

BOLD数据库常用功能操作流程

2021-07-03 首都师范大学 遗传多样性与进化实验室 王瑛 整理，有任何错误和不足，欢迎指出。

如有建议或意见，请邮件联系2200801021@cnu.edu.cn

BOLD官网: <http://www.boldsystems.org/>

1. 条码物种识别



DESIGNED TO SUPPORT THE GENERATION & APPLICATION OF DNA BARCODE DATA

BOLD is a cloud-based data storage and analysis platform developed at the Centre for Biodiversity Genomics in Canada. It consists of four main modules, a data portal, an educational portal, a registry of BINs (putative species), and a data collection and analysis workbench.



Enter fasta formatted sequences in the forward orientation:

```
>BFS150331.014
TGACCAAAAAATCAAAATAATGTTGATAAAGAATAGGATCACCTCCTCAGCTGGATCA
AAAAATGAAGTATTTAAATTCGGTCTGTAATAATAGTAATAGCTCCTGCTAAACT
GGTAATGATAATAATAAAAAATGCTGTAATACCGACAGATCATACAAATAGAGGTATT
TGATCAATGATAAATTTTAAATCGTATATTAATGTTGTAATAAAATTAATAGCT
CCTAAATGGAAGAAATACAGCTAAATGTAGAGAAAAATGCTAAATCTACTGAGCTT
CCACATGAGCAATATTAGATGATAATGGAGGATAAATGTTATCCAGTTCGGGCTCCA
TTTCTCAAAATTCCTTGAATTAAGAGTAATAGAAGGGGTAATATCAAAATCTT
ATATTATTTATTCGTGGGAAAGCTATATCAGAGCTCCTAATTTAAAGGTACTAATCAA
TTTCCAAATCCTCAATTATAATAGGTATACTATAAAAAAATTATAATAAAGCATGT
GCAGTTACAATAGTATTATAAATTTGATCATCTCAATTAAGATCCTGATTACCTAAT
TCTGCAGCAATTAATATCTTAAAGAGTTCCTACTATTCTGCTCAATACCAAAAAATA
AAATATAATGTTCAATATCTTTATG
```

1 粘贴.fasta格式序列

2 点击

SUBMIT

Results Summary

Download

Query ID	Best ID	Search DB	Tree	Top %	Graph	Low %
BFS150331.014	No match	COI SPECIES DATABASE	N/A	N/A	N/A	N/A

Query: BFS150331.014

Top Hit: No match

Unable to match any records in the selected database.

BLAST SEQUENCE ON GENBANK

如果没有匹配
则会显示一个按钮
提示可以使用GenBank BLAST

These results may be due to reverse or reverse complement sequences. Please try again with the forward orientation. The following tool can be used to generate the forward orientation https://www.bioinformatics.org/sms/rev_comp.html

这边BOLD数据库无匹配的一个最大的可能是提交的序列由于拼接过程中正反向序列拼接顺序颠倒，导致序列反向互补，此时最好使用相关软件（例如：seqkit）把序列反向互补，虽然这边BOLD也提供了链接直接跳转GenBank 进行BLAST，但是经测试，结果也不是很准确，最理想的方式还是把序列反向互补回去。

PRINT

Results Summary

Download

Query ID	Best ID	Search DB	Tree	Top %	Graph	Low %
BD110707.003	<i>Euplexia lucipara</i>	COI SPECIES DATABASE		100.00		93.76

1 正常的检索结果报告

Query: BD110707.003

Top Hit: Arthropoda Insecta - Lepidoptera - *Euplexia lucipara* (100%)

Search Result:

A species level match could not be made, the queried specimen is likely to be one of the following:

Euplexia lucipara
Niphonyx segregata

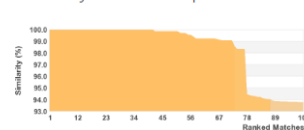
For a hierarchical placement, a neighbor-joining tree is provided:

TREE BASED IDENTIFICATION

Identification Summary

Taxonomic Level	Taxon Assignment	Probability of Placement (%)
Phylum	Arthropoda	100
Class	Insecta	100
Order	Lepidoptera	100
Family	Noctuidae	100
Genus	<i>Euplexia</i>	100

Similarity Scores of Top 100 Matches



Top 20 Matches

Display: Top 20

Phylum	Class	Order	Family	Genus	Species	Subspecies	Similarity (%)	Status
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Euplexia</i>	<i>lucipara</i>		100	Published
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Euplexia</i>	<i>lucipara</i>		100	Published
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Euplexia</i>	<i>lucipara</i>		100	Published
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Euplexia</i>	<i>lucipara</i>		100	Published

这边会显示相似度排名前20的物种鉴定结果，有时你会发现相似度为100%的物种可能不止一种，只是应该校对标本形态数据，不排除别人提交的数据鉴定有误，导致同一条序列会对应不同的物种。

注意：

在考虑存在人为识别错误的情况下，往往匹配到的第一条识别结果不一定是正确的，正如上图，其实除了上面的几条识别结果是*Euplexia lucipara*，其他识别结果都是*Niphonyx segregata*，当然必须参考形态数据。BOLD数据库也提供了物种的标本照供参考。具体操作，见下文。

Top 20 Matches

Display:

Top 20

Phylum	Class	Order	Family	Genus	Species	Subspecies	Similarity (%)	Status
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Euplexia</i>	<i>lucipara</i>		100	Published 🔗
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Euplexia</i>	<i>lucipara</i>		100	Published 🔗
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Euplexia</i>	<i>lucipara</i>		100	Published 🔗
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Euplexia</i>	<i>lucipara</i>		100	Published 🔗
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Euplexia</i>	<i>lucipara</i>		100	Published 🔗
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Niphonyx</i>	<i>segregata</i>		100	Published 🔗
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Niphonyx</i>	<i>segregata</i>		100	Published 🔗
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Niphonyx</i>	<i>segregata</i>		100	Published 🔗
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Niphonyx</i>	<i>segregata</i>		100	Published 🔗
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Niphonyx</i>	<i>segregata</i>		100	Published 🔗
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Niphonyx</i>	<i>segregata</i>		100	Published 🔗
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Niphonyx</i>	<i>segregata</i>		100	Published 🔗
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Niphonyx</i>	<i>segregata</i>		100	Published 🔗
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Niphonyx</i>	<i>segregata</i>		100	Published 🔗
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Niphonyx</i>	<i>segregata</i>		100	Published 🔗
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Niphonyx</i>	<i>segregata</i>		100	Published 🔗
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Niphonyx</i>	<i>segregata</i>		100	Published 🔗
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Niphonyx</i>	<i>segregata</i>		100	Published 🔗
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Niphonyx</i>	<i>segregata</i>		100	Published 🔗
Arthropoda	Insecta	Lepidoptera	Noctuidae	<i>Niphonyx</i>	<i>segregata</i>		100	Published 🔗

2. 物种分类地位查询

BOLDSYSTEMS

DATABASESIDENTIFICATION**TAXONOMY**WORKBENCHRESOURCESLOGIN

分类查询入口

BARCODE OF LIFE DATA SYSTEM v4

Advancing biodiversity science through DNA-based species identification.

EXPLORE THE DATA

DESIGNED TO SUPPORT THE GENERATION & APPLICATION OF DNA BARCODE DATA

BOLD is a cloud-based data storage and analysis platform developed at the Centre for Biodiversity Genomics in Canada. It consists of four main modules, a data portal, an educational portal, a registry of BINs (putative species), and a data collection and analysis workbench.

Niphonyx segregata

SEARCH TAXONOMY

1 输入待检索的物种拉丁名

12,193,872
Specimen Records

2 点击

9,235,973
Specimens with Barcodes

321,110
Species with Barcodes

Animals:

- Acanthocephala [2306]
- Acoelomorpha [20]
- Annelida [102966]
- Arthropoda [10064904]
- Brachiopoda [310]
- Bryozoa [4133]
- Chaetognatha [1743]
- Chordata [839228]
- Cnidaria [29865]
- Ctenophora [511]
- Cyclophora [326]
- Echinodermata [57293]
- Entoprocta [65]
- Gastrotricha [1351]
- Gnathostomulida [24]
- Hemichordata [234]
- Kinorhyncha [720]
- Mollusca [243108]
- Nematoda [34770]
- Nematomorpha [401]
- Nemertea [5669]
- Onychophora [1393]
- Phoronida [160]
- Placozoa [20]
- Platyhelminthes [38589]
- Porifera [7808]
- Priapulida [148]
- Rhombzoa [48]
- Rotifera [12761]
- Sipuncula [1318]
- Tardigrada [2908]
- Xenacoelomorpha [18]

Plants:

- Bryophyta [21902]
- Chlorophyta [14519]
- Lycopodiophyta [1215]
- Magnoliophyta [366736]
- Pinophyta [7068]
- Pteridophyta [11393]
- Rhodophyta [54629]

Fungi:

- Ascomycota [98396]
- Basidiomycota [66496]
- Chytridiomycota [293]
- Glomeromycota [3529]
- Myxomycota [235]
- Zygomycota [3273]

Protists:


- Chlorarachniophyta [67]
- Ciliophora [788]
- Heterokontophyta [7209]
- Pyrrophycomphyta [2337]



目前数据内已有类群分类

Arthropoda / Insecta / Lepidoptera / Noctuidae / Noctuinae / Xylenini / Niphonyx / Niphonyx segregata

2 标本照



1 分类

Taxon Description (Wikipedia)

Niphonyx is a monotypic moth genus of the family Noctuidae erected by Shigero Sugi in 1982. Its only species, *Niphonyx segregata*, the hops angleshade, was first described by Arthur Gardiner Butler in 1878. It is endemic to eastern Asia, including the Russian Far East, the Korean Peninsula, Japan, China and Taiwan. It was introduced to the north-eastern United States in the 1990s and is found from Connecticut south to at least Delaware. The wingspan is 25–30 mm. There are two generations per year in North America. The larvae feed on hop species. [full article at Wikipedia](#)

3 维基百科介绍

维基百科介绍

4 条码组成情况

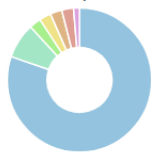
Statistics

Specimen Records:	76
Specimens with Sequences:	76
Specimens with Barcodes:	74
Subspecies:	0
Subspecies with Barcodes:	0
Public Records:	70
Public Subspecies:	0
Public BINs:	3

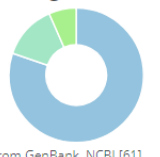
SUBSPECIES LIST

PUBLIC DATA

Specimen Depositories



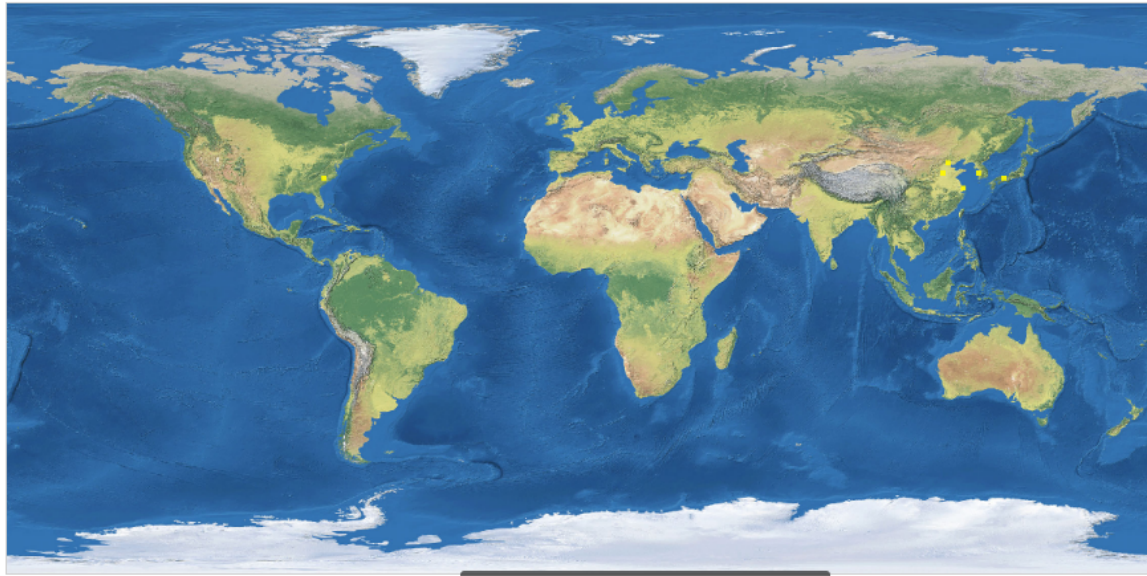
Sequencing Labs



5 条码数据来源分析

条码数据来源分析

Collection Sites



1 数据源在全球的分布

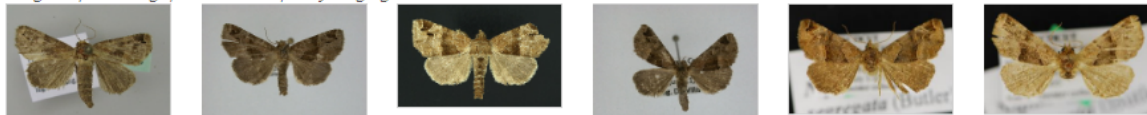
Countries

Collected from 4 countries. Show all countries

China:	9	Japan:	1
United States:	4	Russia:	1

Imagery

Images representing specimens of *Niphonyx segregata*



2 标本照

这些标本照片也是可以通过API批量下载的。

3. 物种条码数据获取

条码数据和下面要提到的物种分布数据都需要从 **Public Data Portal** 模块中下载。

BOLD SYSTEMS

DATABASESIDENTIFICATIONTAXONOMYWORKBENCHRESOURCESLOGINQ

1 公共数据集

Public Data Portal

A data retrieval interface that allows for searching all 1.3M public records in BOLD using multiple search criteria including, but not limited to, geography, taxonomy, and depository. Search results can be summarized, plotted on high-res maps, and downloaded.

2 条形码索引号数据库

BIN Database

A searchable database of Barcode Index Numbers (BINs), sequence clusters that closely approximate species. This system allows for rapid validation and use of barcode data where taxonomic data are lacking or unverified.

3 与条码相关的文献

Publications

A collection of barcode publications and publications that have utilized barcode records.

4 引物数据库

Primer Database

A comprehensive registry of primers used in the generation of barcode sequences. The registry is maintained by users of BOLD.

1 输入需要检索的物种学名

除序列以外的物种信息

文本格式 网页格式

Specimens: [DWC](#) [XML](#) [TSV](#)

Sequences: [FASTA](#) [TRACE](#)

Combined: [XML](#) [TSV](#)

Map: Generate from

只有物种和序列的.fasta格式

上面二者结合

2 该物种的序列

3 可以下载的数据

峰图文件

Results Summary

Found 70 published records, with 70 records with sequences, forming 3 BINs (clusters), with specimens from 5 countries, deposited in 7 institutions.

Of these records, 70 have species names, and represent 1 species.

Specimen Distribution

有4种模式地图可选，和下面的地图效果差不多

4. 物种分布数据获取

见上图。

5. 文章发表前条码数据提交和更新

此模块演示起来比较困难，建议直接参考官网提供的手册：["BOLD Print Handbook for BOLD v4"](#)

6. 其他资源

BOLD SYSTEMS

DATABASES IDENTIFICATION TAXONOMY WORKBENCH RESOURCES LOGIN Q

1 BOLD F.A.Q. BOLD使用可能遇到的常见问题的解答

2 BOLD Print Handbook for BOLD v4 1st Draft 第4版的使用手册

3 BOLD Print Handbook for BOLD v3 第3版的使用手册

4 Citing Use of BOLD Systems 数据库应用格式

5 Data Releases 条码相关的一些重要项目介绍

6 BOLD APIs API 的介绍，不同的模块有不同的API，详细介绍了使用的规则，还有例子演示

这些内容在这里

BOLD F.A.Q.
Frequently Asked Questions for new users and registered users.

BOLD Print Handbook for BOLD v4 1st Draft
This is the PDF version of the handbook for BOLD 4 version, produced in July 2019.

BOLD Print Handbook for BOLD v3
This is the PDF version of the handbook for BOLD 3 version, produced in September 2012.

Citing Use of BOLD Systems
Ratnasingham, S. & Hebert, P. D. N. (2007). BOLD: The Barcode of Life Data System ([www.barcodinglife.org](#)). Molecular Ecology Notes 7, 355-364. DOI: 10.1111/j.1471-8286.2006.01678.x

- 2973 citations circa April 2019

Ratnasingham S, Hebert PDN (2013) A DNA-Based Registry for All Animal Species: The Barcode Index Number (BIN) System. PLoS ONE 8(8): e66213. DOI:10.1371/journal.pone.0066213

- 824 citations circa April 2019

Community Standards
Manuscripts regarding DNA Barcoding standards for Animals, Plants and Fungi.

Data Releases
Data Release packages for the first IBOL project, BARCODE 500K, and the Canadian Barcode of Life Network project.

BOLD APIs
BOLD Web Services that provide the ability to query and retrieve data from the BOLD.