

Meta-analysis of public hypoxic transcriptomes for cancer researches

癌研究のための公共データベースからの低酸素トランスクリプトームのメタ解析

Database Center for Life Science (DBCLS), Research Organization of Information and Systems
情報・システム研究機構 データサイエンス共同利用基盤施設 ライフサイエンス統合データベースセンター

Hidemasa Bono / 坊農 秀雅 <bono@dbcls.rois.ac.jp>



© 2017 DBCLS Togo Picture Gallery / CC-BY 4.0

In order to elucidate functional insights in hypoxic response in data-driven way, we have attempted meta-analysis of hypoxic transcriptome for public expression data which have been archived as microarray and RNA-seq data in public transcriptome databases. While various hypoxic conditions (oxygen concentration and duration of hypoxia) and cell lines are taken in the stored data, we manually curated possible pairs of transcriptome before and after hypoxic stress from microarray and RNA-seq data. The number of experiments for all genes was counted and those were classified into three categories, which are up-regulated, down-regulated, and unchanged. Genes up-regulated in all records contained well-studied hypoxia responsive genes, while down-regulated genes were unfamiliar in hypoxia. Meta-analysis approach to public gene expression database is useful for selecting candidate genes from gene expression profiles of various experimental conditions. The data produced in this study are open and ready for use to all cancer researchers.

Public databases

Nucleotide Sequence DB

- NCBI
- EBI ENA
- DDBJ

DBJ search (DBCLS SRA)

Make Index

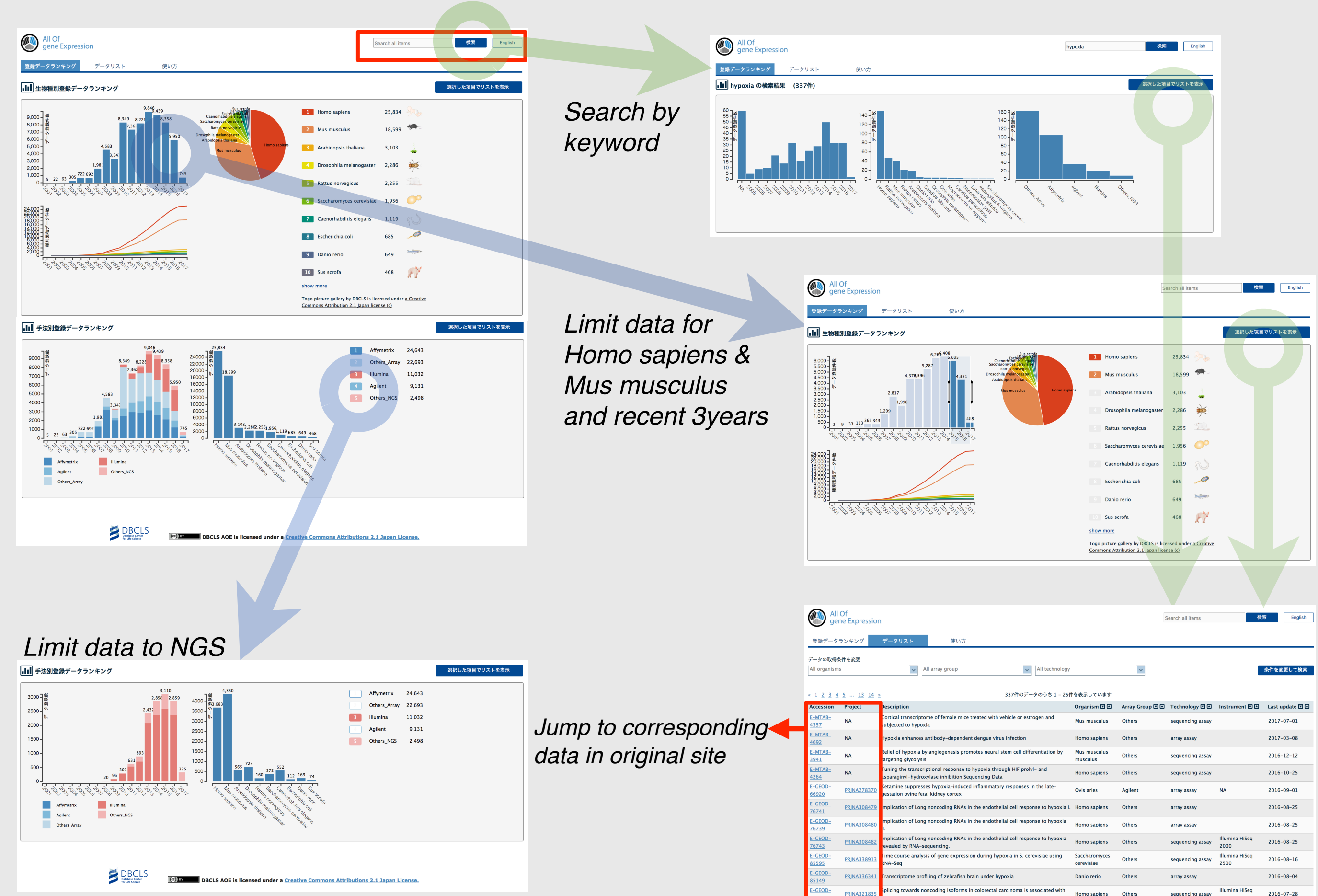
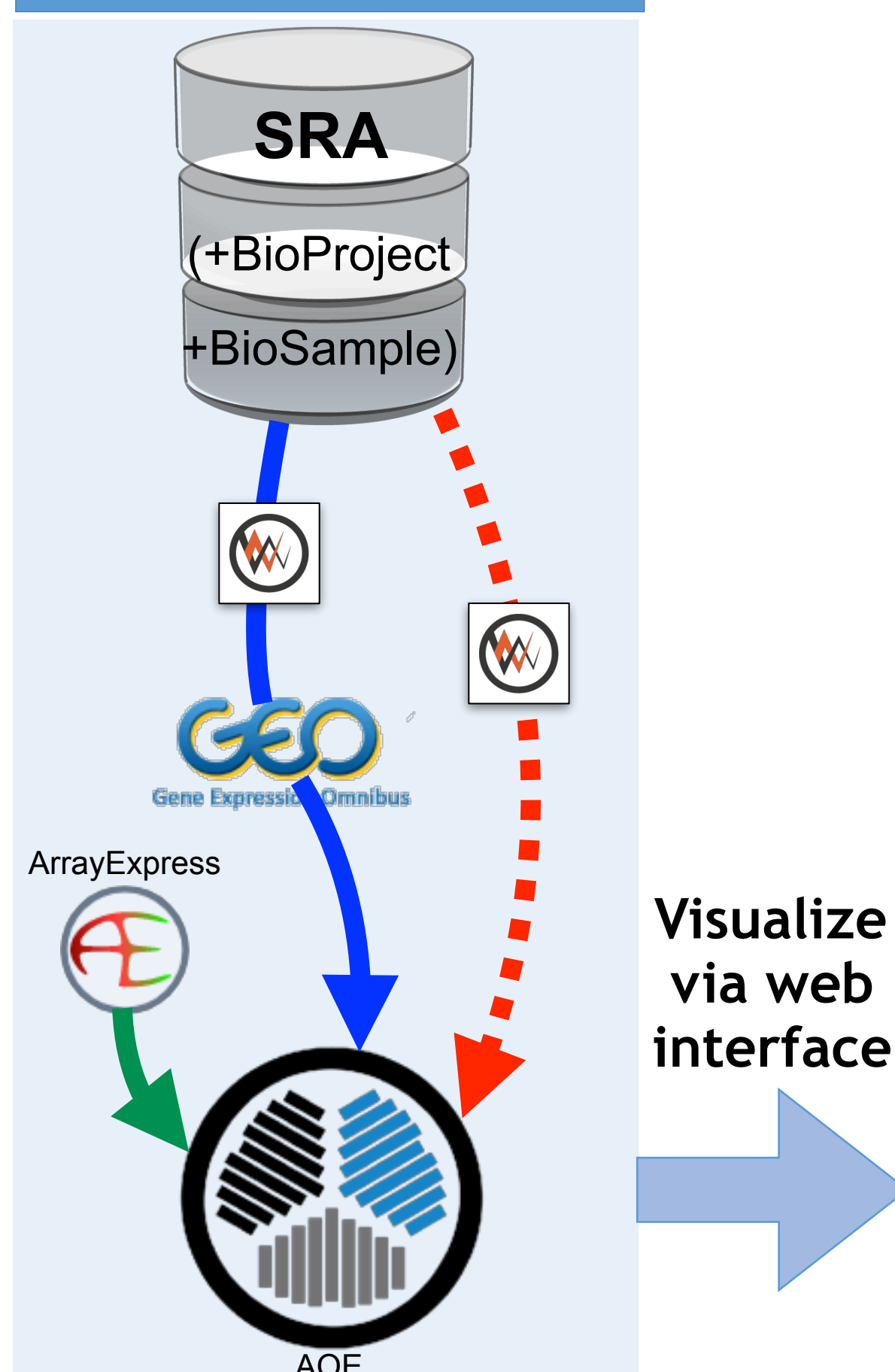
Gene Expression

Gene expression DB

- NCBI Gene Expression Omnibus (GEO)
- EBI ArrayExpress
- DDBJ Genomic Expression Archive (GEA)

All of gene expression (AOE) <http://aoe.dbcls.jp/>

How to make AOE dataset

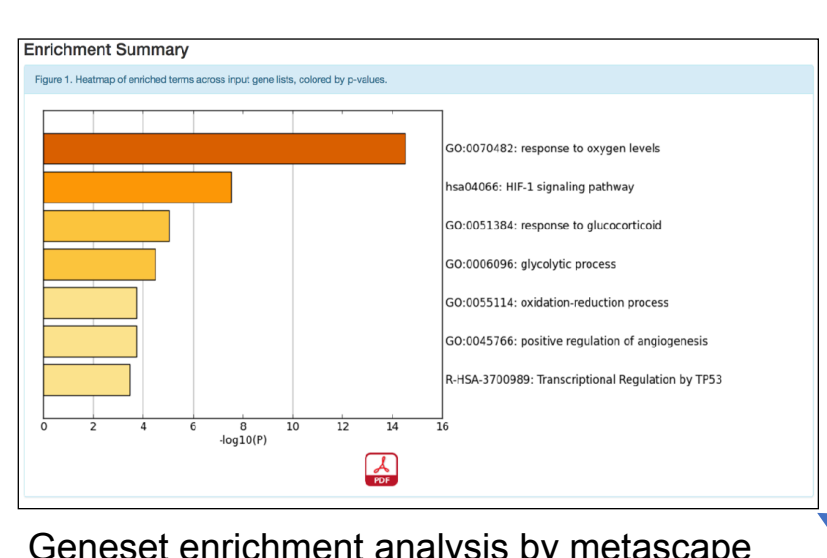


Meta-analysis of hypoxic transcriptomes in public databases

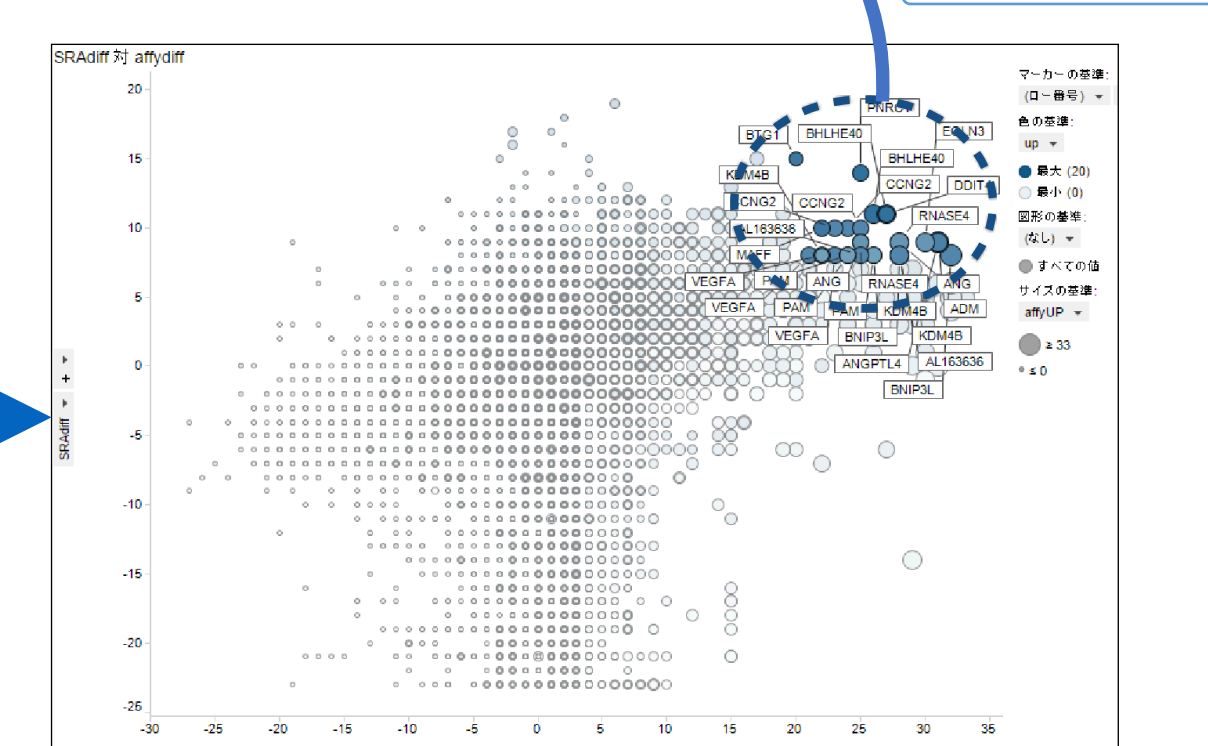
[1] Bono H *Biorxiv* doi: 10.1101/267310

curated hypoxia-normoxia pairs

human GeneChip	37 pairs from 11 data series
mouse GeneChip	53 pairs from 8 data series
Human RNA-seq	23 pairs from 7 data series



Data visualization & evaluation



Use case of results from the meta-analysis

Collaboration with Dr. Keiji Tanimoto at Hiroshima University
[2] Nakamura H *et al. PLoS One* 2018; 13(2):e0192136. doi: 10.1371/journal.pone.0192136

56 Table. Hypoxic down-regulation of DNA damage recognition and repair genes in cancer and non-cancer cells analyzed using the NCBI Gene Expression Omnibus database. <https://doi.org/10.1371/journal.pone.0192136.s006>

No.	Gene Symbol	Description
1	ATM	ataxia telangiectasia mutated
2	ATR	ataxia telangiectasia related
3	BRCA1	breast cancer 1
4	BRCA2	breast cancer 2
5	CHK1	checkpoint kinase 1
6	CHK2	checkpoint kinase 2
7	CTIP1	chromatin target 1
8	CTIP2	chromatin target 2
9	CTIP3	chromatin target 3
10	CTIP4	chromatin target 4
11	CTIP5	chromatin target 5
12	CTIP6	chromatin target 6
13	CTIP7	chromatin target 7
14	CTIP8	chromatin target 8
15	CTIP9	chromatin target 9
16	CTIP10	chromatin target 10
17	CTIP11	chromatin target 11
18	CTIP12	chromatin target 12
19	CTIP13	chromatin target 13
20	CTIP14	chromatin target 14
21	CTIP15	chromatin target 15
22	CTIP16	chromatin target 16
23	CTIP17	chromatin target 17
24	CTIP18	chromatin target 18
25	CTIP19	chromatin target 19
26	CTIP20	chromatin target 20
27	CTIP21	chromatin target 21
28	CTIP22	chromatin target 22
29	CTIP23	chromatin target 23
30	CTIP24	chromatin target 24
31	CTIP25	chromatin target 25
32	CTIP26	chromatin target 26
33	CTIP27	chromatin target 27
34	CTIP28	chromatin target 28
35	CTIP29	chromatin target 29
36	CTIP30	chromatin target 30
37	CTIP31	chromatin target 31
38	CTIP32	chromatin target 32
39	CTIP33	chromatin target 33
40	CTIP34	chromatin target 34
41	CTIP35	chromatin target 35
42	CTIP36	chromatin target 36
43	CTIP37	chromatin target 37
44	CTIP38	chromatin target 38
45	CTIP39	chromatin target 39
46	CTIP40	chromatin target 40
47	CTIP41	chromatin target 41
48	CTIP42	chromatin target 42
49	CTIP43	chromatin target 43
50	CTIP44	chromatin target 44
51	CTIP45	chromatin target 45
52	CTIP46	chromatin target 46
53	CTIP47	chromatin target 47
54	CTIP48	chromatin target 48
55	CTIP49	chromatin target 49
56	CTIP50	chromatin target 50

Application of meta-analyzed data

Secondary analysis of transcriptomes by RNA-seq

Example: Naked mole-rat (*Heterocephalus glaber*), so called 'deba'

GSE30337 (E-GEOD-30337)

Transcriptome sequencing of naked mole rat, *Heterocephalus glaber*

GSM (GEO Sample)	SRR (SRA Run ID)	description
GSM752147	SRR306394	liver pupil
GSM752148	SRR306395	liver 4 years
GSM752149	SRR306396	liver 20 years
GSM752150	SRR306397	kidney pupil
GSM752151	SRR306398	kidney 4 years
GSM752152	SRR306399	kidney 20 years
GSM752153	SRR306400	brain pupil
GSM752154	SRR306401	brain 4 years
GSM752155	SRR306402	brain 20 years
GSM752156	SRR306403	7 tissues mix
GSM752157	SRR306404	liver 4 years hypoxia
GSM752158	SRR306405	kidney 4 years hypoxia
GSM752159	SRR306406	brain 4 years hypoxia

Comparison

1. aging (longevity)

2. kidney 4 vs 20

3. brain 4 vs 20

4. hypoxia

1. liver 4years

2. kidney 4years

3. brain 4years

4. hypoxia

1. liver 4years

2. kidney 4years

3. brain 4years

4. hypoxia

1. liver 4years

2. kidney 4years

3. brain 4years

4. hypoxia

L_brainMat

L_kidneyMat

L_liverMat

HNbrainMat

HNkidneyMat

HNliverMat

Comparison to mouse needed!

ID conversion by Biomart
<http://www.ensembl.org/biomart/>



3 months vs 21 months

mouse

deba

Future work

- Further functional analyses of candidate genes
- More applications for secondary analysis of publicly available transcriptome data (RNA-seq)

RSEM (RNA-Seq by Expectation Maximization)

-> DEG Analysis using EBSeq

https://github.com/bli25ucb/RSEM_tutorial

Licensed under CC-BY 4.0 ©2018 HIDEMASA BONO (DBCLS)

