

NATURAL PRODUCTS DATABASES AS VALUABLE SOURCES OF BIOACTIVE STRUCTURES FOR VIRTUAL SCREENING

Marcus Tullius Scotti

Laboratory of Cheminformatics - Universidade Federal da Paraíba

Marcus Tullius Scotti

Lattes

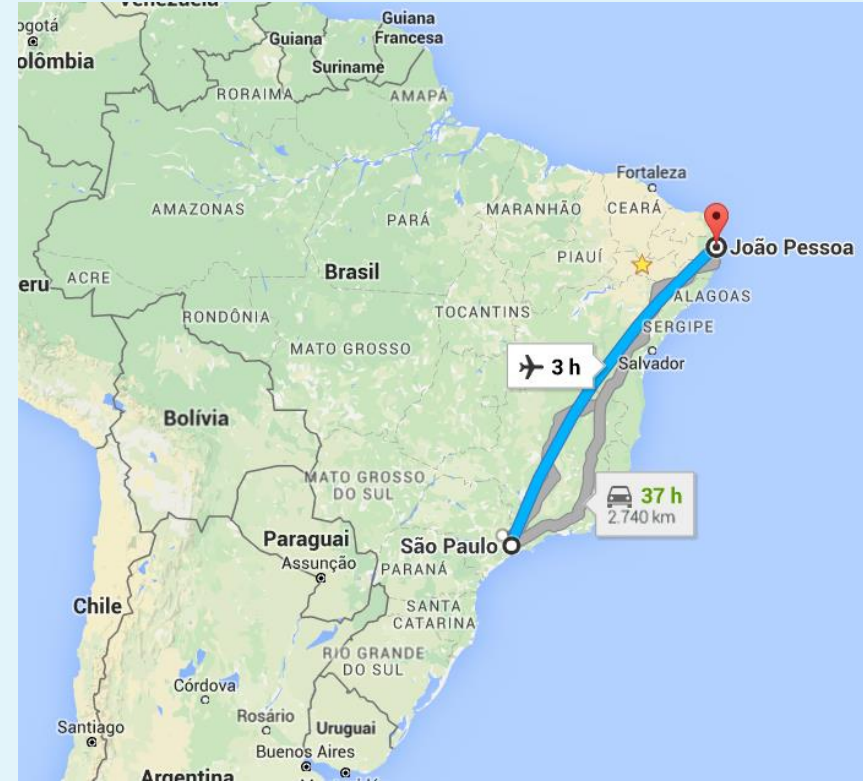
<http://buscatextual.cnpq.br/buscatextual/visualizacv.do?metodo=apresentar&id=K4751704U>

Research Gate:

https://www.researchgate.net/profile/Marcus_Scotti

Google Scholar:

<https://scholar.google.com.br/citations?user=jdVKCTgAAAAJ&hl=es>



Conducts studies in:

Computer-assisted drug design: QSAR, Structure-Base Virtual Screening and Ligand-based VS

Environmental Chemoinformatics

Natural Products Databases:

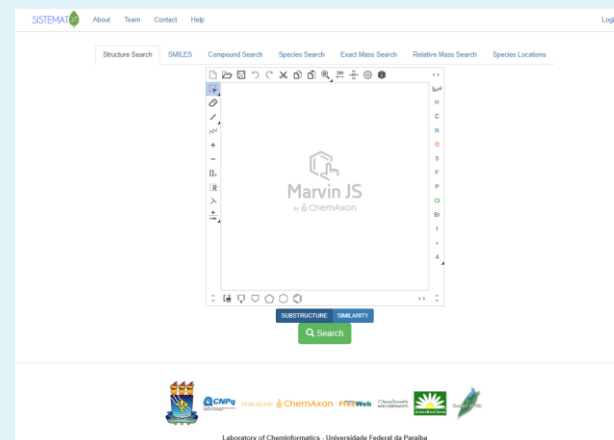


<https://sistemax.ufpb.br>

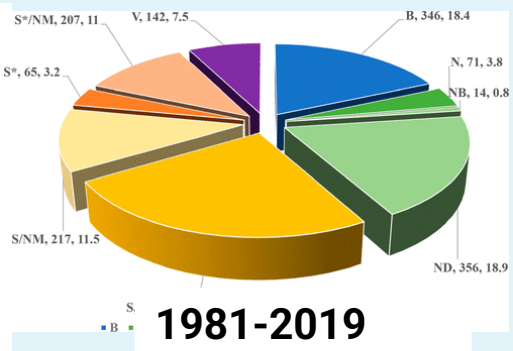
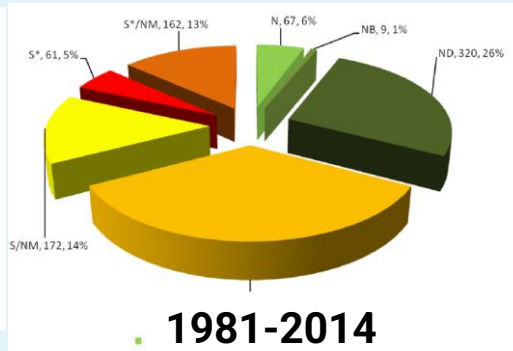
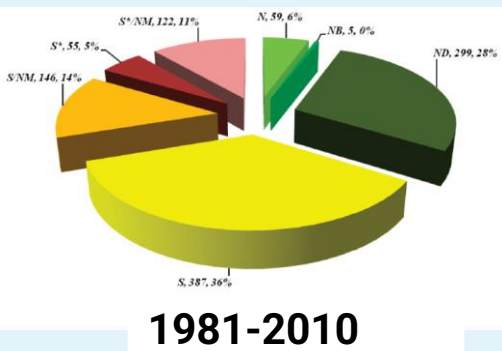
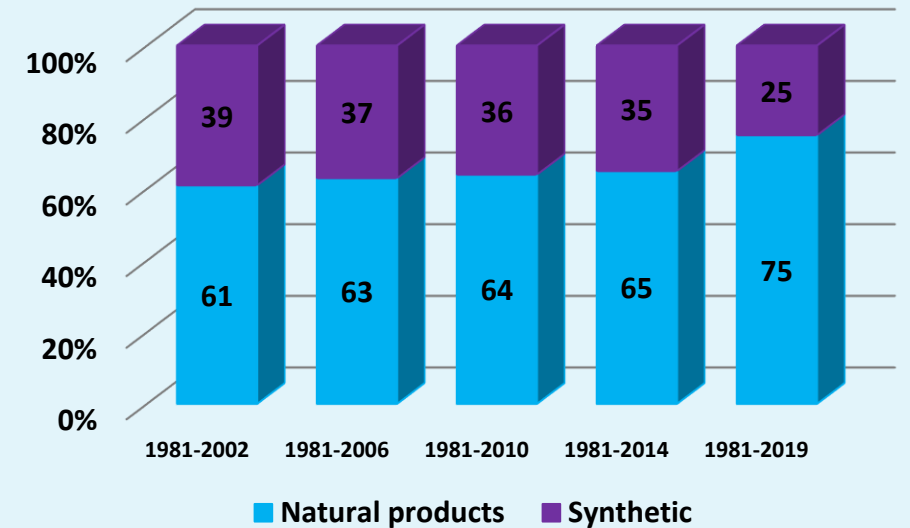
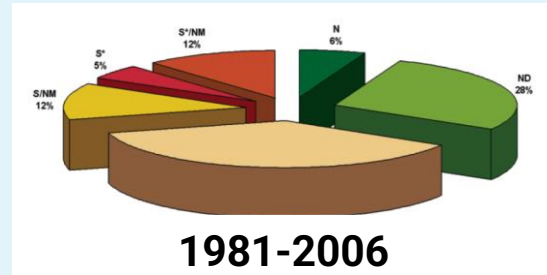
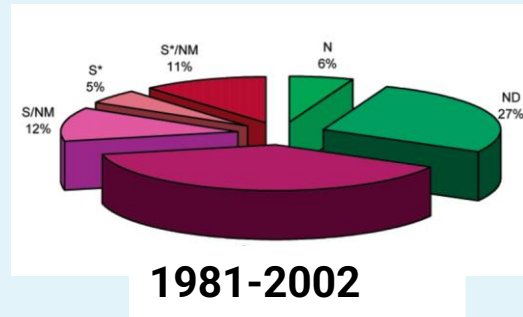
Molecules 2018, 23(1), 103; <https://doi.org/10.3390/molecules23010103> Open Access Technical Note

Sistemax, an Online Web-Based Cheminformatics Tool for Data Management of Secondary Metabolites

Marcus Tullius Scotti ^{1,*}  , Chonny Herrera-Acevedo ¹  , Tiago Branquinho Oliveira ^{2,3}  , Renan Paiva Oliveira Costa ¹ , Silas Yudi Konno de Oliveira Santos ¹, Ricardo Pereira Rodrigues ¹ , Luciana Scotti ¹   and Fernando Batista Da-Costa ³  



Natural products (NP) and their secondary metabolites are promising starting points for the development of prototypes and new drugs, being a large part of the new treatments against countless diseases, directly or indirectly related to them.



Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. David J. Newman and Gordon M. Cragg

Newman, David J.; Cragg, Gordon M.; Snader, Kenneth M. *J. Nat. Prod.*, 75, 1022-1037 (2003).

Newman, David J.; Cragg, Gordon M. *J. Nat. Prod.*, 70, 461-477 (2007).

Newman, David J.; Cragg, Gordon M. *J. Nat. Prod.*, 75, 311-335 (2012).

Newman, David J.; Cragg, Gordon M. *J. Nat. Prod.*, 79, 629-661 (2016).

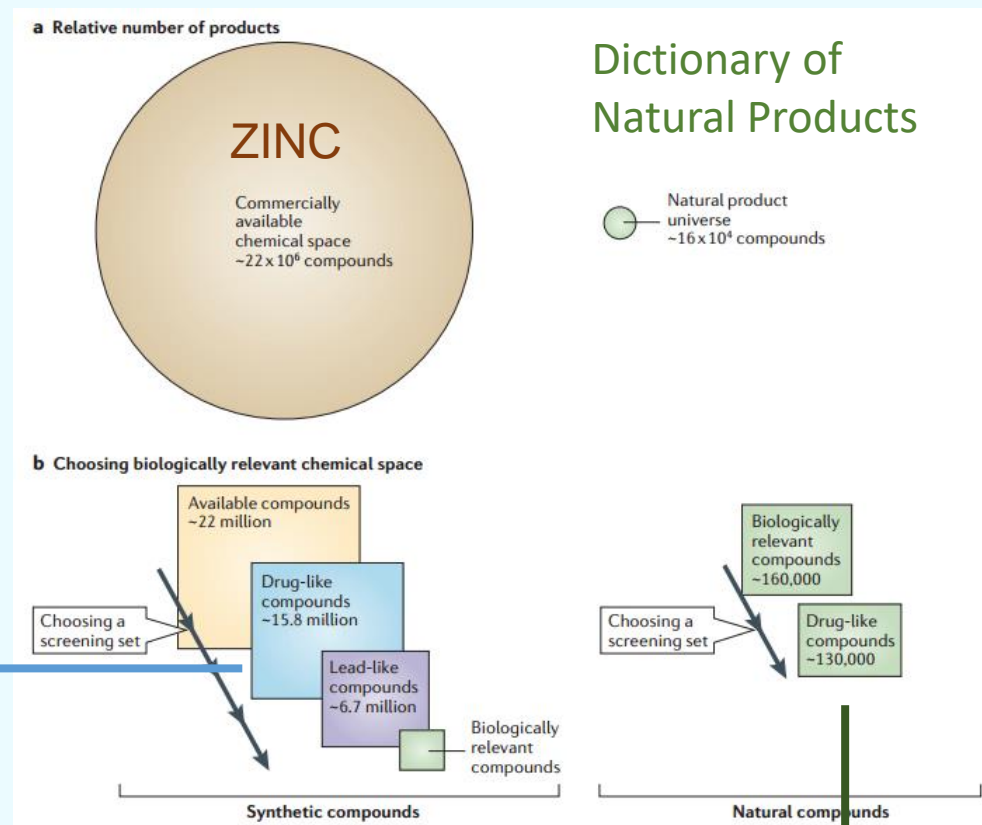
Natural Products

The re-emergence of natural products for drug discovery in the genomics era

Alan L. Harvey, RuAngelie Edrada-Ebel, Ronald J. Quinn

Nature Reviews Drug Discovery 14, 111–129 (2015)

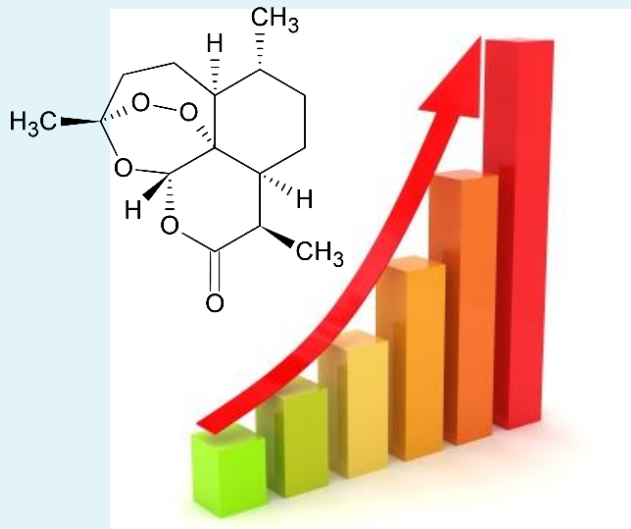
Database	Number of entries	Additional information
Super Natural III	355,000	2D structures; vendor information for over 215,000 compounds
Universal Natural Product Database	197,201	3D structures assembled from other available Chinese databases
Chinese Natural Product Database	53,000	Has been used in a virtual screen for PPAR- γ agonists
Drug Discovery Portal	40,000	Not all natural products, but all based on available samples
iSMART	20,000	Based on components from traditional Chinese medicines
Database from historical medicinal plants, DIOS	6,702	Successfully used in several virtual screening campaigns
AfroDb	1,000	Compounds from African medicinal plants
NuBBE	640	Compounds from Brazilian sources



71.8%

81.2%

Natural Products Databases



The increase in the number of new structures and studies related to biological activity reinforced the development of databases allowing significant support for drug discovery

The classification and organization of information is essential

The databases have provided a collection systematic information on natural products and their derivatives



Computational approaches have played an increasingly prominent role in natural product (NP)-based drug discovery.

For example, the development and use of NP databases allow access to numerous chemical, biological, pharmacological, toxicological, and structural NP data.

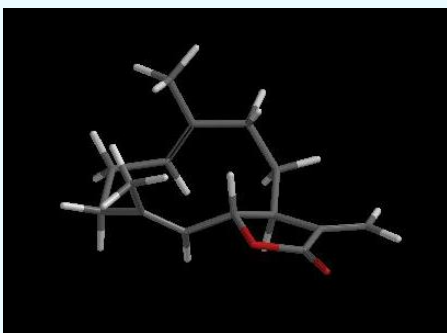
Sorokina and Steinbeck recently reviewed currently available NP databases, citing more than 120 examples, and noted that access to information was limited, and only a few of the databases were sustainably managed and continually developed.



2008

- JAVA
- MySQL database

LAN



Edição de Moléculas

Classe: Sesquiterpeno lactonizado (Nova...) Esqueleto: Germacranolide (Novo...)

Nome: COSTUNOLIDE (NOX: -16)

Nome Comum: COSTUNOLIDE (ID: 11998)

Espectro de Massa | Dados de RMN 1H | HMBC
Dados de RMN 13C | Outros Dados

Solvente: C13 (Novo...)

Átomo	Biogenética	Desloc.

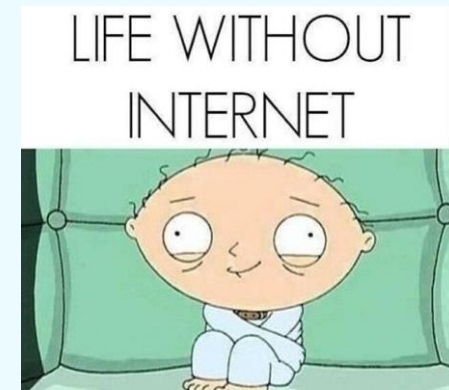
Carregar... | Editar | Exibir MDL... | ver 3D

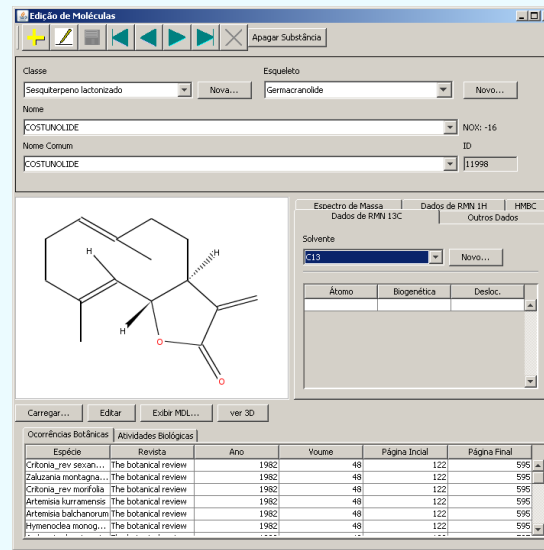
Ocorrências Botânicas | Atividades Biológicas

Espécie	Revista	Ano	Voume	Página Inicial	Página Final
Critonia_rev sexan...	The botanical review	1982	48	122	595
Zaluzania montagna...	The botanical review	1982	48	122	595
Critonia_rev morifolia	The botanical review	1982	48	122	595
Artemisia kurramensis	The botanical review	1982	48	122	595
Artemisia balchanorum	The botanical review	1982	48	122	595
Hymenoclea monog...	The botanical review	1982	48	122	595

Web tools/applications

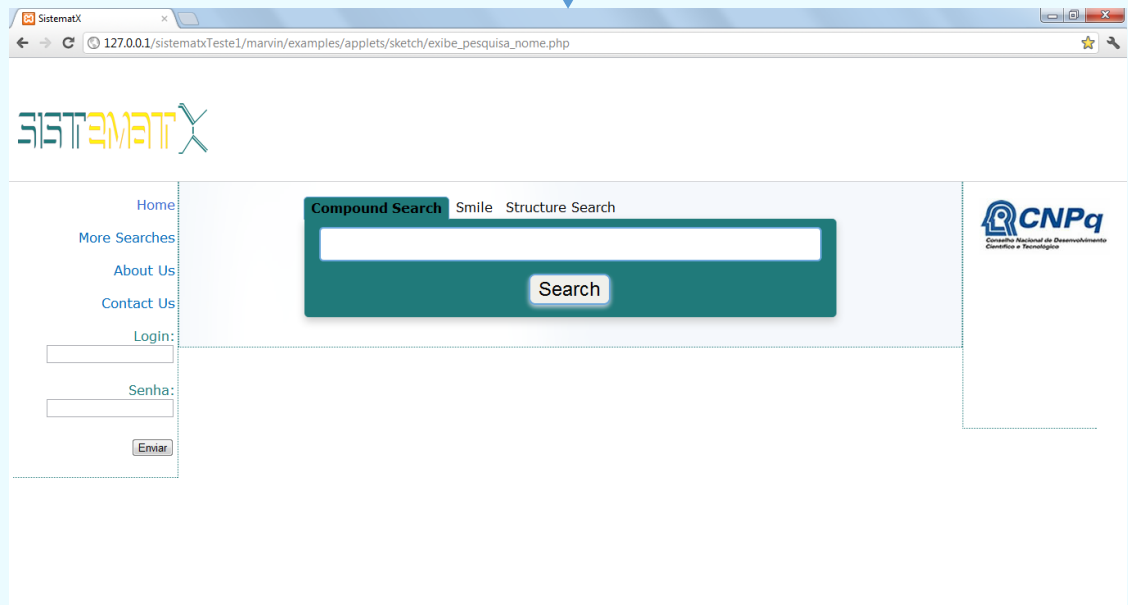
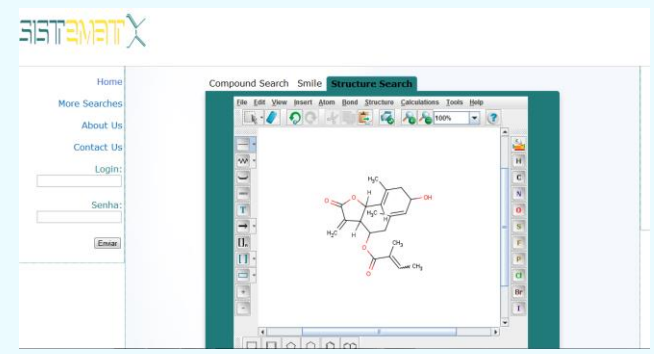
- Don't need to install
- Don't need a powerful computer (use a server)
- You can use in different devices
- Browser
- Internet (sometimes you don't have it)





Espécie	Revista	Ano	Volume	Página Inicial	Página Final
<i>Crinia yev sevan...</i>	The botanical review	1982	48	122	595
<i>Sakania montana...</i>	The botanical review	1982	48	122	595
<i>Crinia yev morfolo...</i>	The botanical review	1982	48	122	595
<i>Artemisia kurramen...</i>	The botanical review	1982	48	122	595
<i>Artemisia balcharo...</i>	The botanical review	1982	48	122	595
<i>Hymenoclea monog...</i>	The botanical review	1982	48	122	595

HTML
CSS
Java
JavaScript

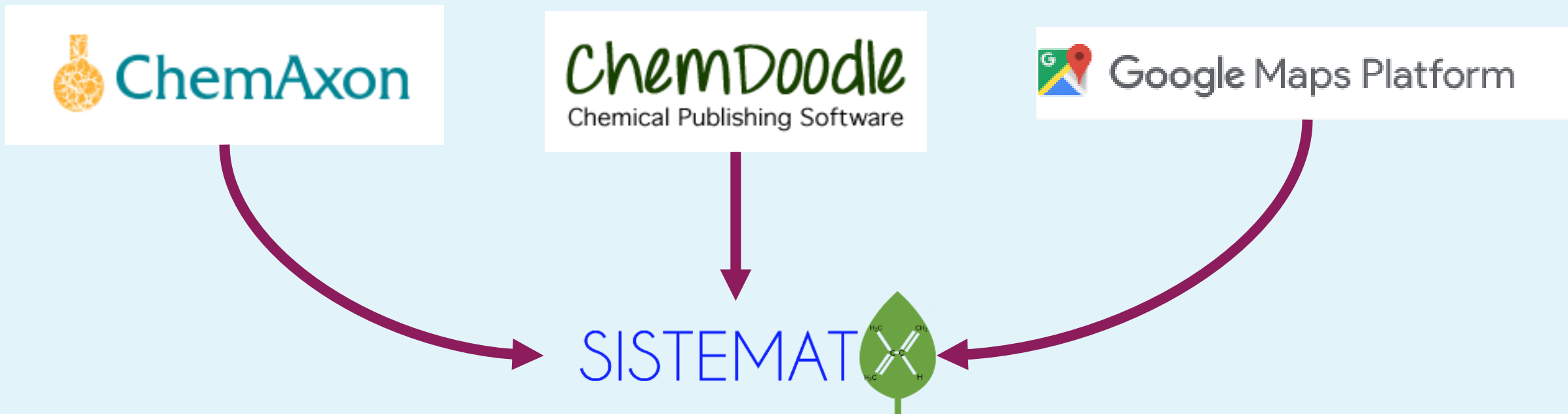



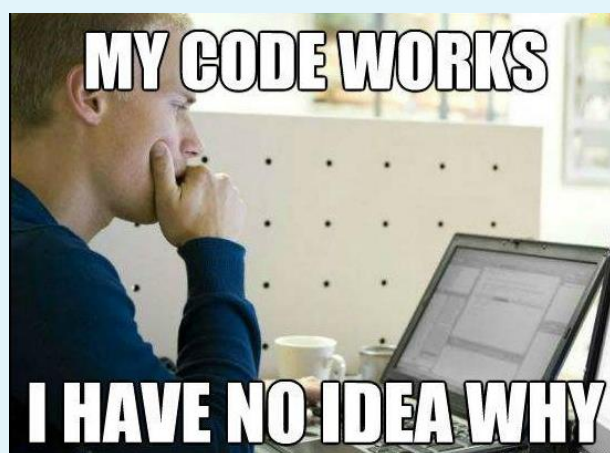
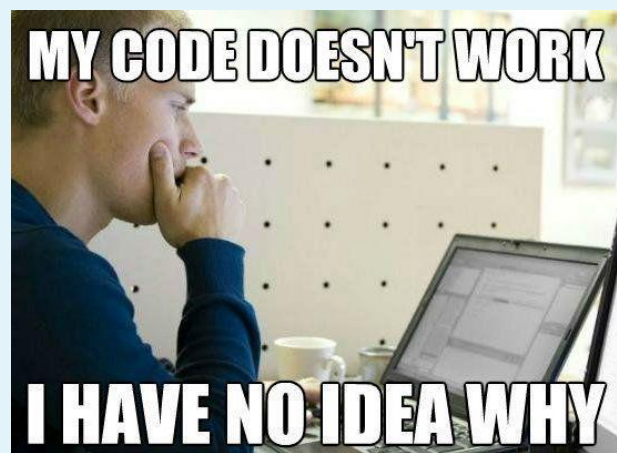
APIs

Several Application Programming Interface (APIs) are used in the Sistemax implementation.

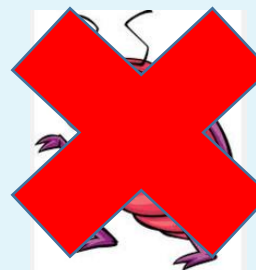
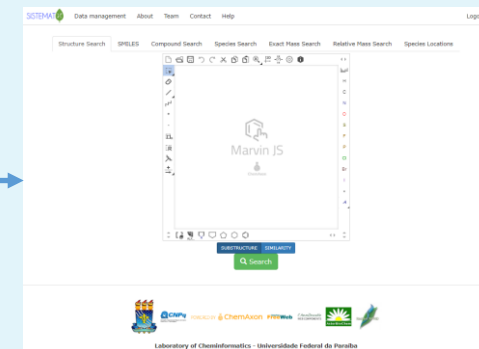
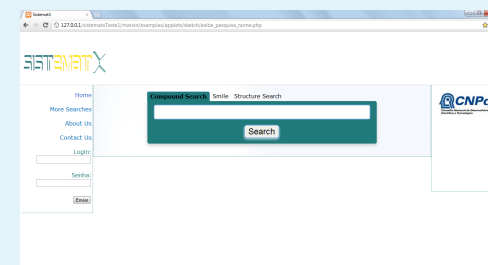
API is code that allows two software programs to communicate with each other.

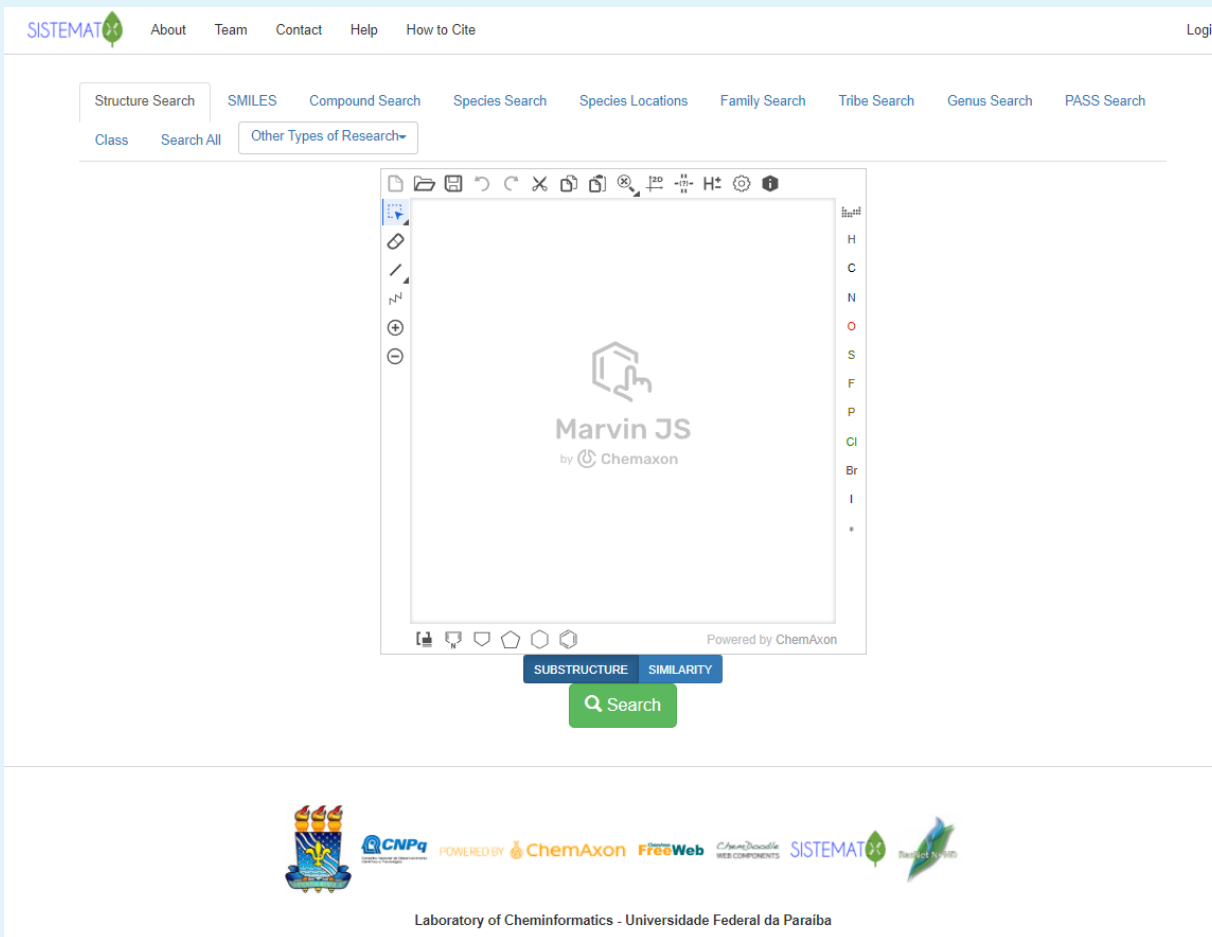
Application Programming Interface (API)





APIs





SISTEMATX About Team Contact Help How to Cite Login

Structure Search SMILES Compound Search Species Search Species Locations Family Search Tribe Search Genus Search PASS Search

Class Search All Other Types of Research

Marvin JS
by Chemaxon

Powered by ChemAxon

SUBSTRUCTURE SIMILARITY

Search

CNPq POWERED BY ChemAxon FreeWeb ChemDoodle WEB COMPONENTS SISTEMAT

Laboratory of Cheminformatics - Universidade Federal da Paraiba

sistematx.ufpb.br

register software: N^o: BR 51 2015 000073 0 e BR51 2016 000371 5

INPI - Instituto Nacional da Propriedade Industrial. (National Institute of Industrial Property.)

- TIPdb

Lin, Y. C.; Wang, C. C.; Chen, I. S.; Jheng, J. L.; Li, J. H.; Tung, C. W., TIPdb: a database of anticancer, antiplatelet, and antituberculosis phytochemicals from indigenous plants in Taiwan. *ScientificWorldJournal* 2013, 2013, 736386.
cwtung.kmu.edu.tw/tipdb/

- TCM database@Taiwan ,

Chen, C. Y., TCM Database@Taiwan: the world's largest traditional Chinese medicine database for drug screening in silico. *PLoS One* 2011, 6(1), e15939.
<http://tcm.cmu.edu.tw/>

- KNApSACk-3D

Nakamura, K.; Shimura, N.; Otabe, Y.; Hirai-Morita, A.; Nakamura, Y.; Ono, N.; Ul-Amin, M. A.; Kanaya, S., KNApSACk-3D: a three-dimensional structure database of plant metabolites. *Plant Cell Physiol.* 2013, 54(2), e4.
knapsack3d.sakura.ne.jp

- 3DMET

Maeda, M. H.; Kondo, K., Three-dimensional structure database of natural metabolites (3DMET): a novel database of curated 3Dstructures. *J. Chem. Inf. Model* 2013
www.3dmet.dna.affrc.go.jp/

- UNPD

Jiangyong Gu, Yuanshen Gui, Lirong Chen, Gu Yuan, Hui-Zhe Lu, Xiaojie Xu. Use of Natural Products as Chemical Library for Drug Discovery and Network Pharmacology. *PLoS ONE.* 2013, 8(4): e62839. doi:10.1371/journal.pone.0062839.

- NuBBE database

Valli, M.; dos Santos, R. N.; Figueira, L. D.; Nakajima, C. H.; Castro-Gamboa, I.; Andricopulo, A. D.; Bolzani, V. S., Development of a natural products database from the biodiversity of Brazil. *J. Nat. Prod.* 2013, 76(3), 439-44.
nubbe.iq.unesp.br/portal/nubbedb.html

Comparison among the queries






	SistematX	3DMET	NuBBE _{DB}	TCM Database @Taiwan	TIPdb	UNPD
Query by structure	YES	YES*	YES	YES*	NO	NO
Name	YES	YES	YES	YES	YES	YES
SMILES	YES	YES	YES	YES	NO	YES
CASRN	YES	YES	NO	NO	NO	YES
InChI	YES	YES	NO	NO	NO	YES
InChIKey	YES	NO	NO	NO	NO	YES
Molecular Formule	YES	YES	YES	YES	NO	YES
Class	YES	NO	YES	NO	YES	NO
Chemical information	YES	YES	YES	YES	YES	YES
Biological Activity	YES	NO	YES	YES	YES	NO
Botanical occurrence	YES	NO	YES	YES	YES	YES
Part of plant	YES	NO	NO	NO	YES	NO
Citação bibliográfica	YES	NO	YES	NO	NO	NO

2 years



Technical Note

SistematX, an Online Web-Based Cheminformatics Tool for Data Management of Secondary Metabolites

Marcus Tullius Scotti ^{1,*} , Chonny Herrera-Acevedo ¹ , Tiago Branquinho Oliveira ^{2,3} ,
Renan Paiva Oliveira Costa ¹, Silas Yudi Konno de Oliveira Santos ¹, Ricardo Pereira Rodrigues ¹,
Luciana Scotti ¹  and Fernando Batista Da-Costa ³ 

¹ Postgraduate Program in Natural Products and Synthetic Bioactive, IPeFarM, Federal University of Paraíba, Campus I, Cidade Universitária, João Pessoa 58051-900, PB, Brazil; chonny622@gmail.com (C.H.-A.); renan0paiva@hotmail.com (R.P.O.C.); syudik12@gmail.com (S.Y.K.d.O.S.); ricardo.p.rodrigues@ufes.br (R.P.R.); luciana.scotti@gmail.com (L.S.)

² Department of Pharmacy, Federal University of Sergipe (UFS-SE), Av. Marechal Rondon s/n, Jd. Rosa Elze, São Cristóvão 49100-000, SE, Brazil; tiago.branquinho@ufs.br

³ AsterBioChem Research Team, Laboratory of Pharmacognosy, School of Pharmaceutical Sciences of Ribeirão Preto, University of São Paulo (USP), Av. do Café s/n, Ribeirão Preto 14040-903, SP, Brazil; febcosta@fcfrp.usp.br

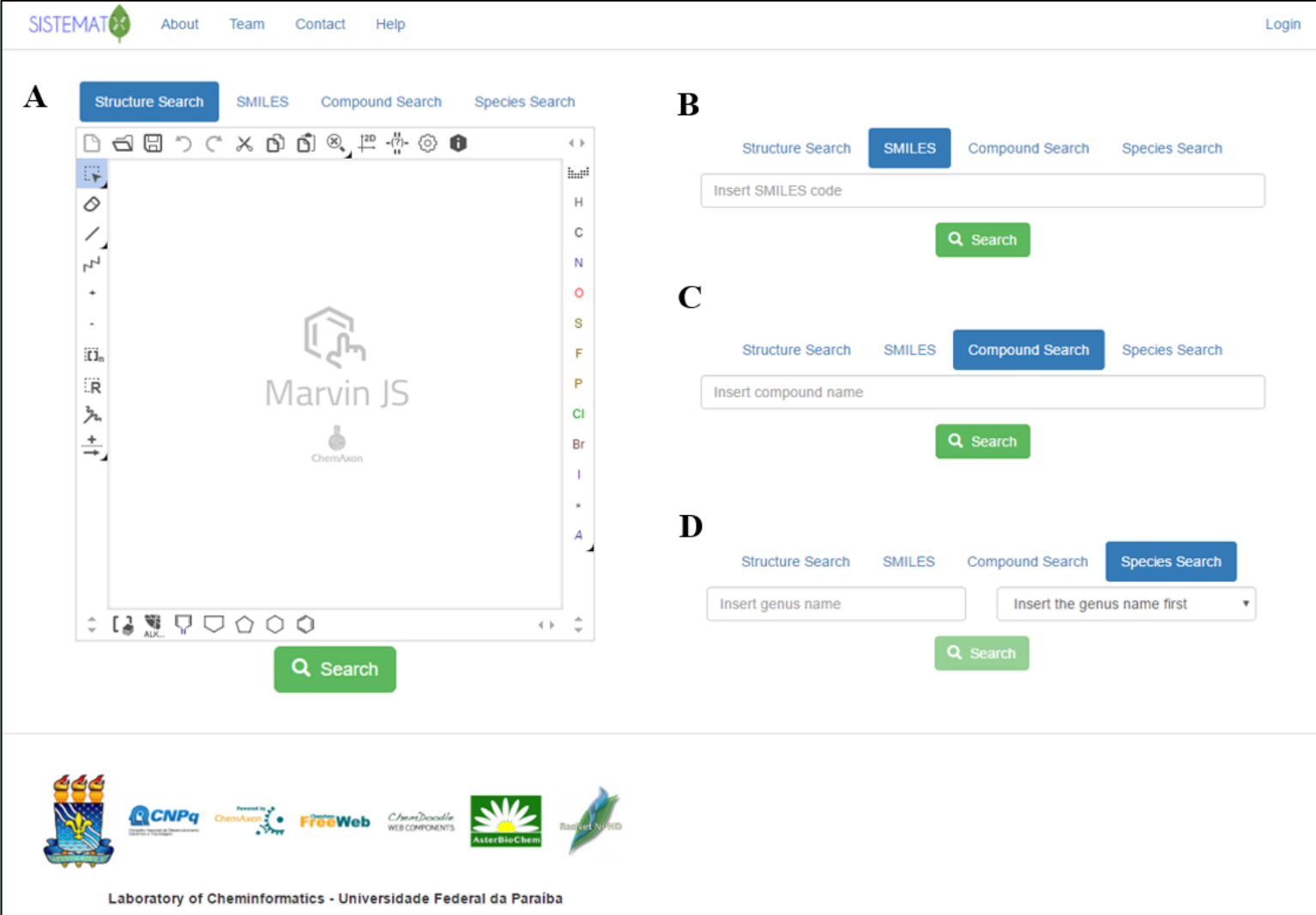
* Correspondence: mtscotti@gmail.com; Tel.: +55-83-99869-0415

Received: 18 November 2017; Accepted: 28 December 2017; Published: 3 January 2018

After the user enters the website, the “Structure search” option is seen with the MarvinJS API (Application Programming Interface) at the top of the screen.

Other search options can be exhibited in the interface.

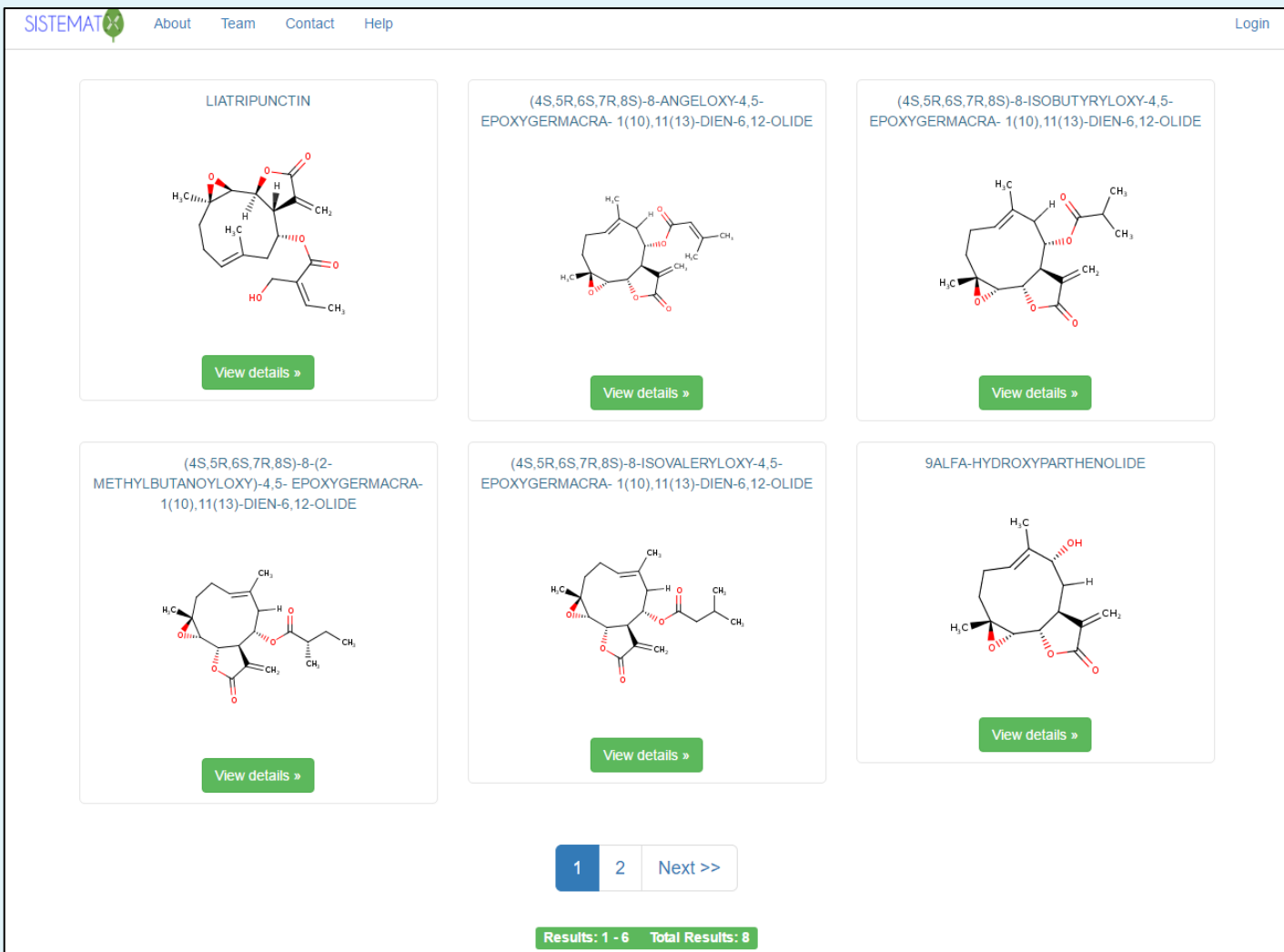
The initial screen of the system also shows the SMILES (Simplified Molecular-Input Line-Entry System) code (B), compound name (C) and plant species search modes (D).



The screenshot displays the SistematX web interface with four search modes labeled A, B, C, and D. At the top, there is a navigation bar with the SistematX logo and links for About, Team, Contact, and Help, along with a Login button. The main content area is divided into four sections:

- A:** MarvinJS interface. It features a toolbar with various drawing tools, a central canvas with the MarvinJS logo and ChemAxon branding, and a vertical element list on the right containing H, C, N, O, S, F, P, Cl, Br, I, and A. A green Search button is located at the bottom.
- B:** SMILES search mode. It has tabs for Structure Search, SMILES (selected), Compound Search, and Species Search. Below the tabs is a text input field labeled "Insert SMILES code" and a green Search button.
- C:** Compound Search mode. It has tabs for Structure Search, SMILES, Compound Search (selected), and Species Search. Below the tabs is a text input field labeled "Insert compound name" and a green Search button.
- D:** Species Search mode. It has tabs for Structure Search, SMILES, Compound Search, and Species Search (selected). Below the tabs are two input fields: "Insert genus name" and "Insert the genus name first" (with a dropdown arrow), and a green Search button.

At the bottom of the page, there is a footer with logos for CNPq, ChemAxon, FiesWeb, ChemDoodle WEB COMPONENTS, AsterBioChem, and SisematX, followed by the text "Laboratory of Cheminformatics - Universidade Federal da Paraiba".



The screenshot shows the SistematX database search results page. At the top, there is a navigation bar with the SistematX logo, links for 'About', 'Team', 'Contact', and 'Help', and a 'Login' button. The main content area displays six search results, each in a white box with a green border. Each result includes the compound name, its chemical structure, and a 'View details »' button. The results are:

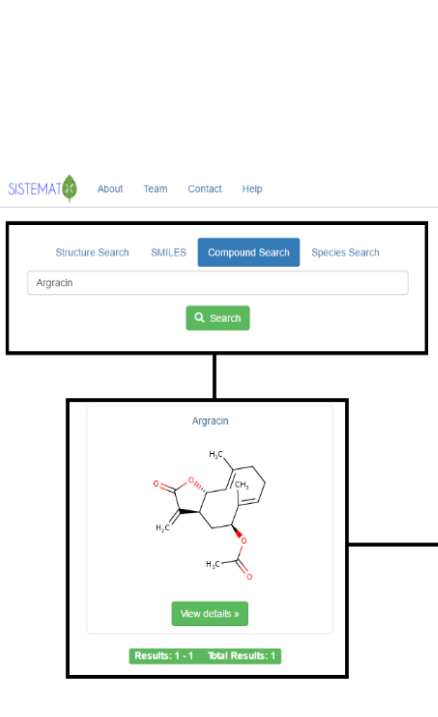
- LIATRIPUNCTIN**: CC1=C(C)C(=O)C(O)C(=O)C1
- (4S,5R,6S,7R,8S)-8-ANGELOXY-4,5-EPOXYGERMACRA- 1(10),11(13)-DIEN-6,12-OLIDE**: CC1=C(C)C(=O)C(O)C(=O)C1
- (4S,5R,6S,7R,8S)-8-ISOBUTYRYLOXY-4,5-EPOXYGERMACRA- 1(10),11(13)-DIEN-6,12-OLIDE**: CC1=C(C)C(=O)C(O)C(=O)C1
- (4S,5R,6S,7R,8S)-8-(2-METHYLBUTANOYLOXY)-4,5-EPOXYGERMACRA- 1(10),11(13)-DIEN-6,12-OLIDE**: CC1=C(C)C(=O)C(O)C(=O)C1
- (4S,5R,6S,7R,8S)-8-ISOVALERYLOXY-4,5-EPOXYGERMACRA- 1(10),11(13)-DIEN-6,12-OLIDE**: CC1=C(C)C(=O)C(O)C(=O)C1
- 9ALFA-HYDROXYPARTHENOLIDE**: CC1=C(C)C(=O)C(O)C(=O)C1

At the bottom of the page, there is a pagination control with buttons for '1', '2', and 'Next >>'. Below the pagination, a green bar indicates 'Results: 1 - 6 Total Results: 8'.

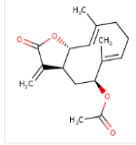
When performing a search, the mechanism generates a results page (six results) using common names; if the compound does not have one, it shows the IUPAC name.

The user can set the number of structure results per page. When a result is selected, the user has access to the data for that compound.

SistematX database



Argracin



[← Back](#) [Amplify](#)

Compound Identification	
Common Name	Argracin
SMILES	<chem>CC(=O)C[C@@H]1C[C@@H]2[C@@H](OC(=O)C2=C)C=C(C)C[C@@H]1C</chem>
IUPAC	(3aS,5S,11aR)-6,10-dimethyl-3-methylenedene-2-oxo-2H,3H,3aH,4H,5H,8H,9H,11aH-cyclodeca[b]furan-5-yl acetate
InChI	1S:C17H22O4:C1-10-6-5-7-11(2)15(20-13(4)18)9-14-12(3)17(19)21-16(14)8-10/17-6,14-16H,3,5-6,9H2,1-2,4H3/b10-6,7+11/4,15-16+10/s1Auxinfo=1/0/N:16,21,13,1,18,17,19,14,6,15,20,12,2,7,6,8,10,3,11,4,9//imvA:21CCOOC.cCC.cC14.s15.s15.s17.s18.s5d+19.s20;/C:.....
InChIKey	DPXBPFOESQDXEOD-YXXPDWHBSA-N
CAS	

Compound Data				
Class	Skeleton	NOX	Exact Mass	Relative Mass
Sesquiterpene lactone	Germacranolide	-14	290.151809188	290.359

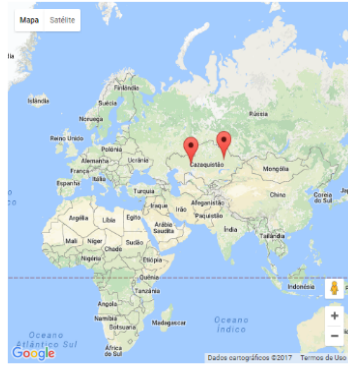
Botanical Data						
Family	Subfamily	Tribe	Subtribe	Genus	Species	Reference
Asteraceae	Unknown	Anthemideae	Artemisinae	Artemisia	aralensis	Chem. Nat. Comp., v 52, p. 417. 2016
Asteraceae	Unknown	Anthemideae	Artemisinae	Artemisia	gracilescens	Chem. Nat. Comp., v 28, p. 387. 1992

Biological Activity

Empty biological activity data.

Geographic Data

Species	Latitude	Longitude	Approximate Location
Artemisia aralensis	47.8509357	59.5757239	Unnamed Road, Shalkar, Cazaquistão
Artemisia gracilescens	49.5059696	76.247557	Unnamed Road, 100000, Cazaquistão




Integrated development environment (IDE)

API	Description	Engine
1. Structure		
2D drawing	Allows drawing and visualization of chemical structures	ChemAxon
3D generator	Uses 2D drawing to generate a 3D representation of the molecule	ChemAxon
3D	Graphical visualization of 3D molecules with JavaScript	ChemDoodl
2. Compound Identification		
SMILES	Simplified Molecular Input Line Entry System	ChemAxon
IUPAC	IUPAC Nomenclature	ChemAxon
InChI	IUPAC International Chemical Identifier	ChemAxon
InChIKey	InChIKey is a compact format of the InChI code	ChemAxon
CAS	Chemical Abstracts Service Registry Number	ChemAxon
3. Compound Data		
NOX	Oxidation number (NOX) of an organic compound	ChemAxon
Exact Mass	Uses the mass of the most abundant isotope of each element	ChemAxon
Relative Mass	Uses the average atomic mass of each element	ChemAxon
4. Geographic data		
Latitude	Can be inserted by the administrator or appears by clicking in the world map	Google Inc.
Longitude	Can be inserted by the administrator or appears by clicking in the world map	Google Inc.
Approximate	Using the latitude and longitude, appears an an approximate location of the specie	Google Inc.
Visualization	Uses the world map to possible to visualize the localization of the species	Google Inc.

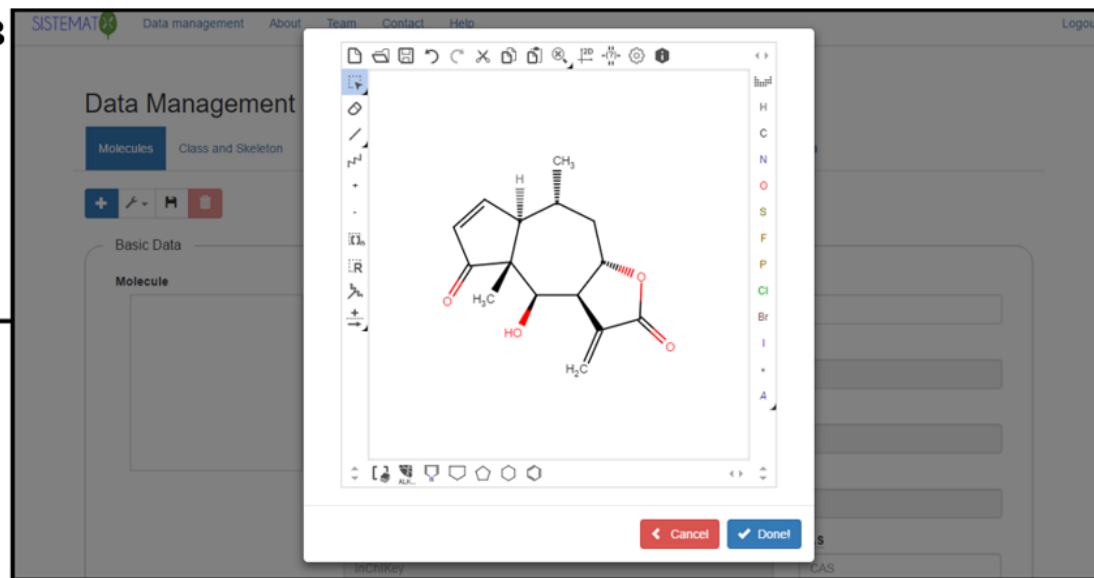
SistematX database

On the SistematX homepage, the user can also log into the data management area using login name and password and from there access the administration pages to edit or register new molecules. Once the corresponding information has been accepted, the data management interface appears.

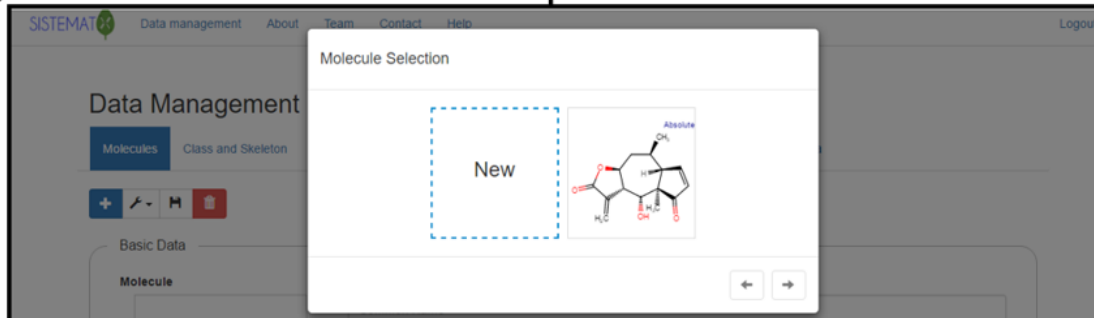
A



B



C



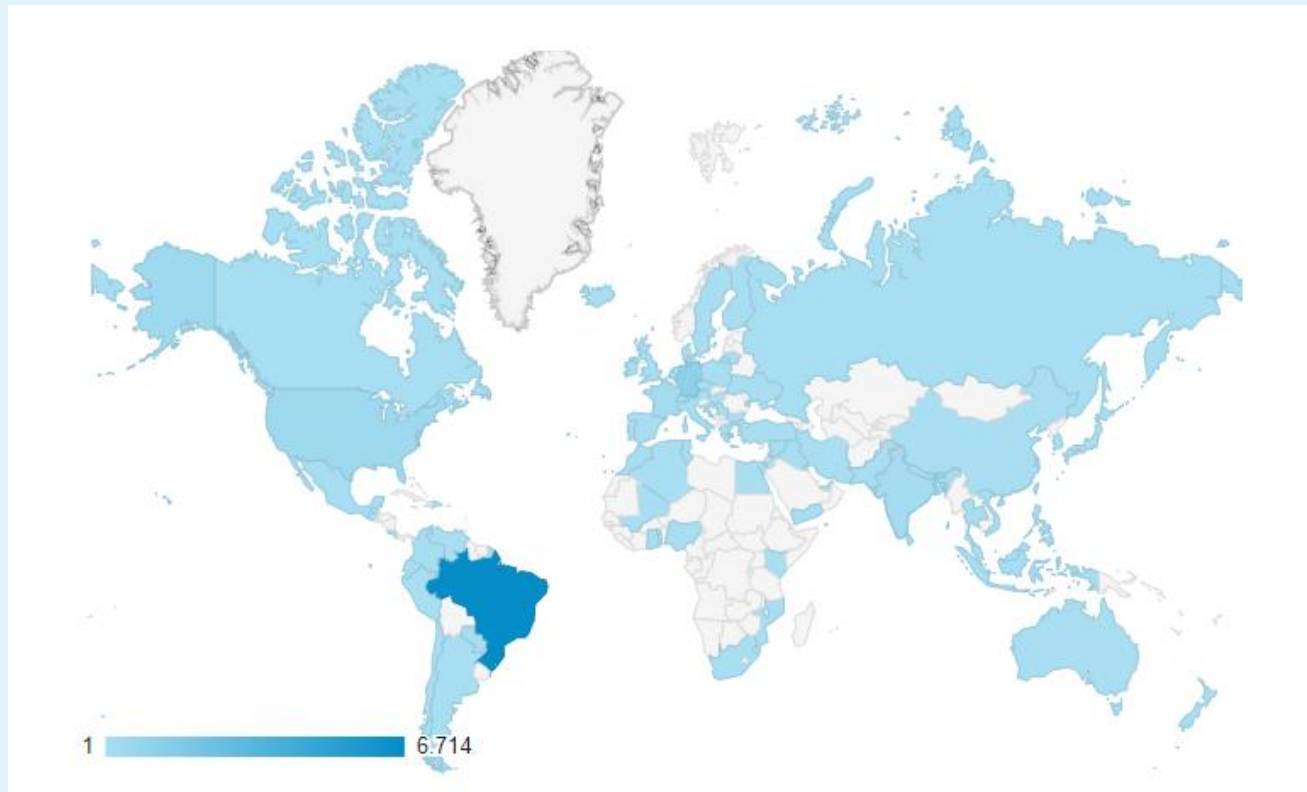
SistematX database

SistematX currently has 9,514 unique secondary metabolites arising from 20,934 botanical occurrences across 5 families:

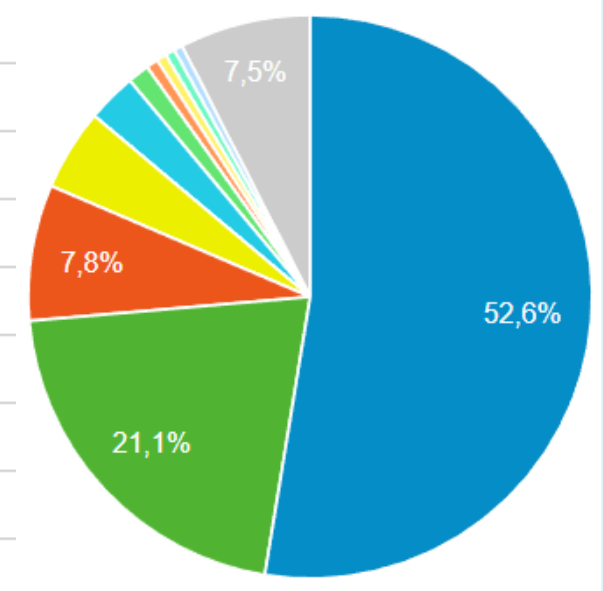
Botanical family	Secondary metabolites	Ocurrences
Asteraceae	2,574	7,879
Apocynaceae	372	620
Annonaceae	1,898	5,335
Lamiaceae	4,097	6,357
Solanaceae	573	743

Sistemax database

Since 2018, Sistemax has reported 5251 users and more than 9200 sessions (47.4% located outside Brazil)



1.  Brazil
2.  Germany
3.  United States
4.  France
5.  India
6.  Colombia
7.  United Kingdom
8.  Russia
9.  Italy
10.  Mexico



Molecular Diversity
<https://doi.org/10.1007/s11030-021-10245-z>

ORIGINAL ARTICLE



Machine learning models to select potential inhibitors of acetylcholinesterase activity from SistemX: a natural products database





Chonny Herrera-Acevedo^{1,4}  · Camilo Perdomo-Madrigal²  · Kenyi Herrera-Acevedo³  · Ericsson Coy-Barrera⁴  · Luciana Scotti¹  · Marcus Tullius Scotti¹ 

Molecular Diversity
<https://doi.org/10.1007/s11030-020-10139-6>

ORIGINAL ARTICLE



Selection of antileishmanial sesquiterpene lactones from SistemX database using a combined ligand-/structure-based virtual screening approach

Chonny Herrera-Acevedo^{1,2}  · Mayara Dos Santos Maia¹ · Élide Batista Vieira Sousa Cavalcanti¹ · Ericsson Coy-Barrera²  · Luciana Scotti¹  · Marcus Tullius Scotti¹ 



DOI: 10.1002/cmdc.201700743

CHEM MED CHEM
Full Papers



In Silico Studies Designed to Select Sesquiterpene Lactones with Potential Antichagasic Activity from an In-House Asteraceae Database

Chonny Herrera Acevedo, Luciana Scotti, and Marcus Tullius Scotti^{*,[a]}



RESEARCH ARTICLE

Current Topics in Medicinal Chemistry, 2020, 20, 2126-2145

Ligand and Structure-based Virtual Screening of Lamiaceae Diterpenes with Potential Activity against a Novel Coronavirus (2019-nCoV)

Gabriela Cristina Soares Rodrigues¹, Mayara dos Santos Maia¹, Renata Priscila Barros de Menezes¹, Andreza Barbosa Silva Cavalcanti¹, Natália Ferreira de Sousa¹, Érika Paiva de Moura¹, Alex France Messias Monteiro¹, Luciana Scotti¹ and Marcus Tullius Scotti^{1,*}

Chagas Diseases, Leishmaniasis, COVID-19, Alzheimer

Sistemax

Chibli et al. profiled over 2000 metabolites from the Asteraceae family while screening for inhibitors of *Leishmania major* dihydroorotate dehydrogenase


Described the challenges of metabolomics methods in a summary of more than 95 metabolomics tools, software programs, and databases, highlighting Sistemax for its use of the Google Maps application program interface (API) as a locational system.

Mishra and Mohapatra noted that Sistemax was an easy tool for metabolomic studies.

The authors criticized the inability to perform bulk downloads in the Sistemax environment, which limited their ability to assess its completeness and cross-reference it with other data.

They suggested that the development of a tool to download publicly available digital data would be beneficial to the field.

Metabolomics (2019) 15:59
<https://doi.org/10.1007/s11306-019-1520-7>

ORIGINAL ARTICLE 

Untargeted LC–MS metabolomic studies of Asteraceae species to discover inhibitors of *Leishmania major* dihydroorotate dehydrogenase

Lucas A. Chibli¹ · Annylory L. Rosa¹ · Maria Cristina Nonato² · Fernando B. Da Costa¹

Quim. Nova, Vol. 43, No. 3, 329–354, 2020 <http://dx.doi.org/10.21577/0100-4042.20170499>

METABOLÔMICA DE PLANTAS: MÉTODOS E DESAFIOS

Alan C. Pilon^{a,b,c,d}, Denise M. Selegato^{b,d,e}, Richard P. Fernandes^b, Paula C. P. Bueno^{a,e}, Danielle R. Pinho^a, Fausto Carnevale Neto^f, Rafael T. Freire^a, Ian Castro-Gamboa^b, Vanderlan S. Bolzan^g e Norberto P. Lopes^{a,h,*}

^aNúcleo de Apoio à Pesquisa em Produtos Naturais e Sintéticos, Faculdade de Ciências Farmacêuticas de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto – SP, Brasil

Electrophoresis 2019, 40, 227–246 227 

Biswapriya B. Misra¹
Subhashree Mohapatra²

Review
Tools and resources for metabolomics research community: A 2017–2018 update

¹Department of Internal Medicine, Section of Molecular Medicine, Medical Center Boulevard, Winston-Salem, NC, USA
²Independent Researcher, 151 Edgeway Drive, Winston-Salem, NC, USA

The scale at which MS- and NMR-based platforms generate metabolomics datasets for both research, core, and clinical facilities to address challenges in the various sciences—ranging from biomedical to agricultural—is underappreciated. Thus, metabolomics efforts regarding methods, equipment, data, and human systems have led to continued

JBC ARTICLE 

Author's Choice

A database-driven approach identifies additional diterpene synthase activities in the mint family (Lamiaceae)

Received for publication, September 27, 2018, and in revised form, November 12, 2018. Published, Papers in Press, November 29, 2018. DOI 10.1074/jbc.RA118.006025

Sean R. Johnson^{1,1}, Wajid Waheed Bhat^{1,2}, Jacob Bibik¹, Aiko Turmo¹, Britta Hamberger¹, Evolutionary Mint Genomics Consortium¹, and Björn Hamberger^{1,3}

From the Departments of ¹Biochemistry and Molecular Biology and ²Pharmacology and Toxicology, ³Michigan State University, East Lansing, Michigan 48824

Edited by John M. Denu

2021: an update




pubs.acs.org/jcim

Application Note


The SistemátX Web Portal of Natural Products: An Update


Renan P. O. Costa, Lucas F. Lucena, Lorena Mara A. Silva, Guilherme Julião Zocolo, Chonny Herrera-Acevedo, Luciana Scotti, Fernando Batista Da-Costa, Nikita Ionov, Vladimir Poroikov, Eugene N. Muratov, and Marcus T. Scotti*


 Cite This: <https://doi.org/10.1021/acs.jcim.1c00083>

 Read Online

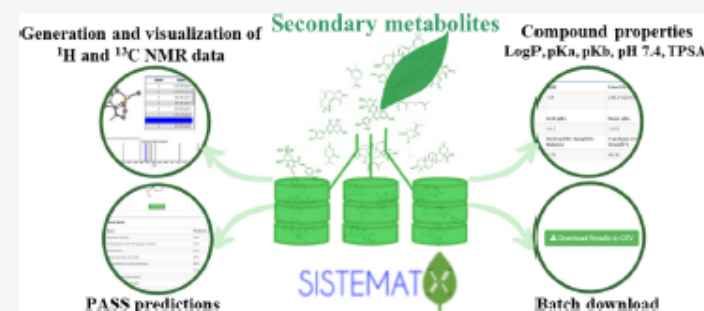
ACCESS |

 Metrics & More

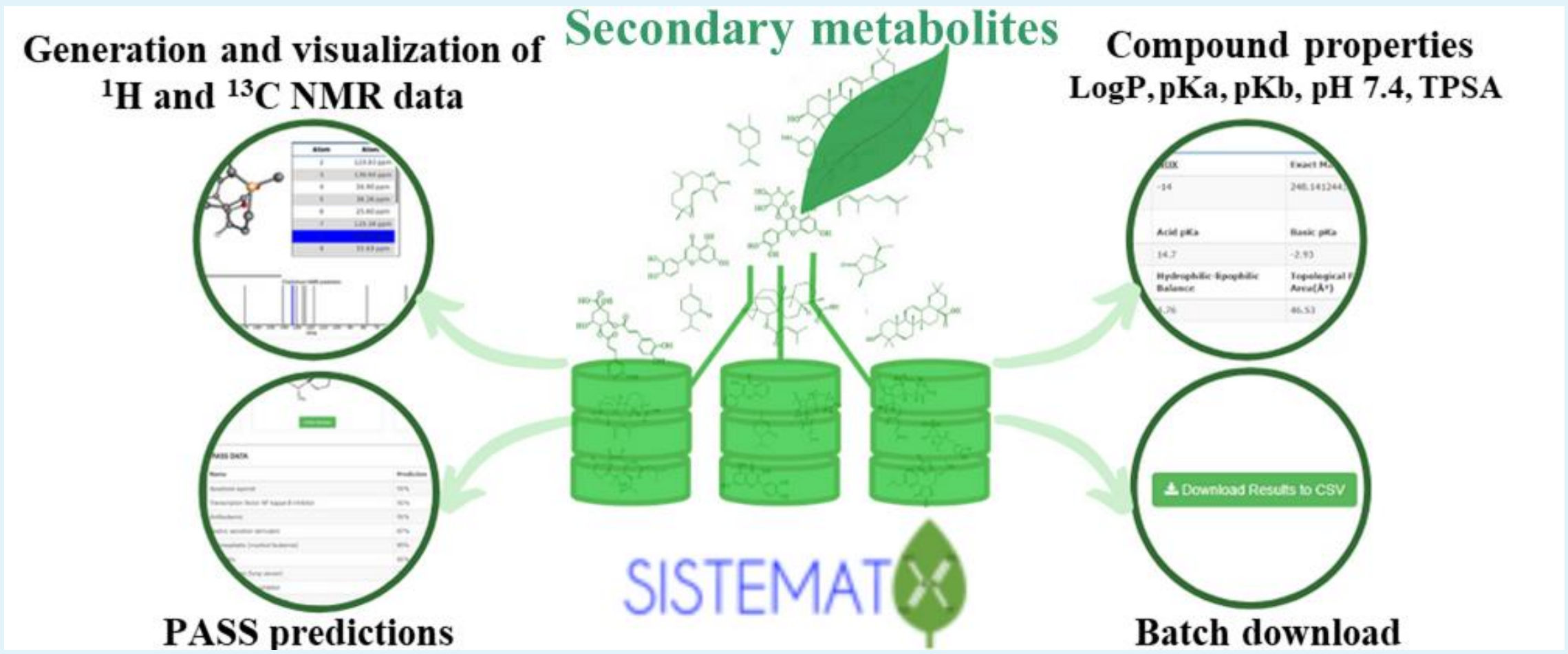
 Article Recommendations

 Supporting Information

ABSTRACT: Natural products and their secondary metabolites are promising starting points for the development of drug prototypes and new drugs, as many current treatments for numerous diseases are directly or indirectly related to such compounds. State-of-the-art, curated, integrated, and frequently updated databases of secondary metabolites are thus highly relevant to drug discovery. The SistemátX Web Portal, introduced in 2018, is undergoing development to address this need and documents crucial information about plant secondary metabolites, including the exact location of the species from which the compounds were isolated. SistemátX also allows registered users to log in to the data management area and gain access to administrative pages. This study reports recent updates and modifications to the SistemátX Web Portal, including a batch download option, the generation and visualization of ^1H and ^{13}C nuclear magnetic resonance spectra, and the calculation of physicochemical (drug-like and lead-like) properties and biological activity profiles. The SistemátX Web Portal is freely available at <http://sistemátx.ufpb.br>.



What's the new?



Compound properties

Summary of Application Programming Interface Calculated Physicochemical (Drug-Like and Lead-Like) Properties Recently Implemented in SistemataX

API	description
$\log P$	octanol–water partition coefficient
pK_a	decimal cologarithm of the acid dissociation constant
pK_b	decimal cologarithm of the base dissociation constant
solubility (pH 7.4)	maximum quantity of a solute that can be dissolved in a pH 7.4 solution
intrinsic solubility	solubility of a compound in its free acid or free base form
TPSA	sum of the surfaces of polar atoms in a molecule
hydrophilic–lipophilic Balance	measurement determining a compound's hydrophilicity or lipophilicity

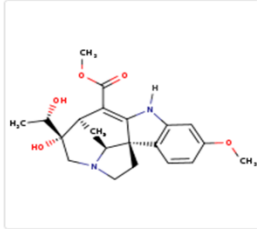
Compound properties


Sistemax implemented new features to meet the demands of the scientific community, users, and our partners.


Physicochemical (druglike and lead-like) properties, specifically logP, pKa, pKb, pH 7.4 solubility, intrinsic solubility, TPSA, and hydrophilic– lipophilic balance, are now calculated for compounds in the database


ALSTOVINE


← Back
Amplify





 NMR 1H Data

 NMR 1H by SPINUS WEB

 NMR 13C Data

 PASS Data

Compound Identification	
Common Name	ALSTOVINE
SMILES	<chem>[H]N1C2=C(C=CC(OC)=C2)[C@]23CCN4C[C@@](O)([C@H](C)O)[C@H](C[C@@H]24)C(C(=O)OC)=C13</chem>
IUPAC	methyl (1R,11R,12R,17S)-12-hydroxy-12-[(1S)-1-hydroxyethyl]-5-methoxy-8,14-diazapentacyclo[9.5.2.0A^1.?.0A^2.?.0A^1?.A^1?]octadeca-2(7),3,5,9-tetraene-10-carboxylate
InChI	1/C21H26N2O5/c1-11(24)21(26)10-23-7-6-20-13-5-4-12(27-2)8-15(13)22-18(20)17(19(25)28-3)14(21)9-16(20)23/h4-5,8,11,14,16,22,24,26H,6-7,9-10H2,1-3H3
InChIKey	NFIFPWIZYHMYPQ-UHFFFAOYNA-N
CAS	

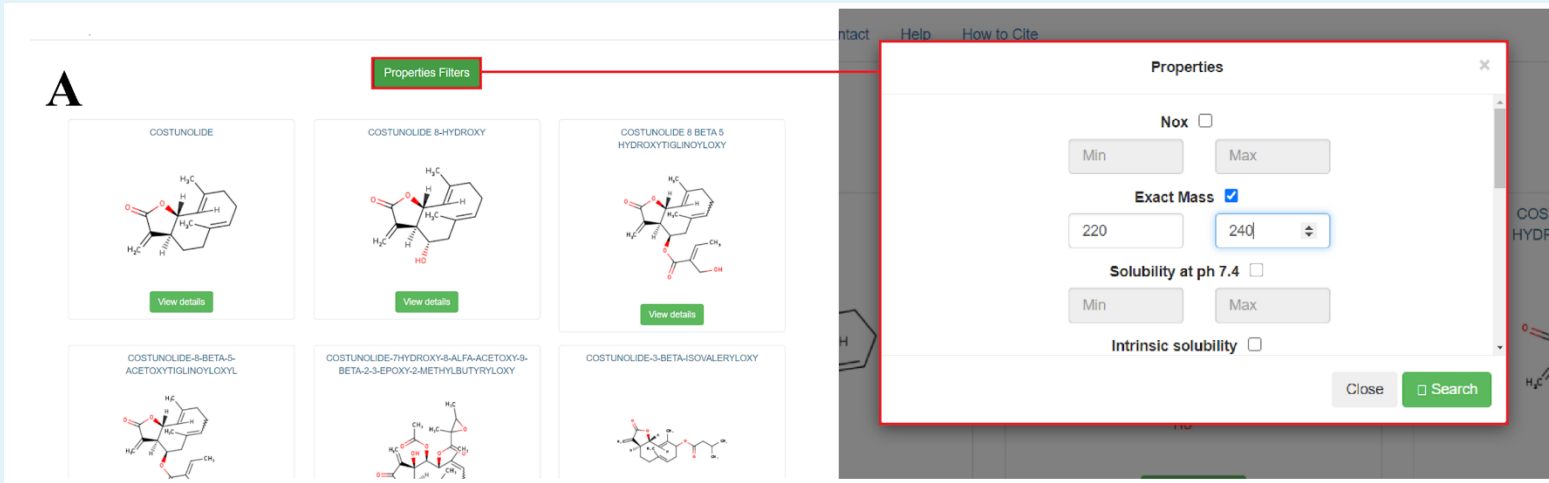
Compound Data			
Class	Skeleton	NOX	Exact Mass 
Alkaloid	Indole	-10	386.184171945
Solubility at pH 7.4	Intrinsic Solubility	Acid pKa	Basic pKa
0.782	-1.18	12.96	9.36
Relative Mass 	LogP	Hydrophilic-lipophilic Balance	Topological Polar Surface Area(Å²)
386.448	-0.06	18.74	91.26

Properties filters

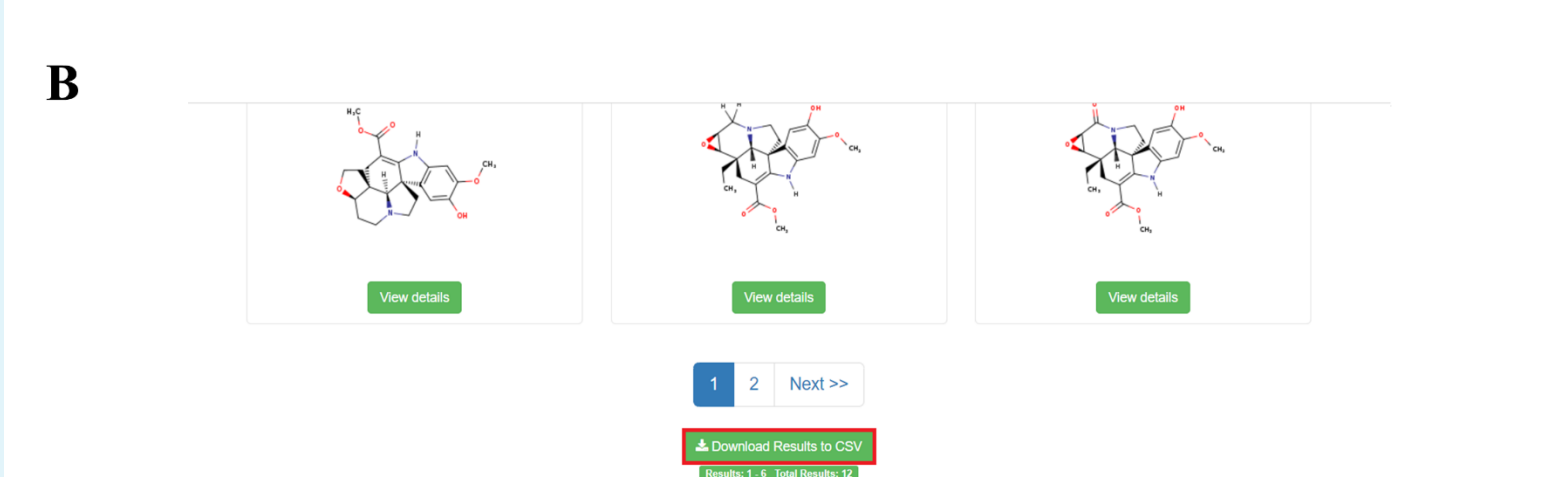
The results page for the molecules of interest now contains new search filters, which are accessible through a new user-responsive button that displays a modal pop-up window

A

Properties Filters



B



1 2 Next >>

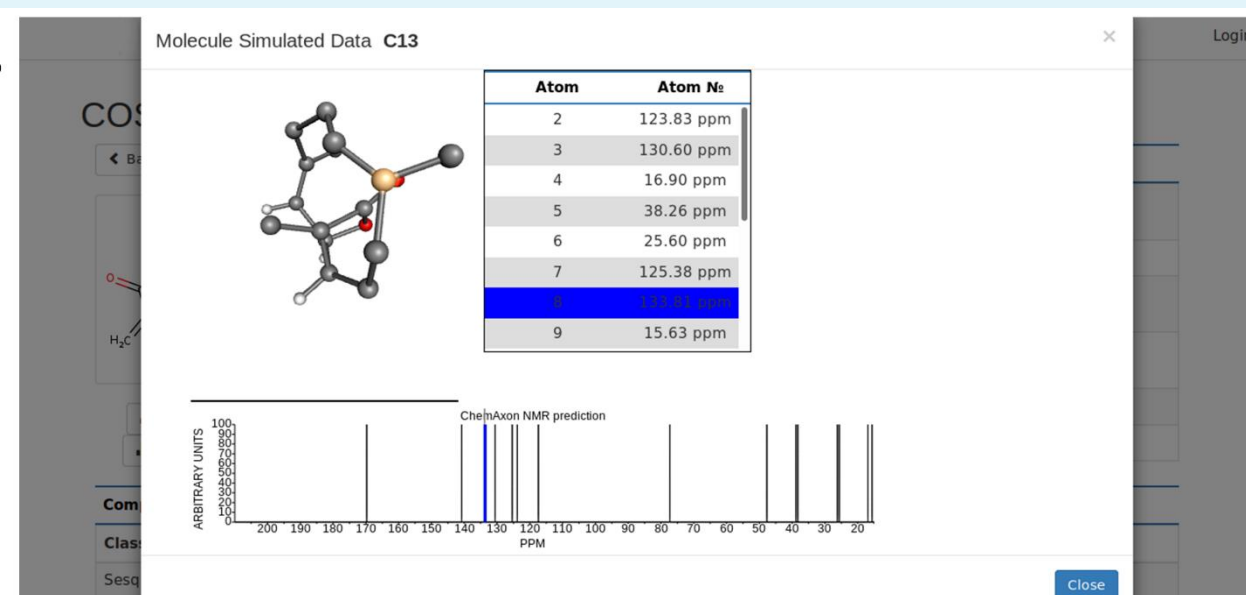
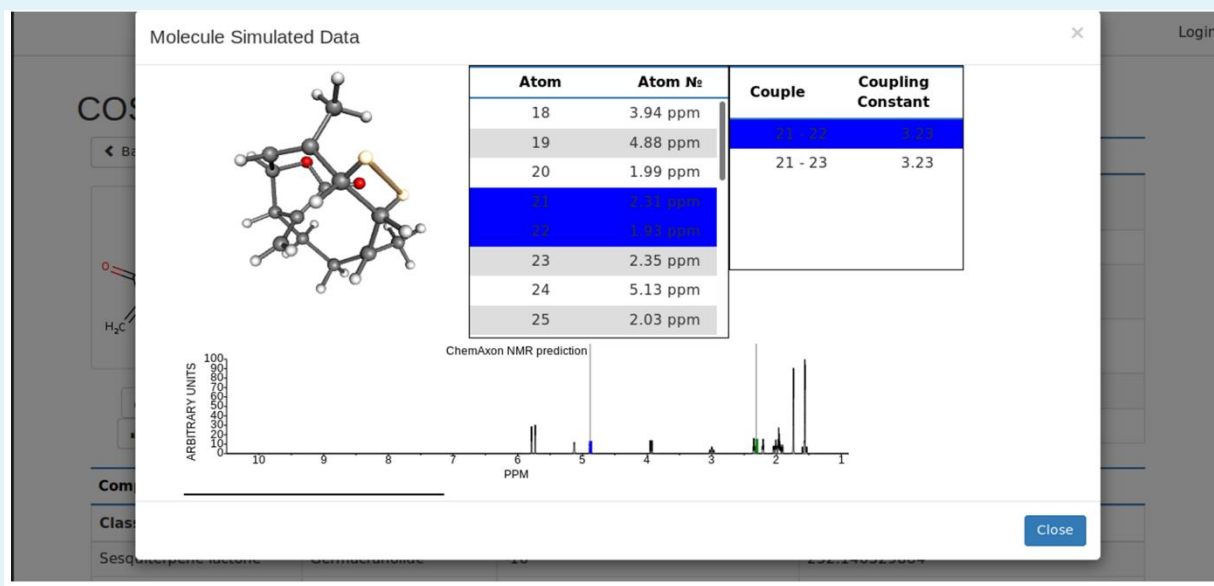
Download Results to CSV

Results: 1 - 6 Total Results: 12

Generation and visualization of ^1H and ^{13}C NMR data

SistematX now includes NMR spectra, coupling constants, and chemical shifts, which are relevant to phytochemical studies.

Utilizing the ChemAxon API, predictions are calculated when a new compound is registered in the system.

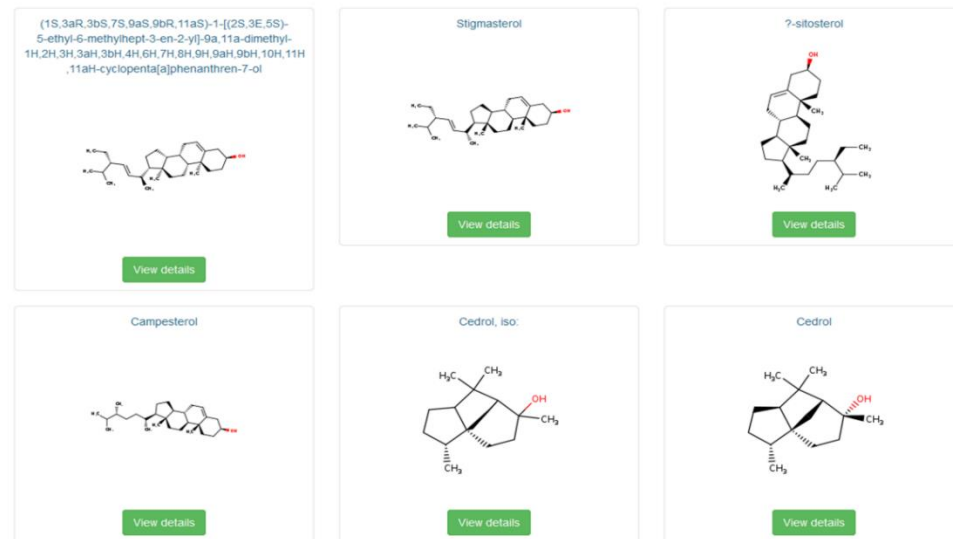
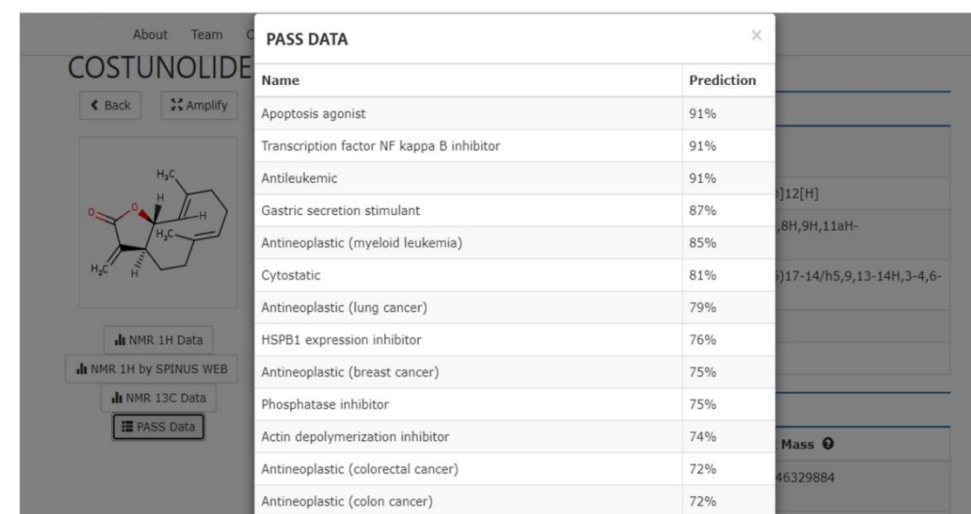


Biological activity profiles of registered structures are estimated using PASS software, which uses NP structural formulas to predict 4686 pharmacological effects and mechanisms of action.

SistematX reports prediction results with a probability of activity cutoff >0.5 , resulting in a total of 1571 predicted pharmacological effects and mechanisms of action.

Saldívar-González, F. I.; Aldas-Bulos, V. D.; Medina-Franco, J. L.; Plisson, F., Natural product drug discovery in the artificial intelligence era. *Chemical Science* 2022, 13 (6), 1526-1546.

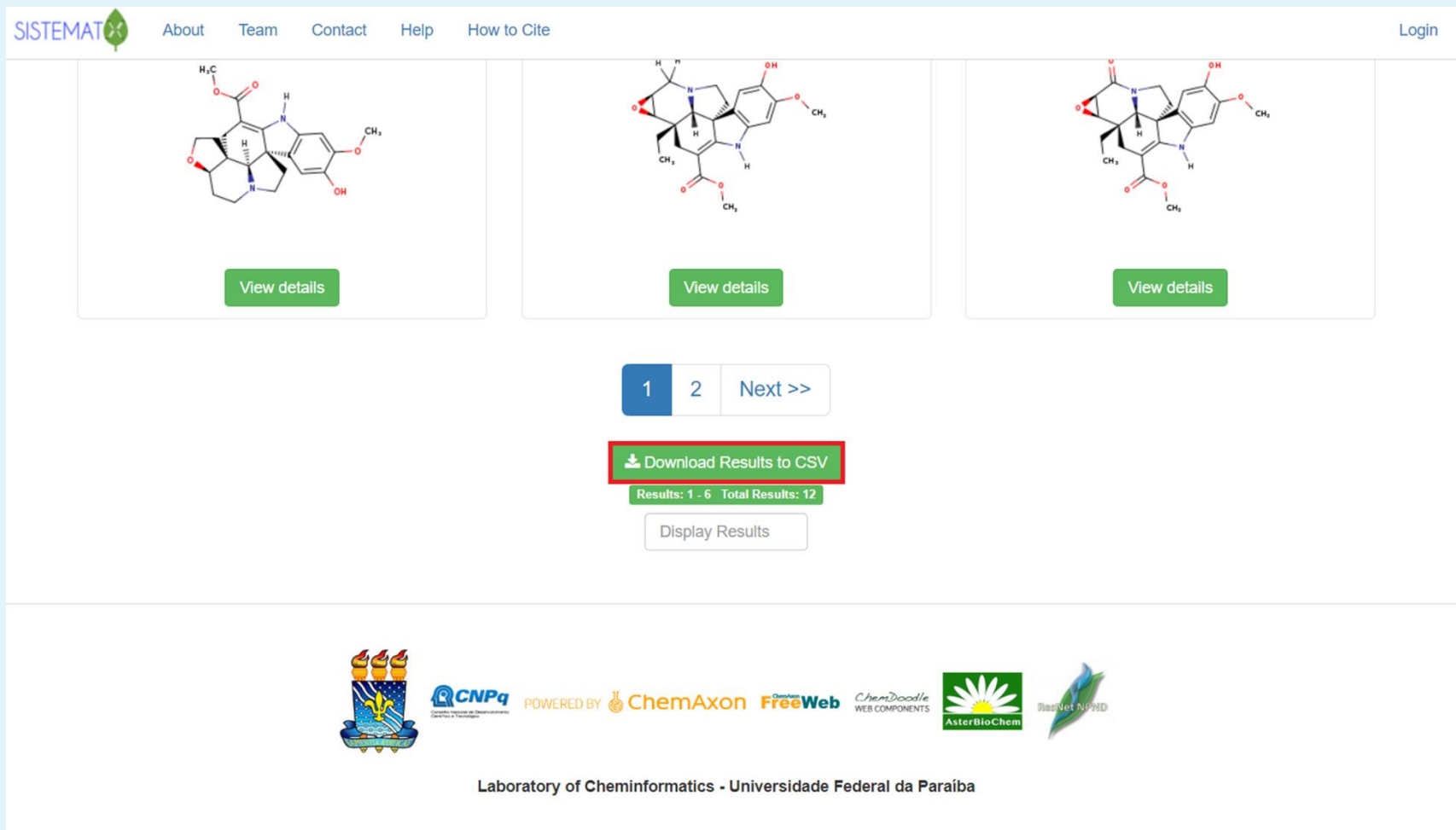
Lagunin A, Stepanchikova A, Filimonov D, Poroikov V. PASS: prediction of activity spectra for biologically active substances. *Bioinformatics*. 2000 Aug;16(8):747-8. doi: 10.1093/bioinformatics/16.8.747. PMID: 11099264.


A**B**

Name	Prediction
Apoptosis agonist	91%
Transcription factor NF kappa B inhibitor	91%
Antileukemic	91%
Gastric secretion stimulant	87%
Antineoplastic (myeloid leukemia)	85%
Cytostatic	81%
Antineoplastic (lung cancer)	79%
HSPB1 expression inhibitor	76%
Antineoplastic (breast cancer)	75%
Phosphatase inhibitor	75%
Actin depolymerization inhibitor	74%
Antineoplastic (colorectal cancer)	72%
Antineoplastic (colon cancer)	72%

Filimonov D.A., Lagunin A.A., Glorizova T.A., Rudik A.V., Druzhilovskii D.S., Pogodin P.V., Poroikov V.V. (2014). Prediction of the biological activity spectra of organic compounds using the PASS online web resource. *Chemistry of Heterocyclic Compounds*, 50 (3), 444-457.

The new batch download feature was developed using Java programming language. Downloading the search results creates a *.CSV file containing all botanical occurrences.



SISTEMAT  About Team Contact Help How to Cite Login

CC(=O)OC1=CC=C(O)C=C1N2C[C@H]3[C@@H](OC(=O)C)C[C@H]4[C@@H](OC(=O)C)C[C@H]5[C@@H](OC(=O)C)C[C@H]4N2C3

View details

CC(=O)OC1=CC=C(O)C=C1N2C[C@H]3[C@@H](OC(=O)C)C[C@H]4[C@@H](OC(=O)C)C[C@H]5[C@@H](OC(=O)C)C[C@H]4N2C3

View details

CC(=O)OC1=CC=C(O)C=C1N2C[C@H]3[C@@H](OC(=O)C)C[C@H]4[C@@H](OC(=O)C)C[C@H]5[C@@H](OC(=O)C)C[C@H]4N2C3







View details

1 2 Next >>

[Download Results to CSV](#)

Results: 1 - 6 Total Results: 12

Display Results

  POWERED BY   WEB COMPONENTS  

Laboratory of Cheminformatics - Universidade Federal da Paraíba

New partnership



Embrapa

A Embrapa / Unidades - Embrapa no Brasil / Unidades / Embrapa Agroindústria Tropical / A Unidade / Infraestrutura / Laboratório Multiusuário de Química e Produtos Naturais

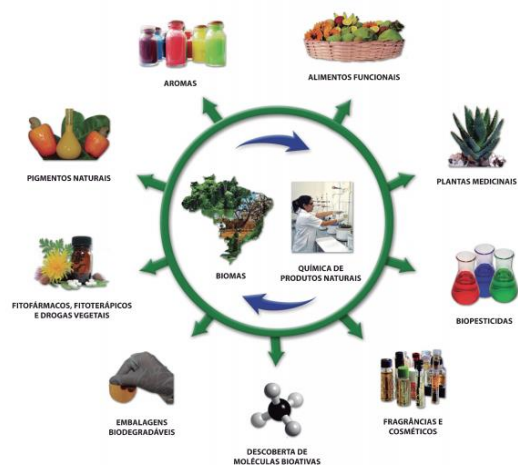
Embrapa Agroindústria Tropical

Encontre nesta Unidade

A Unidade | Soluções Tecnológicas | Projetos | Publicações | Biblioteca | Notícias | Multimídia | Sala de Imprensa

Laboratório
Multiusuário de Química e Produtos Naturais

Multidisciplinaridade de ações da Química de Produtos Naturais



Guilherme Juliao Zocolo



Lorena Mara Alexandre e Silva



AsterBioChem

AsterBioChem

2011

Home | Team | Research | Projects | Publications | Teaching | Facilities | Pictures gallery | AsterDB | Links | Contact

University of São Paulo

AsterDB is the AsterBioChem in house database that contains hundreds of chemical structures reported in species from the family Asteraceae. It is the first database dedicated only to the sunflower family and has free access (see below).

AsterDB is an unique database and contains structures of terpenoids, flavonoids and para-oxo-amic acid derivatives among other chemical classes, including more than 1,000 structures of sesquiterpene lactones. The structures were described in important medicinal (carica, chamomile, dandelion, taraxacum, yarrow, yucca, etc.), ornamental (daisy, zinnia, etc.) or edible plants (artichoke, common lettuce, thistle, etc.) as well as in other taxonomically interesting South American taxa such as *Yucca*, *Opuntia*, *Adiantum* (see [Vijayaraj](#), [Mishra](#), [Bachmann](#), [Barnadesianthus](#), [Epilobium](#), [Crotalaria](#), etc.).

All structures are in 2D mol format with assigned stereochemistry and are currently being used for virtual screening, to build Quantitative Structure-Activity (QSAR) and Quantitative Structure-Relationship (QSAR) models, for derivation of extracts using LC-UV-MS data as well as for chemotaxonomic studies.

Read our most recent publications citing AsterDB:
- Oliveira et al. J. Chem. Inf. Model. 5(9): 20-26, 2015.
- Chagas-Paula et al. Metabolites 5: 404-430, 2015.

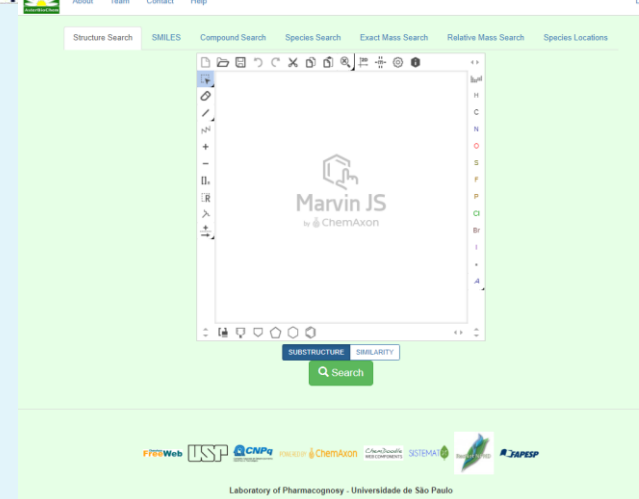
Site version online: click here

Institution: Journal of Pharmacognosy | FreeWeb | ChemDoodle WEBCOMPONENTS | Institution: Chemical Society

Asteraceae database

<http://www.asterbiochem.org/asterdb>

About | Team | Contact | Help | Login



Structure Search | SMILES | Compound Search | Species Search | Exact Mass Search | Relative Mass Search | Species Locations

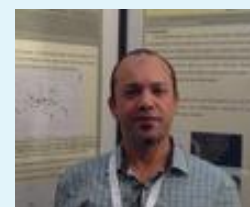
Marvin JS
by ChemAxon

Substructure | Similarity

Search

FreeWeb | USP | CNPq | Inpeco | ChemAxon | ChemDoodle WEBCOMPONENTS | SISTEMATI | JAPESP

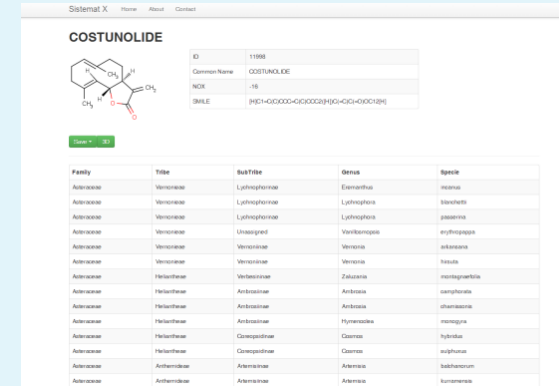
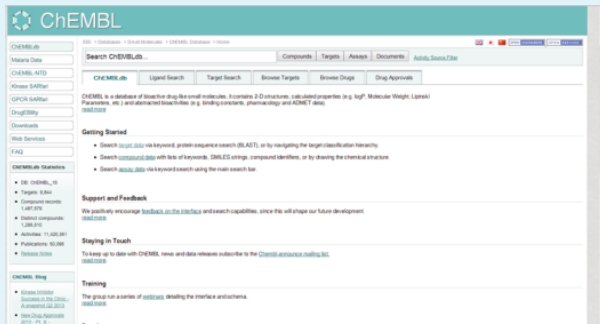
Laboratory of Pharmacognosy - Universidade de São Paulo



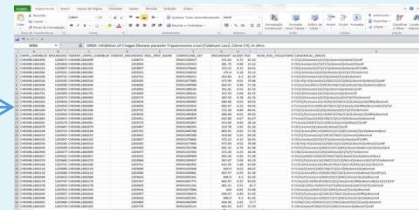
Fernando B. Da Costa

Laboratory of Pharmacognosy
School of Pharmaceutical Sciences of Ribeirão Preto
University of São Paulo





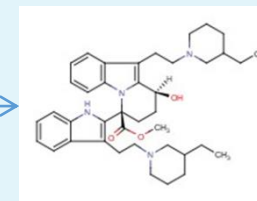
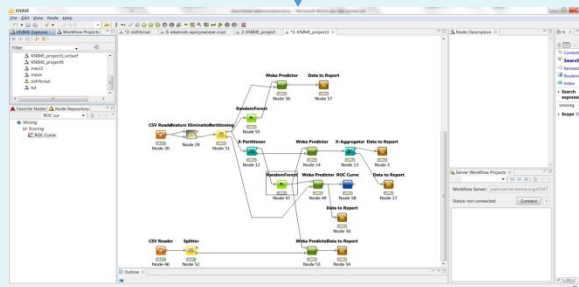
Combine data



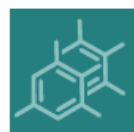
Biological activities
Structures

Structures
Biological activities
NMR data
Geographic location
Botanical data

Virtual Screening









Machine Learning



Article

Identification of Kaurane-Type Diterpenes as Inhibitors of Leishmania Pteridine Reductase I

Chonny Herrera-Acevedo ^{1,2} , Areli Flores-Gaspar ^{3,*} , Luciana Scotti ¹ ,
Francisco Jaime Bezerra Mendonça-Junior ⁴ , Marcus Tullius Scotti ^{1,*}  and Ericsson Coy-Barrera ^{2,3} 

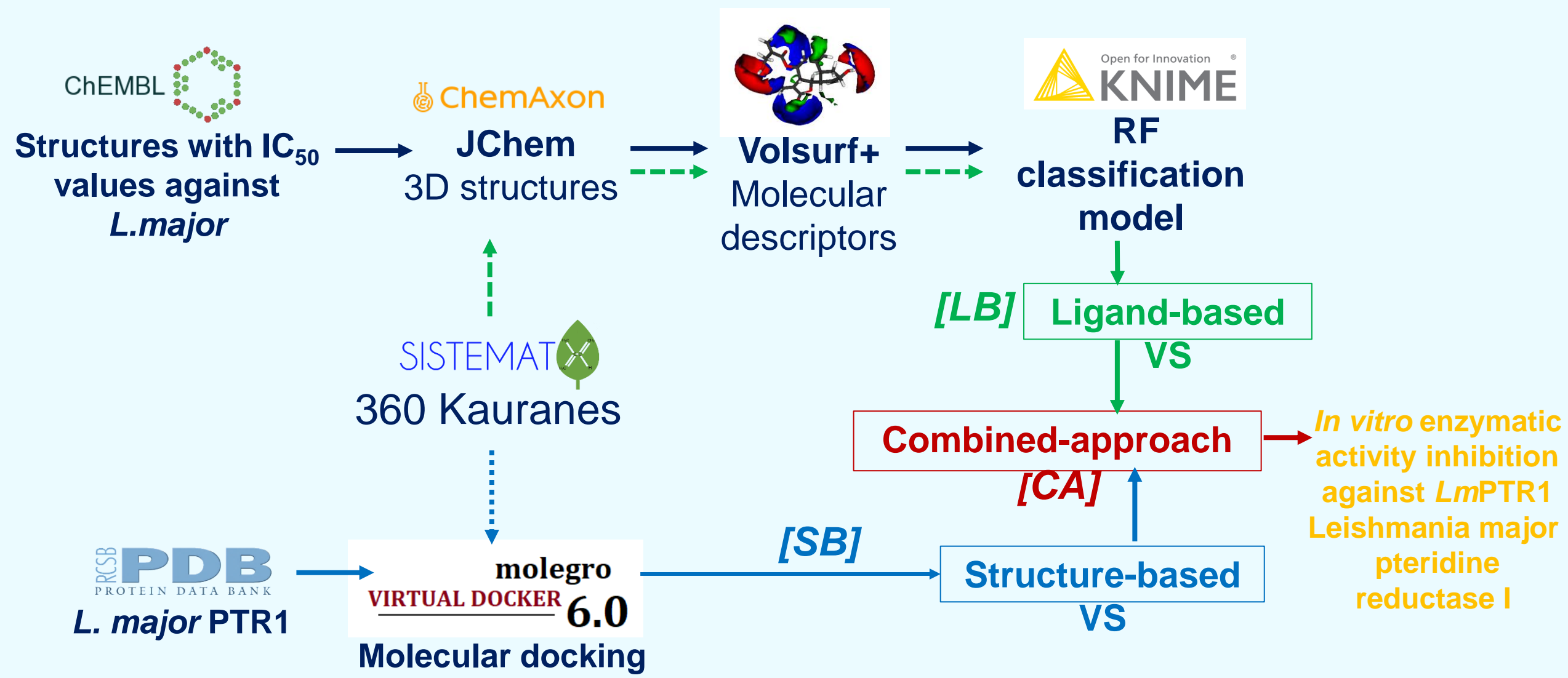
¹ Post-Graduate Program in Natural and Synthetic Bioactive Products, Federal University of Paraíba, João Pessoa 58051-900, PB, Brazil; chonny622@gmail.com (C.H.-A.); luciana.scotti@gmail.com (L.S.)

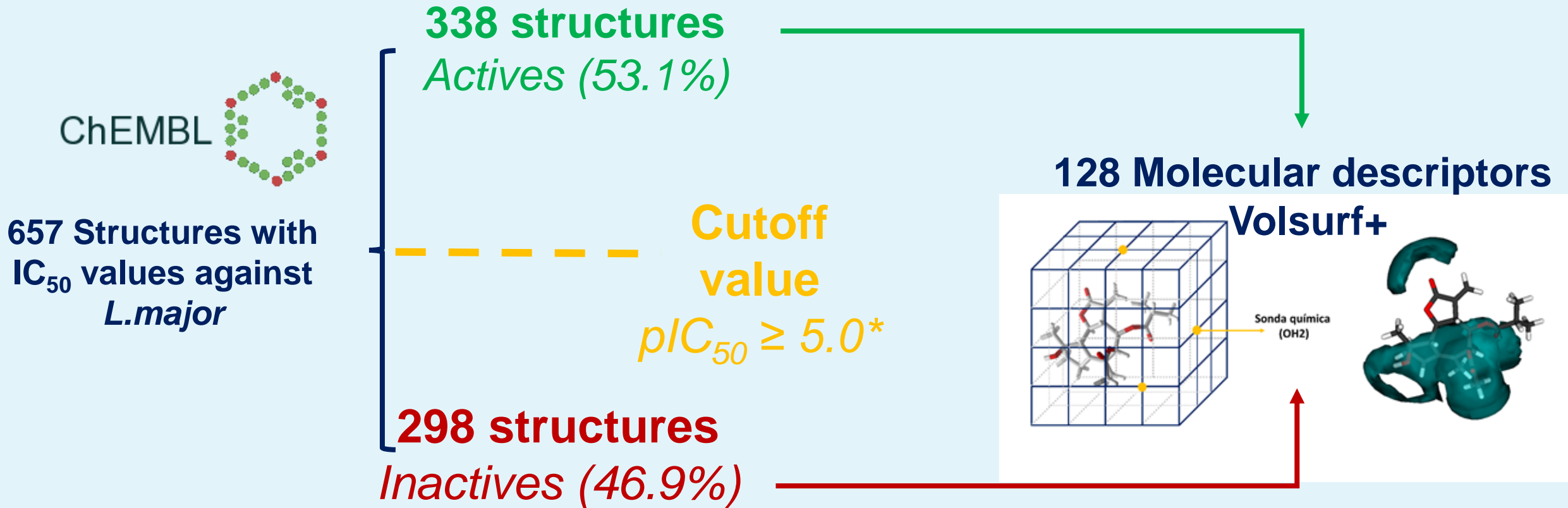
² Bioorganic Chemistry Laboratory, Facultad de Ciencias Básicas y Aplicadas, Universidad Militar Nueva Granada, Cajicá 250247, Colombia; ericsson.coy@unimilitar.edu.co

³ Departamento de Química, Facultad de Ciencias Básicas y Aplicadas, Universidad Militar Nueva Granada, Cajicá 250247, Colombia

⁴ Laboratory of Synthesis and Drug Delivery, State University of Paraíba, João Pessoa 58071-160, PB, Brazil; franciscojbmendonca@yahoo.com.br

* Correspondence: areli.flores@unimilitar.edu.co (A.F.-G.); mtscotti@gmail.com (M.T.S.);
Tel.: +57-1-650-00-00 (ext. 1526) (A.F.-G.); +55-83-99869-0415 (M.T.S.)

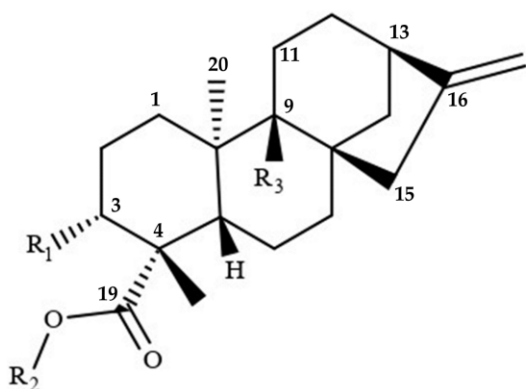


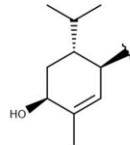
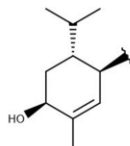
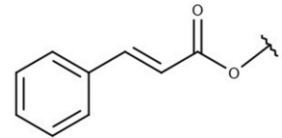
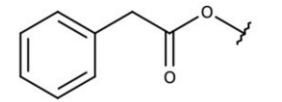


* $4.9 < pIC_{50} < 5.0$ (21 structures were removed)

Ligand-based virtual screening

Only 7 of the 360 structures were classified as active (ligand-based probability value [LB] ≥ 0.5)



ID	R ₁	R ₂	R ₃	LB
135	H		H	0.57
134	H		$\Delta^{9,11}$	0.55
302		H	H	0.54
298		H	H	0.53

Asteraceae

Wedelia chinensis
Asia

Wedelia trilobata
Central
America

$$CA = \frac{SB + (1 + TN) \times LB}{2 + TN}$$

Specificity (TN) 72.7 %

Kaurane	<i>SB</i>	<i>LB</i>	CA_{Lm}
135	0.93	0.57	0.70
101	1.00	0.51	0.69
302	0.94	0.54	0.68
134	0.90	0.55	0.68
298	0.93	0.53	0.68

→ *Ligularia fischeri*
(Asteraceae)

Five kauranes from various Asteraceae species were identified as having promising antileishmanial activity against *LmPTR1* from a dataset of 360 kauranes

In vitro enzymatic activity inhibition against *LmPTR1*

Compound	135	302
IC ₅₀ (μM)	8.6	9.6
Confidence Interval (95%)	9.4–7.9	10.7–8.6
Ki ^{app}	1.88	2.10

Active

Machine learning cutoff value

$$pIC_{50} \geq 5.0$$

$$IC_{50} \leq 1 \times 10^{-5} M (10 \mu M)$$

Natural Products Online is an open-source project for Natural Products (NPs) storage, search and analysis.

<https://naturalproducts.net/>

The image shows a screenshot of the Natural Products Online website. At the top, the text reads "NATURAL PRODUCTS ONLINE" in large white letters on a dark blue background. Below this, a smaller line of text states: "This is a portal for the open-source open-data repository for natural products. The web resources are developed and maintained by the Steinbeck Group." Below the text is a grid of six logos for different tools or databases:

- NAPLES**: Natural Product Likeness Score calculator. The logo features a stylized mountain range in orange and blue.
- COCONUT**: COllection of Open Natural Products. The logo features a brown coconut with a green palm frond and a chemical structure.
- SUGAR**: SUGAR REMOVAL UTILITY. The logo features a 3D grid of white cubes with a red and white chemical structure in the center.
- Watermelon**: The logo features a slice of watermelon with a green rind and red flesh with black seeds.
- LOTUS**: The logo features a purple lotus flower with green leaves.
- DECIMER**: The logo features the word "DECIMER" in a bold, black, sans-serif font.

This is a portal for the open-source open-data repository for natural products.
The web resources are developed and maintained by the Prof. Dr. Steinbeck group.

Laboratory of Cheminformatics UFPB

Main researchers



Prof. Dr. Marcus Tullius Scotti
Prof. Dr. Luciana Scotti

Visiting professor



Prof. Dr. Eugene N. Muratov

Doctoral students

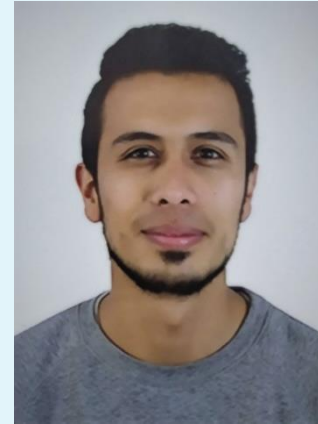
Master students



Natan Dias
Fernandes



Renata Priscila
Barros de Menezes



Chonny
Herrera-Acevedo



Isadora Silva
Luna



Natália Ferreira
de Sousa

SistematiX team



**Emmanuella Faustino
Albuquerque**



**Ávilla Ítalo de
Souza Silva**



Lucas Ferreira Calado

Thanks

mtscotti@gmail.com



Areia Vermelha

