search method." In the "Neighboring search method", one or more molecules are selected as the starting points and a molecular network is generated around the starting molecules within the designated number of paths from the starting point molecules (specified as the "path limit" parameter). In the "N-points to N-points search method," two groups of molecules are selected as the starting points and the end points and a molecular network connecting the starting point molecules to the end point molecules is generated. In the "Common upstream/downstream search method," two or more molecules are selected as the starting points, and paths to the common upstream and/or downstream molecules are searched.

Logical Operation Method

The intersectional or difference networks of the networks generated by the "Network search methods" described above, can be searched and extracted by the logical operation methods "AND" or "DIFF" (meaning subtraction), respectively. Using the "OR" logical operation, multiple networks can be merged.

Importing Experimental Data into Keymolnet

Users can import a list of molecule IDs with the numerical data, such as fold change values and expression levels. Conversion tables for converting molecule/probe/gene IDs for DNA microarray or proteomics etc. into the KeyMolnet IDs have already been built into KeyMolnet. If numerical values are included in the list, they can also be imported and stored in a temporary table. Molecules on the network can be color-coded according to the numerical data by specifying the range of numerical values by corresponding colors.

DNA Microarray Data and Data Processing

Model experiments on all microarray analyses were carried out at the National Institute of Health Sciences, Japan, as follows: The HL-60 cells were exposed to 20 nM TPA for one hour and the biotin-labeled cRNA was prepared. The total RNA was prepared from the dishes of HL-60 cells at 1, 3 and 9 hours after treatment with TPA. The U95A arrays (Affymetrix®) were used for the experiments. Details on the experiment have been published elsewhere [10, 11]. Data processing was performed by the FUMI theory [12] and KeyMolnet.

RESULTS AND DISCUSSION

Here we have the gene expression data obtained by Gene chip® in HL-60 cells at1, 3 and 9 hours after treatment with TPA. We show the mechanism analysis using the DNA microarray data with a time series. First, a file including the gene IDs and their corresponding fold change values was prepared for the dataset at each time point according to the standard format in KeyMolnet. The gene IDs in the list were converted into KeyMolnet IDs with the proprietary conversion table. At 1 hr after treatment with TPA, the molecules with 2-fold or higher gene expression levels in the cells compared to those in the controls, were extracted by specifying the threshold value as the "lower limits of 2 or more" here. The molecules are listed in the Tree View of KeyMolnet (not shown). At 3 and 9 hrs after treatment with TPA, the molecules with 2-fold or higher gene expression levels in the cells compared to those in the controls were also extracted and listed in the same way. Next, the networks were generated based on the molecules with 2-fold or higher gene expression levels in the cells at each time point using the "Neighboring search method." The molecules at each time point (the up-regulated genes induced by TPA in the cells) were selected as the starting points here. After generating the networks, the molecules on the networks were color-coded according to their fold change values by specifying the color based on the range of the fold change values. Figure **1** shows one of the networks generated based on the up-regulated molecules in the cells at 9 hours after treatment with TPA compared to those in the controls. The network was generated using the "Neighboring search method" and was specified by the "path limit of 2." Molecules colored red, orange, light blue, blue and gray indicated expression levels of over 2-fold higher, within 2-fold higher, within 2-fold lower, under 2-fold lower, and not detectable, respectively (see the upper-right corner of Fig. **1**). Molecules represented as nodes and various oval shapes, capsule-like shapes, rectangles and hexagons indicated proteins, complexes, small molecules and exogenous molecules, respectively (Fig. **1**). Each different type of molecular relation is represented by a different type of edge line, an arrow and colors in Figure **1**. Here we have the three sets of specific networks generated based on the up-regulated molecules at the three time points. Further insight could be obtained and discussed by using the logical operations "AND," "OR" and/or "DIFF" between the three sets of networks at each time point. The networks shown in Figure **2** are the ones commonly activated in the cells at 1, 3 and 9 hrs after treatment with TPA, and were created by carrying out the "AND" operation to extract the intersectional networks of the three networks at each time point. As the hundreds of canonical pathways for "signal transduction pathways," "metabolic pathways" and "transcriptional regulations" are manually compiled in the content database, based on the pairs of molecular relations, users can recognize to which canonical pathways a pair of molecules (as a molecular relation on the generated pathways) belong, because the molecular relations on the "Network View" operate simultaneously with the names of the canonical pathways in the "Tree View" of KeyMolnet. In this case, the five names of the canonical pathways, namely, "Transcriptional regulation by NFkB," "Transcriptional regulation by GR," "Transcriptional regulation by NFAT," "Transcriptional regulation by PPARa" and "Wnt signaling pathway," are listed in the "Tree View," and users can recognize to which canonical pathways each molecular relation on the networks belong. For example, by clicking the relation " GR→IkBa" on the network, the dotted line with an arrow of "Molecular Relation" is highlighted in red and, at the same time, the string- "Transcriptional regulation by GR" is highlighted in red (Fig. **2**). Interestingly, the Wnt signaling cascade, well-known for its involvement in colorectal cancer because the APC gene, the target gene of familial adenomatous polyposis, is located at the lower stream of the Wnt in the pathway [13], was suggested to be involved in the HL-60 cells too. Since analysis suggested that the "Wnt signaling pathway" is involved in the action of TPA on the HL-60

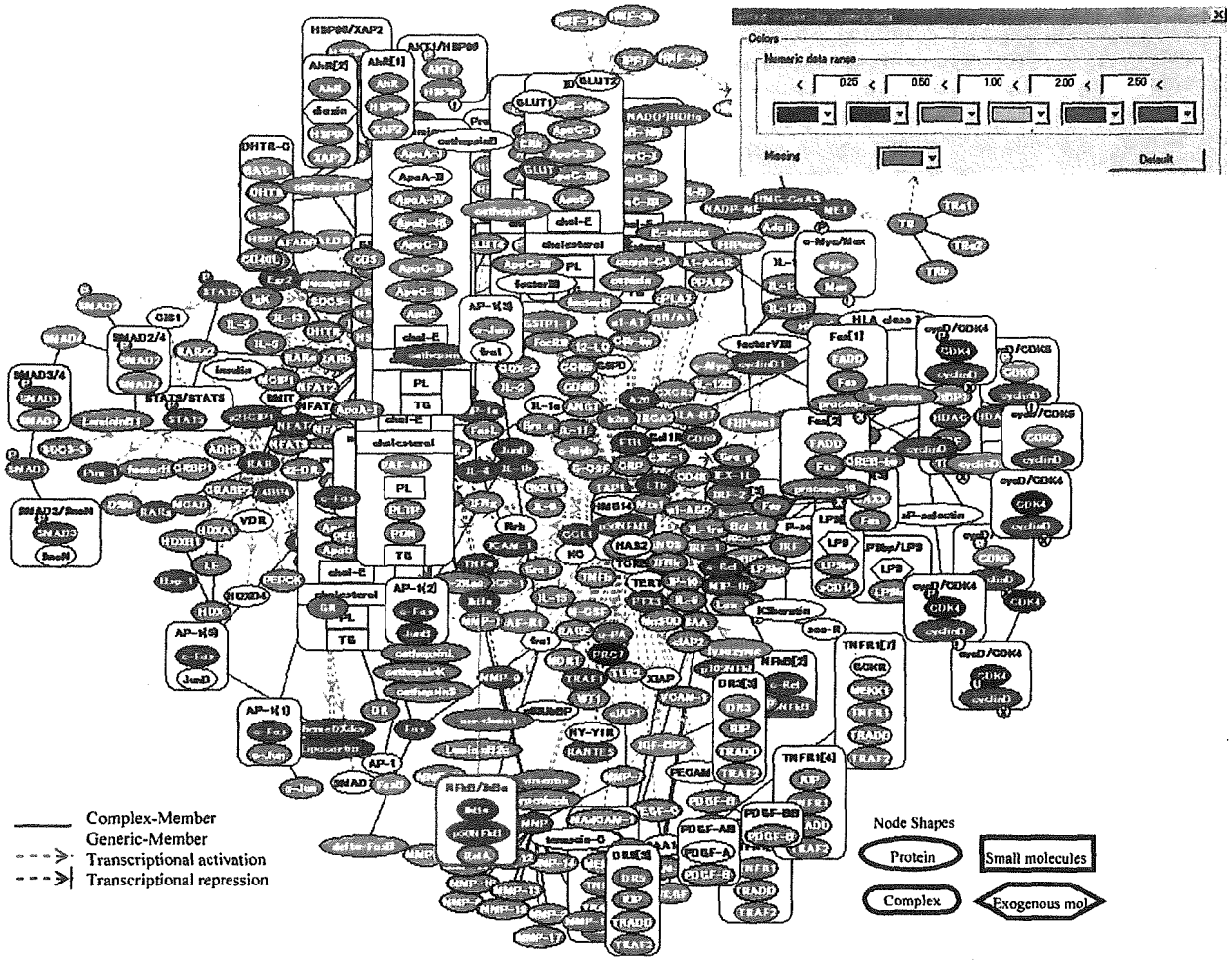

Fig. (1). Network generated based on up-regulated genes in HL-60 cells at 9 hours after treatment with TPA.
A list of gene expression data including the gene IDs and the fold change values corresponding to the IDs were imported into KeyMolnet. The molecules with 2-fold or higher gene expression levels in the HL-60 cells at 9 hrs after treatment with TPA compared to those in the controls, were extracted in KeyMolnet. A network was generated based on the extracted molecules using the "Neighboring search method" (see "MATERIALS AND METHODS"). Oval, rectangular, capsule-like and hexagonal shapes represented proteins, small molecules, complexes and exogenous molecules, respectively. The types of "Molecular Relations" were represented by the different edges, arrows and colors shown in the figure. The color code of the "Molecules" was specified according to the range of fold change values corresponding to the gene IDs imported as the data list. The range of fold change values and their corresponding colors are shown in the upper-right corner of the figure, that is, red indicates 2-fold or higher expression compared to that of the controls; orange indicates within 2-fold higher expression compared to that of the controls; light blue indicates within 2-fold lower expression compared to that of the controls; blue indicates 2-fold lower expression compared to that of the controls; and gray indicates not detectable levels. All the "Molecules" on the network are shown in the "KeyMolnet IDs" on each node in KeyMolnet, and they are used in all figures.

cells, we tried to observe the gene expression in the cells of the entire "Wnt signaling pathway." Figure **3A** shows that the canonical "Wnt signaling pathway" collected in KeyMolnet and the "Mediating Molecules" (see "Materials and Methods") are displayed on the pathway by coloring the molecules that correspond to the "Mediating Molecules" (displayed in blue). The list of diseases is displayed in the "Tree View." Users can see to which diseases the "Mediating Molecules" belong, because disease names and "Mediating Molecules" are linked together (not shown). The "Mediating Molecules" highlighted in orange indicate the molecules related to colorectal cancer. Users can also browse additional information, such as "Pathological Events," on disease (see "MATERIALS AND METHOD). Many molecules are suggested for involvement in colorectal cancer on the "Wnt signaling pathway."

When the gene expression data (in this case, fold change values) at each time point are displayed on the pathway, the changes with time in expression levels could clearly be observed. For example, Figure **3B** shows the gene expression in the HL-60 cells at 3 hrs after treatment with TPA to be in the "Wnt signaling pathway." The Wnt signaling cascade has not been suggested to have any relationship with human promyelocytic leukemia cells so far, but it may be involved in the HL-60 cells differentiation caused by TPA.

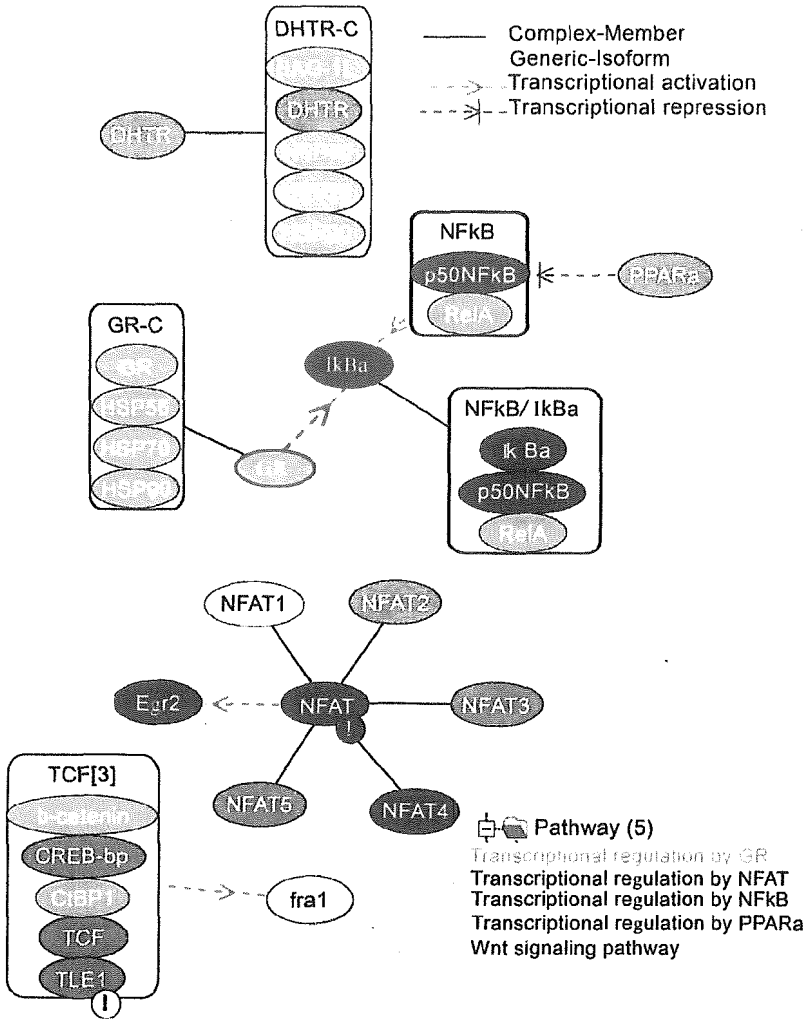

Fig. (2). Intersectional networks extracted from the three time-series networks.
After the networks were generated at each time point in the same way as described in Figure **1**, the intersectional networks of the three time-series networks were extracted. The fold change values are also displayed on the networks by which "Molecules" are color-coded according to the fold change values. In Figure **2**, the fold change values displayed on the "Molecules" of the networks, are used for the ratios of gene expression levels in the cells at 9 hours after treatment with TPA compared to those in the controls. Each "Molecular Relation" belongs to any of five pathways called "Transcriptional regulation by NFkB," "Transcriptional regulation by GR," "Transcriptional regulation by NFAT," "Transcriptional regulation by PPARa" and "Wnt signaling pathway," which can be found in the list in KeyMolnet at the time. By clicking an edge representing a "Molecular Relation," the linked pathways are highlighted in red, because the canonical pathways were manually compiled in the "content database" based on pairs of molecules (see "MATERIALS AND METHODS"). Users can see the correspondence between the "Molecular Relation" and the canonical pathways to which it belongs. In Figure **2**, for example, by clicking the "Molecular Relation" between "GR (glucocorticoid receptor) and IkBa (NFkB inhibitor α)" (highlighted in red), "Transcriptional regulation by GR" in the list is simultaneously highlighted in orange. The colors indicate the ratios of the levels of gene expression in the cells at 9 hours after treatment with TPA and are assigned the same as in Fig. **1**.

The "IFNa, IFNb and IFNg signaling pathways" were suggested to be activated by TPA in the cells. When the network subtraction of the "1 hr after networks" from the "3 hrs after networks" was carried out by the "DIFF" operation function, the "IFNa, IFNb and IFNg signaling pathways" were suggested to be involved in the action of TPA on the cells (now shown). The overview of activation of the overall "IFNa, IFNb and IFNg signaling pathways" in the cells over time was obtained by calling up the three merged pathways and by displaying the fold change values at each time point on the networks separately (Fig. **4**). As shown in the colors representing fold change values, displayed in the "Molecules" on the networks, signals were suggested to transduct through the "IFNg signaling pathway." In fact, Sakamoto *et al.* reported the finding that the TPA-response element (TRE) is present in the -128 to -109 bp region of the interferon-γ receptor 1 chain gene (IFNGR1) in THP-1 cells in 2001 [14]. Only the gene expression data in the HL-60 cells and analysis using KeyMolnet would have suggested the involvement of the "IFNg signaling pathway" in the cell differentiated into a macrophage-like cell by the action of TPA, and the involvement of the TRE in IFGR1 from the

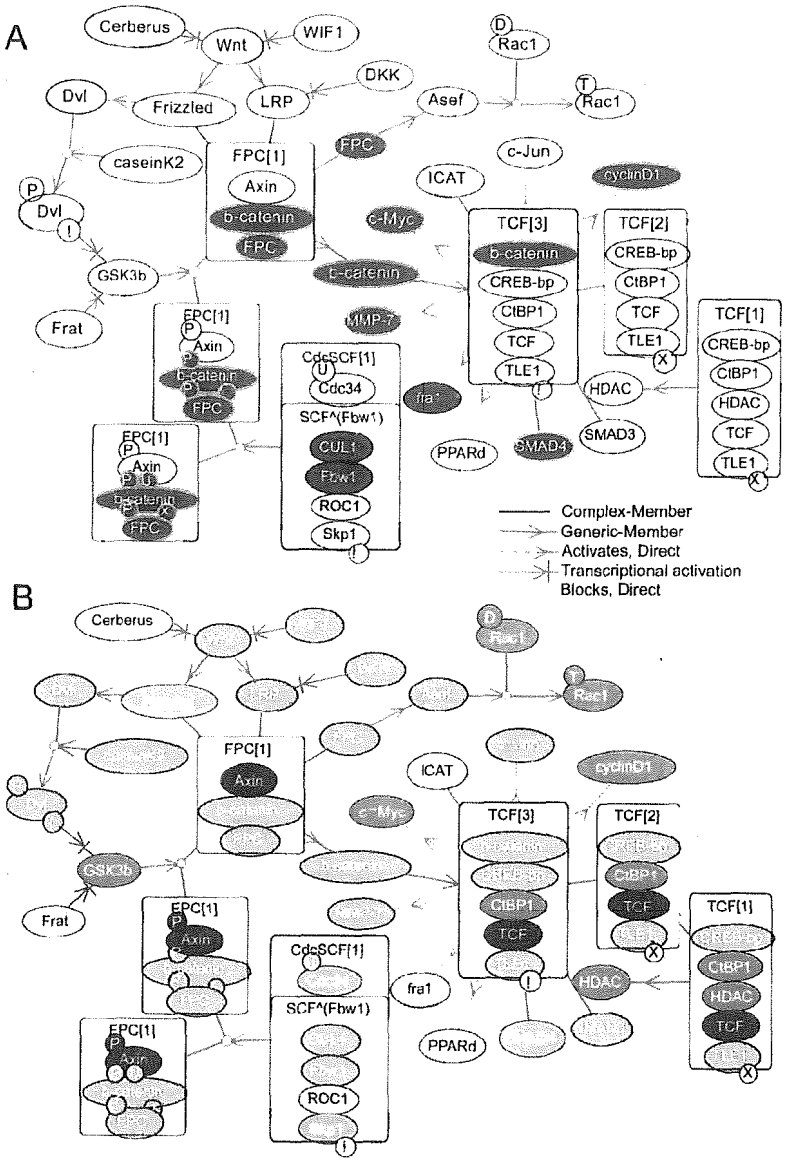

Fig. (3). "Mediating Molecules" and gene expression in HL-60 cells at 3 hours after treatment with TPA on "Wnt signaling pathway"
(A) The canonical "Wnt signaling pathway" was called up by the "Pathway search," because analysis by KeyMolnet suggested that the pathway is involved in the action of TPA on the cells. The "Molecules" shown in blue represent "Mediating Molecules" (see "MATERIALS AND METHODS"). Users can see to which disease in the list the "Mediating Molecules" belong and can also see other information on the disease (not shown). "Molecules" highlighted in orange indicate "Mediating Molecules" related to colorectal cancer. Many molecules are suggested to be involved in colorectal cancer on the "Wnt signaling pathway."
(B) The fold change values of the "3 hrs after" data are displayed on "Molecules" in the canonical "Wnt signaling pathway" shown as in Fig. **3A**. When the fold change values at each time point are displayed on the same pathways, transduction of signals on the pathways could be observed. Each color indicates the same as in Figures **1** and **2**.

activation of the pathway. Thus, using KeyMolnet in the analysis of gene expression data, a possible working hypothesis could be set up.

Next, we show a working hypothesis of the induction of apoptosis in the TPA-treated cells using the same gene expression data. Using the "N-points to N-points search method," the pathways from the "causes" to the "results" could be obtained this time. Here, in order to analyze the mechanisms of the relation between the induction of apoptosis and the molecules up-regulated by the action of TPA in the HL-60 cells, the up-regulated molecules are the ones whose gene expression levels were 2-fold or higher in the cells treated with TPA compared to those in the controls and were specified as the starting points and biological phenomenon -"apoptosis" itself was specified as the end point. The "N-points to N-points search method" was then carried out. As information on the "Biological Process" in

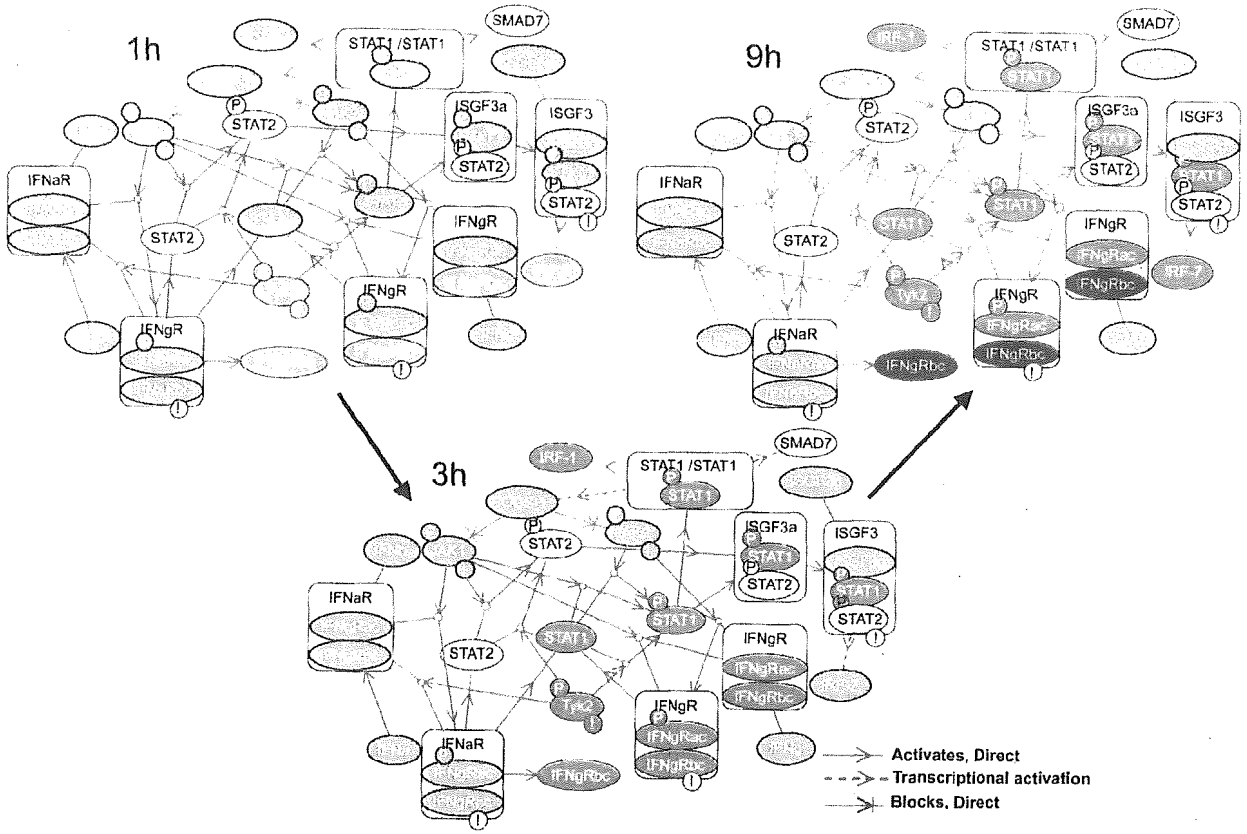

Fig. (4). Change in the levels of gene expression in HL-60 cells over time on "IFNa, IFNb and IFNg signaling pathways".
The IFNa, IFNb and IFNg signaling pathways were merged and called up because the pathways were suggested to be involved in the action of TPA on the cells using the "DIFF" logical operation function of KeyMolnet for the networks of "3 hrs after " and "1 hr after." The fold change values at each time point are displayed separately on "Molecules" in the three merged pathways. The pathways are suggested to be activated with time in the cells 3 hrs after treatment with TPA. Each color indicates the same as in Figures **1**, **2** and **3**.

GO was also collected for the molecules corresponding to their Swiss Prot IDs using GO Annotation (GOA), "Molecules" corresponding to the GO term "Apoptosis" can be set as selected points. In case of a number of molecules corresponding to one GO term, such as "Apoptosis", searched networks could be very large and complicated. Therefore, it could be better to set one or a small number of molecules as selected point(s) to analyze the mechanism of the relation between the "causes" and the "results". For example, one of the molecules involved in apoptosis, "Fas" [15], is selected as the end point here. Figure **5** shows the pathway created based on the molecules with 2-fold or higher gene expression levels in the cells at 1 hr after treatment with TPA compared to those in the controls, using the "N-points to N-points search method," setting the molecules as the starting points and the "Fas" as the end point. The gene expression data were also displayed in the "Molecules" on the network. The network suggests that the pathway from the up-regulated molecules to the "Fas" is mediated by the transcription factors NFkB, AP-1 and/or NFAT (Fig. **5**). Furthermore, three networks were generated at each time point in the same way as described above. The intersectional network was obtained by carrying out the "AND" operation to analyze the commonly related mechanism of induction of apoptosis of the cells caused by TPA at three time points (Fig. **6**). Using the "AND" operation function, networks commonly involved in multiple data sets could be extracted. The analysis also suggested that the pathways of the induction of apoptosis mediated by the Fas in the cells were further mediated by NFkB, AP-1 and/or NFAT (Fig. **6**). The transcription factor NFkB is usually considered to function as an apoptosis blocker [16]. However, if the Fas was transactivated by the NFkB in the TPA-treated HL-60 cells, apoptosis could be induced by the Fas gene, mediated by the NFkB in the cells. In fact, mRNA of the Fas was induced by the NFkB p50-p65 in phorbol myristate acetate plus inomycin-treated Jarkat T-cells [17]. If researchers set up a working hypothesis using KeyMolnet as described above, they would be able to confirm it as the next step by further experiments, and to discuss the results obtained using KeyMolnet.

Other, different pathways were suggested by carrying out the same procedure. They are mediated by 1) NFkB→TNFa→TNF receptor→caspase-8→caspase-3, 2) PI3K→Rac1→MKK→p38, 3) NFkB→TRAF2→TNF receptor1→MKK4→p38, 4) AP-1 →TNFa →TNF receptor1→MKK4→p38, 5) NFkB →c-Myc, 6) ERa →TGFb, 7) IL-2 receptor, 8) IFNg receptor. Among the pathways described above, some had been partly confirmed

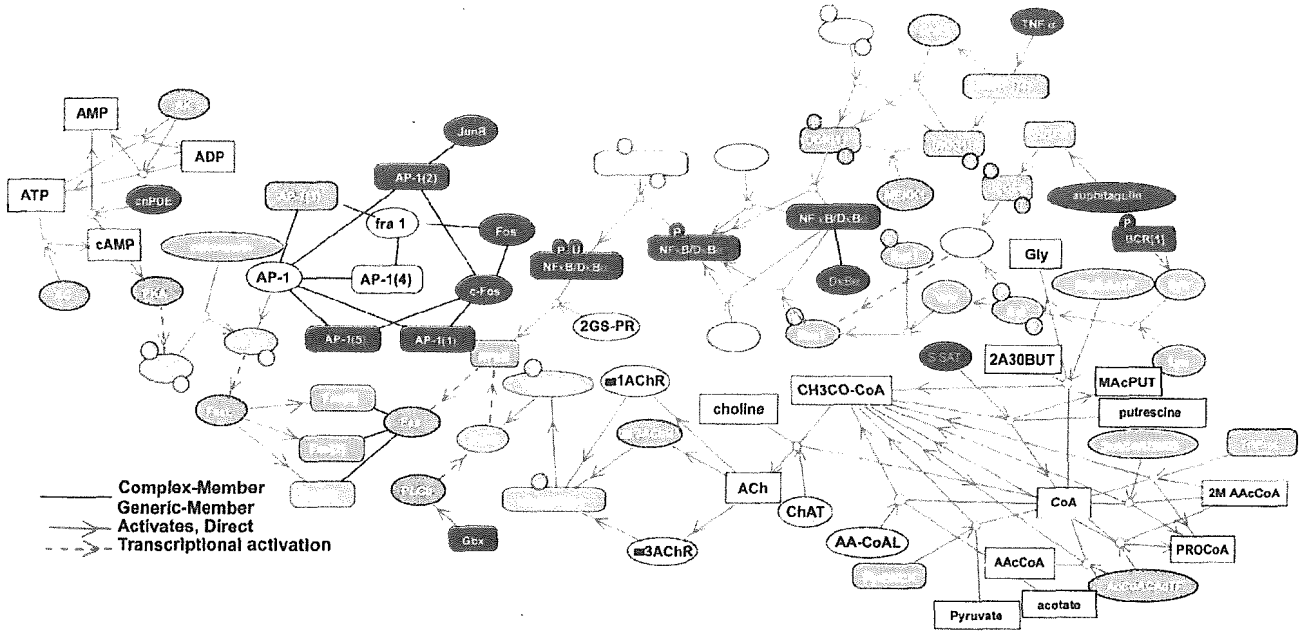

Fig. (5). Suggested pathway of induction of apoptosis mediated by Fas in HL-60 cells at 1 hour after treatment with TPA
The network was generated based on the up-regulated molecules in the cells at 1hr after treatment with TPA, using the "N-points to N-points search" (see "MATERIALS AND METHODS"). The up-regulated molecules were the molecules with 2-fold or higher gene expression levels in the cells at 1 hr after treatment with TPA compared to those in the controls. The starting points were selected as the up-regulated molecules and the end point was selected as the "Fas." The fold change values displayed on "Molecules" in the networks are used for the ratios of the gene expression levels in the cells at 1 hr after treatment with TPA compared to those in the controls. Each color indicates the same as in Figures **1**, **2**, **3** and **4**.

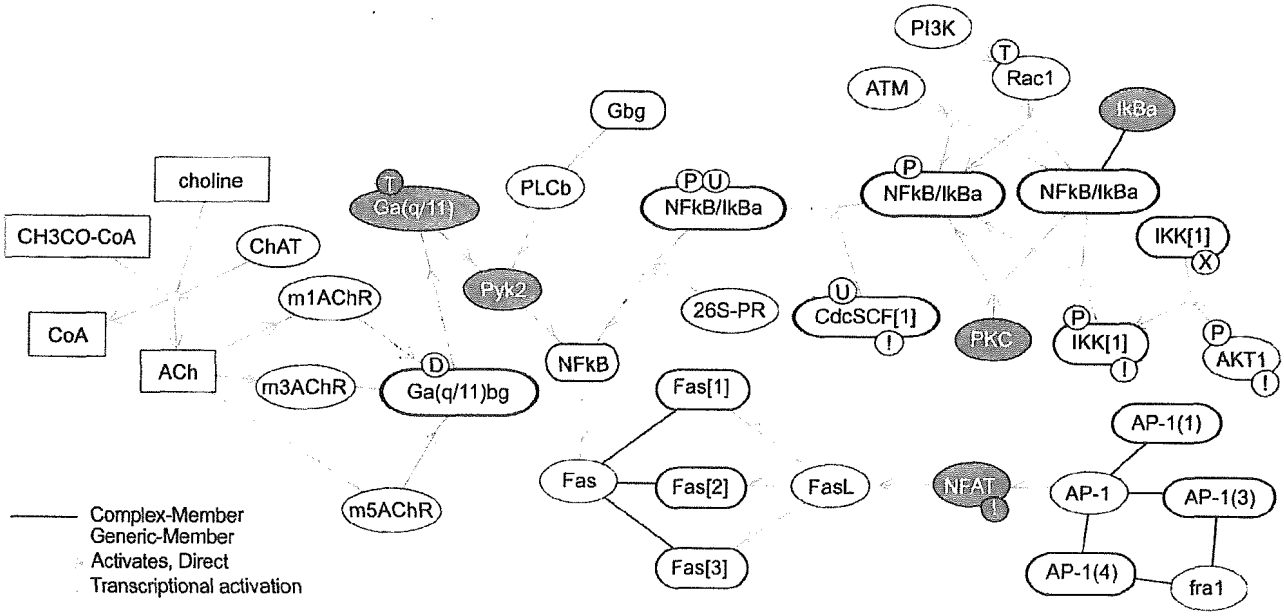

Fig. (6). Suggested network of induction of apoptosis mediated by Fas in HL-60 cells at 1, 3 and 9 hrs after treatment of TPA. After three networks were generated in the same way as shown in Figure **5** at each time point, the intersectional networks were extracted from the three time-series networks using the "AND" logical operation function. The "Molecules" shown in black are the ones with 2-fold or highergene expression levels in the cells at 9 hrs after treatment with TPA compared to those in the controls, and the black "Molecules" were also selected as the starting points when the network was generated based on the up-regulated molecules in the cells at 9 hrs after treatment with TPA.

by previous reports. For example, in myocardial dysfunction, the increase in TNFa induced by the activation of NFkB was suggested to induce apoptosis [18]. Takada *et al.* suggested that apoptosis was mediated by TNFa in TPA-treated human monocytic U937 cells, and that TPA induced apoptosis through the production of TNFa not mediated by protein kinase C in cells deficient in p53 or with a mutated p53 such as U937 cells [19]. In this way, the possible pathways of the induction of apoptosis in the cells treated with TPA were suggested by gene expression data using KeyMolnet. Further experiments are needed to confirm the suggestions of analysis.

CONCLUSION

By using KeyMolnet, complicated networks and large quantities of data could be organized properly and analyzed effectively to help scientists discuss, speculate on and uncover new avenues of research. Therefore, it is suggested to be a powerful tool for high-level research in drug discovery, and the medical and life sciences in the post-genome era.

* As the "content database" is updated and new functions are added, the specifications etc. of the current version of KeyMolnet will not always remain as shown in this paper.

ABBREVIATIONS

AP-1	=	Activator protein 1
APC	=	Adenomatous polyposis coli protein
c-Myc	=	Myc proto-oncogene protein
ERa	=	Eestrogen receptor alpha
GR	=	Glucocorticoid receptor
GO	=	Gene Ontology
GOA	=	Gene Ontology Annotation
IFNa	=	Interferon alpha
IFNb	=	Interferon beta
IFNg	=	Interferon gamma
IFNGR1	=	Interferon-gamma receptor 1 chain gene
IkBa	=	NF-kappaB inhibitor alpha
IL-1	=	Interleukin-1
IL-2	=	Interleukin-2
IMMD	=	Institute of Medicinal Molecular Design
JNK	=	c-Jun N-terminal kinase
MKK	=	MAP kinase kinase
MKK4	=	MAP kinase kinase 4
NFAT	=	Nuclear factor of activated T-cells
NFkB	=	Nuclear factor kappa B
p38	=	Mitogen-activated protein kinase p38
PI3K	=	Phosphatidylinositol 3 kinase
PPARa	=	Peroxisome proliferator activated receptor alpha
TGFb	=	Transforming growth factor beta
TNFa	=	Tumor necrosis factor alpha
TPA	=	12-O-tetradecanoylphorbol 13-acetate
TRAF2	=	TNF receptor associated factor 2
TRE	=	TPA-response element

REFERENCES

[1] Alizadeh A.A., Eisen M.B., Davis R.E., Ma C., Lossos I.S., Rosenwald A., Boldrick J.C., Sabe H., Tran T., Yu X., Powell J.I., Yang L., Marti G.E., Moore T., Hudson J. Jr., Lu L., Lewis D.B., Tibshirani R., Sherlock G., Chan W.C., Greiner T.C., Weisenburger D.D., Armitage J.O., Warnke R., Levy R., Wilson W., Grever M.R., Byrd J.C., Botstein, D., Brown P.O., Staudt L.M.: *Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling.* Nature 403(6769), 503-511, (2000).

[2] http://www.expasy.org/sprot/

[3] http://www.geneontology.org/

[4] Collins S.J: *The HL-60 promyelocytic leukemia cell line: proliferation, differentiation, and cellular oncogene expression.* Blood 70(5), 1233-1244, (1987).

[5] Rovera G., O'Brien T.G., Diamond L.: *Induction of differentiation in human promyelocytic leukemia cells by tumor promoters.* Science 204(4395), 868-870, (1979).

[6] Nguyen H.Q., Hoffman-Liebermann B., Liebermann D.A.: *The zinc finger transcription factor Egr-1 is essential for and restricts differentiation along the macrophage lineage.* Cell 72(2), 197-209, (1993).

[7] Solary E., Bertrand R., Pommier Y.: *Apoptosis of human leukemic HL-60 cells induced to differentiate by phorbol ester treatment.* Leukemia 8(5), 792-797, (1994).

[8] Patent application: PCT/JP01/07830, PCT/JP01/09940, PCT/JP02/08368, PCT/JP03/02847

[9] http://www.ebi.ac.uk/GOA/

[10] Toda K., Ishida S., Nakata K., Matsuda R., Shigemoto-Mogami Y., Fujishita K., Ozawa S., Sawada J., Inoue K., Shudo K., Hayashi Y.: *Test of significant differences with a priori probability in microarray experiments.* Anal. Sci. 19(11), 1-7, (2003).

[11] Toda K., Ishida S., Nakata K., Matsuda R., Shigemoto-Mogami Y., Ozawa S., Sawada J., Ohno Y., Inoue K., Shudo K., Hayashi Y.: *Improvement in reliability of probabilistic test of significant differences in GeneChip experiments.* Anal. Sci. 20(4), 731-733, (2004).

[12] Hayashi Y., Matsuda R.: *Deductive prediction of measurement precision from signal and noise in liquid chromatography.* Anal.Chem. 66, 2874-2881, (1994).

[13] Bienz M., Clevers H.: *Linking colorectal cancer to Wnt signaling.* Cell 103(2), 311-320, (2000).

[14] Sakamoto S., Taniguchi T.: *Identification of a phorbol ester-responsive element in the interferon-gamma receptor 1 chain gene.* J. Biol. Chem. 276(40), 37237-37241, (2001).

[15] Itoh N., Yonehara S., Ishii A., Yonehara M., Mizushima S.I., Sameshima M., Hase A., Seto Y., Nagata S.: *The polypeptide encoded by the cDNA for human cell surface antigen Fas can mediate apoptosis.* Cell 66(2), 233-243, (1991).

[16] Gilmore T.D.: *Multiple mutations contribute to the oncogenicity of the retroviral oncoprotein v-Rel.* Oncogene 18(49), 6925-6937, (1999).

[17] Chan H., Bartos D.P., Owen-Schaub L.B.: *Activation-dependent transcriptional regulation of the human Fas promoter requires NF-kappaB p50-p65 recruitment.* Mol. Cell Biol. 19(3), 2098-2108, (1999).

[18] Meldrum D.R., Shenkar R., Sheridan B.C., Cain B.S., Abraham E., Harken A.H.: *Hemorrhage activates myocardial NFkappaB and increases TNF-alpha in the heart.* J. Mol. Cell. Cardiol. 29(10), 2849-2854, (1997).

[19] Takada Y., Hachiya M., Osawa Y., Hasegawa Y., Ando K., Kobayashi Y., Akashi M.: *12-O-tetradecanoylphorbol-13-acetate-induced apoptosis is mediated by tumor necrosis factor alpha in human monocytic U937 cells.* J. Biol. Chem. 274(40), 28286-28292, (1999).