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.
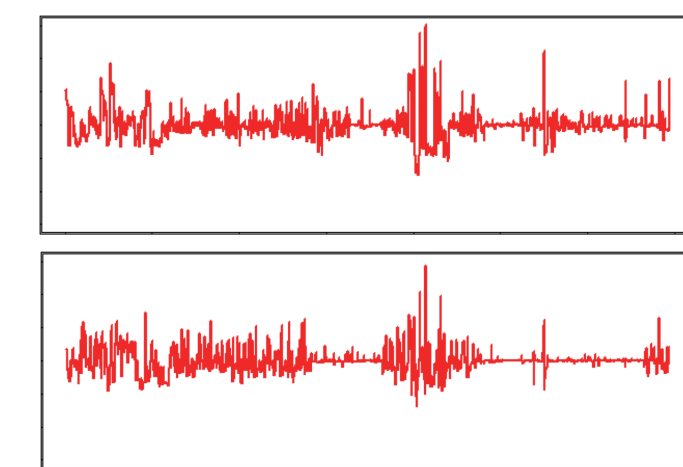
Results: We have constructed a new database named COXPRESdb for coexpressed gene lists and networks in 7 animal species; human, mouse, rat, chicken, zebrafish, fly and nematoda. The coexpression data were calculated from thousands of GeneChip data in ArrayExpress. Comparison of the coexpression data 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)
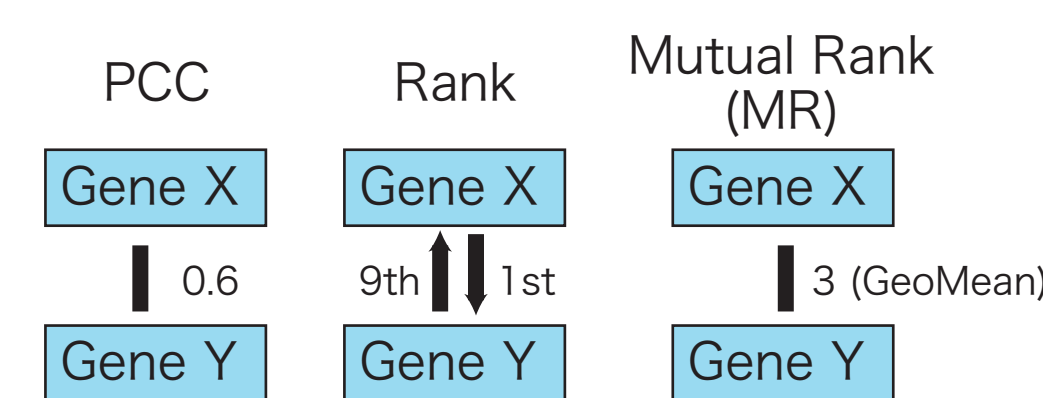
   (3-2) Calculate MR from PCC (Figure 2)

Figure 1: Calculation of coexpression as PCC between expression patterns of any 2 genes.

Many microarray samples

Figure 2: PCC is transformed to MR.

	PCC	Rank	Mutual Rank (MR)
Gene X		Gene X	Gene X
	0.6	9th / 1st	3 (GeoMean)
Gene Y		Gene Y	Gene Y

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

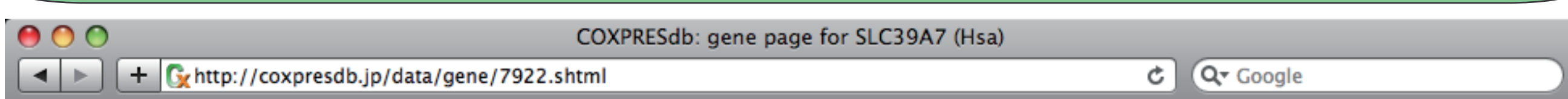Table 1: detail of data source to prepare gene coexpression data.

Usage of COXPRESdb

Gene Search

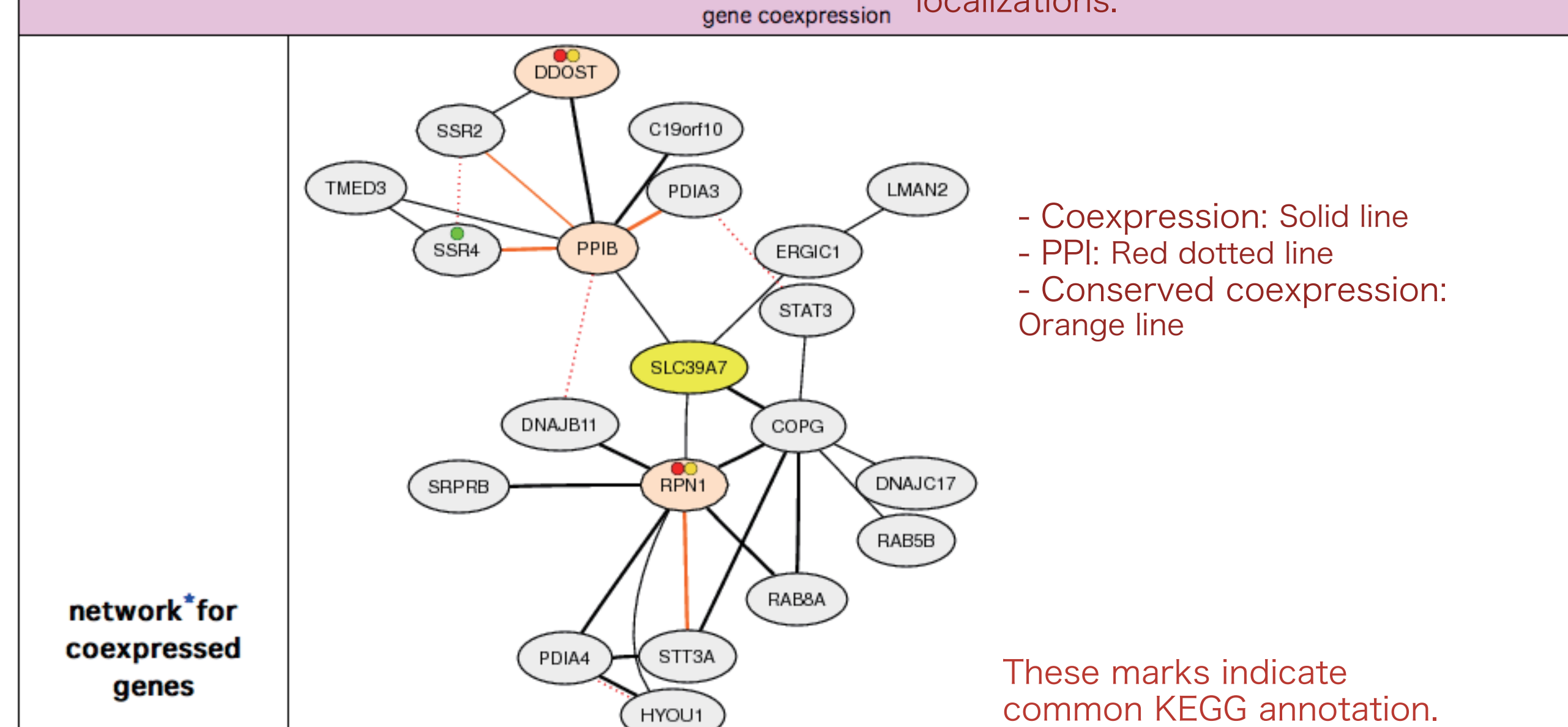Simple search is available in this main search box.

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.

gene coexpression

- Coexpression: Solid line
- PPI: Red dotted line
- Conserved coexpression: Orange line

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	Amyotrophic lateral sclerosis (ALS)	1	

Sample detail is shown in this link.

Genes directly connected with SLC39A7 on the network

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

coexpressed gene list	[coexpressed gene list for SLC39A7]

gene expression

all samples	[expression pattern for all samples]

tissue specificity*	

gene: 7922
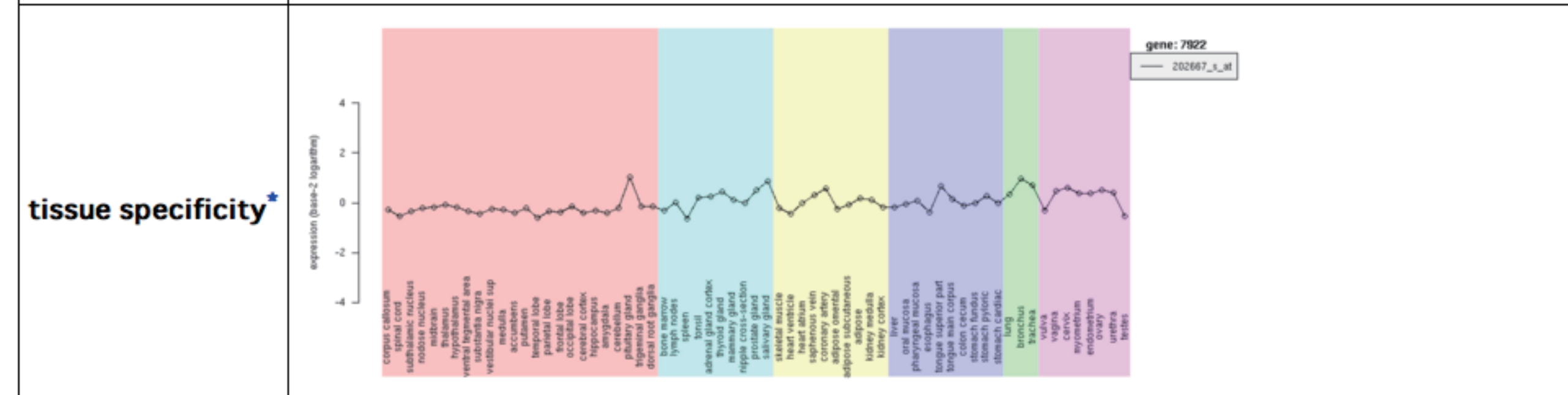
Link to other DBs

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

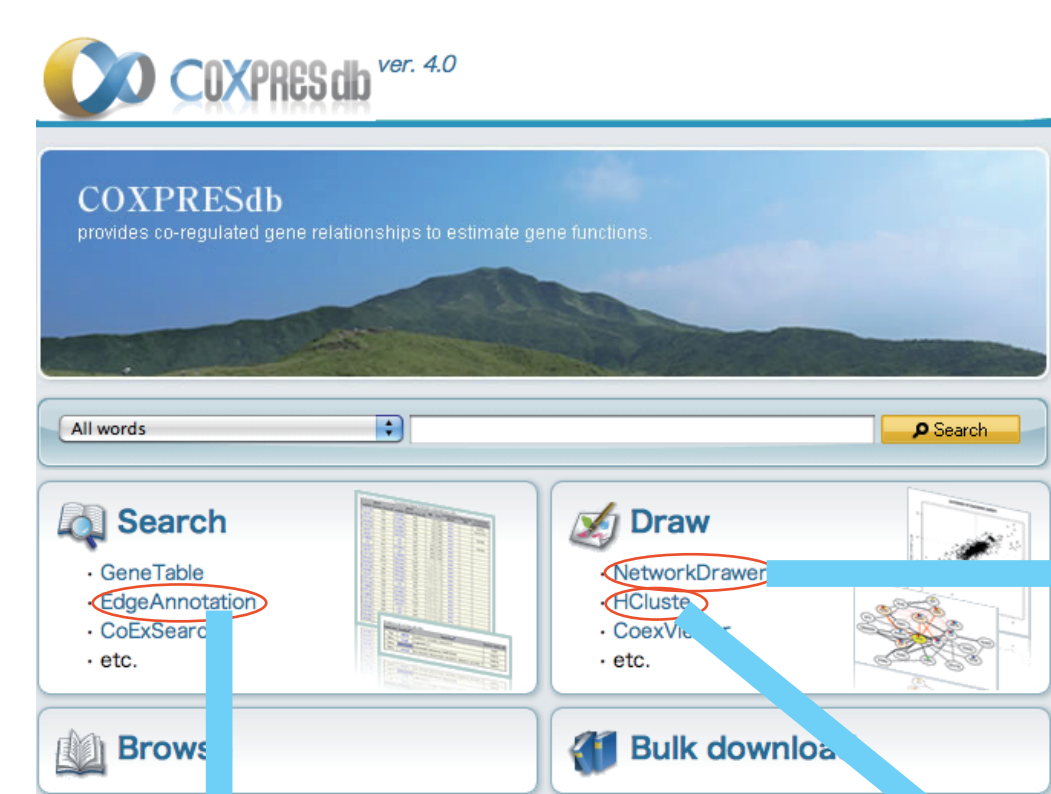| Orthologous gene group in HomoloGene**		
Mmu genes	Slc39a7	solute carrier family 39 (zinc transporter), member 7
Rno genes	Slc39a7	solute carrier family 39 (zinc transporter), member 7
Dre genes	30094	
Dme genes	48805	
Cel genes	181640	

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	coatomer 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	PPIB	5479	peptidylprolyl isomerase B (cyclophilin B)	529.6	13.6	24.4	861.1	8.9 / 17.3
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	C19orf10	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 / 1218.6	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 / 469.6
23	63.3	SEC13	6396	SEC13 homolog (S. cerevisiae)	251.5	354.6	122.7	272.9	2417.5
24	65.1	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	KDELR1	10945	KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 1	1288.2	5.7			18.8 / 2718.1
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			

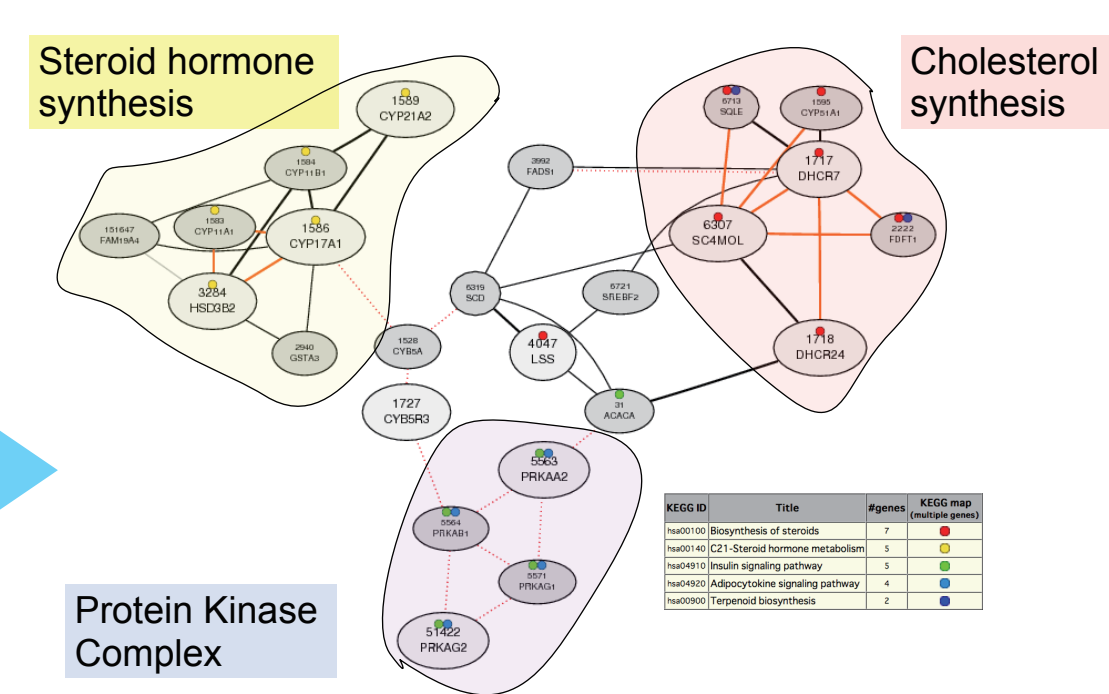Without supports by other species

Coexpressed gene list from SLC39A7 in Human.

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

Tools for multiple gene query

NetworkDrawer

Steroid hormone synthesis

Cholesterol synthesis

Protein Kinase Complex

EdgeAnnotation

HCluster