BIG Search

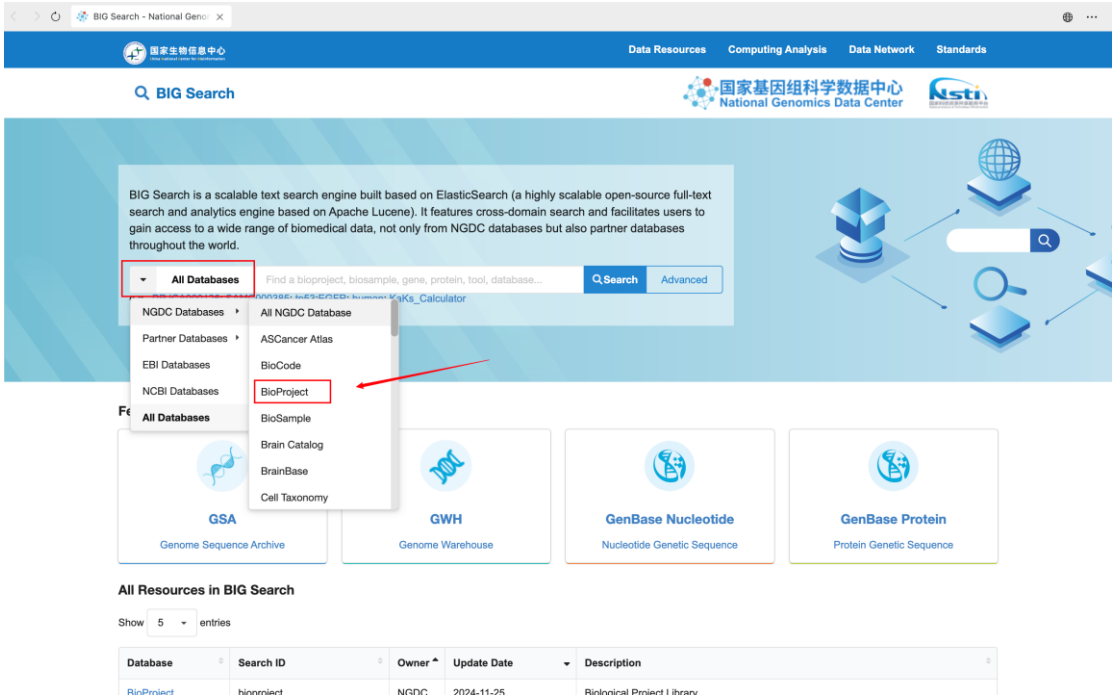
BIG Search 跨库检索服务（BIG Search）是通过建立面向多层次生物数据的通用索引标准，基于 Elasticsearch 等前沿信息技术，实现数据高效可扩展的一站式搜索引擎。目前 BIG Search 已整合中心和华中科技大学、四川大学、北京大学、中国医学科学院、华中农业大学、哈尔滨医科大学、中科院北京生命科学研究院等多个合作结构的 140 余个数据资源库，集成数据索引 15.4 亿条，索引存储达到 1.6TB，同时还整合了 EBI 和 NCBI 的数据资源，能够实现高性能、灵活的全文索引和检索，帮助用户更有效、快速、准确地获取知识。

一、检索

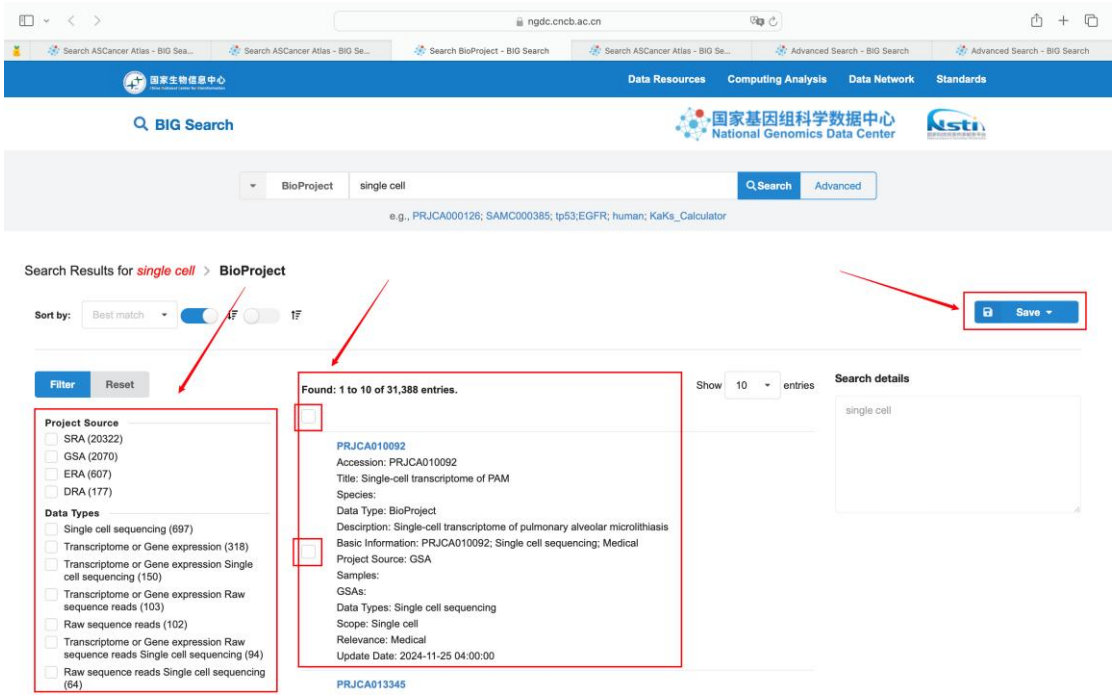
系统首页检索入口

（1）具体数据库检索

用户可以预设检索数据库，并输入关键词进行指定数据库资源检索。



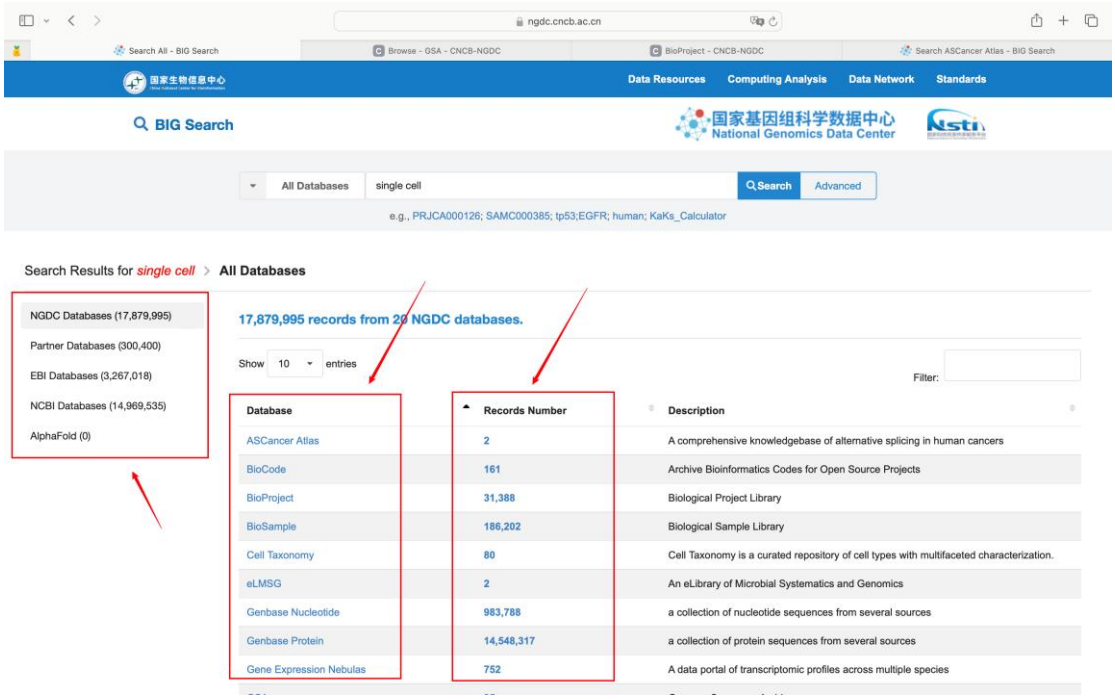

点击 search 按钮后会跳转至对应数据库的检索结果页面，用户可以在左侧筛选框设置详细的筛选条件点击 Filter 按钮筛选。右侧可以浏览检索列表，点击对应名称查看详情，用户也可以选择多个所需的数据资源，点击右上角 Save 按钮进行数据下载。




（2）全局检索

用户可以直接在检索框中输入关键词，点击 search 按钮进行全局资源检索。

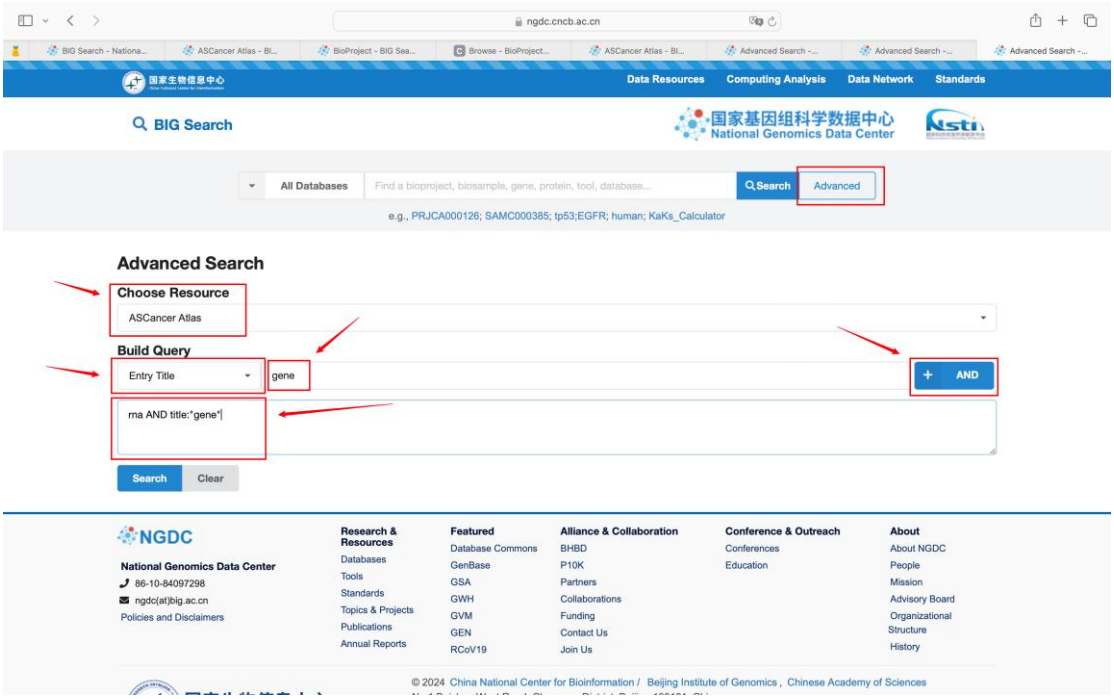
检索后系统会对关键词进行全局检索。用户可以在左侧选择数据库的归属，页面主体会列举出不同数据库中与关键词相关的具体数据库和该数据库中对应的资源数量，点击具体 Database 名称可以跳转至该数据库的首页，点击 Records Number 具体数据可以跳转至该数据中的检索结果列表。



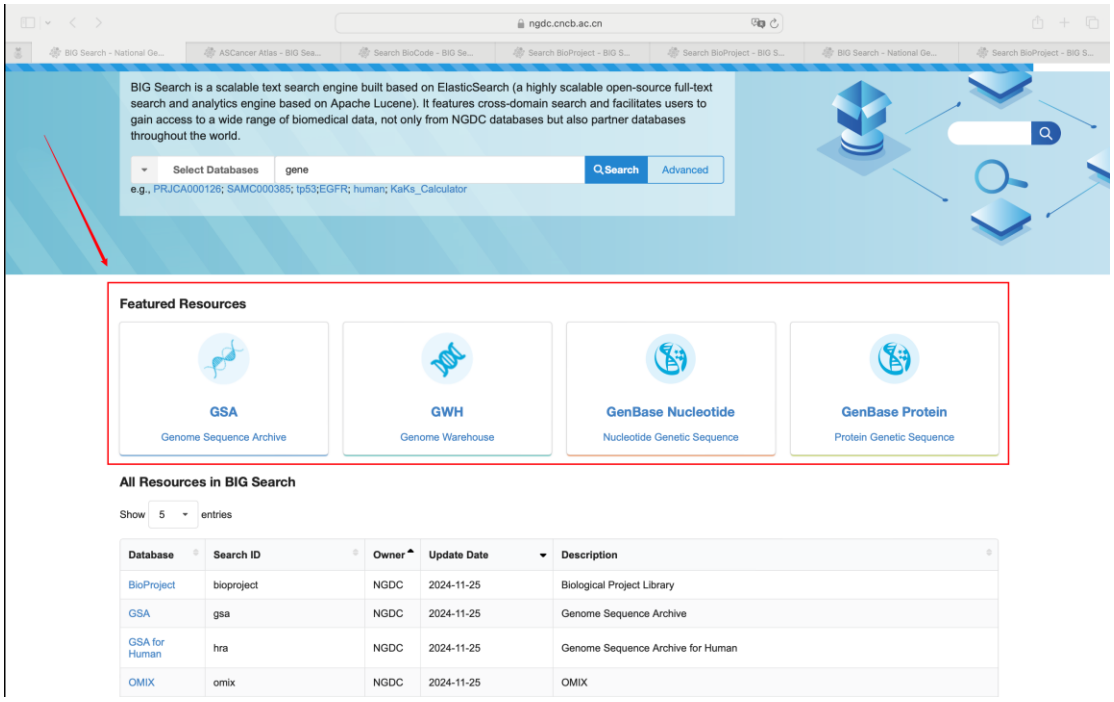

（3）高级检索

点击 Advanced 按钮可进行高级检索。高级检索页面用户可以通过 Choose Resource 筛选检索数据库，在 Build Query 可以拼接关键词匹配的数据字段，例如选择“Entry Title”点击右侧 AND 按钮会拼接基于 Title 的关键词查询。如果有多个筛选条件可通过右侧按钮进行“AND”、

“OR”、“NOT”拼接。

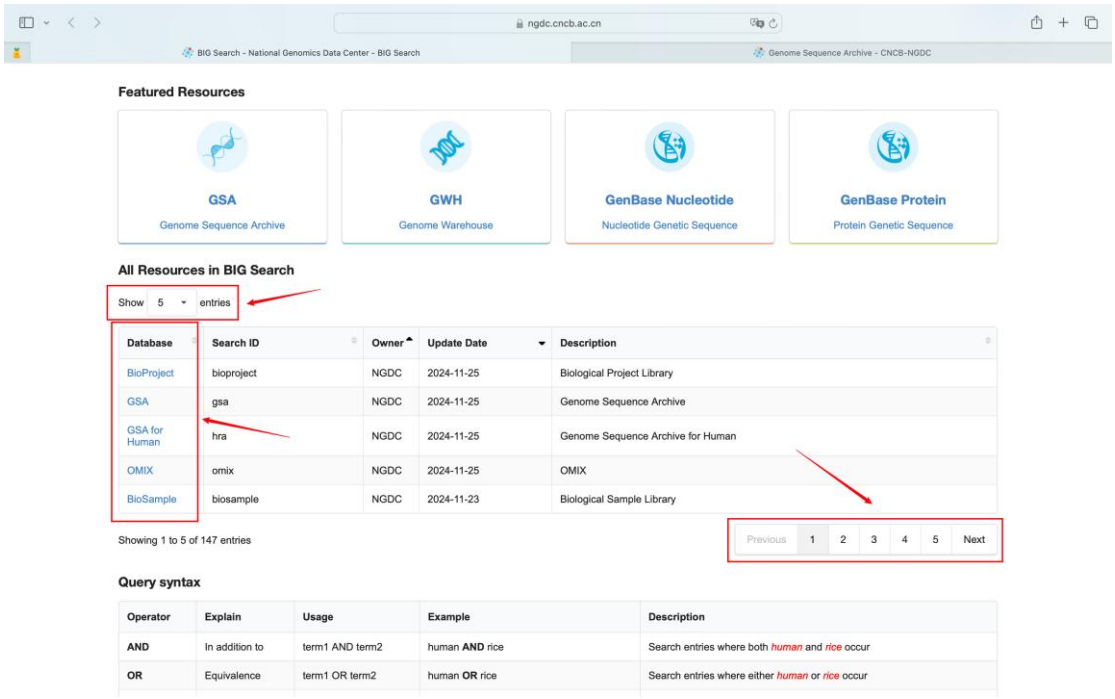



二、Big Search 中的特色资源库，点击可快速跳转至资源库



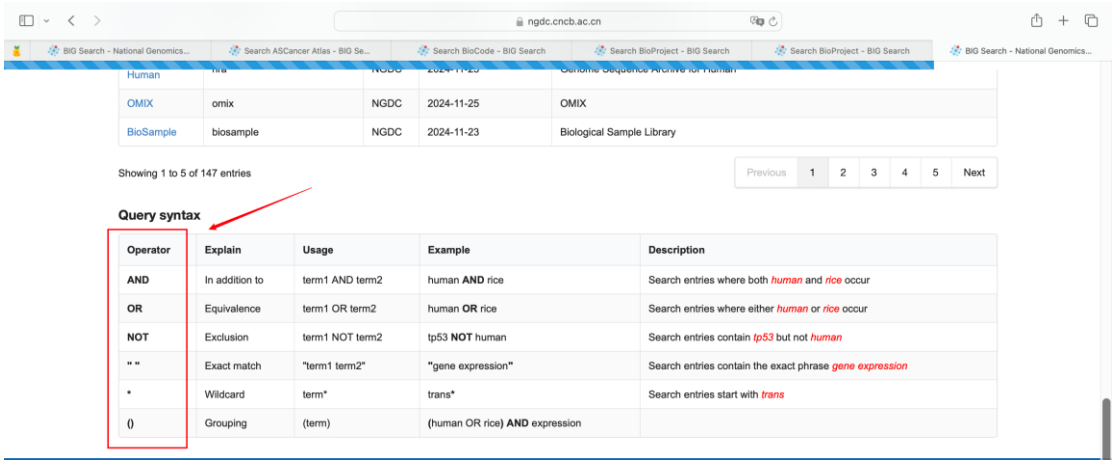

三、Big search 覆盖的数据资源浏览

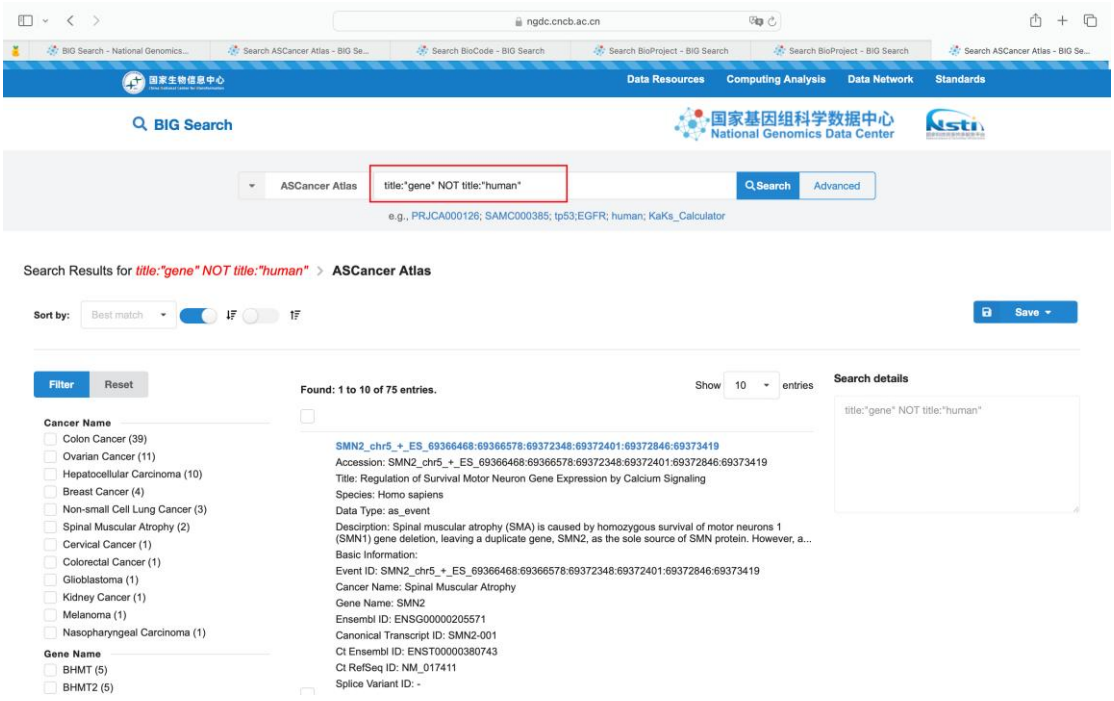
点击“show entries”可以设置每页展示的数据库数量，默认为 5 个。在数据库列表点击具体的数据库名称可跳转至详情页面。




四、查询语法

系统支持 AND、OR、NOT 等多种查询语法，并且查询语法可拼接使用，用户可以在检索模块通过查询语法进行高级检索，例如：title:"gene" NOT title:"human"会筛选标题中含有“gene”并且不含有“human”的数据资源




ngdc.cncb.ac.cn
BIG Search - National Genomics...
Search ASCancer Atlas - BIG Se...
Search BioCode - BIG Search
Search BioProject - BIG Search
Search BioProject - BIG Search
Search ASCancer Atlas - BIG Se...
国家生物信息中心
Data Resources
Computing Analysis
Data Network
Standards
BIG Search
国家基因组科学数据中心
National Genomics Data Center
ASCancer Atlas
title:"gene" NOT title:"human"
Search
Advanced
e.g., PRJCA000126; SAMC000385; tp53;EGFR; human; KaKs_Calculator
Search Results for title:"gene" NOT title:"human" > ASCancer Atlas
Sort by:
Best match
Save
Filter
Reset
Found: 1 to 10 of 75 entries.
Show 10 entries
Search details
title:"gene" NOT title:"human"
Cancer Name
Colon Cancer (39)
Ovarian Cancer (11)
Hepatocellular Carcinoma (10)
Breast Cancer (4)
Non-small Cell Lung Cancer (3)
Spinal Muscular Atrophy (2)
Cervical Cancer (1)
Colorectal Cancer (1)
Glioblastoma (1)
Kidney Cancer (1)
Melanoma (1)
Nasopharyngeal Carcinoma (1)
Gene Name
BHMT (5)
BHMT2 (5)
SMN2_chr5_+_ES_69366468:69366578:69372348:69372401:69372846:69373419
Accession: SMN2_chr5_+_ES_69366468:69366578:69372348:69372401:69372846:69373419
Title: Regulation of Survival Motor Neuron Gene Expression by Calcium Signaling
Species: Homo sapiens
Data Type: as_event
Descirption: Spinal muscular atrophy (SMA) is caused by homozygous survival of motor neurons 1 (SMN1) gene deletion, leaving a duplicate gene, SMN2, as the sole source of SMN protein. However, a...
Basic Information:
Event ID: SMN2_chr5_+_ES_69366468:69366578:69372348:69372401:69372846:69373419
Cancer Name: Spinal Muscular Atrophy
Gene Name: SMN2
Ensembl ID: ENSG00000205571
Canonical Transcript ID: SMN2-001
Ct Ensembl ID: ENST00000380743
Ct RefSeq ID: NM_017411
Splice Variant ID: -