Supplementary materials for Integrating sequence, expression and interaction data to determine condition-specific miRNA regulation

Hai-Son Le[1] and Ziv Bar-Joseph[1,2]

[1] Machine Learning Department, [2] Lane Center for Computational Biology
Carnegie Mellon University
Pittsburgh, PA, USA
{hple,zivbj}@cs.cmu.edu

1 Solving the optimization problem (4)

As discussed in Sect. 2.6, we solve the optimization (4) by an iterative procedure. We show here how to compute the gradients of the objective function, which are required for using `minPQN`. We begin with the objective function:

$$\mathcal{F} = -\log p(\mathbf{Y}|\mathbf{U},\mathbf{V},\mathbf{X},\mu,\mathbf{\Sigma}) - \sum_{i,j}\log p(\phi_{ij}|\mathbf{U},\mathbf{V}) - \sum_{j\neq j'}\log p(\omega_{jj'}\neq 0|\mathbf{V}) \tag{1}$$

The first term can be expanded to :

$$\begin{aligned}
&\log p(\mathbf{Y}|\mathbf{U},\mathbf{V},\mathbf{X},\boldsymbol{\mu},\mathbf{\Sigma})\\
=&\sum_j \log\mathcal{N}(\mathbf{y}_j \mid \boldsymbol{\mu} - \mathbf{X}^{\mathrm{T}}\big((\mathbf{I}_{\mathbf{\Phi}})_{,j}\circ(\mathbf{U}\mathbf{v}_j)\big),\mathbf{\Sigma})\\
=&-\sum_j\sum_p \log(\sqrt{2\pi}\sigma_p) - \frac{1}{2}\Big(\boldsymbol{\mu} - \mathbf{X}^{\mathrm{T}}\big((\mathbf{I}_{\mathbf{\Phi}})_{,j}\circ(\mathbf{U}\mathbf{v}_j)\big) - \mathbf{y}_j\Big)^{\mathrm{T}}\mathbf{\Sigma}^{-1}\Big(\boldsymbol{\mu} - \mathbf{X}^{\mathrm{T}}\big((\mathbf{I}_{\mathbf{\Phi}})_{,j}\circ(\mathbf{U}\mathbf{v}_j)\big) - \mathbf{y}_j\Big)
\end{aligned}$$

(We abuse the notation in $(\boldsymbol{\mu}^{\mathrm{T}} - \mathbf{Y})$ a bit. When subtracting a matrix from a row vector, we need to vertically replicate the row vector.)

$$\begin{aligned}
=&-N\sum_p\log(\sqrt{2\pi}\sigma_p) - \frac{1}{2}\operatorname{Tr}\{(\boldsymbol{\mu}^{\mathrm{T}} - \mathbf{Y})\mathbf{\Sigma}^{-1}(\boldsymbol{\mu}^{\mathrm{T}} - \mathbf{Y})^{\mathrm{T}}\}\\
&+\operatorname{Tr}\{(\boldsymbol{\mu}^{\mathrm{T}} - \mathbf{Y})\mathbf{\Sigma}^{-1}\big((\mathbf{V}\mathbf{U}^{\mathrm{T}}\circ\mathbf{I}_{\mathbf{\Phi}}^{\mathrm{T}})\mathbf{X}\big)^{\mathrm{T}}\} - \frac{1}{2}\operatorname{Tr}\{\big((\mathbf{V}\mathbf{U}^{\mathrm{T}}\circ\mathbf{I}_{\mathbf{\Phi}}^{\mathrm{T}})\mathbf{X}\big)\mathbf{\Sigma}^{-1}\big((\mathbf{V}\mathbf{U}^{\mathrm{T}}\circ\mathbf{I}_{\mathbf{\Phi}}^{\mathrm{T}})\mathbf{X}\big)^{\mathrm{T}}\}
\end{aligned}$$

Define:

$$\begin{aligned}
\mathbf{\Phi}_* &= \mathbf{\Phi} + (1 - \mathbf{I}_{\mathbf{\Phi}})\\
\mathbf{\Omega}_* &= \mathbf{\Omega} + (1 - \mathbf{I}_{\mathbf{\Omega}})
\end{aligned}$$

(basically replacing the zero entries with ones.)

We take the derivatives:

$$\frac{\partial \mathcal{F}}{\partial \mathbf{U}} = \underbrace{-\big(\mathbf{X}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}^{\mathrm{T}} - \mathbf{Y})^{\mathrm{T}}\big) \circ \mathbf{I}_{\boldsymbol{\Phi}}\mathbf{V} + \mathbf{X}\boldsymbol{\Sigma}^{-1}\mathbf{X}^{\mathrm{T}}\big((\mathbf{U}\mathbf{V}^{\mathrm{T}}) \circ \mathbf{I}_{\boldsymbol{\Phi}}\big) \circ \mathbf{I}_{\boldsymbol{\Phi}}\mathbf{V}}_{\text{from } \partial \log p(\mathbf{Y}|\mathbf{U},\mathbf{V},\mathbf{X},\boldsymbol{\mu},\boldsymbol{\Sigma})} + \underbrace{\alpha(\sigma^{\boldsymbol{\Phi}} - \mathbf{I}_{\boldsymbol{\Phi}}) \circ \boldsymbol{\Phi}_*\mathbf{V}}_{\text{from } \partial \sum_{i,j} \log p(I_{\phi_{ij}}|\mathbf{U},\mathbf{V})}$$

where

$$\sigma^{\boldsymbol{\Phi}} = \sigma\big(\alpha\ \boldsymbol{\Phi}_* \circ (\mathbf{U}\mathbf{V}^{\mathrm{T}})\big)$$

$$\frac{\partial \mathcal{F}}{\partial \mathbf{V}} = \underbrace{-\big((\boldsymbol{\mu}^{\mathrm{T}} - \mathbf{Y})\boldsymbol{\Sigma}^{-1}\mathbf{X}^{\mathrm{T}}\big) \circ \mathbf{I}_{\boldsymbol{\Phi}}^{\mathrm{T}}\mathbf{U} + \Big(\big((\mathbf{U}\mathbf{V}^{\mathrm{T}}) \circ \mathbf{I}_{\boldsymbol{\Phi}}\big)^{\mathrm{T}}\mathbf{X}\boldsymbol{\Sigma}^{-1}\mathbf{X}^{\mathrm{T}}\Big) \circ \mathbf{I}_{\boldsymbol{\Phi}}^{\mathrm{T}}\mathbf{U}}_{\text{from } \partial \log p(\mathbf{Y}|\mathbf{U},\mathbf{V},\mathbf{X},\boldsymbol{\mu},\boldsymbol{\Sigma})} + \underbrace{\beta(\sigma^{\boldsymbol{\Omega}} - \mathbf{I}_{\boldsymbol{\Omega}}) \circ \boldsymbol{\Omega}\mathbf{V}}_{\text{from } \partial \sum_{i,j} \log p(I_{\omega_{jj'}=1}|\mathbf{V})} + \underbrace{\alpha(\sigma^{\boldsymbol{\Phi}} - \mathbf{I}_{\boldsymbol{\Phi}})^{\mathrm{T}} \circ {\boldsymbol{\Phi}_*}^{\mathrm{T}}\mathbf{U}}_{\text{from } \partial \sum_{i,j} \log p(I_{\phi_{ij}}|\mathbf{U},\mathbf{V})}$$

where

$$\sigma^{\boldsymbol{\Omega}} = \sigma\big(\beta\ \boldsymbol{\Omega}_* \circ (\mathbf{V}\mathbf{V}^{\mathrm{T}})\big)$$

Procedure ProjectOnSimplex($\mathbf{v}, C$)

output: $\arg\min_{\mathbf{w}} \|\mathbf{w} - \mathbf{v}\|_2$ s.t. $\sum_i w_i \le C, \mathbf{w} \ge 0$

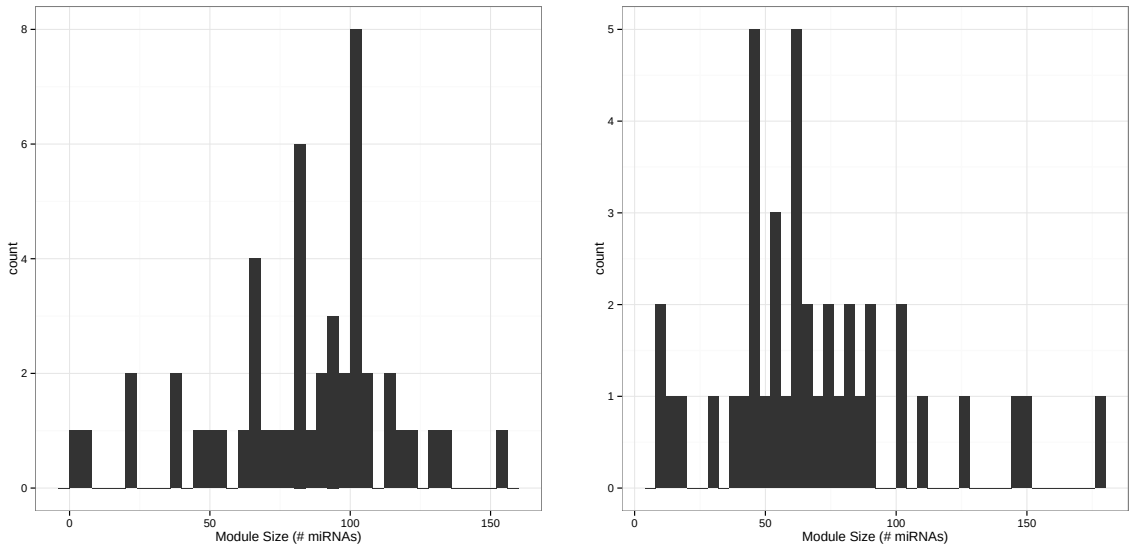
Procedure Project($\mathbf{U}, \mathbf{V}, C_1, C_2$)

```
// Threshold
for i, k do
    u_ik ← 0 if u_ik < ε ;
for j, k do
    v_jk ← 0 if v_jk < ε ;
for k do
    // Remove redundant entries in U
    for i ∈ {i : u_ik > 0  and  u_ik v_{,k}^T φ_i = 0} do
        u_ik ← 0 ;
    // Projection
    u_{,k} ← ProjectOnSimplex(u_{,k}, C_1);
    v_{,k} ← ProjectOnSimplex(v_{,k}, C_2);
```

Fig. 1. Projection procedure to solve the optimization problem (4)

2 Distribution of module sizes

Figure 2 shows the size of modules identified by SNMNMF and PIMiM. Modules identified by both methods have comparable size.



(a) SNMNMF (b) PIMiM

Fig. 2. The histogram of the size of modules of SNMNMF and PIMiM.

3 Choosing the parameter K and α

We select the parameter K of PIMiM that yields the best F1 score as shown in Fig. 3. In addition, we varied the values of α to examine the interplay effect of predictions of miRNA targets and protein interaction data. The result is shown in Fig. 4.

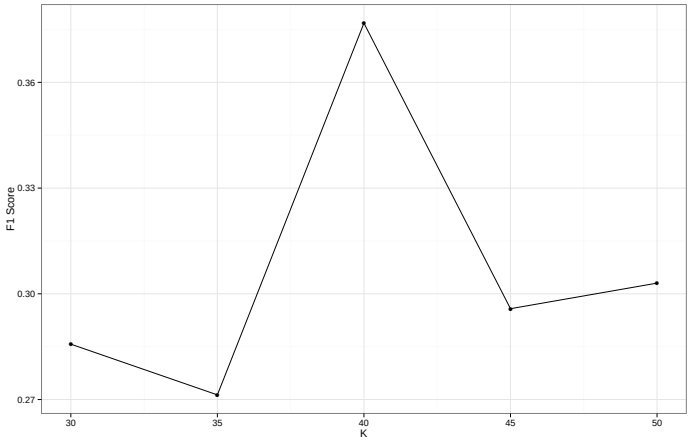


Fig. 3. Performance of PIMiM with different values of K.

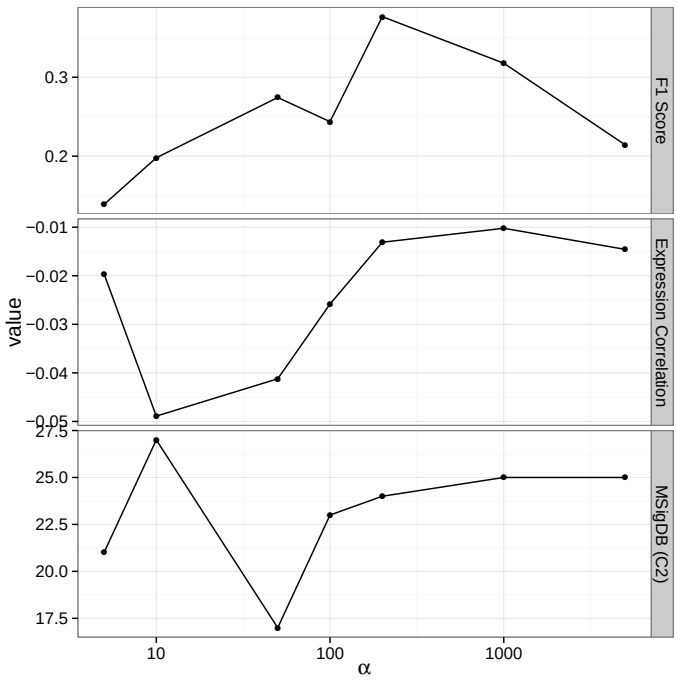


Fig. 4. We varied the value of α and tested the different metrics discussed in Results. On one hand, low values and high values lead to smaller F1 score. On the other hand, small values lead to more coherent gene modules, which explains the better expression correlation.

Table 1: List of modules identified by PIMiM in ovarian cancer (Sect. 3.1): PPI: the number of protein interactions in TRANSFAC between mrnas, Interactions: the number of miRNA-mRNA interactions in MicroCosm.

Module	#mirnas	#mrnas	ExpCorr (mirnas)	ExpCorr (mrnas)	ExpCorr	PPI Interactions		MiRNAs	Enriched GO keywords
1	4	8	0.278082	0.132891	-0.002236	10	11	mir-142-3p;mir-135a;mir-452;mir-144*	
2	4	125	0.251864	0.025514	-0.015311	392	60	mir-514;mir-592;mir-26a-1*;mir-665	tumor necrosis factor, death receptor activity, positive regulation of NF-kappaB transcription factor, regulation of MAPKKK cascade, anti-apoptosis, programmed cell death, immune system process
3	6	81	0.221464	0.067253	-0.013864	1916	100	mir-30b*;mir-29a;mir-29c;mir-512-3p;mir-944;mir-338-3p	ribosome, mitochondrial ribosome, mitochrondrion
4	5	62	0.323695	0.026375	-0.008652	84	103	mir-133b;mir-133a;mir-223;mir-126*;mir-361-3p	Ndc80 complex, blood coagulation, hemostasis
5	4	8	0.408359	0.218362	-0.117000	6	16	mir-516b;mir-183;mir-20a;mir-20b	
6	5	89	0.257539	0.039119	0.003993	3002	110	mir-940;mir-524-3p;mir-411;mir-601;mir-339-3p	RNA splicing, mRNA export from nucleus, miRNA processing, spliceosomal complex
7	5	76	0.207695	0.038756	-0.004794	410	118	mir-23a;mir-23b;mir-125a-5p;mir-512-5p;mir-296-5p	mediator complex, SWI/SNF complex, RNA polymerase II transcription cofactor act.
8	5	49	0.223381	0.042503	-0.008279	62	79	mir-519a;mir-584;mir-193a-5p;mir-215;mir-519c-3p	cysteine-type endopeptidase activity, endoplasmic reticulum membram
9	6	69	0.223090	0.044989	-0.007351	148	133	mir-770-5p;mir-520a-5p;mir-362-5p;mir-525-5p;mir-627;mir-668	CD40 receptor complex, integrin complex, programmed cell death, apoptosis, cell surface receptor linked signaling pathway
10	5	72	0.251736	0.037899	-0.001482	122	103	mir-29b;mir-30c;mir-30b;mir-214;mir-145*	symbiont intracellular protein transport in host, MHC class II protein complex, interferon-gamma-mediated signaling pathway, lymphocyte activation, cytokine stimulus, immune response
11	7	62	0.113588	0.023346	0.008720	150	98	mir-326;mir-219-5p;mir-196b;mir-370;mir-218;mir-656;mir-874	regulation of cellular senescence, cell aging, cyclin-dependent protein kinase holoenzyme complex, stress-activated MAPK cascade, MyD88-independent toll-lie receptor signaling pathway, toll signaling pathway
12	5	18	0.137023	0.066107	-0.016920	30	21	mir-141*;mir-27a;mir-345;mir-509-3-5p;mir-144	catenin-TCFL2 complex
13	6	64	0.168647	0.025391	-0.006487	130	105	mir-100;mir-99a;mir-519b-3p;mir-15b;mir-424;mir-339-3p	cell cycle checkpoint, response to stress

14	8	62	0.176488	0.024091	-0.007109	114	126	mir-423-3p;mir-92b*;mir-183*;mir-484;mir-572;mir-324-3p;mir-324-5p;mir-296-5p	
15	3	38	0.444554	0.036937	0.005140	48	45	mir-760;mir-362-5p;mir-874	
16	4	28	0.337957	0.045733	-0.004327	66	36	mir-129-3p;mir-491-3p;mir-221;mir-631	protein transmembrane transporter activity
17	5	46	0.191255	0.026197	-0.005014	94	78	mir-518f;mir-455-3p;mir-378;mir-661;mir-873	
18	3	63	0.352144	0.036239	-0.010661	92	61	mir-582-3p;mir-141;mir-206	
19	6	80	0.194203	0.028009	-0.008469	1274	118	mir-488;let-7b*;let-7a*;let-7f-1*;mir-361-5p;mir-630	U7 snRNP, spliceosomal complex, catalytic step 2 spliceosome, mRNA processing
20	6	44	0.142625	0.022868	-0.016685	68	92	mir-765;mir-324-3p;mir-137;mir-338-3p;mir-519c-3p;mir-15a*	tumor necrosis factor-mediated signaling pathway, induction of apoptosis, programmed cell death, cellular response to tumor necrosis factor
21	3	109	0.493175	0.080146	-0.003825	6594	34	let-7d;let-7b;mir-34a	cytosolic large ribosomal subunit, viral transcription, viral infectious cycle, ribosome, rRNA processing
22	7	146	0.153077	0.043636	-0.007805	6592	139	mir-185*;mir-200a;mir-744;mir-99a;mir-99b;mir-324-5p;mir-143*	typ2 fibroblast growth factor receptor binding, viral infectious cycle
23	7	103	0.116707	0.036921	-0.002764	360	143	mir-519a;mir-575;mir-518a-3p;mir-501-3p;mir-598;mir-10b;mir-337-5p	insulin receptor substrate binding, regulation of cell motility, cell activation, cell migration, leukocyte migration
24	6	176	0.271370	0.033961	-0.008809	5380	161	mir-454*;mir-297;mir-200c;mir-200b;mir-130b*;mir-429	U7 snRNP, spliceosomal snRNP assembly, U12-type spliceosomal complex
25	6	100	0.189457	0.041726	-0.000940	3340	111	mir-107;mir-103;mir-135a;mir-34c-3p;mir-34b;mir-337-3p	U7 snRNP, spliceosomal snRNP assembly, U12-type spliceosomal complex, cis-trans isomerase activity
26	3	67	0.293141	0.023804	-0.003539	124	84	mir-181c;mir-222;mir-885-5p	CD40 receptor complex, regulation of JNK cascade, JUN kinase activity, regulation of I-kappaB kinase/NF-kappaB cascade
27	4	150	0.385310	0.043335	-0.005580	6420	122	mir-181a;mir-193b*;mir-584;mir-222	cytosolic large ribosomal subunit, translational termination, viral replication, viral infectious cycle, ribosome
28	3	53	0.318788	0.027428	-0.002548	144	41	mir-142-3p;mir-518f;mir-592	eukaryotic cell surface binding, platelet degranulation, response to hormone stimulus, fibrinogen complex
29	5	54	0.176097	0.036215	0.000688	120	66	mir-142-3p;let-7i*;mir-937;mir-602;mir-203	cysteine-type endopeptidase activity, induction to apoptosis by extracellular signals, programmed cell death, I-kappaB kinase/NF-kappaB cascase
30	2	61	0.458172	0.035994	-0.029003	132	20	mir-199a-5p;mir-196b	cholesterol transport, sterol transport, positive regulation of transcription

31	5	45	0.415015	0.052412	-0.012146	50	89	mir-221*;let-7a;let-7f;mir-217;mir-203	
32	6	88	0.144199	0.026152	-0.003597	1234	113	mir-891a;let-7a;mir-193b;mir-425;mir-486-5p;mir-9	laminin-10, laminin-11 complex, U7 snRNP
33	4	72	0.285136	0.033364	0.017618	168	57	mir-34b*;mir-508-3p;mir-126;mir-518d-5p	
34	4	86	0.278431	0.027148	-0.003566	226	83	mir-92b;mir-150*;mir-196a;mir-196b	membrane attack complex, sarcomere, complement activation, titin binding, response to stimulus, chemical response
35	3	47	0.237759	0.067964	-0.002561	132	17	mir-18a;let-7b;mir-375	cell adhesion, hemostasis, response to stimulus
36	4	12	0.256672	0.086306	-0.005786	20	12	mir-450a;mir-520c-3p;mir-133b;mir-659	MAPKKK cascade, response to stress
37	3	57	0.266377	0.038625	-0.005862	98	51	mir-545;mir-135a*;mir-338-3p	transmembrane receptor protein kinase activity
38	5	46	0.210763	0.041924	-0.025455	60	89	mir-518e;mir-105;mir-25;mir-195;mir-365	
39	3	42	0.386768	0.061612	-0.032561	58	28	mir-141*;mir-16;mir-186	response to estradiol stimulus, sterol transport
40	3	54	0.367195	0.034767	-0.019985	92	54	mir-27b*;mir-135a;mir-873	ESC/E(Z) complex, PcG protein complex, histone methyltransferase complex

4 Enrichment results of several modules from TCGA dataset

Module 23 This module (Fig. 6) includes the miR-302 and miR-520 clusters. These two clusters are shown to display similar expression pattern in the differentiation of human embryonic stem cells [5]. Specifically, the miR-302 family is known for coordinately suppressing genes in the CDK2 and CDK4/6 cell cycle pathways [2]. Indeed, miRNAs in the miR-302 family were assigned to the same module by PIMiM indicating that the module-based approach can help in recovering cooperative regulation of groups of miRNAs. Among the top terms and gene sets from Gene Ontology and MSigDB enrichment analysis are: cell death, CD40 receptor complex, regulation of apoptosis, B cell immune response, TNF receptor signaling pathway, ... (Table 3).

Module 48 (Fig. 7 and Table 4) All miRNAs in this module were previously reported as active in cancer: miR-130a/b [6], miR-328 [3] and mirR-504 which negatively regulates tumor suppressor p53 [1]. Mutation of the gene hub CEBPE is shown to increase the risk of acute leukemia[4].

(a) GO

ID	Name	Adj.P-value
GO:0033276	transcription factor TFTC complex	<0.001
GO:0070461	SAGA-type complex	<0.001
GO:0000123	histone acetyltransferase complex	<0.001
GO:0005669	transcription factor TFIID complex	<0.001
GO:0005667	transcription factor complex	<0.001
GO:0044428	nuclear part	<0.001
GO:0016578	histone deubiquitination	<0.001
GO:0006352	transcription initiation, DNA-dependent	<0.001
GO:0019219	regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	<0.001
GO:0051171	regulation of nitrogen compound metabolic process	<0.001

(b) MSigDB

Gene Set Name	Description	P value
MIPS TFTC COMPLEX	TFTC complex (TATA-binding protein-free TAF-II-containing complex)	0E0
MIPS GCN5 TRRAP HISTONE ACETYLTRANSFERASE COMPLEX	GCN5-TRRAP histone acetyltransferase complex	0E0
KEGG BASAL TRANSCRIPTION FACTORS	Basal transcription factors	5.55E-16
MIPS TFIID BETA COMPLEX	TFIID-beta complex	1.86E-13
MIPS TFIID BETA COMPLEX 1	TFIID-beta complex	1.86E-13
MIPS STAGA COMPLEX	STAGA complex (SPT3-TAF9-GCN5 acetyltransferase complex)	3.02E-13
MIPS DA COMPLEX	DA complex	7.05E-13
MIPS PCAF COMPLEX	PCAF complex	7.85E-11
MIPS TFIID COMPLEX	TFIID complex	1.85E-10
MIPS TFIID COMPLEX B CELL SPECIFIC	TFIID complex, B-cell specific	1.85E-10

Table 2. Enrichment analysis of the set of genes in Module 11.

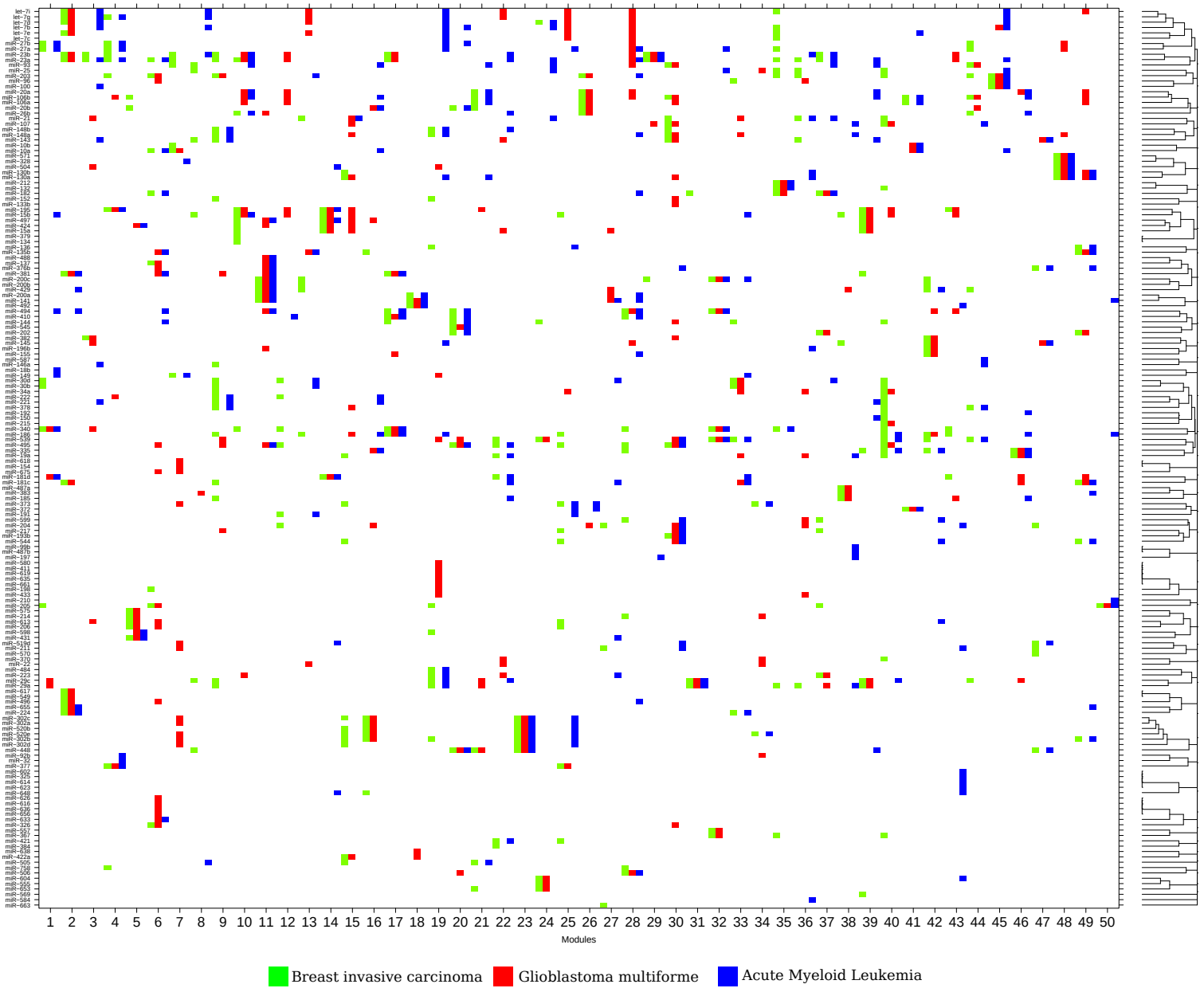


Fig. 5. Inferred miRNA modules of the three cancer types (BRCA, GBM and AML). The x-axis shows the 50×3 modules learned for the three cancer types (each x-axis bar is subdivided into 3 with the color corresponding to the cancer type). The y-axis shows miRNAs ordered by hierarchical clustering of their module membership vector. As can be seen, in several cases the same miRNAs are predicted for all or two of the three cancer types.

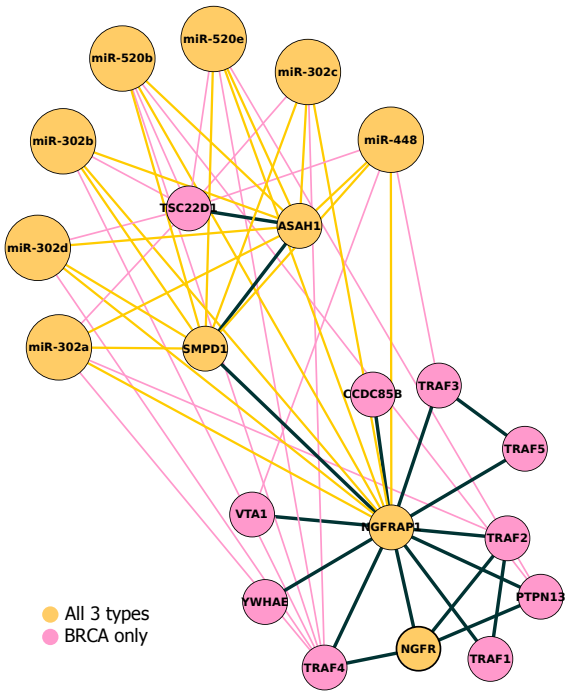


Fig. 6. Network of miRNAs and mRNAs of Module 23.

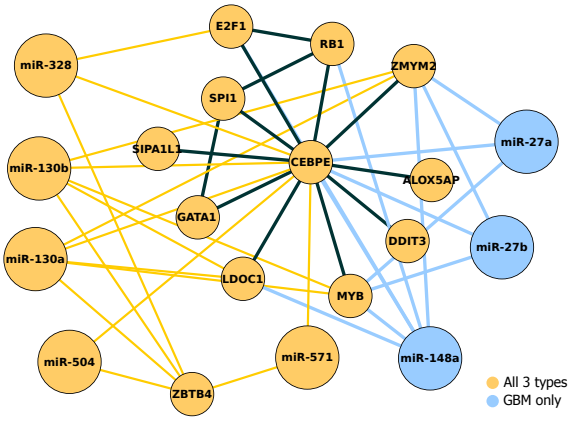


Fig. 7. Network of miRNAs and mRNAs of Module 48.

(a) GO

ID	Name	Adj. P-value
GO:0010941	regulation of cell death	<0.001
GO:0010942	positive regulation of cell death	<0.001
GO:0035631	CD40 receptor complex	<0.001
GO:0042981	regulation of apoptosis	<0.001
GO:0043065	positive regulation of apoptosis	<0.001
GO:0043067	regulation of programmed cell death	<0.001
GO:0043068	positive regulation of programmed cell death	<0.001
GO:0008624	induction of apoptosis by extracellular signals	0.004
GO:0009898	internal side of plasma membrane	0.005
GO:0006917	induction of apoptosis	0.015
GO:0012502	induction of programmed cell death	0.015
GO:0048522	positive regulation of cellular process	0.015
GO:0048518	positive regulation of biological process	0.031
GO:0004842	ubiquitin-protein ligase activity	0.032
GO:0019787	small conjugating protein ligase activity	0.041
GO:0051090	regulation of sequence-specific DNA binding transcription factor activity	0.047
GO:0051092	positive regulation of NF-kappaB transcription factor activity	0.047
GO:0035304	regulation of protein dephosphorylation	0.05

(b) MSigDB

Gene Set Name	Description	P value
KEGG SMALL CELL LUNG CANCER	Small cell lung cancer	1.62E-9
BIOCARTA TALL1 PATHWAY	TACI and BCMA stimulation of B cell immune responses	1.02E-7
BIOCARTA TNFR2 PATHWAY	TNFR2 Signaling Pathway	1.82E-7
PID CD40 PATHWAY	CD40/CD40L signaling	9.97E-7
SIG CD40 PATHWAY MAP	Genes related to CD40 signaling	1.33E-6
KEGG PATHWAYS IN CANCER	Pathways in cancer	1.48E-6
PID TNF PATHWAY	TNF receptor signaling pathway	3.35E-6
PID CERAMIDE PATHWAY	Ceramide signaling pathway	3.81E-6
LAU APOPTOSIS CDKN2A UP	Genes up-regulated by UV-irradiation in cervical cancer cells after knockdown of CDKN2A	5.77E-6
REACTOME CELL DEATH SIGNALLING VIA NRAGE NRIF AND NADE	Genes involved in Cell death signalling via NRAGE, NRIF and NADE	7.51E-6

Table 3. Enrichment analysis of the set of genes in Module 23.

(a) GO

ID	Name	Adj.P-value
GO:0071930	negative regulation of transcription involved in G1/S phase of mitotic cell cycle	0.005
GO:0035189	Rb-E2F complex	0.008
GO:0000122	negative regulation of transcription from RNA polymerase II promoter	0.01
GO:0005634	nucleus	0.032
GO:0003700	sequence-specific DNA binding transcription factor activity	0.033
GO:0001071	nucleic acid binding transcription factor activity	0.033
GO:0003677	DNA binding	0.038
GO:0019219	regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	0.045

(b) MSigDB

Gene Set Name	Description	P value
PID HES HEYPATHWAY	Notch-mediated HES/HEY network	1.1E-8
MARKS ACETYLATED NON HISTONE PROTEINS	Non-histone proteins that are acetylated	6.14E-8
PARK TRETINOIN RESPONSE AND RARA PLZF FUSION	Genes up-regulated by tretinoin (all-trans retinoic acid, ATRA) in U937 cells (acute promyelocytic leukemia, APL) made resistant to the drug by expression of the PLZF-RARA fusion	2.08E-7
PARK TRETINOIN RESPONSE AND PML RARA FUSION	Genes up-regulated by tretinoin (all-trans retinoic acid, ATRA) in U937 cells (acute promyelocytic leukemia, APL) made sensitive to the drug by expression of the PML-RARA fusion	5.46E-7
MAGRANGEAS MULTIPLE MYELOMA IGLL VS IGLK UP	Up-regulated genes discriminating multiple myeloma samples by the ype of immunoglobulin light chain they produce: Ig lambda (IGLL) vs Ig kappa (IGLK)	1.54E-6
TONKS TARGETS OF RUNX1 RUNX1T1 FUSION HSC DN	Genes down-regulated in normal hematopoietic progenitors by RUNX1-RUNX1T1 fusion.	2.67E-6
PID RB 1PATHWAY	Regulation of retinoblastoma protein	5.81E-6
RAMJAUN APOPTOSIS BY TGFB1 VIA MAPK1 DN	Apoptotic genes dependent on MAPK1 and down-regulated in AML12 cells (hepatocytes) after stimulation with TGFB1	8.09E-6
QI PLASMACYTOMA UP	Up-regulated genes that best disciminate plasmablastic plasmacytoma from plasmacytic plasmacytoma tumors	9.69E-6
PID CMYB PATHWAY	C-MYB transcription factor network	1.26E-5

Table 4. Enrichment analysis of the set of genes in Module 48.

References

1. Wenwei Hu, Chang S Chan, Rui Wu, Cen Zhang, Yvonne Sun, Jun S Song, Laura H Tang, Arnold J Levine, and Zhaohui Feng. Negative regulation of tumor suppressor p53 by microrna mir-504. *Molecular cell*, 38(5):689–699, 2010.
2. S.L. Lin, D.C. Chang, S.Y. Ying, D. Leu, and D.T.S. Wu. Microrna mir-302 inhibits the tumorigenecity of human pluripotent stem cells by coordinate suppression of the cdk2 and cdk4/6 cell cycle pathways. *Cancer research*, 70(22):9473–9482, 2010.
3. Yu-Zhuo Pan, Marilyn E Morris, and Ai-Ming Yu. Microrna-328 negatively regulates the expression of breast cancer resistance protein (bcrp/abcg2) in human cancer cells. *Molecular pharmacology*, 75(6):1374–1379, 2009.
4. Pauline Peyrouze, Soizic Guihard, Nathalie Grardel, Céline Berthon, Nicolas Pottier, Arnaud Pigneux, Jean-Yves Cahn, Marie Christine Béné, Véronique Lhéritier, Eric Delabesse, et al. Genetic polymorphisms in arid5b, cebpe, ikzf1 and cdkn2a in relation with risk of acute lymphoblastic leukaemia in adults: a group for research on adult acute lymphoblastic leukaemia (graall) study. *British journal of haematology*, 159(5):599–613, 2012.
5. Jiaqiang Ren, Ping Jin, Ena Wang, Francesco M Marincola, and David F Stroncek. Microrna and gene expression patterns in the differentiation of human embryonic stem cells. *Journal of Translational Medicine*, 7(1):20, 2009.
6. S. Volinia, M. Galasso, S. Costinean, L. Tagliavini, G. Gamberoni, A. Drusco, J. Marchesini, N. Mascellani, M.E. Sana, R.A. Jarour, et al. Reprogramming of mirna networks in cancer and leukemia. *Genome research*, 20(5):589–599, 2010.