“实用生物信息技术”课程小组讨论总结报告

组：G2　　次：1　　组长：边汉青　　执笔：边汉青

1. 时间

2026.3.25

2. 方式

线上

3. 主题

对于NCBI、EBI、CNCB等内容的复习和对序列比对内容的预习。

4. 内容

A 关于NCBI、CNCB、EBI网站基本功能的阐述和讨论

A1　NCBI基本功能的阐述

A2　CNCB基本功能的阐述

A3　EBI基本功能的阐述

B 网络书籍、文档回顾与应用

B1　生物大分子月报（Molecular of the Month）

B2　化学小分子月报（Molecular of the Month）

B3　蛋白质分子精选（Protein Spotlight）

C 序列比对演示。

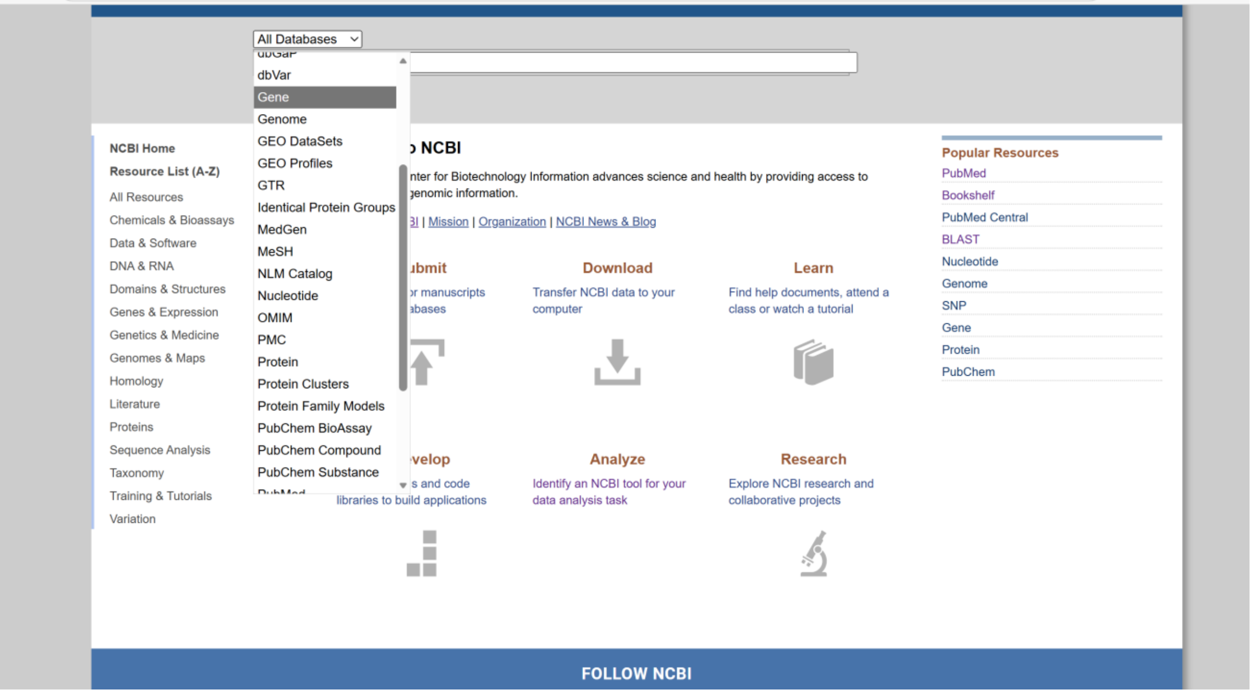
C1　以HBA_HUMAN和HBA_MOUSE为例在Dotlet网站中进行序列比对。

C2　以斑头雁alpha血红蛋白HBA_ANSIN和灰雁alpha血红蛋白序列HBA_ANSAN为例在EMBOSS Explorer平台进行序列比对。

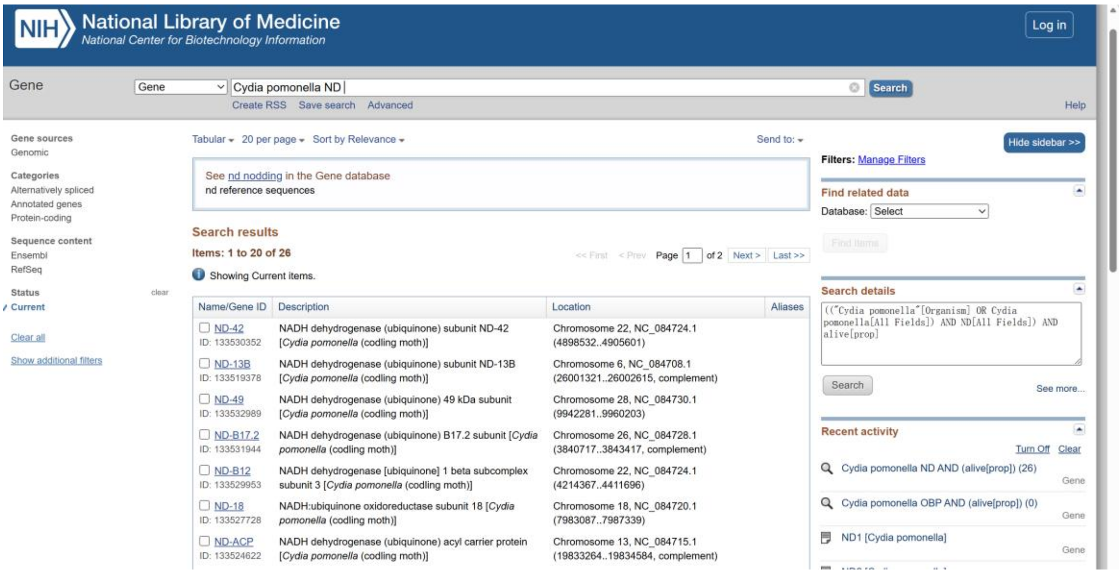
A1 NCBI 基本功能的阐述

（1）序列查找

可在主页搜索栏进行搜索，其中常用数据库中 Nucleotide 方便下载基因序列。Gene 中的基因序列则信息更多更全面，protein 则用于搜索氨基酸序列。例如:在 Gene 中搜索苹果蠹蛾的 ND 基因序列，点击搜索结果可以查看基因详细信息。



例如：在 Gene 中搜索苹果蠹蛾的 ND 基因序列，并点击搜索结果查看基因详细信息。



Gene: Cydia pomonella ND

See nd nodding in the Gene database
nd reference sequences

Search results

Items: 1 to 20 of 26

Page 1 of 2

Showing Current items.

Name/Gene ID	Description	Location	Aliases
☐ ND-42 ID: 133530352	NADH dehydrogenase (ubiquinone) subunit ND-42 [*Cydia pomonella* (codling moth)]	Chromosome 22, NC_084724.1 (4898532..4905601)	
☐ ND-13B ID: 133519378	NADH dehydrogenase (ubiquinone) subunit ND-13B [*Cydia pomonella* (codling moth)]	Chromosome 6, NC_084708.1 (26001321..26002615, complement)	
☐ ND-49 ID: 133532989	NADH dehydrogenase (ubiquinone) 49 kDa subunit [*Cydia pomonella* (codling moth)]	Chromosome 28, NC_084730.1 (9942281..9960203)	
☐ ND-B17.2 ID: 133531944	NADH dehydrogenase (ubiquinone) B17.2 subunit [*Cydia pomonella* (codling moth)]	Chromosome 26, NC_084728.1 (3840717..3843417, complement)	
☐ ND-B12 ID: 133529953	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 3 [*Cydia pomonella* (codling moth)]	Chromosome 22, NC_084724.1 (4214367..4411696)	
☐ ND-18 ID: 133527728	NADH:ubiquinone oxidoreductase subunit 18 [*Cydia pomonella* (codling moth)]	Chromosome 18, NC_084720.1 (7983087..7987339)	
☐ ND-ACP ID: 133524622	NADH dehydrogenase (ubiquinone) acyl carrier protein [*Cydia pomonella* (codling moth)]	Chromosome 13, NC_084715.1 (19833264..19834584, complement)	

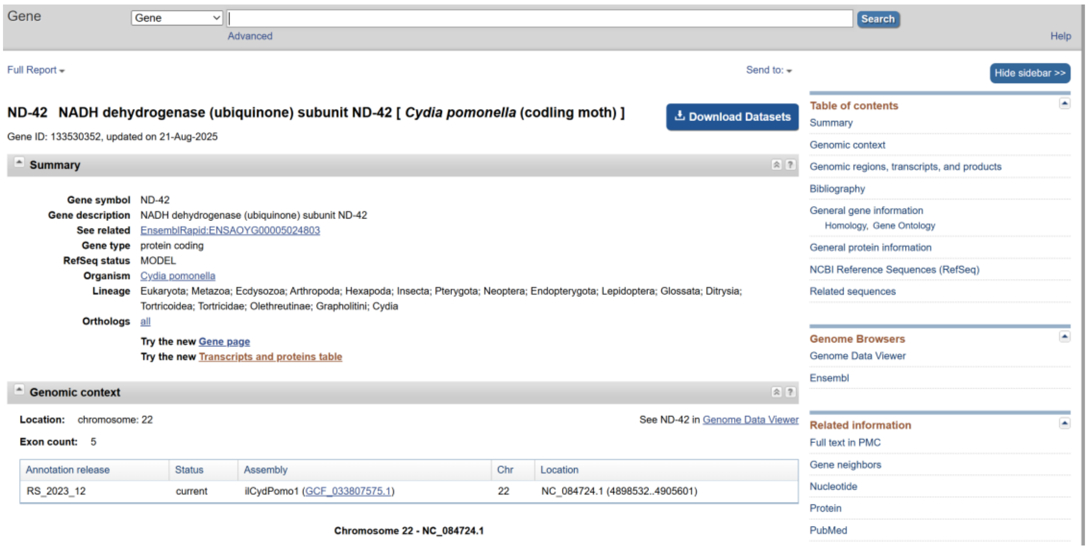
Database: Select

Search details

```
(("Cydia pomonella"[Organism] OR Cydia pomonella[All Fields]) AND ND[All Fields]) AND alive[prop]
```

Recent activity

- Cydia pomonella ND AND (alive[prop]) (26) Gene
- Cydia pomonella OBP AND (alive[prop]) (0) Gene
- ND1 [Cydia pomonella] Gene



ND-42 NADH dehydrogenase (ubiquinone) subunit ND-42 [*Cydia pomonella* (codling moth)]

Gene ID: 133530352, updated on 21-Aug-2025

Summary

Gene symbol ND-42
Gene description NADH dehydrogenase (ubiquinone) subunit ND-42
See related EnsemblRapid:ENSAOYG00005024803
Gene type protein coding
RefSeq status MODEL
Organism Cydia pomonella
Lineage Eukaryota; Metazoa; Ecdysozoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Tortricoidea; Tortricidae; Olethreutinae; Grapholitini; Cydia
Orthologs all

Try the new Gene page
Try the new Transcripts and proteins table

Genomic context

Location: chromosome: 22

See ND-42 in Genome Data Viewer

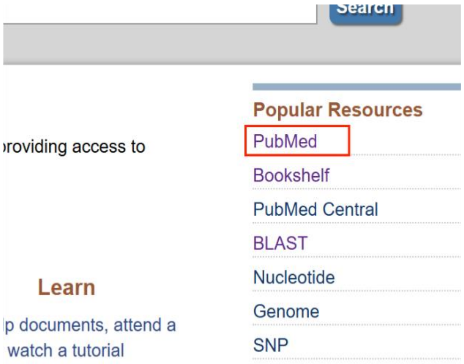
Exon count: 5

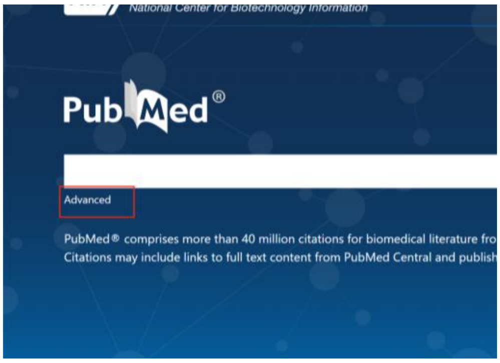
Annotation release	Status	Assembly	Chr	Location
RS_2023_12	current	ilCydPomo1 (GCF_033807575.1)	22	NC_084724.1 (4898532..4905601)

Chromosome 22 - NC_084724.1

（2）文献检索

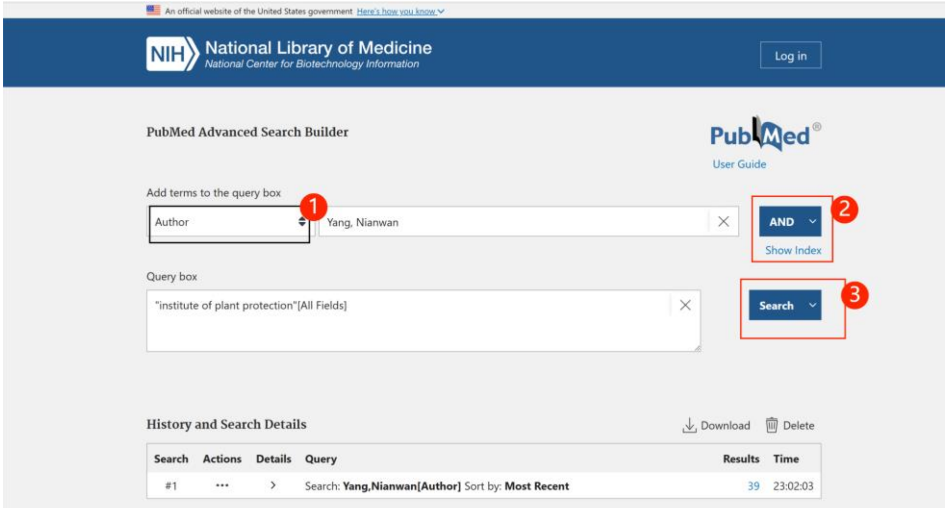
进入 PubMed 进行文献检索，击“advance”可添加更多检索信息进行高级检索。





PubMed® comprises more than 40 million citations for biomedical literature fro
Citations may include links to full text content from PubMed Central and publish

在 1 处可选择添加检索信息的类别，如作者、日期、机构等，输入检索信息后点击 2 处“show index”可出现索引，点击 2 处蓝色方框添加索引信息至 Query box，多个索引信息可以选择是否并列，点击 2 处展开箭头即可选择，添加好索引信息后点击 3 处即可检索。



PubMed Advanced Search Builder

Add terms to the query box

Author: Yang, Nianwan

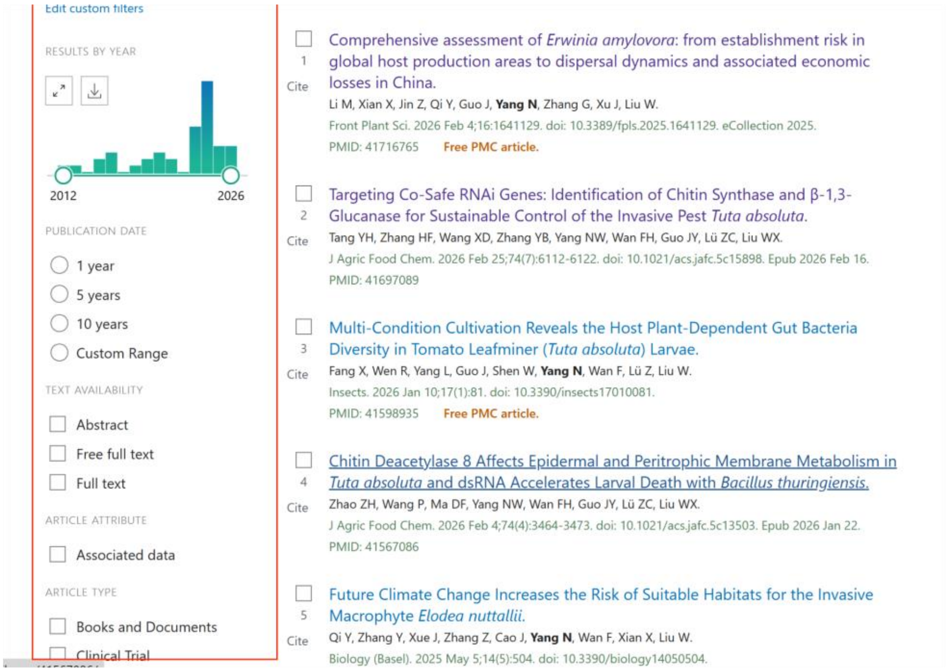
Query box

"institute of plant protection"[All Fields]

History and Search Details

Search	Actions	Details	Query	Results	Time
#1	•••	›	Search: **Yang,Nianwan[Author]** Sort by: **Most Recent**	39	23:02:03

检索结果左侧可进行如年限、文章可用性、文章属性、文章类型等的筛选。



☐ 1 Comprehensive assessment of *Erwinia amylovora*: from establishment risk in global host production areas to dispersal dynamics and associated economic losses in China.
Li M, Xian X, Jin Z, Qi Y, Guo J, **Yang N**, Zhang G, Xu J, Liu W.
Front Plant Sci. 2026 Feb 4;16:1641129. doi: 10.3389/fpls.2025.1641129. eCollection 2025.
PMID: 41716765 Free PMC article.

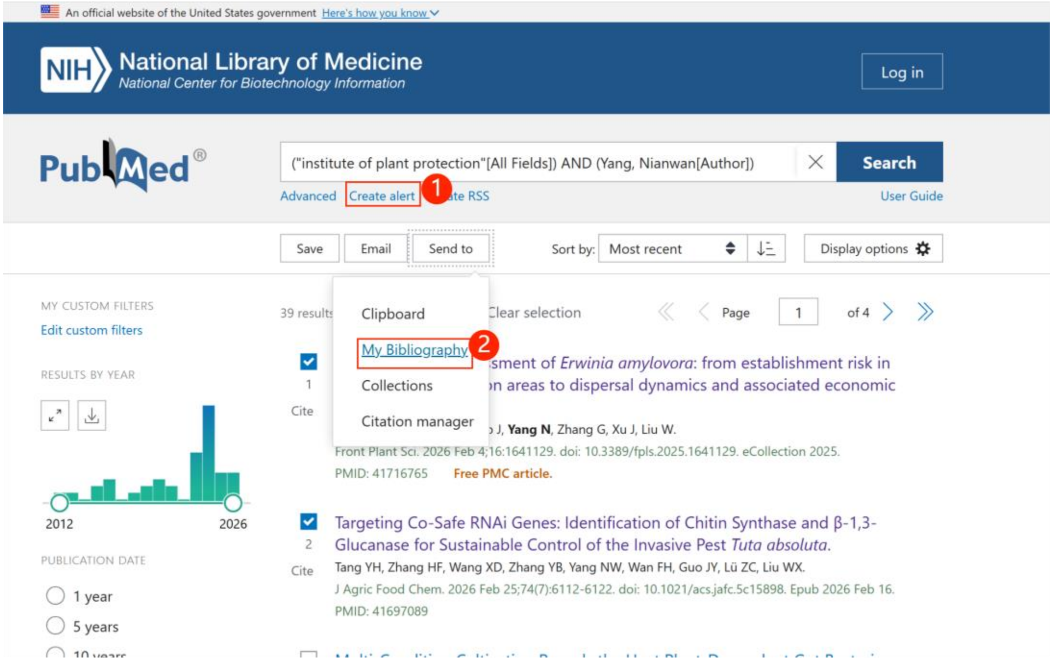
☐ 2 Targeting Co-Safe RNAi Genes: Identification of Chitin Synthase and β-1,3-Glucanase for Sustainable Control of the Invasive Pest *Tuta absoluta*.
Tang YH, Zhang HF, Wang XD, Zhang YB, Yang NW, Wan FH, Guo JY, Lü ZC, Liu WX.
J Agric Food Chem. 2026 Feb 25;74(7):6112-6122. doi: 10.1021/acs.jafc.5c15898. Epub 2026 Feb 16.
PMID: 41697089

☐ 3 Multi-Condition Cultivation Reveals the Host Plant-Dependent Gut Bacteria Diversity in Tomato Leafminer (*Tuta absoluta*) Larvae.
Fang X, Wen R, Yang L, Guo J, Shen W, **Yang N**, Wan F, Lü Z, Liu W.
Insects. 2026 Jan 10;17(1):81. doi: 10.3390/insects17010081.
PMID: 41598935 Free PMC article.

☐ 4 Chitin Deacetylase 8 Affects Epidermal and Peritrophic Membrane Metabolism in *Tuta absoluta* and dsRNA Accelerates Larval Death with *Bacillus thuringiensis*.
Zhao ZH, Wang P, Ma DF, Yang NW, Wan FH, Guo JY, Lü ZC, Liu WX.
J Agric Food Chem. 2026 Feb 4;74(4):3464-3473. doi: 10.1021/acs.jafc.5c13503. Epub 2026 Jan 22.
PMID: 41567086

☐ 5 Future Climate Change Increases the Risk of Suitable Habitats for the Invasive Macrophyte *Elodea nuttallii*.
Qi Y, Zhang Y, Xue J, Zhang Z, Cao J, **Yang N**, Wan F, Xian X, Liu W.
Biology (Basel). 2025 May 5;14(5):504. doi: 10.3390/biology14050504.

勾选想要的文章，点击 1 处可保存文章至个人账号中查看，并且可以勾选通知相关文章的最新发表消息；点击 Send to 可以将文章发送到“My Bibliography”,可以登录后在个人账号中查看，也可以分享个人链接给其他人查看。

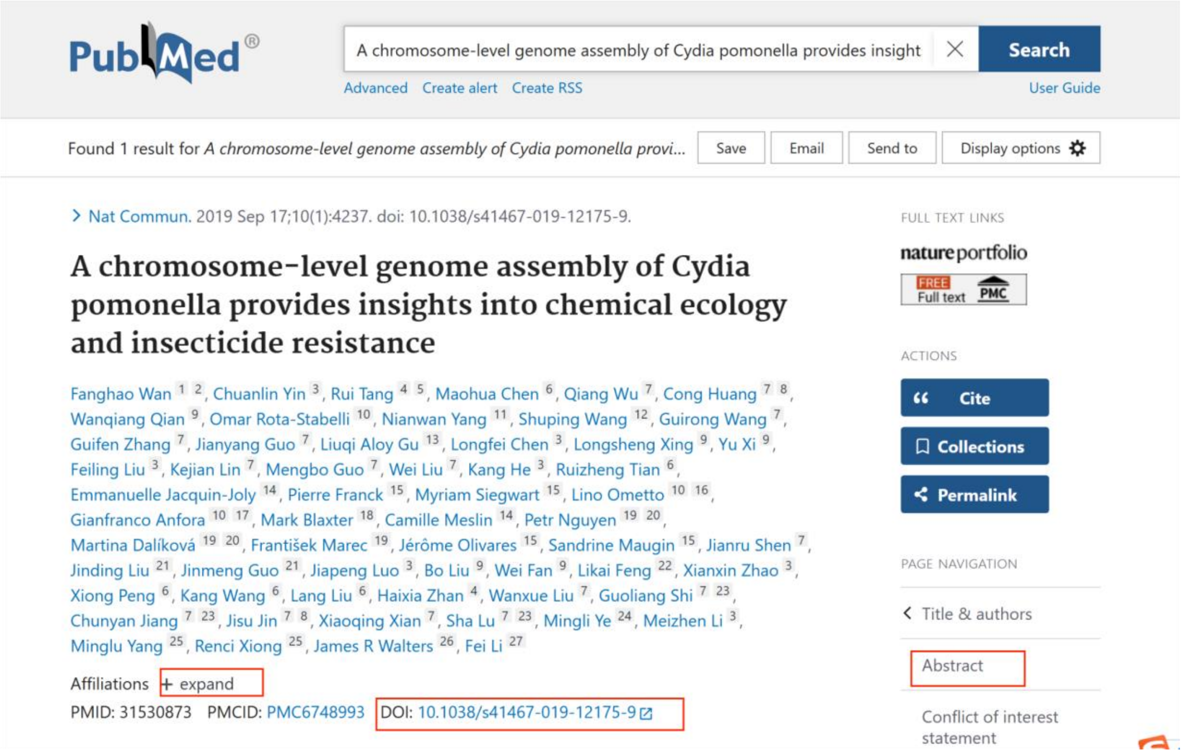


("institute of plant protection"[All Fields]) AND (Yang, Nianwan[Author])

☑ 1 …sment of *Erwinia amylovora*: from establishment risk in …n areas to dispersal dynamics and associated economic
…J, **Yang N**, Zhang G, Xu J, Liu W.
Front Plant Sci. 2026 Feb 4;16:1641129. doi: 10.3389/fpls.2025.1641129. eCollection 2025.
PMID: 41716765 Free PMC article.

☑ 2 Targeting Co-Safe RNAi Genes: Identification of Chitin Synthase and β-1,3-Glucanase for Sustainable Control of the Invasive Pest *Tuta absoluta*.
Tang YH, Zhang HF, Wang XD, Zhang YB, Yang NW, Wan FH, Guo JY, Lü ZC, Liu WX.
J Agric Food Chem. 2026 Feb 25;74(7):6112-6122. doi: 10.1021/acs.jafc.5c15898. Epub 2026 Feb 16.
PMID: 41697089

点开搜索出的文献，可以查看文献摘要、引用文献和被引用的文献等相关信息，点击 expand 可查看隶属关系，点击文章的 DOI 号就可以跳转全文，进行阅读或下载。



A chromosome-level genome assembly of Cydia pomonella provides insight

Found 1 result for *A chromosome-level genome assembly of Cydia pomonella provi...*

Nat Commun. 2019 Sep 17;10(1):4237. doi: 10.1038/s41467-019-12175-9.

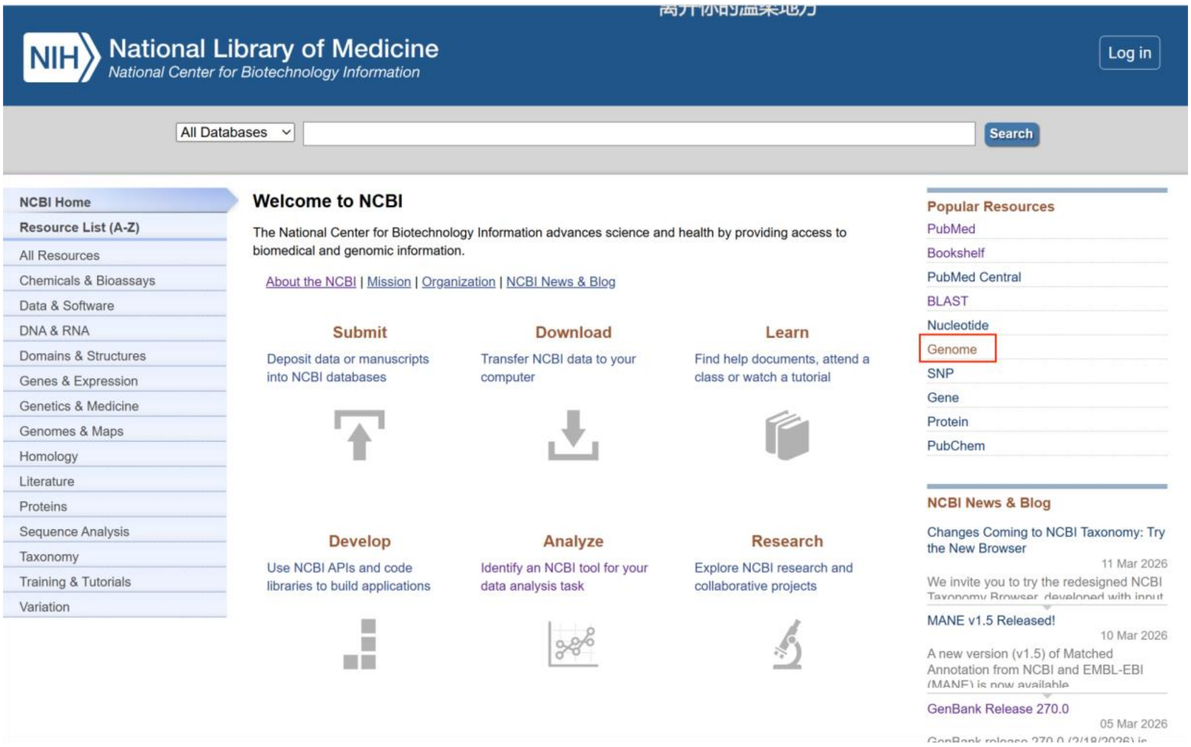
A chromosome-level genome assembly of Cydia pomonella provides insights into chemical ecology and insecticide resistance

Fanghao Wan [1, 2], Chuanlin Yin [3], Rui Tang [4, 5], Maohua Chen [6], Qiang Wu [7], Cong Huang [7, 8], Wanqiang Qian [9], Omar Rota-Stabelli [10], Nianwan Yang [11], Shuping Wang [12], Guirong Wang [7], Guifen Zhang [7], Jianyang Guo [7], Liuqi Aloy Gu [13], Longfei Chen [3], Longsheng Xing [9], Yu Xi [9], Feiling Liu [3], Kejian Lin [7], Mengbo Guo [7], Wei Liu [7], Kang He [3], Ruizheng Tian [6], Emmanuelle Jacquin-Joly [14], Pierre Franck [15], Myriam Siegwart [15], Lino Ometto [10, 16], Gianfranco Anfora [10, 17], Mark Blaxter [18], Camille Meslin [14], Petr Nguyen [19, 20], Martina Dalíková [19, 20], František Marec [19], Jérôme Olivares [15], Sandrine Maugin [15], Jianru Shen [7], Jinding Liu [21], Jinmeng Guo [21], Jiapeng Luo [3], Bo Liu [9], Wei Fan [9], Likai Feng [22], Xianxin Zhao [3], Xiong Peng [6], Kang Wang [6], Lang Liu [6], Haixia Zhan [4], Wanxue Liu [7], Guoliang Shi [7, 23], Chunyan Jiang [7, 23], Jisu Jin [7, 8], Xiaoqing Xian [7], Sha Lu [7, 23], Mingli Ye [24], Meizhen Li [3], Minglu Yang [25], Renci Xiong [25], James R Walters [26], Fei Li [27]

Affiliations + expand

PMID: 31530873 PMCID: PMC6748993 DOI: 10.1038/s41467-019-12175-9

（3）下载基因组学数据



Welcome to NCBI

The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.

Deposit data or manuscripts into NCBI databases

Transfer NCBI data to your computer

Find help documents, attend a class or watch a tutorial

Use NCBI APIs and code libraries to build applications

Identify an NCBI tool for your data analysis task

Explore NCBI research and collaborative projects

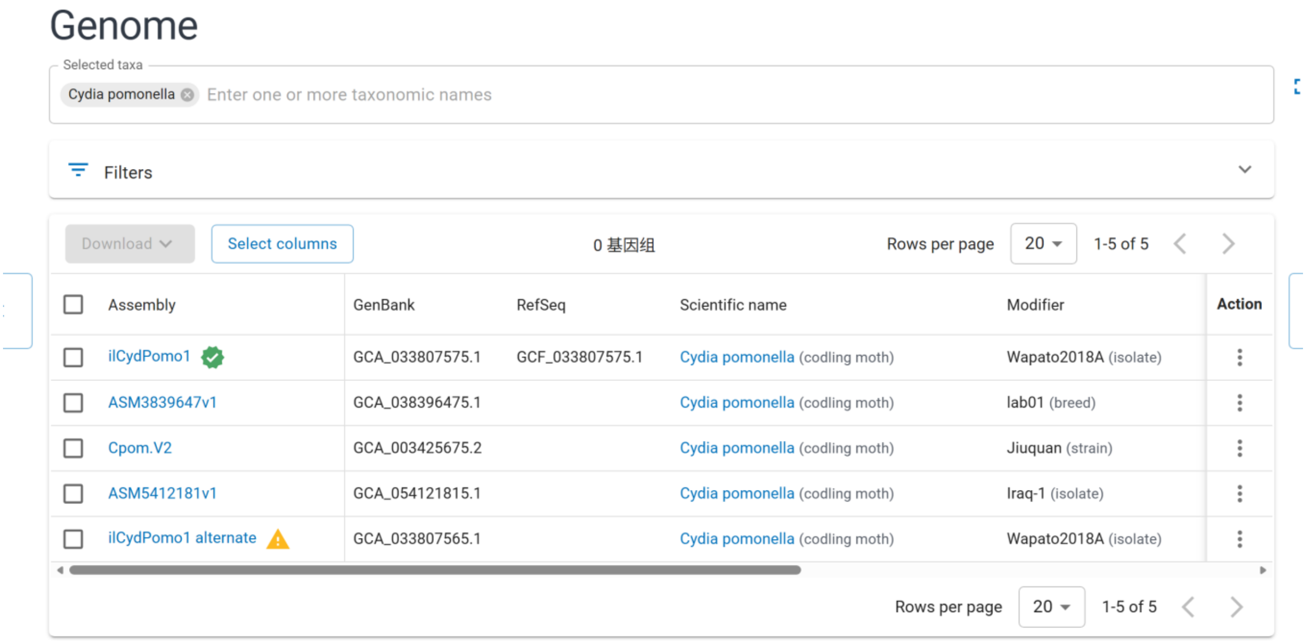
Changes Coming to NCBI Taxonomy: Try the New Browser
11 Mar 2026
We invite you to try the redesigned NCBI

MANE v1.5 Released!
10 Mar 2026
A new version (v1.5) of Matched Annotation from NCBI and EMBL-EBI

GenBank Release 270.0
05 Mar 2026

搜索苹果蠹蛾的基因组数据可以得到以下五个结果，它们可能来自不同地理种群，第一行 iLCydPomo1 是质量最高的版本（唯一一个同时拥有 GCA 和 GCF 编号），通常作为后续研究的参考基因。Assembly 是组装版本名称，第二列 GenBank 以 GCA 开头，是这些数据在 GenBank 数据库中的原始序列号；第三列 RefSeq 以 GCF 开头，是 NCBI 官方审核整理的参考序列号，第四列

Scientific name 是物种学名，第五列 Modifier 说明这些基因组的来源，例如 Wapato2018A 是来自美国华盛顿州的分离株，lab01 是实验室培育品系等等，最后一列 Action 可以点击下载数据或查看详细信息。



Genome

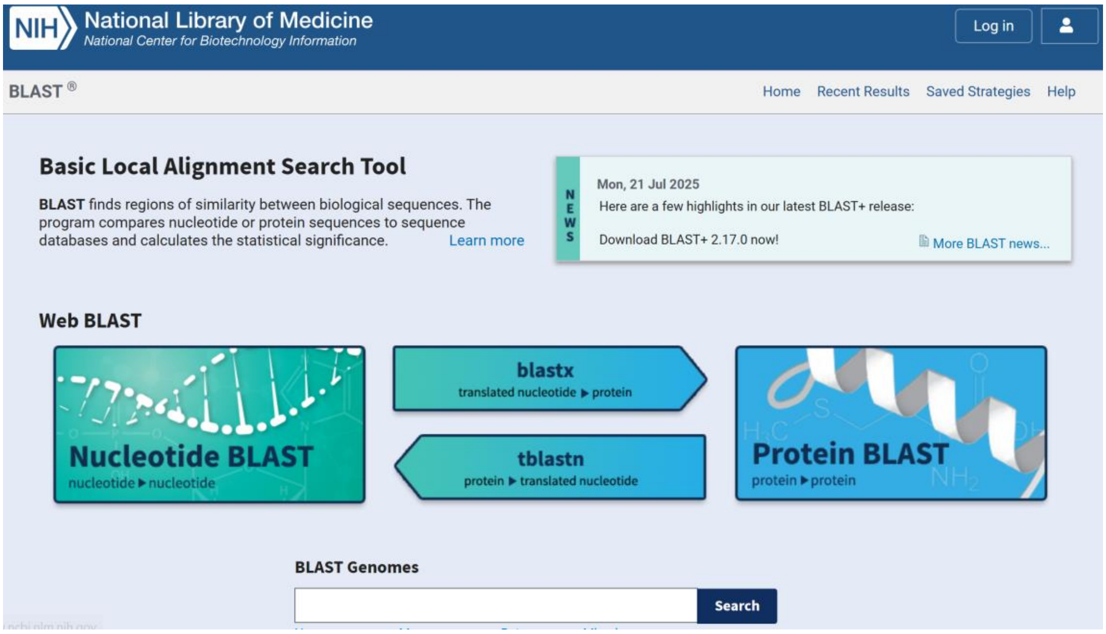
Selected taxa: Cydia pomonella ⊗ Enter one or more taxonomic names

Filters

☐	Assembly	GenBank	RefSeq	Scientific name	Modifier	Action
☐	ilCydPomo1 ✓	GCA_033807575.1	GCF_033807575.1	Cydia pomonella (codling moth)	Wapato2018A (isolate)	⋮
☐	ASM3839647v1	GCA_038396475.1		Cydia pomonella (codling moth)	lab01 (breed)	⋮
☐	Cpom.V2	GCA_003425675.2		Cydia pomonella (codling moth)	Jiuquan (strain)	⋮
☐	ASM5412181v1	GCA_054121815.1		Cydia pomonella (codling moth)	Iraq-1 (isolate)	⋮
☐	ilCydPomo1 alternate ⚠	GCA_033807565.1		Cydia pomonella (codling moth)	Wapato2018A (isolate)	⋮

（4）常见分析

blast 比对分析：Nuclceotide blast 是核苷酸序列比对，protein blast 是蛋白质序列比对



Basic Local Alignment Search Tool

BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance. Learn more

NEWS

Mon, 21 Jul 2025

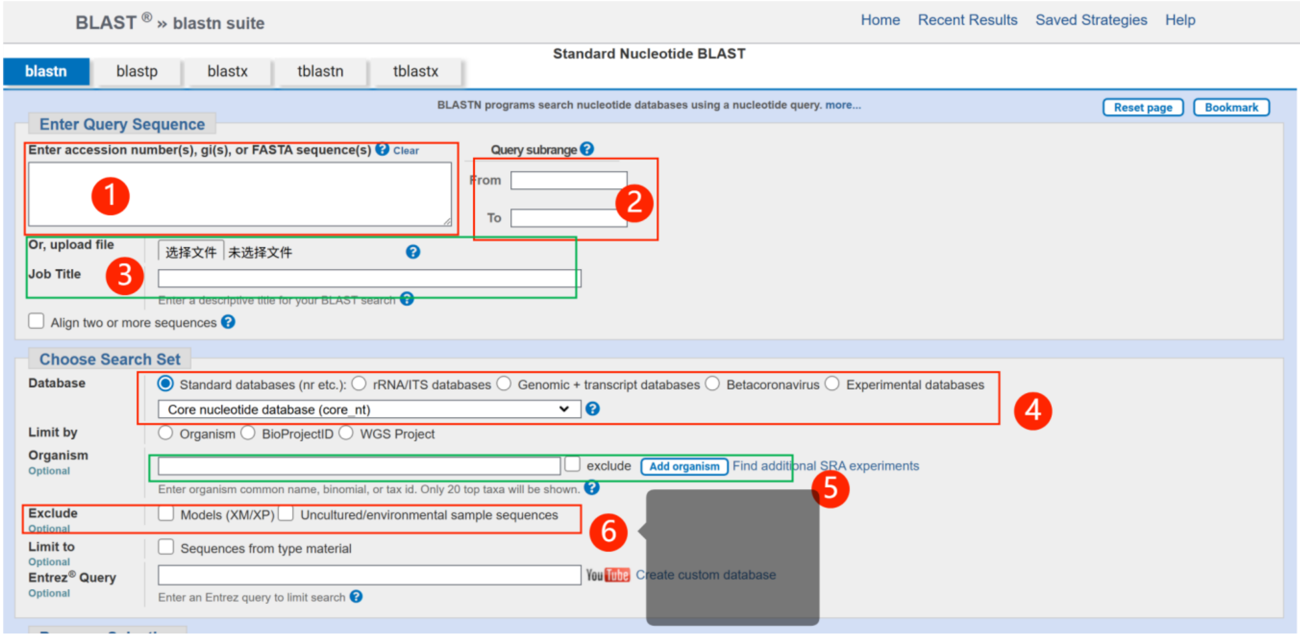
Here are a few highlights in our latest BLAST+ release:

Download BLAST+ 2.17.0 now!

More BLAST news...

Web BLAST

以核苷酸序列比对为例，1 处可粘贴要比对的序列，2 处选择要比对的位置，3 处可以上传序列文件并命名，4 处选择数据库，5 可以选择物种，6 处可以排除特定序列或环境序列。



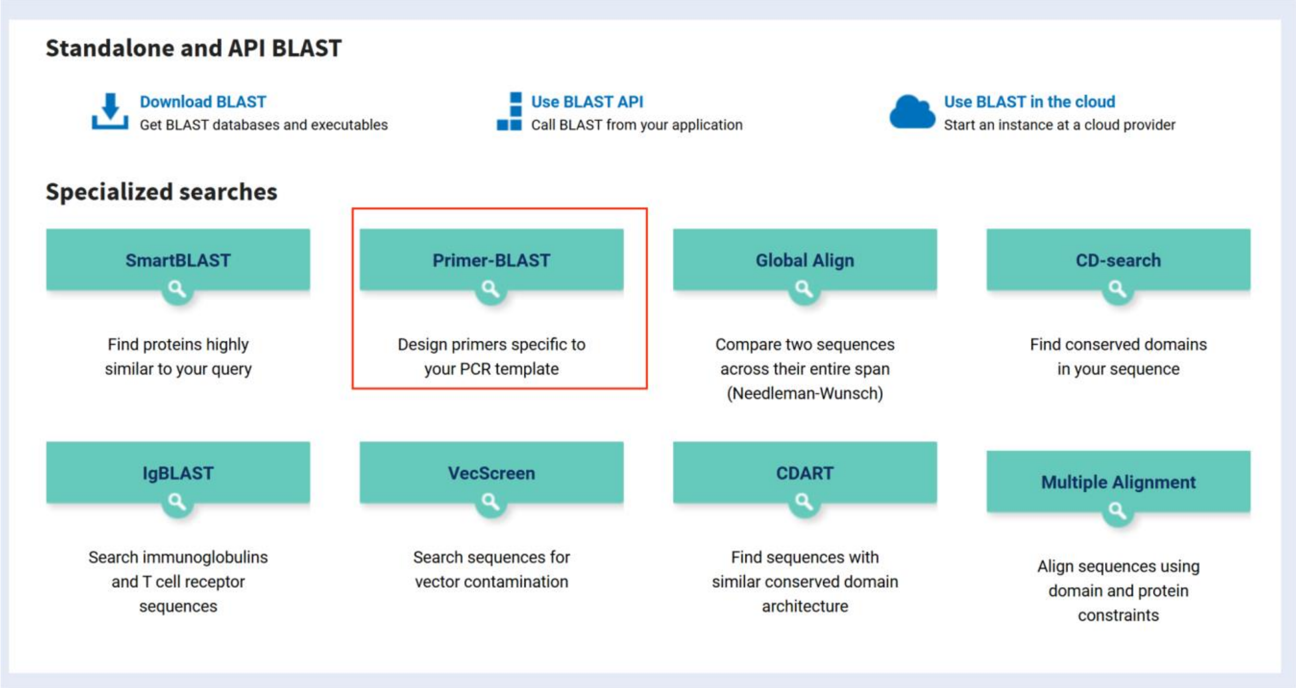
BLASTN programs search nucleotide databases using a nucleotide query. more...

Enter a descriptive title for your BLAST search

Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown.

Enter an Entrez query to limit search

引物设计：



Get BLAST databases and executables

Call BLAST from your application

Start an instance at a cloud provider

Find proteins highly similar to your query

Design primers specific to your PCR template

Compare two sequences across their entire span (Needleman-Wunsch)

Find conserved domains in your sequence

Search immunoglobulins and T cell receptor sequences

Search sequences for vector contamination

Find sequences with similar conserved domain architecture

Align sequences using domain and protein constraints

Finding primers specific to your PCR template (using Primer3 and BLAST).

Primers for target on one template | Primers common for a group of sequences

Retrieve recent results Publication Tips for finding specific primers

Save search parameters Reset page

PCR Template

Enter accession, gi, or FASTA sequence (A refseq record is preferred) Clear

序列

Or, upload FASTA file 选择文件 未选择文件

Range Clear

Forward primer From To

Reverse primer

引物位置范围

Primer Parameters

Use my own forward primer (5'->3' on plus strand) Clear

Use my own reverse primer (5'->3' on minus strand) Clear

PCR product size Min Max

引物长度

\# of primers to return

Primer melting temperatures (T_m) Min Opt Max Max T_m difference

退火温度

Primer Location Preference ☐ Prefer primers at 3' side of the template

内含子、外显子相关设置

Exon/Intron selection A refseq mRNA sequence as PCR template input is required for options in the section

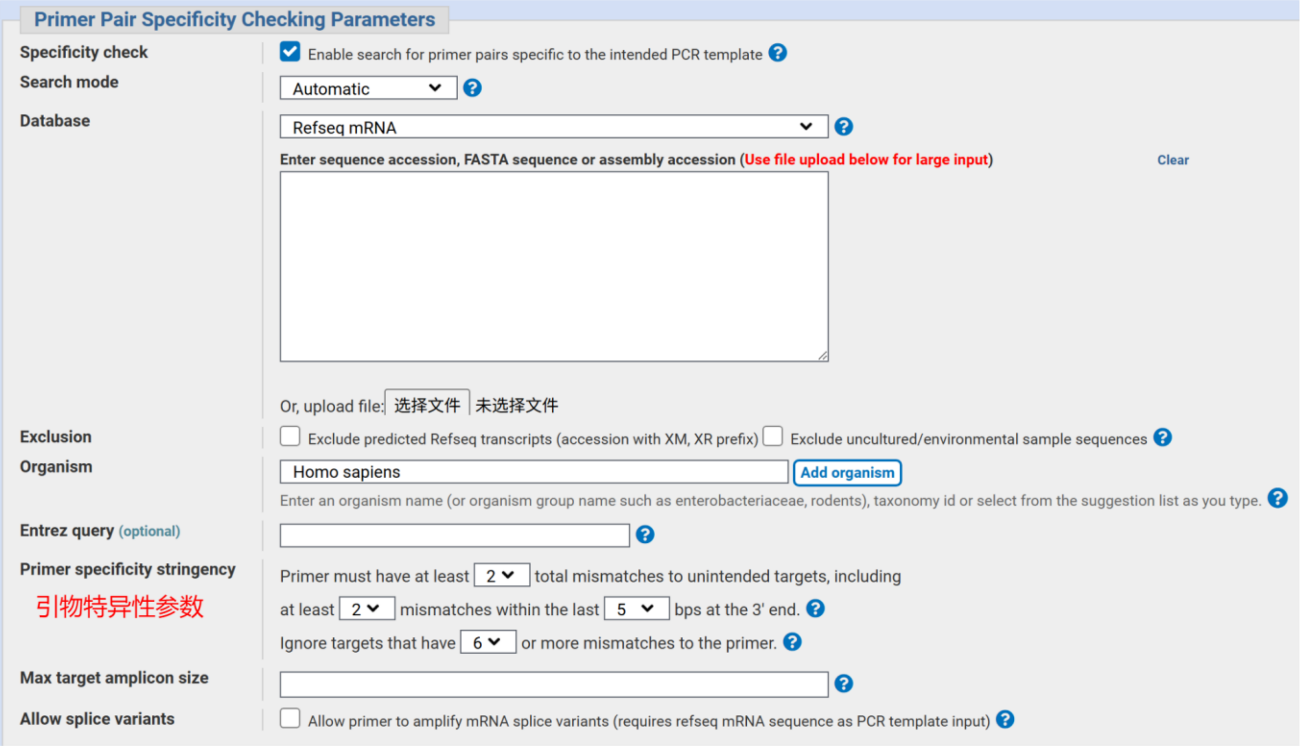
Exon junction span: No preference

Exon junction match: Min 5' match Min 3' match Max 3' match

Minimal and maximal number of bases that must anneal to exons at the 5' or 3' side of the junction

Intron inclusion ☐ Primer pair must be separated by at least one intron on the corresponding genomic DNA

Intron length range Min Max



Primer Pair Specificity Checking Parameters

Specificity check ☑ Enable search for primer pairs specific to the intended PCR template

Search mode: Automatic

Database: Refseq mRNA

Enter sequence accession, FASTA sequence or assembly accession (Use file upload below for large input) Clear

Or, upload file: 选择文件 未选择文件

Exclusion ☐ Exclude predicted Refseq transcripts (accession with XM, XR prefix) ☐ Exclude uncultured/environmental sample sequences

Organism: Homo sapiens Add organism

Enter an organism name (or organism group name such as enterobacteriaceae, rodents), taxonomy id or select from the suggestion list as you type.

Entrez query (optional)

Primer specificity stringency

引物特异性参数

Primer must have at least 2 total mismatches to unintended targets, including

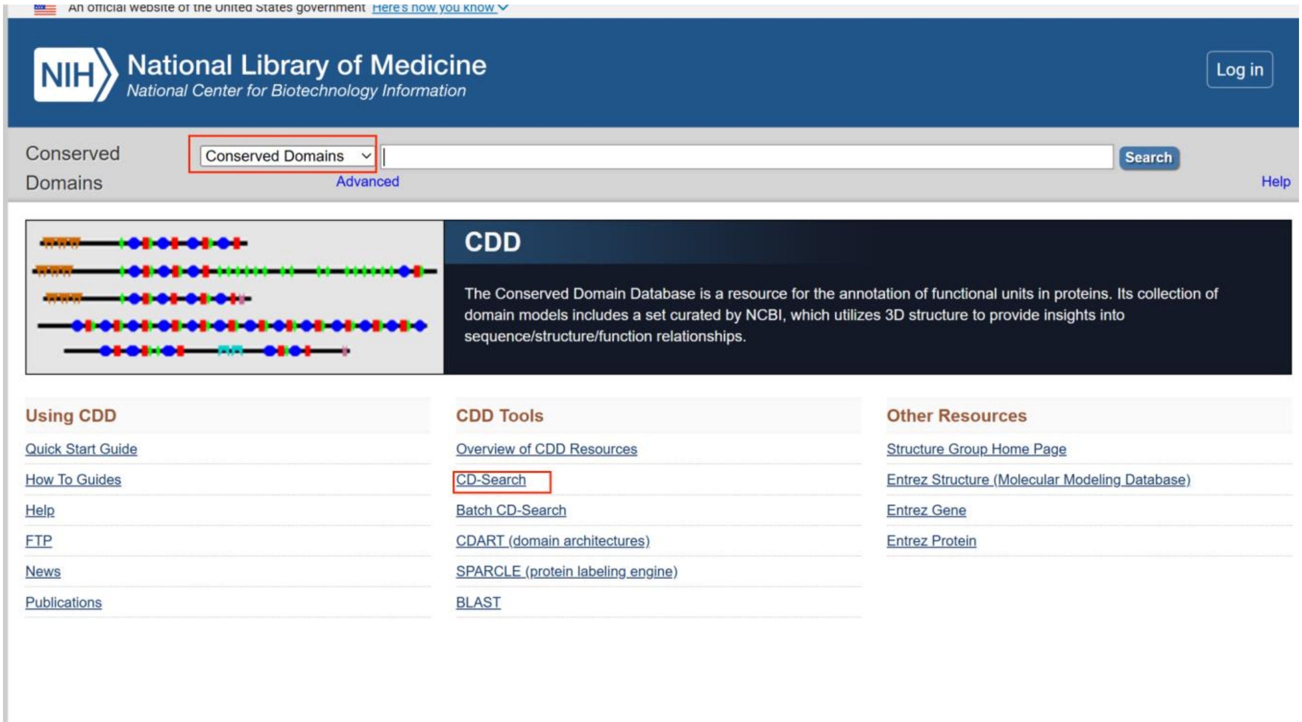
at least 2 mismatches within the last 5 bps at the 3' end.

Ignore targets that have 6 or more mismatches to the primer.

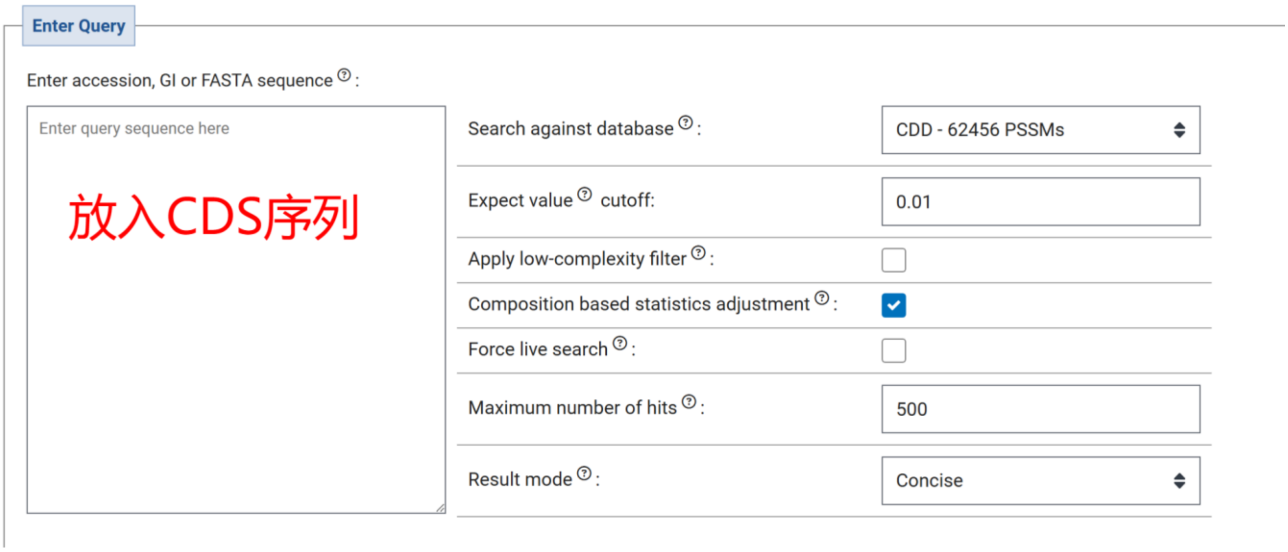
Max target amplicon size

Allow splice variants ☐ Allow primer to amplify mRNA splice variants (requires refseq mRNA sequence as PCR template input)

保守结构域预测：



The Conserved Domain Database is a resource for the annotation of functional units in proteins. Its collection of domain models includes a set curated by NCBI, which utilizes 3D structure to provide insights into sequence/structure/function relationships.



Search against database ⓘ: CDD - 62456 PSSMs

Expect value ⓘ cutoff: 0.01

Apply low-complexity filter ⓘ: ☐

Composition based statistics adjustment ⓘ: ☑

Force live search ⓘ: ☐

Maximum number of hits ⓘ: 500

Result mode ⓘ: Concise

A2 CNCB 基本功能的阐述和讨论

（1）基本介绍

CNCB（国家生物信息中心，China National Center for Bioinformation）于 2019 年 11 月 13 日经中央编办批复，在中国科学院北京基因组研究所（BIG，Beijing Institute of Genomics）加挂成立，是我国生命科学领域国家级生物数据基础设施与核心枢纽。CNCB 以数据覆盖数据资源、计算分析、专项服务三大核心模块。

（2）数据资源模块

CNCB 的核心功能以数据资源模块为基础，涵盖基础科研数据、专项数据与新兴多模态数据，其中基础科研数据包括生物项目（BioProject）、生物样本（BioSample）、组学原始数据，专项数据涵盖人源数据（Human-derived Data）、微生物组（Microbiome）、遗传变异（Genetic Variation），新兴数据则

包括医学影像（Medical Imaging）、算法软件工具等。

同时配套科学项目数据汇交服务系统、人类遗传资源信息管理备份、新冠数据快速递交等专题服务，保障国家重大专项与生物安全数据的规范管理。该模块依托组学原始数据归档库（GSA，Genome Sequence Archive）、基因组序列仓库（GWH，Genome Warehouse）、基因组变异库（GVM，Genome Variation Map）、国家微生物科学数据中心（NMDC，National Microbiology Data Center）等 110 余个核心数据库，实现多组学数据的长期安全存储与结构化分类，截至 2026 年 2 月累计数据量超 100PB；其中 GSA 作为国内首个获国际顶刊认可的组学数据存储库，获 Springer Nature、Elsevier、Wiley、Cell 等国际出版集团认可。

（3）计算分析模块

计算分析模块作为 CNCB 衔接数据资源与科研成果的核心纽带，提供云端化、标准化的全流程生物信息分析服务。

其中序列比对（BLAST，Basic Local Alignment Search Tool）可实现核酸（Nucleotide）与蛋白（Protein）序列的相似性、同源性比对，为基因功能注释与分子进化研究提供基础；

生物信息学云平台（BiT，Bioinformatics Cloud platform）集成常用算法、软件与绘图工具，支持全类型组学数据的规模化分析；

基因组组装与注释可将高通量测序原始数据组装为完整基因组，并对基因结构、功能元件、调控区域进行系统注释，支撑物种基因组图谱构建与功能基因挖掘；

单细胞与空间组学能够解析细胞异质性、表达动态与组织空间分布，为发育生物学、肿瘤研究提供关键数据；系统发生与分子进化可解析物种演化历程、遗传变异规律与适应性特征，揭示生命起源与物种分化机制；

健康与疾病检测则通过多维度分析模型，开展个体健康评估、疾病标志物筛选与治疗进展监测，支撑精准医疗与临床转化研究。

计算分析 查看所有计算分析资源 >

序列比对(BLAST)

基于BLAST算法比对核酸或蛋白序列，鉴定序列间的相似性和同源性，帮助分析序列功能和进化关系

云平台(BiT)

基于云的生物信息学工具包，集成常用生物信息算法、软件与图形绘制方法，提供免费的数据分析云服务

基因组组装与注释

将测序数据组装成连续序列及完整基因组，并对基因组中基因的结构与功能等特征进行注释和分析

单细胞与空间组学

单细胞分类、细胞的组织空间分布及相互作用分析，实现高分辨率的细胞异质性和空间关系表征

系统发生与分子进化

通过比较核酸、蛋白质等分子信息，分析物种演化与遗传变异，揭示进化历程、遗传多样性与适应性

健康与疾病检测

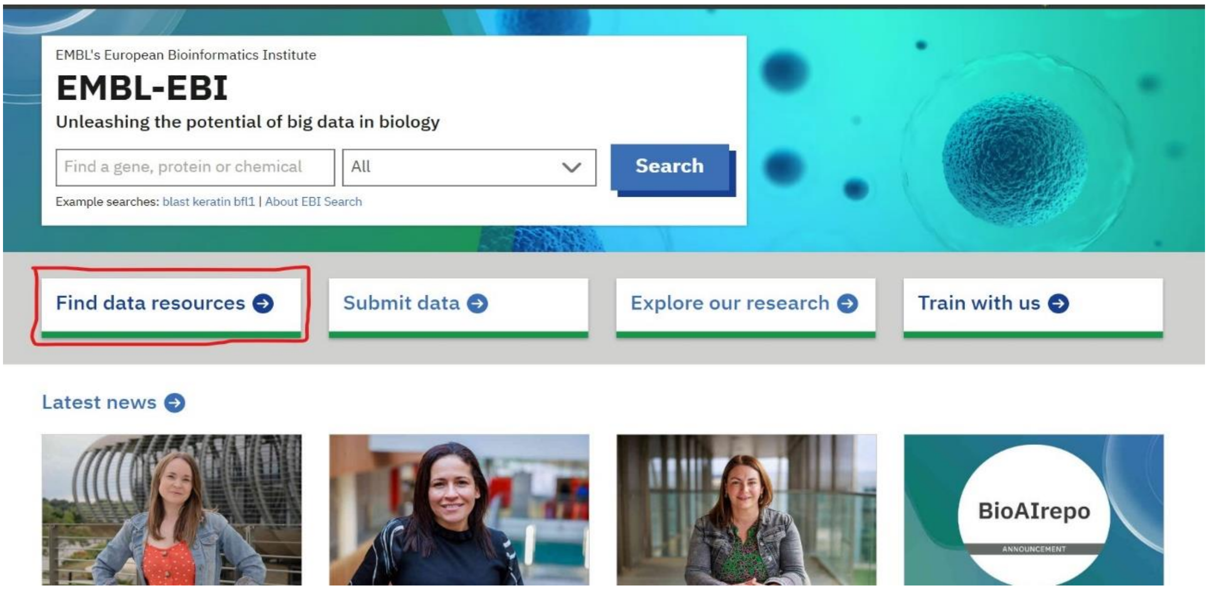
多种定量定性检测分析，以评估和检测个体健康状况、发现疾病严重程度、监测治疗进展等情况

除核心模块外，CNCB 还提供专项服务与安全保障，通过 HGRIP 实现人类遗传资源的规范备份与安全管理，维护国家生物安全；依托 RCoV19 通道，为新冠等突发公共卫生事件提供数据快速递交与变异监测支撑，助力药物与疫苗研发；同时配套汇交证明出具、数据可追溯管理、文献与教育资源等服务，助力生物信息学人才培养。

A3 EBI 基本功能的阐述和讨论

（1）数据资源和工具

在主页面我们可以通过“Find data resources”，找寻数据资源和工具。

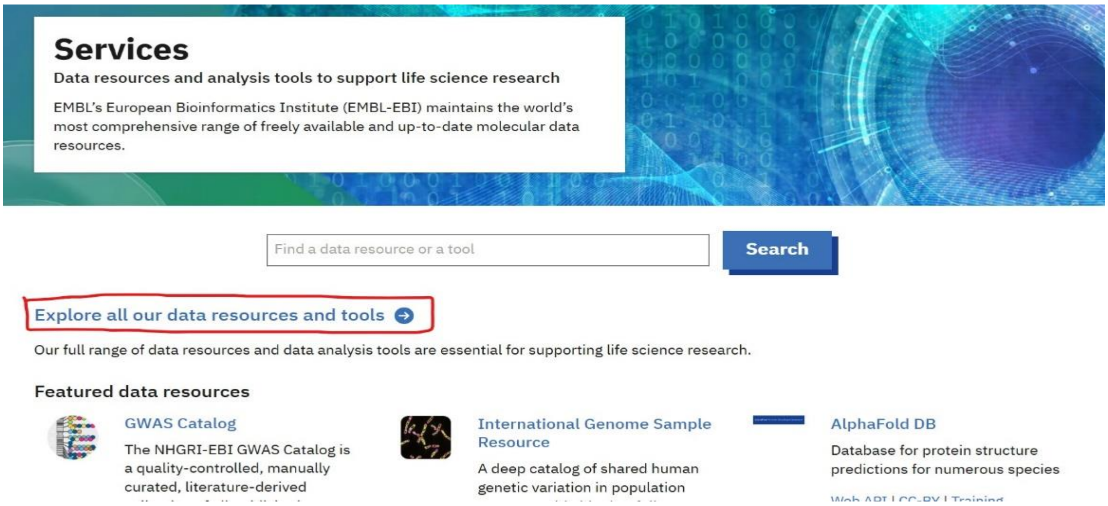


EMBL's European Bioinformatics Institute

EMBL-EBI

Unleashing the potential of big data in biology

主界面有很多常用和特色数据资源，同时也可通过“”explore all our data resources and tools”来探索 EBI 网站所有资源和工具。以常用的 blast 为例，在 EBI 中也可以找到



Services

Data resources and analysis tools to support life science research

EMBL's European Bioinformatics Institute (EMBL-EBI) maintains the world's most comprehensive range of freely available and up-to-date molecular data resources.

Our full range of data resources and data analysis tools are essential for supporting life science research.

Featured data resources

GWAS Catalog

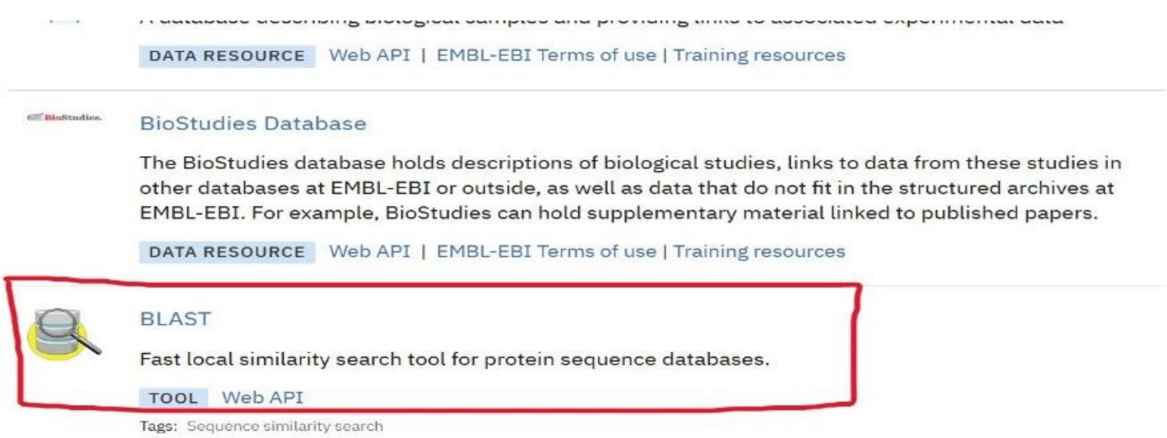
The NHGRI-EBI GWAS Catalog is a quality-controlled, manually curated, literature-derived

International Genome Sample Resource

A deep catalog of shared human genetic variation in population

AlphaFold DB

Database for protein structure predictions for numerous species



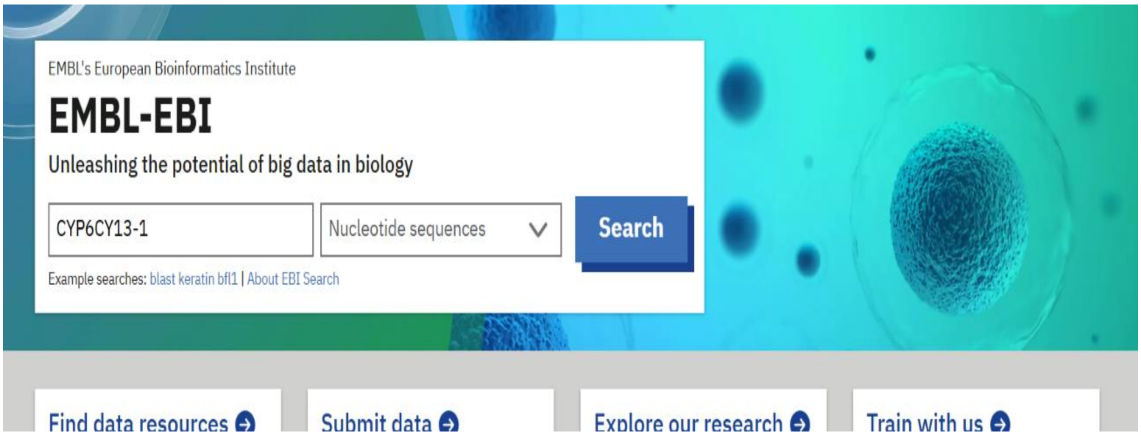
The BioStudies database holds descriptions of biological studies, links to data from these studies in other databases at EMBL-EBI or outside, as well as data that do not fit in the structured archives at EMBL-EBI. For example, BioStudies can hold supplementary material linked to published papers.

Fast local similarity search tool for protein sequence databases.

Tags: Sequence similarity search

（2）基因或蛋白质的查询

以细胞色素 P450 家族的 *CYP6CYP13-1* 为例，在搜索栏中直接查询，并选择查询结果，点击目的基因所在编号,可以查询其序列和所在家族等信息。



Source

All results (3)

Protein sequences (2)
Literature (1)

Source: Europe PMC (ID: PPR801883)

Overexpression of multiple cytochrome P450s contributed to resistance to su
different geographical populations of Aphis gossypii

Wang W, Zhang R, Liu H, Ding R, Lv M, Liang G, Yao J
(2024 Feb 06)

Protein sequences (2 results)

Source: UniRef90 (ID: UniRef90_A0A6G0TZ31)

Cluster: Cytochrome P450

Cross references: Protein sequences (27) Samples & ontologies (3)

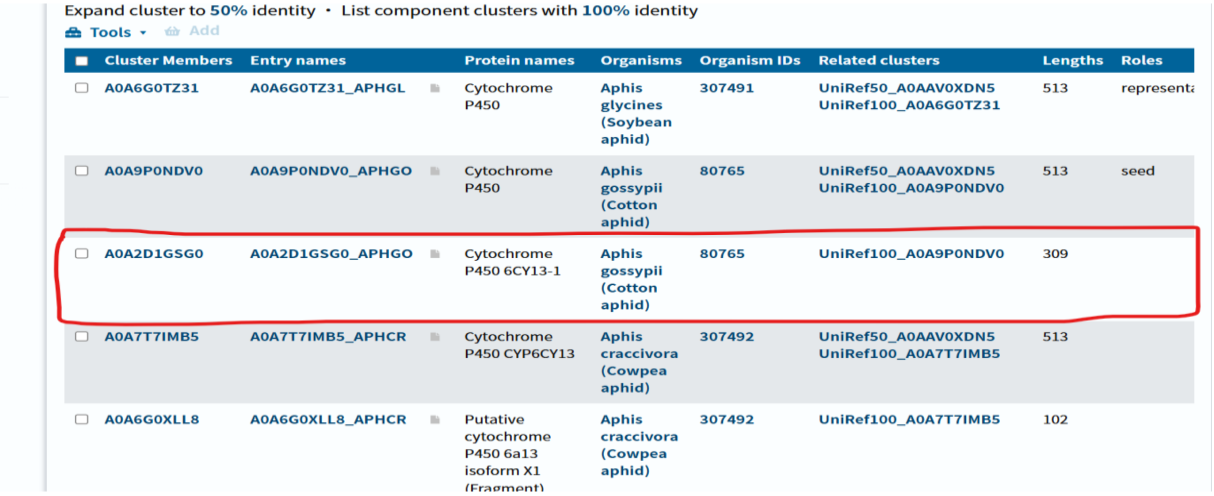
Formats: in FASTA format in UniRef format

Source: UniRef50 (ID: UniRef50_A0AAV0XDN5)

Cluster: Cytochrome P450

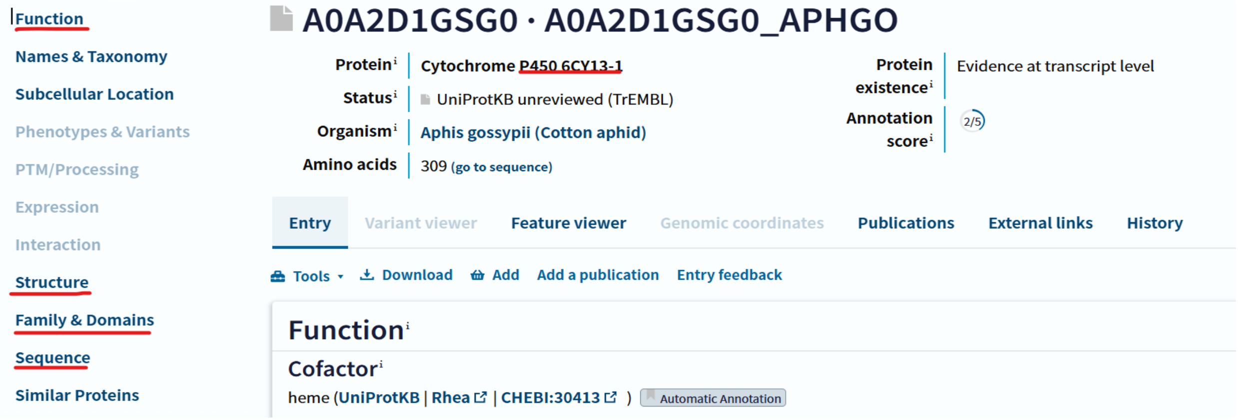
Cross references: Protein sequences (379) Samples & ontologies (16)

Formats: in FASTA format in UniRef format



Expand cluster to 50% identity • List component clusters with 100% identity

☐	Cluster Members	Entry names	Protein names	Organisms	Organism IDs	Related clusters	Lengths	Roles
☐	A0A6G0TZ31	A0A6G0TZ31_APHGL	Cytochrome P450	Aphis glycines (Soybean aphid)	307491	UniRef50_A0AAV0XDN5 UniRef100_A0A6G0TZ31	513	represent
☐	A0A9P0NDV0	A0A9P0NDV0_APHGO	Cytochrome P450	Aphis gossypii (Cotton aphid)	80765	UniRef50_A0AAV0XDN5 UniRef100_A0A9P0NDV0	513	seed
☐	A0A2D1GSG0	A0A2D1GSG0_APHGO	Cytochrome P450 6CY13-1	Aphis gossypii (Cotton aphid)	80765	UniRef100_A0A9P0NDV0	309	
☐	A0A7T7IMB5	A0A7T7IMB5_APHCR	Cytochrome P450 CYP6CY13	Aphis craccivora (Cowpea aphid)	307492	UniRef50_A0AAV0XDN5 UniRef100_A0A7T7IMB5	513	
☐	A0A6G0XLL8	A0A6G0XLL8_APHCR	Putative cytochrome P450 6a13 isoform X1 (Fragment)	Aphis craccivora (Cowpea aphid)	307492	UniRef100_A0A7T7IMB5	102	



Function
Names & Taxonomy
Subcellular Location
Phenotypes & Variants
PTM/Processing
Expression
Interaction
Structure
Family & Domains
Sequence
Similar Proteins

A0A2D1GSG0 · A0A2D1GSG0_APHGO

Protein | Cytochrome P450 6CY13-1
Status | UniProtKB unreviewed (TrEMBL)
Organism | Aphis gossypii (Cotton aphid)
Amino acids | 309 (go to sequence)
Protein existence | Evidence at transcript level
Annotation score | 2/5

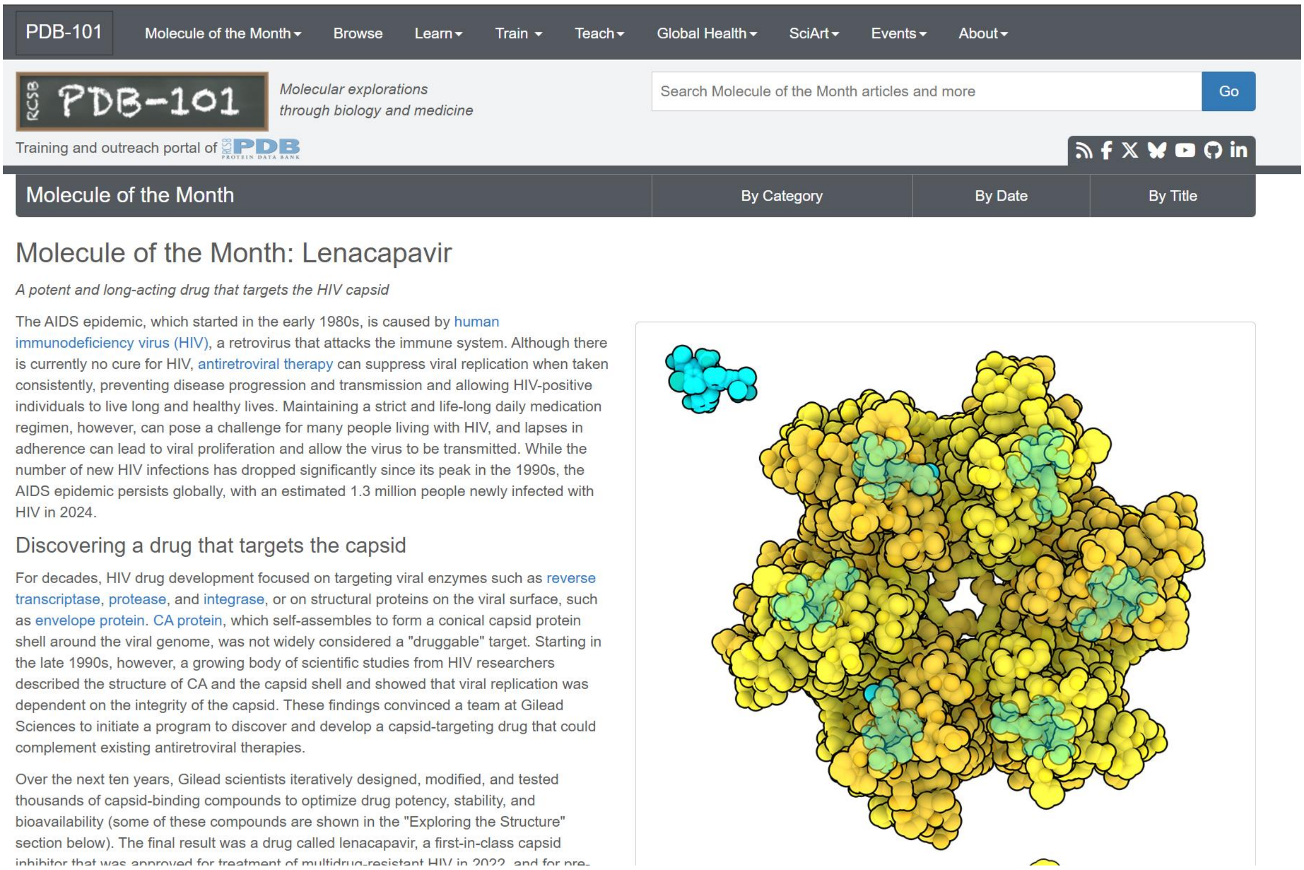
Function

Cofactor

heme (UniProtKB | Rhea | CHEBI:30413) Automatic Annotation

B.1 生物大分子月报（Molecular of the Month）

PDB-101 是何含义？



Molecule of the Month: Lenacapavir

A potent and long-acting drug that targets the HIV capsid

The AIDS epidemic, which started in the early 1980s, is caused by human immunodeficiency virus (HIV), a retrovirus that attacks the immune system. Although there is currently no cure for HIV, antiretroviral therapy can suppress viral replication when taken consistently, preventing disease progression and transmission and allowing HIV-positive individuals to live long and healthy lives. Maintaining a strict and life-long daily medication regimen, however, can pose a challenge for many people living with HIV, and lapses in adherence can lead to viral proliferation and allow the virus to be transmitted. While the number of new HIV infections has dropped significantly since its peak in the 1990s, the AIDS epidemic persists globally, with an estimated 1.3 million people newly infected with HIV in 2024.

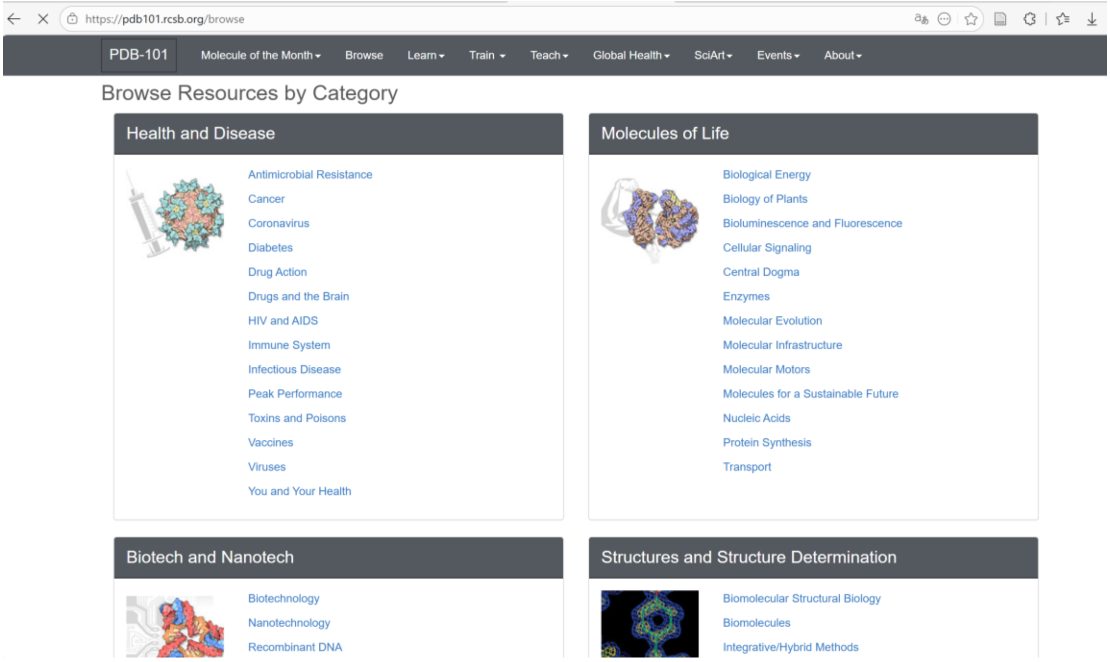
Discovering a drug that targets the capsid

For decades, HIV drug development focused on targeting viral enzymes such as reverse transcriptase, protease, and integrase, or on structural proteins on the viral surface, such as envelope protein. CA protein, which self-assembles to form a conical capsid protein shell around the viral genome, was not widely considered a "druggable" target. Starting in the late 1990s, however, a growing body of scientific studies from HIV researchers described the structure of CA and the capsid shell and showed that viral replication was dependent on the integrity of the capsid. These findings convinced a team at Gilead Sciences to initiate a program to discover and develop a capsid-targeting drug that could complement existing antiretroviral therapies.

Over the next ten years, Gilead scientists iteratively designed, modified, and tested thousands of capsid-binding compounds to optimize drug potency, stability, and bioavailability (some of these compounds are shown in the "Exploring the Structure" section below). The final result was a drug called lenacapavir, a first-in-class capsid inhibitor that was approved for treatment of multidrug-resistant HIV in 2022, and for pre-

PDB-101 是 RCSB PDB（Research Collaboratory for Structural Bioinformatics Protein Data Bank，结构生物信息学研究协作组蛋白质数据库）的培训与科普门户。而 Molecule of the Month 是 PDB-101 教育平台的旗舰栏目，自 2000 年由 David S. Goodsell 创立以来，已发表了 300 多篇文章，旨在通过通俗易懂的方式介绍 PDB 中的重要生物分子结构及其功能。标语 Molecular explorations through biology and medicine 意思是通过生物学与医学探索分子世界。

按功能分类，该网站的生物大分子分为哪几大类？



Browse Resources by Category

Health and Disease

- Antimicrobial Resistance
- Cancer
- Coronavirus
- Diabetes
- Drug Action
- Drugs and the Brain
- HIV and AIDS
- Immune System
- Infectious Disease
- Peak Performance
- Toxins and Poisons
- Vaccines
- Viruses
- You and Your Health

Molecules of Life

- Biological Energy
- Biology of Plants
- Bioluminescence and Fluorescence
- Cellular Signaling
- Central Dogma
- Enzymes
- Molecular Evolution
- Molecular Infrastructure
- Molecular Motors
- Molecules for a Sustainable Future
- Nucleic Acids
- Protein Synthesis
- Transport

Biotech and Nanotech

- Biotechnology
- Nanotechnology
- Recombinant DNA

Structures and Structure Determination

- Biomolecular Structural Biology
- Biomolecules
- Integrative/Hybrid Methods

共分为四类，分别为：

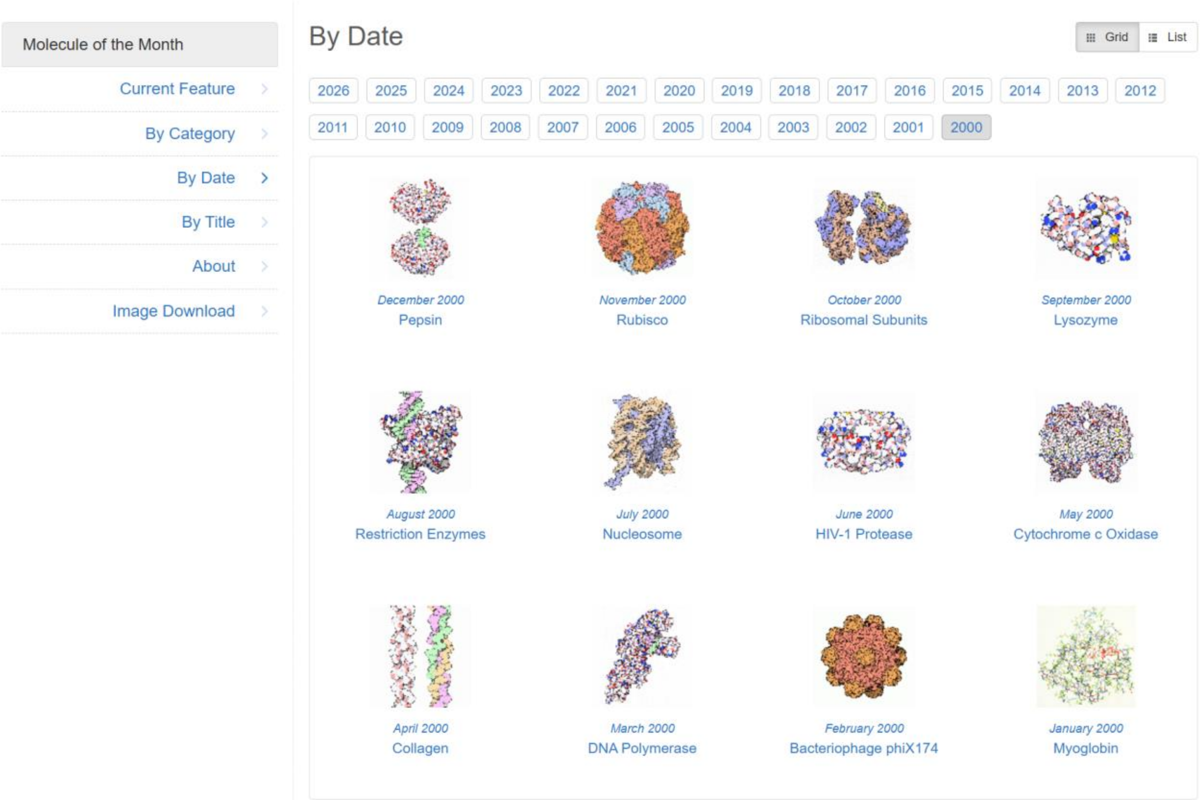
健康与疾病 (Health and Disease)

生命分子 (Molecules of Life)

生物技术与纳米技术 (Biotech and Nanotech)

结构与结构测定 (Structures and Structure Determination)

2000 年 12 篇月报中描述的蛋白质分子，你熟悉的有哪些？



By Date

December 2000 Pepsin	*November 2000* Rubisco	*October 2000* Ribosomal Subunits	*September 2000* Lysozyme
August 2000 Restriction Enzymes	*July 2000* Nucleosome	*June 2000* HIV-1 Protease	*May 2000* Cytochrome c Oxidase
April 2000 Collagen	*March 2000* DNA Polymerase	*February 2000* Bacteriophage phiX174	*January 2000* Myoglobin

Myoglobin（肌红蛋白）、Bacteriophage phiX174（噬菌体 phiX174）、DNA Polymerase（DNA 聚合

酶）、Collagen（胶原蛋白）、Cytochrome c Oxidase（细胞色素 c 氧化酶）、HIV-1 Protease（HIV-1 蛋白酶）、Nucleosome（核小体）、Restriction Enzymes（限制性内切酶）、Lysozyme（溶菌酶）、Ribosomal Subunits（核糖体亚基）、Rubisco（核酮糖 - 1,5 - 二磷酸羧化酶 / 加氧酶）、Pepsin（胃蛋白酶）

其中较为熟悉的有以下蛋白质分子：

Myoglobin（肌红蛋白）：肌肉细胞中的携氧蛋白，能储存并释放氧气，为肌肉运动持续氧供，是肌肉呈现红色的主要原因。

DNA Polymerase（DNA 聚合酶）：催化 DNA 复制的核心酶，能以 DNA 链为模板，将脱氧核苷酸精准拼接成新的 DNA 链，是细胞分裂和遗传信息传递的关键分子。

Collagen（胶原蛋白）：动物体内最丰富的结构蛋白，构成皮肤、骨骼、肌腱等结缔组织的纤维骨架，赋予组织强度和弹性，维持机体结构完整性。

Cytochrome c Oxidase（细胞色素 c 氧化酶）：是一种存在于细菌或线粒体上的大型跨膜蛋白复合物。由于细胞色素氧化酶是呼吸电子传递链的第四个中心酶复合物，因此又被称为复合物 IV。线粒体呼吸链的末端酶，负责将电子传递给氧分子并生成水，同时驱动 ATP 合成，是细胞能量代谢的核心组件。

Nucleosome（核小体）：真核生物染色质的基本结构单位，由 DNA 缠绕组蛋白八聚体形成，能压缩 DNA 长度并调控基因表达，是表观遗传调控的核心结构基础。

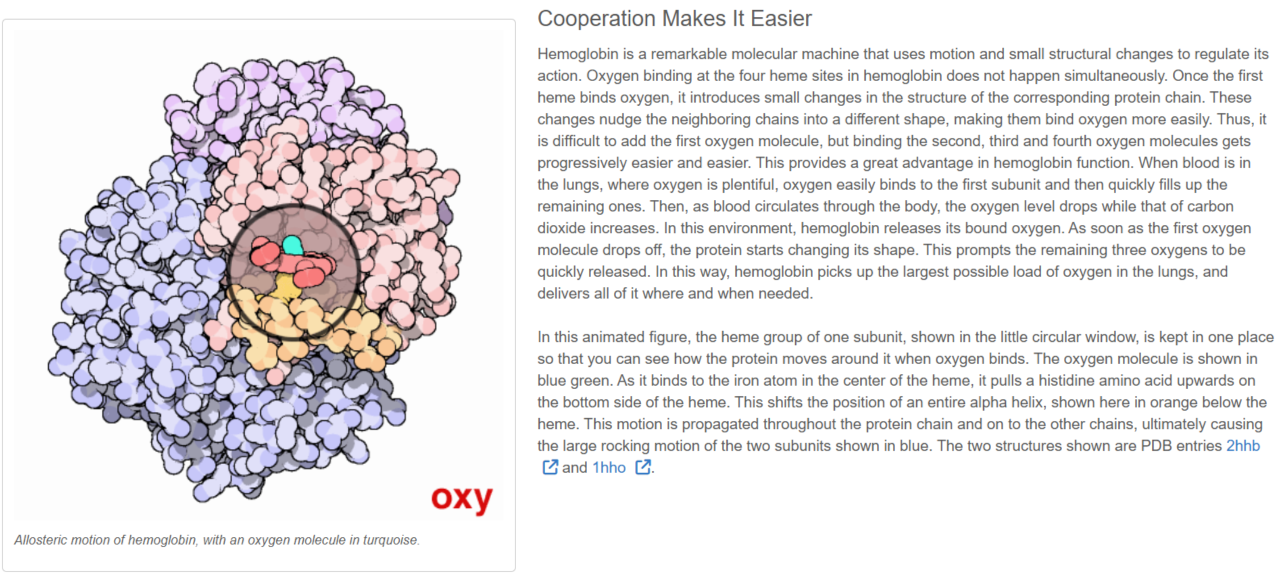
Restriction Enzymes（限制性内切酶）：细菌产生的一类能识别并切割特定 DNA 序列的酶，可保护细菌免受噬菌体感染，也是现代分子克隆和基因编辑技术的核心工具。

Lysozyme（溶菌酶）：广泛存在于黏液、眼泪等体液中的抗菌酶，能破坏细菌细胞壁的肽聚糖结构，裂解细菌，是先天免疫系统的重要防御分子。

Ribosomal Subunits（核糖体亚基）：核糖体的两个功能组件（大亚基与小亚基），协同完成 mRNA 翻译为蛋白质的过程。

Rubisco（核酮糖-1,5-二磷酸羧化酶）：植物叶绿体中含量最丰富的酶，催化光合作用暗反应中 CO_2 的固定，是将无机碳转化为有机碳的核心酶。

阅读该网站关于血红蛋白的短文，说明其结合和释放氧气时不同亚基之间如何协同作用。



Allosteric motion of hemoglobin, with an oxygen molecule in turquoise.

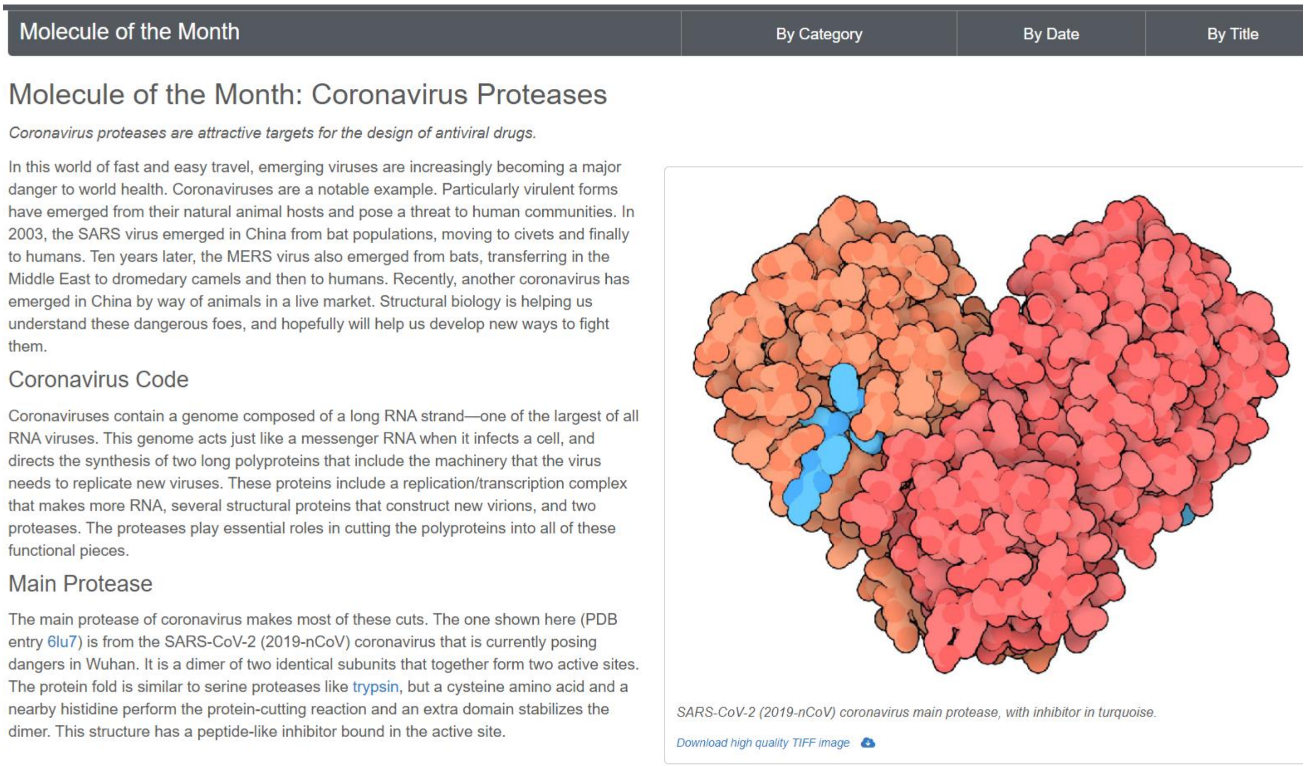
Cooperation Makes It Easier

Hemoglobin is a remarkable molecular machine that uses motion and small structural changes to regulate its action. Oxygen binding at the four heme sites in hemoglobin does not happen simultaneously. Once the first heme binds oxygen, it introduces small changes in the structure of the corresponding protein chain. These changes nudge the neighboring chains into a different shape, making them bind oxygen more easily. Thus, it is difficult to add the first oxygen molecule, but binding the second, third and fourth oxygen molecules gets progressively easier and easier. This provides a great advantage in hemoglobin function. When blood is in the lungs, where oxygen is plentiful, oxygen easily binds to the first subunit and then quickly fills up the remaining ones. Then, as blood circulates through the body, the oxygen level drops while that of carbon dioxide increases. In this environment, hemoglobin releases its bound oxygen. As soon as the first oxygen molecule drops off, the protein starts changing its shape. This prompts the remaining three oxygens to be quickly released. In this way, hemoglobin picks up the largest possible load of oxygen in the lungs, and delivers all of it where and when needed.

In this animated figure, the heme group of one subunit, shown in the little circular window, is kept in one place so that you can see how the protein moves around it when oxygen binds. The oxygen molecule is shown in blue green. As it binds to the iron atom in the center of the heme, it pulls a histidine amino acid upwards on the bottom side of the heme. This shifts the position of an entire alpha helix, shown here in orange below the heme. This motion is propagated throughout the protein chain and on to the other chains, ultimately causing the large rocking motion of the two subunits shown in blue. The two structures shown are PDB entries 2hhb and 1hho.

血红蛋白是一台精巧的分子机器，依靠分子运动和局部结构的微小改变来实现对氧气的高效运输。它的四个亚基并不同时结合氧气，而是表现出典型的协同作用：第一个氧分子与血红素结合最为困难，结合后会引起对应蛋白链发生细微构象变化，这种变化会传递到相邻亚基，使它们结合氧气的能力不断增强，因此第二个、第三个、第四个氧分子的结合会越来越容易。当血液流经肺部时，在高氧环境下，血红蛋白能够迅速结合并满载氧气；而当血液进入氧浓度低、二氧化碳浓度高的组织时，血红蛋白会释放氧气。一旦第一个氧分子脱离，蛋白质构象便开始转变，促使其余氧分子快速释放，从而将氧气精准地输送到需要的部位。在分子层面上，氧与血红素中心的铁原子结合时，会拉动下方的组氨酸残基，使整条 α 螺旋发生位移，这一局部运动通过肽链传递到其他亚基，引发整个四聚体的构象摆动，实现结构状态的转换，这也是血红蛋白协同作用的结构基础。

阅读你最感兴趣的生物大分子，阐述该分子的结构、功能特点。



Molecule of the Month: Coronavirus Proteases

Coronavirus proteases are attractive targets for the design of antiviral drugs.

In this world of fast and easy travel, emerging viruses are increasingly becoming a major danger to world health. Coronaviruses are a notable example. Particularly virulent forms have emerged from their natural animal hosts and pose a threat to human communities. In 2003, the SARS virus emerged in China from bat populations, moving to civets and finally to humans. Ten years later, the MERS virus also emerged from bats, transferring in the Middle East to dromedary camels and then to humans. Recently, another coronavirus has emerged in China by way of animals in a live market. Structural biology is helping us understand these dangerous foes, and hopefully will help us develop new ways to fight them.

Coronavirus Code

Coronaviruses contain a genome composed of a long RNA strand—one of the largest of all RNA viruses. This genome acts just like a messenger RNA when it infects a cell, and directs the synthesis of two long polyproteins that include the machinery that the virus needs to replicate new viruses. These proteins include a replication/transcription complex that makes more RNA, several structural proteins that construct new virions, and two proteases. The proteases play essential roles in cutting the polyproteins into all of these functional pieces.

Main Protease

The main protease of coronavirus makes most of these cuts. The one shown here (PDB entry 6lu7) is from the SARS-CoV-2 (2019-nCoV) coronavirus that is currently posing dangers in Wuhan. It is a dimer of two identical subunits that together form two active sites. The protein fold is similar to serine proteases like trypsin, but a cysteine amino acid and a nearby histidine perform the protein-cutting reaction and an extra domain stabilizes the dimer. This structure has a peptide-like inhibitor bound in the active site.

SARS-CoV-2 (2019-nCoV) coronavirus main protease, with inhibitor in turquoise.

Download high quality TIFF image

结构方面，冠状病毒蛋白酶（以 SARS-CoV-2 主蛋白酶为例）是由两个相同亚基组成的同源二聚体，整体呈现对称的二叶状结构，每个亚基共同形成一个活性位点。蛋白质折叠类似于丝氨酸蛋白酶如胰蛋白酶，但半胱氨酸氨基酸和附近的组氨酸执行蛋白质切割反应，额外的结构域稳定了二聚体。该结构活性位点结合了类似肽的抑制剂。

功能方面，冠状病毒基因组为长链 RNA，感染细胞后翻译出两条多聚蛋白，蛋白酶的核心功能是将这些多聚蛋白切割为具有独立功能的成熟蛋白片段，包括复制/转录复合物、结构蛋白等，是病毒复制的必需步骤；由于蛋白酶是病毒成熟与复制的关键环节，且人类细胞中无同源功能的半胱氨酸蛋白酶，因此成为抗病毒药物设计的理想靶点，即通过小分子抑制剂占据活性位点，可阻断蛋白酶切割功能，从而抑制病毒增殖。

B.2 化学小分子月报（Molecular of the Month）

与生物大分子月报相比，该网站的文章内容和写作风格有何特点？

Welcome to the Molecule of the Month page!

Each month since January 1996 a new molecule has been added to the list on this page, which makes this one of the longest running Chemical websites on the internet! The links will take you to a page at one of the Web sites at a University Chemistry Department or commercial site in the UK, the US, or anywhere in the world, where useful (and hopefully entertaining!), information can be found about a particularly interesting molecule.

Each month's new molecule will be announced on Twitter [@MoleculeM] and via an RSS newsfeed [http://www.chm.bris.ac.uk/motm/rss.xml].

We are always happy to receive *Molecule of the Month* articles from anyone who wishes to submit one. If you wish to contribute a MOTM article, there's a document containing all the information you need online.

What do I need to see the pages properly?

See here for all the details of software and helper files you require. In brief, next to each molecule name (scroll down the page) there will be one or more of the following buttons. Clicking each button takes you to a different version of the page, as described below.

Click for HTML	A simple HTML-only page with text and images.
Click for pdf	A page in pdf format. (Requires a *Adobe Acrobat* viewer).
Click for VRML	A page with optional VRML structure files. (Requires a VRML viewer such as *Cortona*).
Click for Chime	A page with embedded *CHIME* structure files. (Requires the *CHIME* plug-in to be installed on your browser).
Click for Jmol	A page with embedded java structure files. (Requires Java to be enabled on your browser).
Click for JSmol	A page with embedded javascript structure files. (Requires a browser that supports HTML5).

See here for a list of the awards and citations received by this page and a disclaimer.

Select your molecule

Above is a drop-down alphabetical list of all the molecules. Click on your molecule of choice to take you directly to it in the chronological table below.

Buy a book containing many of these molecules!

Paul May and Simon Cotton have compiled a set of some of the most important and interesti
this website, as well as many more that have never been published before, and published th
you can buy from the Taylor and Francis website or from other online retaile

Date	Molecule	Versions	Contributor	Location
March 2026	Peroxyacetyl nitrate (PAN) The eye-irritant in smog	Click for HTML Click for JSmol	Simon Cotton	University of Birmingham
February 2026	Cyclo-octasulfur, S_8 Known as 'brimstone' in ancient times	Click for HTML Click for JSmol	Simon Zhang and Stephen Belding	Rugby School, UK
January 2026	Azelaic acid The skin-whitening agent that can be used to make Nylon	Click for HTML Click for JSmol	Paul May	Bristol University
December 2025	Methyl salicylate The plant-signalling molecule and minty painkiller found in toothpaste and mouthwash	Click for HTML Click for JSmol	Simon Cotton	University of Birmingham
November 2025	*L*-4-Boronophenylalanine An unnatural amino acid used in boron neutron capture therapy for combating cancer	Click for HTML Click for JSmol	Dane Jemc	University of Ljubljana, Slovenia

化学小分子月报是纯化学视角的科普，文章以小分子的化学本质为核心，内容围绕分子的结构、合成、理化性质、工业和生活应用与历史背景展开。写作风格上更偏向化学启蒙与生活化叙事，语言通俗直白，常结合日常场景解释分子的实际意义，同时会提供多种结构可视化版本，让读者直观理解分子的空间构型，，侧重让大众感受化学与生活的紧密联系。

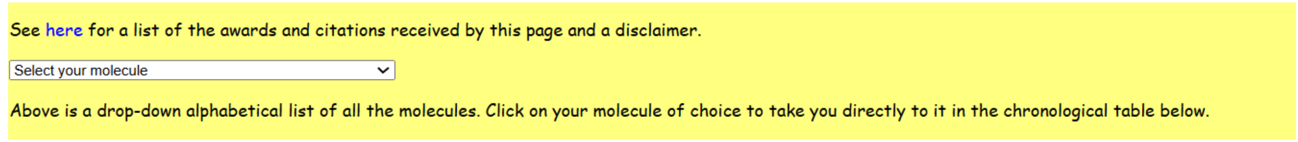
相比之下，**生物大分子月报**则是结构生物学与生物医学视角，核心是生物大分子（蛋白质、核酸等）或与生物靶点结合的药物小分子，内容聚焦分子在生命活动中的功能、三维结构与生理和临床意义，例如介绍血红蛋白时会阐释其协同载氧的构象变化。写作风格更偏向基础生命科学与药物研发叙事，会深入讲解结构与功能的关联、分子在细胞通路或疾病中的角色，语言更具学术科普感，常结合 PDB 结构数据与医学应用，侧重让读者理解“结构决定功能”的生命科学核心逻辑。

如何查看某个特定分子？

下拉菜单快速检索（最推荐）：

在页面中部找到“Select your molecule”下拉框，点击后会显示按字母顺序排列的所有分子列表。

直接在列表中找到目标分子名称，点击即可跳转到该分子的专属介绍页面。



See here for a list of the awards and citations received by this page and a disclaimer.

Select your molecule

Above is a drop-down alphabetical list of all the molecules. Click on your molecule of choice to take you directly to it in the chronological table below.

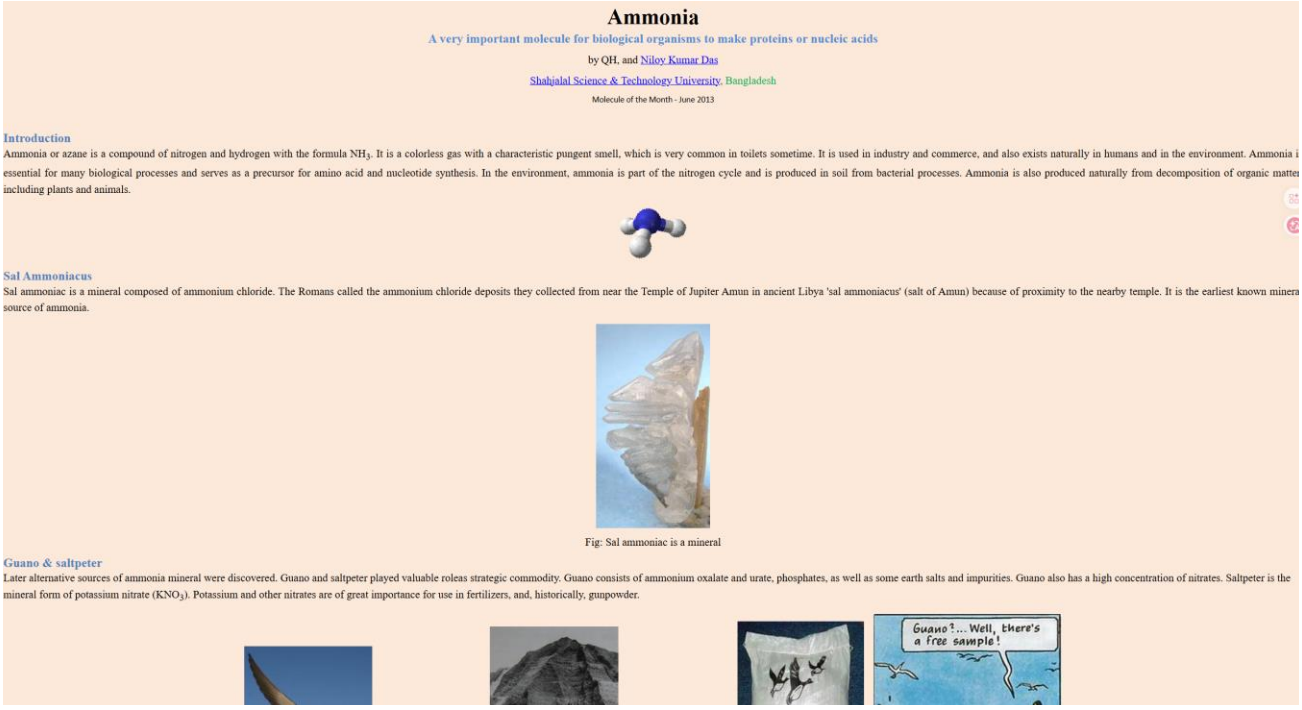
按时间顺序浏览查找：

页面下方的表格按发布时间倒序排列了所有月度分子，包含：

发布日期（Date）、分子名称（Molecule）、页面版本（Versions）、贡献者（Contributor）与机构（Location）

Date	Molecule	Versions	Contributor	Location
March 2026	**Peroxyacetyl nitrate (PAN)** The eye-irritant in smog	Click for HTML Click for JSmol	Simon Cotton	University of Birmingham
February 2026	**Cyclo-octasulfur, S_8** Known as 'brimstone' in ancient times	Click for HTML Click for JSmol	Simon Zhang and Stephen Belding	Rugby School, UK
January 2026	**Azelaic acid** The skin-whitening agent that can be used to make Nylon	Click for HTML Click for JSmol	Paul May	Bristol University
December 2025	**Methyl salicylate** The plant-signalling molecule and minty painkiller found in toothpaste and mouthwash	Click for HTML Click for JSmol	Simon Cotton	University of Birmingham
November 2025	***L*-4-Boronophenylalanine** An unnatural amino acid used in boron neutron capture therapy for combating cancer	Click for HTML Click for JSmol	Dane Jemc	University of Ljubljana, Slovenia
October 2025	**Psittacofulvins** They make parrots colourful!	Click for HTML Click for JSmol	Paul May	Bristol University
September 2025	**Fensic** **Tris[bis(trimethylsilyl)amido]iron(III) or $[Fe(N(SiMe_3)_2)_3]$** A pioneer of three-coordination	Click for HTML Click for JSmol	Simon Cotton	University of Birmingham
August 2025	**5-Aminolevulinic Acid (5-ALA)** The "pink drink" used in fluorescence-guided surgery to identify cancerous cells.	Click for HTML Click for JSmol	Ben Warburton and Claudia Orellana	Hereford Sixth Form College Hereford, UK
July 2025	**Pyridine** The Pungent Powerhouse of Chemistry	Click for HTML Click for JSmol	Shanqiao Yang and Stephen Belding	Rugby School, UK
June 2025	**Ketene** The molecule linked with vaping-associated lung injury (EVALI)	Click for HTML Click for JSmol	Simon Cotton	University of Birmingham
May 2025	**Suzetrigine** A new non-addictive painkiller	Click for HTML Click for JSmol	Paul May	Bristol University
April 2025	**AZT (Zidovudine)** The first antiretroviral medication approved for treating HIV.	Click for HTML Click for JSmol	Henry Goss-Custard	Eton College, UK

阅读你感兴趣的化学小分子，阐述该分子的结构、功能特点。



Ammonia

A very important molecule for biological organisms to make proteins or nucleic acids

by QH, and Niloy Kumar Das

Shahjalal Science & Technology University, Bangladesh

Molecule of the Month - June 2013

Introduction

Ammonia or azane is a compound of nitrogen and hydrogen with the formula NH_3. It is a colorless gas with a characteristic pungent smell, which is very common in toilets sometime. It is used in industry and commerce, and also exists naturally in humans and in the environment. Ammonia i essential for many biological processes and serves as a precursor for amino acid and nucleotide synthesis. In the environment, ammonia is part of the nitrogen cycle and is produced in soil from bacterial processes. Ammonia is also produced naturally from decomposition of organic matte including plants and animals.

Sal Ammoniacus

Sal ammoniac is a mineral composed of ammonium chloride. The Romans called the ammonium chloride deposits they collected from near the Temple of Jupiter Amun in ancient Libya 'sal ammoniacus' (salt of Amun) because of proximity to the nearby temple. It is the earliest known minera source of ammonia.

Fig: Sal ammoniac is a mineral

Guano & saltpeter

Later alternative sources of ammonia mineral were discovered. Guano and saltpeter played valuable roleas strategic commodity. Guano consists of ammonium oxalate and urate, phosphates, as well as some earth salts and impurities. Guano also has a high concentration of nitrates. Saltpeter is the mineral form of potassium nitrate (KNO_3). Potassium and other nitrates are of great importance for use in fertilizers, and, historically, gunpowder.

氨（NH3）是由一个氮原子与三个氢原子构成的三角锥形分子，其独特的结构决定了多样且关键的功能特性。在功能层面，氨的核心价值体现在生物、工业与能源三大领域。生物中，氨是氨基酸和核苷酸合成的前体，是生命制造蛋白质与核酸的关键氮源，也是自然界氮循环的核心中间产物，支撑着生态系统的氮素流转与全球粮食生产；工业上，氨是最重要的化工原料之一：约 83%的氨被制成氮肥（如尿素、硝酸铵等），养活了全球近一半人口；它还通过奥斯特瓦尔德工艺转化为硝酸，进一步生产炸药、塑料与多种化学品；同时氨凭借高效环保的热力学特性，长期作为工业制冷剂，不会破坏臭氧层或加剧温室效应；能源领域，氨被称为“另一种氢”，是理想的氢载体：低压下即可液化，体积氢含量高于压缩氢或液氢，储存与运输更安全便捷，可作为内燃机燃料或燃料电池的氢源，甚至曾为航空航天火箭发动机提供动力，展现出清洁能源的广阔前景。

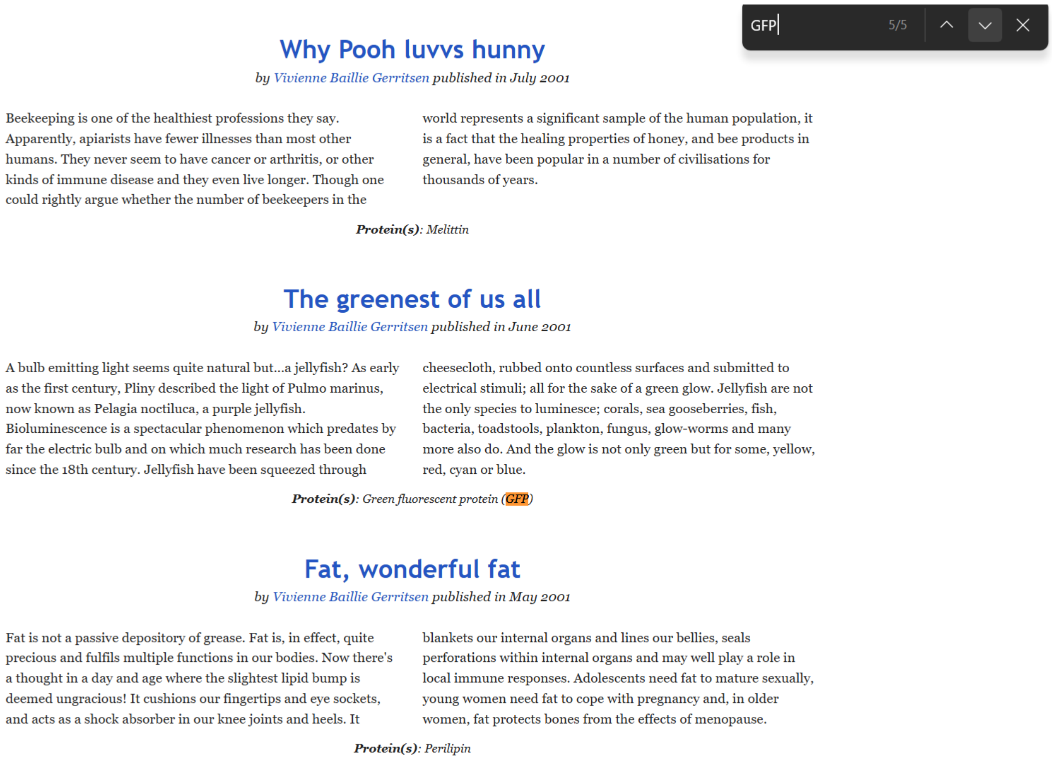
B.3 蛋白质分子精选（Protein Spotlight）

没有外网，如何搜索文献？

点击页面首栏“ARTICLES”（图示蓝框）进入文章库，随后采用 ctrl+F 的方法，也可以根据关键词搜取相关文献。

阅读该网站关于绿色荧光蛋白（GFP）短文（The greenest of us all）以及相关文献，说明该蛋白质分子的结构功能特点。

当直接搜索“GFP”时首先出现的不是目标文章，可点击图中上下箭头找到；或直接搜索文章名称“The greenest of us all”即可找到。



Why Pooh luvvs hunny

by Vivienne Baillie Gerritsen published in July 2001

Beekeeping is one of the healthiest professions they say. Apparently, apiarists have fewer illnesses than most other humans. They never seem to have cancer or arthritis, or other kinds of immune disease and they even live longer. Though one could rightly argue whether the number of beekeepers in the world represents a significant sample of the human population, it is a fact that the healing properties of honey, and bee products in general, have been popular in a number of civilisations for thousands of years.

***Protein(s)**: Melittin*

The greenest of us all

by Vivienne Baillie Gerritsen published in June 2001

A bulb emitting light seems quite natural but...a jellyfish? As early as the first century, Pliny described the light of Pulmo marinus, now known as Pelagia noctiluca, a purple jellyfish. Bioluminescence is a spectacular phenomenon which predates by far the electric bulb and on which much research has been done since the 18th century. Jellyfish have been squeezed through cheesecloth, rubbed onto countless surfaces and submitted to electrical stimuli; all for the sake of a green glow. Jellyfish are not the only species to luminesce; corals, sea gooseberries, fish, bacteria, toadstools, plankton, fungus, glow-worms and many more also do. And the glow is not only green but for some, yellow, red, cyan or blue.

***Protein(s)**: Green fluorescent protein (GFP)*

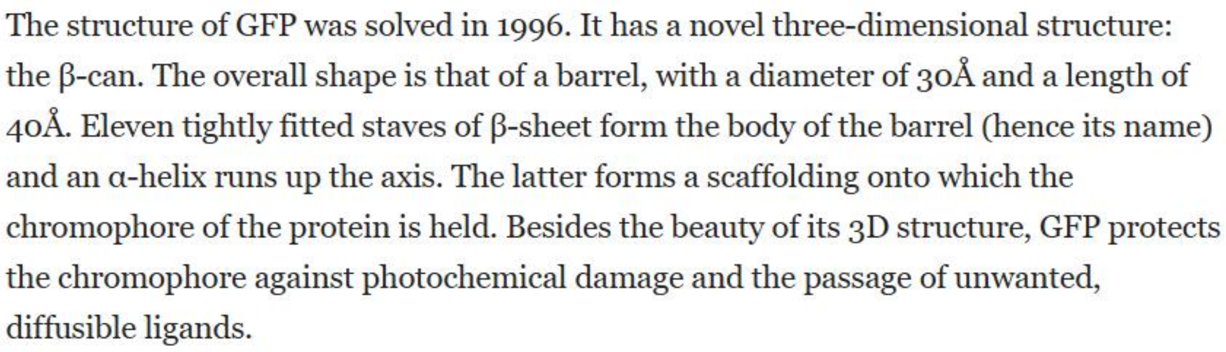
Fat, wonderful fat

by Vivienne Baillie Gerritsen published in May 2001

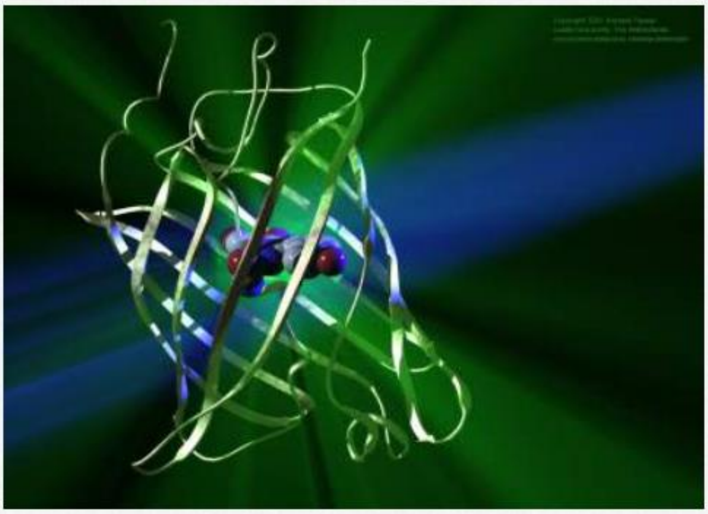
Fat is not a passive depository of grease. Fat is, in effect, quite precious and fulfils multiple functions in our bodies. Now there's a thought in a day and age where the slightest lipid bump is deemed ungracious! It cushions our fingertips and eye sockets, and acts as a shock absorber in our knee joints and heels. It blankets our internal organs and lines our bellies, seals perforations within internal organs and may well play a role in local immune responses. Adolescents need fat to mature sexually, young women need fat to cope with pregnancy and, in older women, fat protects bones from the effects of menopause.

***Protein(s)**: Perilipin*

GFP 的结构于 1996 年被确定。它拥有一种新颖的三维结构：β-桶。整体形状像一个桶，直径 30Å，长度 40Å。桶身由 11 段紧密排列的 β-折叠构成（因此得名），一条 α-螺旋沿中心轴延伸。这条 α-螺旋作为支架，支撑并固定着蛋白质的发色团。除了优美的三维结构之外，GFP 还能保护发色团免受光化学损伤，并阻止无关的可扩散小分子进入其内部。（1Å=10^{-10} 米=0.1 纳米）

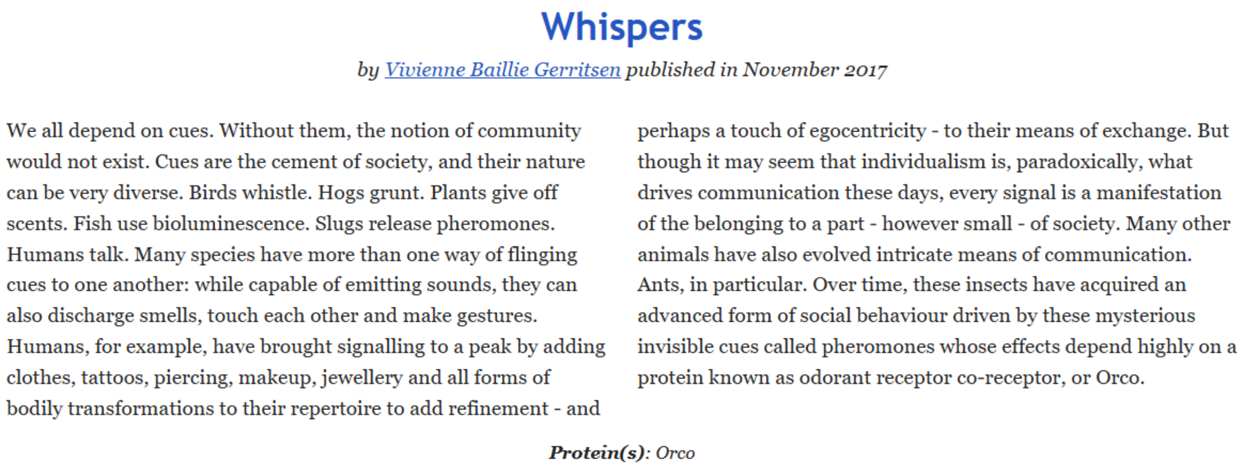


The structure of GFP was solved in 1996. It has a novel three-dimensional structure: the β-can. The overall shape is that of a barrel, with a diameter of 30Å and a length of 40Å. Eleven tightly fitted staves of β-sheet form the body of the barrel (hence its name) and an α-helix runs up the axis. The latter forms a scaffolding onto which the chromophore of the protein is held. Besides the beauty of its 3D structure, GFP protects the chromophore against photochemical damage and the passage of unwanted, diffusible ligands.



Green fluorescent protein
Courtesy of phantatomix.com

该网站中与你研究方向相关的蛋白质分子有哪些？简述它们的生物学功能。



Whispers

by Vivienne Baillie Gerritsen published in November 2017

We all depend on cues. Without them, the notion of community would not exist. Cues are the cement of society, and their nature can be very diverse. Birds whistle. Hogs grunt. Plants give off scents. Fish use bioluminescence. Slugs release pheromones. Humans talk. Many species have more than one way of flinging cues to one another: while capable of emitting sounds, they can also discharge smells, touch each other and make gestures. Humans, for example, have brought signalling to a peak by adding clothes, tattoos, piercing, makeup, jewellery and all forms of bodily transformations to their repertoire to add refinement - and perhaps a touch of egocentricity - to their means of exchange. But though it may seem that individualism is, paradoxically, what drives communication these days, every signal is a manifestation of the belonging to a part - however small - of society. Many other animals have also evolved intricate means of communication. Ants, in particular. Over time, these insects have acquired an advanced form of social behaviour driven by these mysterious invisible cues called pheromones whose effects depend highly on a protein known as odorant receptor co-receptor, or Orco.

Protein(s)*: Orco*

Orco，全称 Odorant receptor co-receptor，气味受体共受体，是昆虫嗅觉系统中不可或缺的核心蛋白。它不直接识别气味分子，但必须与各类特异性气味受体（OR）结合形成复合物，才能使受体正确

折叠、转运并定位于嗅觉神经元的细胞膜上。在功能上，Orco 作为离子通道的重要组成部分，参与气味信号的跨膜传递，当特异性气味受体结合气味配体后，OR-Orco 复合物会开放阳离子通道，引发细胞膜电位变化，进而启动嗅觉信号传导。此外，Orco 还能稳定受体结构、保护功能性复合体，并参与嗅觉适应与信号调控，对昆虫的嗅觉感知、觅食、求偶、定位宿主等关键行为至关重要。

What mosquitoes sniff

by Vivienne Baillie Gerritsen published in March 2004

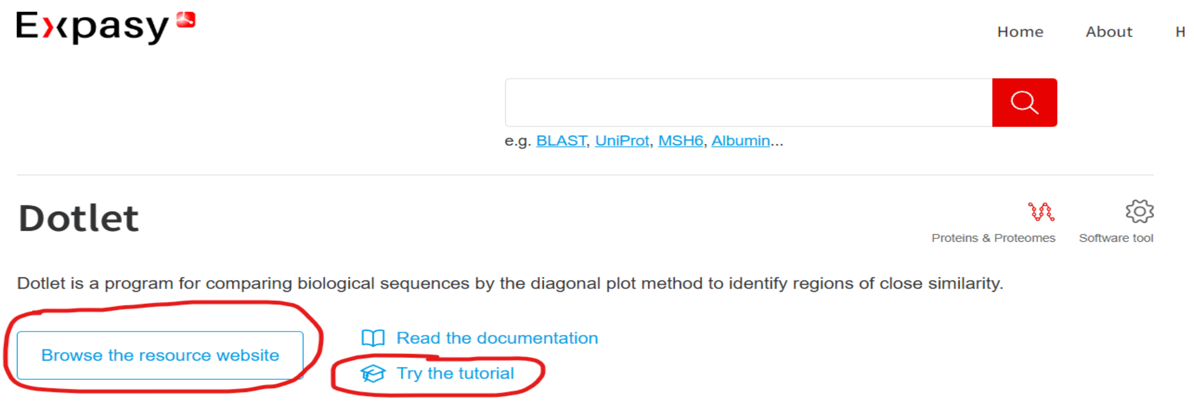
It is hardly the time to talk of mosquitoes when the cold winter winds are still blowing. In milder climates though, mosquitoes are out and about, causing millions of deaths every year through their ability to transmit diseases, such as encephalitis, dengue, yellow fever and, of course, malaria. According to the World Health Organization, malaria alone is the cause of over two million deaths in Africa, one million of which are children under the age of five. The mosquitoes that transmit the disease belong to more than one species and they are collectively known as the anopheline mosquitoes.

***Protein(s)**: Odorant receptor Or1*

Odorant receptor Or1（Or1，气味受体 1），是昆虫（以冈比亚按蚊 *Anopheles gambiae* 为经典研究对象）中一类特异性气味受体，属于昆虫嗅觉受体（OR）家族，是介导按蚊偏好人类宿主、定位吸血目标的关键受体蛋白。

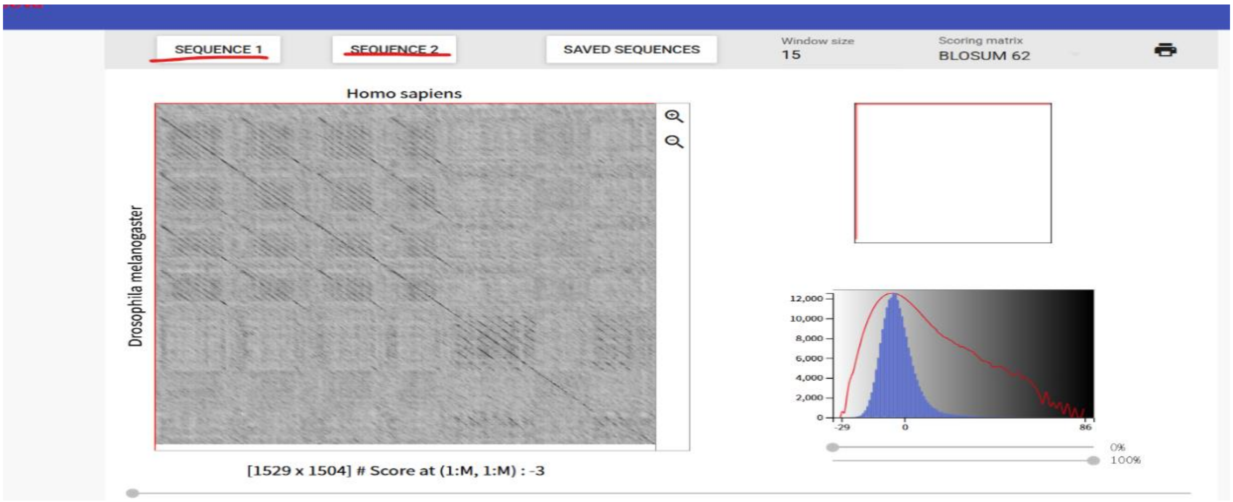
C1 利用 Dotlet 进行序列比对，分析序列的相似性。

1）进入 Dotlet 官网，进入“browse the resource website”（其中 try the tutorial 可以在其中学习如何使用 Dotlet，属于网站的教学指南。）



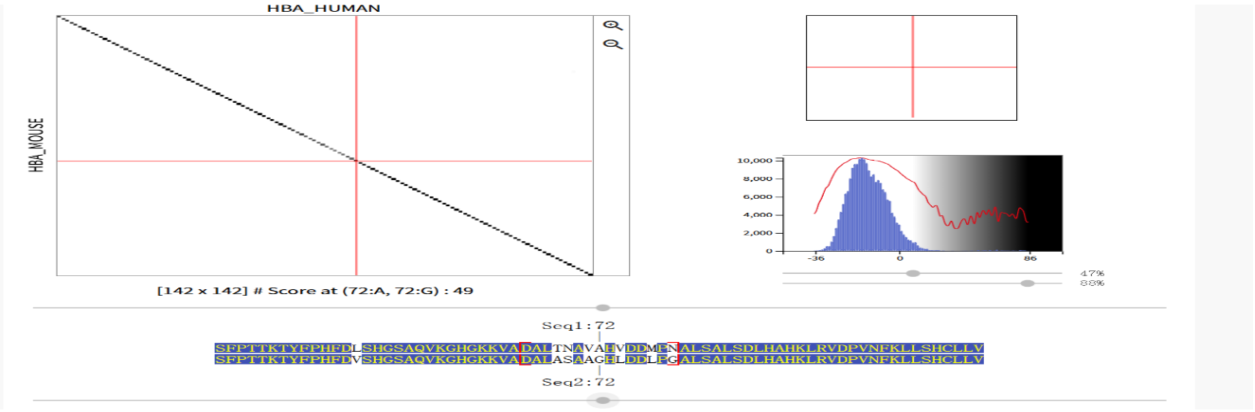
Dotlet is a program for comparing biological sequences by the diagonal plot method to identify regions of close similarity.

2）在 sequence1 和 sequence2 中分别输入 HBA_HUMAN 和 HBA_MOUSE 两个序列。



```
[1529 x 1504] # Score at (1:M, 1:M) : -3
```

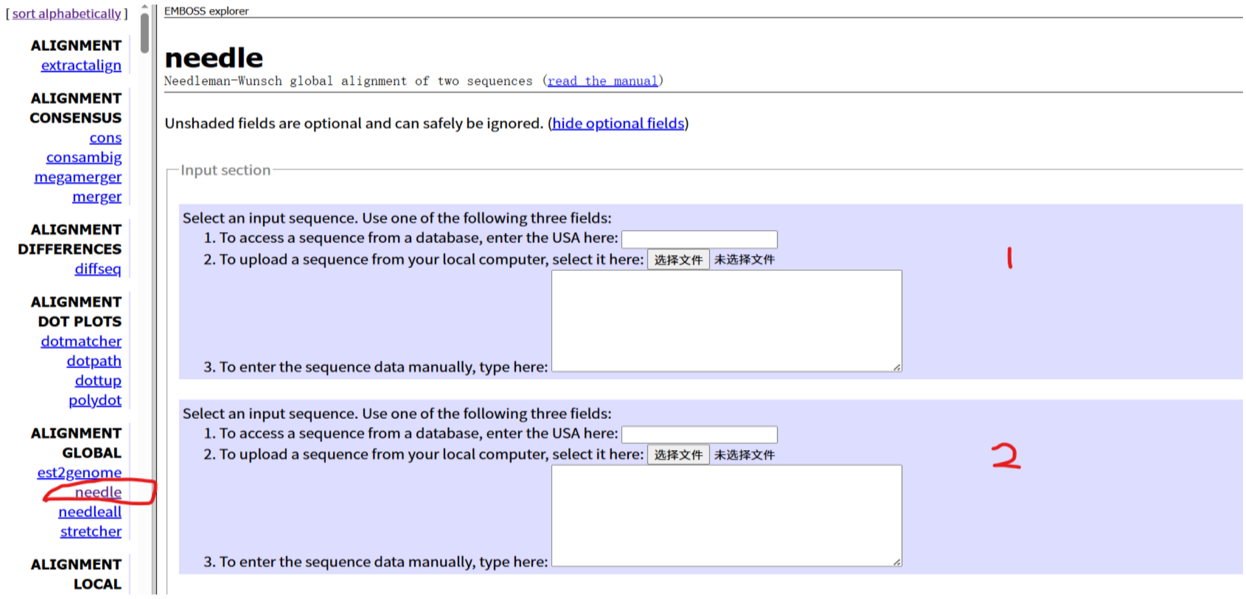
3）之后便可得到比对结果：蓝色的碱基配对区域表示高度匹配的区域。



```
[142 x 142] # Score at (72:A, 72:G) : 49

Seq1:72
SFPTTKTYFPHFDLSHGSAQVKGHGKKVADALTNAVAHVDDMPNALSALSDLHAHKLRVDPVNFKLLSHCLLV
SFPTTKTYFPHFDVSHGSAQVKGHGKKVADALASAAGHLDDLPGALSALSDLHAHKLRVDPVNFKLLSHCLLV
Seq2:72
```

C2 利用 EMBOSS Explorer 进行序列比对

1) 打开 EMBOSS Explorer 分析平台：https://www.bioinformatics.nl/emboss-explorer，在程序导航菜单中找到整体比对 ALIGNMENT GLOBAL，点击程序名 needle。



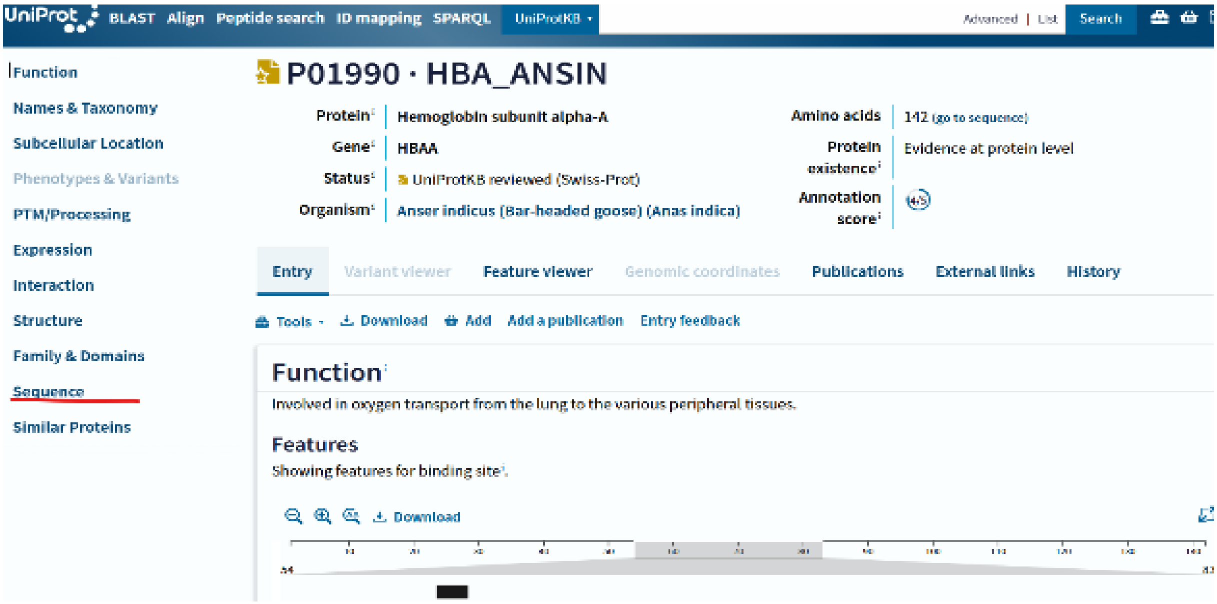
Needleman-Wunsch global alignment of two sequences

Unshaded fields are optional and can safely be ignored.

Select an input sequence. Use one of the following three fields:
1. To access a sequence from a database, enter the USA here:
2. To upload a sequence from your local computer, select it here:
3. To enter the sequence data manually, type here:

Select an input sequence. Use one of the following three fields:
1. To access a sequence from a database, enter the USA here:
2. To upload a sequence from your local computer, select it here:
3. To enter the sequence data manually, type here:

2) 在 UniProt 数据库检索框中输入斑头雁 alpha 血红蛋白序列条目名 HBA_ANSIN，选择 sequence，点击 download，便可得到其蛋白序列，同理可以得到灰雁 alpha 血红蛋白序列条目名为 HBA_ANSAN 的序列。



P01990 · HBA_ANSIN

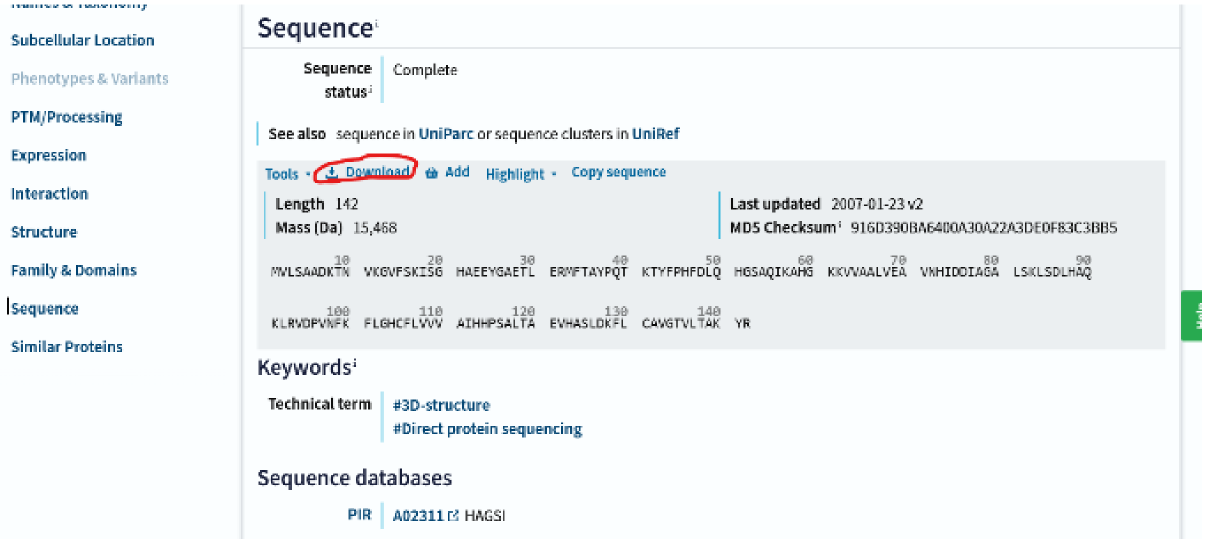
Protein: Hemoglobin subunit alpha-A
Gene: HBAA
Status: UniProtKB reviewed (Swiss-Prot)
Organism: Anser indicus (Bar-headed goose) (Anas indica)
Amino acids: 142 (go to sequence)
Protein existence: Evidence at protein level
Annotation score: 4/5

Function

Involved in oxygen transport from the lung to the various peripheral tissues.

Features

Showing features for binding site.



Sequence

Sequence status: Complete

See also sequence in UniParc or sequence clusters in UniRef

Length 142
Mass (Da) 15,468

Last updated 2007-01-23 v2
MD5 Checksum 916D390BA6400A30A22A3DE0F83C3BB5

```
        10         20         30         40         50         60         70         80         90
MVLSAADKTN VKGVFSKISG HAEEYGAETL ERMFTAYPQT KTYFPHFDLQ HGSAQIKAHG KKVVAALVEA VNHIDDIAGA LSKLSDLHAQ

       100        110        120        130        140
KLRVDPVNFK FLGHCFLVVV AIHHPSALTA EVHASLDKFL CAVGTVLTAK YR
```

Keywords

Technical term #3D-structure
#Direct protein sequencing

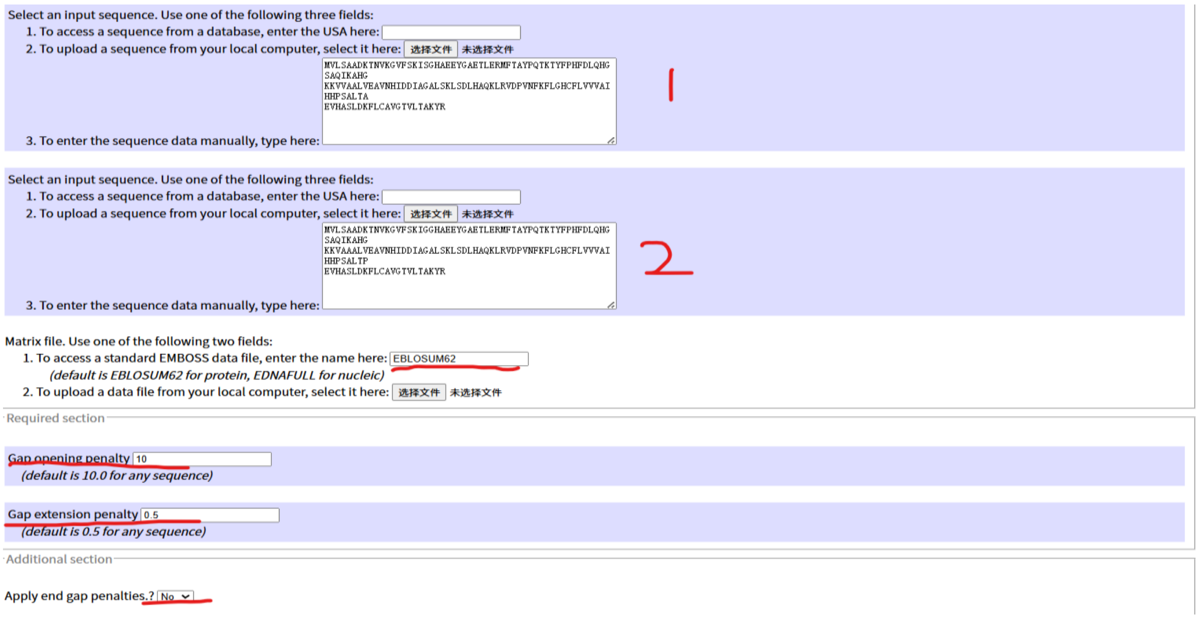
Sequence databases

PIR A02311 HAGSI

```
>sp|P01989|HBA_ANSAN Hemoglobin subunit alpha-A OS=Anser anser anser OX=8844 GN=HBAA PE=1 SV=2
MVLSAADKTNVKGVFSKIGGHAEEYGAETLERMFTAYPQTKTYFPHFDLQHGSAQIKAHG
KKVAAALVEAVNHIDDIAGALSKLSDLHAQKLRVDPVNFKFLGHCFLVVVAIHHPSALTP
EVHASLDKFLCAVGTVLTAKYR
```

```
>sp|P01990|HBA_ANSIN Hemoglobin subunit alpha-A OS=Anser indicus OX=8846 GN=HBAA PE=1 SV=2
MVLSAADKTNVKGVFSKISGHAEEYGAETLERMFTAYPQTKTYFPHFDLQHGSAQIKAHG
KKVVAALVEAVNHIDDIAGALSKLSDLHAQKLRVDPVNFKFLGHCFLVVVAIHHPSALTA
EVHASLDKFLCAVGTVLTAKYR
```

3)将序列填入needle程序输入框中选择默认计分矩阵EBLOSUM62(EMBOSS软件包中所有计分矩阵都冠以字母E)、默认起始空位(GAP OPEN)罚分10和延伸空位(GAP EXTEND)罚分0.5，序列末端空位(END GAP PENALTY)不予罚分(No)。



Select an input sequence. Use one of the following three fields:
1. To access a sequence from a database, enter the USA here:
2. To upload a sequence from your local computer, select it here: 选择文件 未选择文件
3. To enter the sequence data manually, type here:

```
MVLSAADKTNVKGVFSKISGHAEEYGAETLERMFTAYPQTKTYFPHFDLQHG
SAQIKAHG
KKVVAALVEAVNHIDDIAGALSKLSDLHAQKLRVDPVNFKFLGHCFLVVVAI
HHPSALTA
EVHASLDKFLCAVGTVLTAKYR
```

Select an input sequence. Use one of the following three fields:
1. To access a sequence from a database, enter the USA here:
2. To upload a sequence from your local computer, select it here: 选择文件 未选择文件
3. To enter the sequence data manually, type here:

```
MVLSAADKTNVKGVFSKIGGHAEEYGAETLERMFTAYPQTKTYFPHFDLQHG
SAQIKAHG
KKVAAALVEAVNHIDDIAGALSKLSDLHAQKLRVDPVNFKFLGHCFLVVVAI
HHPSALTP
EVHASLDKFLCAVGTVLTAKYR
```

Matrix file. Use one of the following two fields:
1. To access a standard EMBOSS data file, enter the name here: EBLOSUM62
(default is EBLOSUM62 for protein, EDNAFULL for nucleic)
2. To upload a data file from your local computer, select it here: 选择文件 未选择文件

Required section

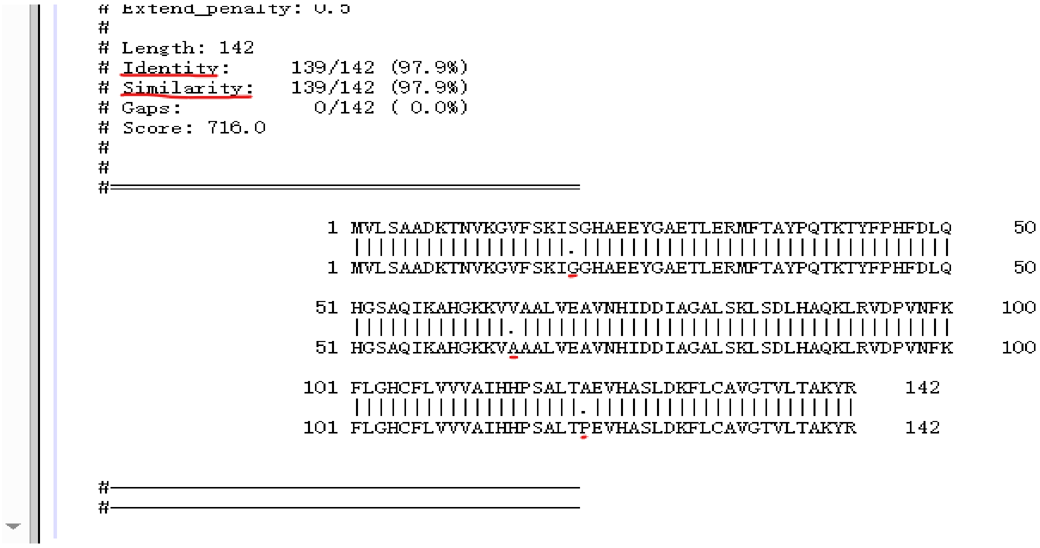
Gap opening penalty: 10
(default is 10.0 for any sequence)

Gap extension penalty: 0.5
(default is 0.5 for any sequence)

Additional section

Apply end gap penalties? No

4) 点击Submit按钮，几秒钟后，输出运行结果。输出结果显示，斑头雁和灰雁血红蛋白alpha亚基共有三个位点差异。

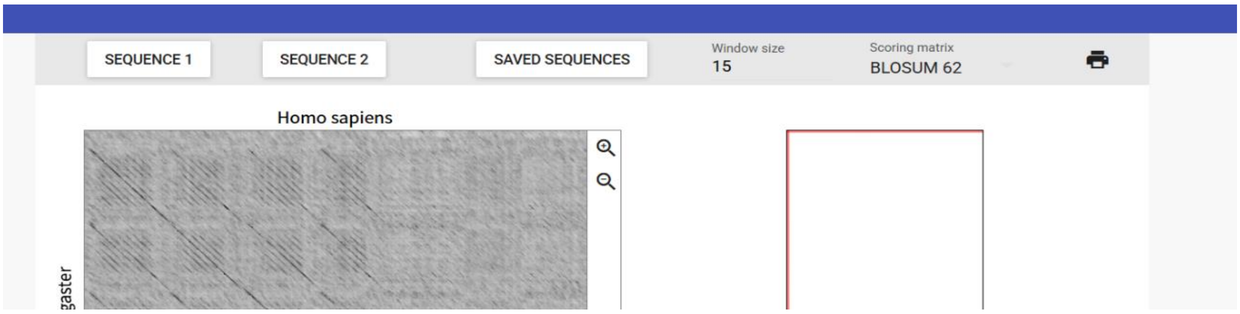


```
# Extend_penalty: 0.5
#
# Length: 142
# Identity:     139/142 (97.9%)
# Similarity:   139/142 (97.9%)
# Gaps:           0/142 ( 0.0%)
# Score: 716.0
#
#
#=======================================


  1 MVLSAADKTNVKGVFSKISGHAEEYGAETLERMFTAYPQTKTYFPHFDLQ     50
    ||||||||||||||||||.|||||||||||||||||||||||||||||||
  1 MVLSAADKTNVKGVFSKIGGHAEEYGAETLERMFTAYPQTKTYFPHFDLQ     50

 51 HGSAQIKAHGKKVVAALVEAVNHIDDIAGALSKLSDLHAQKLRVDPVNFK    100
    |||||||||||||.||||||||||||||||||||||||||||||||||||
 51 HGSAQIKAHGKKVAAALVEAVNHIDDIAGALSKLSDLHAQKLRVDPVNFK    100

101 FLGHCFLVVVAIHHPSALTAEVHASLDKFLCAVGTVLTAKYR    142
    |||||||||||||||||||.||||||||||||||||||||||
101 FLGHCFLVVVAIHHPSALTPEVHASLDKFLCAVGTVLTAKYR    142


#---------------------------------------
#---------------------------------------
```

5. 问题

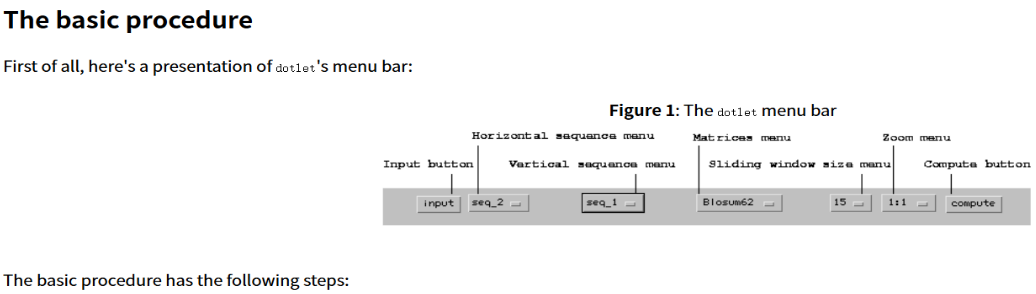
Q1: Dotlet 网站中的教学指南与现在网站存在部分不符合情况，例如操作界面

Q1 建议———以例子进行实操后形成可用的教学指南，帮助操作和举一反三。



Window size 15

Scoring matrix BLOSUM 62



The basic procedure

First of all, here's a presentation of `dotlet`'s menu bar:

Figure 1: The `dotlet` menu bar

The basic procedure has the following steps:

Q2:知网可以批量下载 PDF 版本吗

Q2 建议———可以批量下载 PDF，但知网网页端不直接支持，必须用知网研学（原 E-Study）或浏览器脚本两种方式。

1 官方方法：知网研学 APP

1.1 下载安装知网研学（官网/应用商店）

1.2 登录（用知网账号/机构授权）

1.3 检索→勾选文献→添加到学习单元

1.4 进入学习单元→勾选→下载→选 PDF

1.5 提前设置：设置→知网全文获取→优先 PDF

2 网页端+油猴脚本（快速）

2.1 装浏览器插件 Tampermonkey（油猴）

2.2 去 Greasy Fork 搜知网批量下载 PDF 脚本并安装

2.3 知网检索→勾选→点脚本生成的批量下载 PDF 按钮

2.4 允许浏览器弹窗，即可批量下载

Q3: NCBI 使用起来不是很流畅，有什么可以代替 NCBI 的功能？比起 NCBI 来它有什么短缺？

Q3 建议———文献检索、序列检索、序列比对、基因组浏览、蛋白分析等功能可以使用 EBI 代替，但是相比起来，NCBI 的整合更全面，比如对某个基因进行检索时一个页面可以聚合所有相关信息可以直接跳转，EBI 可能需要多个入口查看不同的信息，并且缺少一体化的引物设计功能，如果需要快速检索综合信息或者设计引物选择 NCBI 更好，其他许多功能可以使用 EBI。

6. 个人总结

G2A 边汉青个人总结

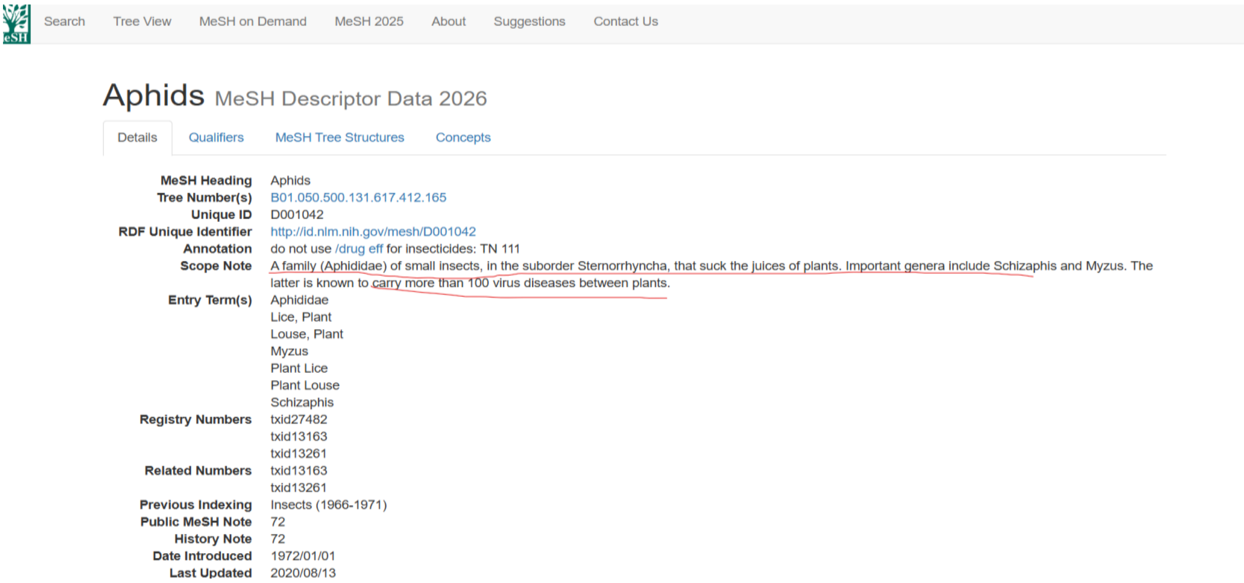
G2B 刘奇个人总结

G2C 高倩个人总结

G2A 边汉青个人总结

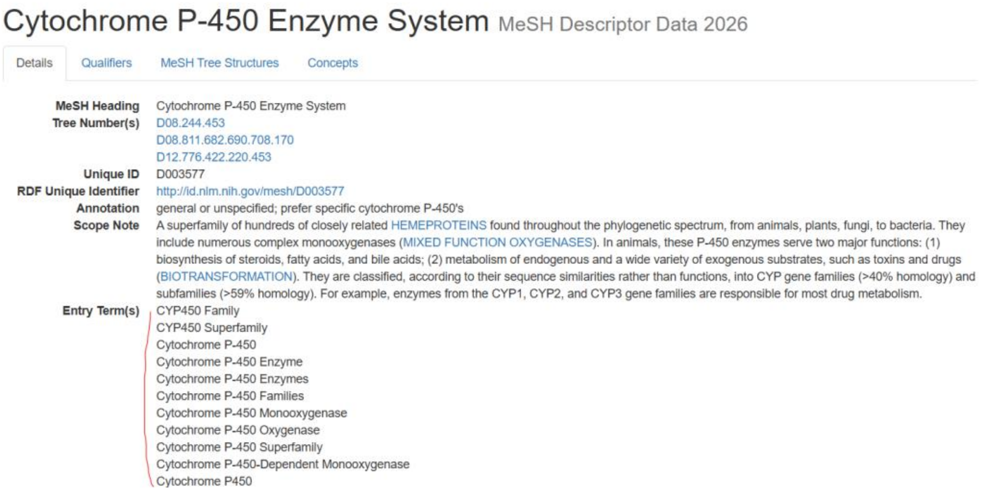
利用 Mesh、Uniprot、EPMC 和生物大分子月报来查询自己方向或感兴趣的信息。

1.确定自己的搜索主题或搜索词，在 Mesh 中找到精确的表达词，例如蚜虫：aphid 和细胞色素氧化酶 P450 Cytochrome P-450 Enzyme System.



Aphids MeSH Descriptor Data 2026

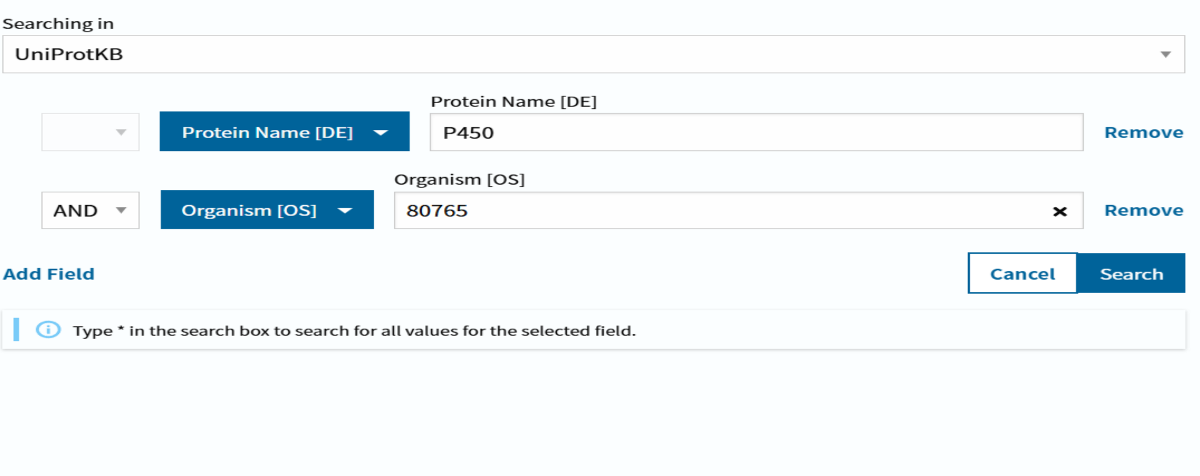
MeSH Heading	Aphids
Tree Number(s)	B01.050.500.131.617.412.165
Unique ID	D001042
RDF Unique Identifier	http://id.nlm.nih.gov/mesh/D001042
Annotation	do not use /drug eff for insecticides: TN 111
Scope Note	A family (Aphididae) of small insects, in the suborder Sternorrhyncha, that suck the juices of plants. Important genera include Schizaphis and Myzus. The latter is known to carry more than 100 virus diseases between plants.
Entry Term(s)	Aphididae Lice, Plant Louse, Plant Myzus Plant Lice Plant Louse Schizaphis
Registry Numbers	txid27482 txid13163 txid13261
Related Numbers	txid13163 txid13261
Previous Indexing	Insects (1966-1971)
Public MeSH Note	72
History Note	72
Date Introduced	1972/01/01
Last Updated	2020/08/13



Cytochrome P-450 Enzyme System MeSH Descriptor Data 2026

MeSH Heading	Cytochrome P-450 Enzyme System
Tree Number(s)	D08.244.453 D08.811.682.690.708.170 D12.776.422.220.453
Unique ID	D003577
RDF Unique Identifier	http://id.nlm.nih.gov/mesh/D003577
Annotation	general or unspecified; prefer specific cytochrome P-450's
Scope Note	A superfamily of hundreds of closely related HEMEPROTEINS found throughout the phylogenetic spectrum, from animals, plants, fungi, to bacteria. They include numerous complex monooxygenases (MIXED FUNCTION OXYGENASES). In animals, these P-450 enzymes serve two major functions: (1) biosynthesis of steroids, fatty acids, and bile acids; (2) metabolism of endogenous and a wide variety of exogenous substrates, such as toxins and drugs (BIOTRANSFORMATION). They are classified, according to their sequence similarities rather than functions, into CYP gene families (>40% homology) and subfamilies (>59% homology). For example, enzymes from the CYP1, CYP2, and CYP3 gene families are responsible for most drug metabolism.
Entry Term(s)	CYP450 Family CYP450 Superfamily Cytochrome P-450 Cytochrome P-450 Enzyme Cytochrome P-450 Enzymes Cytochrome P-450 Families Cytochrome P-450 Monooxygenase Cytochrome P-450 Oxygenase Cytochrome P-450 Superfamily Cytochrome P-450-Dependent Monooxygenase Cytochrome P450

2.在 uniprot 中搜索棉蚜中的 P450 基因，找到感兴趣的，例如：*CYP6CY3*

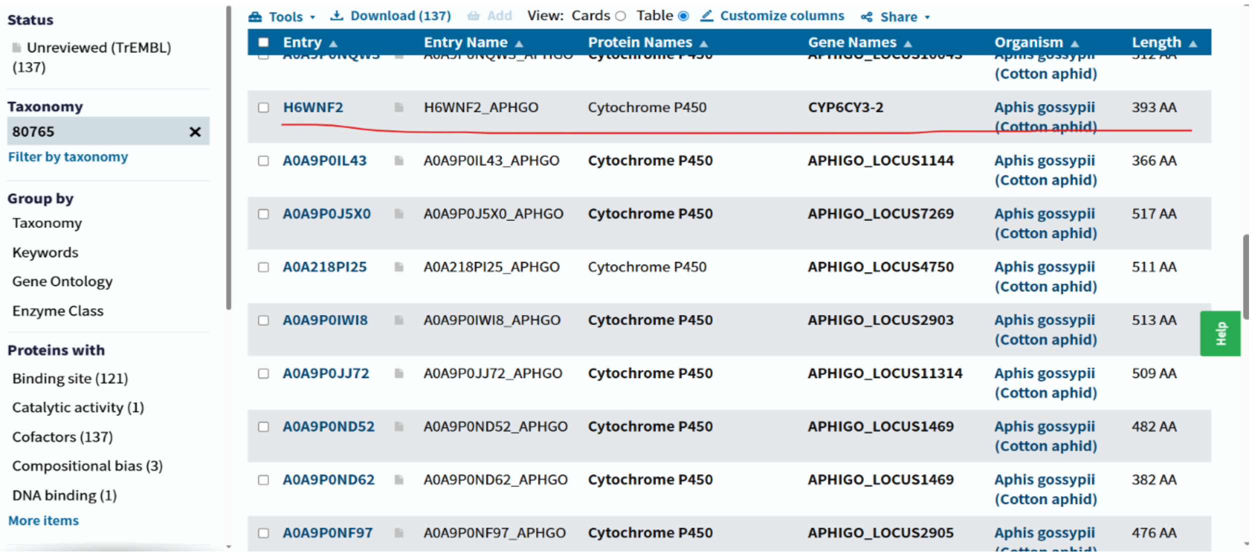


Searching in: UniProtKB

Protein Name [DE]: P450

AND Organism [OS]: 80765

Type * in the search box to search for all values for the selected field.



Status
☐ Unreviewed (TrEMBL) (137)

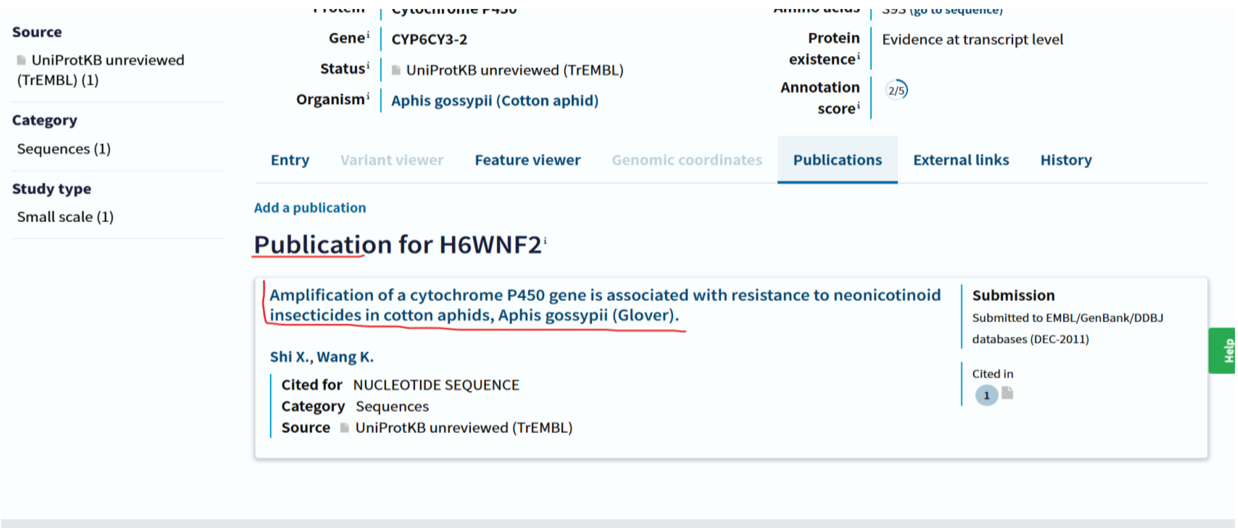
Taxonomy
80765 ✕
Filter by taxonomy

Group by
Taxonomy
Keywords
Gene Ontology
Enzyme Class

Proteins with
Binding site (121)
Catalytic activity (1)
Cofactors (137)
Compositional bias (3)
DNA binding (1)
More items

☐ Entry ▲	Entry Name ▲	Protein Names ▲	Gene Names ▲	Organism ▲	Length ▲
				(Cotton aphid)	
☐ H6WNF2	H6WNF2_APHGO	Cytochrome P450	CYP6CY3-2	Aphis gossypii (Cotton aphid)	393 AA
☐ A0A9P0IL43	A0A9P0IL43_APHGO	Cytochrome P450	APHIGO_LOCUS1144	Aphis gossypii (Cotton aphid)	366 AA
☐ A0A9P0J5X0	A0A9P0J5X0_APHGO	Cytochrome P450	APHIGO_LOCUS7269	Aphis gossypii (Cotton aphid)	517 AA
☐ A0A218PI25	A0A218PI25_APHGO	Cytochrome P450	APHIGO_LOCUS4750	Aphis gossypii (Cotton aphid)	511 AA
☐ A0A9P0IWI8	A0A9P0IWI8_APHGO	Cytochrome P450	APHIGO_LOCUS2903	Aphis gossypii (Cotton aphid)	513 AA
☐ A0A9P0JJ72	A0A9P0JJ72_APHGO	Cytochrome P450	APHIGO_LOCUS11314	Aphis gossypii (Cotton aphid)	509 AA
☐ A0A9P0ND52	A0A9P0ND52_APHGO	Cytochrome P450	APHIGO_LOCUS1469	Aphis gossypii (Cotton aphid)	482 AA
☐ A0A9P0ND62	A0A9P0ND62_APHGO	Cytochrome P450	APHIGO_LOCUS1469	Aphis gossypii (Cotton aphid)	382 AA
☐ A0A9P0NF97	A0A9P0NF97_APHGO	Cytochrome P450	APHIGO_LOCUS2905	Aphis gossypii	476 AA

3.可以在 Publications 中找到其参考的文章，之后可以在 EPMS 上进行搜索，进行阅读了解得更加详细。



Source
UniProtKB unreviewed (TrEMBL) (1)

Category
Sequences (1)

Study type
Small scale (1)

Gene: CYP6CY3-2
Status: UniProtKB unreviewed (TrEMBL)
Organism: Aphis gossypii (Cotton aphid)
Protein existence: Evidence at transcript level
Annotation score: 2/5

Publication for H6WNF2

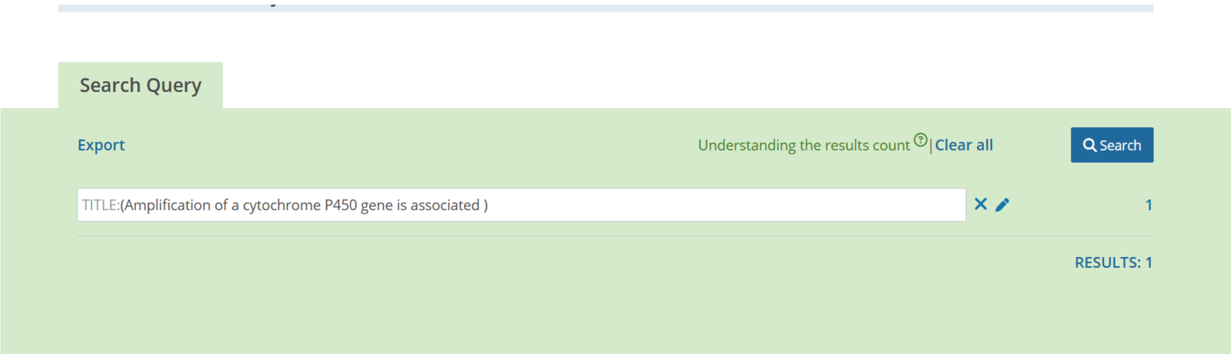
Amplification of a cytochrome P450 gene is associated with resistance to neonicotinoid insecticides in cotton aphids, Aphis gossypii (Glover).

Shi X., Wang K.

Cited for NUCLEOTIDE SEQUENCE
Category Sequences
Source UniProtKB unreviewed (TrEMBL)

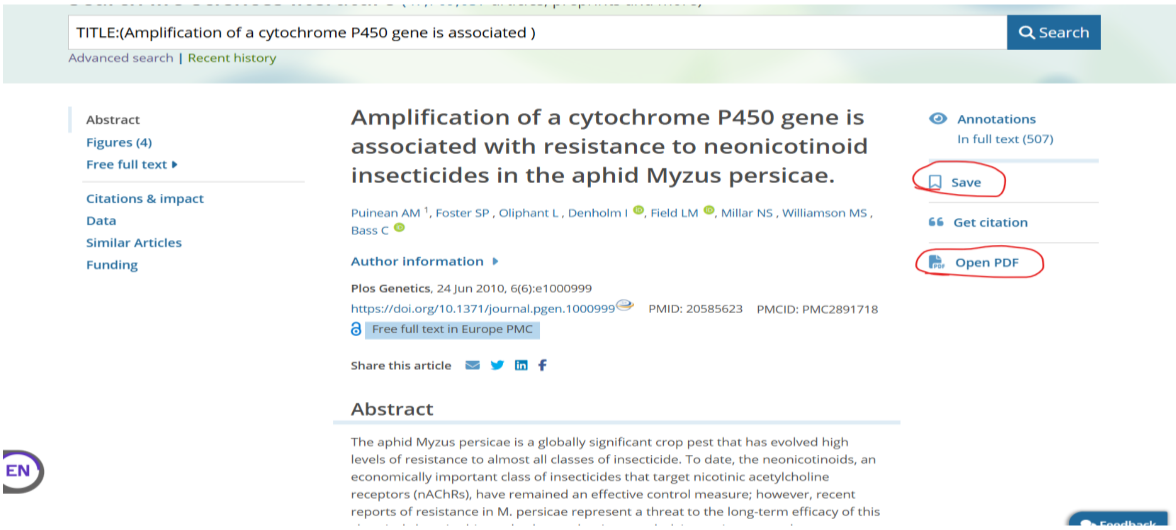
Submission
Submitted to EMBL/GenBank/DDBJ databases (DEC-2011)

Cited in 1



TITLE:(Amplification of a cytochrome P450 gene is associated) 1

RESULTS: 1



TITLE:(Amplification of a cytochrome P450 gene is associated)

Amplification of a cytochrome P450 gene is associated with resistance to neonicotinoid insecticides in the aphid Myzus persicae.

Puinean AM [1], Foster SP, Oliphant L, Denholm I, Field LM, Millar NS, Williamson MS, Bass C

Author information

Plos Genetics, 24 Jun 2010, 6(6):e1000999

https://doi.org/10.1371/journal.pgen.1000999 PMID: 20585623 PMCID: PMC2891718

Free full text in Europe PMC

Abstract

The aphid Myzus persicae is a globally significant crop pest that has evolved high levels of resistance to almost all classes of insecticide. To date, the neonicotinoids, an economically important class of insecticides that target nicotinic acetylcholine receptors (nAChRs), have remained an effective control measure; however, recent reports of resistance in M. persicae represent a threat to the long-term efficacy of this

生物大分子月报关于 P450 的查询与了解。

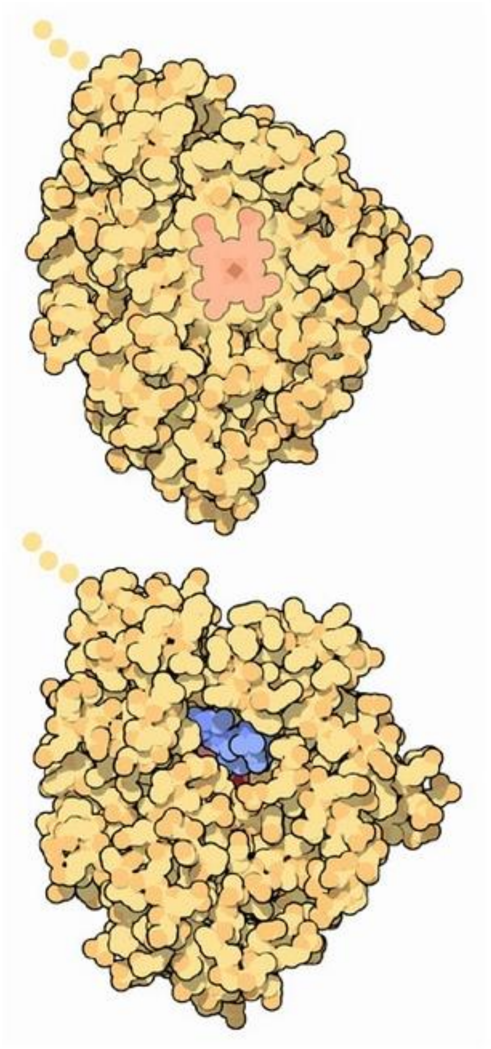
如果你感到头痛，服用药物来缓解疼痛，你会发现药效会在几个小时内逐渐消失。之所以会发生，是因为人体有一套强大的解毒系统，能够识别出不正常的化学物质，例如药物，并将其排出体外。这套系统能应对我们由于进食和呼吸而摄入的各种令人不适的化学物质，包括药物、植物中的有毒物、烹饪过程中形成的致癌物以及环境污染物。细胞色素 P450 酶是我们在这场化学战中的第一道防线。

氧化作用（加氧作用）

细胞色素 P450 酶能够识别出异常分子，并向其添加氧原子。在大多数情况下，这会使分子变得更易溶于水，从而更易排出体内。加氧作用还便于其他解毒酶与其结合，使其能够进一步改造并破坏这些有毒分子。加氧过程颇为棘手，细胞色素 P450 酶利用一种强大的分子工具来完成这一反应：位于血红素基团中的铁原子。

无处不在的 P450

细胞色素 P450 酶存在于所有生物体内。每个生物体都会产生多种不同的酶，每种酶作用于一组特定的分子。通常情况下，细菌会产生约 20 种不同结构的 P450 酶，而我们身体则会产生约 60 种。植物往往会产生数百种，这是因为植物会产生特殊的色素和特殊毒素来保护自己，许多产生这些分子所需的反应都是由专门的细胞色素 P450 酶来完成的。



Cytochrome P450 3A4 (top) bound to erythromycin (bottom).

一把双刃剑

图中所示的分子是 CYP3A4，这是一种在人体内对于药物解毒发挥主要作用的细胞色素 P450。据估计，这种酶能作用于约一半已知的药物。例如，它会改变抗生素红霉素（如图底部所示，呈蓝色），还会对诸如可待因、地西泮（安定）、紫杉醇（泰素）以及几种抗艾滋病毒药物等药物进行解毒。

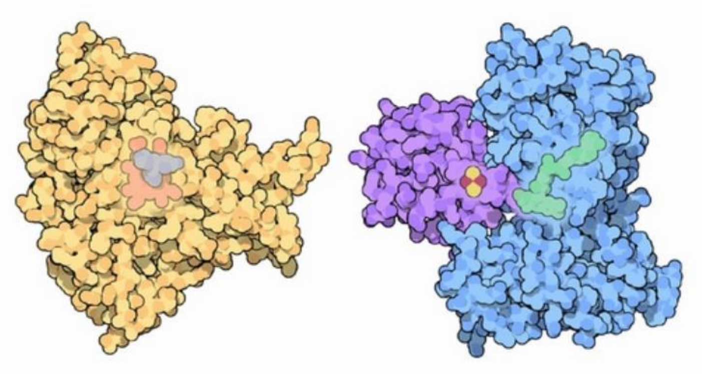
然而，在某些情况下，细胞色素 P450 酶所引发的反应可能会造成危害比益处更多。例如，CYP3A4 是大剂量乙酰氨基酚（泰诺）使身体产生毒性反应的部分原因。这种经过修饰的对乙酰氨基酚具有极强的反应活性，但通常会被其他解毒酶迅速清除。然而，当大剂量使用时，这种活性中间体会累积到危险的水平。

处方与 P450

医生在开具药物处方时必须考虑细胞色素 P450 酶。例如，您可能在处方上看到过这样的警示，告知患者在服用某种药物时不要饮用葡萄柚汁。葡萄柚中含有一个黄酮分子，会抑制细胞色素 P450 酶。这会减缓药物的解毒过程，从而可能导致药物产生的效果比预期的更强烈。

合成高手

细胞色素 P450 酶在正常细胞化合物的合成过程中也发挥着诸多关键作用。例如，部分细胞色素 P450 酶专门被设计参与甾体、维生素 A 和 D 以及进行信号传导的类似脂质的类固醇分子等化合物的化学合成过程中。图左侧所示的这种酶是真菌的细胞色素 P450，它在甾体合成过程中执行一个步骤。我们细胞中的类似酶对于胆固醇的合成也是必需的。右侧的酶复合物为反应提供电子。



Cytochrome P450 14alpha-sterol demethylase (left) and NADPH:adrenodoxin reductase with adrenodoxin (right).

G2B 刘奇个人总结

针对我计划研究的某入侵昆虫物种生育相关基因，以下数据库可协同使用进行分析。

首先 NCBI 主要用于快速获取该基因基础信息。进入 Gene 数据库检索该生育相关基因，可一次性获得参考序列（RefSeq）、相关文献和已知变异等信息。且可以使用 Primer-BLAST 进行引物设计，输入基因名后自动生成 qPCR 引物并验证特异性后，可用于后续检测该基因在不同组织（如卵巢、精巢）和时期的表达。

Europe PMC 作为文献检索的增强版，与 NCBI 的 PubMed 内容高度重叠，但有更强的批量导出和文本挖掘功能，用它批量检索该昆虫物种所有涉及生殖发育的文献，快速建立研究背景。

EBI 的 Ensembl 用于基因组可视化。入侵物种作为非模式生物可能缺乏高质量参考基因，通过 Ensembl 查看该基因在染色体上的精确位置、外显子-内含子结构及相邻基因排布，通过 Compara 模块还可以找到模式生物中的同源基因，根据这些信息可以推测我研究的基因可能有什么功能，是否具有保守性。

CNCB 用于本土数据提交和查找。研究入侵中国的昆虫物种，测序原始数据需提交至 GSA 以满足国内期刊和数据安全要求，且从 CNCB 下载国内项目数据可能更全、速度更快。

UniProt 是蛋白功能注释的权威平台。若该生育相关基因编码蛋白，或需要进行相关蛋白的研究，UniProt 提供结构域注释，可判断是否含有已知的生育相关保守结构域；同时可查找翻译后修饰位点、亚细胞定位和蛋白相互作用网络，为后续蛋白功能实验提供关键线索。

而 PDB 用于蛋白结构可视化。若该生育相关蛋白有三维结构，从 PDB 下载文件可分析关键氨基酸位点在三维空间的位置。

各平台各有侧重：NCBI 解决快速查询和引物设计，Europe PMC 支持文献批量处理，UniProt 提供蛋白深度注释，EBI 的 Ensembl 完成基因组可视化，CNCB 保障数据合规与可及，PDB 补充结构信息。掌握这种跨平台协同能力，是我开展昆虫生育相关基因功能研究的基础。

G2C 高倩个人总结

过去三周，我主要学习了 NCBI、EBI、CNCB 等常用生物信息数据库的基本功能与使用场景，进一步熟悉了 OpenLB、PubMed、EPMC、知网等文献检索平台的操作方法，了解到生物大分子月报，化学小分子月报，蛋白质精选等生物网络文献库，同时结合课程内容自主学习了血红蛋白、绿色荧光蛋白、胰岛素、气味受体等多种生物分子的结构与功能。在 ABC 网页首页，EMBOSS 软件包的主要开发者之一的 Alan Bleseby 曾说：Half day on the Web, saves you half month in the lab! 然而，经过过去三个周的学习，我深刻体会其中深意，也特别荣幸能使用 ABC 这个没有包装却功能齐全，内容翔实的学习工具。

在学习过程中，我也遇到了一些问题。一是专业名词较多，尤其是不同物种的嗅觉受体蛋白概念复杂，记忆和区分较为困难；二是文献检索技巧不够熟练，对高级检索语法和文献管理工具运用不够流畅，还需加强练习；三是自主学习主动性和自律仍需不断提高。

针对这些问题，我计划通过利用 AI 分类归纳相似蛋白概念，每周计划固定时间个人或复习每周内容。希望以上存在的问题在下一次总结中体现进步和改善。