of knowledgebase containing approximately 2,500,000 biological and chemical interactions and functional annotations with definite scientific evidence, curated by expert biologists. By uploading the list of Entrez Gene IDs, the network-generation algorithm identifies focused genes integrated in a global molecular network.

Molecular Network: Molecular network represents the cell-wide map of complex molecular interactions extracted from high-throughput data of the genome, transcriptome and proteome illustrated with the aid of the literature-based knowledgebase of molecular pathways. The logically arranged molecular networks construct the whole system characterized by robustness that maintains the proper function of the system in the face of genetic and environmental perturbations.

Systems Biology: Systems biology is a research field to study the whole system of living organism based on the experimentally and computationally integrated molecular networks by using simulated biological models.

TDP-43: TAR DNA-binding protein-43 (TDP-43) is a nuclear protein, capable of interacting with UG/TG repeat stretches of target RNAs/DNAs, plays a key role in regulation of transcription, alternative splicing, mRNA stability and transport, and microRNA biogenesis. Abnormally phosphorylated, ubiquitinated, and aggregated TDP-43 proteins constitute a principal component of neuronal and glial cytoplasmic and nuclear inclusions in the brains of ALS and FTLD.

Table 1. Top 30 Gene Ontology Categories Enriched in the Set of TDP-43 Target RNAs and Interacting Proteins

Set	Rat TDP-43 Target RNAs			Human TDP-43 Interacting Proteins			The Integration of Human TDP-43 Target RNAs and Interacting Proteins		
Rank	GO Term	Number of Genes	p-Value after Bonferroni Correction	GO Term	Number of Genes	p-Value after Bonferroni Correction	GO Term	Number of Genes	p-Value after Bonferroni Correction
1	GO:0000166~nucleotide binding	624	7.846E-29	GO:0030529~ribonucleoprotein complex	109	9.329E-101	GO:0003723~RNA binding	333	3.995E-38
2	GO:0008104~protein localization	269	1.403E-25	GO:0003723~RNA binding	122	2.460E-98	GO:0005829~cytosol	511	5.094E-34
3	GO:0005829~cytosol	401	1.942E-25	GO:0006414~translational elongation	49	1.252E-60	GO:0000166~nucleotide binding	756	1.365E-27
4	GO:0046907~intracellular transport	214	8.087E-24	GO:0005840~ribosome	58	8.007E-55	GO:0046907~intracellular transport	286	8.787E-26
5	GO:0045184~establishment of protein localization	230	1.979E-22	GO:0006412~translation	64	4.545E-52	GO:0070013~intracellular organelle lumen	610	5.192E-25
6	GO:0015031~protein transport	228	3.148E-22	GO:0033279~ribosomal subunit	46	8.545E-49	GO:0031981~nuclear lumen	513	8.699E-24
7	GO:0016192~vesicle-mediated transport	193	5.554E-22	GO:0022626~cytosolic ribosome	40	1.429E-48	GO:0043233~organelle lumen	615	2.350E-23
8	GO:0045202~synapse	172	3.155E-21	GO:0003735~structural constituent of ribosome	50	2.898E-47	GO:0031974~membrane-enclosed lumen	624	4.034E-23
9	GO:0032553~ribonucleotide binding	494	8.206E-21	GO:0006396~RNA processing	67	1.579E-41	GO:0030529~ribonucleoprotein complex	228	1.356E-22
10	GO:0032555~purine ribonucleotide binding	493	1.296E-20	GO:0043228~non-membrane-bounded organelle	120	8.289E-36	GO:0006396~RNA processing	242	1.717E-22
11	GO:0043005~neuron projection	188	3.111E-20	GO:0043232~intracellular non-membrane-bounded organelle	120	8.289E-36	GO:0045184~establishment of protein localization	312	4.403E-22

12	GO:0070013~intracellular organelle lumen	399	9.446E-20	GO:0044445~cytosolic part	40	8.788E-36	GO:0016071~mRNA metabolic process	181	4.404E-22
13	GO:0031981~nuclear lumen	328	5.256E-19	GO:0005198~structural molecule activity	61	1.009E-28	GO:0015031~protein transport	309	8.598E-22
14	GO:0017076~purine nucleotide binding	506	6.240E-19	GO:0016071~mRNA metabolic process	46	5.746E-27	GO:0008104~protein localization	345	1.927E-21
15	GO:0043233~organelle lumen	404	6.257E-18	GO:0015934~large ribosomal subunit	26	6.441E-27	GO:0006397~mRNA processing	162	2.830E-21
16	GO:0031974~membrane-enclosed lumen	413	6.353E-18	GO:0022625~cytosolic large ribosomal subunit	21	4.693E-25	GO:0008380~RNA splicing	144	7.795E-19
17	GO:0019899~enzyme binding	195	3.107E-17	GO:0006397~mRNA processing	40	6.246E-23	GO:0032555~purine ribonucleotide binding	599	6.424E-17
18	GO:0000267~cell fraction	333	3.419E-17	GO:0005730~nucleolus	53	5.630E-21	GO:0032553~ribonucleotide binding	599	6.424E-17
19	GO:0042995~cell projection	265	9.157E-17	GO:0070013~intracellular organelle lumen	82	1.455E-20	GO:0043232~intracellular non-membrane-bounded organelle	792	1.406E-16
20	GO:0005626~insoluble fraction	271	3.683E-16	GO:0043233~organelle lumen	82	6.435E-20	GO:0043228~non-membrane-bounded organelle	792	1.406E-16
21	GO:0016071~mRNA metabolic process	116	2.963E-15	GO:0015935~small ribosomal subunit	21	1.353E-19	GO:0005730~nucleolus	269	1.642E-16
22	GO:0019904~protein domain specific binding	144	7.230E-15	GO:0008380~RNA splicing	35	1.952E-19	GO:0070727~cellular macromolecule localization	182	8.135E-16
23	GO:0006397~mRNA processing	104	8.461E-15	GO:0031974~membrane-enclosed lumen	82	2.288E-19	GO:0017076~purine nucleotide binding	615	1.439E-15
24	GO:0003723~RNA binding	175	1.066E-14	GO:0005829~cytosol	69	3.767E-19	GO:0045202~synapse	158	2.971E-15
25	GO:0005624~membrane fraction	255	1.423E-14	GO:0031981~nuclear lumen	72	3.995E-19	GO:0034613~cellular protein localization	179	5.267E-15
26	GO:0030163~protein catabolic process	142	7.059E-14	GO:0022627~cytosolic small ribosomal subunit	18	4.886E-19	GO:0016192~vesicle-mediated transport	229	3.924E-14

27	GO:0016023~cytoplasmic membrane-bounded vesicle	196	3.197E-13	GO:0022613~ribonucleoprotein complex biogenesis	28	1.469E-17	GO:0006886~intracellular protein transport	164	7.429E-14
28	GO:0043632~modification-dependent macromolecule catabolic process	125	5.289E-13	GO:0003729~mRNA binding	20	4.647E-17	GO:0030163~protein catabolic process	240	1.039E-12
29	GO:0019941~modification-dependent protein catabolic process	125	5.289E-13	GO:0042254~ribosome biogenesis	23	1.015E-15	GO:0051603~proteolysis involved in cellular protein catabolic process	231	4.158E-12
30	GO:0044456~synapse part	115	1.839E-12	GO:0000377~RNA splicing, via transesterification reactions with bulged adenosine as nucleophile	23	1.867E-13	GO:0043632~modification-dependent macromolecule catabolic process	223	4.158E-12

Entrez Gene IDs of 4,163 rat TDP-43 target RNAs (Sephton et al., 2011), 227 human interacting proteins (Freibaum et al., 2010), and the integration of both 4,063 human TDP-43 target RNAs and 227 human interacting proteins were imported into the Functional Annotation tool of DAVID. Top 30 GO categories enriched in the gene set are listed.

Table 2. The Set of 106 Genes Concurrently Serving as Both Human TDP-43 Target RNAs and Interacting Proteins

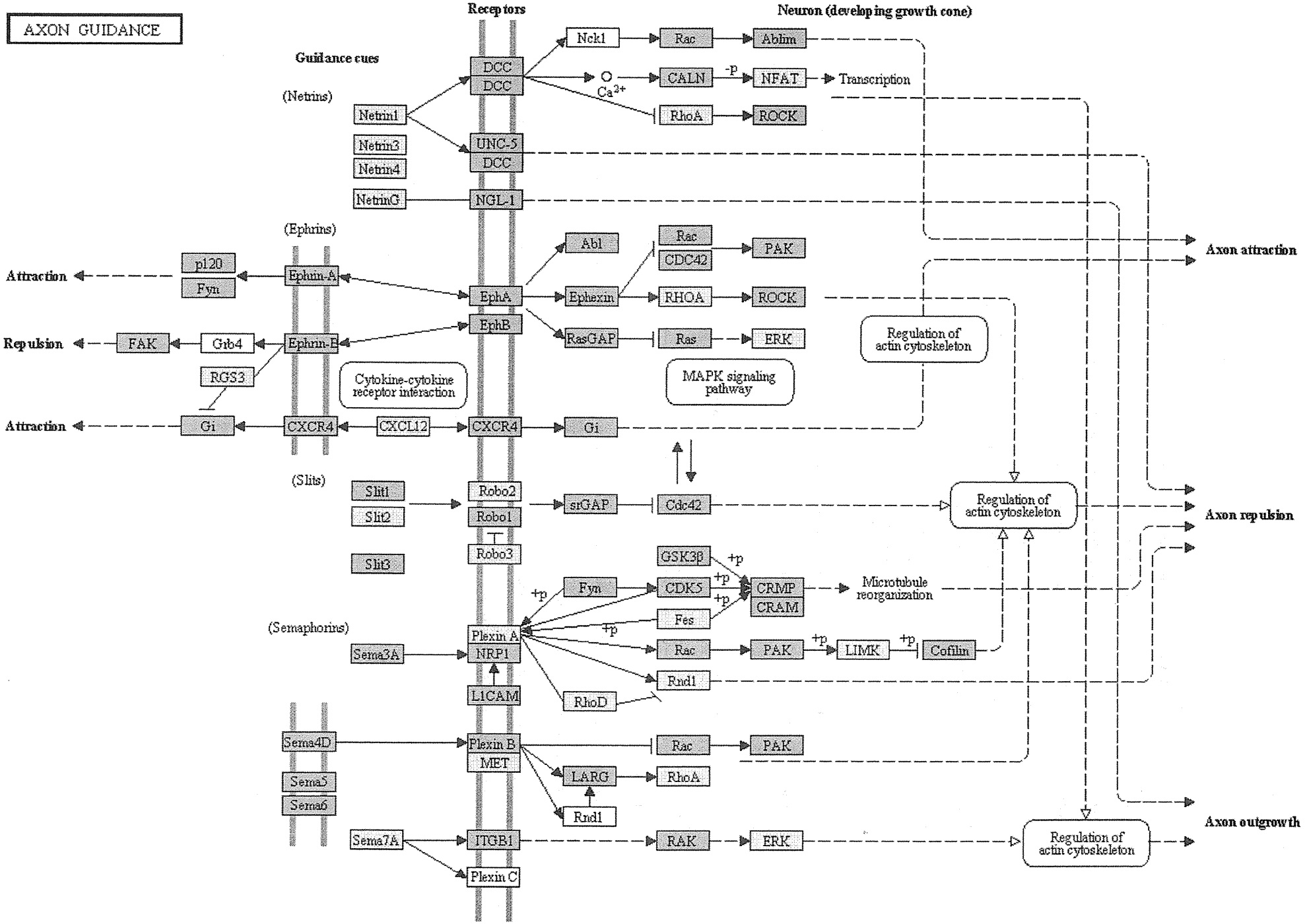
Entrez Gene ID	Gene Symbol	Gene Name
103	ADAR	adenosine deaminase, RNA-specific
444	ASPH	aspartate beta-hydroxylase
11273	ATXN2L	ataxin 2-like
51637	C14orf166	chromosome 14 open reading frame 166
81627	C1orf25	chromosome 1 open reading frame 25
4076	CAPRIN1	cell cycle associated protein 1
988	CDC5L	CDC5 cell division cycle 5-like (S. pombe)
10658	CELF1	CUG triplet repeat, RNA binding protein 1
8531	CSDA	cold shock domain protein A
7818	DAP3	death associated protein 3
10521	DDX17	DEAD (Asp-Glu-Ala-Asp) box polypeptide 17
9188	DDX21	DEAD (Asp-Glu-Ala-Asp) box polypeptide 21
1654	DDX3X	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, X-linked
1655	DDX5	DEAD (Asp-Glu-Ala-Asp) box polypeptide 5
79009	DDX50	DEAD (Asp-Glu-Ala-Asp) box polypeptide 50
1656	DDX6	DEAD (Asp-Glu-Ala-Asp) box polypeptide 6
22907	DHX30	DEAH (Asp-Glu-Ala-His) box polypeptide 30
170506	DHX36	DEAH (Asp-Glu-Ala-His) box polypeptide 36
1660	DHX9	DEAH (Asp-Glu-Ala-His) box polypeptide 9
27292	DIMT1L	DIM1 dimethyladenosine transferase 1-like (S. cerevisiae)
10049	DNAJB6	DnaJ (Hsp40) homolog, subfamily B, member 6
1937	EEF1G	eukaryotic translation elongation factor 1 gamma
27161	EIF2C2	eukaryotic translation initiation factor 2C, 2
8661	EIF3A	eukaryotic translation initiation factor 3, subunit A
8662	EIF3B	eukaryotic translation initiation factor 3, subunit B
8672	EIF4G3	eukaryotic translation initiation factor 4 gamma, 3
1993	ELAVL2	ELAV (embryonic lethal, abnormal vision, Drosophila)-like 2 (Hu antigen B)

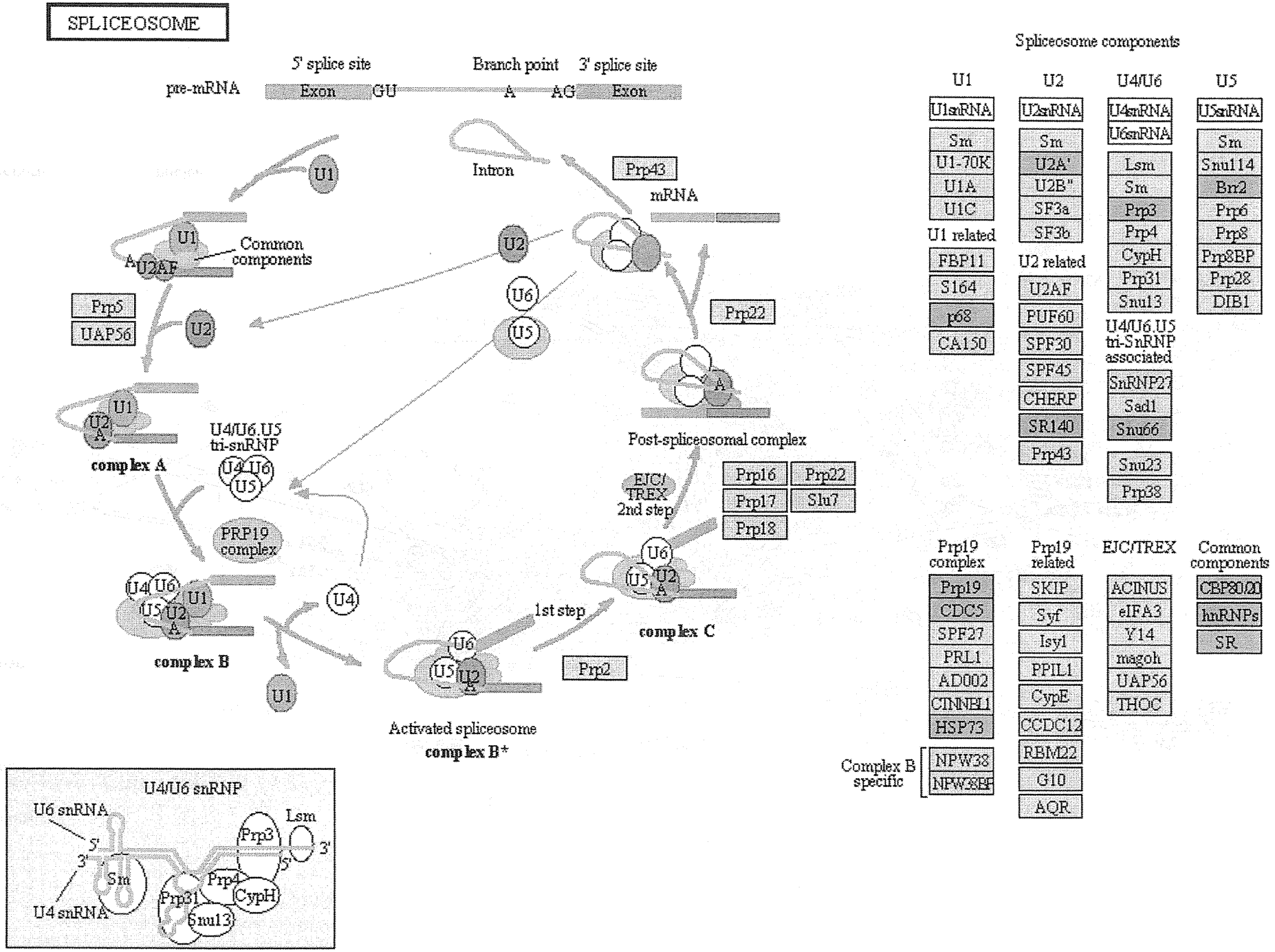
2091	FBL	fibrillarin
1968	FIF2SC	eukaryotic translation initiation factor 2, subunit 3 gamma, 52kDa
10146	G3BP1	GTPase activating protein (SH3 domain) binding protein 1
9908	G3BP2	GTPase activating protein (SH3 domain) binding protein 2
26354	GNL3	guanine nucleotide binding protein-like 3 (nucleolar)
2926	GRSF1	G-rich RNA sequence binding factor 1
9931	HELZ	helicase with zinc finger
3178	HNRNPA1	heterogeneous nuclear ribonucleoprotein A1
3181	HNRNPA2B1	heterogeneous nuclear ribonucleoprotein A2/B1
3182	HNRNPAB	heterogeneous nuclear ribonucleoprotein A/B
3187	HNRNPH1	heterogeneous nuclear ribonucleoprotein H1 (H)
4670	HNRNPM	heterogeneous nuclear ribonucleoprotein M
10236	HNRNPR	heterogeneous nuclear ribonucleoprotein R
3192	HNRNPU	heterogeneous nuclear ribonucleoprotein U (scaffold attachment factor A)
9987	HNRPDL	heterogeneous nuclear ribonucleoprotein D-like
3183	HNRPNC	heterogeneous nuclear ribonucleoprotein C (C1/C2)
3308	HSPA4	heat shock 70kDa protein 4
3309	HSPA5	heat shock 70kDa protein 5 (glucose-regulated protein, 78kDa)
10808	HSPH1	heat shock 105kDa/110kDa protein 1
3608	ILF2	interleukin enhancer binding factor 2, 45kDa
3609	ILF3	interleukin enhancer binding factor 3, 90kDa
8570	KHSRP	KH-type splicing regulatory protein
23185	LARP4B	La ribonucleoprotein domain family, member 4B
124801	LSM12	LSM12 homolog (S. cerevisiae)
51631	LUC7L2	LUC7-like 2 (S. cerevisiae)
4134	MAP4	microtubule-associated protein 4
9782	MATR3	matrin 3
4343	MOV10	Mov10, Moloney leukemia virus 10, homolog (mouse)
65080	MRPL44	mitochondrial ribosomal protein L44
51642	MRPL48	mitochondrial ribosomal protein L48
4440	MSI1	musashi homolog 1 (Drosophila)
92140	MTDH	metadherin

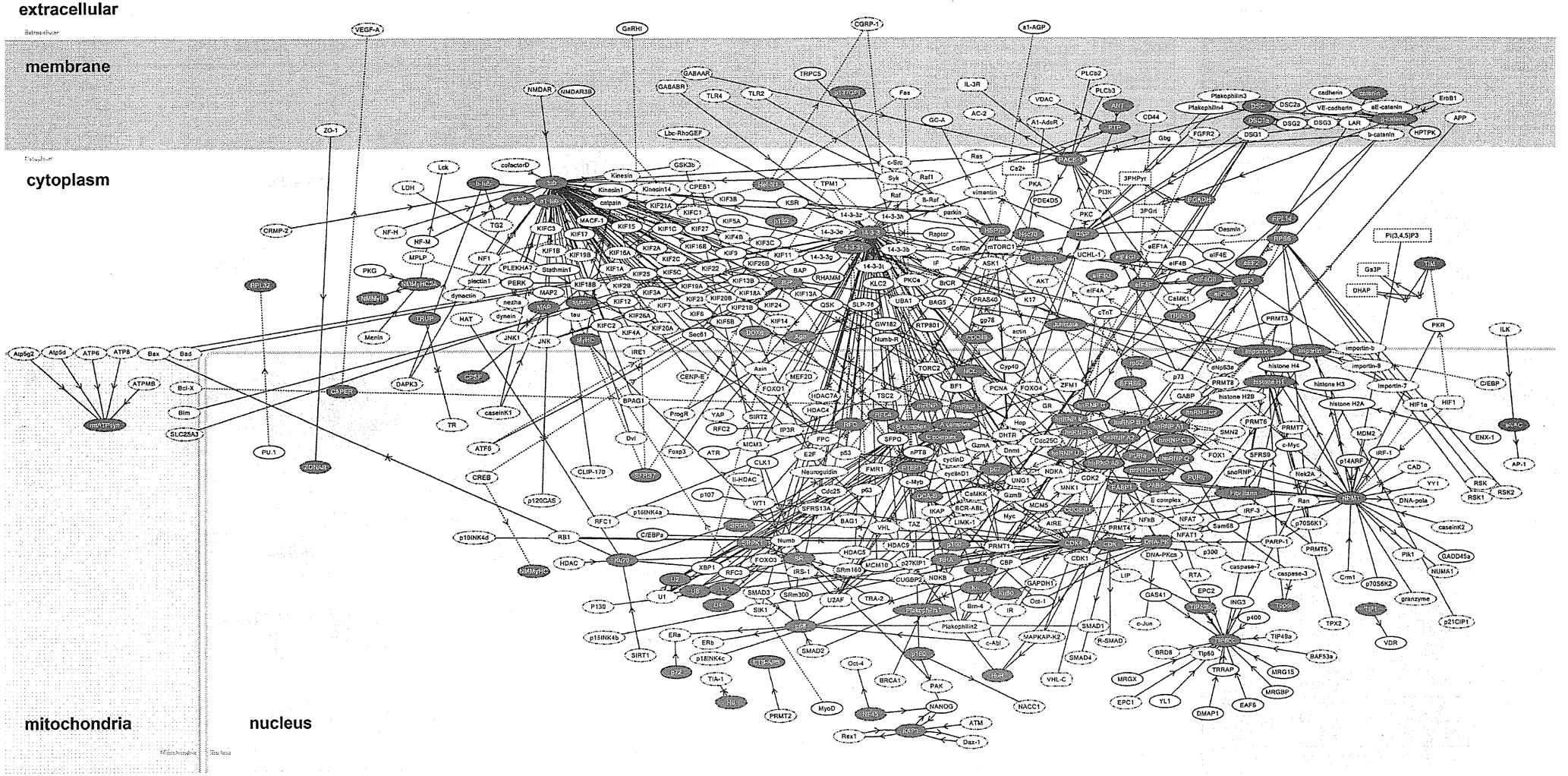
4686	NCBP1	nuclear cap binding protein subunit 1, 80kDa
4691	NCL	nucleolin
51491	NOP16	NOP16 nucleolar protein homolog (yeast)
10528	NOP56	NOP56 ribonucleoprotein homolog (yeast)
51602	NOP58	NOP58 ribonucleoprotein homolog (yeast)
4869	NPM1	nucleophosmin (nucleolar phosphoprotein B23, numatrin)
10482	NXF1	nuclear RNA export factor 1
26986	PABPC1	poly(A) binding protein, cytoplasmic 1
8761	PABPC4	poly(A) binding protein, cytoplasmic 4 (inducible form)
26227	PHGDH	phosphoglycerate dehydrogenase
5317	PKP1	plakophilin 1 (ectodermal dysplasia/skin fragility syndrome)
54814	QPCTL	glutaminyl-peptide cyclotransferase-like
9584	RBM39	RNA binding motif protein 39
9921	RNF10	ring finger protein 10
4736	RPL10A	ribosomal protein L10a
6144	RPL21	ribosomal protein L21
6146	RPL22	ribosomal protein L22
6160	RPL31	ribosomal protein L31
6124	RPL4	ribosomal protein L4
6133	RPL9	ribosomal protein L9
6205	RPS11	ribosomal protein S11
6218	RPS17	ribosomal protein S17
6222	RPS18	ribosomal protein S18
6228	RPS23	ribosomal protein S23
6229	RPS24	ribosomal protein S24
6189	RPS3A	ribosomal protein S3A
6201	RPS7	ribosomal protein S7
6203	RPS9	ribosomal protein S9
3921	RPSA	ribosomal protein SA
9092	SART1	squamous cell carcinoma antigen recognized by T cells
26135	SERBP1	SERPINE1 mRNA binding protein 1
79085	SLC25A23	solute carrier family 25 (mitochondrial carrier; phosphate carrier), member 23

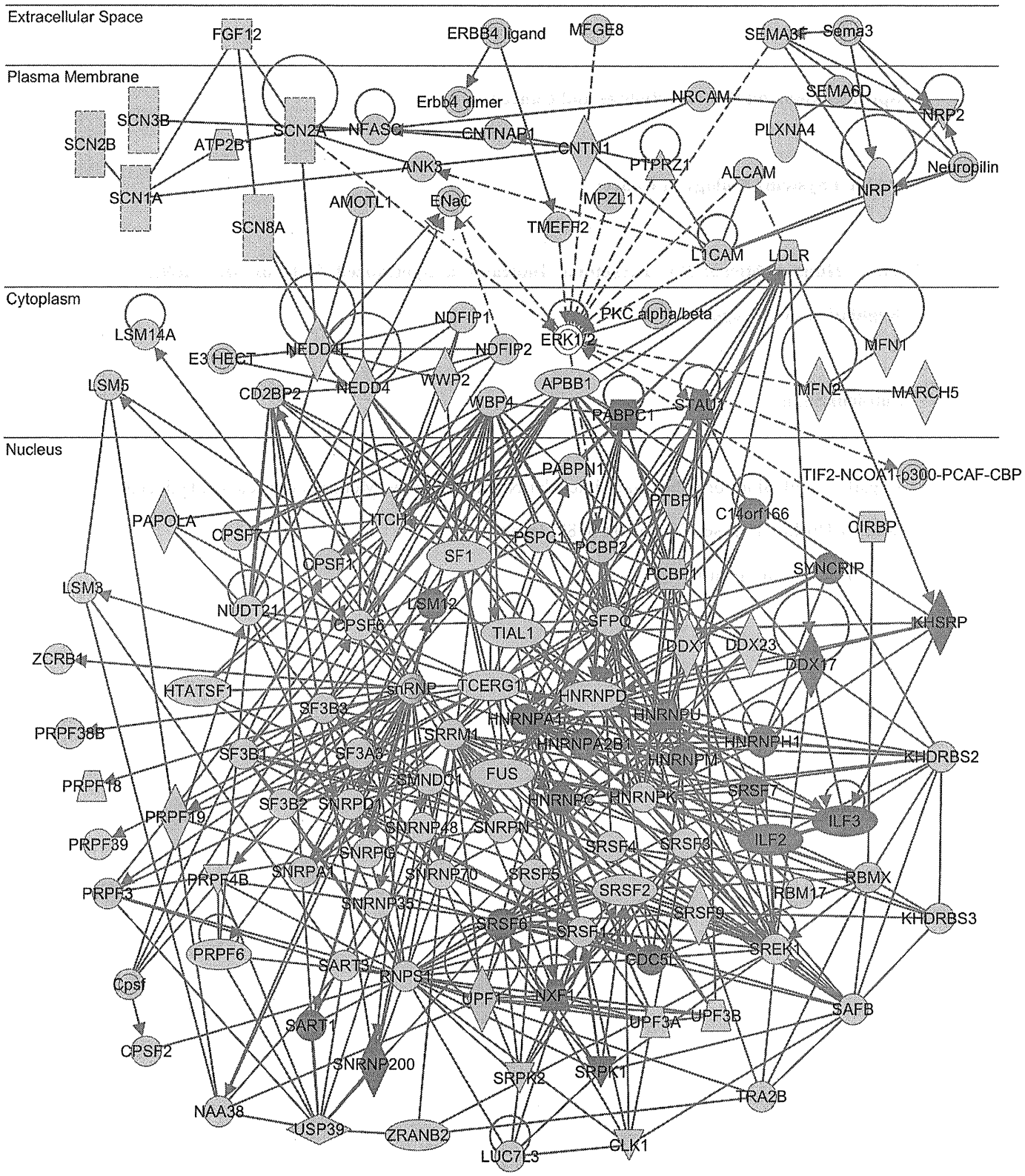
23020	SNRNP200	small nuclear ribonucleoprotein 200kDa (U5)
6732	SRPK1	SRSF protein kinase 1
6431	SRSF6	serine/arginine-rich splicing factor 6
6432	SRSF7	serine/arginine-rich splicing factor 7
6741	SSB	Sjogren syndrome antigen B (autoantigen La)
6780	STAU1	staufen, RNA binding protein, homolog 1 (Drosophila)
10492	SYNCRIP	synaptotagmin binding, cytoplasmic RNA interacting protein
23435	TARDBP	TAR DNA binding protein
7150	TOP1	topoisomerase (DNA) I
9100	USP10	ubiquitin specific peptidase 10
7415	VCP	valosin-containing protein
84305	WIBG	within bgcn homolog (Drosophila)
22803	XRN2	5'-3' exoribonuclease 2
9877	ZC3H11A	zinc finger CCCH-type containing 11A
23567	ZNF346	zinc finger protein 346

The set of 106 genes concurrently serving as both human TDP-43 target RNAs and interacting proteins are listed in an alphabetical order. The genes with Entrez Gene IDs of 3921, 4736, 6124, 6133, 6144, 6146, 6160, 6189, 6201, 6203, 6205, 6218, 6222, 6228, and 6229 are located on "Ribosome", while those with 988, 1655, 3178, 3183, 3192, 4670, 4686, 6431, 6432, 9092, and 23020 are located on "Splicesome" of KEGG pathways.


AXON GUIDANCE
Receptors
Neuron (developing growth cone)
Guidance cues
(Netrins)
Netrin1
Netrin3
Netrin4
NetrinG
DCC
DCC
UNC-5
DCC
NGL-1
Nck1
Rac
Ablim
Ca2+
CALN
-p
NFAT
Transcription
RhoA
ROCK
(Ephrins)
Attraction
p120
Fyn
Ephrin-A
EphA
Abl
Rac
CDC42
PAK
Ephexin
RHOA
ROCK
Repulsion
FAK
Grb4
Ephrin-B
EphB
RasGAP
Ras
ERK
RGS3
Cytokine-cytokine receptor interaction
MAPK signaling pathway
Regulation of actin cytoskeleton
Attraction
Gi
CXCR4
CXCL12
CXCR4
Gi
(Slits)
Slit1
Slit2
Slit3
Robo2
Robo1
Robo3
srGAP
Cdc42
Regulation of actin cytoskeleton
GSK3β
+p
Fyn
CDK5
CRMP
CRAM
Microtubule reorganization
Fes
(Semaphorins)
Plexin A
Sema3A
NRP1
L1CAM
Rac
PAK
LIMK
Cofilin
Rnd1
RhoD
Sema4D
Sema5
Sema6
Plexin B
MET
Rac
PAK
LARG
RhoA
Rnd1
Sema7A
ITGB1
Plexin C
RAK
ERK
Regulation of actin cytoskeleton
Axon attraction
Axon repulsion
Axon outgrowth


SPLICEOSOME
5' splice site
Branch point
3' splice site
pre-mRNA
Exon
GU
A
AG
Exon
Intron
mRNA
Common components
U2AF
Prp5
UAP56
Prp22
U4/U6.U5 tri-snRNP
complex A
Post-spliceosomal complex
EJC/ TREX 2nd step
Prp16
Prp22
Prp17
Slu7
Prp18
PRP19 complex
complex B
complex C
1st step
Prp2
Activated spliceosome
complex B*
Spliceosome components
U1
U2
U4/U6
U5
U1snRNA
U2snRNA
U4snRNA
U6snRNA
U5snRNA
Sm
U1-70K
U1A
U1C
U1 related
FBP11
S164
p68
CA150
U2A'
U2B"
SF3a
SF3b
U2 related
U2AF
PUF60
SPF30
SPF45
CHERP
SR140
Prp43
Lsm
Prp3
Prp4
CypH
Prp31
Snu13
U4/U6.U5 tri-SnRNP associated
SnRNP27
Sad1
Snu66
Snu23
Prp38
Snu114
Brr2
Prp6
Prp8
Prp8BP
Prp28
DIB1
Prp19 complex
Prp19
CDC5
SPF27
PRL1
AD002
CINNEL1
HSP73
Prp19 related
SKIP
Syf
Isyl
PPIL1
CypE
CCDC12
RBM22
G10
AQR
EJC/TREX
ACINUS
eIFA3
Y14
magoh
UAP56
THOC
Common components
CBP80/20
hnRNPs
SR
Complex B specific
NPW38
NPW38BP
U4/U6 snRNP
U6 snRNA
U4 snRNA
Sm
Lsm
Prp3
Prp4
CypH
Prp31
Snu13
5'
3'


extracellular
membrane
cytoplasm
mitochondria
nucleus


Extracellular Space
Plasma Membrane
Cytoplasm
Nucleus
FGF12
ERBB4 ligand
MFGE8
SEMA3F
Sema3
Erbb4 dimer
NRCAM
SEMA6D
NRP2
SCN3B
SCN2B
ATP2B1
SCN2A
NFASC
CNTNAP1
CNTN1
PLXNA4
ANK3
PTPRZ1
ALCAM
Neuropilin
SCN1A
NRP1
AMOTL1
ENaC
MPZL1
SCN8A
TMEFF2
L1CAM
LDLR
LSM14A
NDFIP1
PKC alpha/beta
E3 HECT
NEDD4L
NDFIP2
ERK1/2
MFN1
LSM5
CD2BP2
NEDD4
WWP2
APBB1
MFN2
MARCH5
WBP4
PABPC1
STAU1
PABPN1
TIF2-NCOA1-p300-PCAF-CBP
PTBP1
C14orf166
PAPOLA
CPSF7
ITCH
PSPC1
PCBP2
CIRBP
CPSF1
SF1
PCBP1
SMNCRIP
LSM3
NUDT21
LSM12
CPSF6
TIAL1
SFPQ
KHSRP
DDX1
DDX23
DDX17
ZCRB1
HTATSF1
snRNP
TCERG1
HNRNPD
SF3B3
HNRNPA1
HNRNPU
PRPF38B
SRRM1
HNRNPA2B1
HNRNPH1
SF3B1
SF3A3
HNRNPM
KHDRBS2
PRPF18
SMNDC1
FUS
SF3B2
SNRPD1
HNRNPC
HNRNPK
SRSF7
PRPF19
ILF3
SNRNP48
SNRPN
PRPF39
SRSF4
SRSF3
ILF2
SNRPG
SNRPA1
PRPF4B
SNRNP70
SRSF5
RBMX
PRPF3
SRSF2
RBM17
SNRNP35
SRSF9
KHDRBS3
SRSF6
SRSF1
PRPF6
SREK1
SART3
CDC5L
Cpsf
RNPS1
UPF1
NXF1
SAFB
UPF3A
UPF3B
SART1
CPSF2
SNRNP200
SRPK2
SRPK1
TRA2B
NAA38
USP39
ZRANB2
CLK1
LUC7L3

Springer Book on Systems Biology and Cancer

PART I Systems Biology in Cancer

Title: Human MicroRNA Targetome Indicates a Specialized Role of MicroRNAs in Regulation of Oncogenesis

Jun-ichi Satoh

Department of Bioinformatics and Molecular Neuropathology, Meiji Pharmaceutical University, 2-522-1 Noshio, Kiyose, Tokyo 204-8588, Japan.

E-mail: satoj@my-pharm.ac.jp

Keywords

c-myb

Diana microT

E2F

KeyMolnet

knowledgebase

let-7

microRNA

microRNAome

miR-15a

miR-106b

MITF

molecular network

oncogenesis

targetome

ZEB

Abstract

MicroRNAs (miRNAs), a class of endogenous small noncoding RNAs, mediate posttranscriptional regulation of protein-coding genes by binding to the 3' untranslated region of target mRNAs, leading to translational inhibition, mRNA destabilization or degradation. A single miRNA concurrently downregulates hundreds of target mRNAs, and thereby fine-tunes gene expression involved in diverse cellular functions, such as development, differentiation, proliferation, apoptosis and metabolism. However, it remains unknown whether the set of miRNA target genes designated "targetome" regulated by an individual miRNA constitutes the biological network of functionally-associated molecules or reflects a random set of functionally-independent genes. To address this question, we studied the molecular network of the whole human miRNA targetome. Among 1,223 human miRNAs derived from miRbase Release 16, Diana-microT 3.0, a target prediction program, predicted reliable targets from 273 miRNAs. Among them, KeyMolnet, a bioinformatics tool for analyzing molecular interactions on the comprehensive knowledgebase, successfully extracted molecular networks from 232 miRNAs. In miRNA targetome networks, the most relevant pathway was transcriptional regulation by RB/E2F, important regulators of oncogenic transformation, the disease was adult T cell lymphoma/leukemia, and the pathological event was cancer, indicating that the human miRNA system termed "miRNAome" plays a specialized role in regulation of oncogenesis. The predicted targets derived from approximately 20% of all human miRNAs construct biologically meaningful molecular networks, supporting the view that the miRNA targetome generally constitutes the biological network of functionally-associated molecules in human cells.

Contents

Abbreviations

EMT	Epithelial-Mesenchymal Transition
HPRD	Human Protein Reference Database
IPA	Ingenuity Pathways Analysis
KEGG	Kyoto Encyclopedia of Genes and Genomes
miTG	microRNA-Targeted Gene
MRE	MicroRNA Recognition Elements
PPI	Protein-Protein Interaction
RISC	RNA-Induced Silencing Complex
3'UTR	3' Untranslated Region

1 MicroRNAome and MicroRNA Targetome

MicroRNAs (miRNAs) are a class of endogenous small noncoding RNAs conserved through the evolution. They mediate posttranscriptional regulation of protein-coding genes by binding to the 3' untranslated region (3'UTR) of target mRNAs, leading to translational inhibition, mRNA destabilization or degradation, depending on the degree of sequence complementarity (Guo et al. 2010). During the biogenesis of miRNAs, the pri-miRNAs are transcribed from the intra- and inter-genetic regions of the genome by RNA polymerase II, followed by processing by the RNase III enzyme Drosha into pre-miRNAs. After nuclear export, they are cleaved by the RNase III enzyme Dicer into mature miRNAs consisting of approximately 22 nucleotides. Finally, a single-stranded miRNA is loaded onto the Argonaute-containing RNA-induced silencing complex (RISC), where the seed sequence located at positions 2 to 8 from the 5' end of the miRNA serves as an essential scaffold for recognizing the target mRNA (Bartel 2009).

Currently, more than one thousand of human miRNAs are registered in the miRNA database named miRBase Release 17 (April 2011; www.mirbase.org). In general, the 3'UTR of a single mRNA is targeted by several different miRNAs, while a single miRNA at one time reduces the production of hundreds of target proteins that constitute "targetome" (Selbach et al. 2008). Such redundant interactions between miRNAs and their targets are responsible for the complexity of miRNA-regulated gene expression. Furthermore, certain miRNAs activate transcription and translation of the targets (Vasudevan et al. 2007; Place et al. 2008). Consequently, the whole human miRNA system termed "miRNAome" regulates greater than 60% of all protein-coding genes (Friedman et al. 2009). By targeting multiple transcripts and affecting expression of numerous proteins, miRNAs regulate diverse cellular functions, such as development, differentiation, proliferation, apoptosis and metabolism. Therefore, aberrant regulation of miRNA expression is deeply involved in pathological events that mediate cancers (Blenkiron and Miska 2007; Garzon et al. 2010) and neurodegenerative disorders (Shioya et al. 2010; Wang et al. 2011).

Recent advances in systems biology have made major breakthroughs by illustrating the cell-wide map of complex molecular interactions with the aid of the literature-based knowledgebase of molecular pathways (Viswanathan et al. 2008). The logically arranged molecular networks construct the whole system characterized by robustness, which maintains the proper function of the system in the face of genetic and environmental perturbations (Kitano 2007). In the scale-free molecular network, targeted disruption of limited numbers of critical components designated hubs, on which the biologically important molecular interactions concentrate, efficiently disturbs the whole cellular function by destabilizing the network (Albert et al. 2000). Therefore, the identification and characterization of hub molecules located in the center of the miRNA targetome network would help us to understand biological and pathological roles of individual miRNAs. A recent study determined the human miRNA-regulated protein-protein interaction (PPI) network by utilizing the Human Protein Reference Database (HPRD) and the miRNA target prediction program TargetScan (Hsu et al. 2008). They found that an individual miRNA often targets the hub gene of the PPI network, although they did not attempt to clarify functionally relevant pathways, diseases, and pathological events that play a central role in the miRNA targetome network.

At present, it remains unknown whether the miRNA targetome regulated by an individual miRNA generally constitutes the biological network of functionally-associated molecules or simply reflects a random set of functionally-independent genes. To address this question, we attempted to characterize the molecular network of the whole human miRNA targetome. We found that the miRNA targetome constitutes the biological network of functionally-associated molecules in human cells (Satoh and Tabunoki 2011). Furthermore, functional annotation of the miRNA targetome suggested that the human miRNAome plays a specialized role in regulation of oncogenesis. Importantly, we identified a collaborative regulation of gene expression by transcription factors and miRNAs in cancer-associated miRNA targetome networks.

2 Molecular Network of MicroRNA Targetome

2.1 MicroRNA Target Prediction Programs

First of all, we downloaded the complete list of 1,223 human miRNAs from miRBase Release 16 (September 2010; www.mirbase.org). In general, miRNAs regulate gene expression by forming energetically stable Watson-Crick base pairs with target mRNAs. In most occasions, the seed sequence conserved through evolution located at positions 2 to 8 from the 5' end of the miRNA serves as an essential scaffold for recognizing the target mRNA. The thermodynamic rule makes it possible to fairly accurately predict miRNA target mRNAs by using computational approaches (Bartel 2009). Since open source miRNA target prediction programs, such as TargetScan 5.1 (www.targetscan.org), PicTar (pictar.mdc-berlin.de), miRanda (www.microrna.org) and Diana-microT 3.0 (diana.cslab.ece.ntua.gr/microT), are armed with own unique algorithms, the set of predicted targets often vary among distinct programs utilized (Boross et al. 2009). Furthermore, the lists of predicted targets are mostly cell and tissue-type non-specific, and they inevitably have a risk for containing numerous false positive ones. Recently, to overcome these problems, the miRTarBase (mirtarbase.mbc.nctu.edu.tw) has been established, which represents the largest collection of more than 3,500 manually curated miRNA-target interactions from 985 articles, all of which are experimentally validated by luciferase reporter assay, western blot, quantitative RT-PCR, microarray experiments with overexpression or knockdown of miRNAs, or pulsed stable isotope labeling with amino acids in culture (pSILAC) experiments (Hsu et al. 2011).

We searched the target genes of individual 1,223 miRNAs on the Diana-microT 3.0 target prediction program, which was selected because of the highest ratio of correctly predicted targets over other prediction programs (Maragkakis et al. 2009). Diana-microT 3.0 calculates the miRNA-targeted gene (miTG) score that reflects the weighted sum of the scores of all conserved and non-conserved miRNA recognition elements (MRE) on the 3'UTR of the target mRNA. The miTG score correlates well with fold changes in suppression of protein expression. To optimize the parameter of miRNA-target interaction, we considered the target genes with a

cutoff of the miTG score equal to or larger than 20 as the highly reliable targets, because we found that the targets with the miTG score < 20 exhibited the significantly lower precision score, an indicator of correctness in predicted interactions, compared with those having the score ≥ 20 ($p = 2.78E-08$ by Mann-Whitney's U-test) (Satoh and Tabunoki 2011).

2.2 Molecular Network Analysis Tools

To identify biologically relevant molecular pathways from large-scale data, we could analyze them by using a battery of pathway analysis tools endowed with a comprehensive knowledgebase, such as Kyoto Encyclopedia of Genes and Genomes (KEGG) (www.kegg.jp), Ingenuity Pathways Analysis (IPA) (Ingenuity Systems, www.ingenuity.com), and KeyMolnet (Institute of Medicinal Molecular Design, www.immd.co.jp). KEGG is an open-access database, while both IPA and KeyMolnet are commercial ones, all of which are updated frequently.

KEGG includes manually curated reference pathways that cover a wide range of metabolic, genetic, environmental, and cellular processes, and human diseases (Kanehisa et al. 2010). Currently, KEGG contains 146,294 pathways generated from 406 reference pathways. IPA is a knowledgebase that contains approximately 2,500,000 biological and chemical interactions and functional annotations with definite scientific evidence, curated by expert biologists. KeyMolnet is a tool for analyzing molecular interactions on the literature-based knowledgebase, composed of the contents on 137,300 molecular relationships among human genes, miRNAs, proteins, small molecules, diseases, pathways and drugs. The core contents are collected from selected review articles and textbooks with the highest reliability, curated by expert biologists. The KeyMolnet library contains a panel of human canonical networks constructed by core contents, which represent the gold standard of the networks, composed of 430 pathways, 885 diseases, and 208 pathological events (Satoh et al. 2009; Satoh 2010; Satoh and Tabunoki 2011).

Ensembl Gene IDs of target genes retrieved by Diana-microT 3.0 were converted into the corresponding Entrez Gene IDs by using the DAVID Bioinformatics Resources 6.7 program (david.abcc.ncifcrf.gov) (Huang et al. 2009). Non-annotated IDs were deleted. Then, Entrez Gene IDs of miRNA target genes were uploaded onto KeyMolnet. We utilized the neighboring