



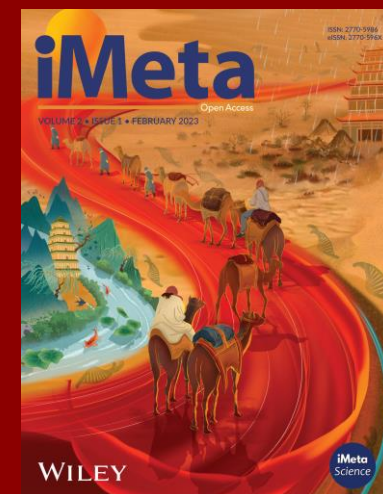
SIDERITE: 通过数字化探索揭示 化学空间中隐藏的铁载体多样性

贺若霖^{1#}, 顾少华^{1,2#}, 许家正³, 李学健⁴, 陈浩然⁴, 邵正英³,
王凡灏¹, 邵吉祺¹, 尹文兵^{5,6}, 钱珑^{1*}, 韦中^{3*}, 李志远^{1,2*}

¹北京大学前沿交叉学科研究院定量生物学中心

²北京大学前沿交叉学科研究院生命联合中心³南京农业大学

⁴北京远轩科技有限公司⁵中国科学院微生物研究所⁶中国科学院大学存济医学院



Ruolin He, Shaohua Gu, Jiazheng Xu, Xuejian Li, Haoran Chen, Zhengying Shao, Fanhao Wang, Jiqi Shao, Wen-Bing Yin, Long Qian, Zhong Wei, Zhiyuan Li. 2024. SIDERITE: Unveiling Hidden Siderophore Diversity in the Chemical Space Through Digital Exploration. *iMeta* 3: e192. <https://doi.org/10.1002/imt2.192>

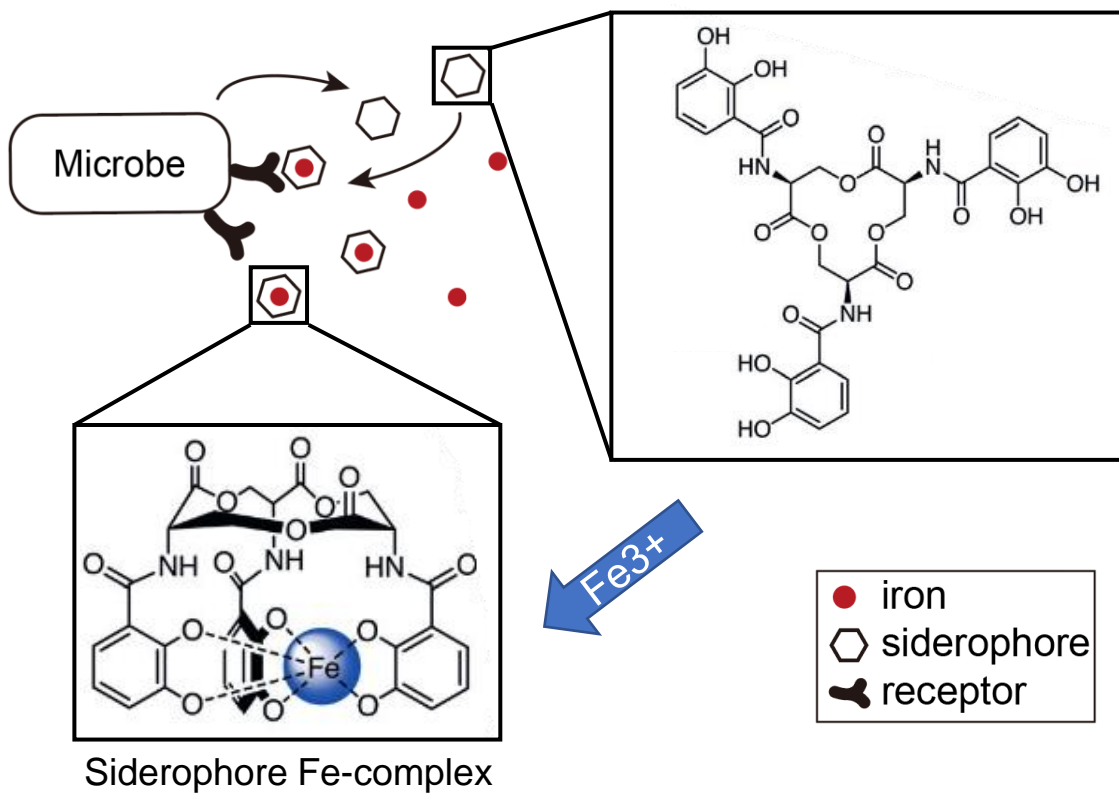
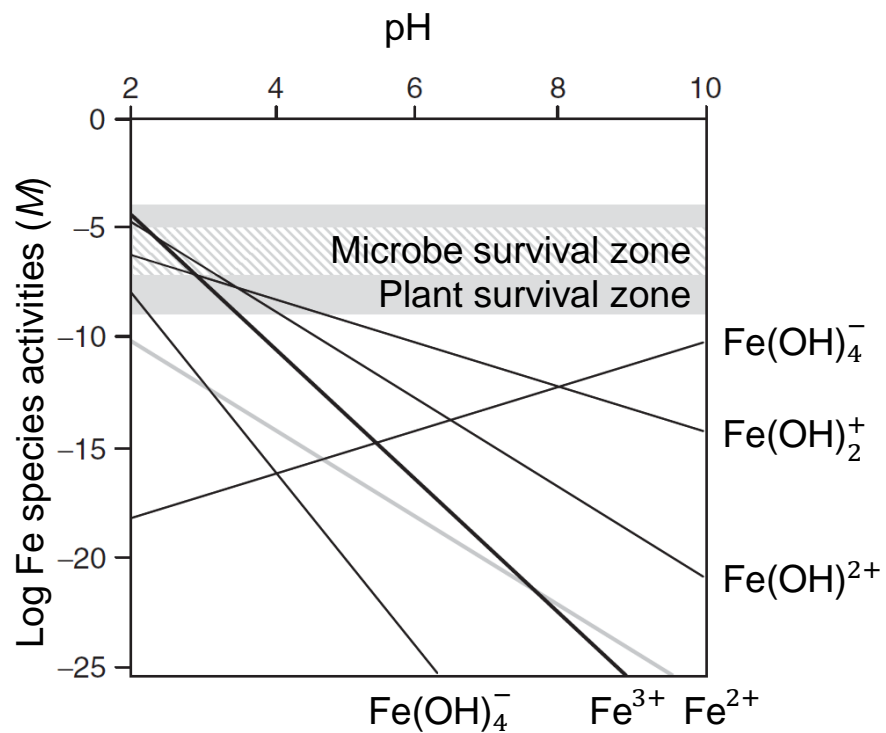


简介

自然界中的铁无法满足微生物的需求。

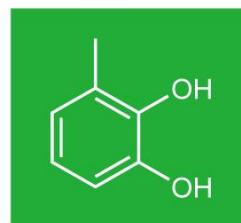
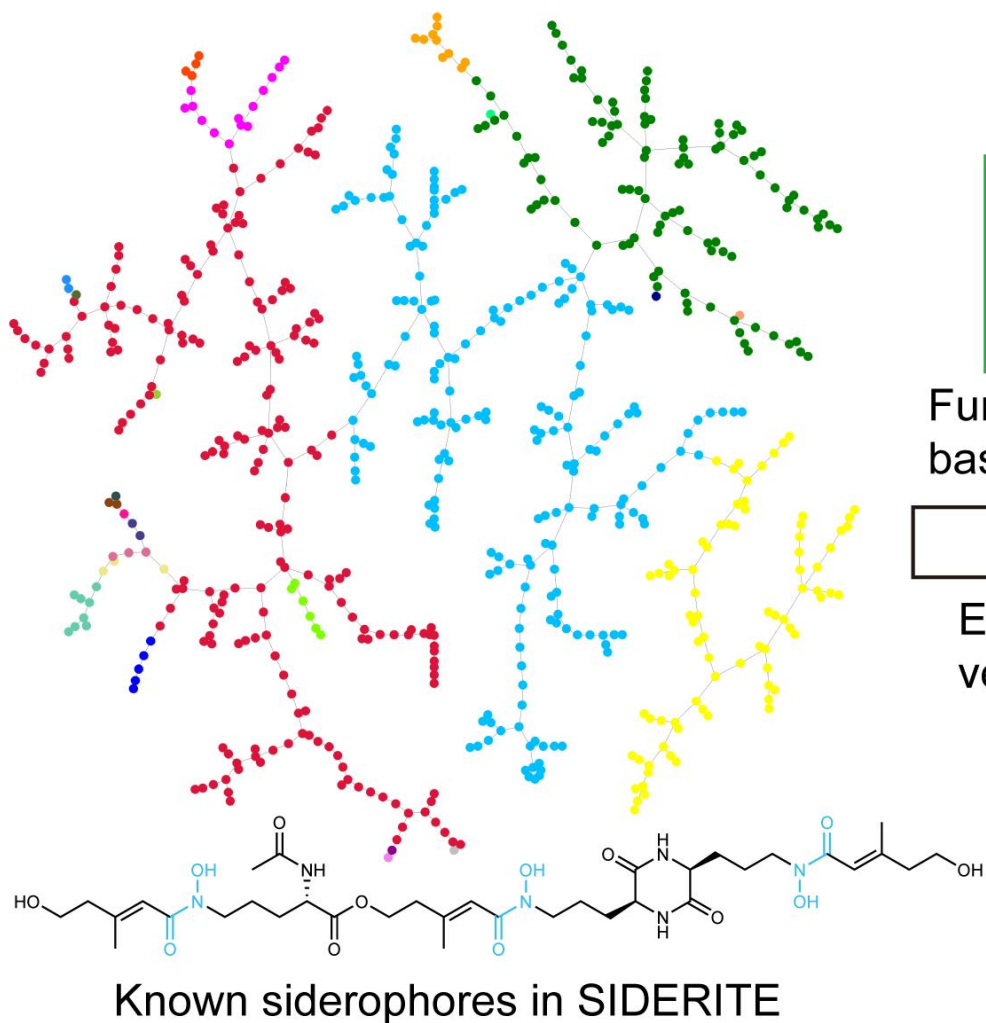
微生物可以产物铁载体以获取环境中的铁。

铁载体是具有高度结构多样性的分子。



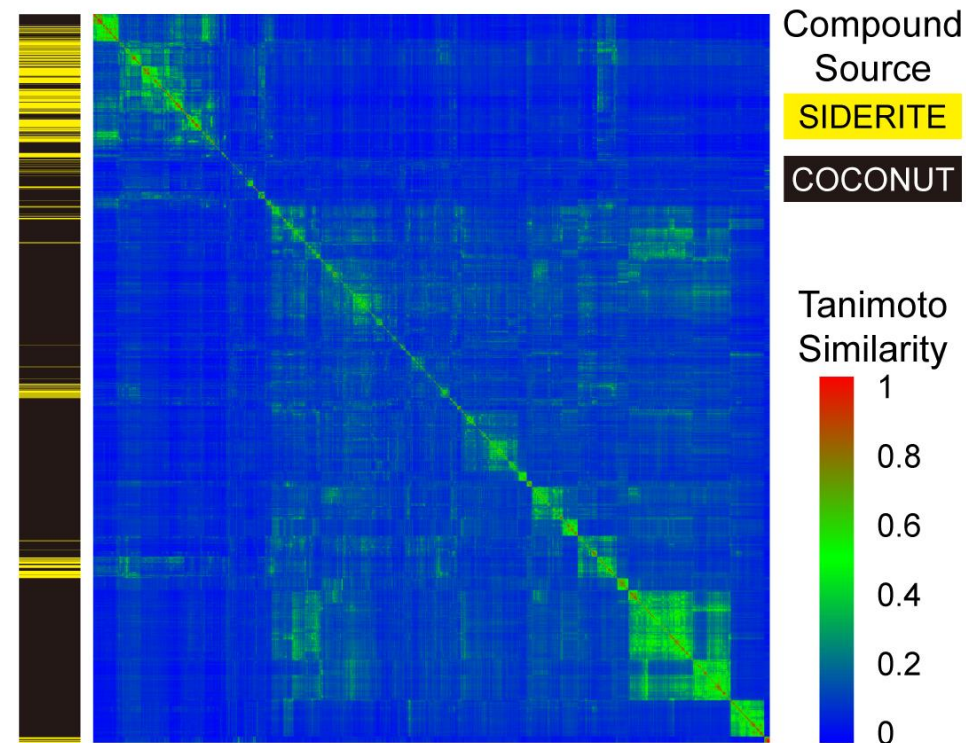


简介



Functional group-based method

Experimental verification



Unexplored siderophore diversity

SIDERITE: Siderophore information database
COCONUT: the collection of open natural products database



亮点



SIDERITE是最全的数字化整合铁载体信息数据库。

数据库涵盖了872个铁载体记录中所有的649个不同的结构。数据库每年至少更新一次。

通过合并具有不同名字但是相同结构的铁载体来解决铁载体的混淆问题。

Bacillibactin / Corynebactin

Siderophore ID: SID00195

Theoretical Denticity: 6

Functional Group: Catecholates (3)

Biosynthetic Type: NRPS

Canonical SMILES: C[C@H]1OC(=O)[C@@H](NC(=O)CNC(=O)c2cccc(O)c2O)[C@@H](C)OC(=O)[C@@H](NC(=O)CNC(=O)c2cccc(O)c2O)[C@@H](C)OC(=O)[C@H]1NC(=O)CNC(=O)c1cccc(O)c1O

Monomers Number: 9

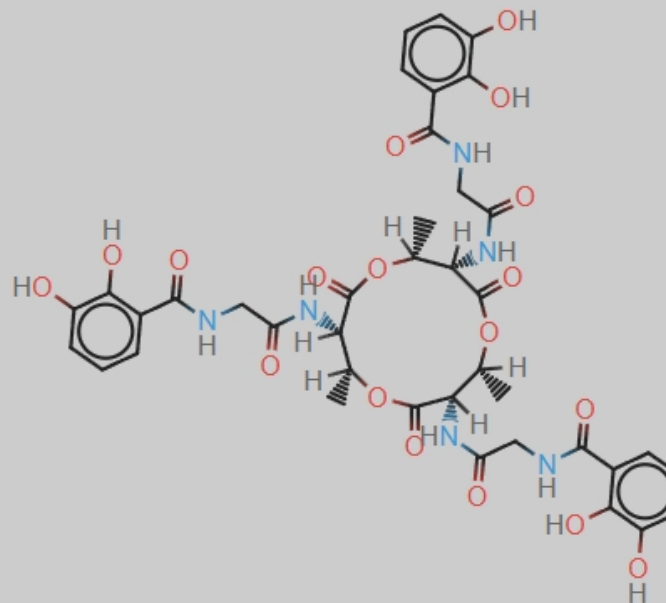
Monomers Name: aThr/Thr, aThr/Thr, Gly, diOH-Bz, aThr/Thr, Gly, diOH-Bz, Gly, diOH-Bz

Coverage of Monomers: 1

Molecular Formula: C₃₉H₄₂N₆O₁₈

References

1. Melissa K. Wilson, Rebecca J. Abergel, Kenneth N. Raymond, Jean E.L. Arceneaux, B. Rowe Byers, "Siderophores of Bacillus anthracis, Bacillus cereus, and Bacillus thuringiensis", Biochemical and Biophysical Research Communications, 2006, 348(1), 320-325

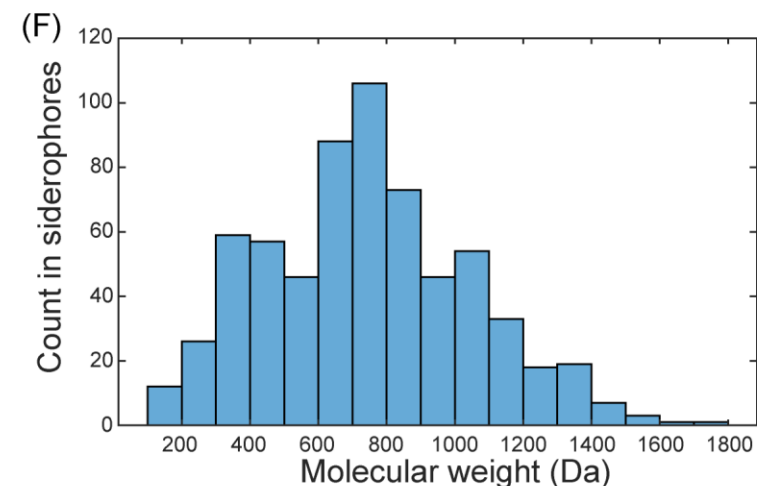
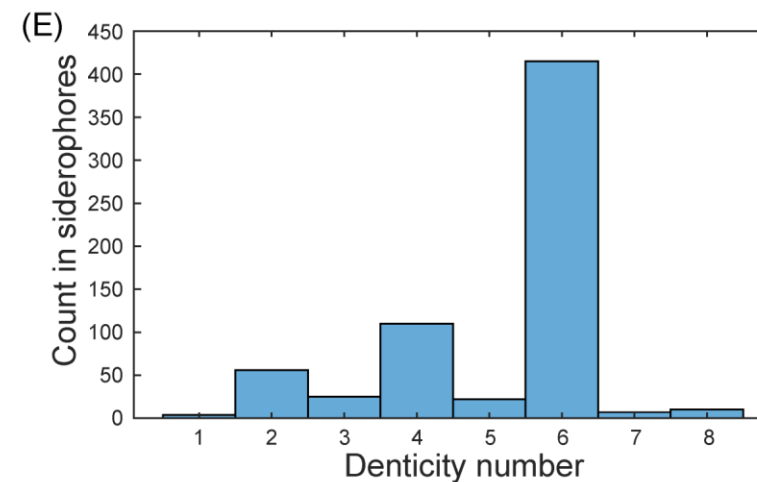
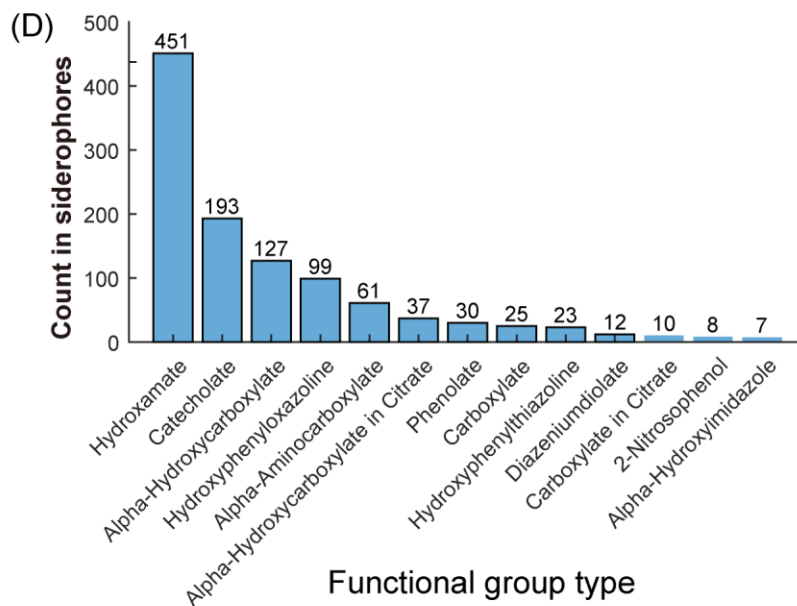
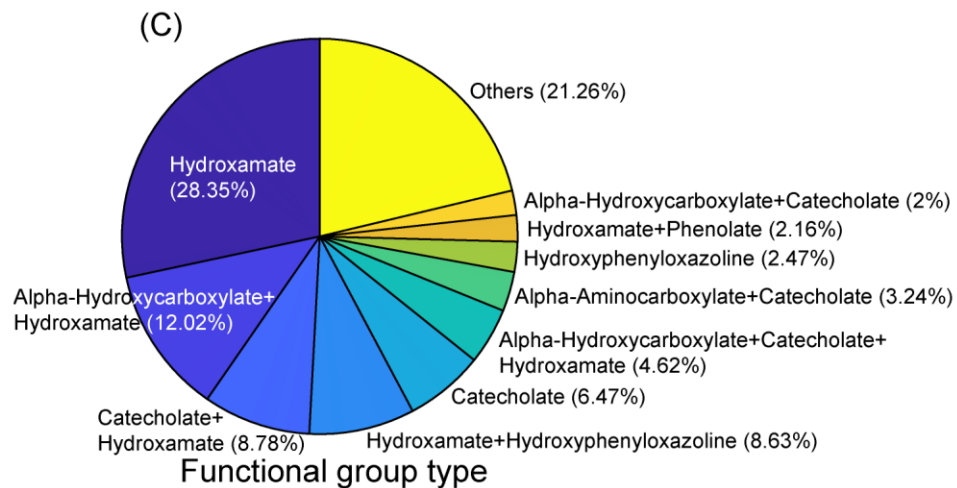
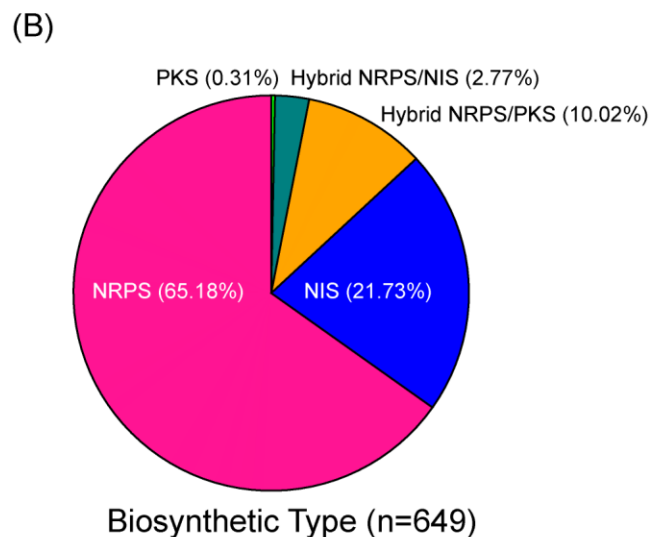
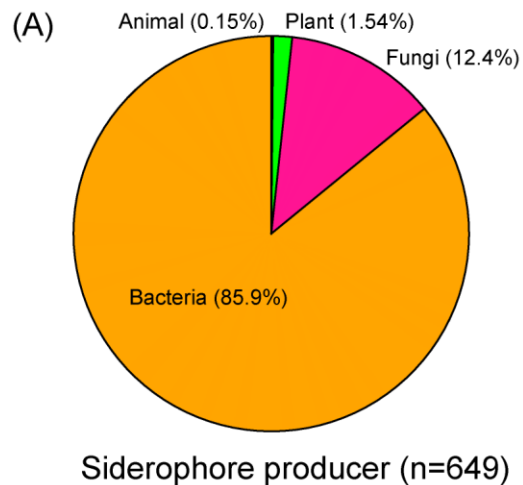


Generated by [Smiles Drawer](#)



SIDERITE数据库概述

SIDERITE数据库主要由微生物来源、NRPS合成、配体为异羟肟酸、齿数为6的铁载体构成。





基于结构相似度对铁载体进行聚类

649铁载体被聚类成25个簇和102个组。

最大的4个簇占据了已知铁载体的^(A) 89.37%。

Cluster 1 (201, 30.97%):
铁载体具有苯环的结构

Cluster 2 (197, 30.35%):
铁载体是使用异羟肟酸或alpha-羟基羧酸作为螯合铁的功能基团的NRPS铁载体。

Cluster 3 (103, 15.87%):
铁载体是使用异羟肟酸或alpha-羟基羧酸作为螯合铁的功能基团的NIS铁载体。

Cluster 4 (79, 12.17%):
由NRPS合成并且具有发色基团的铁载体。

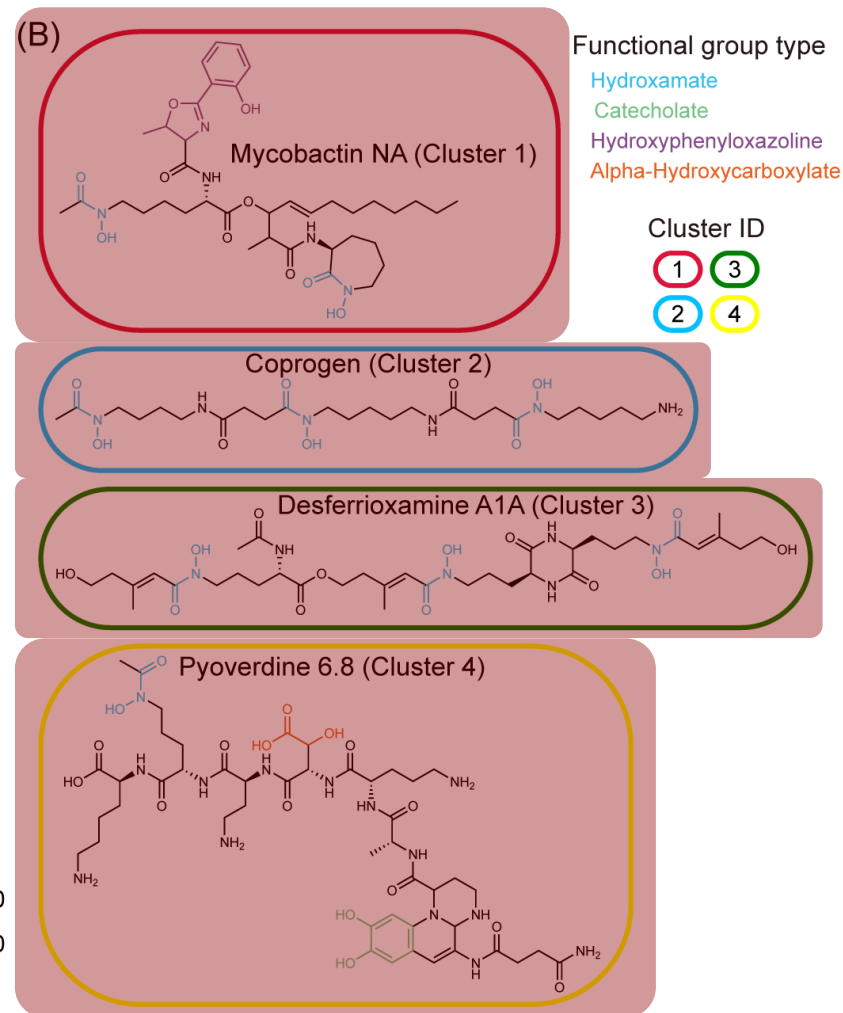
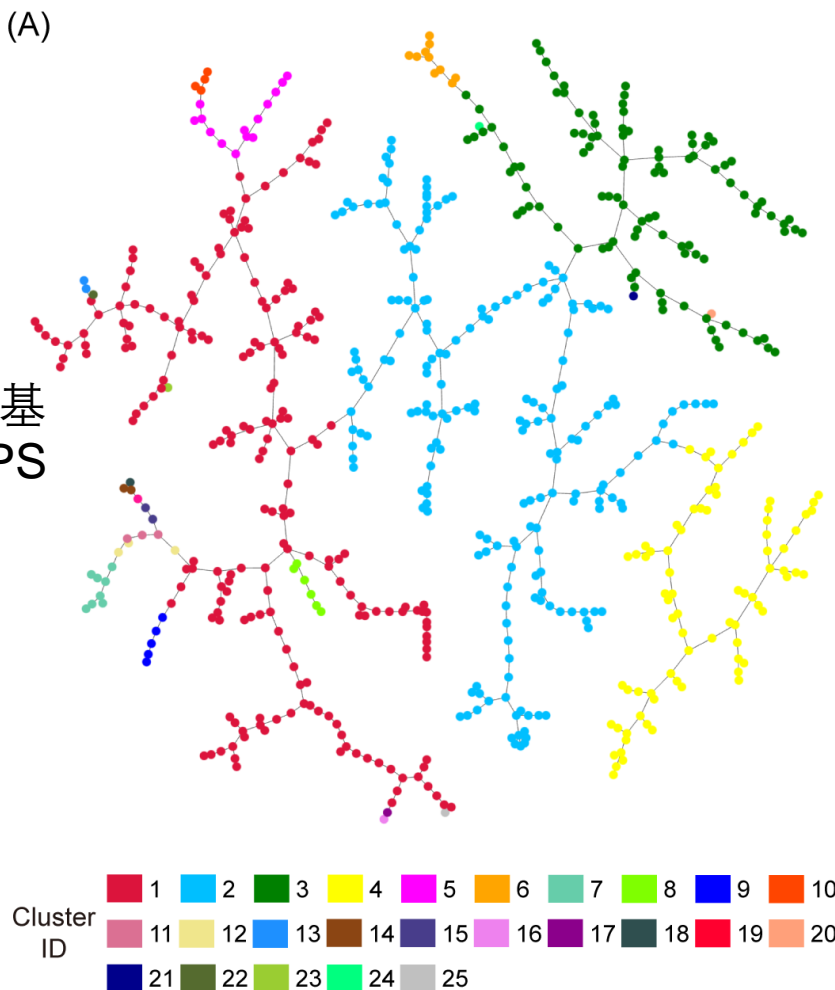


Figure 1 The visualization of 25 siderophore clusters in siderophore information database (SIDERITE).

NRPS: non-ribosomal peptide synthetase
NIS : NRPS-independent siderophore

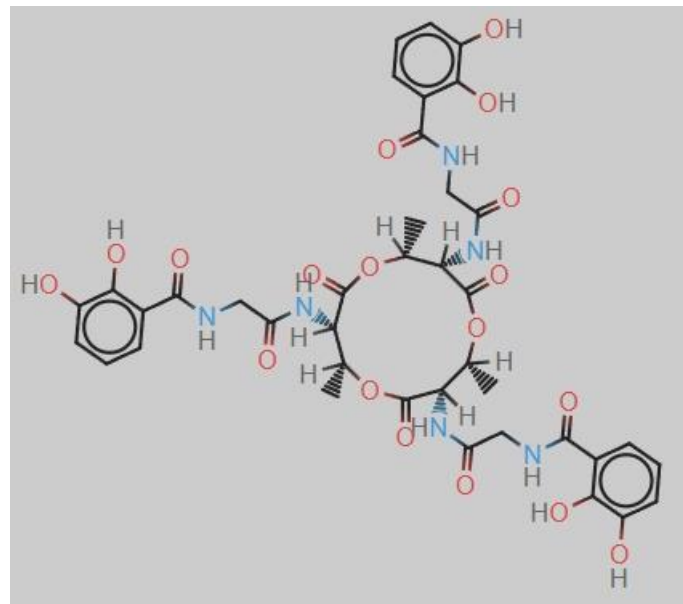
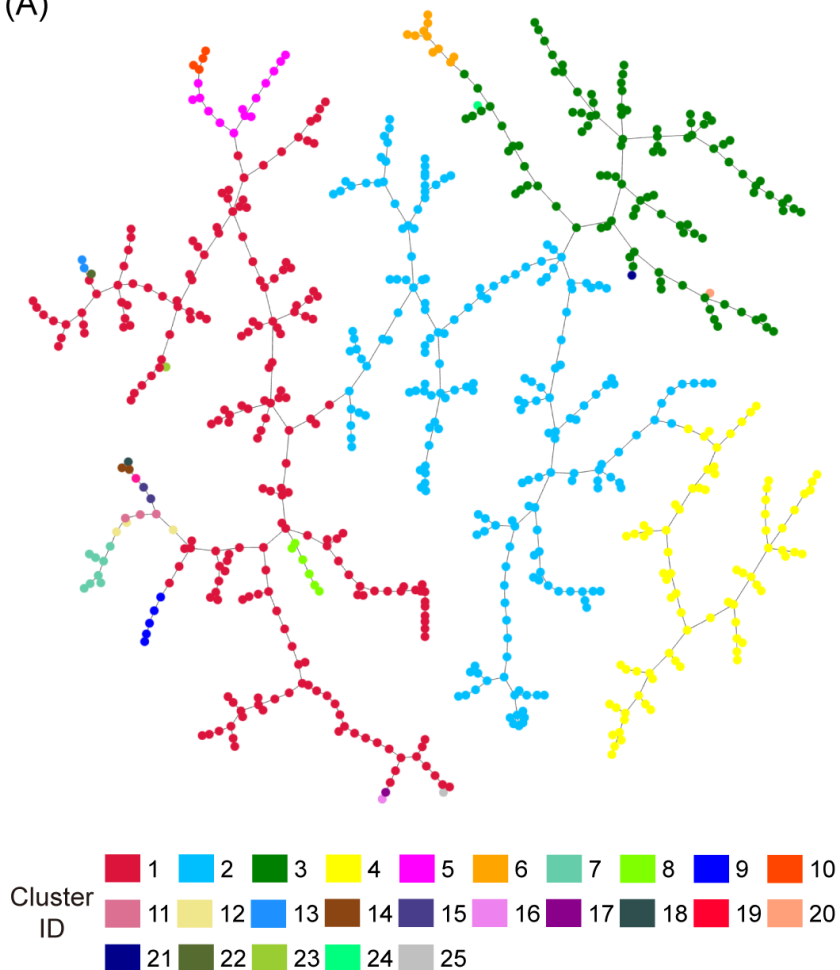


用簇和组命名铁载体

每一个铁载体被分配了一个唯一的ID，称作“类”。

可以通过相同组内已知生物合成类型的铁载体来推断未知生物合成类型的铁载体。

(A)



Example: Bacillibactin

Class: 1.22.3

$\begin{matrix} | & | & | \\ x & y & z \end{matrix}$

我们发现相同组中的铁载体具有相同的生物合成类型。

x : cluster ID

y : group ID

z : the z -th record



用基于功能基团的方法发现潜在的铁载体

规则：
含有至少一种这15种功能基团的分子且排除了含有任何8种被修饰铁载体功能基团

结果：
从COCONUT数据库中识别出了3199种具有潜在的铁结合活性的分子。

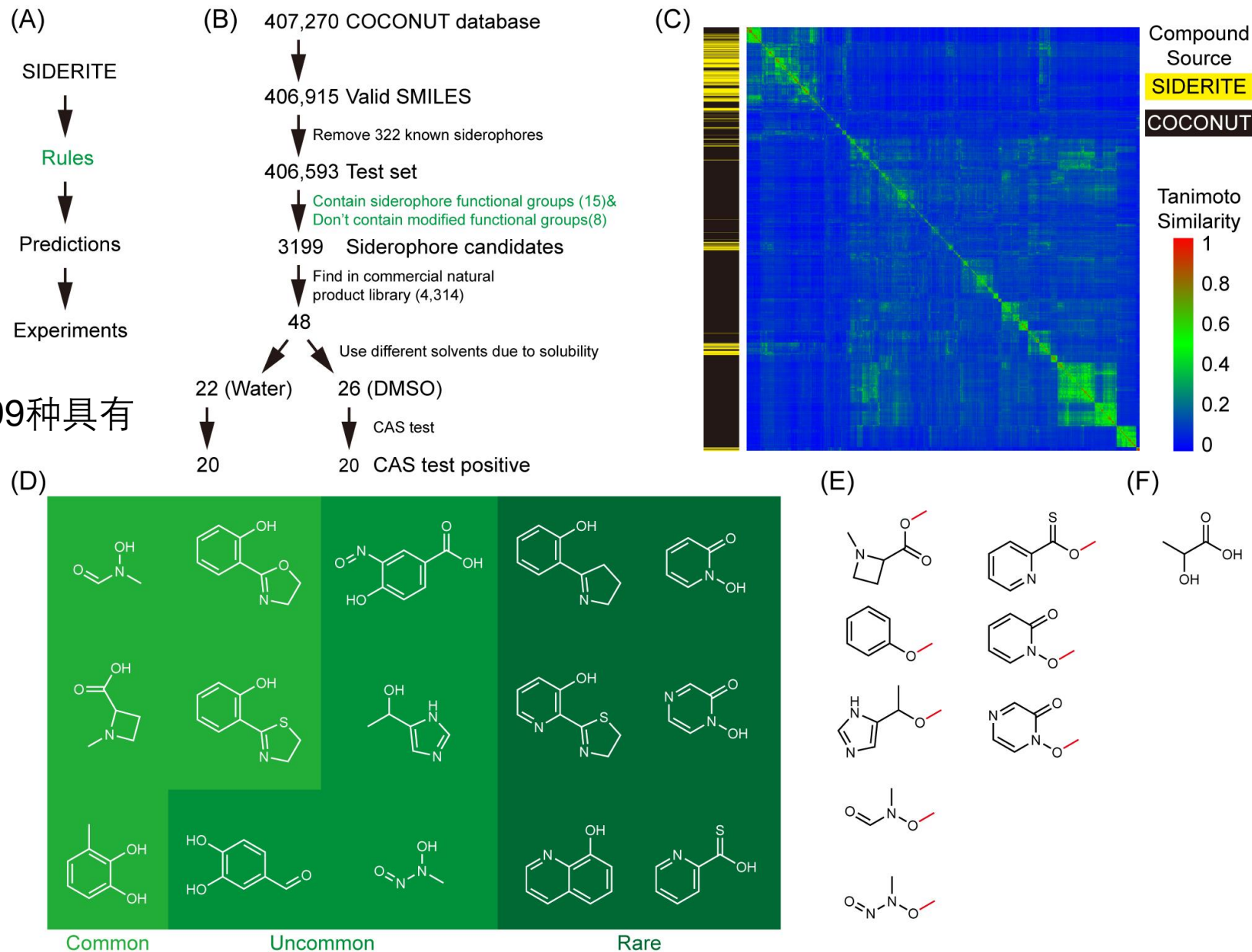


Figure 2 The rule-based siderophore discovery approach and the result of chrome azurol S (CAS) test experiments.



用CAS试验验证潜在的铁载体

3199个候选分子中的48个可购买的分子通过CAS试验进行了验证。

83.3% (40/48) 分子展现出结合铁的活性。

CAS试验中的高阳性率表明我们基于功能基团的方法是有效的。

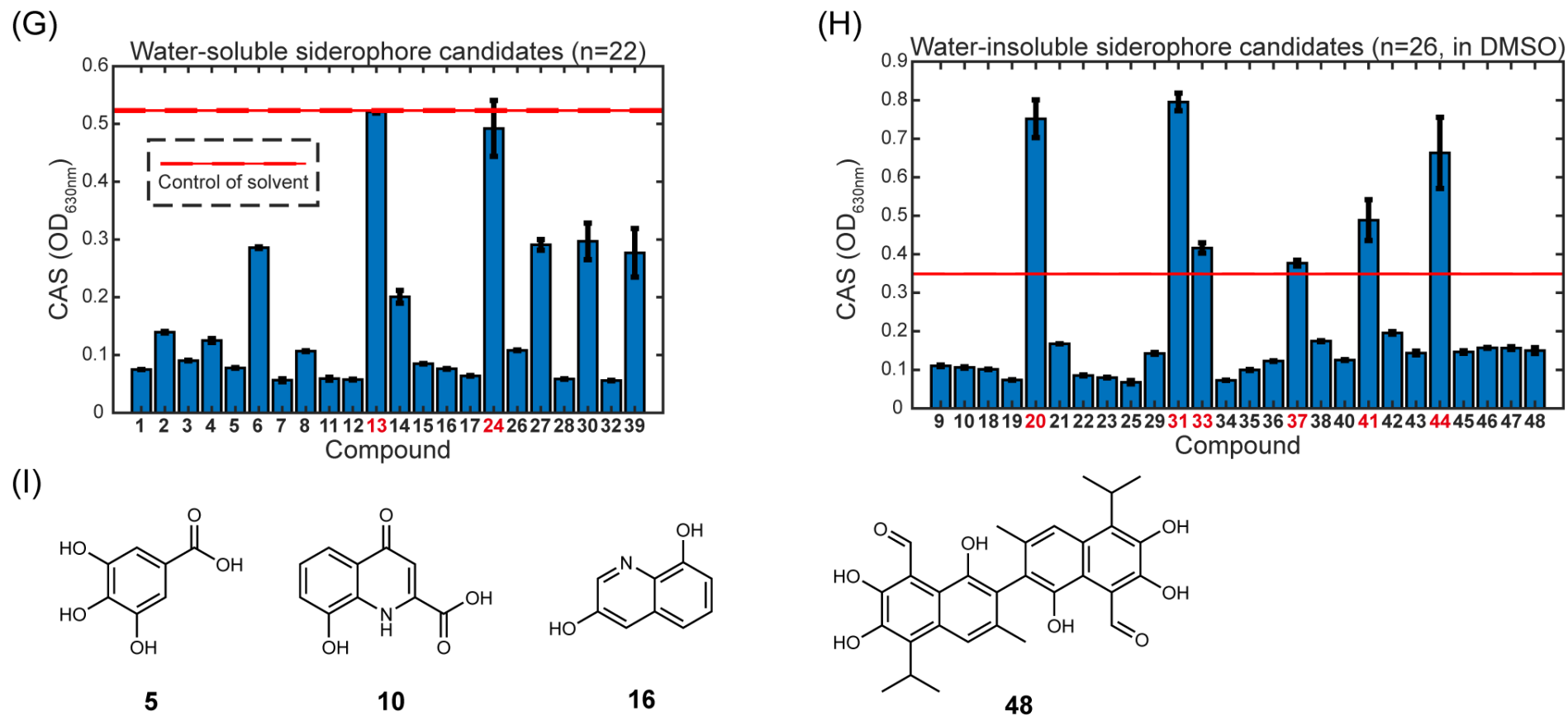


Figure 2 The rule-based siderophore discovery approach and the result of chrome azurol S (CAS) test experiments.



SIDERITE数据库的使用

BDA Informatics Suite



SIDERITE

Siderophore Information Database

文档: <https://github.com/RuolinHe/SIDERITE/wiki>

报告bug: <https://github.com/RuolinHe/SIDERITE/issues>

用于讨论的论坛: <https://groups.google.com/g/siderite-database>

教学视频: <https://www.youtube.com/watch?v=q0EoiRi1zVE>

Database

Similarity Search

TMAP Index

Search

Fast | Advanced

Total: 649

2-deoxymugineic acid

2-hydroxybenzoic acid

2-HydroxyNicotianamine

2-N-hydroxy-3, 4-dihydroisoquinol-2...

2-N-methylcoprogen

2-N-methylcoprogen B

2,3-dihydroxybenzoic acid

2,3-dihydroxybenzoylglycine

2,3-dihydroxybenzoylserine

Welcome to

SIDERITE - Siderophore Information Database

SIDERITE is a digitized integrated siderophore information database. It covers all 649 known siderophores from 872 records as of May 2023.

There are three ways to access the siderophore resource in the SIDERITE, which are as follows:

1. Search Based on Conditions

- **Fast Search:** Provides global real-time fuzzy search
- **Advanced Search:** Provides precise search with support for composite conditions

To initiate your search, please enter your queries on the search sidebar located on the left-hand side of the screen.

2. Similarity Search

Find similar siderophores in the SIDERITE with the desired molecule.

3. Indexing Interface

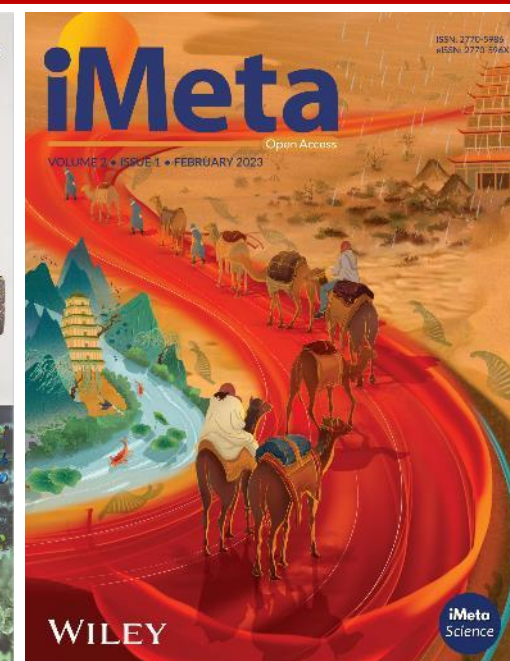
A visualized indexing interface of SIDERITE. You can set desired property of siderophores to label all siderophores in the database and select the specific siderophore with a single click.



总结

- ❑ 在我们的工作中，我们建立了最全面的铁载体信息数据库（SIDERITE），这是第一个采用SMILES格式具有649种独特结构的数字化铁载体存储库。
- ❑ 基于这些数字化结构，我们开发了一种基于功能基团的方法，用于发现具有高准确度的新型结合铁分子，并发现在铁载体研究领域尚未完全探索的结构多样性。
- ❑ 我们在数据库或GitHub页面提供了教程材料和反馈渠道，并持续维护SIDERITE数据库，并根据用户的反馈进行更新。
- ❑ 平台网址: <https://siderite.bdainformatics.org/>

Ruolin He, Shaohua Gu, Jiazheng Xu, Xuejian Li, Haoran Chen, Zhengying Shao, Fanhao Wang, Jiqi Shao, Wen-Bing Yin, Long Qian, Zhong Wei, Zhiyuan Li. 2024. SIDERITE: Unveiling Hidden Siderophore Diversity in the Chemical Space Through Digital Exploration. *iMeta* 3: e192. <https://doi.org/10.1002/imt2.192>



“**iMeta**”由威立、肠菌分会和华人科学家出版的开放获取期刊，主编由中科院微生物所刘双江和荷兰格罗宁根大学傅静远教授共同担任。目的是发表原创研究、方法和综述以促进宏基因组学、微生物组和生物信息学发展。目标是发表前10%(IF>20)的高影响力论文。期刊特色包括视频投稿、可重复分析、图片打磨、青年编委、中英双语、50万用户的社交媒体宣传等。2022年2月发行，相继被**ESCI**、**Google Scholar**、**DOAJ**、**Scopus**等数据库收录，发文161篇，被引2316次(**Dimension**, 2024/2/19)!



主页: <http://www.imeta.science>

出版社: <https://wileyonlinelibrary.com/journal/imeta>



投稿: <https://wiley.atyponrex.com/journal/IMT2>



office@imeta.science



宣传片



[iMeta](#)

