Data Mining of MicroRNAs in Breast Carcinogenesis Which May Be a Potential Target for Cancer Prevention

Jin-Wook Kang^{1,2}, Min-Ji Kim^{1,2}, Hyun-Ah Baek³, and Jeong-Sang Lee^{1,2,3,*}

¹Molecular Mechanisms of Functional Foods Laboratory, Jeonju University, Jeonju, Jeonbuk 55069, Korea

²Department of Functional Food and Biotechnology, College of Medical Science, Jeonju University, Jeonju, Jeonbuk 55069, Korea

³Food Industry Research Institute, Jeonju University, Jeonju, Jeonbuk 55069, Korea

Received December 13, 2015
Revised February 29, 2016
Accepted March 2, 2016
Published online March 31, 2016

*Corresponding Author
Tel: +82-63-220-4660
Fax: +82-63-220-2054
E-mail: jslee11@jj.ac.kr

pISSN 1226-7708
eISSN 2092-6456

© KoSFoST and Springer 2016

Abstract The microRNAs (miRNAs) negatively regulate the stability and translation of target messenger RNAs by selectively binding. It has been implicated in diverse processes such as cellular differentiation, cell-cycle control, apoptosis, and carcinogenesis. Examination of tumor-specific miRNA expression profiles has revealed wide spread dysregulation of these molecules in diverse cancers. The available genomic bulk evidences were extracted from The Cancer Genome Atlas by using IlluminaGA_miRNASeq platform in human breast cancer samples. After mining collected data, group of each miRNA ID was analyzed through five D/Bs (mirWalk, miranda, mirDB, RNA22, and TargetScan) on predicted and validated miRNA targets. Oncogenes known to have a high correlation with breast cancer (*C-myc*, *HER2*, *cyclin D-1*, *N-RAS*, *FGF-4*, *FGF-3*, *BRCA1*, and *BRCA2*) are subject in this study to select their relevant miRNAs. Function of miRNA regulation will be essential to achieve a complete understanding of carcinogenesis and these miRNAs would be potential target for breast cancer prevention.

Keywords: microRNA, breast cancer, data mining, chemoprevention, target

Introduction

Breast cancer is a highly heterogeneous disease that has multiple subtypes with distinct clinical outcomes. Clinically, breast cancers are classified by hormone receptor status, including expression of estrogen receptor (ER), progesterone receptor (PR), and human EGF-like receptor type 2 (HER2) receptor (1). Oncogenes such as *FGF-3*, *FGF-4*, *Cyclin D1* (*PRAD-1*), *HER2*, *N-RAS*, *c-Myc*, *BRCA1*, and *BRCA2* have a high correlation with breast cancer. Fibroblast growth factor 3/4 (*FGF3/4*) is member of the fibroblast growth factor (FGF) families. FGF families are involved in mitogenic, cell survival activities, biological processes, cell growth, morphogenesis, tissue repair, tumor growth, and invasion (2). Frequent amplification of this *FGF* family gene has been found in human tumors, which may be important for neoplastic transformation and tumor progression (2). *RAS*-gene family is associated with a variety of human cancer. *N-RAS* is an oncogene encoding a membrane protein that shuttles between the Golgi apparatus and the plasma membrane (3). This shuttling is regulated through palmitoylation and depalmitoylation by the ZDHHC9-GOLGA7 complex (4). A variable form of *RAS* is mainly found in the growth sites of breast cancer epithelial cells. Mutation of this gene, easily happen on the amino acid residues 12 and 13, promotes the cell transformation (5). Cyclin D1, which is also known as *PRAD-1*,

forms a complex and works as a regulatory subunit of CDK4 or CDK6 (6). Cyclin D1 is required for cell cycle G1/S transition (7). Mutation, amplification, and over-expression of this gene, which alters cell cycle progression, are observed frequently in a variety of tumors and may contribute to tumorigenesis (7). *Cyclin D1* over-expression is strongly correlated to ER(+) breast cancer and deregulation of *cyclin D1* is associated with hormone therapy resistance in breast cancer (8). *Human epidermal growth factor receptor* type 2 (*HER2*) over-expression has been known to play an important role in the development and progression of breast cancer (9). *c-Myc* is involved that regulation of cellular proliferation, differentiation, apoptosis, and angiogenesis (10). Over-expression of *c-Myc* was found to be associated with its amplification, suggesting that *c-Myc* amplification may have a key oncogenic role in the pathology of breast cancer (10). *Breast Cancer 1/2* (*BRCA1/BRCA2*) act as a tumor suppressor encodes a nuclear phosphoprotein that plays a role in maintaining genomic stability (11). *BRCA1/BRCA2* associates with RNA polymerase II by the c-terminal domain and also interacts with histone deacetylase complexes (11). Thus, this protein plays a role in transcription, DNA repair of double-stranded breaks, and recombination.

MicroRNAs (miRNAs) are a class of non-coding small RNA molecules (20-25 nucleotides) that negatively regulate the gene expression by binding their target mRNA. *Lin-4*, the first miRNA, discovered by the

Victor Ambros group in 1993 and this small molecule regulates the generation time of *Caenorhabditis elegans* (*C. elegans*) by affecting the synthesis of LIN-14 protein. The miRNAs comprise 1 to 3% of human genome but they are known to regulate more than 30% of human genes (12). The miRNAs have been known to control the cell growth, differentiation, cell proliferation, cell-cycle, apoptosis, and carcinogenesis (13). Interestingly, change of miRNA expression is related with various diseases (14). A single miRNA can regulate a variety of target genes and a single gene can be simultaneously regulated by multiple miRNAs vice versa (13,15-17). In this aspect, the miRNAs may have a co-targeting network with their multi-targeting capability. However, most of studies demonstrated validation of single miRNA-single target. These studies require a lot of time and cost (18). In an era of big data, we already have vast data such as miRNA expression profiling and predicted or validated miRNA-target data. Amount of accumulating biological data give us opportunity to find a new valuable meanings. Therefore, analyzing tool and method that can extract refined data from bulk data are very important to understand biological complexity and to find new meaning. A number of array-platform-based studies showed that miRNA expression profiling results between array studies are not consistent and huge (19). Compared to miRNA array-platform, the miRNA deep sequencing technology provides massive results of genome-widely expressed miRNAs, expression level of individual miRNAs, and novel miRNAs.

The Cancer Genome Atlas (TCGA), open access portal is a research network that has analyzed large number of human cancer sample related with DNA, RNA, protein, and miRNA. The TCGA miRNA sequencing data includes resource for both tumor and adjacent normal samples clinically. Big data about miRNA expression change in human sample which is obtained from the TCGA offer better information of miRNAs dysregulation in breast cancer. Changed miRNAs may mean reliable biomarker and potential therapeutic targets (20). To understand the mechanism of miRNA regulation, we need targets that are regulated by changed miRNAs. The miRNAs usually bind to the 3'-untranslated region (3'UTR) of their mRNA target and this condition triggers the repression of translation or degradation of the target according to the complementary sequencing match between the target 3'UTR and seed site of miRNA. Many of algorithms, like mirWalk (21), Miranda (22), mirDB (23), RNA22 (24), and TargetScan (25) seek putative miRNA for given target sequences based on canonical miRNA binding. However, since these programs have slightly different criteria, each of them predicts different results (26). So, finding a functional miRNA and their target is a challenging. To identify more reliable putative miRNA and target, we need a cross-check between available data bases. In this study, we were focusing on the regulation networking of miRNAs which can regulate multi-oncogenes. The aim of this study was to elucidate the common miRNAs among the target oncogenes which may be critical resource for breast cancer chemoprevention.

Materials and Methods

miRNA bulk data collection from human cancer sample TCGA (<http://cancergenome.nih.gov>) provides opportunity for researchers to download a wide range of information by analyzing the clinical cancer sample. Five columns to select such as cancer type, center, platform, data type, and archive type exist in the section of bulk download. We chose breast invasive carcinoma as cancer type and used the Illumina Genome Analyzer miRNA sequencing as platform in the column. There are four types of information: miRNA ID, read count, reads per million (RPM), and cross-mapped.

miRNA sequencing data contain the quantitative expression levels of miRNA that are computed based on the read counts in one sample. For comparison of miRNA expression in two or more samples, expression levels in each sequenced sample are normalized by the total number of reads in the sample. RPM is a value of normalized expression level that were produced for each sample. Therefore, we use miRNA ID and RPM because we need an expression level of miRNAs in data.

Sorting Single-molecule-based platforms such as Illumina Genome analyzer platform usually show an inherent high error because this platform detected all sequences containing the errors. High error rates cause low read counts. According to the difference of miRNA expression level above five-folds (≥ 5), two and half-folds (≥ 2.5), and another (< 2.5), the miRNA IDs were divided into 3 groups. The low level expression (< 2.5) and/or no changes miRNA IDs in the collected bulk data were removed to avoid error.

Double-checking The mirWalk is supporting not only a documentary system of predicted miRNA-target interactions, but also comparing this information with another miRNA-target prediction programs (DIANA-microT, miranda, miRBridge, miRDB, miRmap, miRNA Map, PITA, RNA22, and Targetscan). To investigate the interaction of miRNA IDs and molecular target, the available miRNA IDs were double-checked through five reliable D/Bs (mirWalk, miRanda, miRDB, RNA22, and TargetScan). Based on the SUM value, over 3 in the five D/Bs, miRNA IDs is sorting subject in this study. By double-checking with comparing five D/Bs, we could obtain three facts. The first thing was a correlation between target gene and miRNA IDs. The second thing was reliability of predicted miRNAs belong to the specific gene. The last thing was completed removal of meaningless miRNA IDs.

Clustering Each target gene have multiple miRNA IDs that expressed over five-folds (≥ 5), two and half-folds (≥ 2.5), and below that. Extracted miRNAs were speculated that those are the oncogenic miRNAs which may promote the breast carcinogenesis via target gene. Each miRNA ID data was clustering to drain the common one

Table 1. Raw data downloaded by TCGA

miRNA ID	Read count	Reads per million miRNA mapped	miRNA ID	Read count	Reads per million miRNA mapped	miRNA ID	Read count	Reads per million miRNA mapped	miRNA ID	Read count	Reads per million miRNA mapped	miRNA ID	Read count	Reads per million miRNA mapped
let-7a-1	16795	6280.683542	miR-136	123	45.997266	miR-205	6584	2462.162575	miR-3174	1	0.373962	miR-3200	11	4.113577
let-7a-2	33593	12562.48897	miR-138-1	1	0.373962	miR-206	37	13.836576	miR-3177	1	0.373962	miR-320a	1227	458.850771
let-7a-3	16999	6356.97169	miR-138-2	2	0.747923	miR-20a	241	90.124724	miR-3190	1	0.373962	miR-320b-2	20	7.47923
let-7b	118911	44468.13698	miR-139	151	56.468188	miR-20b	91	34.030497	miR-32	47	17.576191	miR-320c-1	1	0.373962
let-7c	11387	4258.299702	miR-140	952	356.011356	miR-21	556738	208198.5825	miR-3200	11	4.113577	miR-320c-2	4	1.495846
let-7d	1107	413.97539	miR-141	3958	1480.139652	miR-210	1090	407.618045	miR-320a	1227	458.850771	miR-320d-1	6	2.243769
let-7e	2123	793.920283	miR-142	3180	1189.197598	miR-2110	9	3.365654	miR-320b-2	20	7.47923	miR-320d-2	4	1.495846
let-7f-1	108	40.387843	miR-143	107718	40282.38582	miR-2114	2	0.747923	miR-320c-1	1	0.373962	miR-323	3	1.121885
let-7f-2	10381	3882.094424	miR-144	93	34.77842	miR-2115	1	0.373962	miR-320c-2	4	1.495846	miR-323b	18	6.731307
let-7g	1000	373.961509	miR-145	3391	1268.103477	miR-2116	1	0.373962	miR-320d-1	6	2.243769	miR-324	49	18.324114
let-7i	1013	378.823009	miR-146	9	3.365654	miR-2117	1	0.373962	miR-320d-2	4	1.495846	miR-326	9	3.365654
miR-1-2	24	8.975076	miR-146a	146	54.59838	miR-212	2	0.747923	miR-323	3	1.121885	miR-328	55	20.567883
miR-100	11132	4162.939517	miR-146b	898	335.817435	miR-214	61	22.811652	miR-323b	18	6.731307	miR-329-2	1	0.373962
miR-101-1	55076	20596.30406	miR-147b	2	0.747923	miR-215	6	2.243769	miR-324	49	18.324114	miR-330	38	14.210537
miR-101-2	237	88.628878	miR-148a	115226	43090.08882	miR-216a	1	0.373962	miR-326	9	3.365654	miR-331	58	21.689768
miR-103-1	33856	12660.84085	miR-148b	729	272.61794	miR-217	18	6.731307	miR-328	55	20.567883	miR-335	232	86.75907
miR-103-2	20	7.47923	miR-149	290	108.448838	miR-218-1	5	1.869808	miR-329-2	1	0.373962	miR-337	108	40.387843
miR-106a	36	13.462614	miR-150	988	369.473971	miR-218-2	59	22.063729	miR-330	38	14.210537	miR-338	523	195.581869
miR-106b	1056	394.903353	miR-151	5644	2110.638756	miR-219-1	3	1.121885	miR-331	58	21.689768	miR-339	59	22.063729
miR-107	161	60.207803	miR-152	642	240.083289	miR-219-2	1	0.373962	miR-335	232	86.75907	miR-33a	23	8.601115
miR-10a	42900	16042.94873	miR-153-1	2	0.747923	miR-22	129277	48344.62199	miR-337	108	40.387843	miR-340	44	16.454306
miR-10b	440954	164899.8232	miR-153-2	35	13.088653	miR-221	126	47.11915	miR-338	523	195.581869	miR-342	374	139.861604
miR-1180	17	6.357346	miR-154	27	10.096961	miR-222	64	23.933537	miR-339	59	22.063729	miR-345	29	10.844884
miR-1185-2	1	0.373962	miR-155	388	145.097065	miR-223	128	47.867073	miR-33a	23	8.601115	miR-34a	134	50.110842
miR-1226	2	0.747923	miR-15a	246	91.994531	miR-224	14	5.235461	miR-33b	14	5.235461	miR-34b	7	2.617731
miR-1228	1	0.373962	miR-15b	347	129.764644	miR-227	1	0.373962	miR-340	44	16.454306	miR-34c	22	8.227153
miR-1229	1	0.373962	miR-16-1	2314	865.346932	miR-2355	127	47.493112	miR-342	374	139.861604	miR-3605	7	2.617731
miR-1245	21	7.853192	miR-16-2	19	7.105269	miR-23a	9321	3485.695224	miR-345	29	10.844884	miR-3607	246	91.994531
miR-1247	389	145.471027	miR-17	554	207.174676	miR-23b	2616	978.283307	miR-34a	134	50.110842	miR-361	504	188.4766
miR-1248	4	1.495846	miR-181a-1	3560	1331.302972	miR-23c	3	1.121885	miR-34b	7	2.617731	miR-3610	7	2.617731
miR-1249	2	0.747923	miR-181a-2	1358	507.839729	miR-24-1	27	10.096961	miR-34c	22	8.227153	miR-3613	24	8.975076
miR-1250	2	0.747923	miR-181b-1	653	244.196865	miR-24-2	1859	695.194445	miR-3605	7	2.617731	miR-3614	4	1.495846
miR-1251	5	1.869808	miR-181b-2	8	2.991692	miR-25	22649	8469.854215	miR-3607	246	91.994531	miR-3615	8	2.991692
miR-1258	2	0.747923	miR-181c	345	129.016721	miR-26a-1	39	14.584499	miR-361	504	188.4766			
miR-125a	1910	714.266482	miR-181d	56	20.941844	miR-26a-2	5656	2115.126294	miR-3610	7	2.617731			
miR-125b-1	1551	580.0143	miR-182	235196	87954.25105	miR-26b	1653	618.158374	miR-3613	24	8.975076			
miR-125b-2	65	24.307498	miR-183	71825	26859.78538	miR-27a	1732	647.701333	miR-3614	4	1.495846			
miR-126	7203	2693.644749	miR-184	63	23.559575	miR-27b	1288	481.662423	miR-3615	8	2.991692			

Table 1. Continued

miRNA ID	Read count	Reads per million miRNA mapped	miRNA ID	Read count	Reads per million miRNA mapped	miRNA ID	Read count	Reads per million miRNA mapped	miRNA ID	Read count	Reads per million miRNA mapped
miR-1262	5	1.869808	miR-185	179	66.93911	miR-28	15433	5771.347967	miR-362	2	0.747923
miR-1266	44	16.454306	miR-186	538	201.191292	miR-296	9	3.365654	miR-363	46	17.202229
miR-1269	3	1.121885	miR-187	7	2.617731	miR-299	13	4.8615	miR-3647	22	8.227153
miR-127	2222	830.942473	miR-188	3	1.121885	miR-29a	9670	3616.207791	miR-3648	1	0.373962
miR-1270-1	3	1.121885	miR-18a	6	2.243769	miR-29b-1	427	159.681564	miR-365-1	89	33.282574
miR-1271	14	5.235461	miR-18b	1	0.373962	miR-29b-2	500	186.980754	miR-365-2	104	38.891997
miR-1274b	3	1.121885	miR-190	15	5.609423	miR-29c	4618	1726.954248	miR-3651	1	0.373962
miR-1277	5	1.869808	miR-190b	252	94.2383	miR-301a	9	3.365654	miR-3652	1	0.373962
miR-128-1	163	60.955726	miR-191	2105	787.188976	miR-301b	1	0.373962	miR-3653	14	5.235461
miR-128-2	120	44.875381	miR-192	1012	378.449047	miR-305	286	106.952992	miR-3656	1	0.373962
miR-1287	54	20.193921	miR-193a	698	261.025133	miR-3074	49	18.324114	miR-3661	1	0.373962
miR-129-1	1	0.373962	miR-193b	187	69.930802	miR-30a	233790	87428.46116	miR-3664	4	1.495846
miR-129-2	2	0.747923	miR-194-1	309	115.554106	miR-30b	779	291.316015	miR-3676	6	2.243769
miR-1291	1	0.373962	miR-194-2	293	109.570722	miR-30c-1	2	0.747923	miR-3677	34	12.714691
miR-1292	2	0.747923	miR-195	228	85.263224	miR-30c-2	2519	942.009041	miR-3678	1	0.373962
miR-1293	1	0.373962	miR-196a-1	1554	581.136185	miR-30d	28088	10503.83086	miR-3679	1	0.373962
miR-1294	1	0.373962	miR-196a-2	92	34.404459	miR-30e	37296	13947.26844	miR-3680	1	0.373962
miR-1295	8	2.991692	miR-196b	547	204.556945	miR-31	20	7.47923	miR-3682	2	0.747923
miR-1296	4	1.495846	miR-197	251	93.864339	miR-3125	1	0.373962	miR-3684	1	0.373962
miR-1301	15	5.609423	miR-1976	20	7.47923	miR-3127	54	20.193921	miR-369	27	10.096961
miR-1306	12	4.487538	miR-199a-1	5997	2242.647169	miR-3128	1	0.373962	miR-3690	1	0.373962
miR-1307	1501	561.316225	miR-199a-2	10528	3937.066766	miR-3129	1	0.373962	miR-370	30	11.218845
miR-130a	123	45.997266	miR-199b	13648	5103.826673	miR-3130-1	11	4.113577	miR-374a	1894	708.283098
miR-130b	34	12.714691	miR-19a	61	22.811652	miR-3136	3	1.121885	miR-374b	169	63.199495
miR-132	121	45.249343	miR-19b-1	17	6.357346	miR-3148	1	0.373962	miR-374c	14	5.235461
miR-133a-1	5	1.869808	miR-19b-2	206	77.036071	miR-3156-1	4	1.495846	miR-375	11025	4122.925636
miR-133a-2	1	0.373962	miR-200a	3091	1155.915024	miR-3156-2	1	0.373962	miR-376a-1	2	0.747923
miR-133b	2	0.747923	miR-200b	2021	755.776209	miR-3156-3	4	1.495846	miR-376a-2	2	0.747923
miR-134	1475	551.593226	miR-200c	26793	10019.55071	miR-3158-1	1	0.373962	miR-376b	3	1.121885
miR-135a-1	40	14.95846	miR-202	6	2.243769	miR-3161	1	0.373962	miR-376c	9	3.365654
miR-135a-2	3	1.121885	miR-203	6557	2452.065614	miR-3170	2	0.747923	miR-377	9	3.365654
miR-135b	15	5.609423	miR-204	8	2.991692	miR-3173	1	0.373962	miR-378	236	88.254916
miR-378c	11	4.113577	miR-450b	140	52.354611	miR-518f	1	0.373962	miR-616	12	4.487538
miR-379	2966	1109.169835	miR-451	192	71.80061	miR-519a-1	1	0.373962	miR-618	2	0.747923
miR-380	2	0.747923	miR-452	23	8.601115	miR-520g	1	0.373962	miR-624	1	0.373962
miR-381	96	35.900305	miR-454	12	4.487538	miR-520h	1	0.373962	miR-625	738	275.983594
miR-382	93	34.77842	miR-455	388	145.097065	miR-526b	2	0.747923	miR-627	4	1.495846
miR-383	3	1.121885	miR-483	5	1.869808	miR-532	779	291.316015	miR-628	108	40.387843

Table 1. Continued

miRNA ID	Read count	Reads per million miRNA mapped	miRNA ID	Read count	Reads per million miRNA mapped	miRNA ID	Read count	Reads per million miRNA mapped	miRNA ID	Read count	Reads per million miRNA mapped
miR-3909	2	0.747923	miR-484	91	34.030497	miR-539	35	13.088653	miR-629	386	144.349142
miR-3912	1	0.373962	miR-485	7	2.617731	miR-542	1331	497.742768	miR-632	3	1.121885
miR-3913-1	9	3.365654	miR-486	59	22.063729	miR-543	2	0.747923	miR-636	1	0.373962
miR-3917	4	1.495846	miR-487a	2	0.747923	miR-545	1	0.373962	miR-642a	6	2.243769
miR-3923	1	0.373962	miR-487b	13	4.8615	miR-548a-3	1	0.373962	miR-643	1	0.373962
miR-3926-1	3	1.121885	miR-491	9	3.365654	miR-548b	2	0.747923	miR-651	10	3.739615
miR-3928	3	1.121885	miR-493	32	11.966768	miR-548e	3	1.121885	miR-652	45	16.828268
miR-3934	4	1.495846	miR-494	6	2.243769	miR-548j	1	0.373962	miR-653	18	6.731307
miR-3936	1	0.373962	miR-495	17	6.357346	miR-548t	2	0.747923	miR-654	42	15.706383
miR-3940	1	0.373962	miR-496	5	1.869808	miR-548v	5	1.869808	miR-655	5	1.869808
miR-409	86	32.16069	miR-497	65	24.307498	miR-550a-1	2	0.747923	miR-656	5	1.869808
miR-410	28	10.470922	miR-499	1	0.373962	miR-550a-2	3	1.121885	miR-658	1	0.373962
miR-411	21	7.853192	miR-500a	343	128.268798	miR-551a	1	0.373962	miR-659	3	1.121885
miR-412	26	9.722999	miR-500b	4	1.495846	miR-551b	1	0.373962	miR-660	43	16.080345
miR-421	6	2.243769	miR-501	74	27.673152	miR-556	9	3.365654	miR-664	84	31.412767
miR-423	314	117.423914	miR-502	9	3.365654	miR-561	1	0.373962	miR-671	17	6.357346
miR-424	660	246.814596	miR-503	121	45.249343	miR-574	139	51.98065	miR-675	42	15.706383
miR-425	780	291.689977	miR-504	4	1.495846	miR-576	27	10.096961	miR-676	2	0.747923
miR-429	318	118.91976	miR-505	67	25.055421	miR-577	2	0.747923	miR-7-1	75	28.047113
miR-4306	1	0.373962	miR-508	13	4.8615	miR-578	1	0.373962	miR-7-3	3	1.121885
miR-431	15	5.609423	miR-509-1	2	0.747923	miR-581	1	0.373962	miR-708	146	54.59838
miR-432	11	4.113577	miR-509-2	2	0.747923	miR-582	294	109.944684	miR-720	1	0.373962
miR-4326	8	2.991692	miR-511-1	6	2.243769	miR-584	100	37.396151	miR-744	113	42.257651
miR-433	1	0.373962	miR-511-2	7	2.617731	miR-585	1	0.373962	miR-758	44	16.454306
miR-449a	21	7.853192	miR-513c	1	0.373962	miR-589	266	99.473761	miR-760	8	2.991692
miR-449b	5	1.869808	miR-514-1	1	0.373962	miR-590	32	11.966768	miR-766	6	2.243769
miR-449c	2	0.747923	miR-514-2	1	0.373962	miR-592	2	0.747923	miR-767	1	0.373962
miR-450a-1	17	6.357346	miR-514-3	1	0.373962	miR-598	23	8.601115	miR-769	53	19.81996
miR-450a-2	12	4.487538	miR-518a-1	1	0.373962	miR-615	15	5.609423	miR-874	30	11.218845
miR-877	1	0.373962	miR-92a-2	10702	4002.136068	miR-942	6	2.243769	miR-9-2	342	127.894836
miR-887	13	4.8615	miR-92b	381	142.479335	miR-944	16	5.983384	miR-92a-1	617	230.734251
miR-889	38	14.210537	miR-93	8778	3282.634125	miR-95	12	4.487538	miR-939	3	1.121885
miR-891a	1	0.373962	miR-934	3	1.121885	miR-96	85	31.786728	miR-940	8	2.991692
miR-9-1	373	139.487643	miR-937	1	0.373962	miR-98	112	41.883689	miR-99a	3865	1445.361232
miR-99b	89093	33317.35271									

which will be more effective and new target for breast cancer prevention.

Results and Discussion

The breast cancer is mainly related with estrogen receptor, progesterone receptor, and HER2 receptor. We selected 8 oncogenes (*BRCA1*, *BRCA2*, *HER2*, *Cyclin D1*, *N-RAS*, *FGF3*, *FGF4*, and *c-Myc*) and their miRNAs have high correlation molecular target for breast cancer. Table 1 has raw data about breast cancer downloaded by TCGA. Predicted miRNA list concerning the each oncogenes are classified according to the expression level (Table 2 to 5). Putative miRNA list that may regulate the *BRCA1* are as follows: miRNA above five-folds (≥ 5) occupied 11% [52 miRNAs] of the total followed by two and half-folds (≥ 2.5) occupied 1.5% [7 miRNAs] and regulate the *BRCA2* above five folds (≥ 5) occupied 5% [26 miRNAs] of the total followed by two and half-folds (≥ 2.5) occupied 0.4% [2 miRNAs] (Table 2). In the Table 3, *Cyclin D1* regulating miRNAs were suggested that above five-folds (≥ 5) occupied 18% [83 miRNAs] of the total followed by two and half-folds (≥ 2.5) occupied 2% [9 miRNAs]. Regarding oncogene *N-RAS*-regulating miRNAs were listed up Table 4, miRNAs above five-folds (≥ 5) occupied 23% [102 miRNAs] of the total followed by two and half-folds (≥ 2.5) occupied 2.9% [13 miRNAs]. Table 5 suggested that *FGF3* and *FGF4* regulating miRNA. miRNA above five-folds (≥ 5)

occupied 0.4% [2 miRNAs] of the total followed by two and half-folds (≥ 2.5) occupied 0.2% [1 miRNA]. In case of oncogene *FGF4*, above five-folds (≥ 5) miRNAs occupied 0.4% [2 miRNAs] of the total followed by two and half-folds (≥ 2.5) occupied 0.2% [1 miRNA]. In the oncogene *HER2* case, miRNAs above five-folds (≥ 5) occupied 2.9% [13 miRNAs] subsequently followed by two and half-folds (≥ 2.5) occupied 0.4% [2 miRNA]. In the case of *c-Myc*, 0.2% of miRNAs were above five-folds (≥ 5) [1 miRNAs].

Although there was a no universal miRNA targeting all 8 oncogenes, however, there were revealed some meaningful common miRNAs. As illustrated in Table 6, results of overlapped common miRNAs among the 8 oncogenes showed that 28 miRNAs were commonly regulating three above oncogenes target. Let-7c is controlling the 5 oncogenes (*BRCA1*, *BRCA2*, *HER2*, *N-RAS*, and *Cyclin D1*) among the miRNA above five-folds occupied. In five-folds (≥ 5) raw data 4 miRNAs, let-7g, miR-186, miR-589, and miR-224 were commonly regulating among four oncogenes target. 21 miRNAs (let-7b, let-7e, let-7i, miR-30d, miR-93, miR-205, miR-106b, miR-146b, miR-424, miR-17, miR-20a, miR-146a, miR-20b, miR-769, miR-26a-1, miR-106a, miR-625, miR-98, miR-19a, miR-3613, and miR-361) were commonly regulating among three oncogenes target. In Table 6, in two and half-folds (≥ 2.5) raw data miR-651 was repeating on targeting the 3 oncogenes (*BRCA1*, *BRCA2*, and *Cyclin D1*). miR-1468 was also repeating 3 oncogenes (*BRCA1*, *N-RAS*, and *c-Myc*).

The miRNA can regulate many cell mechanisms, including cell

Table 2. *BRCA1/2* related microRNAs list up

Oncogene	$\geq 2.5 \sim < 5$			≥ 5		
<i>BRCA1</i>	miR-1306	let-7a-2	miR-106b	miR-629	miR-221	miR-758
	miR-95	miR-30d	let-7g	miR-589	miR-132	miR-654
	miR-651	miR-28	miR-150	miR-197	miR-584	miR-26a-1
	miR-1468	miR-199b	miR-140	miR-20a	miR-20b	miR-106a
	miR-3615	let-7c	miR-146b	miR-335	miR-484	miR-493
	miR-760	miR-93	miR-424	miR-193b	miR-222	miR-345
	miR-187	miR-205	miR-17	miR-146a	miR-328	miR-576
		miR-26a-2	miR-186	miR-708	miR-769	miR-1976
		miR-127	miR-29b-2	miR-450b	miR-324	miR-671
		miR-125a	miR-29b-1	miR-34a	miR-1266	miR-944
		miR-615	miR-224			
	7			52		
	miR-616	miR-21	miR-429	miR-425		
	miR-651	miR-182	miR-146a	miR-625		
<i>BRCA2</i>		let-7b	miR-744			
		miR-200c	miR-98			
		let-7c	miR-628			
		miR-29a	miR-19a			
		miR-205	miR-769			
		let-7e	miR-369			
		let-7i	miR-323b			
		miR-192	miR-653			
		let-7g	miR-224			
		miR-146b	miR-532			
	2			26		

growth, cell proliferation, cell-cycle, apoptosis, and carcinogenesis (13). Over-expressed miRNA in specific diseases suggest that miRNA can promote related diseases via suppressing target gene (27). Since the disease specific-miRNA is found, miRNA is receiving attention as a potential chemoprevention agent (27). It is important to understanding the biological role of miRNA which regulate their corresponding mRNA targets. However, due to the various miRNAs and potential targets, validating the miRNA-target would be extremely difficult and not economical. Various computational

analysis methods have been generated the putative miRNA target. The miRNA data base such as miRWalk, miRanda, miRDB, RNA22, and TargetScan provide predicted miRNA-target data and use different algorithm. These programs often predict distinct miRNA target. Kuhn *et al.* (28) reported that when Angiotensin II Receptor Type 1 (AGTR1) was analyzed for identifying predicted miRNA, miRanda predicted 27 miRNA within the AGTR1 3'-UTR and TargetScan showed 37 miRNA. In contrast to two data base, PicTar show no miRNA list. Therefore, cross-checking the bioinformatics

Table 3. *Cyclin D1* related microRNA list up

Oncogene	$\geq 2.5 \sim < 5$			≥ 5		
<i>Cyclin D1</i>	miR-1306	miR-21	miR-106b	miR-500a	miR-503	miR-1266
	miR-95	let-7b	let-7i	miR-3065	miR-128-2	miR-340
	miR-432	let-7c	let-7g	miR-589	miR-98	miR-758
	miR-651	miR-92a-2	miR-150	miR-197	miR-584	miR-654
	miR-296	miR-23a	miR-140	miR-15a	miR-20b	miR-26a-1
	miR-326	miR-93	miR-625	miR-3607	miR-96	miR-330
	miR-502	miR-26a-2	miR-193a	miR-20a	miR-7-1	miR-206
	miR-204	miR-145	miR-424	miR-195	miR-497	miR-106a
	miR-187	miR-142	miR-17	miR-193b	miR-19a	miR-3677
		miR-23b	miR-186	miR-374b	miR-214	miR-370
		miR-127	miR-338	miR-128-1	miR-486	miR-874
		let-7e	miR-361	miR-450b	miR-1287	miR-154
		miR-374a	miR-155	miR-34a	miR-3127	miR-576
		miR-26b	miR-342	miR-221	miR-769	miR-3613
		miR-134	miR-15b	miR-130a	miR-324	miR-33a
		miR-452	miR-495	miR-224	miR-1271	
		miR-449a	miR-615	miR-374c	miR-653	
	9			83		

Table 4. *N-RAS* related microRNA list up

Oncogene	$\geq 2.5 \sim < 5$			≥ 5		
<i>N-RAS</i>	miR-1306	miR-182	miR-145	miR-17	miR-337	miR-889
	miR-450a-2	miR-30a	miR-200a	miR-196b	miR-381	miR-206
	miR-616	miR-22	miR-379	miR-186	miR-144	miR-106a
	miR-95	let-7b	miR-16-1	miR-338	miR-382	miR-539
	miR-378c	miR-148a	let-7e	miR-361	miR-20b	miR-3677
	miR-1468	miR-143	miR-200b	miR-629	miR-96	miR-590
	miR-2110	miR-183	miR-374a	miR-92b	miR-7-1	miR-874
	miR-377	miR-30e	miR-26b	miR-181c	miR-501	miR-369
	miR-502	miR-30d	miR-134	miR-3065	miR-125b-2	miR-3613
	miR-204	miR-200c	miR-181a-2	miR-589	miR-497	miR-33a
	miR-760	miR-25	let-7d	miR-20a	miR-222	miR-452
	miR-34b	miR-28	miR-106b	miR-335	miR-19a	miR-411
	miR-485	let-7c	let-7i	miR-195	miR-331	miR-1976
		miR-92a-2	let-7g	miR-19b-2	miR-181d	miR-16-2
		let-7f-2	miR-146b	miR-374b	miR-32	miR-19b-1
		miR-29a	miR-425	miR-139	miR-363	miR-495
		miR-93	miR-625	miR-146a	miR-340	miR-944
		miR-126	miR-148b	miR-708	miR-660	miR-1301
		miR-205	miR-424	miR-503	miR-26a-1	miR-1271
		miR-29c	miR-152	miR-98	miR-330	miR-224
		miR-3653	miR-374c			
	13			102		

Table 5. *FGF3/4* and *HER2* related microRNA list up

Oncogene	≥2.5~<5	≥5	
FGF3	miR-485	miR-589 miR-324	
	1	2	
FGF4	miR-651	miR-200b miR-361	
	1	2	
HER2	miR-4326	miR-30a	miR-193a
	miR-485	miR-30e	miR-486
		miR-30d	miR-331
		let-7c	miR-3127
		miR-375	miR-660
		miR-125a	miR-3613
		miR-532	
	2	13	
c-Myc	miR-1468	miR-186	
	miR-377		
	miR-34b		
	3	1	

tools are essential for more accurate prediction. We conduct data mining to elucidate common miRNA which regulate multi-oncogene.

Table 6. Oncogenes related microRNAs clustering

miRNA ID \ Oncogene		<i>BRCA1</i>	<i>BRCA2</i>	<i>HER2</i>	<i>N-RAS</i>	<i>Cyclin D1</i>	<i>FGF3</i>	<i>FGF4</i>	<i>c-Myc</i>	Sum
≥2.5~<5	miR-651	○	○				○			3
	miR-1468	○			○				○	
	let-7c	○	○	○	○	○				5
	let-7g	○	○		○	○				4
	miR-224	○	○		○	○				
	miR-186	○			○	○			○	
	miR-589	○			○	○	○			
≥5	miR-30d	○			○					3
	miR-93	○			○	○				
	miR-205	○	○		○					
	miR-106b	○			○	○				
	miR-146b	○	○		○					
	miR-424	○			○	○				
	miR-17	○			○	○				
	miR-20a	○			○	○				
	miR-146a	○	○		○					
	miR-20b	○			○	○				
	miR-769	○	○			○				
	miR-26a-1	○			○	○				
	miR-106a	○			○	○				
	let-7b		○		○	○				
	let-7e		○		○	○				
	let-7i		○		○	○				
	miR-625		○		○	○				
	miR-98		○		○	○				
	miR-19a		○		○	○				
	miR-3613			○	○	○				
	miR-361				○	○		○		

Because bulk data provided by TCGA contains over-expressed miRNA information from clinical breast cancer sample, it can be a new tool to increase the accuracy and specificity of computational methods. Selected 8 oncogenes have more high correlation with breast cancer. So, when we investigated the miRNA predicted data base, gene-miRNA target module is used inversely. Adding the clinical information and process of cross-check with inverse target module between the miRNA data base, data mining would be more reliable. Elucidated common miRNAs means that they are likely to be more effective than miRNA provided from previous study, because common miRNAs control the multi-gene related with breast cancer and have regulation-networking among the miRNAs.

In general, there are two approaches to applying miRNA as a therapeutic or biomarker. Anti-miRNA oligonucleotides (AMOs) are blocking reagents that inhibit the activity of mature miRNA by hybridizing and sense oligonucleotides (mimics) are double-stranded miRNA-like RNA fragment. The miRNA-based therapeutic has already passed Phase 2 clinical trials (29-32). Altered miRNAs can be utilized for diagnosis and prognosis of a variety of diseases, such as cancer, cardiovascular, inflammatory, and neurodevelopmental disease (33, 34). However, there is a little interest about miRNA utilizing for food or animal feed. Various food components including Vitamin A,

Vitamin B, Curcumin, Resveratrol, Ellagitannins, and Isoflavones can affect the miRNA expression (35-39). These findings suggest that food components either indirectly participate in the physiologic process by regulating miRNA expression or directly regulate the physiologic process by itself. It is worthy to check the possibility of combination between the miRNA and food component, because of potential synergic effect. Plant is a main resource of food. Zhang *et al.* have demonstrated that ingested miRNAs through plant were detected in human serum and exogenous plant miR-168a specifically could inhibit mammalian low-density lipoprotein receptor adapter 1 (LDLRAP1) mRNA (40). Indeed, exogenous plant miRNA were stable in cooked food and could endure the mouse G1 tract (40, 41).

Strategy of miRNAs-containing food may be an effective method for disease prevention by combination of miRNA and food components. This high-tech food can be achieved by using the ASO and/or mimic of miRNA that involved in specific disease. Common miRNA with multi-targeting ability will be a critical resource for development of food.

Acknowledgements This study was supported by the National Research Foundation of Korea (KRF) grant funded by the Korea government (Ministry of Science, ICT & Future Planning) (No. 2015R1C1A101053520). Authors thank to Jae-Joong Yun and Soo-Ji Jeong for their initial preliminary works and You-Kyung Son for the kind cooperation of references sorting and searching.

Disclosure The authors declare no conflict of interest.

References

- Graveel CR, Calderone HM, Westerhuis JJ, Winn ME, Sempere LF. Critical analysis of the potential for microRNA biomarkers in breast cancer management. *Breast Cancer* 7: 59-79 (2015)
- André F, Cortés J. Rationale for targeting fibroblast growth factor receptor signaling in breast cancer. *Breast Cancer Res. Tr.* 150: 1-8 (2015)
- Rebecca Siegel, Jiemin Ma, Zhaohui Zou, Ahmedin Jemal. *Cancer statistics, 2014*. CA-Cancer J. Clin. 64: 9-29 (2014)
- Somnay Y, Schneider D, Mazeh H. Thyroid: Medullary carcinoma. *Atlas Genet. Cytogenet. Oncol. Haematol.* 17: 291-296 (2013)
- Zheng ZY, Tian L, Bu W, Fan C, Gao X, Wang H, Liao YH, Li Y, Lewis MT, Edwards D, Zwaka TP, Hilsenbeck SG, Medina D, Perou CM, Creighton CJ, Zhang XH, Chang EC. Wild-type N-Ras, overexpressed in basal-like breast cancer, promotes tumor formation by inducing IL-8 secretion via JAK2 activation. *Cell Rep.* 12: 511-524 (2015)
- Chung IF, Chen CY, Su SC, Li CY, Wu KJ, Wang HW, Cheng WC. DriverDBv2: A database for human cancer driver gene research. *Nucleic Acids Res.* 44: in press (2015)
- Baldin V, Lukas J, Marcote MJ, Pagano M, Draetta G. Cyclin D1 is a nuclear protein required for cell cycle progression in G1. *Gene. Dev.* 7: 812-821 (1993)
- He Y, Liu Z, Qiao C, Xu M, Yu J, Li G. Expression and significance of Wnt signaling components and their target genes in breast carcinoma. *Mol. Med. Rep.* 9: 137-143 (2014)
- Mitri Z, Constantine T, O'Regan R. The HER2 receptor in breast cancer: Pathophysiology, clinical use, and new advances in therapy. *Chemother. Res. Pract.* 2012: 743193
- Lee M, Beggs SM, Gildea D, Bupp S, Lichtenberg J, Trivedi NS, NISC Comparative Sequencing Program, Hu Y, Bodine DM, Crawford NP. Necdin is a breast cancer metastasis suppressor that regulates the transcription of *c-Myc*. *Oncotarget* 6: 31557-31568 (2015)
- Welch PL, Owens KN, King MC. Insights into the functions of BRCA1 and BRCA2. *Trends Genet.* 16: 69-74 (2000)
- Kuhn DE, Martin MM, Feldman DS, Terry AV, Nuovo GJ, Elton TS. Experimental validation of miRNA targets. *Methods* 44: 47-54 (2008)
- Mathe A, Scott RJ, Avery-Kiejda KA. MiRNAs and other epigenetic changes as biomarkers in triple negative breast cancer. *Int. J. Mol. Sci.* 16: 28347-28376 (2015)
- Kloosterman WP, Plasterk RH. The diverse functions of microRNAs in animal development and disease. *Dev. Cell* 11: 441-450 (2006)
- Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, Sweet-Cordero A, Ebert BL, Mak RH, Ferrando AA, Downing JR, Jacks T, Horvitz HR, Golub TR. MicroRNA expression profiles classify human cancers. *Nature* 435: 834-838 (2005)
- Sui X, Wang X, Han W, Li D, Xu Y, Lou F, Zhou J, Gu X, Zhu J, Zhang C, Pan H. MicroRNAs-mediated cell fate in triple negative breast cancers. *Cancer Lett.* 361: 8-12 (2015)
- Garnier D, Magnus N, D'Asti E, Hashemi M, Meehan B, Milsom C, Rak J. Genetic pathways linking hemostasis and cancer. *Thromb. Res.* 129: S22-S29 (2012)
- Motameny S, Wolters S, Nürnberg P, Schumacher B. Next generation sequencing of miRNAs-strategies, resources and methods. *Genes* 1: 70-84 (2010)
- Osanto S, Qin Y, Buermans HP, Berkers J, Lerut E, Goeman JJ, van Poppel H. Genome-wide microRNA expression analysis of clear cell renal cell carcinoma by next generation deep sequencing. *PLoS ONE* 7: e38298 (2012)
- Heneghan HM, Miller N, Kerin MJ. MiRNAs as biomarkers and therapeutic targets in cancer. *Curr. Opin. Pharmacol.* 10: 543-550 (2010)
- Dweep H, Sticht C, Pandey P, Gretz N. miRWalk-database: Prediction of possible miRNA binding sites by "walking" the genes of three genomes. *J. Biomed. Inform.* 44: 839-847 (2011)
- Betel D, Wilson M, Gabow A, Marks DS, Sander C. The microRNA.org resource: Targets and expression. *Nucleic Acids Res.* 36: D149-D153 (2008)
- Wang X. miRDB: A microRNA target prediction and functional annotation database with a wiki interface. *RNA* 14: 1012-1017 (2008)
- Loher P, Rigoutsos I. Interactive exploration of RNA22 microRNA target predictions. *Bioinformatics* 28: 3322-3323 (2012)
- Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. *Elife* 4: e05005 (2015)
- Witkos TM, Koscińska E, Krzyżosiak WJ. Practical aspects of microRNA target prediction. *Curr. Mol. Med.* 11: 93-109 (2011)
- Alvarez-Garcia I, Eric EA. MicroRNA functions in animal development and human disease. *Development* 132: 4653-4662 (2005)
- Kuhn DE, Martin MM, Feldman DS, Terry Jr. AV, Nuovo GJ, Elton TS. Experimental validation of miRNA targets. *Methods* 44: 47-54 (2008)
- Soifer HS, Rossi JJ, Sætrum P. MicroRNAs in disease and potential therapeutic applications. *Mol. Ther.* 15: 2070-2079 (2007)
- Choi SW, Friso S. Epigenetics: A new bridge between nutrition and health. *Adv. Nutr.* 1: 8-16 (2010)
- Landecker H. Food as exposure: Nutritional epigenetics and the new metabolism. *Biosocieties* 6: 167-194 (2011)
- Paul B, Barnes S, Demark-Wahnefried W, Morrow C, Salvador C, Skibola C, Tollefsbol TO. Influences of diet and the gut microbiome on epigenetic modulation in cancer and other diseases. *Clin. Epigenetics* 7: 112 (2015)
- Mishra PJ. MicroRNAs as promising biomarkers in cancer diagnostics. *Biomark. Res.* 2: 19 (2014)
- Ali MA, Mozghan MN. The role of microRNAs in human diseases. *Avicenna J. Med. Biotechnol.* 2: 161-180 (2010)
- Garzon R, Pichiorri F, Palumbo T, Visentini M, Aqeilan R, Cimmino A, Wang H, Sun H, Volinia S, Alder H, Calin GA, Liu CG, Andreeff M, Croce CM. MicroRNA gene expression during retinoic acid-induced differentiation of human acute promyelocytic leukemia. *Oncogene* 26: 4148-4157 (2007)
- Marsit CJ, Eddy K, Kelsey KT. MicroRNA responses to cellular stress. *Cancer Res.* 66: 10843-10848 (2006)
- Mudduluru G, George-William JN, Muppala S, Asangani IA, Kumarswamy R, Nelson LD, Allgayer H. Curcumin regulates miR-21 expression and inhibits invasion and metastasis in colorectal cancer. *Biosci. Rep.* 3: 185-197 (2011)
- Dhar S, Hicks C, Levenson AS. Resveratrol and prostate cancer: Promising role for microRNAs. *Mol. Nutr. Food Res.* 55: 1219-1229 (2011)
- Li Y, Vandenboom TG 2nd, Wang Z, Kong D, Ali S, Philip PA, Sarkar FH. miR-146a suppresses invasion of pancreatic cancer cells. *Cancer Res.* 70: 1486-1495 (2010)
- Zhang L, Hou D, Chen X, Li D, Zhu L, Zhang Y, Li J, Bian Z, Liang X, Cai X, Yin Y, Wang C, Zhang T, Zhu D, Zhang D, Xu J, Chen Q, Ba Y, Liu J, Wang Q, Chen J, Wang J, Wang M, Zhang Q, Zhang J, Zen K, Zhang CY. Exogenous plant MIR168a specifically targets mammalian LDLRAP1: Evidence of cross-kingdom regulation by microRNA. *Cell Res.* 22: 107-126 (2012)
- Philip A, Ferro VA, Tate RJ. Determination of the potential bioavailability of plant microRNAs using a simulated human digestion process. *Mol. Nutr. Food Res.* 59: 1962-1972 (2015)