

LXgeo 的使用

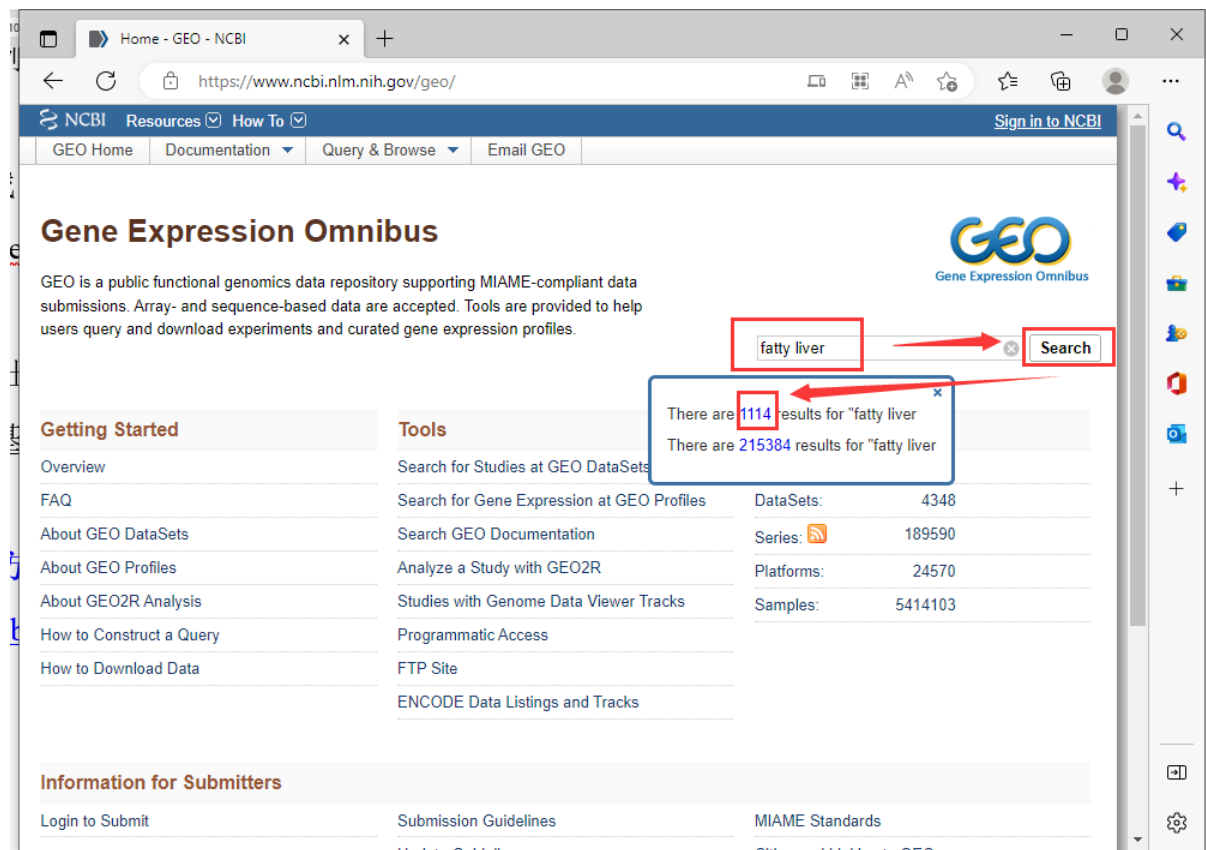
林兴 2022.12.10

一、制作 LXgeo 的目的： 在没有开展转录组测序的情况下，如何快速知道哪些基因参与了疾病的调控。例如，想开展延龄草苷抗非酒精性脂肪肝（NAFLD）的研究，可以这么操作：

- （1）利用 LXgeo 找出正常组与模型组的差异基因，即疾病相关基因；
- （2）用 pharmpaper、targetnet、CTD、SEA 等在线数据库查找延龄草苷 targets，即药物靶基因；
- （3）用 LXvenn 筛出 疾病相关基因 和 药物靶基因 的共同基因，即两者的交集基因，然后，用交集基因作 LXkegg 分析，确定研究哪个代谢通路/信号通路。

二、LXgeo 的使用方法：

- （1）<https://www.ncbi.nlm.nih.gov/geo/> 查找疾病关键词，设定查找条件：



fat liver - GEO DataSets - NCBI

https://www.ncbi.nlm.nih.gov/gds

National Library of Medicine
National Center for Biotechnology Information

筛选条件: 如 小鼠, 基因列表、转录组

GEO DataSets

Search: fatty liver

Entry type: DataSets (0)
Series (245)
Samples (0)
Platforms (0)

Organism: Customized ...

Study type: clear
Expression profiling by array
Methylation profiling by array
Customize ...

Author: Customize ...

Attribute name: Tissue (180)
Strain (120)
Customize ...

Publication dates: 30 days
1 year

Search results
Items: 1 to 20 of 245

Filters: Manage Filters

Top Organisms [Tree]
Mus musculus (180)
Homo sapiens (42)
Rattus norvegicus (17)
Danio rerio (2)
Mesocricetus auratus (1)
More...

Find related data
Database: Select
Find items

Search details
("fatty liver"[MeSH Terms] OR fatty liver[All Fields]) AND ("gse"[Filter]) AND "Expression profiling by array"[Filter]

1. NIK/MAP3K14 in hepatocytes orchestrates NASH to hepatocellular carcinoma progression via JAK2/STAT5 inhibition
(Submitter supplied) Nonalcoholic fatty liver disease (NAFLD) ranges from steatosis to nonalcoholic steatohepatitis (NASH), that often progresses to hepatocellular carcinoma (HCC) through a largely undefined mechanism. NASH and HCC both depend on inflammatory signaling whose master regulator is the NFkB transcription factor family, activated by canonical and non-canonical pathways. Here, we investigated non-canonical NFkB-inducing kinase (NIK/MAP3K14) in metabolic NASH, NASH to HCC transition as well as in DEN-induced HCC. more...

Organism: Homo sapiens
Type: Expression profiling by array
Platform: GPL10558 12 Samples
Download data: TXT
Series Accession: GSE146049 ID: 200146049
PubMed Analyze with GEO2R

(2) 选出合适的芯片 (GSE): 同时含有基因表达矩阵和基因 symbol

基因表达矩阵: 芯片 GSE205390, 检测平台 GPL1261

(fatty liver) AND "Mus musculus"

https://www.ncbi.nlm.nih.gov/gds

Download data: CEL, CHP
Series Accession: GSE198173 ID: 200198173
PubMed Full text in PMC Similar studies Analyze with GEO2R

4. IkB ζ Regulates the Development of Nonalcoholic Fatty Liver Disease through the Attenuation of Hepatic Steatosis in Mice
(Submitter supplied) IkB ζ is a transcriptional regulator that augments inflammatory responses from the Toll-like receptor or interleukin signaling. These innate immune responses contribute to the progression of nonalcoholic fatty liver disease (NAFLD); however, the role of IkB ζ in the pathogenesis of NAFLD remains elusive. We investigated whether IkB ζ was involved in the progression of NAFLD in mice. We generated hepatocyte-specific IkB ζ -deficient mice (Alb-Cre; Nfkbiz1/fl) by crossing Nfkbiz1/fl mice with Alb-Cre transgenic mice. more...

Organism: Mus musculus
Type: Expression profiling by array
Platform: GPL1261 6 Samples
Download data: CEL
Series Accession: GSE205390 ID: 200205390
Analyze with GEO2R

检测平台

芯片号

内容合适

GSM6211295 Liver_Luciferase_4w_rep3

Relations
BioProject PRJNA844960

Analyze with GEO2R

Download family
SOFT formatted family file(s)
MINIML formatted family file(s)
Series Matrix File(s)

Format
SOFT [?] [?] [?]
MINIML [?]
TXT [?]

Supplementary file

Supplementary file	Size	Download	File type/resource
GSE205390_RAW.tar	19.4 Mb	(http)(custom)	TAR (of CEL)

Raw data provided as supplementary file
Processed data included within Sample table

基因表达矩阵

NLM | NIH | GEO Help | Disclaimer | Accessibility |
HHS Vulnerability Disclosure

GEO Accession viewer x Index of /geo/series/GSE205nnn/GSE205390/matrix/

https://ftp.ncbi.nlm.nih.gov/geo/series/GSE205nnn/GSE205390/matrix/

Index of /geo/series/GSE205nnn/GSE205390/matrix

Name	Last modified	Size
Parent Directory		-
GSE205390_series_matrix.txt.gz	2022-12-06 00:25	1.3M
HHS Vulnerability Disclosure		

基因表达矩阵文件
一般较大 (>1M),
如果只有几个k, 说明可不
含基因表达信息

基因 symbol: 查看 GPL1261

Contributor(s) Hayakawa M
Citation missing Has this study been published? Please login to update or notify GEO.
Submission date Jun 02, 2022
Last update date Jun 07, 2022
Contact name Morisada Hayakawa
E-mail(s) morisada@jichi.ac.jp
Organization name Jichi Medical University
Street address 3311-1
City Yakushiji Shimotsuke-shi
State/province Tochigi
ZIP/Postal code 329-0498
Country Japan

Platforms (1) **GPL1261** [Mouse430_2] Affymetrix Mouse Genome 430 2.0 Array
Samples (6) [GSM6211293](#) Liver_Luciferase_4w_rep1
[GSM6211294](#) Liver_Luciferase_4w_rep2
[GSM6211295](#) Liver_Luciferase_4w_rep3

平台信息
查看基因symbol
形式: gene_symbol
gene symbol
symbol

GEO Accession viewer x

https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GPL1261

Alternative to [GPL29970](#) (Alternative CDF [Mouse4302_Mm_ENTREZG_24.0.0])

Data table header descriptions

ID	Affymetrix Probe Set ID
GB_ACC	GenBank Accession Number
SPOT_ID	Identifies controls
Species Scientific Name	The genus and species of the organism represented by the probe set.
Annotation Date	The date that the annotations for this probe array were last updated. It will generally be earlier than the date when the annotations were posted on the Affymetrix web site.
Sequence Type	
Sequence Source	The database from which the sequence used to design this probe set was taken.
Target Description	
Representative Public ID	The accession number of a representative sequence. Note that for consensus-based probe sets, the representative sequence is only one of several sequences (sequence sub-clusters) used to build the consensus sequence and it is not directly used to derive the probe sequences. The representative sequence is chosen during array design as a sequence that is best associated with the transcribed region being interrogated by the probe set. Refer to the "Sequence Source" field to determine the database used.
Gene Title	Title of Gene represented by the probe set.
Gene Symbol	A gene symbol, when one is available (from UniGene).
ENTREZ_GENE_ID	Entrez Gene Database UID
RefSeq Transcript ID	References to multiple sequences in RefSeq. The field contains the ID and Description for each entry, and there can be multiple entries per ProbeSet.
Gene Ontology Biological Process	Gene Ontology Consortium Biological Process derived from LocusLink. Each annotation consists of three parts: "Accession Number // Description // Evidence". The description corresponds directly to the GO ID. The evidence can be "direct", or "extended".
Gene Ontology Cellular Component	Gene Ontology Consortium Cellular Component derived from LocusLink. Each annotation consists of three parts: "Accession Number // Description // Evidence". The description corresponds directly to the GO ID. The evidence can be "direct", or "extended".
Gene Ontology Molecular Function	Gene Ontology Consortium Molecular Function derived from LocusLink. Each annotation consists of three parts: "Accession Number // Description // Evidence". The description corresponds directly to the GO ID. The evidence can be "direct", or "extended".

	Gene Symbol	ENTREZ_GENE_ID	RefSeq Transcript ID
complex, subunit gamma 1	Copg1	54161	NM_017477 /// NM_201244 /// XM_006506386
orting, lysosomal V0 subunit D1	Atp6v0d1	11972	NM_013477
golain subfamily a. 7	Golga7	57437	NM_001042484 /// NM_020585 /// XM_006509179

(3) 安装和运行 LXgeo

```
# 安装 LXgeo

if(!requireNamespace("devtools"))
  install.packages("devtools")
library(devtools)

install_github("gluck4668/LXgeo")

library(LXgeo)

??LXgeo #查看使用方法

rm(list=ls())

setwd("D:/Desktop/R_example/LXgeo_example")

GSE_id="GSE205390" # https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi

Platforms="GPL1261" # Please check if there is the gene symbol.

LXgeo(GSE_id, Platforms)
```

(4) 最后，查看结果 “GSE205390_gene_expression_Set.xlsx”

	Liver_Luciferase_4w_rep1	Liver_Luciferase_4w_rep2	Liver_Luciferase_4w_rep3	Liver_Nfkbiz_4w_rep1	Liver_Nfkbiz_4w_rep2	Liver_Nfkbiz_4w_rep3
Copg1	1562.015	1433.58	1085.045	1076.428	1399.348	1463.481
Atp6v0d1	5969	5805.005	6304.025	4946.33	6249.883	4702.768
Golga7	5407.256	5598.942	6536.159	5809.478	6553.324	5882.507
PspH	291.2404	170.861	252.5936	278.7615	502.9802	409.5932
Trappc4	1435.994	1443.274	1056.69	1632.825	1609.278	1668.318
Dpm2	744.6266	785.4747	820.2359	686.087	820.7115	924.1483
Psmb5	4904.045	4674.425	4937.641	5230.911	6939.502	5580.984
Dhrs1	2873.656	3559.34	2734.757	2907.811	3429.717	3543.358
Ppm1a	4240.086	4458.673	4227.981	4213.678	4312.568	3877.447
Psenen	6070.457	6344.689	5238.878	6505.458	7190.709	6746.589
Anapc1	685.7303	617.2585	850.9307	446.7304	583.2481	607.7861
Mrpl43	2621.367	2434.505	2461.823	2874.059	3204.043	2883.717
Xpo7	659.1478	702.8918	614.1038	429.252	639.1219	514.8187
Nmt1	3245.977	2752.552	2939.25	2779.031	3175.648	3093.897