

Parallel Multi-Omics in High-Risk Subjects for the Identification of Integrated Biomarker Signatures of Type 1 Diabetes

Oscar Alcazar^{1,#}, Luis F. Hernandez^{2,#}, Ernesto S. Nakayasu³, Carrie D. Nicora⁴, Charles Ansong⁵, Michael J. Muehlbauer⁶, James R. Bain⁷, Ciara J. Myer⁸, Sanjoy K. Bhattacharya⁹, Peter Buchwald^{10,#,*}, and Midhat H. Abdulreda^{11,#,*}

¹ University of Miami Miller School of Medicine, Diabetes Research Institute, Miami, FL, USA; email: o.alcazar@miami.edu

² University of Miami Miller School of Medicine, Diabetes Research Institute, Miami, FL, USA; email: lfh34@miami.edu

³ Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA; email: ernesto.nakayasu@pnl.gov

⁴ Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA; email: carrie.nicora@pnl.gov

⁵ Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA; email: Charles.Ansong@pnl.gov

⁶ Duke Molecular Physiology Institute, Duke University Medical Center, Durham, NC, USA. email: michael.muehlbauer@duke.edu

⁷ Duke Molecular Physiology Institute, Duke University Medical Center, Durham, NC, USA. email: james.bain@duke.edu

⁸ University of Miami Miller School of Medicine, Department of Ophthalmology and Miami Integrative Metabolomics Research Center, Miami, FL, USA; email: cxm1496@miami.edu

⁹ University of Miami Miller School of Medicine, Department of Ophthalmology and Miami Integrative Metabolomics Research Center, Miami, FL, USA. email: sbhattacharya@miami.edu

¹⁰ University of Miami Miller School of Medicine, Diabetes Research Institute, and Department of Molecular and Cellular Pharmacology, Miami, FL, USA. email: pbuchwald@miami.edu

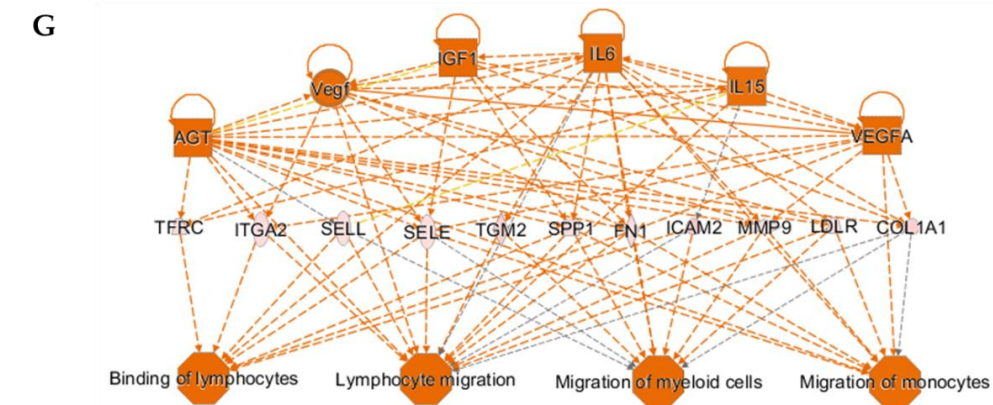
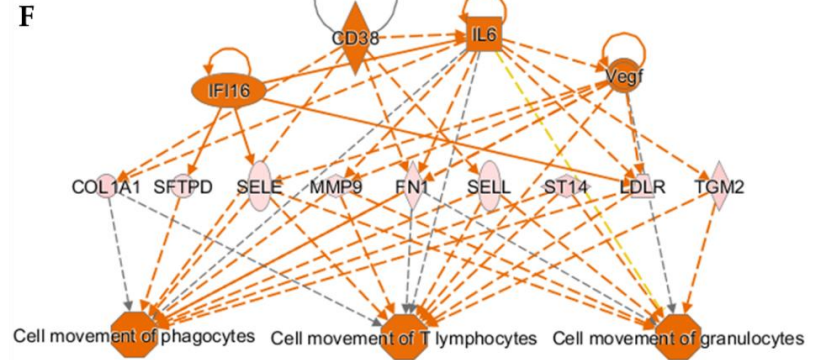
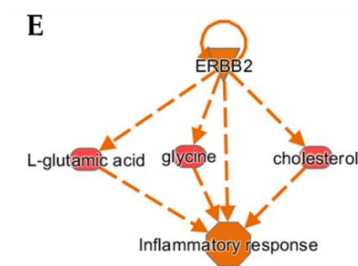
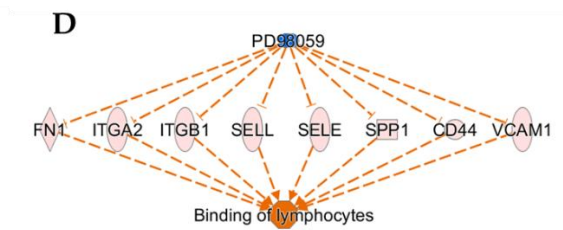
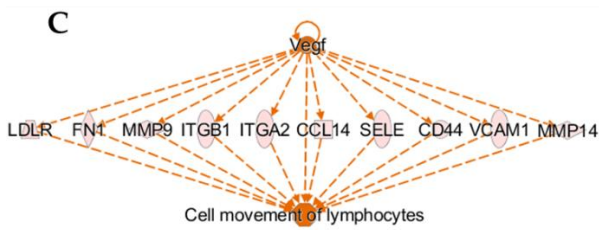
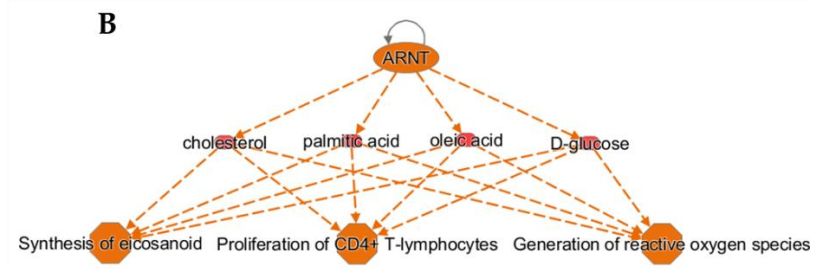
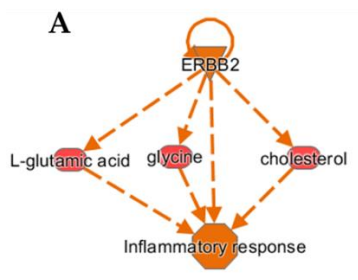
¹¹ University of Miami Miller School of Medicine, Diabetes Research Institute, Department of Surgery, Department of Microbiology and Immunology, and Department of Ophthalmology, Miami, FL, USA; email: mabdulreda@miami.edu

[#] These authors contributed equally

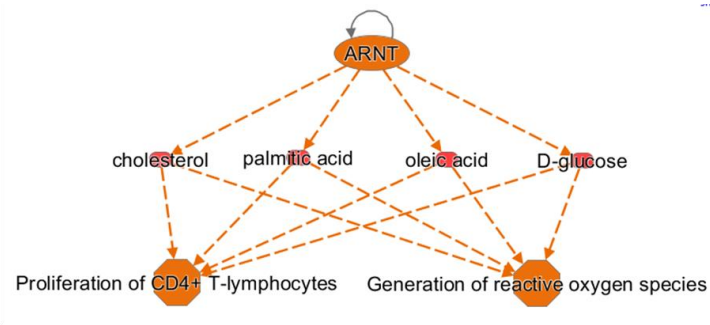
^{*} Correspondence emails mabdulreda@miami.edu or pbuchwald@miami.edu

Supplementary Materials Content

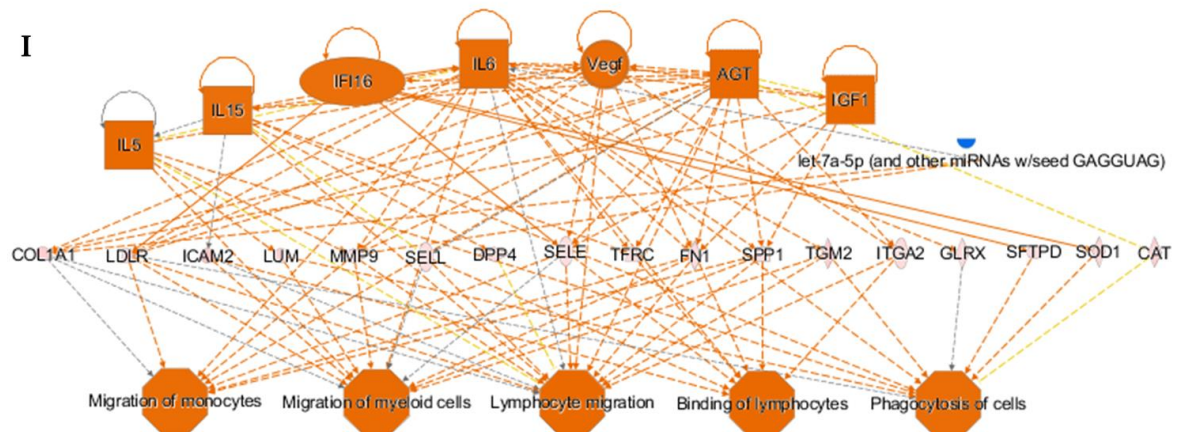
- **Figure S1.** Regulator Effects three-tier diagrams (networks) relevant to immune activation and inflammatory processes predicted at different analysis cutoff threshold values.
- **Figure S2.** High-resolution versions of the networks shown in **Figure 5**.
- **Figure S3.** Networks predicted at various levels of analysis stringency summarized in **Figure 6**.
- **Figure S4.** Scatter plot of p -values comparing (by t-test) the variance within the high-risk T1D group with and without the one new-onset subject.
- **Table S1.** Demographic, serological, and other information of subjects from whom blood samples were obtained in this study.
- **Table S2.** Transcriptomics data quality control information.
- **Table S3.** Complete list of candidate biomarkers predicted by *IPA* in the individual omics datasets without any curation in the context of the global molecular network predicted in the integrated multi-omics dataset at fold-change cutoff value of 1.1 during feature selection in the integrative analysis.



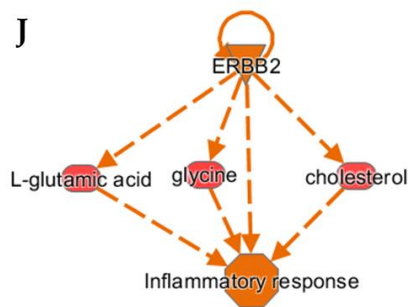
H



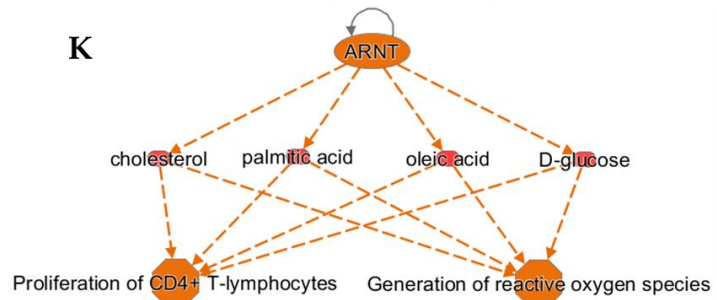
I



J



K



L

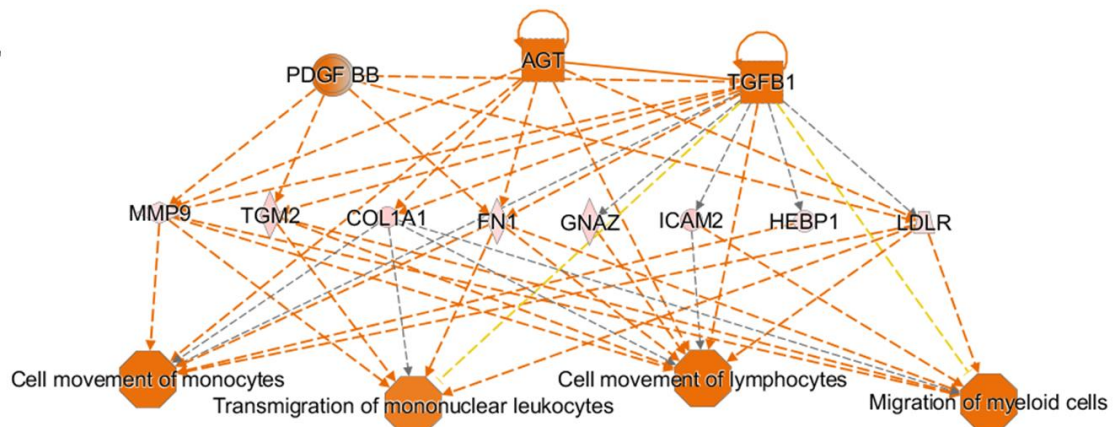
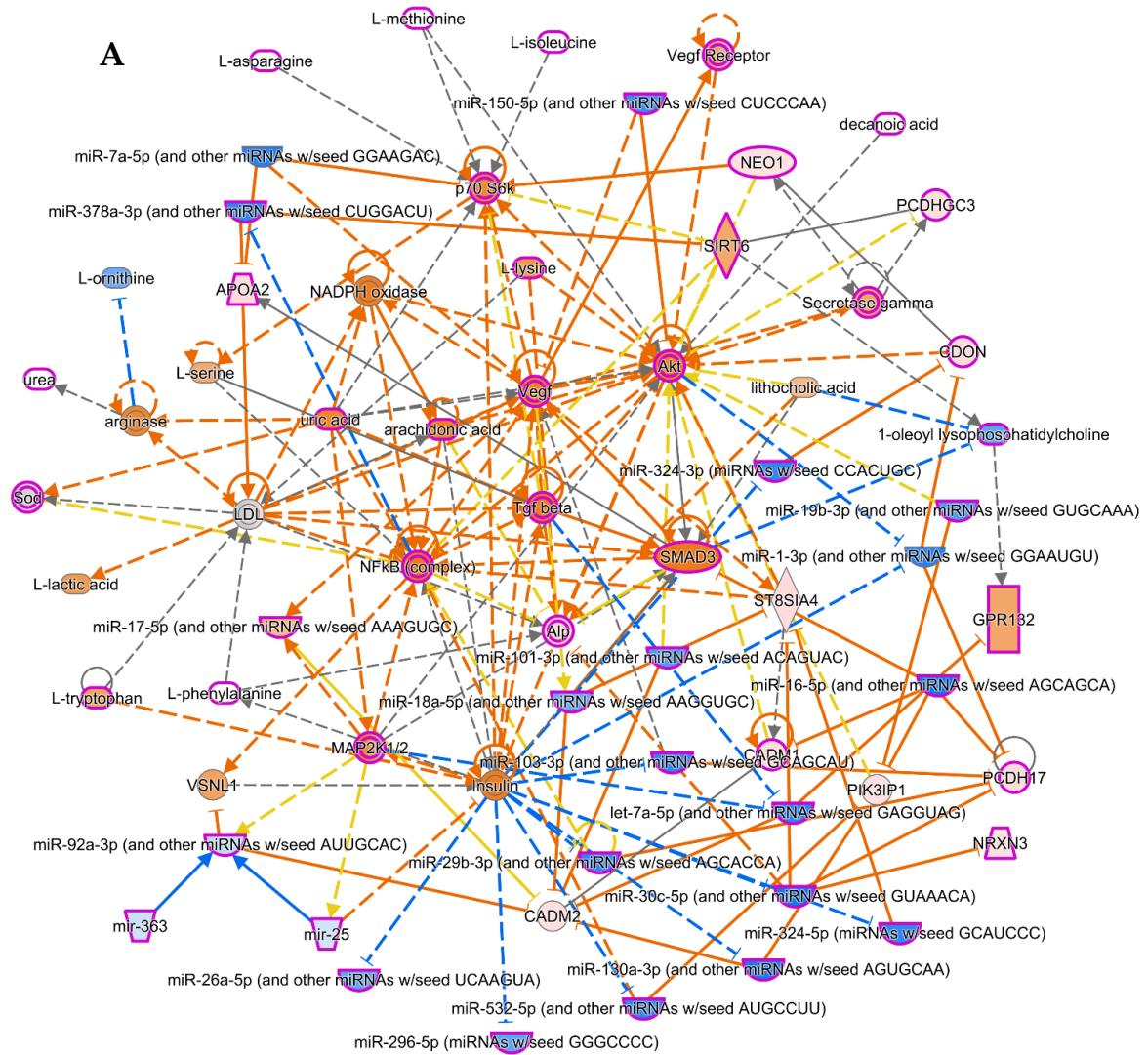
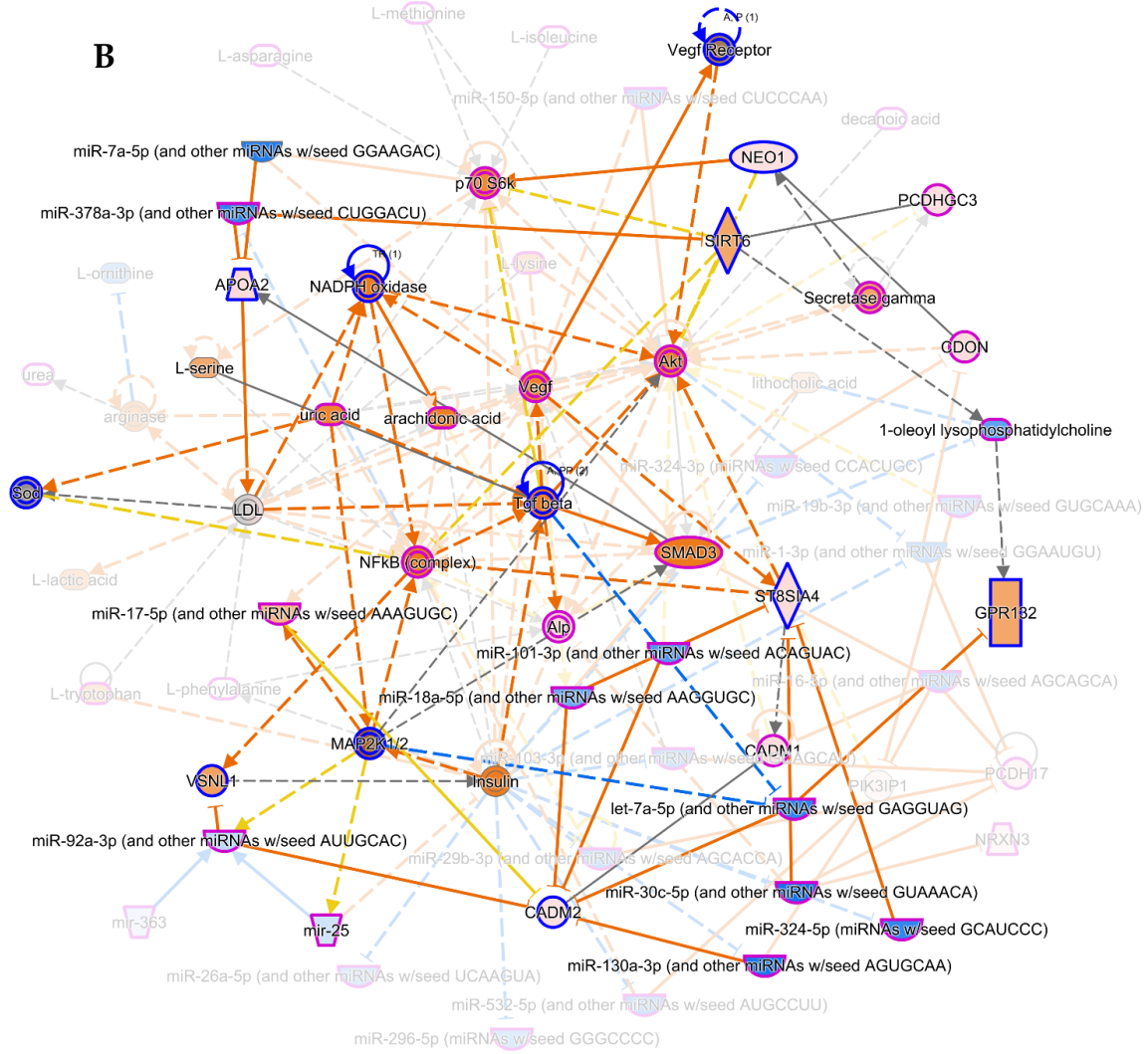


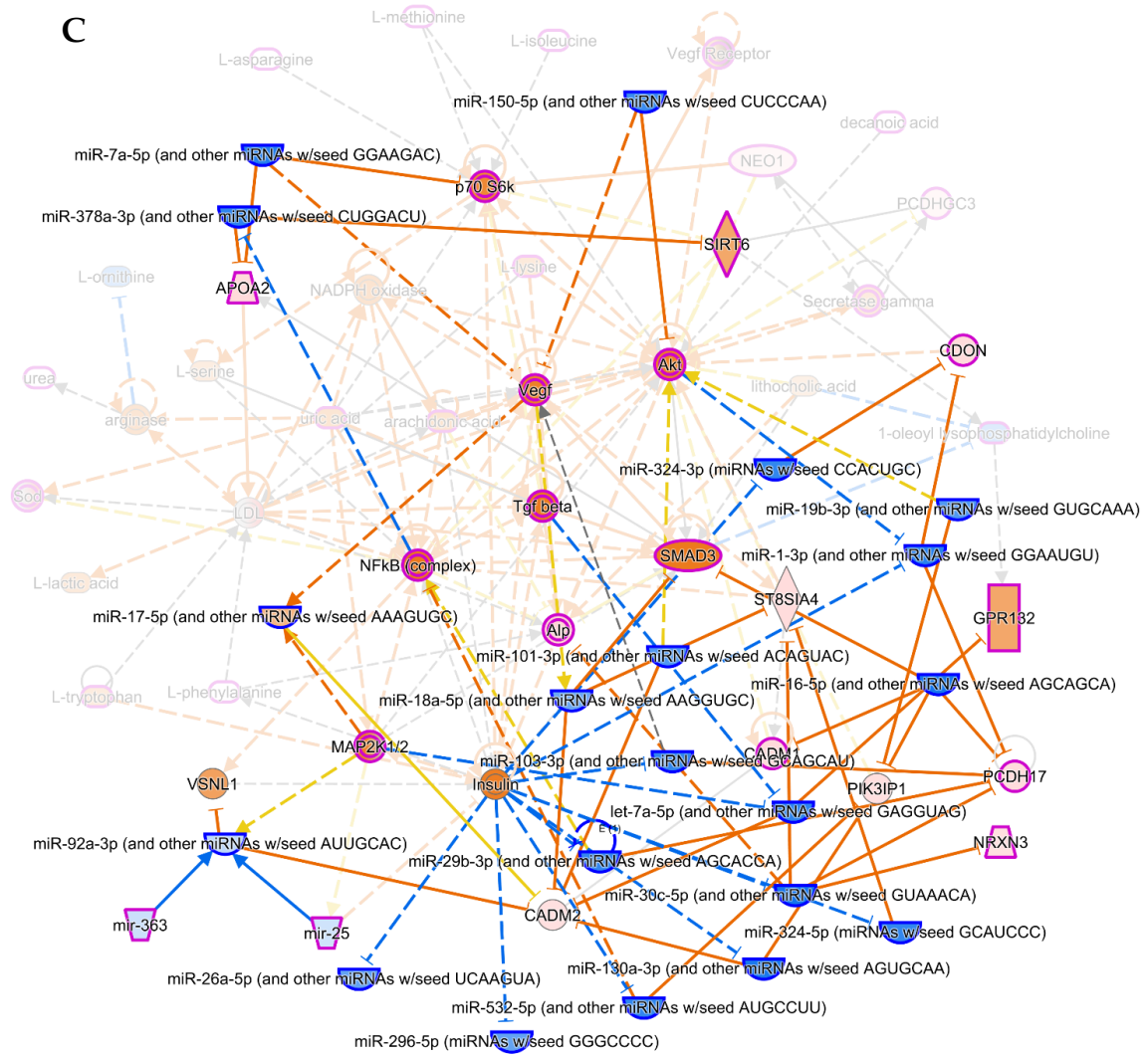
Figure S1 (corresponds to Figure 4). Regulator Effects three-tier diagrams (networks) predicted at the different feature selection thresholds (fold-change cutoff values). Predictions at cutoff values (A–D) 1.1, (E–I) 1.2, and (J–L) 1.3 using the “Regulator Effects” module of the IPA software. Analyses at the cutoff values 2.0 or 3.0 did not yield any predictions. Marker key – Triangle: kinase horizontal-oval: transcription regulator; vertical-oval: transmembrane receptor; diamond: enzyme; square: cytokine; horizontal ellipse: metabolite; circle: other; octagon: function. Color key – orange: predicted activation, blue: predicted inhibition. Connecting line color key – orange: activation; blue: inhibition; yellow: findings are inconsistent with the state of the downstream molecule; and gray: not predicted.

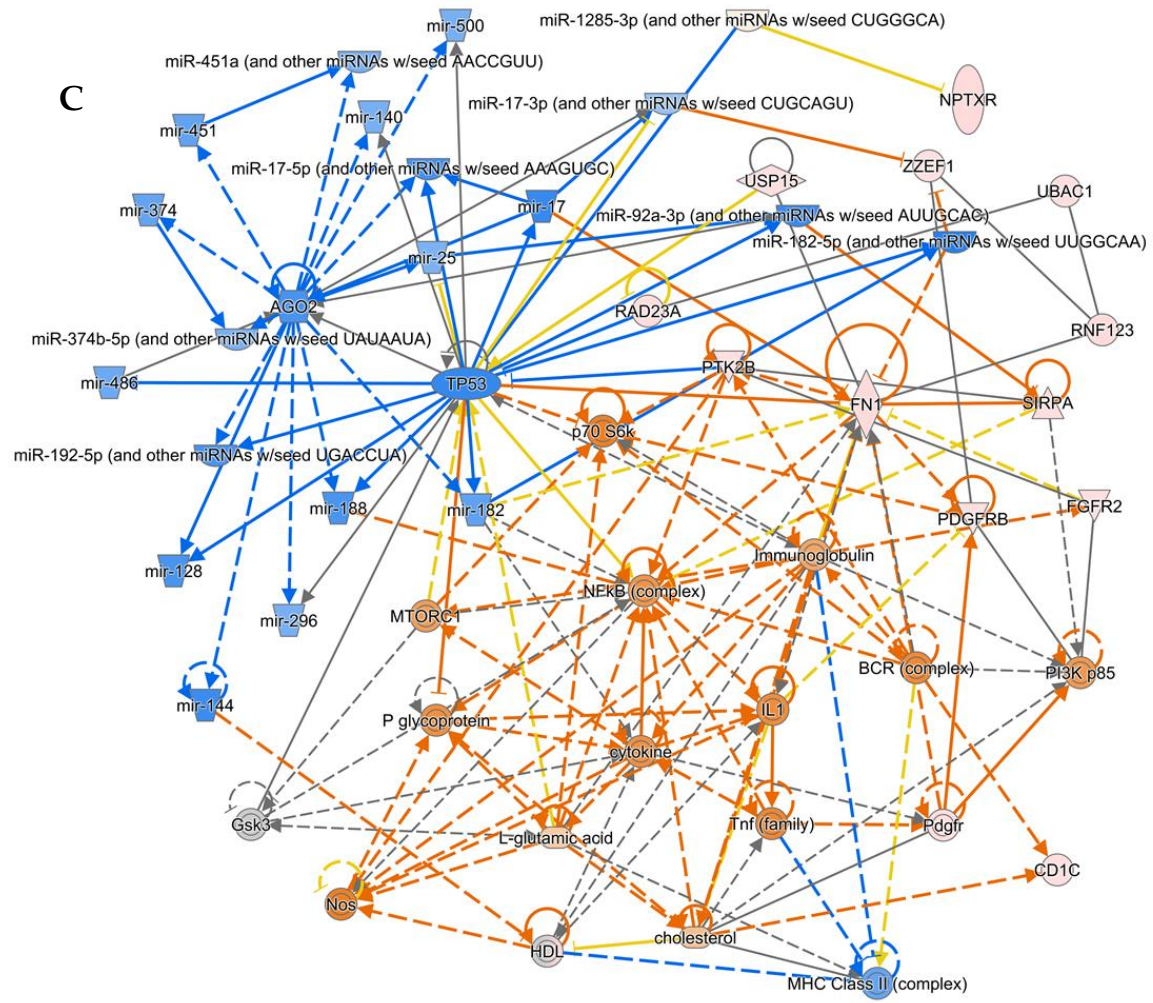


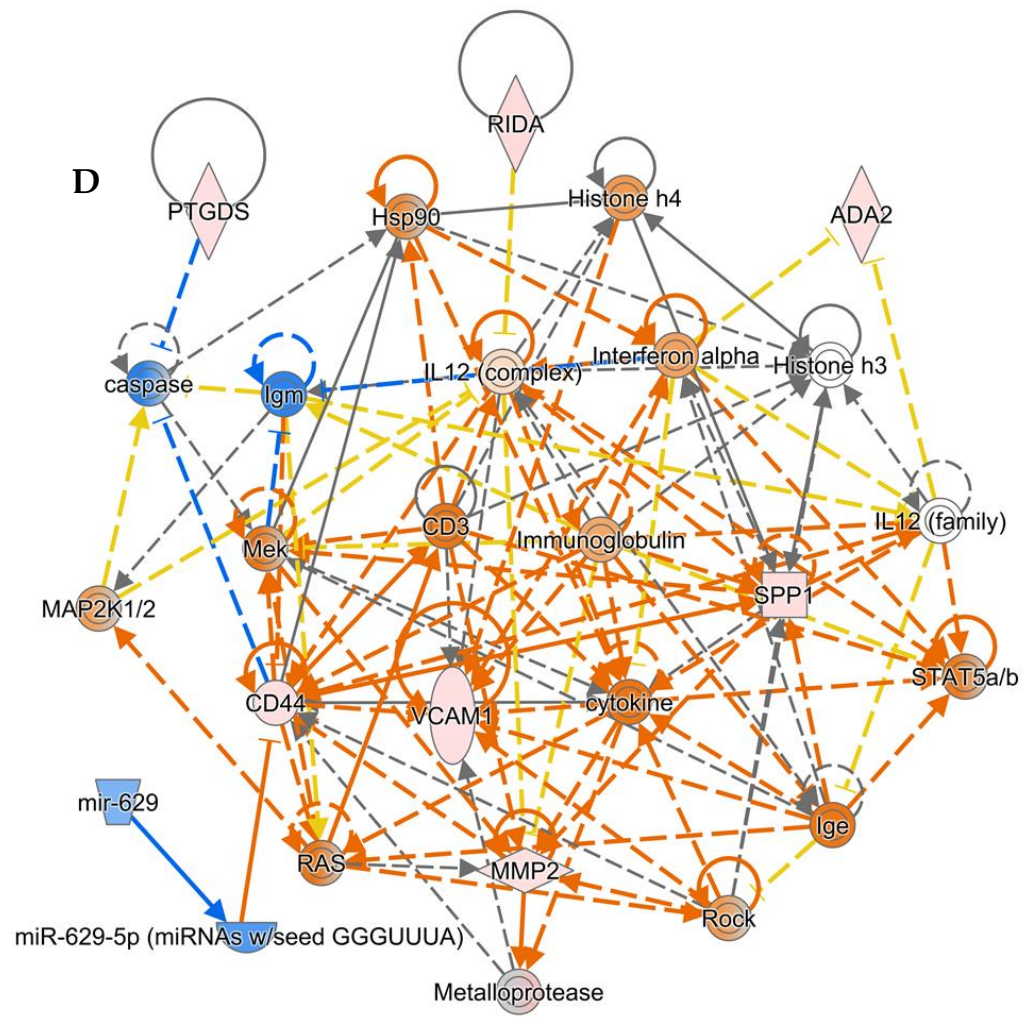
B

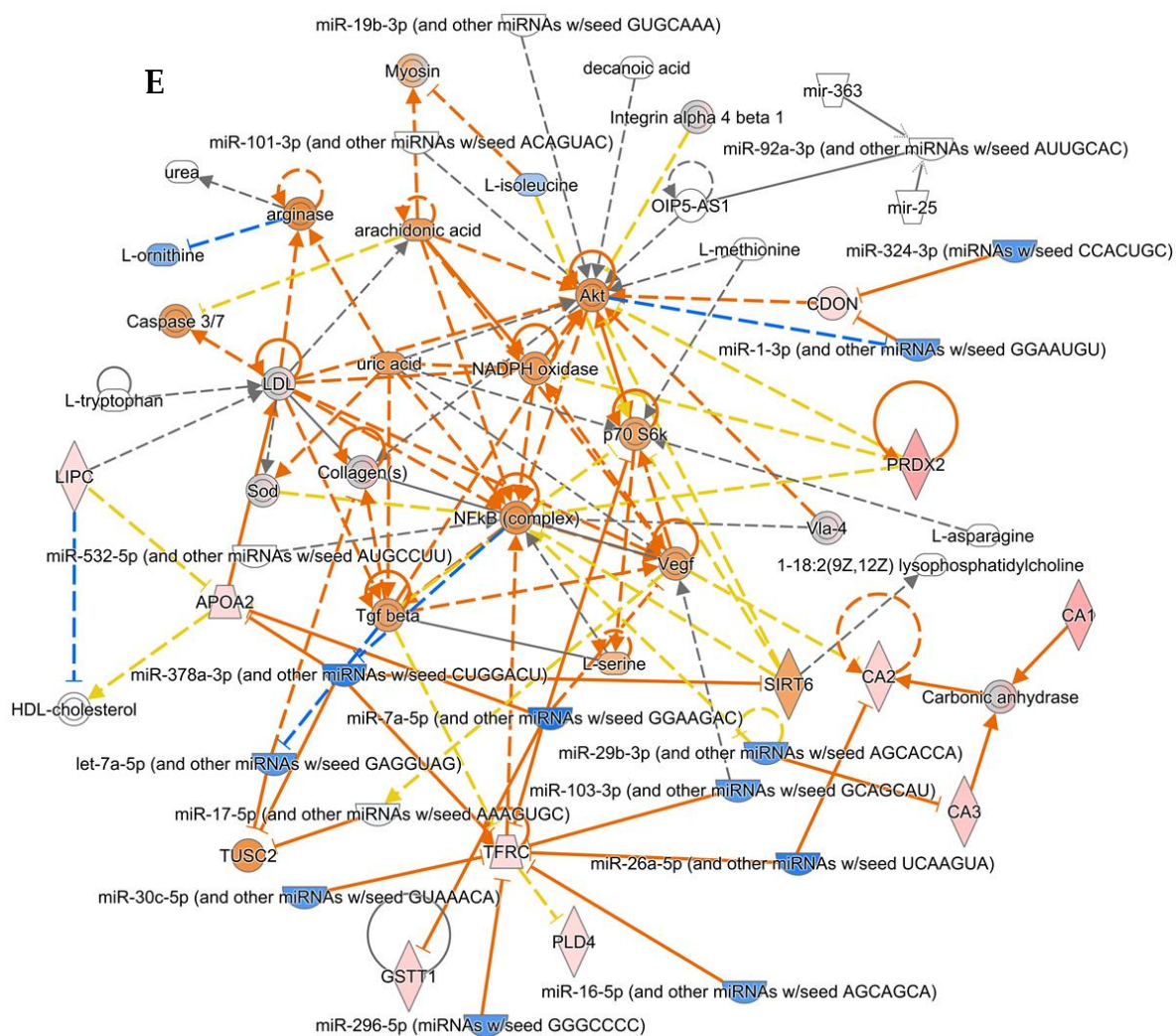


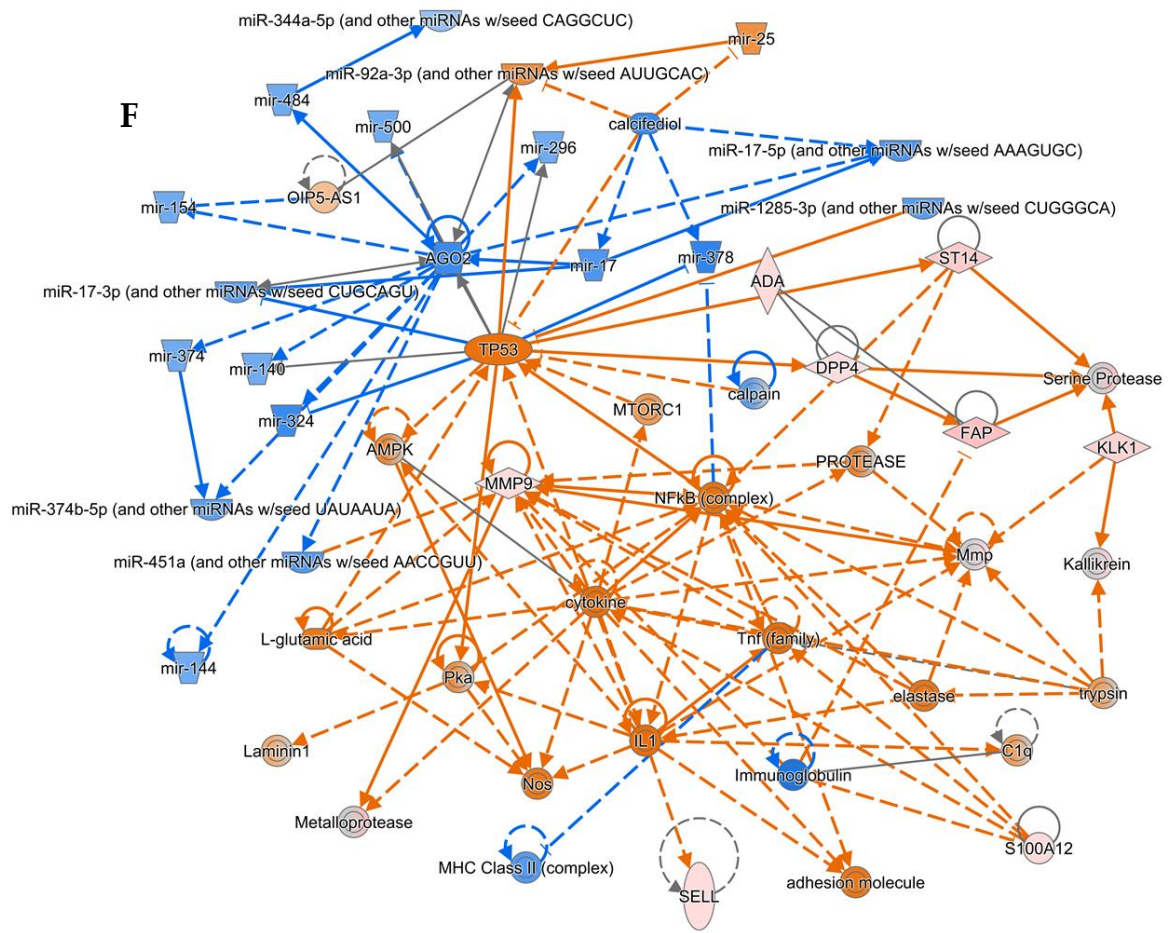
C

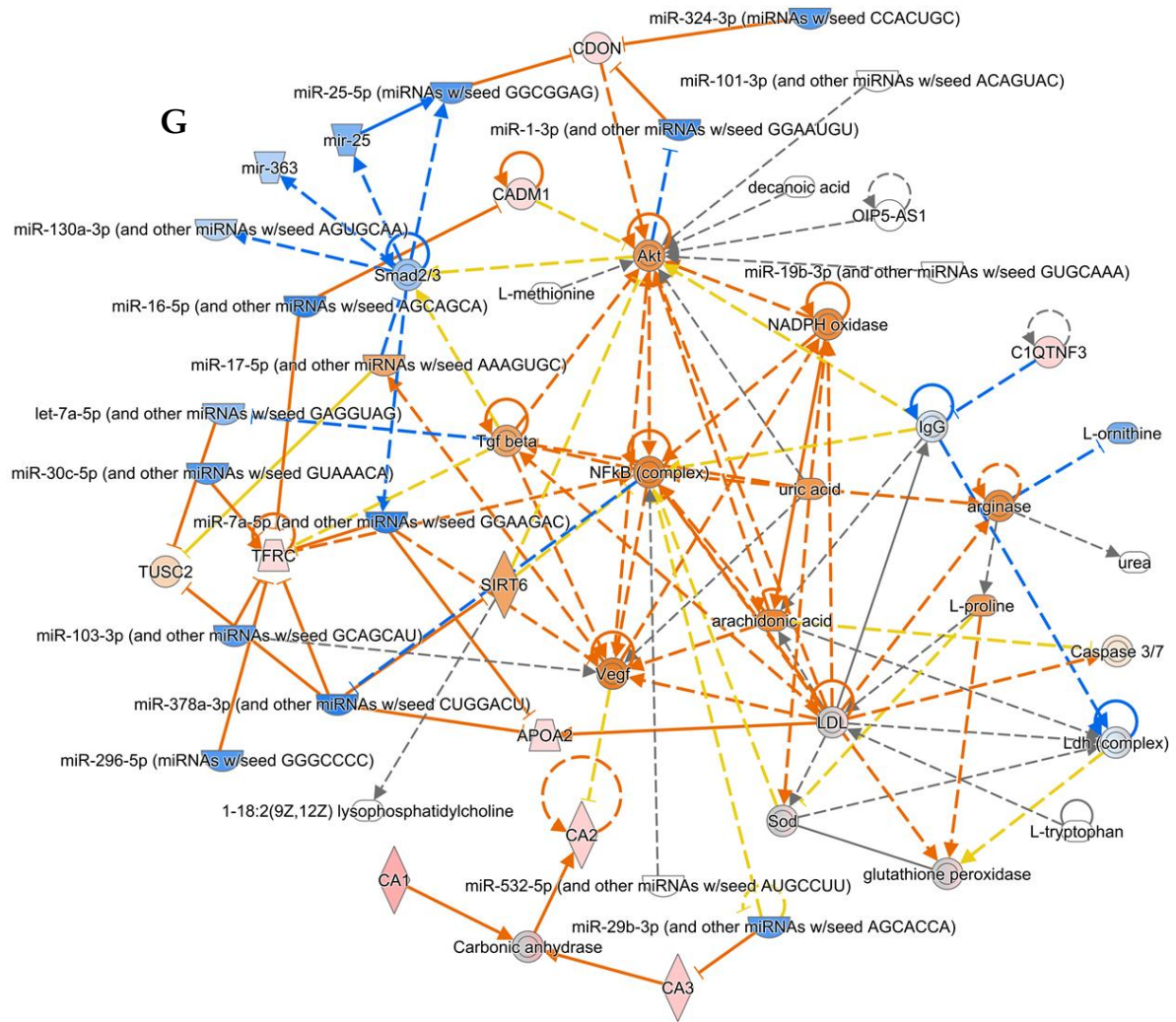


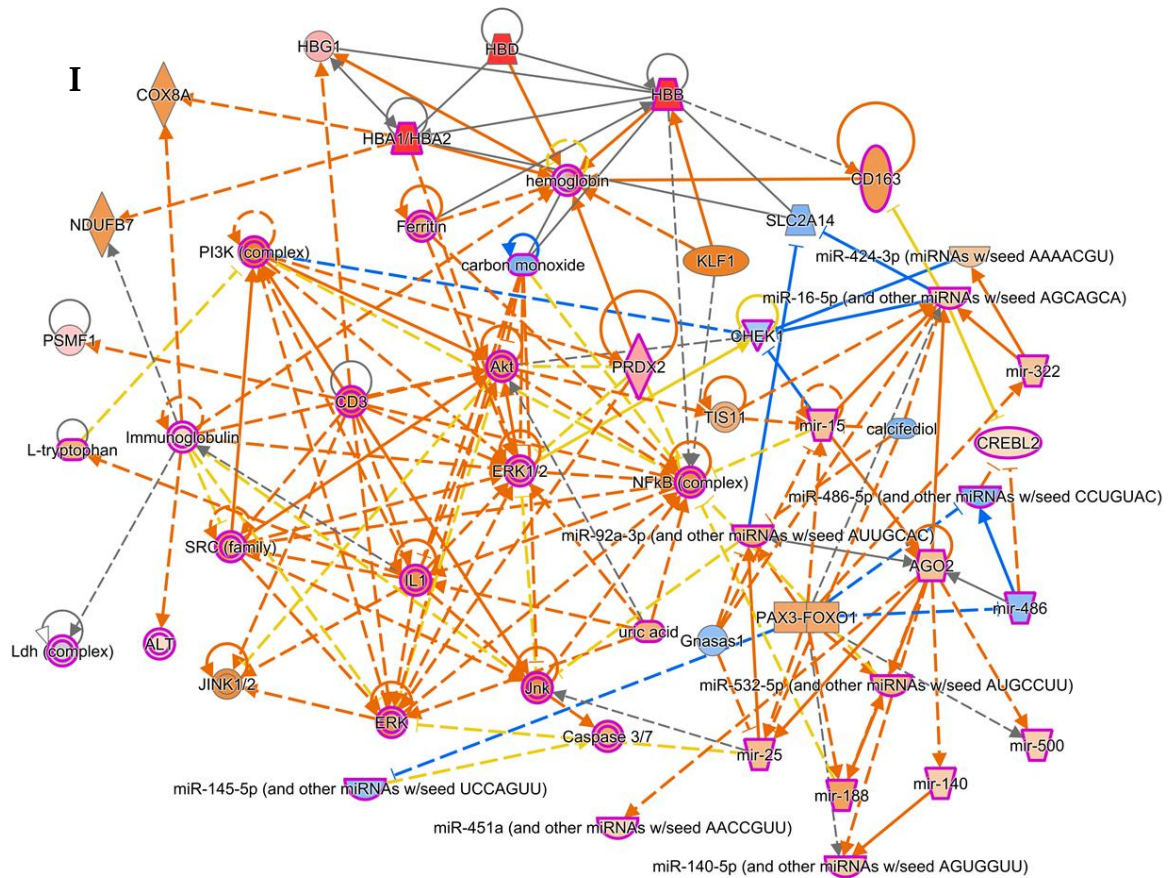
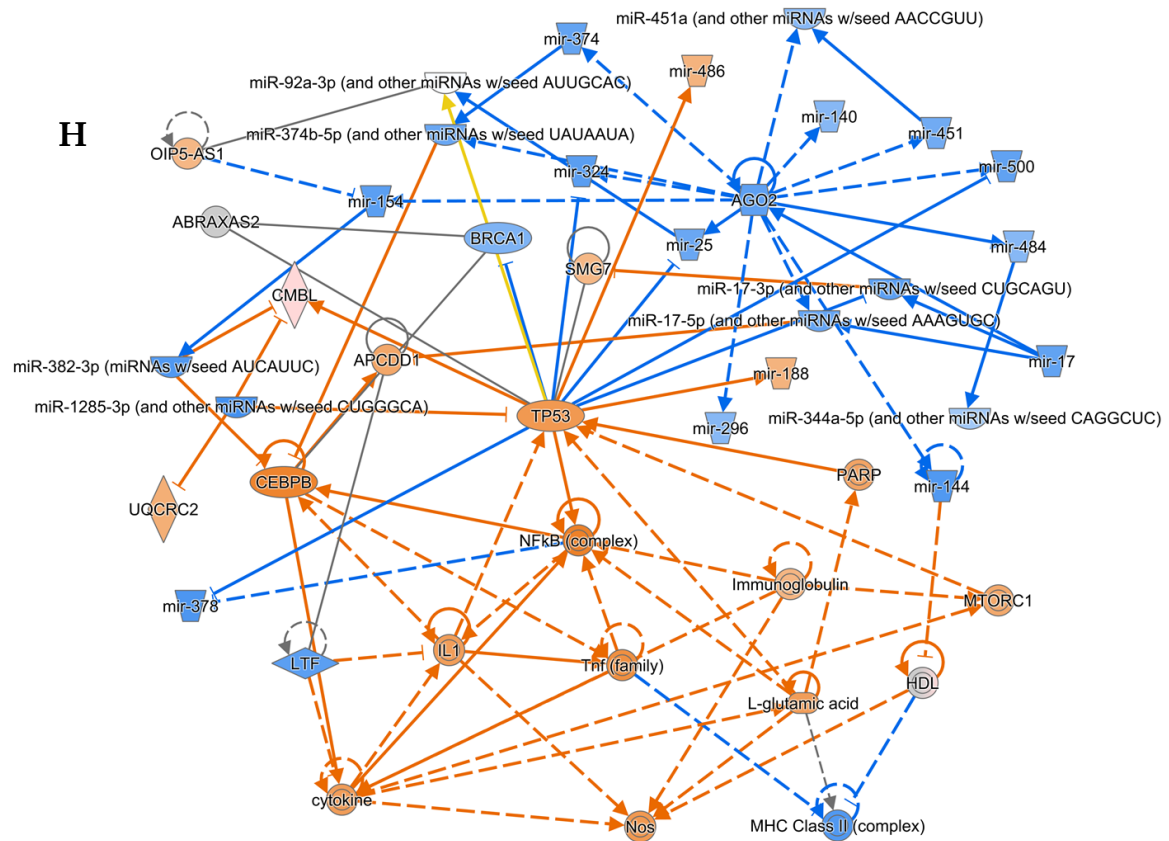












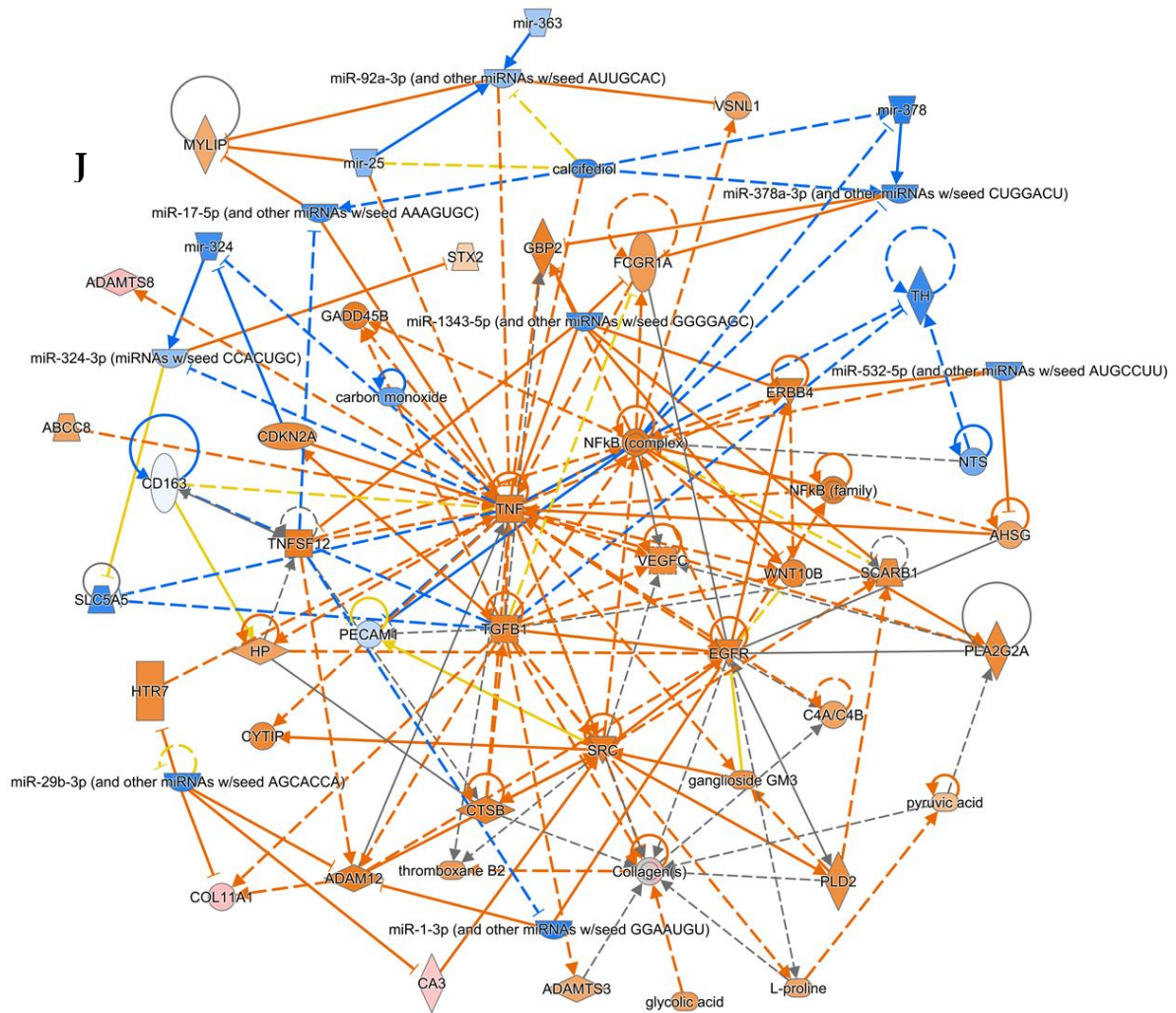


Figure S3 (corresponds to Figure 6). Networks predicted at various stringency levels of feature selection in the analyses shown in Figure 6. Networks predicted at the cutoff values of (A–D) 1.1, (E,F) 1.2, (G,H) 1.3, and (I,J) 2.0. See Figure 6 for a summary of key molecules consistently predicted at the different cutoff values in these analyses. Marker shape key – horizontal-oval: transcription regulator; vertical-oval: transmembrane receptor; diamond: enzyme; square: cytokine; vertical-rectangle: G-protein coupled receptor; broken-lined vertical-rectangle: ion channel; horizontal-diamond: peptidase; trapezoid: transporter; concentric circle: complex; circle: other; horizontal ellipse: metabolite; cut ellipsoid: microRNA; and inverted trapezoid: microRNA family. Marker color key – orange: predicted activation, blue: predicted inhibition. Connecting line color key – orange: activation; blue: inhibition; yellow: findings are inconsistent with the state of the downstream molecule; and gray: not predicted.

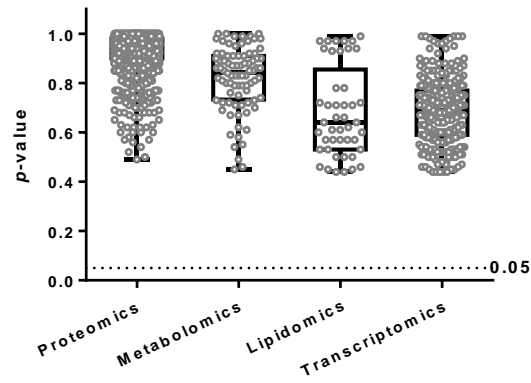


Figure S4. Variance within the high-risk T1D group with and without including the one new-onset subject. Shown are p -values of the comparison (by t -test) within the high-risk T1D group showing that no significant change ($p < 0.05$; dotted horizontal line) within the group is caused by inclusion of the one new-onset subject. Data points correspond to each analyte/feature identified by each omics approach in all samples without exception.

Table S1. Demographic, serological, and other information of subjects from whom blood samples were obtained in this study. Subjects are divided in two groups, namely the high-risk T1D and healthy (control) subjects, in accordance with the TrialNet staging system for T1D risk.

High-risk T1D subjects													
Subject #	Age (years)	Sex	Race/Ethnicity	1 st degree relatives with T1D	AAbs #	Autoantibodies (years since positive until sample collection)					HLA typing		
						GAD65/ GAD65H	ICA512/ IA-2H	mIAA	ICA	Znt8	DQA1*0120/DQB1*0602 protective allele	DQB1*0302	OGTT
1	13	M	White Caucasian	Yes	5	Positive (<1 year) ¹	Positive (<1 year) ¹	Positive (<1 year) ¹	Positive (<1 year) ¹	Positive (<1 year) ¹	ND	ND	Abnormal ²
2	14	M	White Caucasian	Yes	4 ³	Positive (9 years)	Positive (6 years)	Positive (9 years)	Positive (5 years)	Questionable ³	Negative	Positive	Normal
3	14	M	White Hispanic	Yes	5	Positive (9 years)	Positive (9 years)	Positive (12 years)	Positive (12 years)	Positive (3 years)	Negative	Negative	Normal
4	11	F	White Caucasian	Yes	5	Positive (6 years)	Positive (3 years)	Positive (5 years)	Positive (6 years)	Positive (6 years)	Negative	Positive	Normal
Healthy subjects													
1	22	F	White Caucasian	No	0	Negative	Negative	Negative	Negative	Negative	ND	ND	Normal
2	31	F	White Caucasian	No	0	Negative	Negative	Negative	Negative	Negative	ND	ND	Normal
3	19	M	White Hispanic	No	0	Negative	Negative	Negative	Negative	Negative	ND	ND	Normal
4	34	M	Asian	No	0	Negative	Negative	Negative	Negative	Negative	ND	ND	Normal

¹ Subject confirmed positive for all 5 autoantibodies (AABs) twice before sample collection for this study. 1st and 2nd positive confirmations were respectively 4 and 1 months prior to sample collection.

² Two samples collected for this study two weeks apart, 1st sample collected before OGTT (fasting) and 2nd after OGTT (post-prandial) when diagnosis was confirmed.

³ ZnT8 was positive in 2017 and then negative in 2018.

Table S2. Transcriptomics data quality control. All samples showed a minimum of 14.65 million passed-filter 75 nt single-end reads in the RNA sequencing. After removal of reads comprised substantially of adaptor sequences and trimming partial adapter sequences, the reads aligned at an average efficiency of 63% +/- 15% (SD) to the hg38 human reference genome and hg38 RNA database.

High-risk T1D subjects								
Subject #	PF Clusters	Yield (MB)	% Perfect Index Reads	% of \geq Q30 Bases (PF)	Mean Quality Score (PF)	Reads Passing All Filters	Reads Mapped to genome with Mismatch	Reads Mapped / Passed All Filters Read (%)
1*	16202916	1227	93.96%	91.39%	33.823	13878248	6415772	46.23%
1*	14646469	1109	92.76%	84.67%	32.474	12038011	8359355	69.44%
2	27600411	2093	93.24%	89.14%	33.384	23046222	16995316	73.74%
3	27696989	2100	94.04%	90.91%	33.730	22411557	17273116	77.07%
4	20079312	1521	90.63%	78.94%	31.271	17627203	15119578	85.77%
Healthy subjects								
Subject #	PF Clusters	Yield (MB)	% Perfect Index Reads	% of \geq Q30 Bases (PF)	Mean Quality Score (PF)	Reads Passing All Filters	Reads Mapped to genome with Mismatch	Reads Mapped / Passed All Filters Read (%)
1	18547614	1406	95.11%	90.25%	33.595	15982462	7903801	49.45%
2	32940064	2500	91.34%	88.95%	33.313	21716390	14897669	68.60%
3	17248712	1306	92.78%	90.37%	33.610	12379105	6222993	50.27%
4	23922839	1814	85.17%	86.55%	32.808	17548491	8120520	46.27%

* Two samples collected for this study from this subject two weeks apart, 1st sample collected before OGTT (fasting) and 2nd after OGTT (post-prandial) when diagnosis was confirmed

Table S3. Complete list of candidate biomarkers in high-risk T1D predicted by IPA in the individual quadra-omics datasets. Biomarker prediction by *IPA* was performed in the individual omics datasets with feature selection threshold at cutoff value of 1.1. The top 10 features (separated by a double line) were selected based on their involvement in central nodes within the global molecular network generated from the combined quadra-omics dataset (also shown in Table 2).

Feature	Accession no. ¹	Omics	Expression		IPA-predicted Function/Disorder	Link to T1D
			Fold-change ²	p-value ³		
let-7a-5p	MIMAT0000062	Transcriptomics	1.834	0.133	Inflammation of organ	
miR-130a-3p	MIMAT0000425	Transcriptomics	1.503	0.745	Chronic inflammatory disorder	
miR-16-5p	MIMAT0000069	Transcriptomics	2.954	0.151	Maturation of monocytes, chronic inflammatory disorder, inflammation of organ	[109]
miR-296-5p	MIMAT0000690	Transcriptomics	1.444	0.696	Inflammation of organ	[110]
miR-30c-5p	MIMAT0000244	Transcriptomics	1.39	0.326	Chronic inflammatory disorder, inflammation of organ	
miR-324-3p	MIMAT0000762	Transcriptomics	2.682	0.468	Inflammation of organ	[108]
miR-532-5p	MIMAT0002888	Transcriptomics	2.484	0.122	Chronic inflammatory disorder, inflammation of organ	
miR-92a-3p	MIMAT0000092	Transcriptomics	21.029	0.146	Induction of follicular T helper precursor cells, chronic inflammatory disorder, inflammation of organ	
Uric acid	69-93-2	Metabolomics	2.008	0.872	Activation of macrophages, immune response of cells, induction of peripheral blood lymphocytes, inflammation of organ	[111]
LPC 1-18:1(11Z)		Lipidomics	1.549	0.00949	None	
HBA1/HBA2	P69905	Proteomics	54.293	0.383	Chronic inflammatory disorder, diabetes mellitus, inflammation of organ, systemic autoimmune syndrome	
miR-3131	MIMAT0014996	Transcriptomics	44.427	0.125	Inflammation of organ	
miR-150-3p	MIMAT0004610	Transcriptomics	5.17	0.181	Chronic inflammatory disorder, inflammation of organ	[115]
miR-486-5p	MIMAT0002177	Transcriptomics	3.268	0.00272	Inflammation of organ	[116]
miR-629-5p	MIMAT0004810	Transcriptomics	2.625	0.0861	Inflammation of organ	
miR-500a-3p	MIMAT0002871	Transcriptomics	2.144	0.561	Inflammation of organ	
Citric acid	5949-29-1	Metabolomics	1.997	0.959	Accumulation of pyruvic acid, inflammation of organ	
D-glucose	50-99-7	Metabolomics	1.997	0.636	Accumulation of pyruvic acid, activation of macrophages, cell viability of islet cells, immune response of cells, proliferation of CD4+ T-lymphocytes, release of arachidonic acid, inflammation of organ	
Urea	57-13-6	Metabolomics	1.986	0.444	Inflammation of organ	
Cholesterol	57-88-5	Metabolomics	1.964	0.153	Activation of macrophages, cell spreading of transitional B lymphocytes, chemotaxis of pro-B lymphocytes, immune response of cells, inflammation of organ, proliferation of CD4+ T-lymphocytes, release of arachidonic acid, ruffling of transitional B lymphocytes, inflammation of organ, systemic inflammatory response syndrome	[117]

COL1A1	P02452	Proteomics	1.811	0.306	Cell movement of T lymphocytes, transmigration of lymphocytes, transmigration of monocytes, chronic inflammatory disorder, diabetes mellitus, diabetic complication, inflammation of organ, systemic autoimmune syndrome	
miR-1285-3p	MIMAT0005876	Transcriptomics	1.753	0.57	Inflammation of organ	
miR-344a-5p	MIMAT0002174	Transcriptomics	1.748	0.191	Inflammation of organ	
CA2	P54289	Proteomics	1.678	0.39	Chronic inflammatory disorder, diabetes mellitus, diabetic complication, inflammation of organ, systemic autoimmune syndrome	[118]
CAT	<u>P04040</u>	Proteomics	1.669	0.412	Diabetes mellitus, inflammation of organ, systemic autoimmune syndrome	[119]
CDH11	<u>P55287</u>	Proteomics	1.464	0.422	Inflammation of organ, chronic inflammatory disorder, diabetes mellitus, diabetic complication, inflammation of organ, systemic autoimmune syndrome	
ALDH1A1	P00352	Proteomics	1.446	0.47	Diabetes mellitus	[120]
LDLR	P01130	Proteomics	1.443	0.532	Cell movement of T lymphocytes, diabetes mellitus, inflammation of organ, quantity of IL-10 in blood, transmigration of lymphocytes, transmigration of monocytes, diabetic complication	[121]
PC 18:2(9Z,12Z)/ 18:2(9Z,12Z)		Lipidomics	1.387	0.186	None	
TFRC	P02786	Proteomics	1.358	0.584	Inflammation of organ, chronic inflammatory disorder, systemic autoimmune syndrome	[122]
FN1	P02751	Proteomics	1.317	0.553	Cell movement of T lymphocytes, transmigration of lymphocytes, transmigration of monocytes, chronic inflammatory disorder, diabetes mellitus, diabetic complication, inflammation of organ, systemic autoimmune syndrome	[123]
BGLAP	P02818	Proteomics	1.311	0.58	Chronic inflammatory disorder, diabetes mellitus, systemic autoimmune syndrome	[124]
MMP9	P14780	Proteomics	1.307	0.582	Cell movement of T lymphocytes, chronic inflammatory disorder, inflammation of organ, quantity of IL-10 in blood, transmigration of monocytes, diabetes mellitus, systemic autoimmune syndrome	[125]

¹ Accession numbers corresponding to mature miRNAs in the miRbase database, proteins in the human SwissProt database (from Uniprot Knowledgebase), and metabolites in the American Chemical Society (CAS = Chemical Abstract Service). Lipids were annotated in the Thermo Scientific proprietary Lipid Search 4.1 database built into their software without access to accession numbers.

² Predicted expression fold-change in the high-risk T1D subjects relative to healthy controls.

³ Non-adjusted *p*-value as obtained from IPA software based on comparisons of high-risk T1D subjects vs. healthy controls in each omics-type dataset independently (*t*-test). Only LPC from the lipidomics and miR-486-5p from transcriptomics datasets had *p* < 0.05.