

COXPRESdb

a gene coexpression database for 7 animal species



Takeshi Obayashi
(Tohoku Univ, Japan)



Kengo Kinoshita
(Tohoku Univ, Japan)



<http://coxpresdb.jp/>

Summary

Background: Publicly available databases of coexpressed gene sets are a valuable resource for a wide variety of experimental studies, including gene targeting for functional identification, and for investigations of regulatory mechanisms.

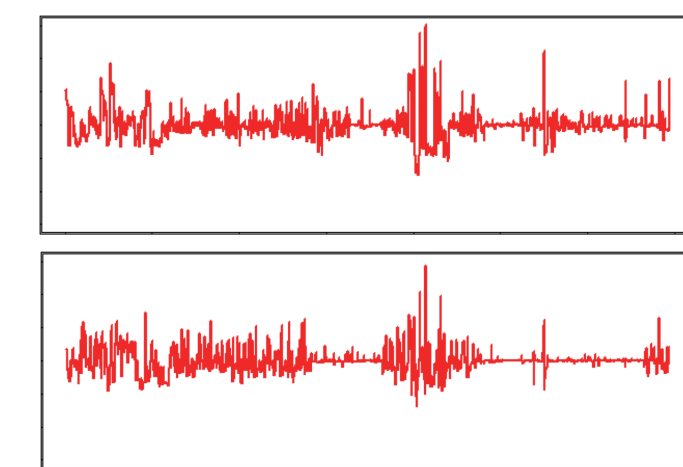
Results: We have constructed a new database named COXPRESdb for coexpressed gene lists and networks in 7 animal species; human, mouse, rat, chicken, zebrafish, fly and nematoda. The coexpression data were calculated from thousands of GeneChip data in ArrayExpress. Comparison of the coexpression data of multiple species is a key concept to enhance the reliability of gene coexpression, and the integration of different information can reduce the noise inherent in the information.

Preparation of gene coexpression data

- (1) Microarray data are downloaded from ArrayExpress (Table 1).
- (2) Normalize GeneChip data by RMA and gene centering.
- (3) Measure of coexpression
 - (3-1) Calculate Pearson correlation for any gene pairs (Figure 1)
 - (3-2) Calculate MR from PCC (Figure 2)

Figure 1:

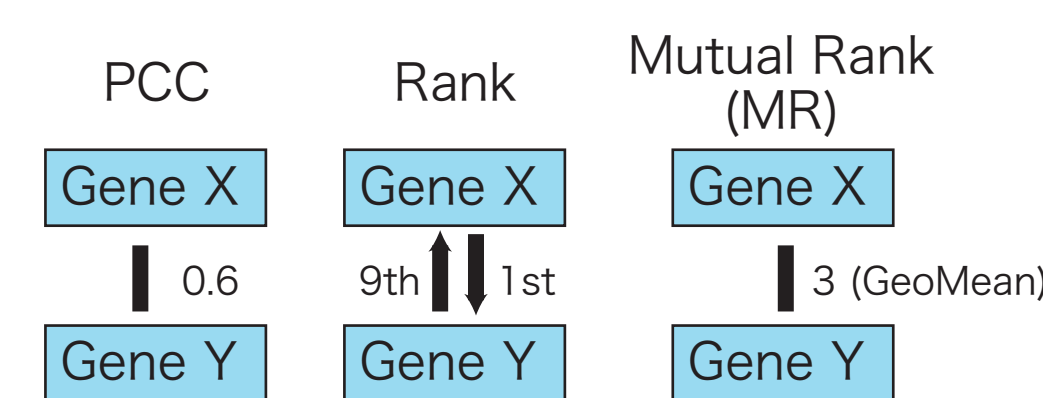
Calculation of coexpression as PCC between expression patterns of any 2 genes.



Many microarray samples

Figure 2:

PCC is transformed to MR.



Species	Version	Platform	Genes	Experiments	GeneChips
Human	c3.1	GPL570	19777 genes	123 experiments	4401 GeneChips
Mouse	c2.1	GPL1261	21036 genes	154 experiments	2226 GeneChips
Rat	c2.0	GPL1355	13751 genes	164 experiments	3526 GeneChips
Chicken	c1.0	GPL3213	13757 genes	22 experiments	352 GeneChips
Fly	c1.0	GPL1319	10112 genes	41 experiments	590 GeneChips
Zebrafish	c1.0	GPL1322	12626 genes	52 experiments	1022 GeneChips
Nematoda	c1.0	GPL200	17256 genes	35 experiments	514 GeneChips

Table 1: detail of data source to prepare gene coexpression data.

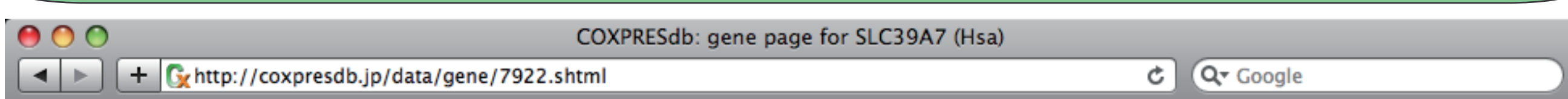
Usage of COXPRESdb

Gene Search



Simple search is available in this main search box.

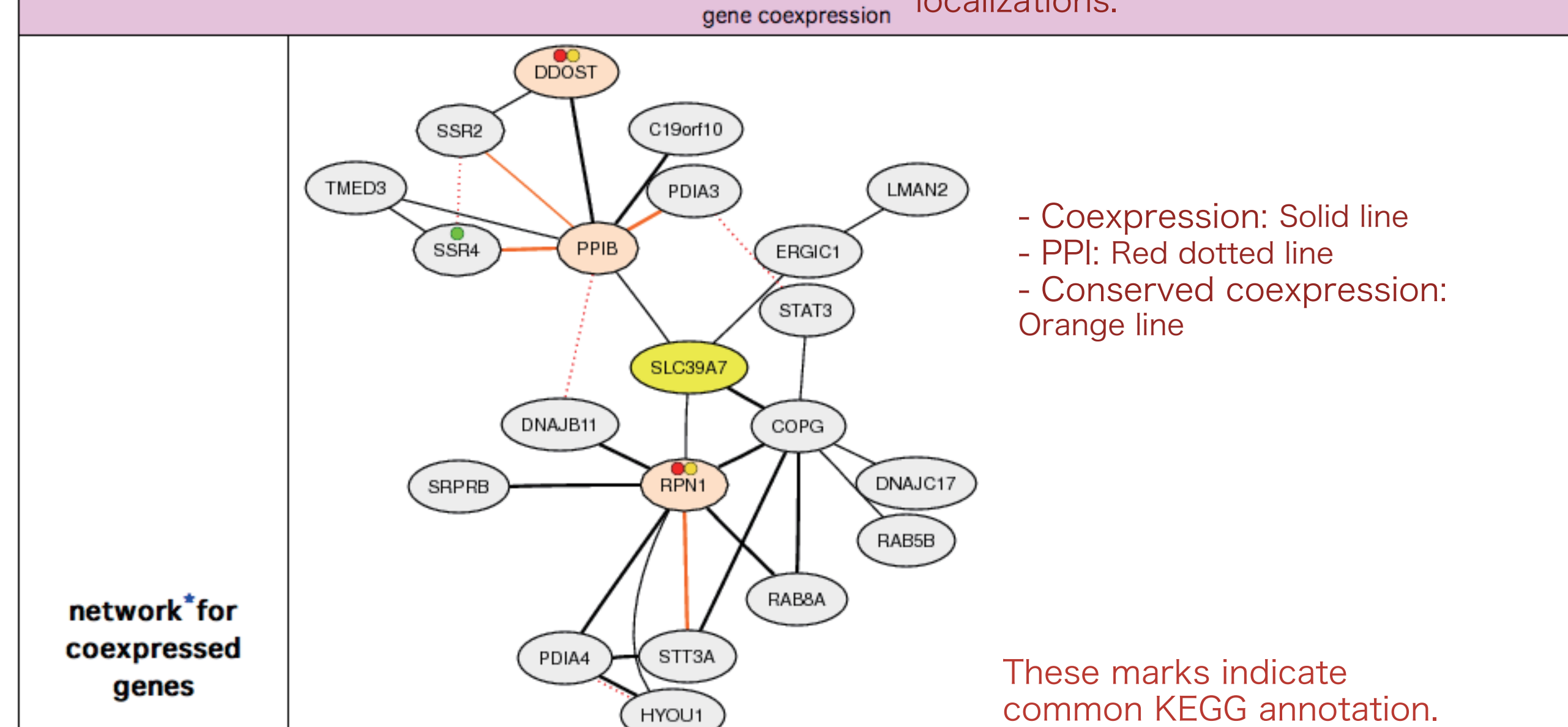
Gene Page



Hsa SLC39A7 Gene

functional annotation		
function*	solute carrier family 39 (zinc transporter), member 7	
GO BP*	GO:0006829 [list] [network] zinc ion transport (24 genes) IEA	
	GO:0006811 [list] [network] ion transport (717 genes) IEA	
GO CC*	GO:0005624 [list] [network] membrane fraction (585 genes) TAS	
	GO:0016021 [list] [network] integral to membrane (5266 genes) IEA	
	GO:0016020 [list] [network] membrane (7099 genes) IEA	
GO MF*	GO:0008270 [list] [network] zinc ion binding (2357 genes) IEA	
	GO:0005515 [list] [network] protein binding (7230 genes) IPI	
KEGG*		
orthologous	[ortholog page] Slc39a7 (Mmu) Slc39a7 (Rno)	
subcellular localization*	plas 6 (prediction for NP_001070984.1) plas 6 (prediction for NP_008910.2)	

We use WoLF PSORT to predict protein subcellular localizations.



network* for coexpressed genes

These marks indicate common KEGG annotation.

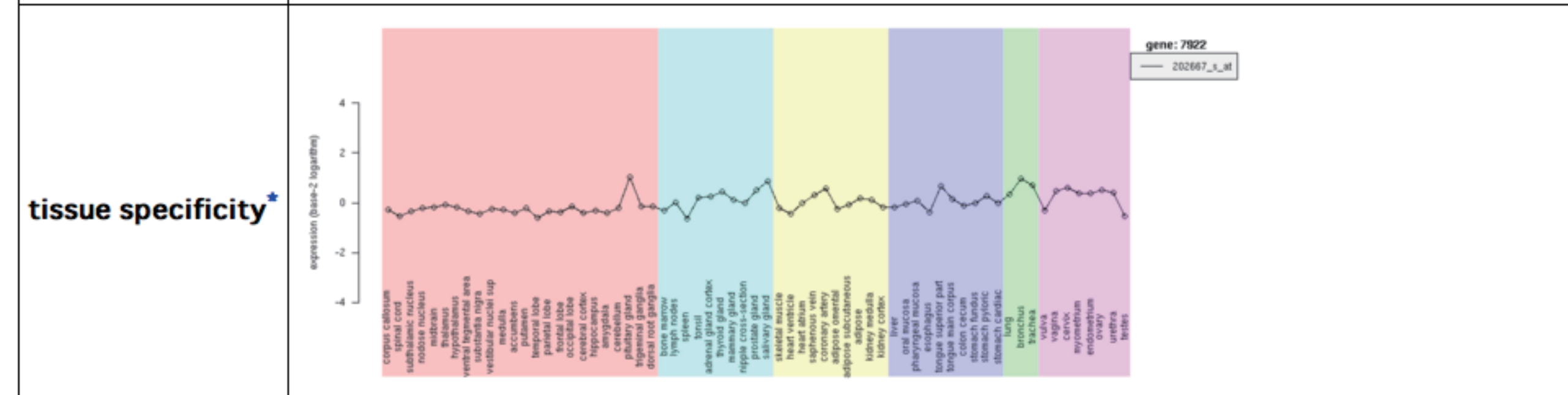
KEGG ID	Title	#genes	Link to the KEGG map (multiple genes)
hsa01030	Glycan structures - biosynthesis 1	2	
hsa00510	N-Glycan biosynthesis	2	
hsa05030	Amotrophic lateral sclerosis (ALS)	1	

Genes directly connected with SLC39A7 on the network

MR*	Cor*	symbol	function	coexpression detail	Entrez Gene ID
3.9	0.52	COPG	coatamer protein complex, subunit gamma	[detail]	22820
9.8	0.49	RPN1	ribophorin I	[detail]	6184
10.6	0.50	PIIB	peptidylprolyl isomerase B (cyclophilin B)	[detail]	5479
11.5	0.45	ERGIC1	endoplasmic reticulum-golgi intermediate compartment (ERGIC) 1	[detail]	57222

Sample detail is shown in this link.

coexpressed gene list	[coexpressed gene list for SLC39A7]	
all samples	[expression pattern for all samples]	



Link to other DBs		
Entrez Gene ID	7922	
HPRD ID	9028	
RGD ID	1344767	
refseq ID (mRNA)	NM_001077516.1 NM_006979.2	
refseq ID (protein)	NP_001070984.1 NP_008910.2	
GeneChip probe (GPL570)	202667_s_at	

Coexpressed gene list

Hsa: SLC39A7	
Orthologous gene group in HomoloGene**	
Mmu genes	Slc39a7 solute carrier family 39 (zinc transporter), member 7
Rno genes	Slc39a7 solute carrier family 39 (zinc transporter), member 7
Dre genes	30094
Dme genes	48805
Cel genes	181640

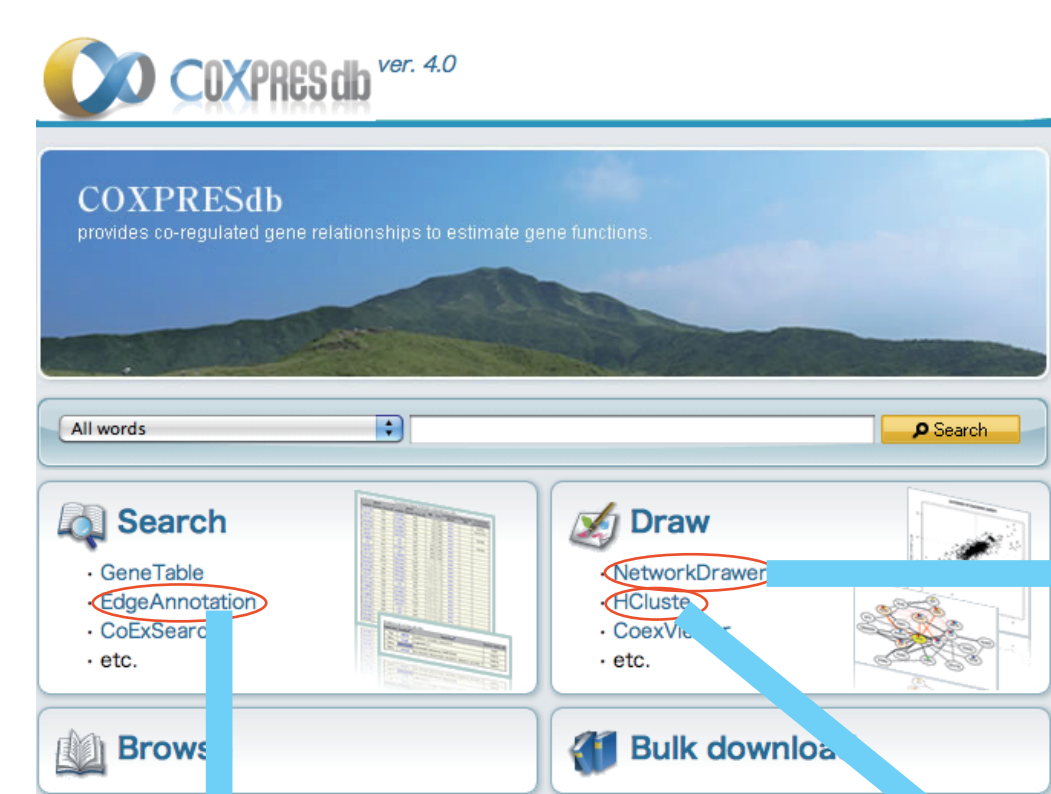
Top 300 coexpressed genes to SLC39A7 (Hsa c3.1 coexpression data)

	MR	Gene	Entrez Gene ID	Function	Mmu MR for SLC39A7 [list]	Rno MR for SLC39A7 [list]	Dre MR for 30094 [list]	Dme MR for 48805 [list]	Cel MR for 181640 [list]
0	0.0	SLC39A7	7922	solute carrier family 39 (zinc transporter), member 7	0.0	0.0	0.0	0.0	0.0
1	3.9	COPG	22820	coatamer protein complex, subunit gamma	38.7	125.5			
2	9.8	RPN1	6184	ribophorin I	11.2	22.4	68.6		3547.2
3	10.6	PIIB	5479	peptidylprolyl isomerase B (cyclophilin B)	529.6	13.6	24.4	861.1	8.9
4	11.5	ERGIC1	57222	endoplasmic reticulum-golgi intermediate compartment (ERGIC) 1	886.9	844.3			191.3
5	11.8	YIPF3	25844	Yip1 domain family, member 3	1322.0	1038.0	1047.5		11640.2
6	12.0	STT3A	3703	STT3, subunit of the oligosaccharyltransferase complex, homolog A (S. cerevisiae)	75.3	1300.7	49.2	1909.6	
7	26.7	YIF1A	10897	Yip1 interacting factor homolog A (S. cerevisiae)	480.8		623.3		3911.9
8	31.4	PFDN6	10471	prefoldin subunit 6	2028.3	3948.7	653.6	4657.3	16867.6
9	32.9	GMPPB	29925	GDP-mannose pyrophosphorylase B	32.7	127.2	1191.9	214.8	11204.9
10	32.9	HM13	81502	histocompatibility (minor) 13	360.0	4.0	139.3	32.6	3836.4
11	33.0	CT9orf10	56005	chromosome 19 open reading frame 10	157.9	65.9	8.9		
12	33.0	PDIA4	9601	protein disulfide isomerase family A, member 4	389.9	30.6	280.1		26.0
13	34.8	TMED3	23423	transmembrane emp24 protein transport domain containing 3	433.8	36.5	349.0		
14	41.8	TMED9	54732	transmembrane emp24 protein transport domain containing 9	131.2	52.8	667.1		
15	42.9	GMPPA	29926	GDP-mannose pyrophosphorylase A	187.6	252.1	1478.7	102.0	11689.1
16	46.1	ALDH18A1	5832	aldehyde dehydrogenase 18 family, member A1	257.5	641.1	3249.2	2132.9	23.2
17	47.4	SLC35B2	347734	solute carrier family 35, member B2	35.8	6.6	5502.3	5421.3	4513.9
18	48.4	PDIA3	2923	protein disulfide isomerase family A, member 3	223.6	25.0		1.7	952.6
19	54.2	DAP	1611	death-associated protein	4439.1	159.2	267.2		
20	55.3	MOGS	7841	mannosyl-oligosaccharide glucosidase	1003.9	17.7		5715.9	15829.4
21	58.5	SRPRB	58477	signal recognition particle receptor, B subunit	360.4	79.7	849.4	627.2	2891.8
22	61.2	HYOU1	10525	hypoxia up-regulated 1	1100.6	880.1	939.8	157.0	9270.5
23	63.3	SEC13	6396	SEC13 homolog (S. cerevisiae)	251.5	354.6	122.7	272.9	2417.5
24	64.1	YIPF2	78992	Yip1 domain family, member 2	7729.2	9099.4			
25	65.0	LMAN2	10960	lectin, mannose-binding 2	244.1		2143.3	86.7	140.5
26	65.0	SEC61A1	29927	Sec61 alpha 1 subunit (S. cerevisiae)	150.9	9.9	138.1	974.0	
27	65.1	C6orf89	221477	chromosome 6 open reading frame 89	3391.7	1257.4			
28	68.1	KDEL1	10945	KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 1	1288.2	5.7			18.8
29	69.3	LSM2	57819	LSM2 homolog, U6 small nuclear RNA associated (S. cerevisiae)	780.9		1346.5	4546.0	15769.2
30	70.0	SDF2L1	23753	stromal cell-derived factor 2-like 1	278.4	408.1			

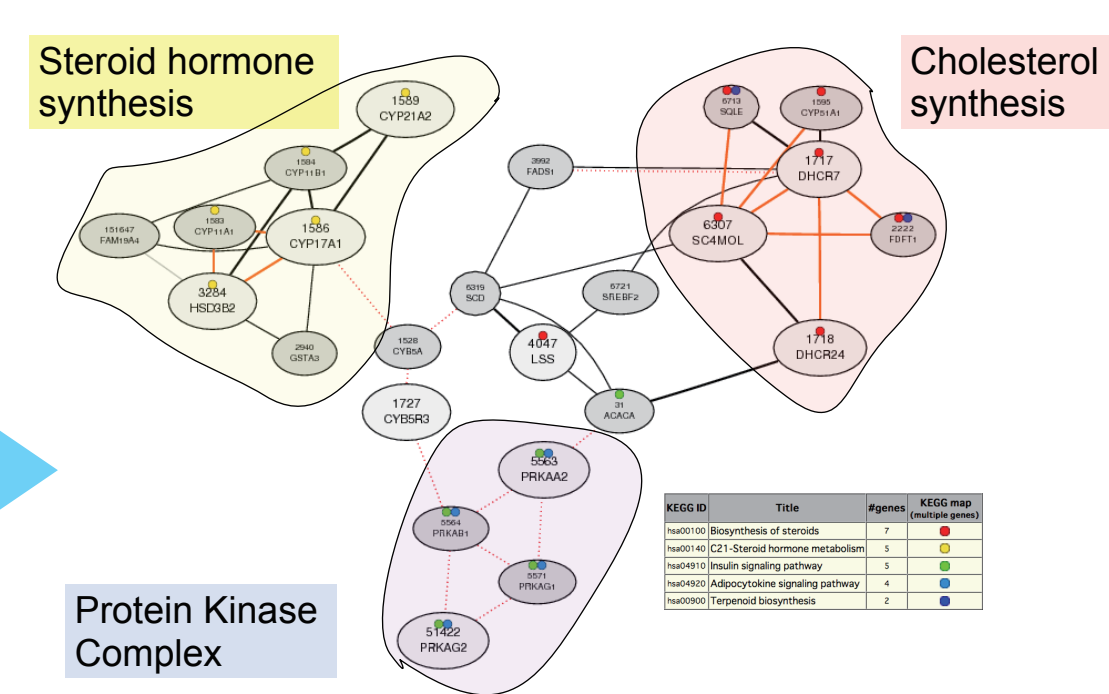
Coexpressed gene list from SLC39A7 in Human.

Coexpression in other species. (Smaller MR indicates stronger coexpression)

Tools for multiple gene query



NetworkDrawer



EdgeAnnotation

symbol	function	Entrez Gene ID
CDT1	chromosome P450, family 17, subfamily A, polypeptide 1	1308
CDT2	chromosome P450, family 21, subfamily A, polypeptide 2	1312
CDT3	4-dehydrocholesterol reductase	1317
CDT4	4-dehydrocholesterol reductase	1318
CDT5	4-dehydrocholesterol reductase	1319
CDT6	4-dehydrocholesterol reductase	1320
CDT7	4-dehydrocholesterol reductase	1321
CDT8	4-dehydrocholesterol reductase	1322
CDT9	4-dehydrocholesterol reductase	1323
CDT10	4-dehydrocholesterol reductase	1324
CDT11	4-dehydrocholesterol reductase	1325
CDT12	4-dehydrocholesterol reductase	1326
CDT13	4-dehydrocholesterol reductase	1327
CDT14	4-dehydrocholesterol reductase	1328
CDT15	4-dehydrocholesterol reductase	1329
CDT16	4-dehydrocholesterol reductase	1330
CDT17	4-dehydrocholesterol reductase	1331
CDT18	4-dehydrocholesterol reductase	1332
CDT19	4-dehydrocholesterol reductase	1333
CDT20	4-dehydrocholesterol reductase	1334
CDT21	4-dehydrocholesterol reductase	1335
CDT22	4-dehydrocholesterol reductase	1336
CDT23	4-dehydrocholesterol reductase	1337
CDT24	4-dehydrocholesterol reductase	1338
CDT25	4-dehydrocholesterol reductase	1339
CDT26	4-dehydrocholesterol reductase	1340
CDT27	4-dehydrocholesterol reductase	1341
CDT28	4-dehydrocholesterol reductase	1342
CDT29	4-dehydrocholesterol reductase	1343
CDT30	4-dehydrocholesterol reductase	1344
CDT31	4-dehydrocholesterol reductase	1345
CDT32	4-dehydrocholesterol reductase	1346
CDT33	4-dehydrocholesterol reductase	1347
CDT34	4-dehydrocholesterol reductase	1348
CDT35	4-dehydrocholesterol reductase	1349
CDT36	4-dehydrocholesterol reductase	1350
CDT37	4-dehydrocholesterol reductase	1351
CDT38	4-dehydrocholesterol reductase	1352
CDT39	4-dehydrocholesterol reductase	1353
CDT40	4-dehydrocholesterol reductase	1354
CDT41	4-dehydrocholesterol reductase	1355
CDT42	4-dehydrocholesterol reductase	1356
CDT43	4-dehydrocholesterol reductase	1357
CDT44	4-dehydrocholesterol reductase	1358
CDT45	4-dehydrocholesterol reductase	1359
CDT46	4-dehydrocholesterol reductase	1360
CDT47	4-dehydrocholesterol reductase	1361
CDT48	4-dehydrocholesterol reductase	1362
CDT49	4-dehydrocholesterol reductase	1363
CDT50	4-dehydrocholesterol reductase	1364
CDT51	4-dehydrocholesterol reductase	1365
CDT52	4-dehydrocholesterol reductase	1366
CDT53	4-dehydrocholesterol reductase	1367
CDT54	4-dehydrocholesterol reductase	1368
CDT55	4-dehydrocholesterol reductase	1369
CDT56	4-dehydrocholesterol reductase	1370
CDT57	4-dehydrocholesterol reductase	1371
CDT58	4-dehydrocholesterol reductase	1372
CDT59	4-dehydrocholesterol reductase	1373
CDT60	4-dehydrocholesterol reductase	1374
CDT61	4-dehydrocholesterol reductase	1375
CDT62	4-dehydrocholesterol reductase	1376
CDT63	4-dehydrocholesterol reductase	1377
CDT64	4-dehydrocholesterol reductase	1378
CDT65	4-dehydrocholesterol reductase	1379
CDT66	4-dehydrocholesterol reductase	1380
CDT67	4-dehydrocholesterol reductase	1381
CDT68	4-dehydrocholesterol reductase	1382
CDT69	4-dehydrocholesterol reductase	1383
CDT70	4-dehydrocholesterol reductase	1384
CDT71	4-dehydrocholesterol reductase	1385
CDT72	4-dehydrocholesterol reductase	1386
CDT73	4-dehydrocholesterol reductase	1387
CDT74	4-dehydrocholesterol reductase	1388
CDT75	4-dehydrocholesterol reductase	1389
CDT76	4-dehydrocholesterol reductase	1390
CDT77	4-dehydrocholesterol reductase	1391
CDT78	4-dehydrocholesterol reductase	1392
CDT79	4-dehydrocholesterol reductase	1393
CDT80	4-dehydrocholesterol reductase	1394
CDT81	4-dehydrocholesterol reductase	1395
CDT82	4-dehydrocholesterol reductase	1396
CDT83	4-dehydrocholesterol reductase	1397
CDT84	4-dehydrocholesterol reductase	1398
CDT85	4-dehydrocholesterol reductase	1399
CDT86	4-dehydrocholesterol reductase	1400
CDT87	4-dehydrocholesterol reductase	1401
CDT88	4-dehydrocholesterol reductase	1402
CDT89	4-dehydrocholesterol reductase	1403
CDT90	4-dehydrocholesterol reductase	1404
CDT91	4-dehydrocholesterol reductase	1405
CDT92	4-dehydrocholesterol reductase	1406
CDT93	4-dehydrocholesterol reductase	1407
CDT94	4-dehydrocholesterol reductase	1408
CDT95	4-dehydrocholesterol reductase	1409
CDT96	4-dehydrocholesterol reductase	1410
CDT97	4-dehydrocholesterol reductase	1411
CDT98	4-dehydrocholesterol reductase	1412
CDT99	4-dehydrocholesterol reductase	1413
CDT100	4-dehydrocholesterol reductase	1414

HCluster

