Supplementary Material

Description of Datasets

This section presents the brief descriptions of nine microarray data sets, which are used to evaluate the performance of different criteria, namely, maximum relevance-maximum functional similarity (MRMFS), maximum relevance (MR), minimum redundancy-maximum relevance (mRMR) [3], and maximum relevance-maximum significance (MRMS) [9], and different integrated methods for disease gene selection, namely, Rel-Sim, MR+PPIN [10], mRMR+PPIN [8], MRMS+PPIN [12], CLAIM [2], and GenePEN [15]. Experimental results on first five data sets are reported in the main paper, while that of the remaining four data sets are presented in this supplementary material. The descriptions of these nine microarray data sets used are as follows:

(i) *GSE25070:* It is the gene expression data retrieved from study of Hinoue et al. [6]. The data set contains the expression profiles of 26 colorectal tumors matched histological to normal adjacent colonic tissue samples. Illumina Ref-8 whole- genome expression BeadChip with 24526 probes corresponding to 18491 genes was used to obtaing the gene expression profiles.

(ii) *GSE10950:* The gene expression data comapres 24 colon normal to tumor pairs, making use of the Illumina BeadChip Human Ref8-v2. The data had been put forward by Jiang et al. to study the role of DACT3 as an epigenetic regulator of Wnt/β-catenin signaling in colorectal cancer [7]. It contains the gene expression profiles of 22184 genes.

(iii) *GSE4988:* Collado et al. put forward this gene expression data set to measure the relative abundance of the different RNA species in plasma, using the cDNA microarray hybridization. It contains the expression profiles of 12 colorectal cancer patients and 8 healthy persons [1].

(iv) *GSE44861:* GSE44861 is an Affymetrix expression data from colon cancer patient tissues. Affymetrix U113A arrays were used to contain the expression profiles of 111 colonic samples obtained from 56 tumors and 55 adjacent noncancerous tissues [13].

(v) *GSE11223:* Thid data contains the transcriptional profiling of colon epithelial biopsies from ulcerative colitis patients and healthy control donors. It contains the expression profiles of 44290 probes corresponding to 40991 genes, for the 202 samples, of which 129 are diseased and 73 are non-diseased [11]. Each data set is pre-processed by standardizing each sample to zero mean and unit variance.

(vi) *Breast Cancer:* The breast cancer data set contains expression levels of 7129 genes in 49 breast tumor samples [16]. The samples are classified according to their estrogen receptor (ER) status: 25 samples are ER positive while other 24 samples are ER negative.

(vii) *Leukemia:* It is an Affymetrix high density oligonucleotide array that contains 7070 genes and 72 samples from two classes of leukemia [4]: 47 acute lymphoblastic leukemia and 25 acute myeloid leukemia.

(viii) *Diffuse Large B-Cell Lymphoma:* This data set consists of gene expression profiles of 77 patients: 58 with diffuse large b-cell lymphoma (DLBCL) and 19 with follicular lymphoma morphology [14]. The expression profile contains 7129 genes.

(ix) *Lung Cancer:* There are 181 tissue samples: 31 mesothelioma and 150 adenocarcinoma. Each sample is described by 12533 genes [5].

Experimental Results

Table 1
Performance Analysis for Different Values of Weight Parameter Using Gene Ontology, Disease Ontology, and KEGG Pathway

Data	0.1		0.2		0.3		0.4		0.5		0.6		0.7		0.8		0.9	
Sets	Term	P-Value	Term	P-Value	Term	P-Value	Term	P-Value	Term	P-Value	Term	P-Value	Term	P-Value	Term	P-Value	Term	P-Value
Breast	P35	5.78E-40	P43	3.85E-35	P43	3.60E-35	P43	3.66E-21	P39	3.61E-10	P39	2.14E-08	P39	1.88E-07	P45	3.44E-07	P45	2.89E-07
	C25	1.15E-14	C25	1.25E-12	C25	1.16E-11	C21	2.73E-10	C24	1.91E-03	C15	4.47E-03	C27	5.22E-03	C2	1.19E-03	C4	9.46E-04
	F25	2.15E-19	F8	1.29E-15	F8	4.60E-12	F8	4.22E-10	F13	1.88E-05	F13	1.49E-04	F13	1.49E-04	F3	4.09E-05	F3	3.18E-05
	K24	6.08E-20	K24	8.31E-23	K24	1.72E-21	K24	3.30E-12	K29	2.27E-05	K26	1.28E-03	K18	2.84E-04	K18	2.22E-04	K18	1.70E-03
	D3	7.72E-21	D3	5.22E-21	D3	3.26E-19	D3	6.00E-13	D3	6.97E-06	D3	1.05E-04	D3	9.79E-05	D3	2.99E-05	D3	1.47E-04
DLBCL	P3	1.57E-30	P30	4.14E-34	P4	3.94E-59	P5	1.78E-44	P10	3.91E-20	P47	1.16E-05	P47	9.65E-05	P47	6.70E-04	P6	1.89E-03
	C8	5.57E-20	C8	9.22E-23	C3	2.78E-38	C3	4.13E-33	C8	1.55E-06	C28	5.34E-05	C28	5.34E-05	C28	5.34E-05	C28	5.34E-05
	F25	2.06E-19	F25	3.33E-15	F4	1.65E-11	F9	1.29E-08	F15	1.91E-11	F10	1.89E-07	F10	1.76E-07	F10	1.56E-07	F10	1.76E-07
	K24	2.70E-18	K8	4.77E-16	K3	4.68E-21	K9	8.89E-15	K5	9.63E-09	K21	2.19E-02	K21	1.35E-02	K21	9.72E-03	K21	6.48E-03
	D9	1.79E-10	D9	1.02E-12	D9	9.34E-03	D9	1.08E-02	D9	4.58E-02	D11	8.24E-04	D10	8.96E-03	D4	2.98E-02	D4	1.88E-02
Leukemia	P44	2.39E-17	P18	1.09E-18	P15	6.05E-15	P41	5.72E-12	P22	4.35E-07	P22	4.38E-05	P9	1.22E-04	P9	1.22E-04	P9	1.11E-04
	C10	4.72E-12	C10	1.12E-14	C10	8.26E-12	C10	7.08E-12	C16	2.42E-06	C16	1.21E-06	C16	1.01E-06	C16	1.01E-06	C16	1.01E-06
	F3	1.96E-11	F3	6.23E-10	F19	1.55E-07	F19	8.23E-08	F19	5.52E-05	F19	9.62E-04	F19	9.62E-04	F19	8.56E-04	F19	8.56E-04
	K13	6.05E-11	K24	6.44E-11	K22	5.44E-11	K22	2.04E-11	K22	8.22E-10	K22	2.65E-06	K22	2.37E-05	K22	2.37E-05	K22	1.48E-04
	D7	3.19E-20	D7	4.87E-20	D7	3.33E-16	D7	1.77E-11	D7	1.64E-06	D7	1.04E-04	D7	2.63E-04	D7	2.63E-04	D7	8.76E-04
Lung	P37	2.73E-43	P37	1.27E-44	P34	7.63E-15	P48	2.28E-09	P19	9.55E-07	P19	6.93E-06	P19	5.92E-05	P19	5.92E-05	P19	5.92E-05
	C25	5.15E-28	C25	2.36E-29	C25	1.46E-11	C14	2.13E-07	C14	3.48E-06	C14	3.98E-06	C14	4.48E-06	C14	3.98E-06	C14	3.48E-06
	F8	8.73E-16	F8	5.14E-17	F20	2.61E-05	F22	3.21E-05	F5	1.08E-04	F5	1.08E-04	F28	3.57E-03	F28	3.57E-03	F11	5.62E-05
	K8	2.80E-23	K8	1.03E-24	K24	3.08E-10	K23	1.32E-04	K23	6.27E-04	K23	6.27E-04	K23	5.76E-04	K7	1.49E-03	K7	1.49E-03
	D8	8.63E-21	D8	4.17E-19	D8	8.54E-12	D8	2.48E-05	D8	1.25E-04	D8	1.69E-03	D8	1.70E-03	D8	1.59E-03	D8	6.51E-04

Table 2
Performance Analysis of Different Criteria Using Gene Ontology, Disease Ontology, and KEGG Pathway

Data	Different	Biological process		Cellular Components		Molecular Function		KEGG Pathway		Disease Ontology	
Used	Criteria	Term	P-Value	Term	P-Value	Term	P-Value	Term	P-Value	Term	P-Value
Breast	*t*-test	P17	7.29E-05	C9	3.90E-05	F6	1.77E-05	K12	2.60E-02	D3	9.14E-04
	MR	P40	8.99E-07	C17	5.76E-07	F18	2.13E-06	K18	7.22E-06	D3	8.04E-03
	mRMR	P40	4.25E-08	C17	2.94E-05	F18	1.56E-04	K18	1.66E-04	D3	1.08E-03
	MRMS	P40	1.37E-05	C17	6.37E-07	F18	2.04E-06	K18	8.46E-06	D3	3.29E-02
	MRMFS	P43	3.85E-35	C25	1.25E-12	F8	1.29E-15	K24	8.31E-23	D3	5.22E-21
DLBCL	*t*-test	P25	2.72E-03	C26	4.94E-03	F7	4.46E-03	K11	1.05E-02	*	*
	MR	P14	6.07E-04	C15	2.78E-03	F10	2.48E-05	K14	1.79E-04	*	*
	mRMR	P29	7.50E-06	C11	2.54E-03	F10	1.79E-05	K2	6.47E-04	D10	4.94E-02
	MRMS	P47	1.44E-04	C22	4.42E-03	F10	2.15E-05	K14	1.09E-03	*	*
	MRMFS	P30	4.14E-34	C8	9.22E-23	F25	3.33E-15	K8	4.77E-16	D9	1.02E-12
Leukemia	*t*-test	P21	3.02E-05	C19	5.88E-07	F12	5.40E-04	K11	1.24E-03	D7	4.43E-03
	MR	P22	1.88E-05	C16	1.41E-06	F21	3.01E-03	K19	7.60E-03	D7	4.87E-02
	mRMR	P9	6.41E-05	C16	1.41E-06	F19	9.36E-03	K19	8.44E-03	D7	5.48E-03
	MRMS	P22	1.79E-05	C16	1.01E-06	F19	7.49E-03	K19	1.00E-02	D15	4.54E-02
	MRMFS	P18	1.09E-18	C10	1.12E-14	F3	6.23E-10	K24	6.44E-11	D7	4.87E-20
Lung	*t*-test	P8	6.55E-03	C20	2.37E-03	F23	4.58E-03	K7	7.59E-03	D13	3.65E-02
	MR	P20	6.13E-06	C14	2.07E-06	F28	1.51E-05	K20	1.09E-03	D8	1.52E-02
	mRMR	P20	6.35E-05	C14	1.03E-06	F28	1.29E-05	K28	1.07E-04	D8	3.87E-02
	MRMS	P48	6.56E-07	C14	2.41E-06	F28	1.61E-05	K28	1.41E-03	D8	1.35E-02
	MRMFS	P37	1.27E-44	C25	2.36E-29	F8	5.14E-17	K8	1.03E-24	D8	4.17E-19

Definitions of Different Terms Used in Table 1, Table 2, and Table 3

- **Terms of Biological Processes: P1:** branch elongation involved in mammary gland duct branching; **P2:** calcium-independent cell-cell adhesion via plasma membrane cell-adhesion molecules; **P3:** cell activation; **P4:** cell cycle; **P5:** cell cycle process; **P6:** cell-cell recognition; **P7:** cell-type specific apoptotic process; **P8:** cellular response to catecholamine stimulus; **P9:** chaperone-mediated autophagy; **P10:** chemokine-mediated signaling pathway; **P11:** endodermal cell differentiation; **P12:** female pregnancy; **P13:** fructose 1,6-bisphosphate metabolic process; **P14:** hexose catabolic process; **P15:** inflammatory response; **P16:** inter-species interaction between organisms; **P17:** intrinsic apoptotic signaling pathway in response to endoplasmic reticulum stress; **P18:** leukocyte activation; **P19:** lung development; **P20:** morphogenesis of a branching epithelium; **P21:** myeloid dendritic cell activation; **P22:** myeloid leukocyte activation; **P23:** negative regulation of apoptotic signaling pathway; **P24:** negative regulation of cellular component movement; **P25:** nucleoside diphosphate metabolic process; **P26:** Pathways in cancer; **P27:** platelet degranulation; **P28:** positive regulation of cell activation; **P29:** positive regulation of cell-cell adhesion; **P30:** positive regulation of immune system process; **P31:** positive regulation of lymphocyte activation; **P32:** positive regulation of neuron death; **P33:** regulation of carbohydrate biosynthetic process; **P34:** regulation of cell adhesion; **P35:** regulation of cell proliferation; **P36:** regulation of cell-cell adhesion; **P37:** regulation of immune system process; **P38:** regulation of tissue remodeling; **P39:** response to alcohol; **P40:** response to amino acid; **P41:** response to bacterium; **P42:** response to laminar fluid shear stress; **P43:** response to lipid; **P44:** response to oxidative stress; **P45:** response to vitamin; **P46:** stimulatory C-type lectin receptor signaling pathway; **P47:** T cell receptor signaling pathway; **P48:** vascular endothelial growth factor receptor signaling pathway;

Table 3
Performance Analysis of Different Integrated Methods Using Gene Ontology, Disease Ontology, and KEGG Pathway

Data Used	Integrated Methods	Biological Processes		Cellular Components		Molecular Functions		KEGG Pathway		Disease Ontology	
		Term	P-Value	Term	P-Value	Term	P-Value	Term	P-Value	Term	P-Value
Breast	MR+PPIN	P42	2.93E-04	C22	3.16E-04	F17	4.35E-05	K17	6.21E-03	D3	9.41E-03
	mRMR+PPIN	P33	1.36E-04	*	*	F17	2.18E-06	K30	2.75E-04	D6	1.68E-02
	MRMS+PPIN	P31	1.32E-05	C5	1.55E-03	F6	1.14E-03	K17	1.76E-04	D3	9.41E-03
	CLAIM	P16	3.15E-21	C7	7.22E-15	F24	2.20E-11	K25	1.79E-08	*	*
	GenePEN	P26	1.29E-12	*	*	*	*	*	*	D3	4.89E-12
	RelSim	P43	3.85E-35	C25	1.25E-12	F8	1.29E-15	K24	8.31E-23	D3	5.22E-21
DLBCL	MR+PPIN	P13	1.46E-04	C23	2.62E-04	F9	3.50E-04	K9	3.38E-03	D9	1.30E-02
	mRMR+PPIN	P32	3.87E-04	C23	7.64E-04	F5	3.84E-05	K27	2.17E-02	D5	7.16E-03
	MRMS+PPIN	P1	3.62E-04	C23	9.05E-04	F5	3.03E-05	K27	8.36E-04	D1	7.16E-05
	CLAIM	P11	2.77E-04	C13	4.43E-03	F2	1.35E-03	K6	1.44E-02	*	*
	GenePEN	P28	8.93E-12	C22	1.32E-04	F1	1.10E-03	K31	2.01E-04	D9	4.68E-04
	RelSim	P30	4.14E-34	C8	9.22E-23	F25	3.33E-15	K8	4.77E-16	D9	1.02E-12
Leukemia	MR+PPIN	P27	6.93E-04	C12	3.98E-04	F16	9.82E-05	K16	1.08E-01	*	*
	mRMR+PPIN	P38	8.08E-05	C18	1.00E-04	F16	1.97E-06	K20	1.04E-04	*	*
	MRMS+PPIN	P9	3.23E-05	C12	3.98E-04	F16	8.84E-05	K1	1.16E-03	D7	1.19E-02
	CLAIM	P46	1.97E-04	C10	6.67E-05	F26	1.68E-03	K11	2.45E-05	D12	9.63E-04
	GenePEN	P7	1.81E-19	C25	1.44E-10	F25	4.44E-14	K24	6.83E-17	D14	3.77E-28
	RelSim	P18	1.09E-18	C10	1.12E-14	F3	6.23E-10	K24	6.44E-11	D7	4.87E-20
Lung	MR+PPIN	P24	8.15E-05	C14	9.49E-05	F21	6.28E-04	K4	3.27E-04	D8	2.36E-04
	mRMR+PPIN	P23	7.99E-06	C22	9.53E-05	F27	4.34E-04	K15	9.09E-04	D8	8.88E-05
	MRMS+PPIN	P12	1.51E-04	C14	3.97E-06	F21	5.23E-04	K4	3.93E-04	D8	1.20E-05
	CLAIM	P2	1.88E-04	C1	4.68E-04	F14	3.96E-03	K10	2.07E-02	D2	2.15E-02
	GenePEN	P36	2.19E-13	C6	1.99E-07	F22	4.86E-07	K15	1.74E-05	D8	1.57E-10
	RelSim	P37	1.27E-44	C25	2.36E-29	F8	5.14E-17	K8	1.03E-24	D8	4.17E-19

Table 4
Classification Accuracy of Different Integrated Methods for Four Microarray Data Sets

Microarray Data Sets	Accuracy (%) of Different Integrated Methods					
	MR+PPIN	mRMR+PPIN	MRMS+PPIN	CLAIM	GenePEN	RelSim
Breast	93.88	93.88	95.92	67.35	97.96	97.96
Leukemia	98.61	97.22	98.61	98.61	98.61	98.61
DLBCL	96.10	96.10	97.40	94.81	93.51	93.51
Lung	100	100	98.90	100	98.34	99.45

- **Terms of Cellular Component: C1:** basal plasma membrane; **C2:** basement membrane; **C3:** chromosome; **C4:** costamere; **C5:** cytoplasmic membrane-bounded vesicle lumen; **C6:** cytoplasmic side of plasma membrane; **C7:** cytosolic small ribosomal subunit; **C8:** external side of plasma membrane; **C9:** extracellular matrix component; **C10:** I-kappaB/NF-kappaB complex; **C11:** immunological synapse; **C12:** integrin complex; **C13:** intercellular bridge; **C14:** lateral plasma membrane; **C15:** lysosomal lumen; **C16:** mast cell granule; **C17:** MHC class II protein complex; **C18:** multivesicular body; **C19:** nuclear chromatin; **C20:**

plasma membrane raft; **C21:** plasma membrane receptor complex; **C22:** platelet alpha granule; **C23:** rough endoplasmic reticulum; **C24:** sarcolemma; **C25:** side of membrane; **C26:** small ribosomal subunit; **C27:** T-tubule; **C28:** zona pellucida receptor complex;

- **Terms of Molecular Function: F1:** activating transcription factor binding; **F2:** amino acid binding; **F3:** antioxidant activity; **F4:** ATPase activity; **F5:** cell adhesion molecule binding; **F6:** core promoter binding; **F7:** core promoter sequence-specific DNA binding; **F8:** cytokine receptor binding; **F9:** damaged DNA binding; **F10:** DNA insertion or deletion binding; **F11:** endopeptidase inhibitor activity; **F12:** enhancer binding; **F13:** enhancer sequence-specific DNA binding; **F14:** extracellular matrix binding; **F15:** G-protein coupled peptide receptor activity; **F16:** inositol trisphosphate kinase activity; **F17:** integrin binding; **F18:** MHC class II receptor activity; **F19:** peroxidase activity; **F20:** phosphatidylinositol 3-kinase binding; **F21:** protease binding; **F22:** protein tyrosine kinase activity; **F23:** scaffold protein binding; **F24:** structural constituent of ribosome; **F25:** transcription factor binding; **F26:** transcriptional activator activity, RNA polymerase II distal enhancer sequence-specific binding; **F27:** transmembrane receptor protein kinase activity; **F28:** transmembrane receptor protein tyrosine kinase activity;
- **Terms of KEGG Pathway: K1:** B cell receptor signaling pathway; **K2:** Cell adhesion molecules (CAMs); **K3:** Cell cycle; **K4:** Chagas disease (American trypanosomiasis); **K5:** Chemokine signaling pathway; **K6:** Citrate cycle (TCA cycle); **K7:** Complement and coagulation cascades; **K8:** Cytokine-cytokine receptor interaction; **K9:** DNA replication; **K10:** ECM-receptor interaction; **K11:** Epstein-Barr virus infection; **K12:** Estrogen signaling pathway; **K13:** FoxO signaling pathway; **K14:** Glycolysis / Gluconeogenesis; **K15:** HIF-1 signaling pathway; **K16:** Inositol phosphate metabolism; **K17:** Intestinal immune network for IgA production; **K18:** Leishmaniasis; **K19:** Lysosome; **K20:** Malaria; **K21:** Mismatch repair; **K22:** Osteoclast differentiation; **K23:** Pathogenic Escherichia coli infection; **K24:** Pathways in cancer; **K25:** Ribosome; **K26:** Small cell lung cancer; **K27:** TGF-beta signaling pathway; **K28:** Tight junction; **K29:** TNF signaling pathway; **K30:** Toxoplasmosis; **K31:** Transcriptional misregulation in cancer;
- **Terms of Disease Ontology: D1:** Alzheimer's disease; **D2:** benign neoplasm; **D3:** breast cancer; **D4:** esophagus squamous cell carcinoma; **D5:** gastrointestinal system cancer; **D6:** hereditary breast ovarian cancer; **D7:** leukemia; **D8:** lung cancer; **D9:** lymphoma; **D10:** melanoma; **D11:** nervous system cancer; **D12:** neurodegenerative disease; **D13:** non-small cell lung carcinoma; **D14:** organ system cancer; **D15:** prostate cancer;

References

[1] M. Collado, V. Garcia, J. M. Garcia, I. Alonso, L. Lombardia, R. Diaz-Uriarte, L. A. Lpez Fernndez, A. Zaballos, F. Bonilla, M. Serrano, Genomic profiling of circulating plasma rna for the analysis of cancer, Clinical Chemistry 53 (10) (2007) 1860–1863.

[2] S. Daniele, S. Aleksandra, Z. Agnieszka, K. Marta, B. Marek, B. Paola, F. Giovanni, An Integrated Approach (CLuster Analysis Integration Method) to Combine Expression Data and Protein-Protein Interaction Networks in Agrigenomics: Application on Arabidopsis thaliana, OMICS: A Journal of Integrative Biology 18 (2) (2014) 155–165.

[3] C. Ding, H. Peng, Minimum Redundancy Feature Selection from Microarray Gene Expression Data, Journal of Bioinformatics and Computational Biology 3 (2) (2005) 185–205.

[4] T. R. Golub, D. K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J. P. Mesirov, H. Coller, M. L. Loh, J. R. Downing, M. A. Caligiuri, C. D. Bloomfield, E. S. Lander, Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring, Science 286 (5439) (1999) 531–537.

[5] G. J. Gordon, R. V. Jensen, L. L. Hsiao, S. R. Gullans, J. E. Blumenstock, S. Ramaswamy, W. G. Richards, D. J. Sugarbaker, R. Bueno, Translation of Microarray Data into Clinically Relevant Cancer Diagnostic Tests Using Gene Expression Ratios in Lung Cancer and Mesothelioma, Cancer Research 62 (2002) 4963–4967.

[6] T. Hinoue, D. J. Weisenberger, C. P. Lange, H. Shen, H. M. Byun, D. Van Den Berg, S. Malik, F. Pan, H. Noushmehr, C. M. van Dijk, R. A. E. M. Tollenaar, P. W. Laird, Genome-Scale Analysis of Aberrant DNA Methylation in Colorectal Cancer, Genome Research 22 (2) (2012) 271–282.

[7] X. Jiang, J. Tan, J. Li, S. Kivime, X. Yang, L. Zhuang, P. L. Lee, M. T. Chan, L. W. Stanton, E. T. Liu, B. N. Cheyette, Q. Yu, {DACT3} is an epigenetic regulator of wnt/-catenin signaling in colorectal cancer and is a therapeutic target of histone modifications, Cancer Cell 13 (6) (2008) 529 – 541.

[8] B.-Q. Li, T. Huang, L. Liu, Y.-D. Cai, K.-C. Chou, Identification of Colorectal Cancer Related Genes with mRMR and Shortest Path in Protein-Protein Interaction Network, PLoS ONE 7 (4) (2012) e33393.

[9] P. Maji, S. Paul, Rough Set Based Maximum Relevance-Maximum Significance Criterion and Gene Selection from Microarray Data, International Journal of Approximate Reasoning 52 (3) (2011) 408–426.

[10] P. Maji, S. Paul, Scalable Pattern Recognition Algorithms: Applications in Computational Biology and Bioinformatics, Springer-Verlag, London, 2014, p. 304.

[11] C. L. Noble, A. R. Abbas, J. Cornelius, C. W. Lees, G.-T. Ho, K. Toy, Z. Modrusan, N. Pal, F. Zhong, S. Chalasani, H. Clark, I. D. Arnott, I. D. Penman, J. Satsangi, L. Diehl, Regional variation in gene expression in the healthy colon is dysregulated in ulcerative colitis, Gut 57 (10) (2008) 1398–1405.

[12] S. Paul, P. Maji, Gene Expression and Protein-Protein Interaction Data for Identification of Colon Cancer Related Genes Using f-Information Measures, Natural Computing DOI:10.1007/s11047-015-9485-6 (2016) 1–15.

[13] B. M. Ryan, K. A. Zanetti, A. I. Robles, A. J. Schetter, J. Goodman, R. B. Hayes, W.-Y. Huang, M. J. Gunter, M. Yeager, L. Burdette, S. I. Berndt, C. C. Harris, Germline variation in ncf4, an innate immunity gene, is associated with an increased risk of colorectal cancer, International Journal of Cancer 134 (6) (2014) 1399–1407.

[14] M. A. Shipp, K. N. Ross, P. Tamayo, A. P. Weng, J. L. Kutok, R. C. Aguiar, M. Gaasenbeek, M. Angelo, M. Reich, G. S. Pinkus, T. S. Ray, M. A. Koval, K. W. Last, A. Norton, T. A. Lister, J. Mesirov, D. S. Neuberg, E. S. Lander, J. C. Aster, T. R. Golub, Diffuse Large B-Cell Lymphoma Outcome Prediction by Gene Expression Profiling and Supervised Machine Learning, Nature Medicine 8 (1) (2002) 68–74.

[15] N. Vlassis, E. Glaab, GenePEN: Analysis of Network Activity Alterations in Complex Diseases via the Pairwise Elastic Net, Statistical Applications in Genetics and Molecular Biology 14 (2) (2015) 221–224.

[16] M. West, C. Blanchette, H. Dressman, E. Huang, S. Ishida, R. Spang, H. Zuzan, J. A. Olson, J. R. Marks, J. R. Nevins, Predicting the Clinical Status of Human Breast Cancer by Using Gene Expression Profiles, Proceedings of the National Academy of Sciences, USA 98 (20) (2001) 11462–11467.