Literature-based discovery of diabetes- and ROS-related targets

- ROS-Diabetes webpage manual -

Authors:

Junguk Hur[1,2]	(juhur@umich.edu)
Kelli A. Sullivan[2]	(ksulliva@umich.edu)
Adam D. Schuyler[4]	(schuyler@uchc.edu)
Yu Hong[2]	(hongyu@umich.edu)
Manjusha Pande[2,3]	(mpande@umich.edu)
David J. States[5]	(David.J.States@uth.tmc.edu)
H. V. Jagadish[1,3]	(jag@umich.edu)
Eva L. Feldman[1,2,3*]	(efeldman@umich.edu)

Affiliations:
1) Bioinformatics Program
2) Department of Neurology
3) National Center for Integrative Biomedical Informatics
University of Michigan
Ann Arbor, MI 48109, USA
4) Department of Molecular, Microbial, and Structural Biology,
University of Connecticut Health Center
Farmington, CT 06030, USA
5) School of Health Information Science
University of Texas at Houston
Houston, TX 77030, USA

* To whom correspondence should be made

http://jdrf.neurology.med.umich.edu/ROSDiabetes/

ROS-Diabetes web-page manual

The ROS-Diabetes website (http://jdrf.neurology.med.umich.edu/ROSDiabetes/) provides comprehensive lists of the ROS-Diabetes associated targets (genes/proteins) mined from the biomedical literature by SciMiner. Users can navigate the website to obtain details of ROS-Diabetes targets such as frequency of targets, actual terms in the literature, and over-represented biological functions in the ROS-Diabetes targets in terms of gene ontology (GO) and pathways.

The website also allows users to predict putative ROS-Diabetes targets based on either the enriched GO and pathway terms of 53 core ROS-Diabetes targets or user-provided GO and pathway terms. Using this feature, users can navigate putative ROS-Diabetes related targets sharing similar biological functions with the core 53 ROS-Diabetes (or user-provided) targets but that have not been identified in ROS-Diabetes related literature yet. These putative ROS-Diabetes targets will augment the literature of derived ROS-Diabetes targets.

S1. Explore Menu

The Explore menu has four options available.

- (S1.1.) Full 1,026 ROS-Diabetes targets identified by SciMiner
- (S1.2.) Enriched biological functions of 1,026 ROS-Diabetes targets
- (S1.3.) 53 highly-overrepresented ROS-Diabetes targets
- (S1.4.) Enriched biological functions of 53 ROS-Diabetes targets

S1.1. Full 1,026 ROS-Diabetes targets identified by SciMiner

Users can see details of all 1,026 ROS-Diabetes targets identified by SciMiner. The main page contains a link to a document-based summary page and the top 10 most frequent ROS-Diabetes targets.

Summary

SciMiner has analyzed 1154 articles and identified 1026 targets. [Documents Summary] [PMIDs] Citations

Top 10 most frequent targets ... (All 1026 genes)

[HIDE Top 10 most frequent targets]

Rank	HUGO	Symbol	Target Name	#Occur	#Paper	MiMI
1	6081	INS	insulin	4605	503	
2	11179	SOD1	superoxide dismutase 1, soluble (amyotrophic lateral sclerosis 1 (adult))	2479	368	
3	1516	CAT	catalase	1424	241	
4	9393	PRKCA	protein kinase C, alpha	491	194	
5	399	ALB	albumin	550	179	
6	14874	NOX5	NADPH oxidase, EF-hand calcium binding domain 5	1393	177	
7	7873	NOS2A	nitric oxide synthase 2A (inducible, hepatocytes)	932	144	
8	12805	XDH	xanthine dehydrogenase	375	133	
9	333	AGT	angiotensinogen (serpin peptidase inhibitor, clade A, member 8)	1997	131	
10	11892	TNF	tumor necrosis factor (TNF superfamily, member 2)	1299	120	

Full result (HTML, TXT, EXCEL)

Figure 1. Main Display of Full 1,026 ROS-Diabetes targets

- Link to document-based pages
  - [Document Summary]: The list of 1,154 ROS-Diabetes literature with details.
  - [PMIDs]: The list of 1,154 ROS-Diabetes literature in PMIDs (PubMed)
  - [Citations]: An EndNote citation for all the 1,154 papers
- Top 10 most frequent targets.

S1.1.1. Link to [document summary]

Clicking [Document Summary] will display a table of 1,154 ROS-Diabetes papers.

PMID	TITLE	# of Targets	Links	Journal
18984854	Proteomic changes associated with diabetes in the BB-DP rat.	106	M E H P	Am J Physiol Endocrinol Metab
17979836	From death receptor to reactive oxygen species and c-Jun N-terminal protein kinase: the receptor-interacting protein 1 odyssey.	97	M E H P	Immunol Rev
14567954	The contribution of mitochondria to common disorders.	70	M E H P	Mol Genet Metab
18423414	Oxidative damage to extracellular matrix and its role in human pathologies.	67	M E H P	Free Radic Biol Med
17586614	Renin-angiotensin-aldosterone system and oxidative stress in cardiovascular insulin resistance.	59	M E H P	Am J Physiol Heart Circ Physiol
17337232	Nitric oxide: ocular blood flow, glaucoma, and diabetic retinopathy.	57	M E H P	Prog Retin Eye Res
12372842	Oxidative stress and stress-activated signaling pathways: a unifying hypothesis of type 2 diabetes.	55	M E H P	Endocr Rev
17584843	Diabetes associated cell stress and dysfunction: role of mitochondrial and non-mitochondrial ROS production and activity.	54	M E H P	J Physiol
18793162	NOX family NADPH oxidases in liver and in pancreatic islets: a role in the metabolic syndrome and diabetes?	52	M E H P	Biochem Soc Trans

Figure 2. Document summary

- **# of Targets:** The total number of targets found in each article (Figure 3)
- **Links:**
  1. **M:** MEDLINE (text format)
  2. **E:** Export EndNote citation
  3. **H:** HTML (full text) file
  4. **P:** PubMed summary (NCBI PubMed page)
- At the bottom of the page, there is a clickable link for downloading the citation (EndNote) of all 1,154 ROS-Diabetes papers.

Document Information

PMID	18984854 (M E H P)
Title	Proteomic changes associated with diabetes in the BB-DP rat.
Abstract	These studies were structured with the aim of utilizing emerging technologies in two-dimensional (2D) gel electrophoresis and mass spectrometry to evaluate protein expression changes associated with type 1 diabetes. We reasoned that a broad examination of diabetic tissues at the protein level might open up novel avenues of investigation of the metabolic and signaling pathways that are adversely affected in type 1 diabetes. This study compared the protein expression of the liver, heart, and skeletal muscle of diabetes-prone rats and matched control rats by semiquantitative liquid chromatography-mass spectrometry and differential in-gel 2D gel electrophoresis. Differential expression of 341 proteins in liver, 43 in heart, and 9 (2D gel only) in skeletal muscle was detected. These data were assembled into the relevant metabolic pathways affected primarily in liver. Multiple covalent modifications were also apparent in 2D gel analysis. Several new hypotheses were generated by these data, including mechanisms of net cytosolic protein oxidation, formaldehyde generation by the methionine cycle, and inhibition of carbon substrate oxidation via reduction in citrate synthase and short-chain acyl-CoA dehydrogenase.

NOTE: Color highlight is limited to the **abstract** and **SciMiner text-mining mode**. If you see much more identified targets below from "Targets by SciMiner Summary" and "Targets by SciMiner Full list", they may have been identified from the **full text**.

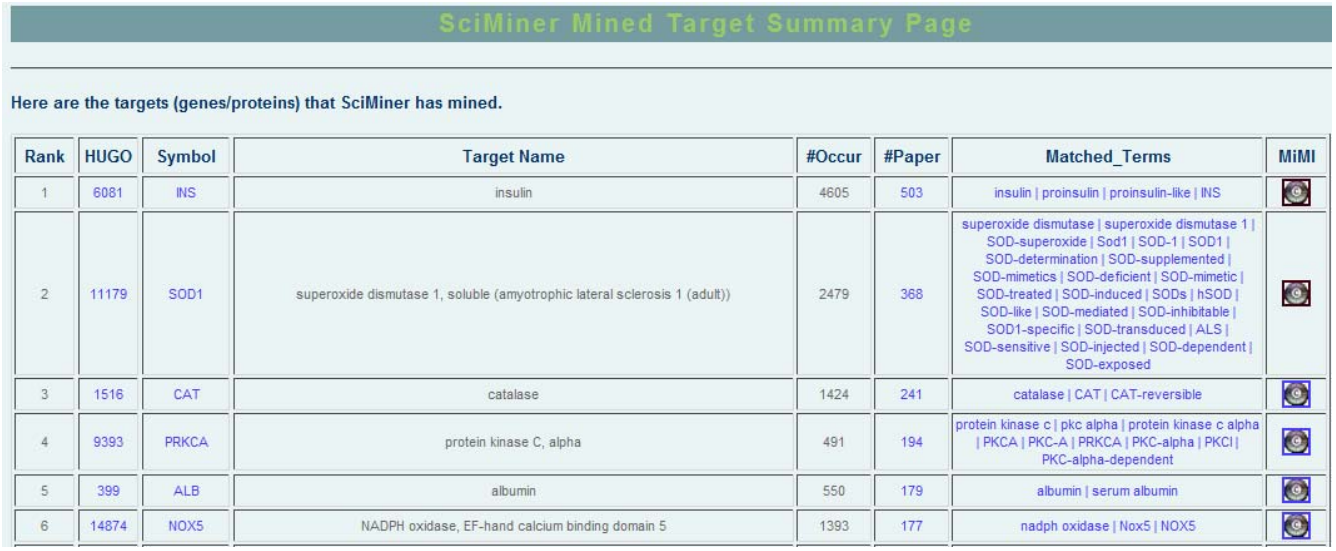
Targets by SciMiner Summary

HUGO ID	Symbol	Target Name	#Occur	ActualStr
6081	INS	insulin	11	insulin |
4196	GCKR	glucokinase (hexokinase 4) regulator	10	glucokinase regulatory protein | GKRP |
3700	FH	fumarate hydratase	7	fumarate hydratase | fumarase |
90	ACADS	acyl-Coenzyme A dehydrogenase, C-2 to C-3 short chain	6	short chain acyl coa dehydrogenase | SCAD |
21210	LPAL2	lipoprotein, Lp(a)-like 2	6	apolipoprotein a | apolipoprotein a |
11180	SOD2	superoxide dismutase 2, mitochondrial	5	mn sod | Mn-SOD | mn superoxide dismutase |
4458	GPI	glucose phosphate isomerase	4	glucose 6 phosphate isomerase | PGI |
2228	COMT	catechol-O-methyltransferase	3	COMT | catechol o methyltransferase |
9726	PYGM	phosphorylase, glycogen; muscle (McArdle syndrome, glycogen storage disease type	3	glycogen phosphorylase |

Figure 3. Individual document detail providing abstract and identified targets

S1.1.2. Top 10 most frequent targets

The second section of the main page for Full 1,026 ROS-Diabetes targets identified by SciMiner shows a list of the top 10 most frequent targets in 1,154 ROS-Diabetes literature including insulin (INS), superoxide dismutase 1 (SOD1), catalase (CAT), etc. The full list of 1,026 ROS-Diabetes genes is available at the following hyperlinks; "**All 1,026 genes**", "**Full result HTML**".



SciMiner Mined Target Summary Page

Here are the targets (genes/proteins) that SciMiner has mined.

Rank	HUGO	Symbol	Target Name	#Occur	#Paper	Matched_Terms	MiMI
1	6081	INS	insulin	4605	503	insulin | proinsulin | proinsulin-like | INS	
2	11179	SOD1	superoxide dismutase 1, soluble (amyotrophic lateral sclerosis 1 (adult))	2479	368	superoxide dismutase | superoxide dismutase 1 | SOD-superoxide | Sod1 | SOD-1 | SOD1 | SOD-determination | SOD-supplemented | SOD-mimetics | SOD-deficient | SOD-mimetic | SOD-treated | SOD-induced | SODs | hSOD | SOD-like | SOD-mediated | SOD-inhibitable | SOD1-specific | SOD-transduced | ALS | SOD-sensitive | SOD-injected | SOD-dependent | SOD-exposed	
3	1516	CAT	catalase	1424	241	catalase | CAT | CAT-reversible	
4	9393	PRKCA	protein kinase C, alpha	491	194	protein kinase c | pkc alpha | protein kinase c alpha | PKCA | PKC-A | PRKCA | PKC-alpha | PKCI | PKC-alpha-dependent	
5	399	ALB	albumin	550	179	albumin | serum albumin	
6	14874	NOX5	NADPH oxidase, EF-hand calcium binding domain 5	1393	177	nadph oxidase | Nox5 | NOX5	

Figure 4. Full 1,026 ROS-Diabetes genes (truncated)

- **Field description**
  1. **Rank**
  2. **HUGO:** HGNC (Human Gene Nomenclature Consortium) HUGO ID
  3. **Symbol:** Official HUGO symbol
  4. **Target Name:** Official HUGO description
  5. **#Occur:** Total number of occurrences in all of the papers
  6. **#Paper:** Total number of papers in which target has been found
  7. **Matched_terms:** Actual matched terms in articles
  8. **MiMI:** JAVA Web Start link to launch Cytoscape with MiMI plugin to display protein-protein interaction network of the given targets.
- **Further links**
  1. #Paper: Lists all documents that the specific target was identified in (Figure 5)
  2. Matched_Terms: Lists all the incidences of actual matching with up to 100 characters at each side. (flanking text in Figure 6)

List of PMIDs

Current Target Information

HUGO ID	11179
Symbol	SOD1
Name	superoxide dismutase 1, soluble (amyotrophic lateral sclerosis 1 (adult))
#Occurrence	2475
#Paper	368

PMID	TITLE	# of Targets	Links	Journal
18984854	Proteomic changes associated with diabetes in the BB-DP rat.	106	M E H P	Am J Physiol Endocrinol Metab
17979836	From death receptor to reactive oxygen species and c-Jun N-terminal protein kinase: the receptor-interacting protein 1 odyssey.	97	M E H P	Immunol Rev
14567954	The contribution of mitochondria to common disorders.	70	M E H P	Mol Genet Metab
18423414	Oxidative damage to extracellular matrix and its role in human pathologies.	67	M E H P	Free Radic Biol Med
17586614	Renin-angiotensin-aldosterone system and oxidative stress in cardiovascular insulin resistance.	59	M E H P	Am J Physiol Heart Circ Physiol
17337232	Nitric oxide: ocular blood flow, glaucoma, and diabetic retinopathy.	57	M E H P	Prog Retin Eye Res

Figure 5. List of PMIDs for each target

HUGO ID Detailed Result 11179

HUGO ID	11179
Symbol	SOD1
Name	superoxide dismutase 1, soluble (amyotrophic lateral sclerosis 1 (adult))
#Occurrence	2475
#Paper	368

PMID	Match String	Actual String	Score	Flanking text
1646926	SOD	SOD	0.9	The effect of superoxide dismutase (SOD) SOD on the inhibition of peroxidation of unsaturated lipids within liposomal
2171534	SOD	SOD	0.9	Allopurinol and SOD inhibited cytochrome c reduction in a hypoxanthine-xanthine oxidase superoxide generating
3623415	SOD	SOD	0.9	evaluated the following red blood cell parameters superoxide dismutase (SOD), SOD glutathione peroxidase (GSH.Px), GSH.Px catalase (C-ase), C-ase glutathione content (GSH)
6324223	SOD	SOD	0.9	peroxide was determined by the activities of superoxide dismutase (SOD) SOD and catalase respectively
6324223	SOD	SOD	0.9	SOD activity was not altered in the diabetic rat skeletal muscles
6327438	SOD	SOD	0.9	the activities of both cytoplasmic and mitochondrial superoxide dismutase (SOD), SOD the effect being more pronounced in the cytoplasmic fraction
6327438	SOD	SOD	0.9	PMNL obtained from insulin-treated diabetic patients showed considerable alleviation of SOD levels
7742795	SOD	SOD	0.9	completely inhibited the DNA strand breaks but superoxide dismutase (SOD) SOD did not suggesting that the strand breaks are induced by

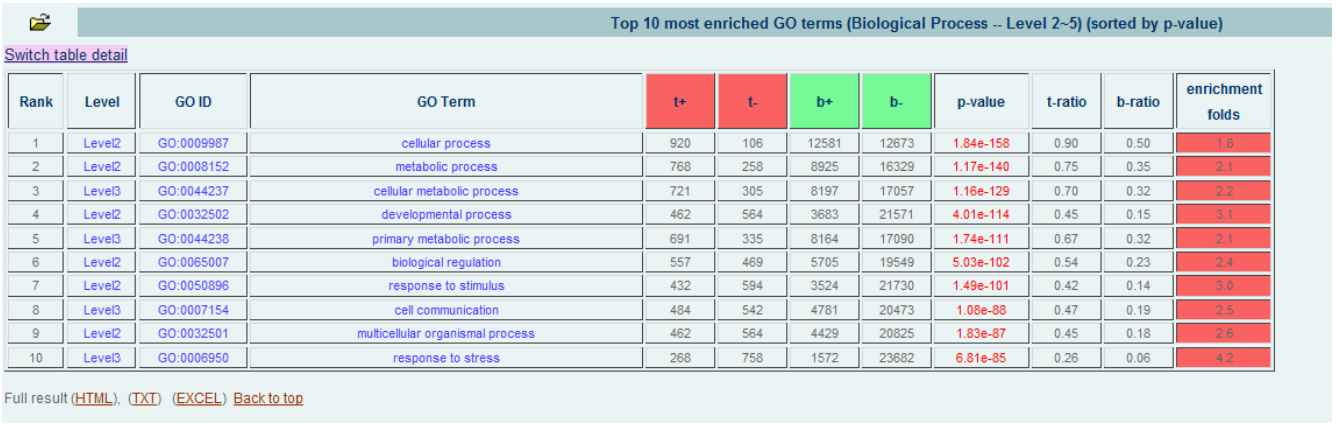
Figure 6. Matched terms detail with flanking text

S1.2. Enriched biological functions of 1,026 ROS-Diabetes targets

This section displays enriched biological functions of the ROS-Diabetes genes in terms of Gene Ontology (GO), pathways, and protein-protein interaction networks. Fisher's exact test is used to assess statistical significance of the literature-derived ROS-Diabetes target set (using a 2x2 contingency table (T: 1,026 ROS-Diabetes, B: Full HUGO genes)).

S1.2.1. Gene Ontology

There are three GO categories (Biological Processes, Molecular Functions, and Cellular Components). Each GO category has complete enrichment analysis results based on three different sets of GO terms; (1) Level 2~5 (using only level 2~5 GO terms); (2) Full GO (using all of the explicitly assigned GO and implicitly assigned GO obtained from the GO tree structure); (3) Explicit GO (using only the explicitly assigned GO terms)



Top 10 most enriched GO terms (Biological Process -- Level 2~5) (sorted by p-value)

Switch table detail

Rank	Level	GO ID	GO Term	t+	t-	b+	b-	p-value	t-ratio	b-ratio	enrichment folds
1	Level2	GO:0009987	cellular process	920	106	12581	12673	1.84e-158	0.90	0.50	1.8
2	Level2	GO:0008152	metabolic process	768	258	8925	16329	1.17e-140	0.75	0.35	2.1
3	Level3	GO:0044237	cellular metabolic process	721	305	8197	17057	1.16e-129	0.70	0.32	2.2
4	Level2	GO:0032502	developmental process	462	564	3683	21571	4.01e-114	0.45	0.15	3.1
5	Level3	GO:0044238	primary metabolic process	691	335	8164	17090	1.74e-111	0.67	0.32	2.1
6	Level2	GO:0065007	biological regulation	557	469	5705	19549	5.03e-102	0.54	0.23	2.4
7	Level2	GO:0050896	response to stimulus	432	594	3524	21730	1.49e-101	0.42	0.14	3.0
8	Level3	GO:0007154	cell communication	484	542	4781	20473	1.08e-88	0.47	0.19	2.5
9	Level2	GO:0032501	multicellular organismal process	462	564	4429	20825	1.83e-87	0.45	0.18	2.6
10	Level3	GO:0006950	response to stress	268	758	1572	23682	6.81e-85	0.26	0.06	4.2

Full result (HTML), (TXT) (EXCEL) Back to top

Figure 7. GO Biological Process category (Level 2~5)

- Level: level of GO terms. Clicking will show all the GO terms in the given level (Figure 8)
- GO ID: GO ID linking to all targets (both ROS-Diabetes and HUGO in Figure 9)
- GO Term: GO term linking to detail at EBI eGO web page (http://www.ebi.ac.uk/ego/)
- t+: the number of genes with the given GO terms in the ROS-Diabetes set
- t-: the number of genes without the given GO terms in the ROS-Diabetes set
- b+: the number of genes with the given GO terms in the full HUGO gene set
- b-: the number of genes without the given GO terms in the full HUGO gene set
- P-value: p-value from Fisher's exact test
- T-ratio: t+ / (t+ and t-)
- B-ratio: b+ / (b- and t-)
- Enrichment folds: t-ratio / b-ratio

#GOID	GOTerm	t-total	t+	t-	b+	b-	b-total	p-value	t-ratio	b-ratio	Enrichment
GO:0048522	positive regulation of cellular process	1026	237	789	1298	23956	25254	2.26E-79	0.23	0.05	4.5
GO:0012501	programmed cell death	1026	188	838	934	24320	25254	7.93E-68	0.18	0.04	5
GO:0019752	carboxylic acid metabolic process	1026	142	884	629	24625	25254	6.10E-56	0.14	0.02	5.6
GO:0043067	regulation of programmed cell death	1026	143	883	654	24600	25254	7.26E-55	0.14	0.03	5.4
GO:0048523	negative regulation of cellular process	1026	202	824	1352	23902	25254	3.83E-54	0.2	0.05	3.7
GO:0007242	intracellular signaling cascade	1026	226	800	1789	23465	25254	1.23E-49	0.22	0.07	3.1
GO:0042127	regulation of cell proliferation	1026	134	892	644	24610	25254	3.32E-49	0.13	0.03	5.1
GO:0051093	negative regulation of developmental process	1026	107	919	437	24817	25254	7.80E-45	0.1	0.02	6
GO:0006796	phosphate metabolic process	1026	175	851	1222	24032	25254	2.76E-44	0.17	0.05	3.5
GO:0044267	cellular protein metabolic process	1026	325	701	3588	21666	25254	1.11E-43	0.32	0.14	2.2
GO:0043412	biopolymer modification	1026	217	809	2096	23158	25254	3.26E-35	0.21	0.08	2.5

Figure 8. Level based Gene Ontology (Level 5)

Clicking each level in the GO summary table (Figure 7) will open an Excel file containing all the GO terms at the selected level.

GO:0012501 (ROSDiabetesFullAnalysis/GO/BP_SelectedLevel_DIR/25_Level5_GO 0012501_programmed cell dea.txt) Download Text

#CLASS HugoID Symbol NCBIGeneID Name
#TESTONLY: 0
#COMMON: 188
#BACKGROUNDONLY: 746

Class	HUGOID	Symbol	NCBI_Gene	Name
COMMON	959	BAX	581	BCL2-associated X protein
COMMON	8912	PHB	5245	prohibitin
COMMON	9666	PTPRC	5788	protein tyrosine phosphatase, receptor type, C
COMMON	10618	CCL2	6347	chemokine (C-C motif) ligand 2
COMMON	11180	SOD2	6648	superoxide dismutase 2, mitochondrial
COMMON	1785	CDKN1B	1027	cyclin-dependent kinase inhibitor 1B (p27, Kip1)
COMMON	634	AQP2	359	aquaporin 2 (collecting duct)
COMMON	3537	F2R	2149	coagulation factor II (thrombin) receptor
COMMON	9113	PML	5371	promyelocytic leukemia

Figure 9. An example of GO detail (GO:0012501)

Clicking each GO ID will bring up a table with all the genes annotated with the selected GO ID. The class “COMMON” refers to the targets belonging to the literature derived 1,026 ROS-Diabetes genes, while the class “BACKGROUNDONLY” refers to the targets that were not identified from the 1,154 ROS-Diabetes literature by SciMiner.

S1.2.2. Pathway

Pathway information from KEGG (http://www.genome.jp/kegg/) and Reactome (http://www.reactome.org/) was used in the analysis. Fisher’s exact test was used to identify significantly over-represented pathways in the ROS-Diabetes targets (Figure 10).

Top 10 most enriched Pathways (sorted by p-value)

# of targets in test set 1026 # of targets in background set 25254 Switch table detail

Rank	PathwayID	Title	t+	t-	b+	b-	p-value	t-ratio	b-ratio	enrichment folds
1	hsa04510	Focal adhesion - Homo sapiens (human)	75	951	197	25057	2.36e-42	0.07	0.01	9.4
2	hsa04210	Apoptosis - Homo sapiens (human)	49	977	83	25171	6.68e-35	0.05	0.00	14.5
3	hsa04010	MAPK signaling pathway - Homo sapiens (human)	73	953	259	24995	4.29e-34	0.07	0.01	6.9
4	109582	Hemostasis	62	964	172	25082	4.54e-34	0.06	0.01	8.9
5	hsa05215	Prostate cancer - Homo sapiens (human)	49	977	93	25161	3.78e-33	0.05	0.00	13.0
6	hsa04920	Adipocytokine signaling pathway - Homo sapiens (human)	45	981	72	25182	5.99e-33	0.04	0.00	15.4
7	hsa04910	Insulin signaling pathway - Homo sapiens (human)	53	973	137	25117	1.68e-30	0.05	0.01	9.5
8	hsa05212	Pancreatic cancer - Homo sapiens (human)	42	984	74	25180	1.42e-29	0.04	0.00	14.0
9	168256	Signaling in Immune system	72	954	312	24942	5.52e-29	0.07	0.01	5.7
10	hsa05222	Small cell lung cancer - Homo sapiens (human)	43	983	87	25167	2.28e-28	0.04	0.00	12.2

Full result (HTML, TXT, EXCEL) Back to top

Figure 10. Top 10 most over-represented pathways

S1.2.3. Protein-Protein Interaction (PPI) Network

This section shows how 1,026 ROS-Diabetes targets identified by SciMiner are closely related with respect to the topic (ROS-Diabetes). **100** sets of size **1,026** are randomly drawn from all HUGO genes. The PPI network of these 100 sets is compared with the PPI of the ROS-Diabetes target. The **red bar** in the figures (below) represents the ROS-Diabetes set.

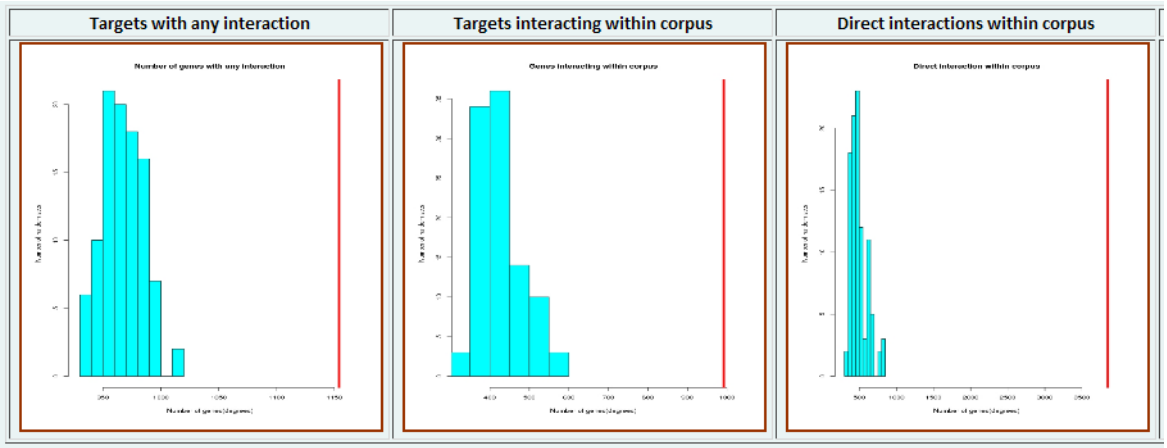


Figure 11. Protein-Protein Interaction network of 1,026 ROS-Diabetes targets

S1.3. 53 highly-overrepresented ROS-Diabetes targets

Please refer to S4.1.1. for more details.

S1.4. Enriched biological functions of 53 ROS-Diabetes targets

Please refer to S4.1.2. for more details.

S2. Search Menu

In ‘Search Menu’, users can search the literature derived ROS-Diabetes targets (either the full 1,026 and the core 53 targets) and predicted ROS-Diabetes targets by Gene ID (HUGO and Entrez Gene), symbol, name, and GO IDs/terms. Users can also create lists of putative ROS-Diabetes targets that share biological functions of the core 53 ROS-Diabetes targets in terms of GO terms and pathways. In addition to the over-represented biological functions of the 53 core ROS-Diabetes targets, users may also supply their own sets of GO/pathway IDs.

S2.1. Search predefined ROS-Diabetes targets

Figure 12 illustrates the first search section. Users can search against the following 5 different sets of ROS-Diabetes targets (two literature derived and three predicted sets). The predicted sets are based on statistically significant biological functions (Benjamini-Hochberg adjusted p-value < 0.05) of core 53 ROS-Diabetes targets.

- 1,026 ROS-Diabetes targets by SciMiner
- Highly over-represented 53 ROS-Diabetes targets
- Predicted targets with significant GO terms in ALL 3 categories
- Predicted targets with significant GO terms in ANY 3 categories
- Predicted targets with significant terms in ALL 3 GO categories + pathway

Clicking the ‘Search ROS-Diabetes targets’ button without choosing any filter will display all the available targets in a given set.

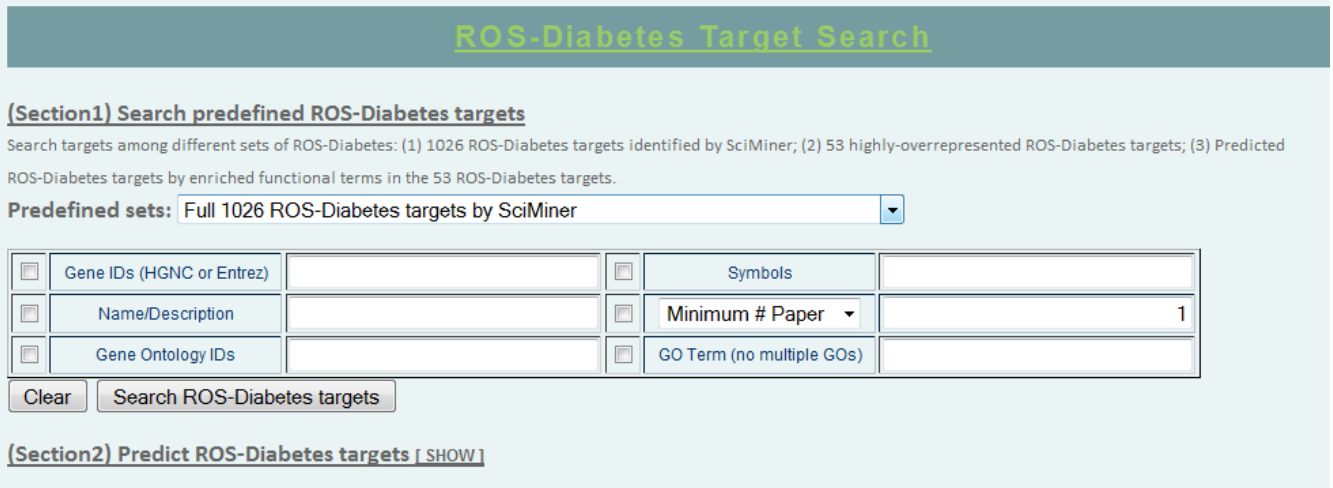


Figure 12. Search predefined ROS-Diabetes targets menu

S2.1.1. Predicted Target Summary Table

By clicking [Search ROS-Diabetes targets] a list of targets based on the user-provided filters in Figure 12 will be displayed in a summary table (Figure 13).

There are 3 possible 'Classes' for each target:

(1) CORE: This target belongs to the 53 core ROS-Diabetes set

(2) FULL: This target belongs to the 1,026 ROS-Diabetes set but NOT the 53 core set. This implies that this target was mentioned in the literature at least once, but the frequency is not significantly different from the background sets. This target might be an important ROS-Diabetes gene if it began to appear in the ROS-Diabetes literature only recently, as the number of papers mentioning this target could be relatively low.

(3) NOVEL: This target doesn't belong to the any of the literature derived ROS-Diabetes target sets. However, this target shares similar biological functions with the 53 core ROS-Diabetes set.

Predefined Set	Highly-overrepresented 53 ROS-Diabetes targets
Min # sig terms	10

Here are the targets (ROSDiabetes by SciMiner and/or by prediction). Download in Excel

Rank	HUGO	Symbol	Target Name	Class	# sig terms	MiMI
1	11179	SOD1	superoxide dismutase 1, soluble (amyotrophic lateral sclerosis 1 (adult))	CORE	57	
2	391	AKT1	v-akt murine thymoma viral oncogene homolog 1	CORE	56	
3	9801	RAC1	ras-related C3 botulinum toxin substrate 1 (rho family, small GTP binding protein Rac1)	CORE	40	
4	12680	VEGFA	vascular endothelial growth factor A	CORE	39	
5	672	RHOG	ras homolog gene family, member G (rho G)	NOVEL	37	
6	3176	EDN1	endothelin 1	CORE	36	
7	9393	PRKCA	protein kinase C, alpha	CORE	35	
8	8975	PIK3CA	phosphoinositide-3-kinase, catalytic, alpha polypeptide	FULL	30	
9	333	AGT	angiotensinogen (serpin peptidase inhibitor, clade A, member 8)	CORE	29	
10	7872	NOS1	nitric oxide synthase 1 (neuronal)	CORE	29	
11	8977	PIK3CD	phosphoinositide-3-kinase, catalytic, delta polypeptide	NOVEL	29	
12	7876	NOS3	nitric oxide synthase 3 (endothelial cell)	CORE	28	
13	21899	ATG9B	ATG9 autophagy related 9 homolog B (S. cerevisiae)	NOVEL	27	
14	7873	NOS2A	nitric oxide synthase 2A (inducible, hepatocytes)	CORE	27	
15	7989	NRAS	neuroblastoma RAS viral (v-ras) oncogene homolog	NOVEL	27	

Figure 13. A summary table for predicted ROS-Diabetes targets

GO/Pathway ID	GO/Pathway Terms
GO:0045859	regulation of protein kinase activity
GO:0046620	regulation of organ growth
GO:0006879	cellular iron ion homeostasis
GO:0008217	regulation of blood pressure
GO:0006801	superoxide metabolic process
GO:0005615	extracellular space
GO:0000303	response to superoxide
GO:0001890	placenta development

Figure 14. List of significant terms (truncated) by clicking '# sig terms' in the table in Figure 13

S2.2. Predict ROS-Diabetes targets

Users can also predict ROS-Diabetes by using either core 53 ROS-Diabetes genes or user-defined GO/Pathways. Clicking [SHOW], located to the right of **(Section2)** Predict ROS-Diabetes targets, will display a new query table as in Figure 15. **Details of this literature-based target prediction strategy are out of the scope of the current manuscript entitled "Literature-based discovery of diabetes- and ROS-related targets", but will be available in a separate paper (in preparation).**

If there are GO terms or pathways that are of specific interest, users can also run the prediction process using only these terms. GO IDs and pathway IDs (from KEGG or Reactome) can be used.

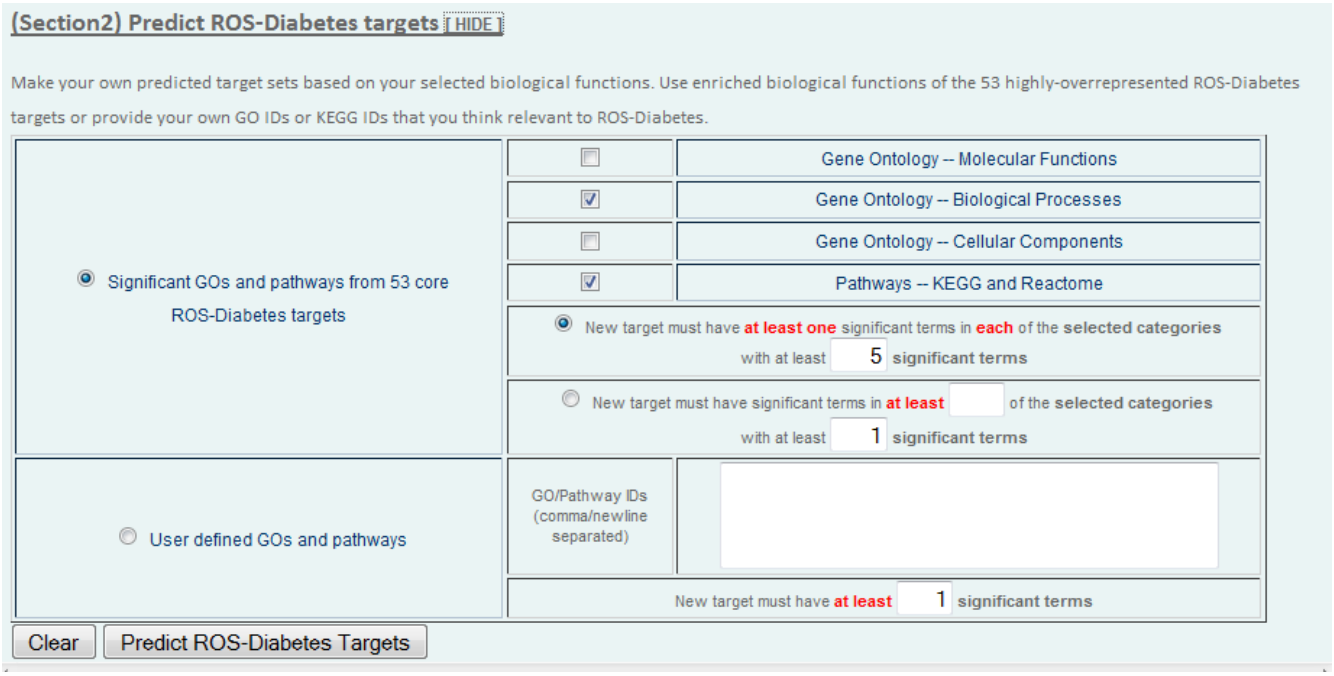
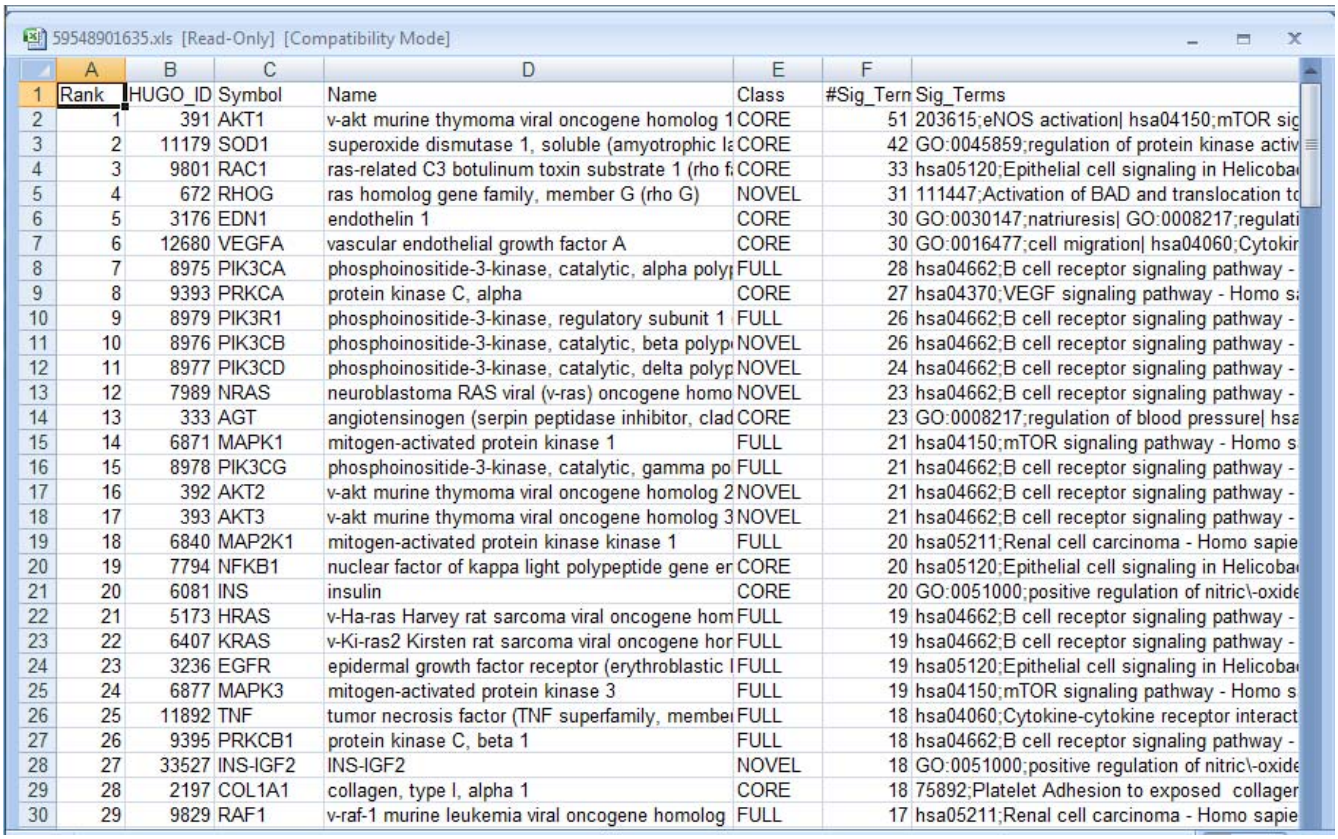


Figure 15. Menu for predicting novel ROS-Diabetes targets

In Figure 15, the prediction process has been set to use only biological processes GO terms and pathways. Each predicted target should have at least one significant term in each selected category (biological process and pathway) with a minimum of 5 significant terms in total. This prediction results in 378 targets with 195 novel targets. Figure 16 illustrates the top portion of the prediction results in an

Excel file.



59548901635.xls [Read-Only] [Compatibility Mode]

	A	B	C	D	E	F	
1	Rank	HUGO_ID	Symbol	Name	Class	#Sig_Tern	Sig_Terms
2	1	391	AKT1	v-akt murine thymoma viral oncogene homolog 1	CORE	51	203615;eNOS activation| hsa04150;mTOR sig
3	2	11179	SOD1	superoxide dismutase 1, soluble (amyotrophic la	CORE	42	GO:0045859;regulation of protein kinase activ
4	3	9801	RAC1	ras-related C3 botulinum toxin substrate 1 (rho f	CORE	33	hsa05120;Epithelial cell signaling in Helicoba
5	4	672	RHOG	ras homolog gene family, member G (rho G)	NOVEL	31	111447;Activation of BAD and translocation to
6	5	3176	EDN1	endothelin 1	CORE	30	GO:0030147;natriuresis| GO:0008217;regulati
7	6	12680	VEGFA	vascular endothelial growth factor A	CORE	30	GO:0016477;cell migration| hsa04060;Cytokin
8	7	8975	PIK3CA	phosphoinositide-3-kinase, catalytic, alpha polyp	FULL	28	hsa04662;B cell receptor signaling pathway -
9	8	9393	PRKCA	protein kinase C, alpha	CORE	27	hsa04370;VEGF signaling pathway - Homo sa
10	9	8979	PIK3R1	phosphoinositide-3-kinase, regulatory subunit 1	FULL	26	hsa04662;B cell receptor signaling pathway -
11	10	8976	PIK3CB	phosphoinositide-3-kinase, catalytic, beta polyp	NOVEL	26	hsa04662;B cell receptor signaling pathway -
12	11	8977	PIK3CD	phosphoinositide-3-kinase, catalytic, delta polyp	NOVEL	24	hsa04662;B cell receptor signaling pathway -
13	12	7989	NRAS	neuroblastoma RAS viral (v-ras) oncogene homo	NOVEL	23	hsa04662;B cell receptor signaling pathway -
14	13	333	AGT	angiotensinogen (serpin peptidase inhibitor, clad	CORE	23	GO:0008217;regulation of blood pressure| hsa
15	14	6871	MAPK1	mitogen-activated protein kinase 1	FULL	21	hsa04150;mTOR signaling pathway - Homo s
16	15	8978	PIK3CG	phosphoinositide-3-kinase, catalytic, gamma po	FULL	21	hsa04662;B cell receptor signaling pathway -
17	16	392	AKT2	v-akt murine thymoma viral oncogene homolog 2	NOVEL	21	hsa04662;B cell receptor signaling pathway -
18	17	393	AKT3	v-akt murine thymoma viral oncogene homolog 3	NOVEL	21	hsa04662;B cell receptor signaling pathway -
19	18	6840	MAP2K1	mitogen-activated protein kinase kinase 1	FULL	20	hsa05211;Renal cell carcinoma - Homo sapie
20	19	7794	NFKB1	nuclear factor of kappa light polypeptide gene er	CORE	20	hsa05120;Epithelial cell signaling in Helicoba
21	20	6081	INS	insulin	CORE	20	GO:0051000;positive regulation of nitric\-oxide
22	21	5173	HRAS	v-Ha-ras Harvey rat sarcoma viral oncogene hom	FULL	19	hsa04662;B cell receptor signaling pathway -
23	22	6407	KRAS	v-Ki-ras2 Kirsten rat sarcoma viral oncogene hor	FULL	19	hsa04662;B cell receptor signaling pathway -
24	23	3236	EGFR	epidermal growth factor receptor (erythroblastic l	FULL	19	hsa05120;Epithelial cell signaling in Helicoba
25	24	6877	MAPK3	mitogen-activated protein kinase 3	FULL	19	hsa04150;mTOR signaling pathway - Homo s
26	25	11892	TNF	tumor necrosis factor (TNF superfamily, membe	FULL	18	hsa04060;Cytokine-cytokine receptor interact
27	26	9395	PRKCB1	protein kinase C, beta 1	FULL	18	hsa04662;B cell receptor signaling pathway -
28	27	33527	INS-IGF2	INS-IGF2	NOVEL	18	GO:0051000;positive regulation of nitric\-oxide
29	28	2197	COL1A1	collagen, type I, alpha 1	CORE	18	75892;Platelet Adhesion to exposed collagen
30	29	9829	RAF1	v-raf-1 murine leukemia viral oncogene homolog	FULL	17	hsa05211;Renal cell carcinoma - Homo sapie

Figure 16. Custom prediction of ROS-Diabetes targets (in Excel file)

S3. Contact

Junguk Hur
Department of Neurology
University of Michigan
5380 Biomedical Science Research Building
109 Zina Pitcher Place
Ann Arbor, MI 48109
juhur@umich.edu
http://www.umich.edu/~juhur

Group Homepages

H. V. Jagadish Lab	**http://www.eecs.umich.edu/~jag/**
Eva Feldman Lab	**http://www.med.umich.edu/PNRD/index.html**
David States Lab	**http://www.stateslab.org/**