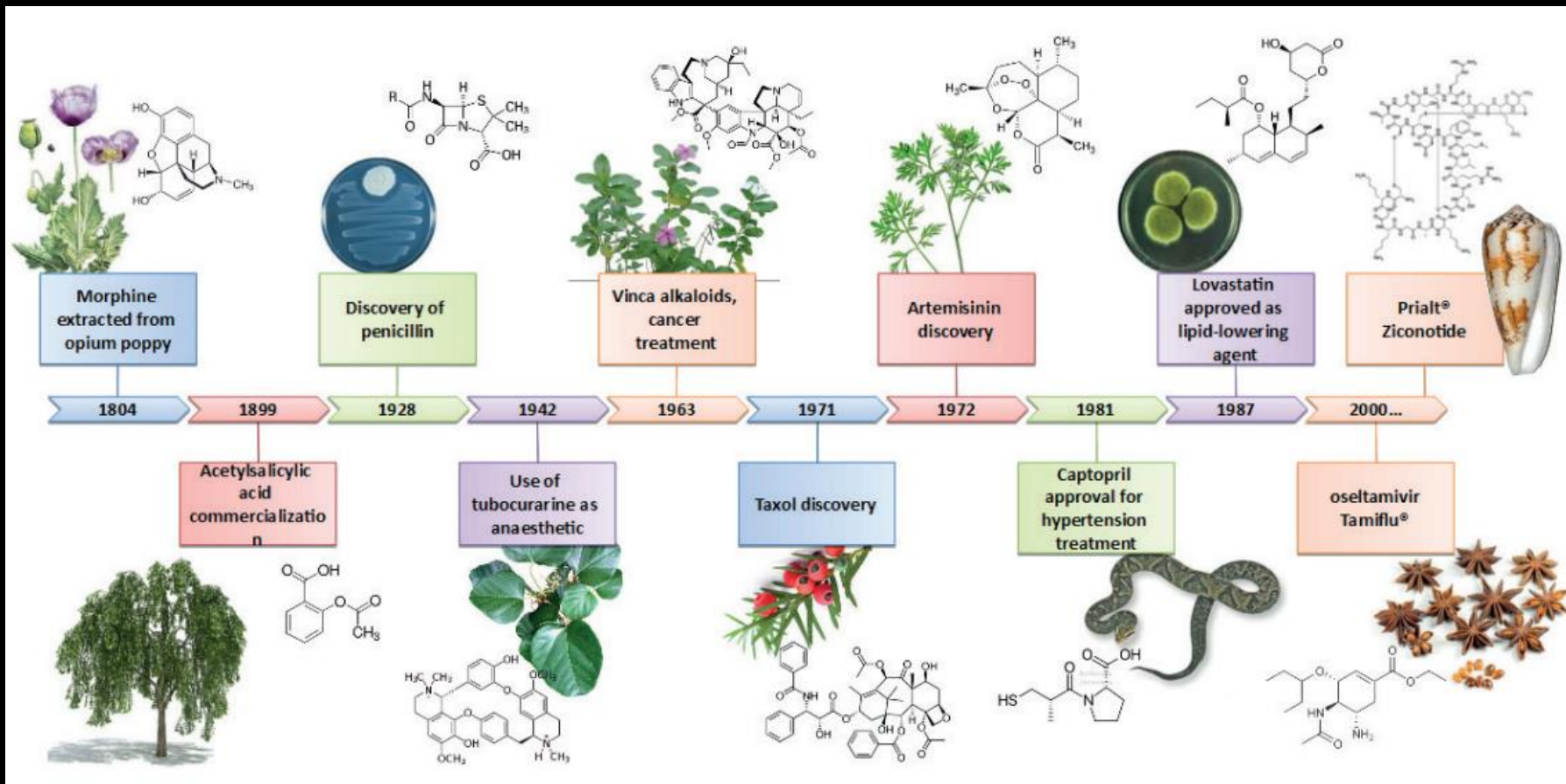
Chemical space: key concept in chemoinformatics*

Author(s)	Chemical space definitions
Varnek and Baskin	The ensemble of graphs or descriptor vectors forms a chemical space in which some relations between the objects must be defined.
Lipinski and Hopkins	Chemical space can be viewed as being analogous to cosmological universe in its vastness, with chemical compounds populating space instead of stars.
Reymond et al.	Ensemble of all known and possible molecules described by their chemical properties.
Vogt	Comprehensive collection of all possible small molecules under some reasonable restrictions considering size and composition.

Applications

- Structure-property relationships
- Structural diversity analysis
- Library design
- Compound selection

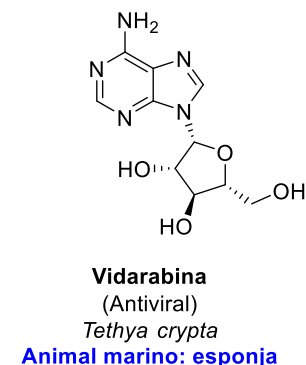
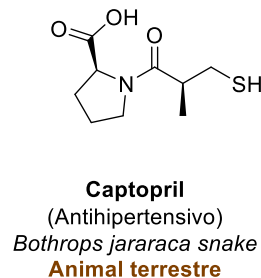
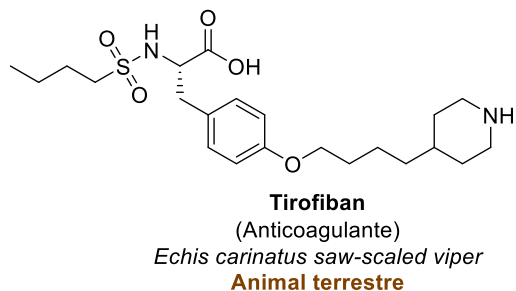
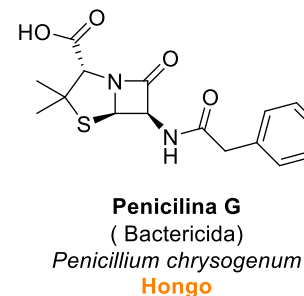
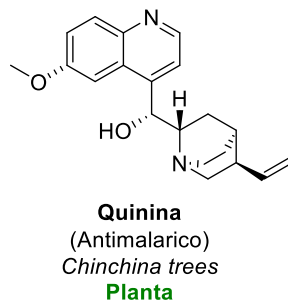
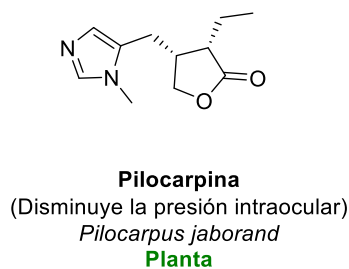
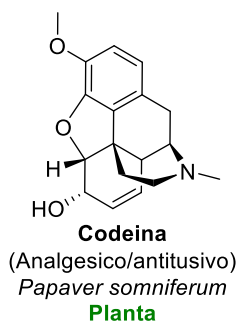
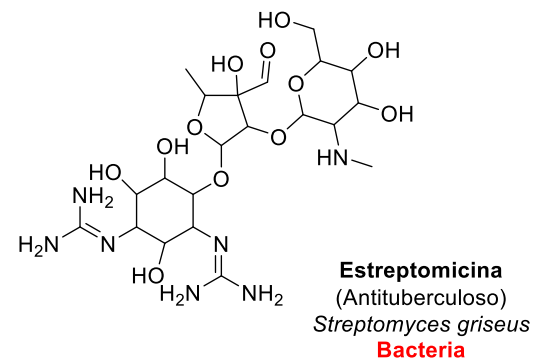
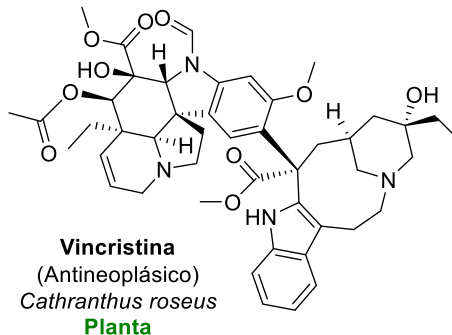
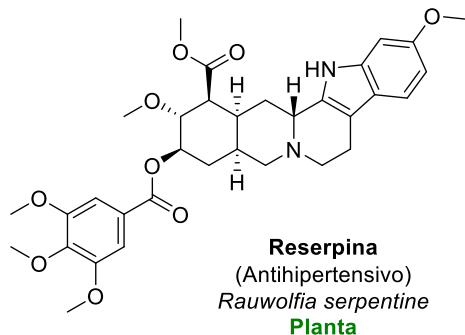




Valli and Bolzani. *Anais da Academia Brasileira de Ciências* 2019 91: e20190208

Natural products and databases in drug development

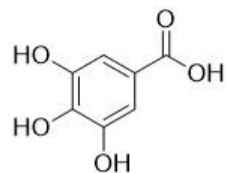
Natural products as drugs



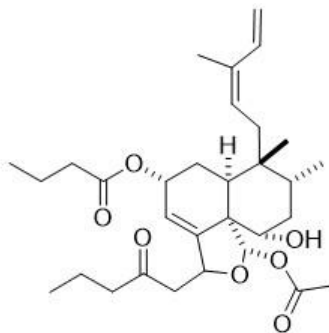
Atanasov-G, A. et al. Natural products in drug discovery: advances and opportunities.
Nature Rev. Drug Discov. 2021 20:200-216

Examples of
**bioactive
 compounds**
 obtained in Latin
 American
medicinal plants.

Brazil

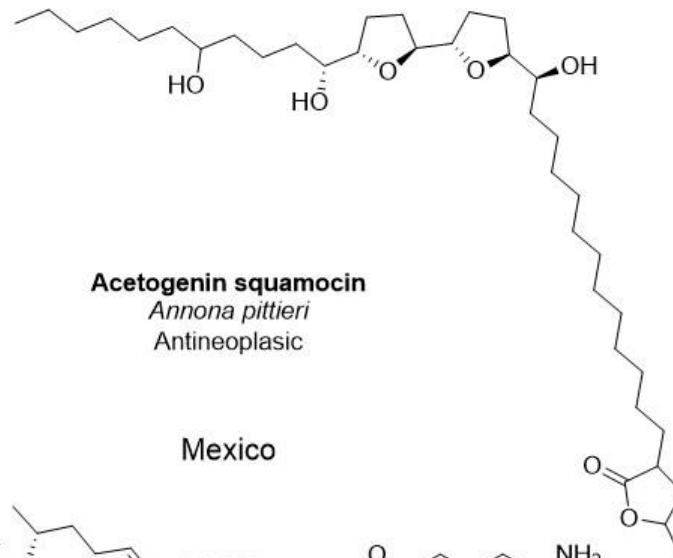


Gallic acid
Stryphnodendron adstringens
 Antioxidant and antineoplastic



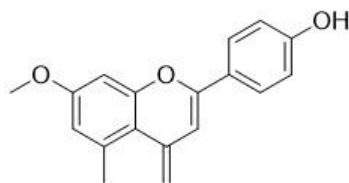
Casearin X
Casearia sylvestris
 Antineoplastic

Costa Rica



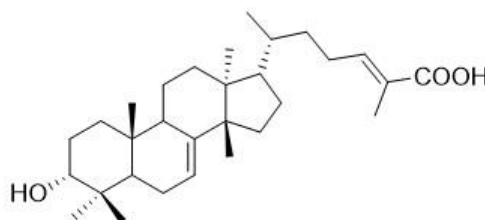
Acetogenin squamocin
Annona pittieri
 Antineoplastic

El Salvador

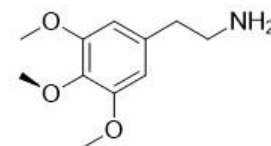


Genkwanin
Calea tenuifolia
 Antiplasmodial

Mexico

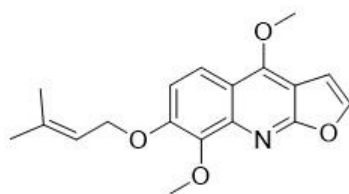


3α-Hydroxy masticadienoic acid
Amphipterygium adstringens
 Antineoplastic, antimicrobial and
 wound healing



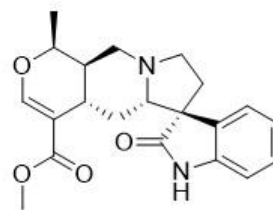
Mescaline
Lophophora williamsii
 Improvement in psychiatric conditions

Panama



Furoquinoline alkaloid
Desmotes incomparabilis
 Antileishmanial

Peru



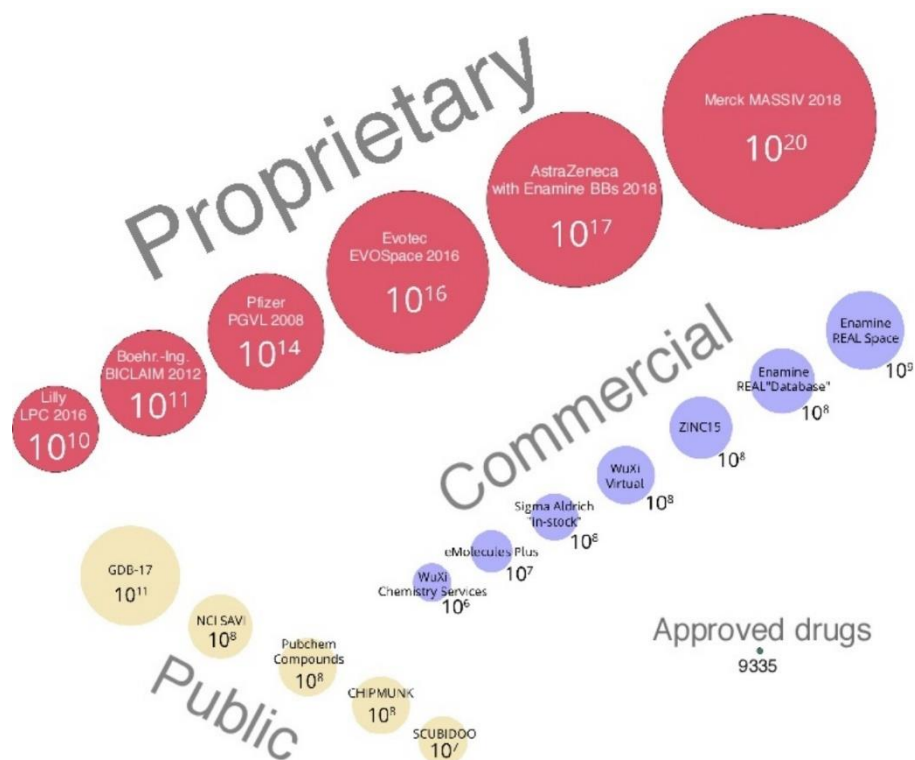
Mitraphylline
Uncaria tomentosa
 Antineoplastic



Benzylglucosinolate
Lepidium meyenii
 Antineoplastic

Compound collections: types and size

- Commercial: for purchase and testing.
 - General screening (diverse libraries).
 - Focused on a specific target or indication.
- *De novo*.
- On-demand.
- Natural products and food chemicals.



Chemical Multiverse and Diversity of Food Chemicals

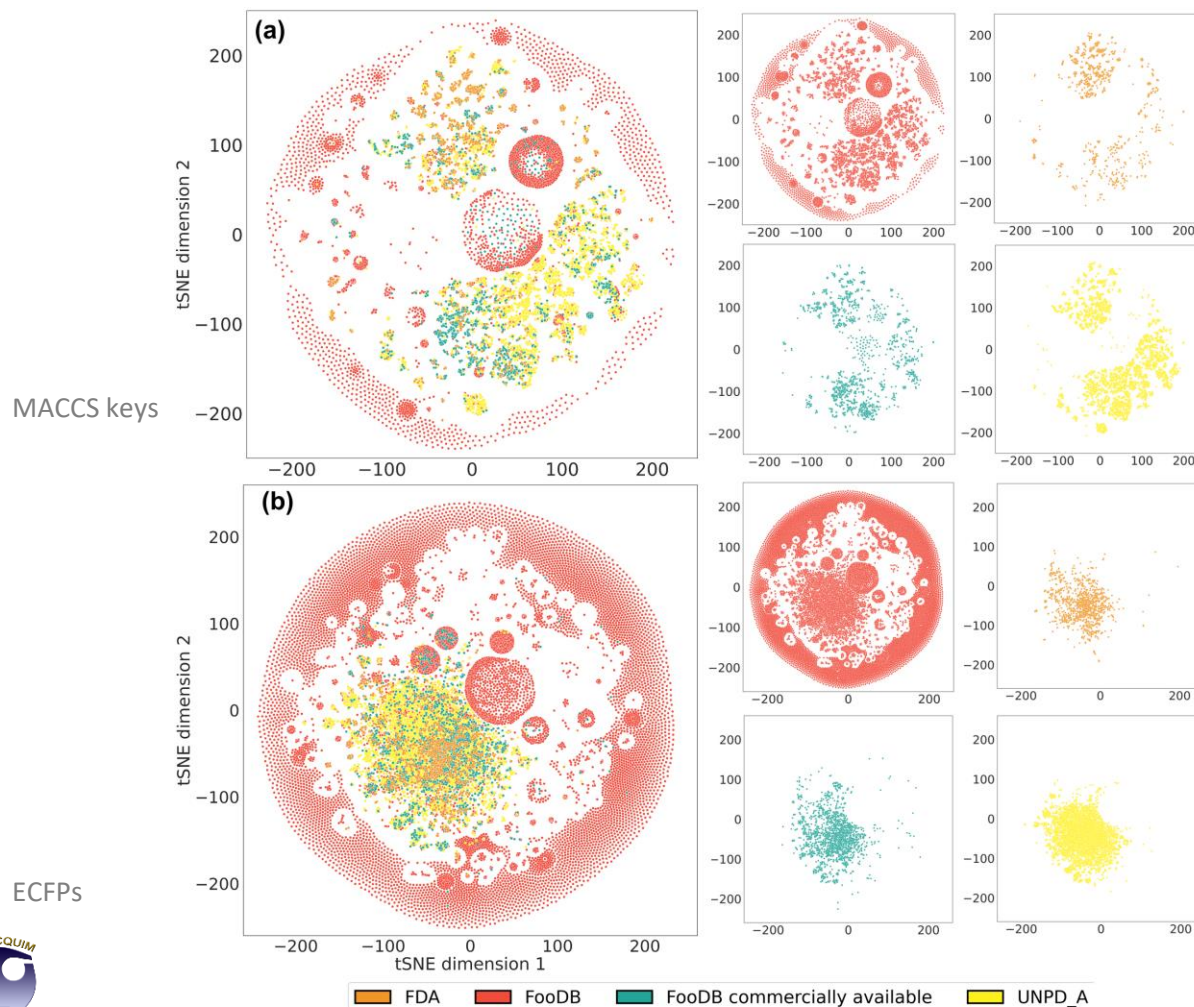
Juan F. Avellaneda-Tamayo, Ana L. Chávez-Hernández, Diana L. Prado-Romero, and José L. Medina-Franco*

Cite This: *J. Chem. Inf. Model.* 2024, 64, 1229–1244

Read Online



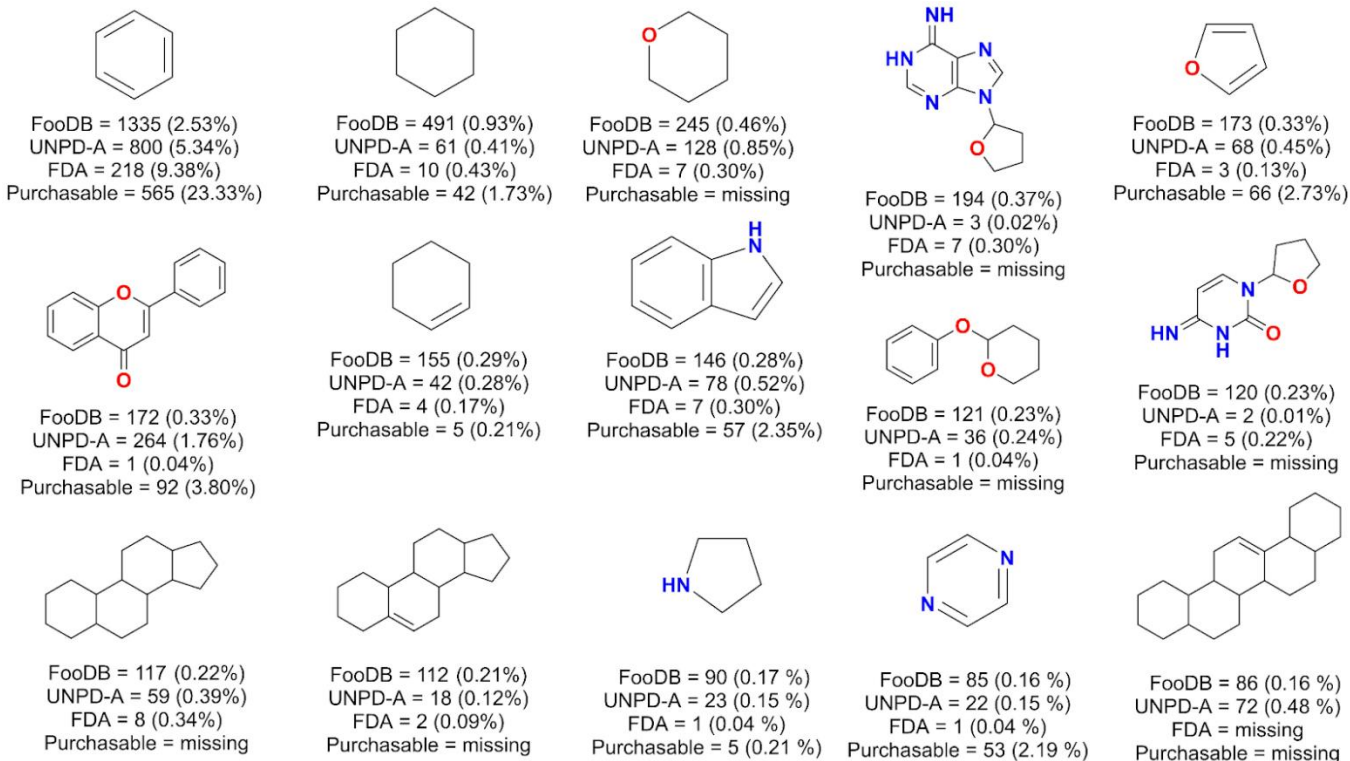
~77,000 food chemicals



Food chemicals expand the traditional **chemical multiverse**

of bioactive compounds
Chemical multiverse:
multiple chemical spaces

Characterization of food chemicals



Fifteen most frequent scaffolds from FooDB and their presence in the four reference (bioactive) data sets: natural products and approved drugs:

➤ Large overlap of bioactive rings.

CN1C=NC2=C1C(=O)N(C(=O)N2C)
 SMARTS: [#6]-[#7]-1-[#6]-[#7]-[#6]-2-[#6]-1-[#6](-O)-[#7]-[#6](-O)-[#7]-2-[#6]
 InChI:
 InChI=1S/C8H10N4O2/c1-10-4-9-6-5(10)/13/12/3(14)/11/8/2(8)H,1-3H3
 InChI Key:
 RYYVLZVUUVJVGH-UHFFFAOYSA-N

Natural products databases

Contributions from Latin America

COCONUT: the COLleCtion of Open NatUral productTs

A comprehensive platform facilitating natural product research by providing data, tools, and services for deposition, curation, and reuse.

[Molecules](#) [Organism](#) [Citations](#)

Search compound name, SMILES, InChI, InChI Key

Search

Try: [Caffeine](#), [CNP0228556](#)

Report bugs: [Issue Tracker](#)



Draw Structure



Browse Data



Submit Data

Total Molecules

695133

Total Collections

63

Unique Organisms

53092

Citations Mapped





34991



<https://coconut.naturalproducts.net/>

Nuclei of Bioassays, Biosynthesis and Ecophysiology of Natural Products (NuBBE)



 GENERAL INFORMATION	 SPECIES	 SPECIES LOCATION
<input type="text" value="Common Name"/>	<input type="text" value="Choose an option, or ref"/>	<input type="text" value="Brasil"/>
<input type="text" value="Please Choose..."/>	<input type="text" value="Family"/>	 BIOLOGICAL PROPERTIES
<input type="text" value="Fórmula Molecular"/>	<input type="text" value="Genus"/>	
	<input type="text" value="Specie"/>	

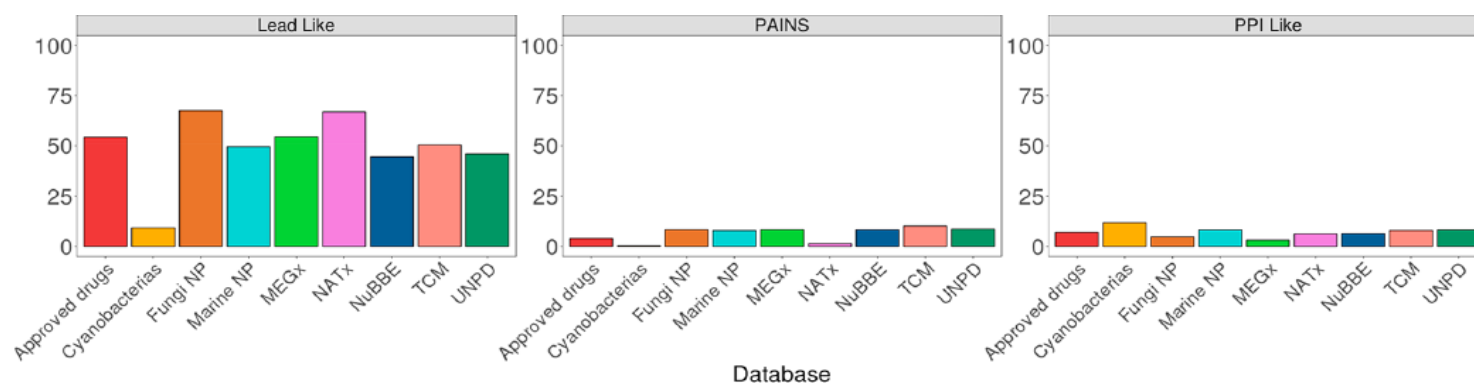
- 2223 compounds
- Plants, terrestrial and marine animals; microorganisms
- São Paulo State University; University of São Paulo

<http://www.nubbe.iq.unesp.br/portal/nubbe-search.html>

Pilon AC et al. NuBBEDB: An updated database to uncover chemical and biological information from Brazilian biodiversity. *Sci. Rep.* 2017 7:7215

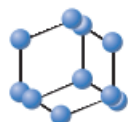
Chemical Space and Diversity of the NuBBE Database: A Chemoinformatic Characterization

Fernanda I. Saldívar-González,[†] Marília Valli,^{‡,ID} Adriano D. Andricopulo,[§] Vanderlan da Silva Bolzani,[‡] and José L. Medina-Franco^{*,†,ID}



Anti-Infective Agents, 2019, 17, 138-149

RESEARCH ARTICLE



BENTHAM SCIENCE

New *Trypanosoma cruzi* Trypanothione Reductase Inhibitors Identification using the Virtual Screening in Database of Nucleus Bioassay, Biosynthesis and Ecophysiology (NuBBE)



Nelcí do Carmo Santos¹, Vinícius G. da Paixão¹ and Samuel S. da Rocha Pita^{1,*}

¹Bioinformatics and Molecular Modeling Laboratory (LaBiMM), Pharmacy College, Federal University of Bahia, Salvador-BA, Brazil

PERUNPDB

Peruvian Natural Products Database

Compound

List

Image

hydroxybenzoic, etc...

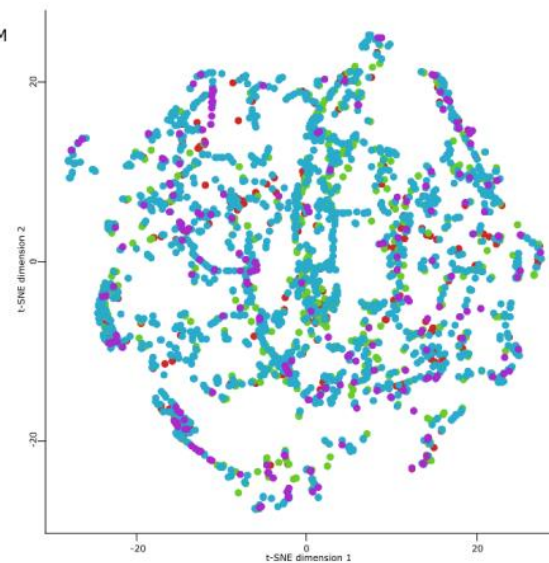
Search

<https://perunpdb.com.pe/>

Dr. Miguel Angel Chávez

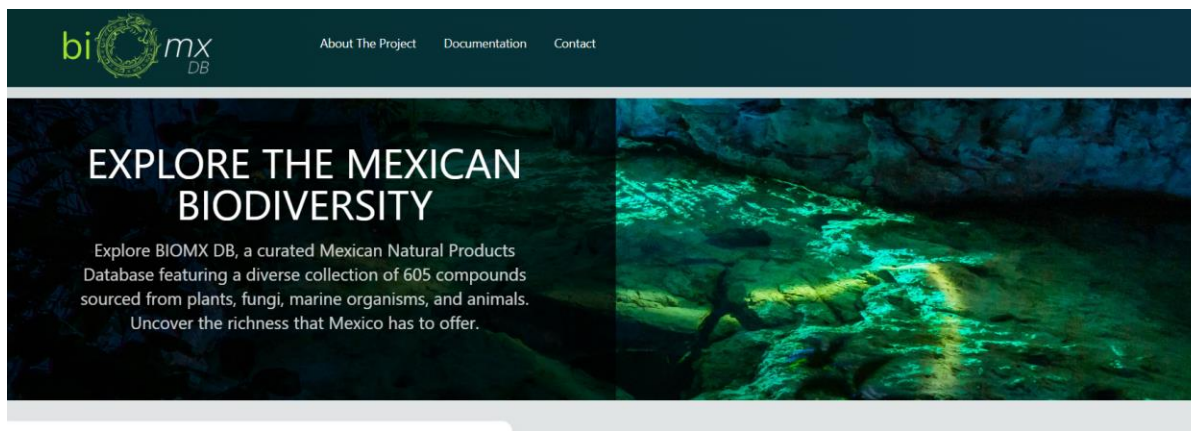

- 280 compounds
- Plants & animals
- Universidad Católica de Santa María,
Arequipa, Perú

● BIOFAQUIM
● AfroDB
● NuBBEDB
● PeruNPDB



PeruNPDB large structural diversity

Barazorda-Ccahuana HL et al. PeruNPDB: The Peruvian Natural Products Database for in silico drug screening. *Sci. Rep.* 2023 13:7577

**APPLICATION NOTE****BIOMX-DB: A web application for the BIOFACQUIM natural product database**Fernando Martínez-Urrutia | José L. Medina-Franco 

biomx DB

About The Project Documentation Contact

EXPLORE THE MEXICAN BIODIVERSITY

Explore BIOMX DB, a curated Mexican Natural Products Database featuring a diverse collection of 605 compounds sourced from plants, fungi, marine organisms, and animals. Uncover the richness that Mexico has to offer.

Advance Search

Search by Name

Source of NP

Potential Bioactivity

(4R 5S 6R)-6-Chloro-4 5-dihydroxy-3-methoxy-5-methylcyclohex-2-enone
Bioactivity: Phytotoxic
Origin: Fungus Xylaria feejeensis

Acacatin
Bioactivity: Hypoglycemic
Origin: Plant Anoda cristata

- 605 compounds
- Plants, fungi, marine animals

www.biomx-db.com

Other databases under update and development



About Team Contact Help How to Cite

Dr. Marcus Tullius Scotti
Coordinator
Universidad Federal de Paraiba

About

SISTEMAT X Web, an acronym of "SISTEMAT eXtended Webservices", is suite of tools to manage databank of secondary metabolites that is available to consult for entire scientific community, through a "Web" interface. SISTEMAT X Web manages databank of secondary metabolites, including information about the species and respective taxonomy data (genus, tribe, family). Additionally, it is possible to manage physical-chemical properties and experimental data, as NMR spectra or biological activity. The structures of secondary metabolites are available in MDL format (.mol) in 2 dimensions (2D) or (3D).



Powered by Chemaxon



ChemDoodle
WEB COMPONENTS



PASS

PLATO

MolPredict



Costa RPO et al. *J. Chem. Inf. Model.* 2021: 61 2516-2522

<https://sistemax.ufpb.br/>

ChemRxiv®

How To Submit Browse About News

Organic Chemistry

ChemRxiv®

How To Submit Browse About News

Theoretical and Computational Chemistry

NaturAr, a collaborative, open source, database of natural products from Argentinian biodiversity for drug discovery and bioprospecting

15 July 2024, Version 1

Working Paper

[Leandro Martínez-Heredia](#) , [Patricia Quispe](#) , [Julián Fernández](#) , [Martin Lavecchia](#) [Hide author details](#)

NPDBEjeCol: a natural products database from Colombia

09 August 2024, Version 1

Working Paper

[Johny Roberto Rodríguez-Pérez](#) , [Hoover Albeiro Valencia-Sanchez](#) ,
[Oscar Marino Mosquera-Martínez](#) , [Alejandro Gómez-García](#) , [Jose Luis Medina-Franco](#) ,
[Hector Fabio Cortes-Hernandez](#)

[Show author details](#)





Towards a unified chemolibrary

Latin American Natural Product Database

Latin American Natural Product Database

LANaPDB



LANaPDB ~13,000 compounds




pharmaceuticals

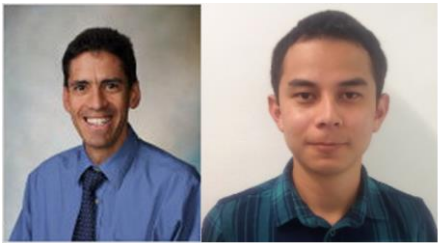


Article

Navigating the Chemical Space and Chemical Multiverse of a Unified Latin American Natural Product Database: LANaPDB

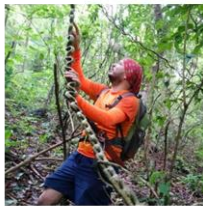
Alejandro Gómez-García ¹, Daniel A. Acuña Jiménez ², William J. Zamora ^{2,3,4}, Haruna L. Barazorda-Ccahuana ⁵, Miguel Á. Chávez-Fumagalli ⁵, Marilia Valli ⁶, Adriano D. Andricopulo ⁶, Vanderlan da S. Bolzani ⁷, Dionisio A. Olmedo ⁸, Pablo N. Solís ⁸, Marvin J. Núñez ⁹, Johnny R. Rodríguez Pérez ^{10,11}, Hoover A. Valencia Sánchez ¹⁰, Héctor F. Cortés Hernández ¹⁰ and José L. Medina-Franco ^{1,*}

Latin American Natural Product Database work team



Dr. José Luis Medina Franco MSc. Alejandro Gómez García

México
BIOFACQUIM



Dr. Marvin José Núñez
El Salvador
LAIPNUDESAL



Dr. Héctor Fabio Cortés Hernández MSc. Johny Roberto Rodríguez Pérez

Colombia
CIFPMA



Dr. Miguel Ángel Chávez Fumagalli

Perú



Dr. Willam Zamora
Costa Rica
NAPRORE-CR



Dr. Dionisio Antonio Olmedo
Panamá
CIFPMA



Dra. Vanderlan da Silva Bolzani Dr. Adriano D. Andricopulo Dra. Marilia Valli

Brasil



Latin America Databases at LANaPDB (v.2)

Database	Number of compounds ^a	Country	Source	General description
NuBBEDB	2223	Brazil	Plants Microorganisms Terrestrial animals Marine animals	Natural products of Brazilian biodiversity. Developed by the São Paulo State University and the University of São Paulo.
SistematX	9514	Brazil	Plants	Database composed of secondary metabolites and developed at the Federal University of Paraiba.
UEFS	503	Brazil	Plants	Natural products that have been separately published. Developed at the State University of Feira de Santana.
NPDB EjeCol	200	Colombia	Plants	Natural products and foods from plants in the Eje Cafetero Region. Database from the Technological University of Pereira.
NAPRORE-CR	1600	Costa Rica	Plants Microorganisms	Developed in the University of Costa Rica.
LAIPNUDELSAV	214	El Salvador		Developed by the Research Laboratory in Natural Products of the University of El Salvador.

Latin America Databases at LANaPDB (v.2)

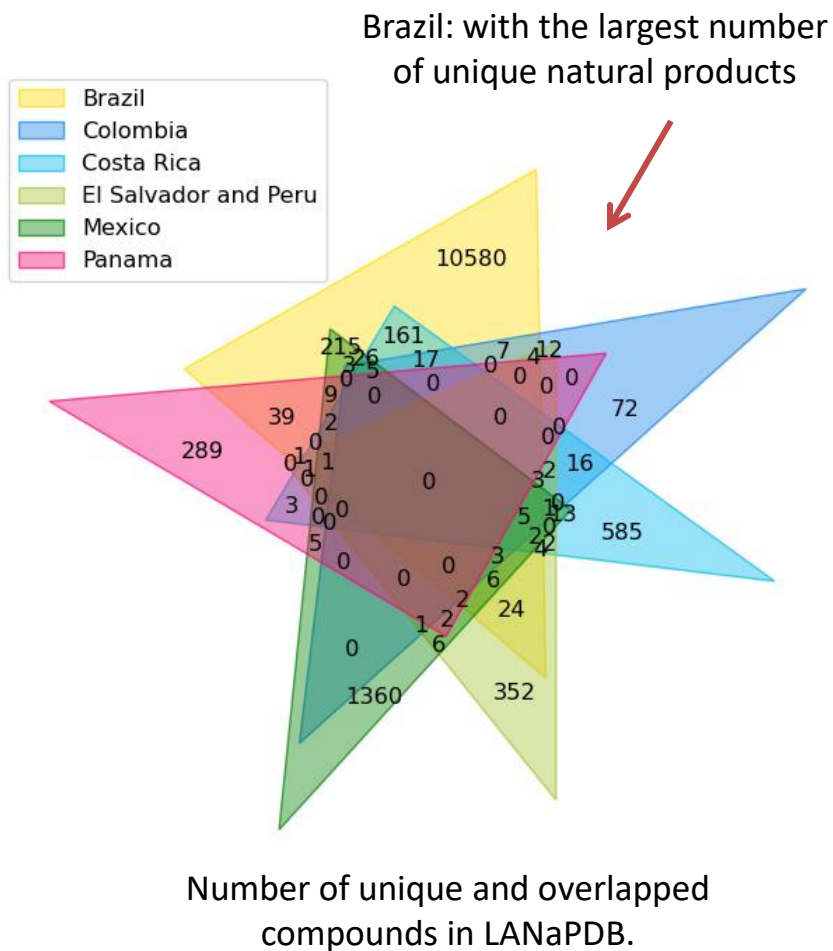
Database	Number of compounds ^a	Country	Source	General description
UNIQUIM	1112	Mexico	Plants	Natural products mainly isolated and characterized at the Institute of Chemistry of the National Autonomous University of Mexico.
BIOFACQUIM	750	Mexico	Plants Fungus Propolis Marine animals	Natural products isolated and characterized in Mexico at the School of Chemistry of the National Autonomous University of Mexico and other Mexican institutions.
CIFPMA	363	Panama	Plants	Natural products that have been tested in over 25 <i>in vitro</i> and <i>in vivo</i> bioassays, for different therapeutic targets. Developed at the University of Panama.
PeruNPDB	280	Peru	Animals Plants	Created at the Catholic University of Santa Maria.



Gómez-González A. et al. *Pharmaceuticals* 2023 16: 1388

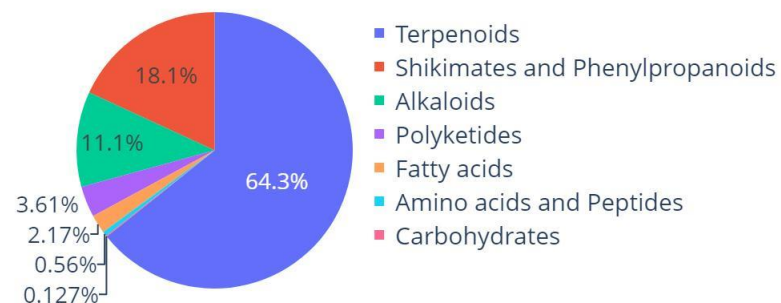
Gómez-González A. et al. *Mol. Inf.* 2024 43:e202400052

Unique compounds and structure classification

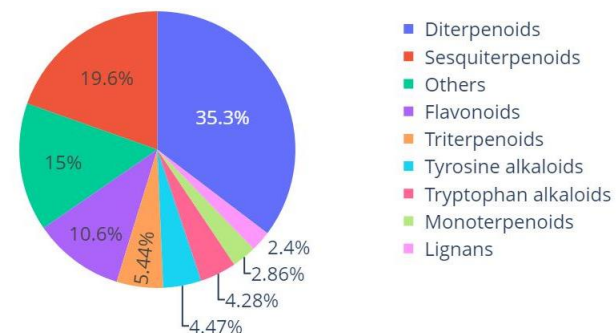


Terpenoids and diterpenoids are the most abundant

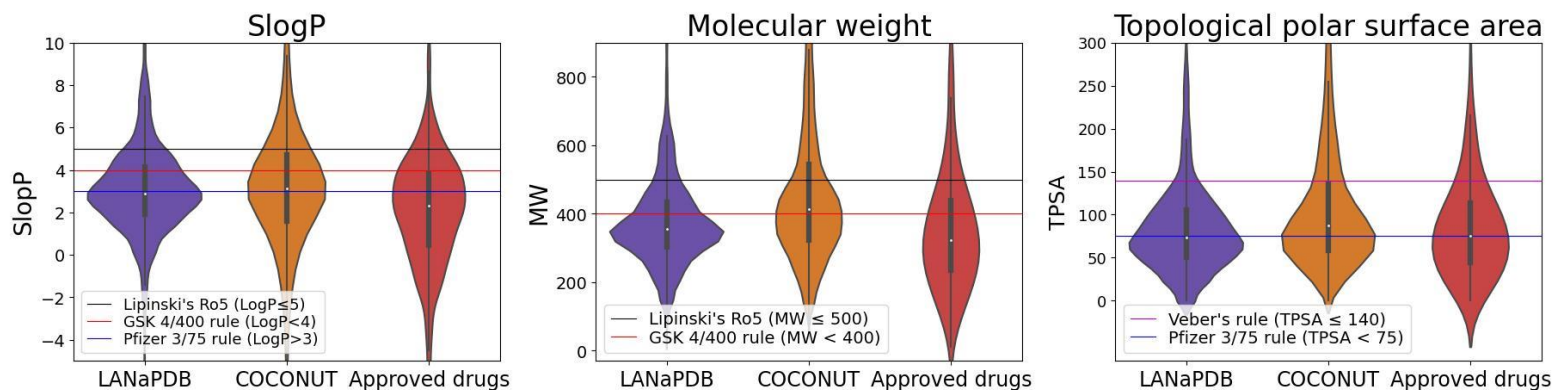
Pathway



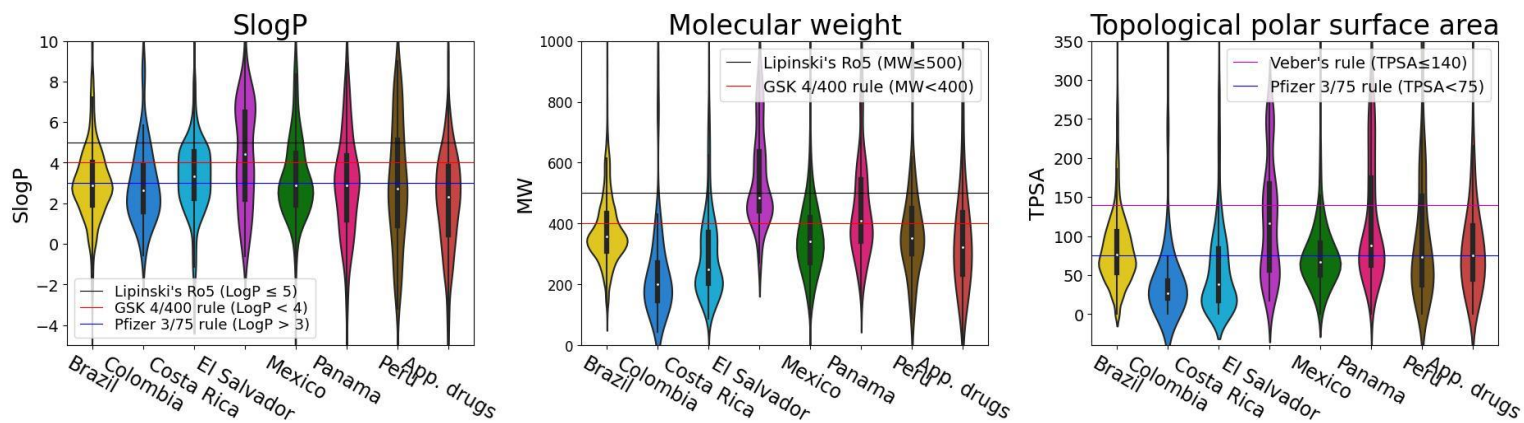
SuperClass



Property profile

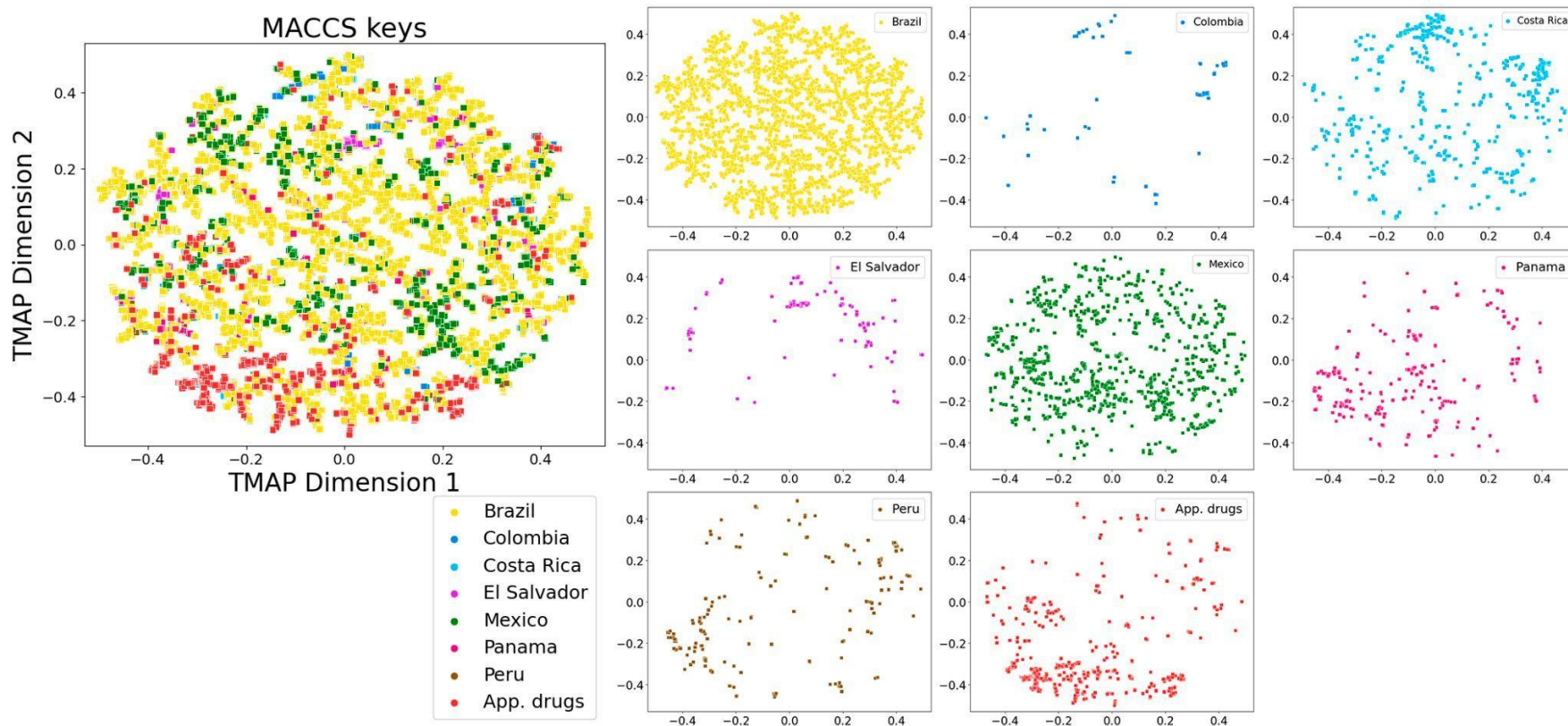


LANaPDB has similar profile to COCONUT (and distinct from approved drugs)



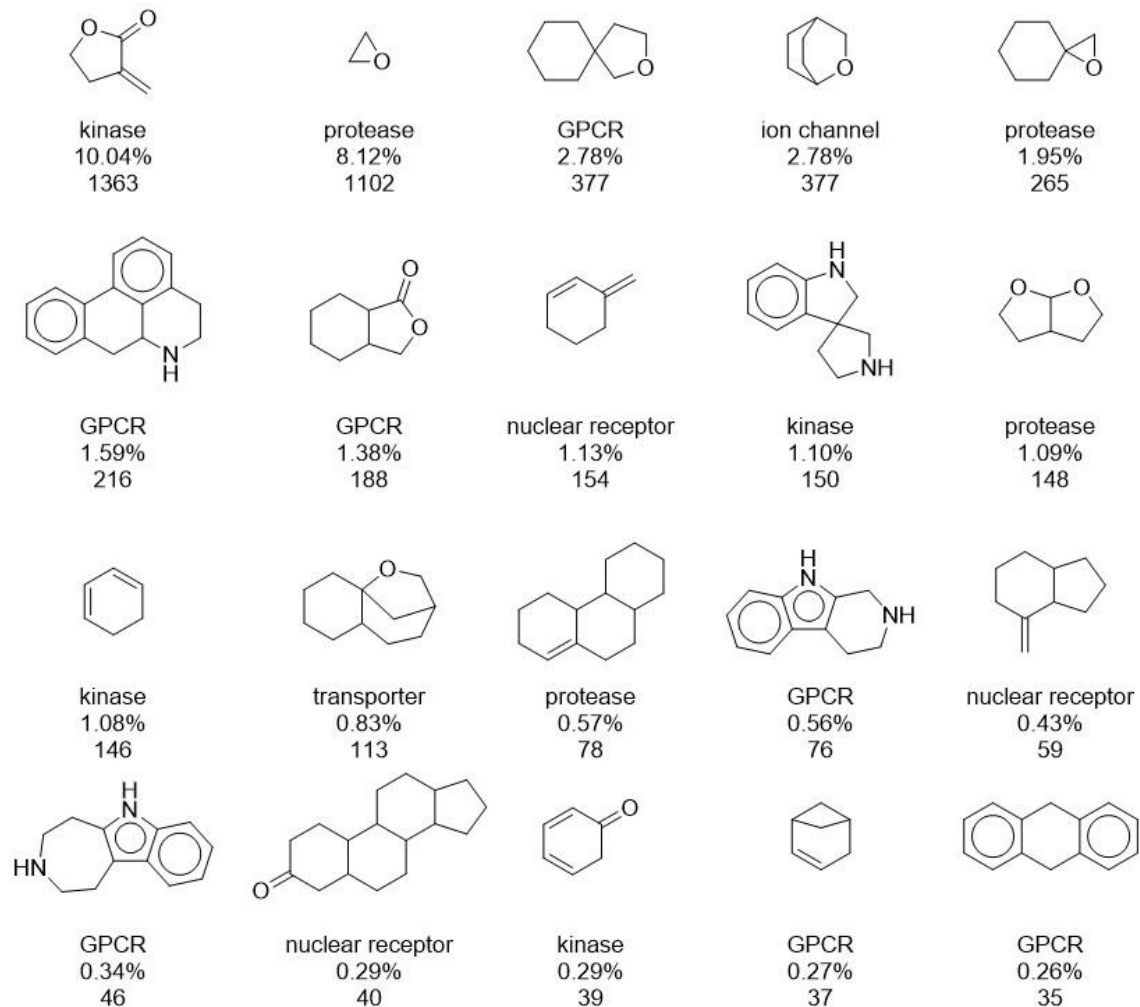
Natural products from El Salvador and Panama have distinct property profile.

Visualization of chemical space



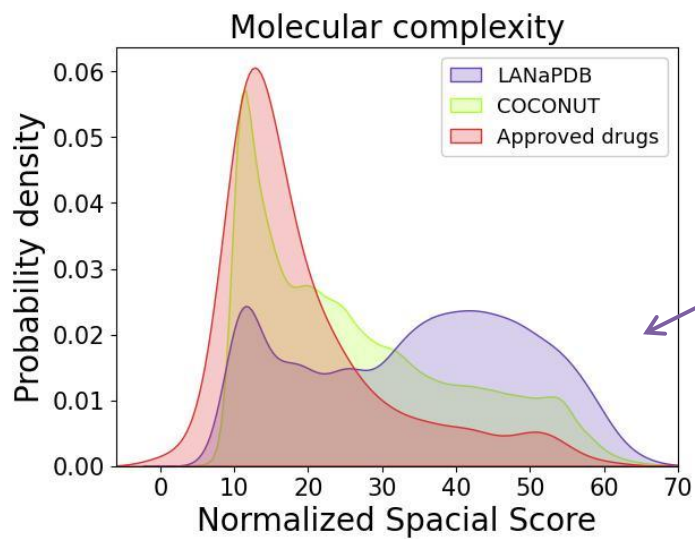
Natural products from **Brazil** are the most diverse, followed by **Mexico** and **Costa Rica**.

Bioactive (“magic”) rings in LANA-PDB



The twenty most abundant **bioactive ring systems** in LANA-PDB, their biological target and percentage of occurrence.


Molecular complexity



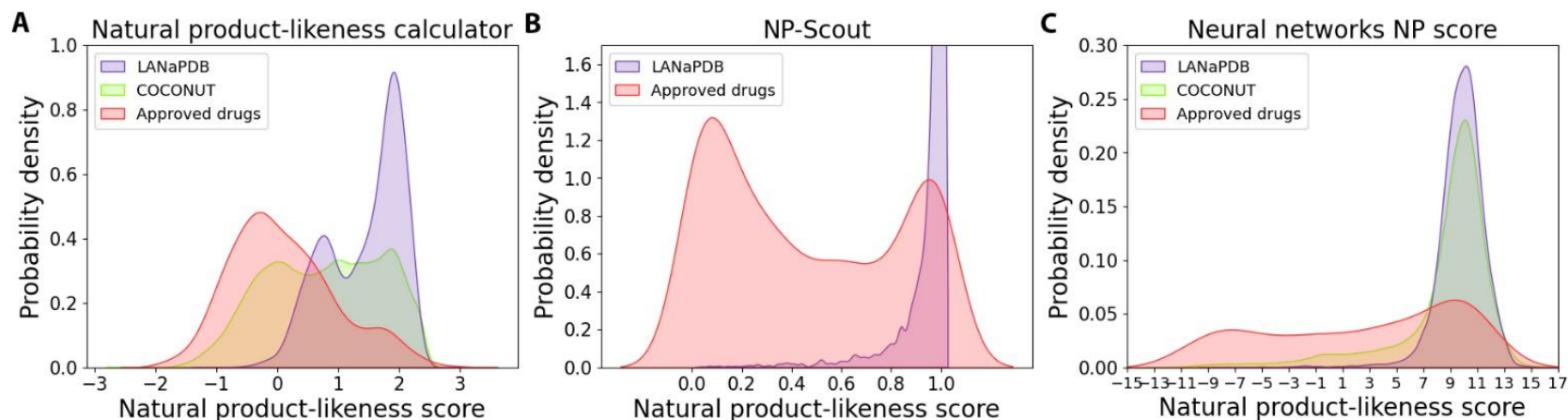
- **LANaPDB**: large fraction of compounds potentially selective.
- Higher structural complexity than COCONUT.

Krzyzanowski A et al. Spatial Score—A Comprehensive Topological Indicator for Small-Molecule Complexity. *J. Med. Chem.* 2023 66:12739.

Updating and profiling the natural product-likeness of Latin American compound libraries

Alejandro Gómez-García¹  | Ann-Kathrin Prinz² | Daniel A. Acuña Jiménez³ | William J. Zamora^{3,4,5} | Haruna L. Barazorda-Ccahuana⁶ | Miguel Á. Chávez-Fumagalli⁶ | Marilia Valli⁷ | Adriano D. Andricopulo⁷ | Vanderlan da S. Bolzani⁸ | Dionisio A. Olmedo⁹ | Pablo N. Solís⁹ | Marvin J. Núñez¹⁰ | Johny R. Rodríguez Pérez^{11,12} | Hoover A. Valencia Sánchez¹¹ | Héctor F. Cortés Hernández¹¹ | Oscar M. Mosquera Martínez¹³ | Oliver Koch² | José L. Medina-Franco¹

- NPLC (Ertl's) distinguishes LANaPDB from COCONUT and approved drugs.
- NP-Scout sharp distinction between databases.
- Neural-network-based score does not distinguish NP collections (based on structural and property descriptors).

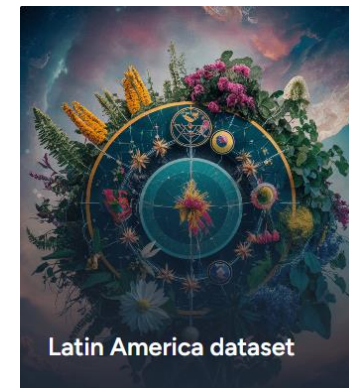


Kernel density estimate plots that represent the distribution of the natural product-likeness scores of LANaPDB, COCONUT and approved drugs calculated with: A) [Natural product-likeness calculator](#) B) [NP-Scout](#) and C) [Neural networks NP score](#).

Summary and outlook

- Natural products and food chemicals expand the bioactive chemical space.
- **LANaPDB**: Ten Latin American databases; **13,580** compounds.
- Open access at <https://github.com/alexgoga21/LaNaPDB>
- Available in COCONUT (under Collections)
- For each compound:
 - Linear notation.
 - Reference (peer-reviewed paper).
 - Taxonomic and structural classification.
 - Physicochemical properties and chirality.
 - Molecular complexity and synthetic feasibility scores.
 - Commercial availability and bioactivity data.

COCONUT



In progress:

- Physical collection for experimental screening.
- Several virtual screenings vs. multiple endpoints.

Chemical
Science

ROYAL SOCIETY
OF CHEMISTRY

PERSPECTIVE

View Article Online
View Journal | View Issue



Cite this: *Chem. Sci.*, 2022, **13**, 1526

All publication charges for this article have been paid for by the Royal Society

Natural product drug discovery in the artificial intelligence era

F. I. Saldívar-González,^a V. D. Aldas-Bulos,^b J. L. Medina-Franco^{*a} and F. Plisson^{id}^{*c}



Acknowledgements: DIFACQUIM research group

Computer-Aided Drug Design at School of Chemistry

Bachelors

Jazmín Miranda

Jimena García

Fernando Martínez

Gabriela Valle

Pedro A. Laurel

Armando Rufino

Oswaldo Sánchez

MSc

Brayan Raziel

Massyel S. Martínez

Juan F. Avellaneda

Daniel Márquez

Israel Castañeda

Verónica Ramírez

www.difacquim.com

medinajl@unam.mx

PhD

Alejandro Gómez

Ana Chávez, PhD

Fernanda Saldívar

Diana Prado

Edgar López (CINVESTAV)

Co-director: Carlos C. Rojas

Postdoctoral fellow

Alberto Marbán



Visiting scholars

Johny R. Rodríguez (Colombia)

Daniela Quadros (Brazil)

Samuel Homberg (Germany)

Naicolette Agudo (Panama)



Funding



2nd School of Chemoinformatics in Latin America

December 4-6, 2024

Speakers



Carolina Andrade



Adriano Andricopolo



Luis Bahena



Jürgen Bajorath



Haruna Barazorda



Miguel Angel Chávez



José Correa Basurto



Gonzalo Colmenarejo



Kimito Funatsu



Abraham Madariaga



Vinicius Maltarollo



Filip Miljković



Patricia Ruiz



Norberto Sánchez



Hanoch Senderowitz



Alan Talevi



Fidele Ntie-Kang



David Ramírez



Alexander Tropsha



Alexandre Varnek



William Zamora

Organizers

Jose Luis Medina-Franco



Karina Martinez-Mayorga

Free Online event



Registration

<https://schoolchilatina.com/>