

## Introduction

MicroRNAs (miRNAs) are a class of endogenous small noncoding RNAs conserved through the evolution. They mediate posttranscriptional regulation of protein-coding genes by binding to the 3' untranslated region (3'UTR) of target mRNAs, leading to translational inhibition, mRNA destabilization or degradation, depending on the degree of sequence complementarity [1]. During the biogenesis of miRNAs, the primary miRNAs (pri-miRNAs) are transcribed from the intra- and inter-genetic regions of the genome by RNA polymerase II, followed by processing by the RNase III enzyme Droscha into pre-miRNAs. After nuclear export, they are cleaved by the RNase III enzyme Dicer into mature miRNAs consisting of approximately 22 nucleotides. Finally, a single-stranded miRNA is loaded onto the RNA-induced silencing complex (RISC), where the seed sequence located at positions 2 to 8 from the 5' end of the miRNA plays a pivotal role in recognition of the target mRNA [2]. At present, more than one thousand of human miRNAs are registered in miRBase Release 16 <http://www.mirbase.org>. The 3'UTR of a single mRNA is often targeted by several different miRNAs, while a single miRNA concurrently reduces the production of hundreds of target proteins [3]. Consequently, the whole miRNA system (microRNAome) regulate greater than 60% of all protein-coding genes in a human cell [4]. By targeting multiple transcripts and affecting expression of numerous proteins, miRNAs play a key role in fine-tuning of diverse cellular functions, such as development, differentiation, proliferation, apoptosis and metabolism. Therefore, aberrant regulation of miRNA expression is deeply involved in pathological events that mediate cancers [5] and neurodegenerative disorders [6].

Recent advances in systems biology have made major breakthroughs by illustrating the cell-wide map of complex molecular interactions with the aid of the literature-based knowledgebase of molecular pathways [7]. The logically arranged molecular networks construct the whole system characterized by robustness, which maintains the proper function of the system in the face of genetic and environmental perturbations [8]. In the scale-free molecular network, targeted disruption of limited numbers of critical components designated hubs, on which the biologically important molecular interactions concentrate, efficiently disturbs the whole cellular function by destabilizing the network [9]. Therefore, the identification of the hub in the molecular network constructed by target genes of a particular miRNA helps us to understand biological and pathological roles of individual miRNAs. Recently, Hsu et al. studied the human microRNA-regulated protein-protein interaction (PPI) network by utilizing the Human Protein Reference Database (HPRD) and the miRNA target prediction program TargetScan [10]. They found that an individual miRNA often targets the hub gene of the PPI network, although they did not attempt to characterize relevant pathways, diseases, and pathological events regulated by miRNA target genes.

At present, the question remains to be fully elucidated whether a set of miRNA target genes regulated by an individual miRNA in the whole human microRNAome generally constitute the biological network of functionally-associated molecules or simply reflect a random set of functionally-independent genes. To address this question, we attempted to characterize molecular networks of target genes of all human miRNAs by using KeyMolnet, a bioinformatics tool for analyzing molecular interactions on the comprehensive knowledgebase.

## Materials and methods

### MicroRNA Target Prediction

The complete list of 1,223 human miRNAs was downloaded from miRBase Release 16 <http://www.mirbase.org>. We searched the target genes of individual miRNA on the Diana-microT 3.0 target prediction program ([diana.cslab.ece.ntua.gr/microT](http://diana.cslab.ece.ntua.gr/microT)), which was selected because of the highest ratio of correctly predicted targets over other prediction tools [11]. Diana-microT 3.0 calculates the miRNA-targeted gene (miTG) score that reflects the weighted sum of the scores of all conserved and non-conserved miRNA recognition elements (MRE) on the 3'UTR of the target mRNA. The miTG score correlates well with fold changes in suppression of protein expression [11]. To optimize the parameter of miRNA-target interaction, we considered the target genes with a cutoff of the miTG score equal to or larger than 20 as the highly reliable targets, because we found that the targets with the miTG score < 20 exhibited the significantly lower precision score, an indicator of correctness in predicted interactions [11], compared with those having the score  $\geq 20$  ( $p = 2.78E-08$  by Mann-Whitney's U-test).

### Molecular Network Analysis

Ensembl Gene IDs of target genes retrieved by Diana-microT 3.0 were converted into the corresponding Entrez Gene IDs by using the DAVID Bioinformatics Resources 6.7 program <http://david.abcc.ncifcrf.gov>[12], where non-annotated IDs were deleted. Then, Entrez Gene IDs of miRNA target genes were uploaded onto KeyMolnet.

KeyMolnet is a tool for analyzing molecular interactions on the literature-based knowledgebase that contains the contents on 123,000 molecular relationships among human genes and proteins, small molecules, diseases, pathways and drugs, established by the Institute of Medicinal Molecular Design (IMMD) (Tokyo, Japan) [13-15]. The core contents are collected from selected review articles and textbooks with the highest reliability, regularly updated and carefully curated by a team of expert biologists. KeyMolnet contains a panel of human canonical networks constructed by core contents in the KeyMolnet library. They represent the gold standard of the networks, composed of 430 pathways, 885 diseases, and 208 pathological events. Detailed information on all the contents is available from IMMD <http://www.immd.co.jp/en/keymolnet/index.html> upon request.

We utilized the neighboring network-search algorithm that selects the set of miRNA target genes as starting points to generate the network around starting points within one path, composed of all kinds of molecular interactions, including direct activation/inactivation, transcriptional activation/repression, and the complex formation. By uploading the list of Entrez Gene IDs onto KeyMolnet, it automatically provides corresponding molecules and a minimum set of intervening molecules as a node on networks. The generated network was compared side by side with human canonical networks described above. The algorithm that counts the number of overlapping molecules and/or molecular relations between the extracted network and the canonical network identifies the canonical network showing the most statistically significant contribution to the extracted network. This algorithm is essentially based on that of the GO::TermFinder [16]. The significance in the similarity between the extracted network and the canonical network is scored following the formula, where  $O$  = the number of overlapping molecules and molecular relations for the pathway or overlapping molecules alone for the disease and the pathological event between the extracted network and the canonical network,  $V$  = the number of molecules

and/or molecular relations located in the extracted network,  $C$  = the number of molecules and/or molecular relations located in the canonical network,  $T$  = the number of total molecules and/or molecular relations of KeyMolnet, currently composed of approximately 15,700 molecules and 123,000 molecular relations, and the  $x$  = the sigma variable that defines coincidence.

$$\text{Score} = -\log_2(\text{Score}(p)) \quad \text{Score}(p) = \sum_{x=0}^{\text{Min}(C,V)} f(x) \quad f(x) = \frac{C C_x \cdot T - C C_{V-x}}{T C_V} \quad (1)$$

## Results

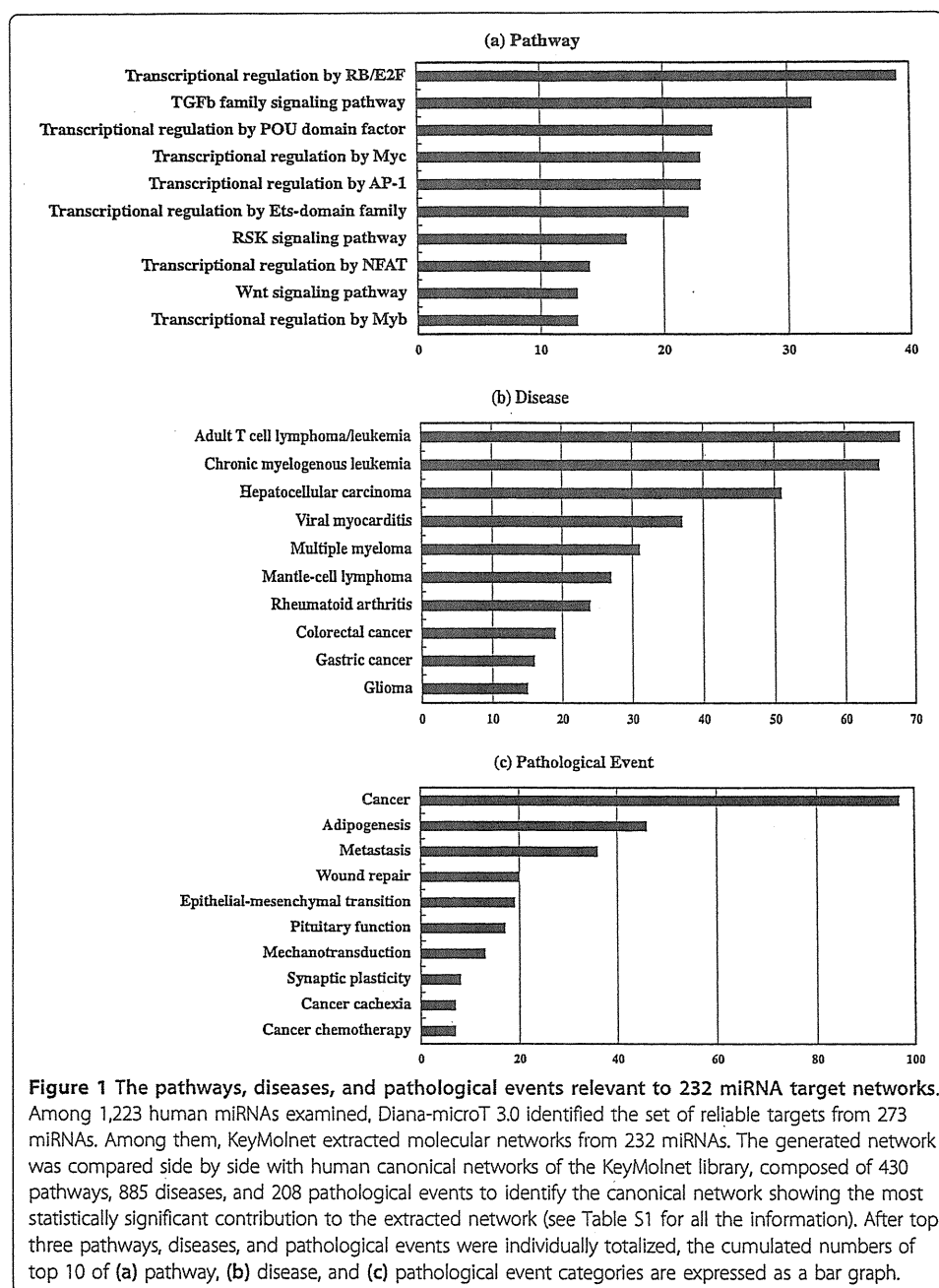
### Molecular Network of MicroRNA Target Genes

Among 1,223 human miRNAs examined, Diana-microT 3.0 predicted the targets from 532 miRNAs (43.5%). Among the 532 miRNAs, 273 miRNAs contained a set of highly reliable targets showing the miTG score  $\geq 20$ . Among 273 miRNAs having reliable targets, KeyMolnet successfully extracted molecular networks from 232 miRNAs. They are comprised of 19% of total human miRNAs (microRNAome). Then, the generated network was compared side by side with human canonical networks of the KeyMolnet library, composed of 430 pathways, 885 diseases, and 208 pathological events. We found that not all 232 miRNAs contained entire categories of canonical networks because several miRNAs comprised relatively small numbers of targets. See Additional file 1 for all the information on 232 miRNAs and their target networks. When top three pathways, diseases, and pathological events were individually totalized, the most relevant pathway is 'transcriptional regulation by RB/E2F' ( $n = 39$ ; 6.8% of total), followed by 'TGF-beta family signaling pathway' ( $n = 32$ ; 5.6%) and 'transcriptional regulation by POU domain factor' ( $n = 24$ ; 4.2%), the most relevant disease is 'adult T cell lymphoma/leukemia' ( $n = 68$ ; 12.1%), followed by 'chronic myelogenous leukemia' ( $n = 65$ ; 11.5%) and 'hepatocellular carcinoma' ( $n = 51$ ; 9.1%), and the most relevant pathological event is 'cancer' ( $n = 97$ ; 24.7%), followed by 'adipogenesis' ( $n = 46$ ; 11.7%) and 'metastasis' ( $n = 36$ ; 9.2%) (Figure 1 and Additional file 1).

Next, we identified the large-scale miRNA target networks by uploading targets greater than 100 per individual miRNA onto KeyMolnet (Table 1). Fifty-two miRNAs that construct such a large-scale miRNA target network include let-7, miR-9, 17, 19, 20, 26, 27, 29, 30, 32, 92, 93, 96, 98, 101, 106b, 124, 137, 147, 153, 218, 372, 429, 495, 506, 519, 520, 603, and their closely-related family members. The targets of these miRNAs established highly complex molecular networks, in which the pathways of 'transcriptional regulation by RB/E2F', 'transcriptional regulation by Ets-domain family', and 'transcriptional regulation by p53', the diseases of 'chronic myelogenous leukemia' and 'viral myocarditis', and the pathological event of 'cancer' were notably accumulated (Table 1). Importantly, distinct members belonging to the same miRNA family, for example, five miR-30 family members ranging from miR-30a to miR-30e constructed a virtually identical molecular network (Table 1).

### Biological Implications of MicroRNA Target Networks

As described above, the present observations indicated that a set of miRNA target genes regulated by an individual miRNA generally constitute the biological network of



functionally-associated molecules in human cells. Therefore, it is highly important to obtain deeper insights into biological implications of miRNA target networks.

The protooncogene *c-myc* is a key transcription factor for normal development of hematopoietic cells. A recent study showed that miR-15a targets *c-myc*, while *c-myc* binds to the promoter of miR-15a, providing an autoregulatory feedback loop in human hematopoietic cells [17]. Consistent with this study, we found ‘transcriptional regulation by myb’ as the most relevant pathway to the miR-15a target network (the score = 602; the score p-value = 7.39E-182) (Figure 2 and Additional file 1). These observations propose a scenario that miR-15a synchronously downregulates both *c-myc* itself and downstream genes transcriptionally regulated by *c-myc*, resulting in

**Table 1 The large-scale human microRNA target networks**

MicroRNA	Number of Targets	Molecules in KeyMolnet Networks	Top Pathway	Score	p-Value	Top Disease	Score	p-Value	Top Pathological Event	Score	p-Value
hsa-let-7a	244	1022	Transcriptional regulation by p53	593	2.69E-179	Viral myocarditis	113	1.21E-34	Cancer	206	1.31E-62
hsa-let-7b	242	1016	Transcriptional regulation by p53	594	1.83E-179	Viral myocarditis	113	9.32E-35	Cancer	206	7.66E-63
hsa-let-7c	243	1020	Transcriptional regulation by p53	593	2.49E-179	Viral myocarditis	113	1.11E-34	Cancer	206	1.10E-62
hsa-let-7d	145	885	Transcriptional regulation by RB/E2F	836	2.18E-252	Chronic myelogenous leukemia	72	1.95E-22	Cancer	130	9.68E-40
hsa-let-7e	236	1111	Transcriptional regulation by p53	575	8.90E-174	Viral myocarditis	116	1.20E-35	Cancer	175	1.86E-53
hsa-let-7f	244	1022	Transcriptional regulation by p53	593	2.69E-179	Viral myocarditis	113	1.21E-34	Cancer	206	1.31E-62
hsa-let-7g	245	1022	Transcriptional regulation by p53	593	2.69E-179	Viral myocarditis	113	1.21E-34	Cancer	206	1.31E-62
hsa-let-7i	245	1022	Transcriptional regulation by p53	593	2.69E-179	Viral myocarditis	113	1.21E-34	Cancer	206	1.31E-62
hsa-miR-9	352	1115	Transcriptional regulation by PPARα	340	5.28E-103	Hepatocellular carcinoma	72	1.69E-22	Cancer	171	3.50E-52
hsa-miR-17	195	961	Transcriptional regulation by RB/E2F	971	3.27E-293	Chronic myelogenous leukemia	92	2.83E-28	Cancer	181	3.58E-55
hsa-miR-19a	226	1094	Transcriptional regulation by RB/E2F	760	2.10E-229	Chronic myelogenous leukemia	113	1.26E-34	Cancer	253	7.04E-77
hsa-miR-19b	225	1094	Transcriptional regulation by RB/E2F	760	2.10E-229	Chronic myelogenous leukemia	113	1.26E-34	Cancer	253	7.04E-77
hsa-miR-20a	165	1038	Transcriptional regulation by RB/E2F	856	1.64E-258	Chronic myelogenous leukemia	87	6.09E-27	Cancer	85	3.33E-26
hsa-miR-20b	198	981	Transcriptional regulation by RB/E2F	962	2.35E-290	Chronic myelogenous leukemia	98	3.39E-30	Cancer	183	6.98E-56
hsa-miR-26a	148	672	Transcriptional regulation by RB/E2F	919	1.76E-277	Chronic myelogenous leukemia	107	6.15E-33	Cancer	181	3.20E-55
hsa-miR-26b	148	672	Transcriptional regulation by RB/E2F	919	1.76E-277	Chronic myelogenous leukemia	107	6.15E-33	Cancer	181	3.20E-55
hsa-miR-27a	229	1192	Transcriptional regulation by CREB	1022	2.23E-308	Chronic myelogenous leukemia	95	1.96E-29	Cancer	194	3.05E-59
hsa-miR-27b	261	1337	Transcriptional regulation by CREB	1022	2.23E-308	Chronic myelogenous leukemia	94	4.51E-29	Cancer	211	4.11E-64
hsa-miR-29a	119	543	Transcriptional regulation by Ets-domain family	430	4.36E-130	Glioma	85	3.46E-26	Cancer	139	1.41E-42

**Table 1 The large-scale human microRNA target networks (Continued)**

hsa-miR-29b	118	578	Transcriptional regulation by Ets-domain family	422	1.15E-127	Glioma	82	1.55E-25	Cancer	146	1.44E-44
hsa-miR-29c	118	543	Transcriptional regulation by Ets-domain family	430	4.36E-130	Glioma	85	3.46E-26	Cancer	139	1.41E-42
hsa-miR-30a	455	1494	Transcriptional regulation by RB/E2F	777	9.43E-235	Chronic myelogenous leukemia	86	1.11E-26	Cancer	195	2.39E-59
hsa-miR-30b	455	1480	Transcriptional regulation by RB/E2F	781	1.08E-235	Chronic myelogenous leukemia	87	7.01E-27	Cancer	188	1.92E-57
hsa-miR-30c	454	1495	Transcriptional regulation by RB/E2F	778	6.13E-235	Chronic myelogenous leukemia	86	1.15E-26	Cancer	191	3.63E-58
hsa-miR-30d	452	1491	Transcriptional regulation by RB/E2F	778	7.28E-235	Chronic myelogenous leukemia	86	1.01E-26	Cancer	195	1.96E-59
hsa-miR-30e	455	1481	Transcriptional regulation by RB/E2F	780	1.29E-235	Chronic myelogenous leukemia	87	7.25E-27	Cancer	188	2.05E-57
hsa-miR-32	261	905	Transcriptional regulation by RB/E2F	842	2.74E-254	Gastric cancer	80	8.85E-25	Cancer	157	4.19E-48
hsa-miR-92a	219	642	Transcriptional regulation by MEF2	335	1.51E-101	Viral myocarditis	59	1.62E-18	Epithelial-mesenchymal transition	83	7.76E-26
hsa-miR-92b	258	701	Transcriptional regulation by MEF2	328	1.59E-99	Viral myocarditis	60	1.23E-18	Cancer	94	3.97E-29
hsa-miR-93	195	958	Transcriptional regulation by RB/E2F	972	2.37E-293	Chronic myelogenous leukemia	92	2.47E-28	Cancer	181	2.77E-55
hsa-miR-96	142	688	Transcriptional regulation by Ets-domain family	407	3.42E-123	Viral myocarditis	36	1.06E-11	Cancer	106	1.37E-32
hsa-miR-98	162	671	Transcriptional regulation by Myb	549	4.73E-166	Viral myocarditis	85	2.66E-26	Cancer	126	1.42E-38
hsa-miR-101	188	806	Transcriptional regulation by AP-1	492	1.10E-148	Hepatocellular carcinoma	70	6.40E-22	Cancer	127	4.26E-39
hsa-miR-106b	164	1028	Transcriptional regulation by RB/E2F	854	7.21E-258	Chronic myelogenous leukemia	87	5.48E-27	Cancer	85	2.93E-26
hsa-miR-124	285	1346	Transcriptional regulation by RB/E2F	756	3.57E-228	Chronic myelogenous leukemia	83	9.34E-26	Cancer	185	1.90E-56
hsa-miR-137	288	941	Transcriptional regulation by MTF family	339	1.19E-102	Adult T cell lymphoma/leukemia	66	1.30E-20	Cancer	179	1.00E-54
hsa-miR-147	199	867	Transcriptional regulation by RB/E2F	805	4.06E-243	Chronic myelogenous leukemia	113	6.60E-35	Cancer	132	2.57E-40
hsa-miR-153	154	1019	Transcriptional regulation by Myb	507	2.35E-153	Multiple myeloma	60	6.44E-19	Cancer	174	4.31E-53

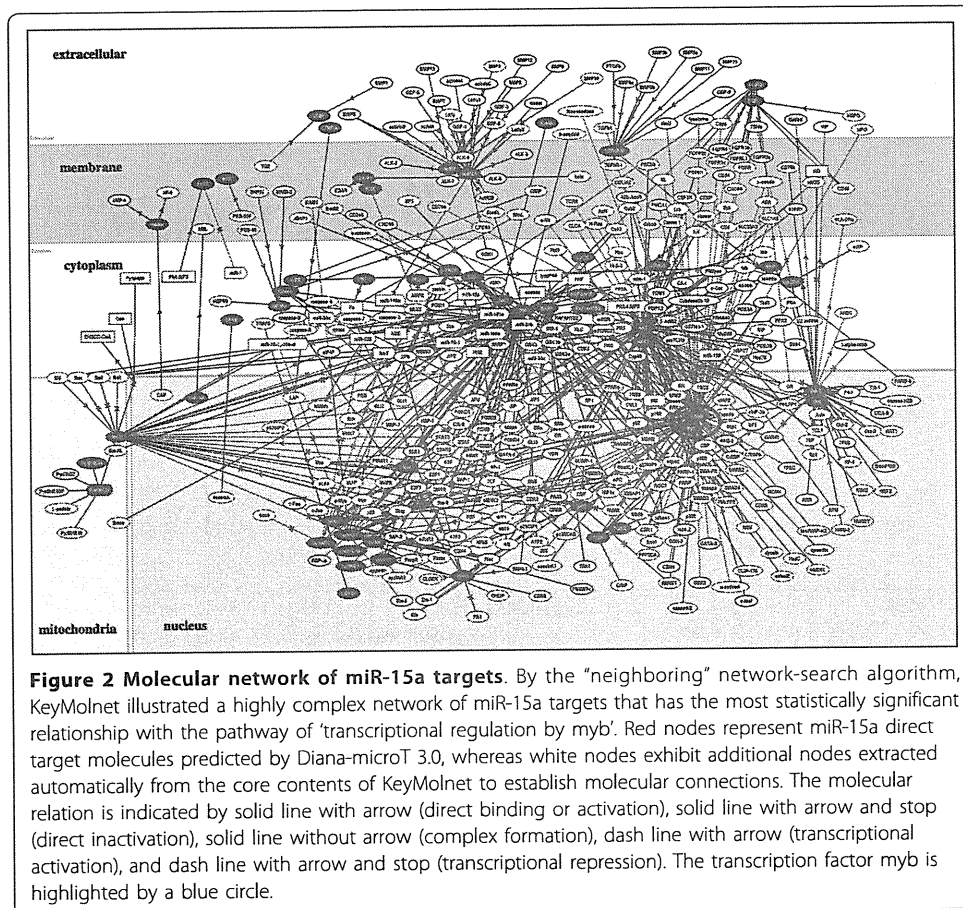
**Table 1 The large-scale human microRNA target networks (Continued)**

hsa-miR-218	155	830	Transcriptional regulation by AP-1	344	2.28E-104	Hepatocellular carcinoma	69	1.63E-21	Cancer	136	1.52E-41
hsa-miR-372	101	562	Transcriptional regulation by RB/E2F	1022	2.23E-308	Chronic myelogenous leukemia	85	1.90E-26	Cancer	144	2.75E-44
hsa-miR-429	123	634	Transcriptional regulation by RB/E2F	918	2.45E-277	Chronic myelogenous leukemia	76	1.71E-23	Cancer	130	5.28E-40
hsa-miR-495	156	601	Transcriptional regulation by Ets-domain family	431	2.14E-130	Rheumatoid arthritis	77	5.90E-24	Adipogenesis	79	1.32E-24
hsa-miR-506	394	1536	Transcriptional regulation by Ets-domain family	317	4.69E-96	Viral myocarditis	99	1.73E-30	Cancer	172	1.43E-52
hsa-miR-519a	281	1256	Transcriptional regulation by RB/E2F	811	5.32E-245	Chronic myelogenous leukemia	106	1.34E-32	Cancer	220	8.03E-67
hsa-miR-519b-3p	281	1256	Transcriptional regulation by RB/E2F	811	5.32E-245	Chronic myelogenous leukemia	106	1.34E-32	Cancer	220	8.03E-67
hsa-miR-519c-3p	281	1256	Transcriptional regulation by RB/E2F	811	5.32E-245	Chronic myelogenous leukemia	106	1.34E-32	Cancer	220	8.03E-67
hsa-miR-520a-3p	184	690	Transcriptional regulation by RB/E2F	1022	2.23E-308	Chronic myelogenous leukemia	94	6.95E-29	Cancer	146	1.12E-44
hsa-miR-520b	182	690	Transcriptional regulation by RB/E2F	1022	2.23E-308	Chronic myelogenous leukemia	94	6.95E-29	Cancer	146	1.12E-44
hsa-miR-520c-3p	182	690	Transcriptional regulation by RB/E2F	1022	2.23E-308	Chronic myelogenous leukemia	93	9.28E-29	Cancer	145	1.77E-44
hsa-miR-520d-3p	183	690	Transcriptional regulation by RB/E2F	1022	2.23E-308	Chronic myelogenous leukemia	94	6.95E-29	Cancer	146	1.12E-44
hsa-miR-520e	184	690	Transcriptional regulation by RB/E2F	1022	2.23E-308	Chronic myelogenous leukemia	94	6.95E-29	Cancer	146	1.12E-44
hsa-miR-603	252	1150	Transcriptional regulation by Ets-domain family	344	3.26E-104	Multiple myeloma	84	4.36E-26	Cancer	161	4.24E-49

Among 1,223 human miRNAs examined, Diana-microT 3.0 predicted reliable targets from 273 miRNAs. Among them, KeyMolnet extracted molecular networks from 232 miRNAs. The generated network was compared side by side with human canonical networks of the KeyMolnet library, composed of 430 pathways, 885 diseases, and 208 pathological events. The canonical pathways, diseases, and pathological events with the most statistically significant contribution to the extracted network are shown. The table contains only the large-scale miRNA target networks generated by importing targets greater than 100 per individual miRNA into KeyMolnet. See Additional file 1 for all the information on 232 miRNAs and their target networks.

efficient inactivation of the whole molecular network governed by the hub gene *c-myb*. These results suggest a collaborative regulation of gene expression at both transcriptional and posttranscriptional levels that involve coordinated regulation by miRNAs and transcription factors.

The retinoblastoma protein Rb/E2F pathway acts as a gatekeeper for G1/S transition in the cell cycle. The Rb/E2F-regulated G1 checkpoint control is often disrupted in cancer cells. A recent study showed that miR-106b is directly involved in posttranscriptional regulation of E2F1 [18]. E2F1 activates transcription of miR-106b, while miR-

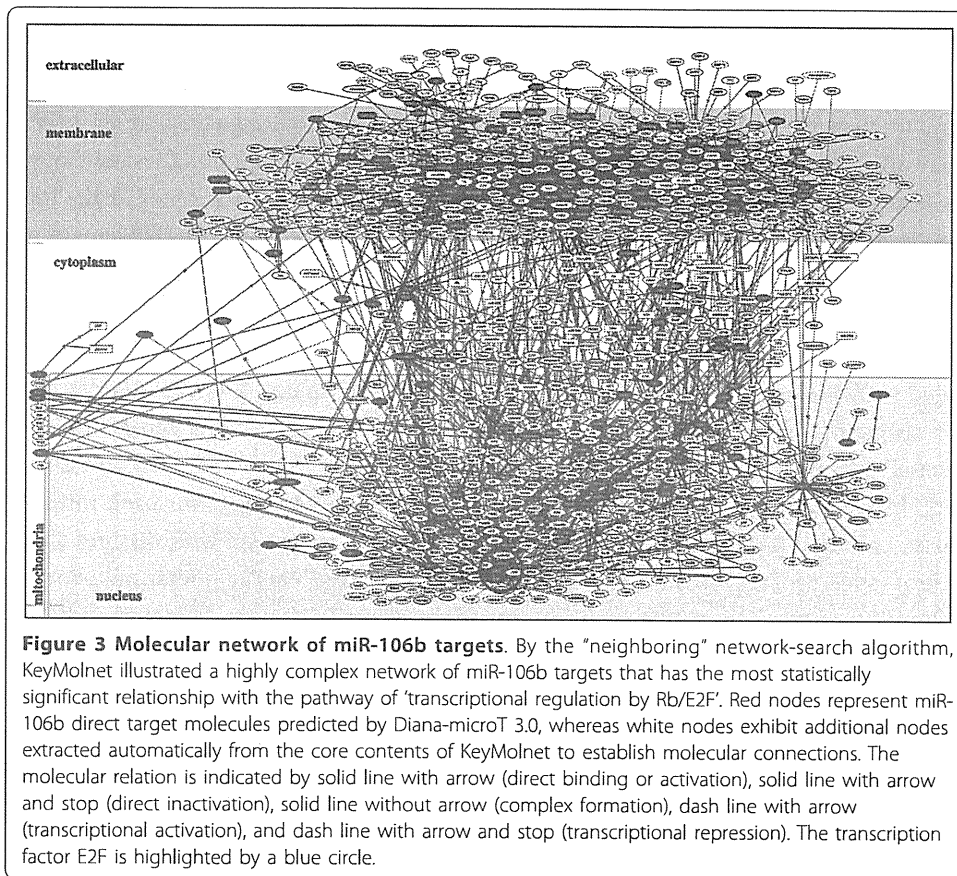


106b targets E2F1, serving as a miRNA-directed negative feedback loop in gastric cancer cells [18]. Supporting these findings, we identified 'transcriptional regulation by Rb/E2F' as the most relevant pathway to the miR-106b target network (the score = 854; the score p-value = 7.21E-258) (Figure 3, Table 1 and Additional file 1). The relationship between miR-106b and Rb/E2F would provide another example of coordinated regulation of gene expression by miRNAs and transcription factors.

We found 'transcriptional regulation by p53' as the most relevant pathway to the target network of all let-7 family members except for let-7d (Table 1). It is worthy to note that the tumor suppressor p53 regulates the expression of components of the miRNA-processing machinery, such as Drosha, DGCR8, Dicer, and TARBP2, all of which have p53-responsive elements in their promoters [19]. Furthermore, Dicer and TARBP2, along with p53, serve as a target of the let-7 family miRNAs, suggesting a close link between p53 and let-7 in miRNA biogenesis [19]. The expression of let-7 family members was greatly reduced in certain cancer cells [20].

The microphthalmia associated transcription factor (MITF), a basic helix-loop-helix zipper (bHLH-Zip) transcription factor, acts as not only a master regulator of melanocyte differentiation but also an oncogene promoting survival of melanoma. Recent studies indicate that MITF is a direct target of both miR-137 and miR-148b [21,22]. Again, we identified 'transcriptional regulation by MITF family' as the most relevant pathway to both miR-137 (the score = 339; the score p-value = 1.19E-102) and miR-





148b (the score = 40; the score p-value = 3.91E-142) target networks (Table 1 and Additional file 1).

Cellular responsiveness to glucocorticoids (GCs) is regulated by the delicate balance of the glucocorticoid receptor (GR) protein, GR coactivators and corepressors, GR splice variants and isoforms, and regulators of GR retrograde transport to the nucleus. A recent study showed that miR-18a targets the GR protein, and thereby inhibits GR-mediated biological events in neuronal cells [23]. Consistent with this, we found 'transcriptional regulation by GR' as the most relevant pathway to the miR-18a target network (the score = 1022; the score p-value = 2.23E-308) (Additional file 1).

Zinc finger transcription factors ZEB1 and ZEB2 act as a transcriptional repressor of E-cadherin. A recent study showed that the expression of miR-200b, which targets both ZEB1 and ZEB2, was downregulated in the cells that undergo TGF-beta-induced epithelial to mesenchymal transition (EMT), and was lost in invasive breast cancer cells [24]. We identified 'transcriptional regulation by ZEB' as the third-rank significant pathway (the score = 155; the score p-value = 1.88E-47) and 'EMT' as the third-rank significant pathological event relevant to the miR-200b target network (the score = 61; the score p-value = 4.15E-19) (Additional file 1).

## Discussion

In general, a single miRNA concurrently downregulates hundreds of target mRNAs by binding to the corresponding 3'UTR of mRNA via either perfect or imperfect sequence complementarity [3]. Such fuzzy mRNA-miRNA interactions result in the redundancy

of miRNA-recognized targets. By targeting multiple transcripts and affecting expression of numerous proteins at one time, miRNAs regulate a wide range of cellular functions, such as development, differentiation, proliferation, apoptosis and metabolism. Therefore, we have the question whether a set of miRNA target genes regulated by an individual miRNA generally constitute the biological network of functionally-associated molecules or simply reflect a random set of functionally-independent genes. If the former is the case, what kind of biological networks does the human microRNAome most actively regulates?

To address these questions, first we identified the set of credible target genes for all individual human miRNAs by using the Diana-microT 3.0 program. Then, we investigated miRNA target networks by applying them to KeyMolnet, a bioinformatics tool for analyzing molecular interactions on the comprehensive knowledgebase. Diana-microT 3.0 identified highly reliable targets from 273 miRNAs out of 1,223 all human miRNAs. Previous studies showed that the list of predicted targets for each miRNA varies among different miRNA target prediction programs armed with distinct algorithms, such as TargetScan 5.1 <http://www.targetscan.org>, PicTar ([pictar.mdc-berlin.de](http://pictar.mdc-berlin.de)), miRanda <http://www.microrna.org> and Diana-microT 3.0 [25]. Therefore, miRNA target networks are to some extent flexible, depending on the target prediction program employed. Among the programs described above, we have chosen Diana-microT 3.0 because of the highest ratio of correctly predicted targets over other prediction tools and the simplicity of setting a cut-off point for detection of reliable miRNA-target interactions based on the miTG score [11].

Here we found that highly reliable targets of substantial numbers of human miRNAs actually constructed biologically meaningful molecular networks. These observations strongly supported the theoretical view that miRNA target genes regulated by an individual miRNA in the whole human microRNAome generally constitute the biological network of functionally-associated molecules. A recent study showed that interacting proteins in the human PPI network tend to share restricted miRNA target-site types than random pairs, being consistent with our observations [26].

We also found that there exists a coordinated regulation of gene expression at the transcriptional level by transcription factors and at the posttranscriptional level by miRNAs in miRNA target networks. Recently, Cui et al. investigated the relationship between miRNA and transcription factors in gene regulation [27]. Importantly, they found that the genes with more transcription factor-binding sites have a higher probability of being targeted by miRNAs and have more miRNA-binding sites.

A recent study by miRNA expression profiling of thousands of human tissue samples revealed that diverse miRNAs constitute a complex network composed of coordinately regulated miRNA subnetworks in both normal and cancer tissues, and they are often disorganized in solid tumors and leukemias [28]. During carcinogenesis, various miRNAs play a central role, acting as either oncogenes named oncomir or tumor suppressors termed anti-oncomir, by targeting key molecules involved in apoptosis, cell cycle, cell adhesion and migration, chromosome stability, and DNA repair [5]. Many miRNA gene loci are clustered in cancer-associated genomic regions [29]. Furthermore, miRNA expression signatures well discriminate different types of cancers with distinct clinical prognoses [30]. In the present study, KeyMolnet analysis of miRNA target networks showed that the most relevant pathological event is 'cancer', when top three

pathological events were overall cumulated. Furthermore, the highly relevant diseases include 'adult T cell lymphoma/leukemia', 'chronic myelogenous leukemia', and 'hepatocellular carcinoma'. These observations suggest that the human microRNAome plays a more specialized role in regulation of oncogenesis. Therefore, the miRNA-based therapy directed to targeting multiple cancer-associated pathways simultaneously might serve as the most effective approach to suppressing the oncogenic potential of a wide range of cancers.

## Conclusion

The reliable targets predicted by Diana microT 3.0 derived from approximately 20% of all human miRNAs constructed biologically meaningful molecular networks by KeyMolnet. These observations support the view that miRNA target genes regulated by an individual miRNA in the whole human microRNAome generally constitute the biological network of functionally-associated molecules. In the human miRNA target networks, the most relevant pathway is transcriptional regulation by transcription factors RB/E2F, the disease is adult T cell lymphoma/leukemia, and the pathological event is cancer. In miRNA target networks, there exists a coordinated regulation of gene expression at the transcriptional level by transcription factors and at the posttranscriptional level by miRNAs.

## Additional material

**Additional file 1: KeyMolnet identifies microRNA target networks in 232 human miRNAs.** The prediction of target genes of individual miRNA was performed by Diana-microT 3.0. Entrez Gene IDs of miRNA target genes were uploaded onto KeyMolnet. The generated network was compared side by side with human canonical networks composed of 430 pathways, 885 diseases, and 208 pathological events of the KeyMolnet library. Top-three pathways, diseases, and pathological events with the statistically significant contribution to the extracted network are shown.

## Acknowledgements

This work was supported by grants from the Research on Intractable Diseases (H22-Nanchi-Ippan-136), the Ministry of Health, Labour and Welfare (MHLW), Japan and the High-Tech Research Center Project (S0801043) and the Grant-in-Aid (C22500322), the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan.

## Authors' contributions

JS designed the methods, analyzed the data, and drafted the manuscript. HT helped the data analysis. All authors have read and approved the final manuscript.

## Competing interests

The authors declare that they have no competing interests.

Received: 13 October 2010 Accepted: 17 June 2011 Published: 17 June 2011

## References

1. Guo H, Ingolia NT, Weissman JS, Bartel DP: Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* 2010, **466**:835-40.
2. Bartel DP: MicroRNAs: target recognition and regulatory functions. *Cell* 2009, **136**:215-33.
3. Selbach M, Schwanhäusser B, Thierfelder N, Fang Z, Khanin R, Rajewsky N: Widespread changes in protein synthesis induced by microRNAs. *Nature* 2008, **455**:58-63.
4. Friedman RC, Farh KK, Burge CB, Bartel DP: Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* 2009, **19**:92-105.
5. Garzon R, Marcucci G, Croce CM: Targeting microRNAs in cancer: rationale, strategies and challenges. *Nat Rev Drug Discov* 2010, **9**:775-89.
6. Shiota M, Obayashi S, Tabunoki H, Arima K, Saito Y, Ishida T, Satoh J: Aberrant microRNA expression in the brains of neurodegenerative diseases: miR-29a decreased in Alzheimer disease brains targets neurone navigator 3. *Neuropathol Appl Neurobiol* 2010, **36**:320-30.
7. Viswanathan GA, Seto J, Patil S, Nudelman G, Sealfon SC: Getting started in biological pathway construction and analysis. *PLoS Comput Biol* 2008, **4**:e16.

8. Kitano H: A robustness-based approach to systems-oriented drug design. *Nat Rev Drug Discov* 2007, **6**:202-10.
9. Albert R, Jeong H, Barabasi AL: Error and attack tolerance of complex networks. *Nature* 2000, **406**:378-82.
10. Hsu CW, Juan HF, Huang HC: Characterization of microRNA-regulated protein-protein interaction network. *Proteomics* 2008, **8**:1975-79.
11. Maragkakis M, Alexiou P, Papadopoulos GL, Reczko M, Dalamagas T, Giannopoulos G, Goumas G, Koukis E, Kourtis K, Simossis VA, Sethupathy P, Vergoulis T, Koziris N, Sellis T, Tsanakas P, Hatzigeorgiou AG: Accurate microRNA target prediction correlates with protein repression levels. *BMC Bioinformatics* 2009, **10**:295.
12. Huang da W, Sherman BT, Lempicki RA: Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009, **4**:44-57.
13. Sato H, Ishida S, Toda K, Matsuda R, Hayashi Y, Shigetaka M, Fukuda M, Wakamatsu Y, Itai A: New approaches to mechanism analysis for drug discovery using DNA microarray data combined with KeyMolnet. *Curr Drug Discov Technol* 2005, **2**:89-98.
14. Satoh JI, Tabunoki H, Yamamura T: Molecular network of the comprehensive multiple sclerosis brain-lesion proteome. *Mult Scler* 2009, **15**:531-41.
15. Satoh J, Tabunoki H, Arima K: Molecular network analysis suggests aberrant CREB-mediated gene regulation in the Alzheimer disease hippocampus. *Dis Markers* 2009, **27**:239-52.
16. Boyle EI, Weng S, Gollub J, Jin H, Botstein D, Cherry JM, Sherlock G: GO:TermFinder—open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics* 2004, **20**:3710-15.
17. Zhao H, Kalota A, Jin S, Gewirtz AM: The c-myc proto-oncogene and microRNA-15a comprise an active autoregulatory feedback loop in human hematopoietic cells. *Blood* 2009, **113**:505-16.
18. Petrocca F, Visone R, Onelli MR, Shah MH, Nicoloso MS, de Martino I, Iliopoulos D, Pilozzi E, Liu CG, Negrini M, Cavazzini L, Volinia S, Alder H, Ruco LP, Baldassarre G, Croce CM, Vecchione A: E2F1-regulated microRNAs impair TGF $\beta$ -dependent cell-cycle arrest and apoptosis in gastric cancer. *Cancer Cell* 2008, **13**:272-86.
19. Boominathan L: The tumor suppressors p53, p63, and p73 are regulators of microRNA processing complex. *PLoS One* 2010, **5**:e10615.
20. Takamizawa J, Konishi H, Yanagisawa K, Tomida S, Osada H, Endoh H, Harano T, Yatabe Y, Nagino M, Nimura Y, Mitsudomi T, Takahashi T: Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. *Cancer Res* 2004, **64**:3753-56.
21. Bemis LT, Chen R, Amato CM, Classen EH, Robinson SE, Coffey DG, Erickson PF, Shellman YG, Robinson WA: MicroRNA-137 targets microphthalmia-associated transcription factor in melanoma cell lines. *Cancer Res* 2008, **68**:1362-68.
22. Hafliðadóttir BS, Bergsteinsdóttir K, Praetorius C, Steingrímsson E: miR-148 regulates Mitf in melanoma cells. *PLoS One* 2010, **5**:e11574.
23. Vreugdenhil E, Verissimo CS, Mariman R, Kamphorst JT, Barbosa JS, Zweers T, Champagne DL, Schouten T, Meijer OC, de Kloet ER, Fitzsimons CP: MicroRNA 18 and 124a down-regulate the glucocorticoid receptor: implications for glucocorticoid responsiveness in the brain. *Endocrinology* 2009, **150**:2220-28.
24. Gregory PA, Bert AG, Paterson EL, Barry SC, Tsykin A, Farshid G, Vadas MA, Khew-Goodall Y, Goodall GJ: The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat Cell Biol* 2008, **10**:593-601.
25. Boross G, Orosz K, Farkas I: Human microRNAs co-silence in well-separated groups and have different predicted essentialities. *Bioinformatics* 2009, **25**:1063-69.
26. Liang H, Li WH: MicroRNA regulation of human protein protein interaction network. *RNA* 2007, **13**:1402-8.
27. Cui Q, Yu Z, Pan Y, Purisima EO, Wang E: MicroRNAs preferentially target the genes with high transcriptional regulation complexity. *Biochem Biophys Res Commun* 2007, **352**:733-38.
28. Volinia S, Galasso M, Costinean S, Tagliavini L, Gamberoni G, Drusco A, Marchesini J, Mascellani N, Sana ME, Abu Jarour R, Despons C, Teitell M, Baffa R, Aqeilan R, Iorio MV, Taccioli C, Garzon R, Di Leva G, Fabbri M, Catozzi M, Previati M, Ambros S, Palumbo T, Garofalo M, Veronese A, Bottoni A, Gasparini P, Harris CC, Visone R, Pekarsky Y, de la Chapelle A, Bloomston M, Dillhoff M, Rassenti LZ, Kipps TJ, Huebner K, Pichiorri F, Lenze D, Cairo S, Buendia MA, Pineau P, Dejean A, Zanesi N, Rossi S, Calin GA, Liu CG, Palatini J, Negrini M, Vecchione A, Rosenberg A, Croce CM: Reprogramming of miRNA networks in cancer and leukemia. *Genome Res* 2010, **20**:589-99.
29. Calin GA, Sevignani C, Dumitru CD, Hyslop T, Noch E, Yendamuri S, Shimizu M, Rattan S, Bullrich F, Negrini M, Croce CM: Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc Natl Acad Sci USA* 2004, **101**:2999-3004.
30. 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* 2005, **435**:834-38.

doi:10.1186/1756-0381-4-17

Cite this article as: Satoh and Tabunoki: Comprehensive analysis of human microRNA target networks. *BioData Mining* 2011 **4**:17.

# 神経変性と神経炎症の分子ネットワーク解析

Molecular Network Analysis of Neurodegenerative and Neuroinflammatory Diseases

佐藤 準一

Jun-ichi Satoh

近年、ヒトゲノムの解読が完了し、個々の細胞における遺伝子やタンパク質の発現情報を網羅的に解析可能なポストゲノム時代が到来した。創薬研究の中心は網羅的発現解析を統合したオミックス研究に基盤を置くゲノム創薬へとパラダイムシフトした。ヒトは大規模な分子ネットワークで精密に構築された複雑系であり、多くの難病がシステム固有の防御機構であるロバストネスの破綻に起因する。いまだ特効薬がない神経変性疾患 アルツハイマー病や炎症性脱髄疾患 多発性硬化症では、分子機序の解明および新規の標的分子に対する画期的な創薬が待望されている。最近、筆者らは神経疾患のオミックスデータに関して、分子ネットワークを詳細に解析して、創薬標的分子を同定した。今後はゲノムワイドの分子ネットワーク解析が、神経難病の病態解明や治療薬開発のためにますます重要な研究戦略となると思われる。



KeyMolnet, 分子ネットワーク, 創薬標的分子, システムバイオロジー

## はじめに

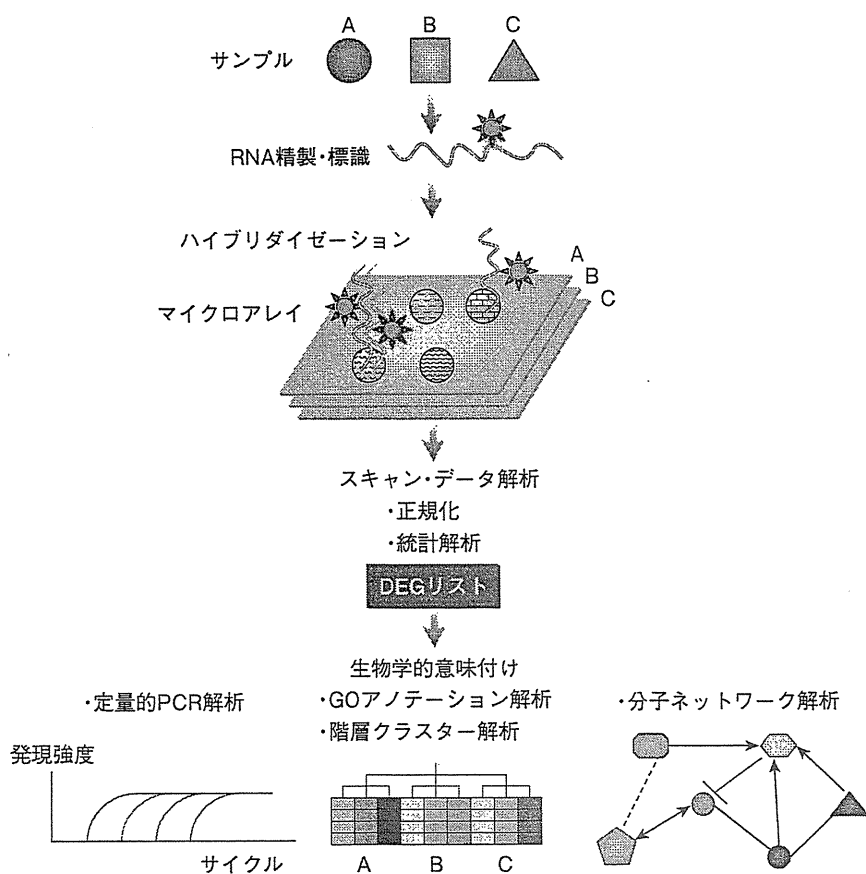
アルツハイマー病 (Alzheimer's disease ; AD) は、中高年期に発症し、進行性の認知機能障害を呈する神経変性疾患で、いまだ特効薬がない難病である。病理学的には、海馬や大脳皮質を中心に A  $\beta$  (amyloid beta) の蓄積と異常リン酸化タウを含む神経原線維変化の出現を主徴とし、広汎な神経細胞死を認める。若年発症家族性 AD では、プレセニリン (PSEN1, PSEN2) やアミロイド前駆体タンパク質 (APP) の遺伝子変異を認める。一方、大多数の AD は遺伝子変異がなく孤発性であり、いまだ不明の機序により A  $\beta$  産生増大、分解低下、凝集促進を来して、脳に大量の A  $\beta$  が蓄積し、タウの異常リン酸化と神経細胞死が誘導されると考えられている。特に早期から蓄積する A  $\beta$  オリゴマーは神経毒性が強い。一方、多発性硬化症 (multiple sclerosis ; MS) は、若年期に好発し、中枢神経系白質に炎症性脱髄巣が多発し、様々な神経症状が再発を繰り返して進行する難病である。MS では、遺伝的要因と環境因子の複雑な相互作用を背景に出現した活性化自己反応性 Th17 細胞や Th1 細胞が、血液脳関門を通過して脳や脊髄に浸潤し、マクロファージやミクログリアを活性化してサイトカインや活性酸素の産生を誘導し、脱髄を惹起すると考えられている。MS では、IFN- $\beta$  などの免疫調節薬が投与されているが、ノンレスポンド (無効例) も多い。現在まで、神経細胞や軸索・髄鞘の再生促進

薬はなく、新規の標的分子に対する画期的な創薬が待望されている。

2003年にヒトゲノムの解読が完了し、マイクロアレイ、質量分析装置、次世代シーケンサーを用いて、個々の細胞における遺伝子やタンパク質の発現情報を網羅的に解析可能なポストゲノム時代が到来した。近年、創薬研究の中心は網羅的発現解析を統合したオミックス研究に基盤を置くゲノム創薬へとパラダイムシフトした。同時に薬理ゲノミクスの分野は急成長を遂げ、薬物応答性の個人差をある程度予測可能となり、テーラーメイド医療 (personalized medicine) の樹立に道が開かれた。システムバイオロジー (systems biology) の観点からは、ヒトは大規模な分子ネットワークで精密に構築された複雑系であり、多くの難病がシステム固有の防御機構であるロバストネス (robustness) の破綻に起因すると考えられている<sup>1)</sup>。したがって神経難病の病態解明のためには、オミックス研究に直結したゲノムワイドの分子ネットワーク解析が重要な研究手段となりうる。最近、筆者らは神経疾患のオミックスデータに関して、分子ネットワークを詳細に解析することにより、創薬標的分子を同定した<sup>2)</sup>。以下に筆者らの研究を中心に分子ネットワーク解析の意義について概説する。

## I 網羅的発現解析から分子ネットワーク解析へ

2003年にヒト全遺伝子塩基配列が解読され、DNAマイク



■図1 網羅的発現解析から分子ネットワーク解析への流れ

比較対象となる遺伝子発現レベルが異なる数種類以上の細胞や組織からRNAを精製し、蛍光標識し、アレイとハイブリダイゼーションを行う。スキャン後に、シグナル強度を正規化し、サンプル間の遺伝子発現プロフィールを統計学的に比較解析し、有意な発現差異を呈する遺伝子群 (DEG) を抽出し、定量的PCRで検証する。生物学的意味付けのため、GO (Gene Ontology) のアノテーション (annotation) を調べ、階層クラスター解析を行い、KEGG, PANTHER, STRING, IPA, KeyMolnetを利用して分子ネットワークを解析する。

マイクロアレイを用いて、個々の細胞における数万遺伝子の発現情報を包括的に解析することが可能になった。最近では、高速次世代シーケンサーを用いて、発現量の低い遺伝子も含めて、一度に全遺伝子の発現解析が可能になっている。ヒト以外では、マウス・ラット・アカゲザル・イヌ・ウシ・イネ・ゼブラフィッシュ・ショウジョウバエ・酵母・線虫・大腸菌でも、マイクロアレイ解析が可能である。DNAマイクロアレイは、cDNAを基盤上にスポットで固定するスタンフォード方式と、直接基盤上でオリゴヌクレオチドを合成・伸長するフォトログラフ方式のGeneChip® (Affymetrix社) に大別される。さらに、スプライスバリエーションの網羅的解析が可能なエクソンアレイ、遺伝子多型マッピングや染色体コピー数を解析できるジェノタイプングアレイ、ChIP

(chromatin immunoprecipitation) on Chip解析に用いるゲノムタイリングアレイが市販されている。一方、プロテインマイクロアレイは、基盤上にリコンビナントタンパク質を高密度に固定してあり、タンパク質間相互作用 (protein-protein interaction; PPI) を網羅的に解析可能なチップである。この解析手法は、酵母two-hybrid法に比較して偽陽性率が低く、翻訳後修飾を受けたタンパク質との結合に関しても、鋭敏に検出できる利点がある。

マイクロアレイでは、比較対象となる遺伝子発現レベルが異なる2種類以上の細胞や組織 (例えば、正常細胞と癌細胞、治療前後の細胞など) からtotal RNAまたはmRNAを抽出し、cDNAやcRNAに変換して蛍光色素で標識後、フラグメントに切断してハイブリダイゼーションを行う (図1)。1色法では1サンプルに1アレイを使用し、アレイ間の発現レベルを比較解析する。一般に、同じ実験条件のサンプルに対して、アレイを2~3枚 (レプリケート) 使用する。プロテインマイクロアレイでは、タグを付加したプローブタンパク質をアレイ上のターゲットタンパク質と反応させ、蛍光標識した抗タグ抗体で検出する。アレイを専用の

スキャナーでスキャン後に、シグナル強度を正規化 (normalization) して、サンプル間の遺伝子発現プロフィールを統計学的に比較解析する。マイクロアレイ解析では、一度に非常に多くの遺伝子の発現レベルを解析するため、遺伝子ごとにt検定などで評価すると、偽陽性遺伝子を多数拾ってしまう。通常は多重検定を行いBonferroniの補正を付加するか、または偽陽性率 (false discovery rate; FDR) を評価する。最終的に、サンプル間で有意な発現差異を呈する遺伝子群 (differentially expressed genes; DEG) を抽出し、発現レベルを定量的PCRで検証する。

次に、DEGに関して生物学的意味付けを行う。初めに個々の遺伝子のアノテーション (annotation) を調べる。NCBI (National Center for Biotechnology Information) のデータベースEntrez Geneを利用して、1つずつGO

(Gene Ontology)のCellular Function, Cellular Process, Cellular Componentを調べることも可能だが, DAVID Bioinformatics Resources (david.abcc.ncifcrf.gov)のFunctional Annotationツールを用いると, 膨大な遺伝子セットのアノテーションを一括して解析できる<sup>3)</sup>. 多数のサンプルを比較解析する場合は, データセットの要素特性を分類するために, GeneSpring<sup>®</sup> (Agilent社)やCluster 3.0 (bonsai.ims.u-tokyo.ac.jp/~mdehoon/software/cluster)などのツールを用いて, DEGを指標に階層クラスター解析(hierarchical clustering analysis)を行うと, 発現プロフィールのビジュアルな比較ができる.

さらに, DEGが構成する分子ネットワークを解析すると, 生物学的意味をより明確に把握することができる(図1). 体内では, 遺伝子でコードされたタンパク質は複雑なネットワークから成るシステムを構築している<sup>1)</sup>. PPIには, 直接的結合関係のみならず, 活性化, 不活性化, 酵素反応, 運搬, 複合体形成など多彩な相互作用様式が存在する. 複雑多岐のオミックスデータに関連している分子ネットワークを同定するためには, 精査された文献情報に裏付けられた専用の解析ツールを使う必要がある. すなわち, 膨大な文献情報から様々な分子間相互作用を抽出し, 信頼性が高い知識を整理して, コンテンツとして収録した知識データベース(knowledgebase)を用いて, 既知のどのネットワークやパスウェイ(canonical network/pathway)に最も高い類似性を呈しているかについて, 統計学的手法で解析する方法である. 無償で利用できる代表的なデータベースには, KEGG (Kyoto Encyclopedia of Genes and Genomes) (www.kegg.jp)<sup>4)</sup>, PANTHER (the Protein Analysis Through Evolutionary Relationships) classification system (www.pantherdb.org)<sup>5)</sup>, STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) (string.embl.de)<sup>6)</sup>がある. KEGGとPANTHERは, キュレーターと呼ばれる専門家により精査された遺伝子や代謝物に関する情報を収録している. 2011年4月現在, KEGG PATHWAYには392 reference pathwaysから構成される134,607種類のパスウェイが収録されている. 目的とする遺伝子やタンパク質のセットを, DAVID Functional Annotationツールに入力すると, 統計学的検定を行い, 最も密接に関連しているKEGGパスウェイを同定できる. PANTHERでも同様にリファレンスセットとの比較により, 類似性の統計学的有意差を多重検定で評価することができる. STRINGはKEGG, HPRD (Human Protein Reference Database), BIND (Biomolecular Interaction Network

Database), IntAct Molecular Interaction Databaseに登録されている情報も統合して収録している.

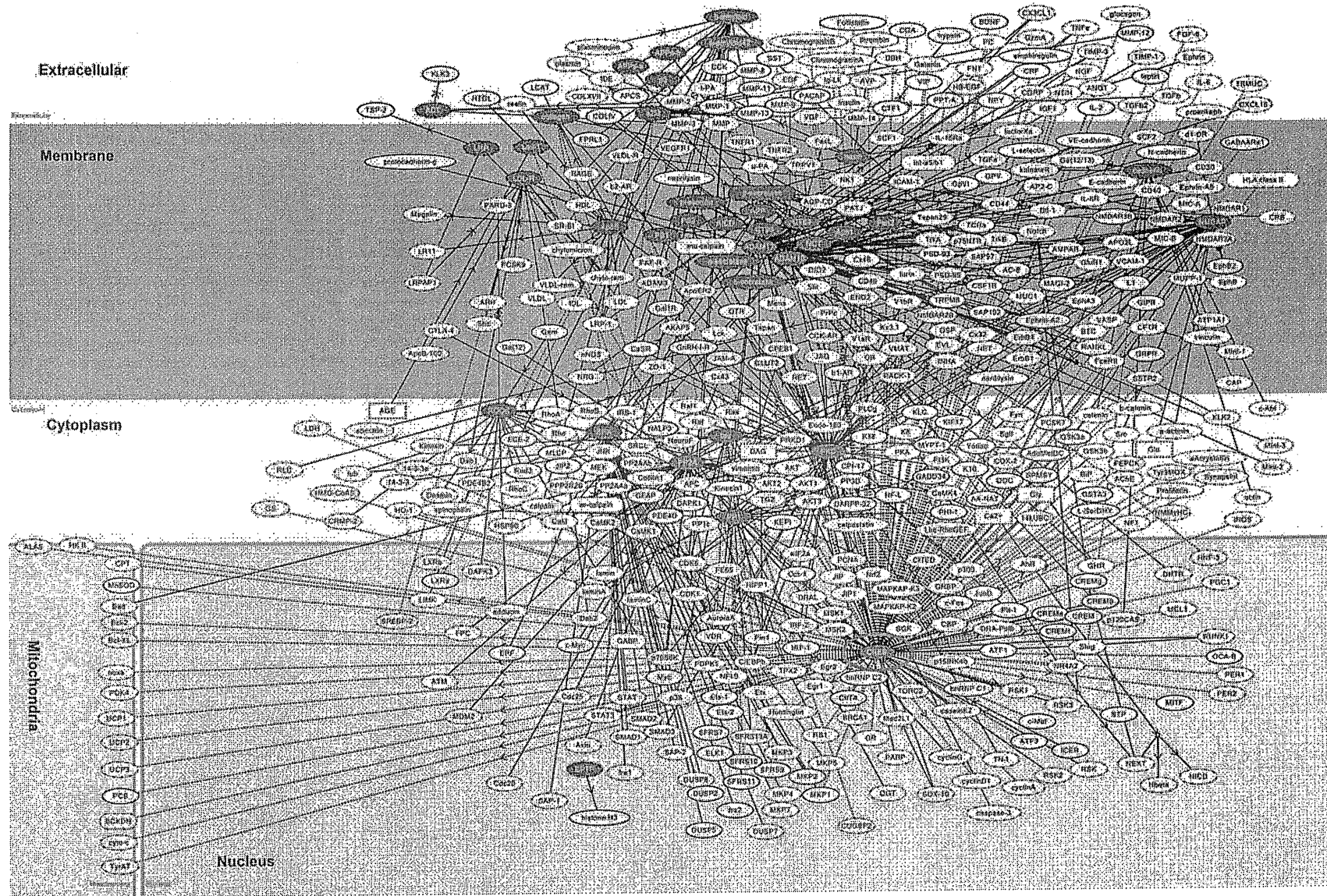
また, 有償ツールとしては, IPA (Ingenuity Pathways Analysis)<sup>®</sup> (Ingenuity Systems, Redwood City, CA)やKeyMolnet<sup>®</sup> (Institute of Medicinal Molecular Design, Tokyo)などがある. これらは精選された文献を専門家が精読して, 分子間相互作用に関する信頼性の高い情報を選択して収集しており, 定期的にアップデートされている. KeyMolnetは日本語入力にも対応しており, 種々の疾患のメディエート分子を整理して収録している. また, 臨床試験中を含む既存の医薬品の標的分子も明示されている. 検索法として, 結合・発現制御・複合体形成を包括的に調べる周辺検索(neighboring search), 発現制御に関与する転写因子群を調べる共通上流検索(common upstream search), 始点と終点間のネットワークを調べる始点終点検索(N-points to N-points search), 複数の端点を始点として, 最多数の始点を含む最小の分子ネットワークを調べる相互関係検索(interrelation search)を選択できる<sup>2)</sup>.

解析ツールで描画した分子ネットワークから, 創薬標的分子を探索する場合は, 多数の分子からのリレーションが集中しているハブ(hub)と呼ばれる中心分子を同定することが重要である. ハブの抑制薬または活性化薬は, ネットワークのロバストネス維持に重大な影響(治療効果や毒性・副作用)を及ぼす<sup>2)</sup>.

## II 分子ネットワーク解析から見たADの創薬標的分子

2004年にBlalockらは, 年齢を一致させた31例の高齢者の剖検海馬CA1脳組織から抽出したRNAを用いて, GeneChip<sup>®</sup>HG-U133Aで遺伝子発現を網羅的に解析した<sup>7)</sup>. 生前に施行した知能検査MMSE (mini-mental state examination)のスコアに従って, 正常9例, 早期AD7例, 中等症AD8例, 重症AD7例にグループを分類した. 彼らは3,413種類の全AD関連遺伝子(発現上昇1,977, 発現低下1,436)と609種類の早期AD関連遺伝子(発現上昇431, 発現低下178)を同定し, 公開した. 早期AD関連遺伝子には癌抑制遺伝子やオリゴデンドロサイト成長因子が集積していたが, これらの分子が構成するネットワークは解析されなかった. 筆者らは彼らのデータセットを用いて, KeyMolnetの共通上流検索法でAD脳における病態形成に関与してい





■図2 アルツハイマー病の分子ネットワーク解析

KeyMolnetに収録されている42種類のアルツハイマー病疾患メディエート分子(●)を入力して、上下流1パス周辺検索法で、分子ネットワークを解析した。CREB(●)による発現調節との関連性が最も強く示唆された。矢印付き実線は直接結合および活性化、ストップ矢印付き実線は不活性化、矢印付き点線は転写活性化、ストップ矢印付き点線は転写抑制化を示す。



る分子ネットワークを解析した<sup>8)</sup>。その結果、全ADおよび早期ADの関連遺伝子群が構成する分子ネットワークは、両者とも転写因子CREB (cAMP-response element-binding protein) による発現調節と密接に関連していることがわかった。CREBは、成長因子やホルモンの刺激で133位のセリン残基(Ser133)がプロテインキナーゼA (PKA) によりリン酸化されて活性化し、標的遺伝子プロモーターのCRE (cAMP response element) に結合し、標的遺伝子の転写を活性化する。筆者らは、AD海馬脳組織の抗pCREB抗体による免疫組織染色からpCREBが顆粒空胞変性 (granulovacuolar degeneration ; GVD) に集積していることを見いだした。GVDはオートファゴソームとして働いている可能性があり、筆者らの知見は、ADにおけるオートファジーによるタンパク質代謝制御系の異常を示唆している。

また、KeyMolnetには、精選された文献から専門家により収集された42種類のAD疾患メディエート分子が収録されている。KeyMolnetの上下流1パス周辺検索法で、これら42分子が構成するネットワークを解析したところ、再びCREBによる発現調節との関連性が最も強く示唆された ( $p = 2.225E-308$ ) (図2)。以上の結果は、CREBがAD病態ネットワークのハブとして働くことを示している。KeyMolnetライブラリーのCREBによる発現調節 canonical pathwayは、162分子 (CREBおよび入力27分子と出力134分子) で構成されている (図3)。このうち65分子 (図3●) は、現在市販薬や開発薬が存在している薬の標的分子である。cAMP/PKA/CREBシグナル系活性化を促進するホスフォジエステラーゼ4 (PDE4) 抑制薬であるロリプラム (Rolipram) をAD動物モデルAPP/PS1トランスジェニックマウスに投与すると、認知機能を改善した<sup>9)</sup>。

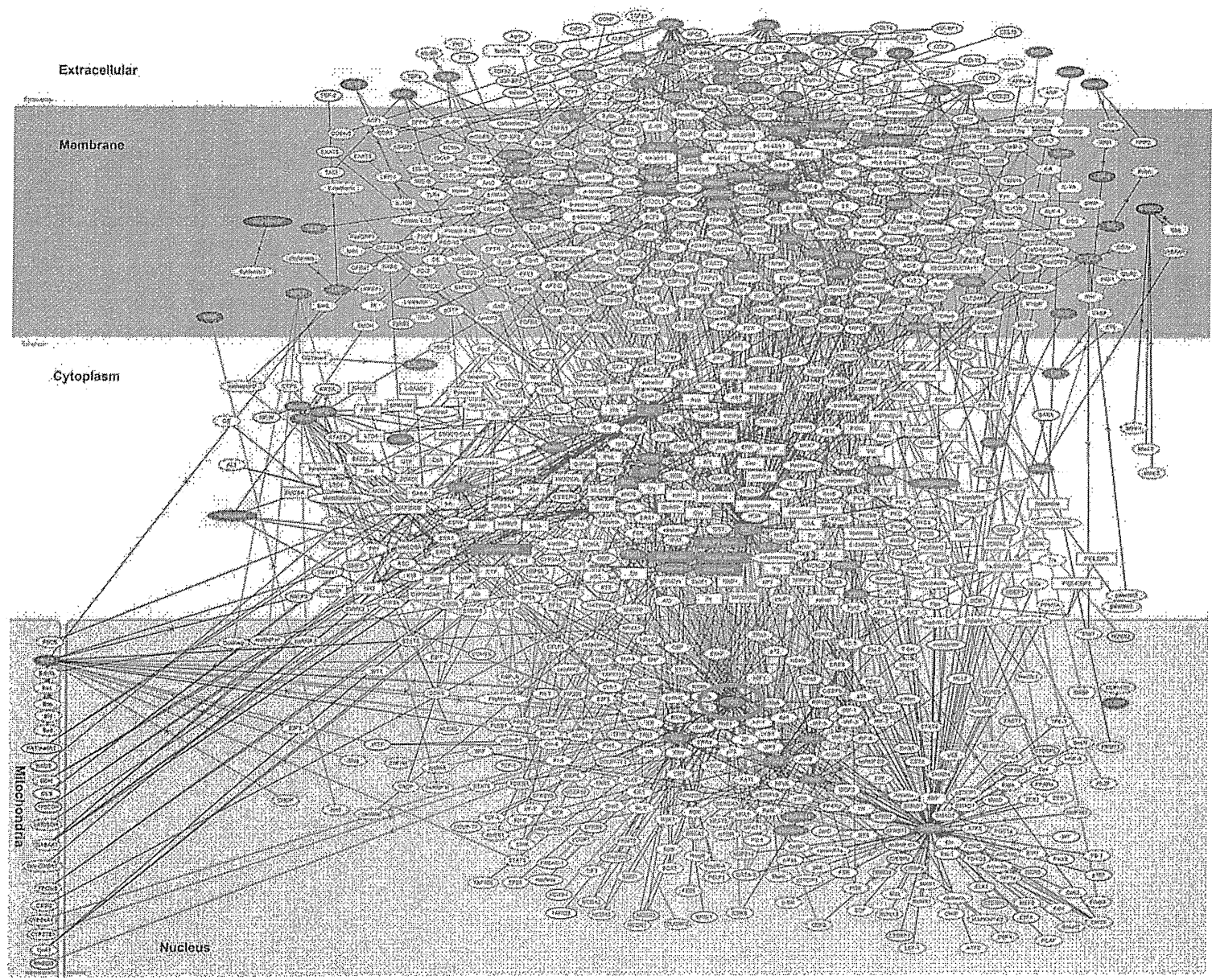
### III 分子ネットワーク解析から見たMSの創薬標的分子

2008年にHanらは、6例のMS凍結脳を用いて、病理学的ステージを確認した脳病巣からレーザーマイクロダイセクションで採取したサンプルを、SDS-PAGEで分離後に、タンパク質を抽出し、トリプシン消化ペプチド断片を質量分析で網羅的に解析した<sup>10)</sup>。病理学的ステージに関しては、炎症性細胞浸潤と浮腫を主徴とする急性脱髄巢 (active plaque ; AP)、炎症が脱髄巢辺縁部に局限している慢性活動性脱髄巢 (chronic active plaque ; CAP)、炎症所見に乏しくグリア瘢痕を主徴とする慢性非活動性脱髄巢 (chronic plaque ; CP)

に分類した。同時に2例の健常脳のプロテオームも解析した。健常脳では検出されず、かつ各ステージ特異的なタンパク質を選出し、AP 158, CAP 416, CP 236種類のプロテオームデータを公開した<sup>10)</sup>。彼らはCAPにおいて5種類の血液凝固系タンパク質の発現を認めた。この所見に基づいて、抗凝固薬であるトロンビン阻害剤Hirudinおよび活性化プロテインCを用いて、MS動物モデルであるマウス自己免疫性脳脊髄炎 (experimental autoimmune encephalomyelitis ; EAE) を治療した。どちらの抗凝固薬も、抗原特異的リンパ球の増殖とIL-17, TNF- $\alpha$ 産生を抑制した。以上の結果より、血液凝固系タンパク質が新規MS創薬標的分子となることが示唆された。しかしながら大多数を占める凝固系以外のタンパク質に関しては、MS脳分子病態における意義は明らかされなかった。

筆者らは、HanらのプロテオームデータをKEGG, PANTHER, IPA, KeyMolnetに入力し、ステージ特異のプロテオームを最もよく反映している分子ネットワークを同定した<sup>11)</sup>。4種類の異なるツールは様々な独自の分子ネットワークを抽出したが、共通してCAP, CPプロテームにおける細胞外基質 (extracellular matrix ; ECM) -インテグリンシグナル伝達系の中心的役割を示唆した<sup>11)</sup>。インテグリンは複数の $\alpha$ ,  $\beta$ サブユニットから構成される24種類のヘテロダイマー分子で、ECMリガンドとして働く。 $\beta$ 1インテグリンファミリーはコラーゲン, フィブロネクチン, ラミニンと結合し、 $\alpha$ <sub>v</sub>インテグリンファミリーはビトロネクチンと結合する。ECM-インテグリン系は細胞接着, 遊走, 分化, 増殖に必須なシグナルを伝達する。MS慢性病巣においては、髄鞘や軸索の再生が著しく乏しい。その理由として、グリア瘢痕に含まれているECMが再生阻害因子として働く可能性や、活性化マクロファージやミクログリアが産生するタンパク質分解酵素がECMに結合して長期に保持され、その結果、髄鞘崩壊が遷延化している可能性がある<sup>12)</sup>。現在欧米においてMS再発抑制のために、 $\alpha$ 4 $\beta$ 1インテグリン (VLA4) に対するヒト化モノクローナル抗体ナタリズマブ (Natalizumab) が臨床で用いられている。しかしながら、ナタリズマブは致死的な進行性多巣性白質脳症 (progressive multifocal leukoencephalo pathy ; PML) を惹起する危険性があり、安全な薬の開発が必要である。ECM-インテグリンシグナル伝達系では、FAK (focal adhesion kinase) がハブとして働く。低分子化合物TAE226は、ECMによるFAKの自己リン酸化を選択的に抑制し、*in vivo*モデル系では経口投与により腫瘍細胞の増殖と血管新生を抑制する<sup>13)</sup>。





■図4 多発性硬化症の分子ネットワーク解析

KeyMolnetに収録されている91種類の多発性硬化症疾患メドエート分子(●)を入力して、上下流1パス周辺検索法で、分子ネットワークを解析した。VDR(○)による発現調節との関連性が最も強く示唆された。

したがってMSにおいて、TAE226はFAKを標的分子とする慢性炎症性脱髄抑制薬の候補となる可能性がある。

また、KeyMolnetには、精選された文献から専門家により収集された91種類のMS疾患メディエート分子が収録されている。KeyMolnetの上下流1パス周辺検索法で、これら91分子が構成するネットワークを解析したところ、転写因子であるビタミンD受容体(vitamin D receptor: VDR)による発現調節との関連性が最も強く示唆された( $p = 5.793E-237$ ) (図4)。以上の結果は、VDRがMS病態ネットワークのハブとして働くことを示している。日照時間が短い高緯度地域ではMSの発症頻度が高いという疫学的所見より、ビタミンDはMS発症抑制因子として働くと考えられている<sup>14)</sup>。VDRは活性型ビタミンD(1,25-ジヒドロキシビタミンD)と結合すると、レチノイドX受容体(retinoid X receptor: RXR)とヘテロダイマーを形成して、標的遺伝子プロモーターのVDRE(vitamin D response element)に結合し、標的遺伝子の転写を活性化し、免疫調節作用など多彩な生理活性を呈する。しかしながら現在まで、MSにおいて、ビタミンDの有効性が明示された臨床試験は見当たらない。

## おわりに

膨大なオミックスデータに関与する分子ネットワークを手際よく解析するためには、精査された文献情報に基づく解析

ツールを使う必要がある。しかしながら解析ツールはまだまだ発展途上かつ日進日歩であり、現時点では、どのツールもスプライスバリエーション、翻訳後修飾、細胞特異的発現、細胞内局在化、動的特性に関しては十分対応できていない。生命現象を複雑なシステムとして捉えるシステムバイオロジーの観点からすると、神経難病では分子ネットワークを詳細に解析することにより、初めて論理的な仮説に裏付けられた創薬標的分子を効率的に同定することができると思われる。

謝辞 本稿で紹介した研究は、国立精神・神経医療研究センター神経研究所免疫研究部山村隆部長、明治薬科大学バイオインフォマティクス天竺桂弘子助教との共同研究でなされ、文部科学省基盤研究(C22500322)と私立大学戦略的研究基盤形成支援事業明治薬科大学ハイテクリサーチセンター研究事業(S0801043)および厚生労働科学難治性疾患克服研究事業(H21-難治-一般-201; H22-難治-一般-136)の補助を受けた。

## PROFILE 佐藤準一

- 明治薬科大学薬学部 生命創薬科学科 バイオインフォマティクス 教授
- E-mail: satoj@my-pharm.ac.jp
- 趣味: 海釣り と 化石発掘

1988年東京医科歯科大学大学院医学研究科博士課程修了、医学博士、日本神経学会専門医。現在の研究テーマ: 神経疾患病態関連遺伝子群の網羅的解析。

## 文献

- 1) Kitano H: Nat Rev Drug Discov (2007) 6: 202-210
- 2) Satoh J: Clin Exp Neuroimmunol (2010) 1: 127-140
- 3) Huang da W, et al: Nat Protoc (2009) 4: 44-57
- 4) Kanehisa M, et al: Nucleic Acids Res (2010) 38: D355-D360
- 5) Mi H, et al: Nucleic Acids Res (2010) 38: D204-D210
- 6) Szklarczyk D, et al: Nucleic Acids Res (2011) 39: D561-D568
- 7) Blalock EM, et al: Proc Natl Acad Sci USA (2004) 101: 2173-2178
- 8) Satoh J, et al: Dis Markers (2009) 27: 239-252
- 9) Gong B, et al: J Clin Invest (2004) 114: 1624-1634
- 10) Han MH, et al: Nature (2008) 451: 1076-1081
- 11) Satoh JI, et al: Mult Scler (2009) 15: 531-541
- 12) van Horsen J, et al: J Neurochem (2007) 103: 1293-1301
- 13) Liu TJ, et al: Mol Cancer Ther (2007) 6: 1357-1367
- 14) Ascherio A, et al: Lancet Neurol (2010) 9: 599-612

## for beginners

- ・「マイクロアレイデータ統計解析プロトコール」藤淵 航・堀本勝久 編, 羊土社 (2008) (マイクロアレイデータ解析入門書)
- ・「創薬・タンパク質研究のためのプロテオミクス解析」小田吉哉・長野光司 編, 羊土社 (2010) (プロテオーム解析入門書)