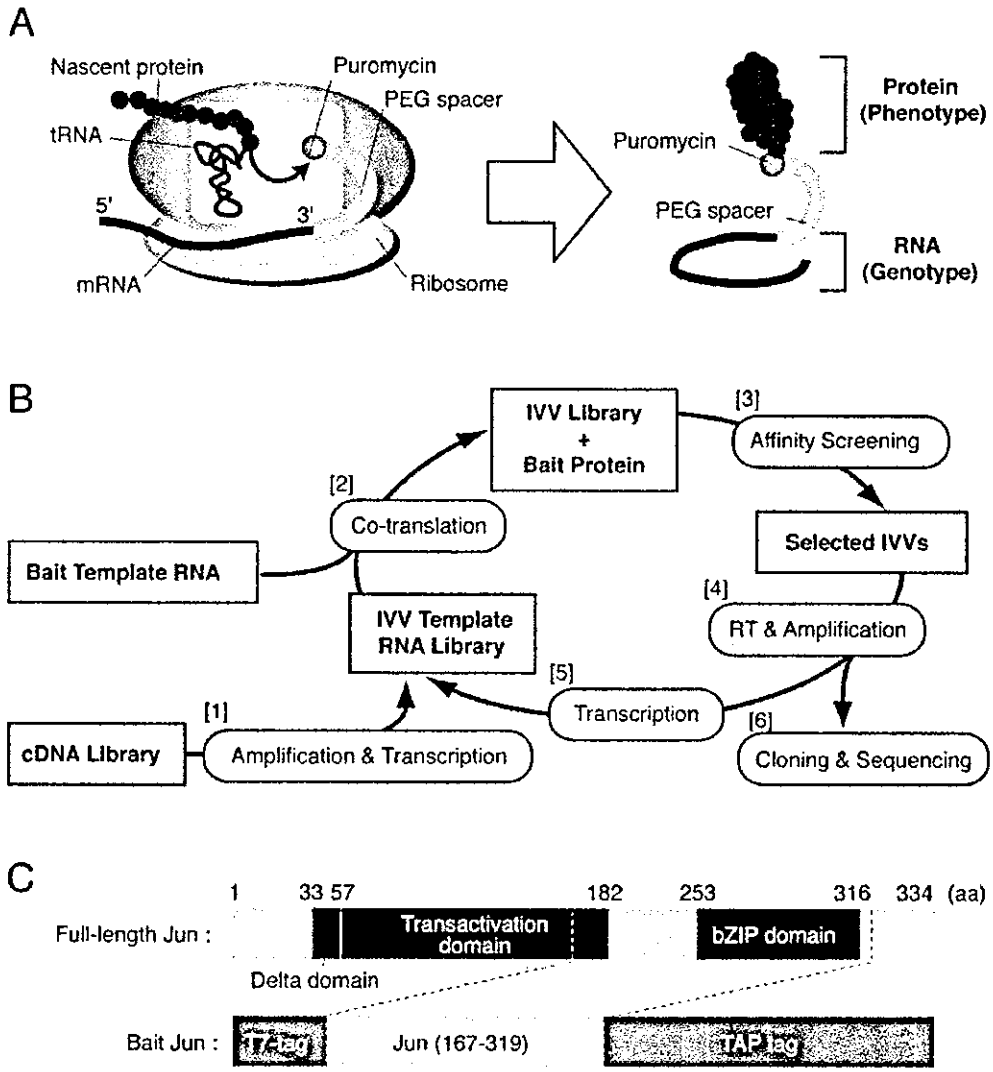

Figure 1. IVV selection of Jun-associated proteins. (**A**) Principle of IVV formation on the ribosome in a cell-free translation system (8). Puromycin ligated to the 3′-terminal end of mRNA through the polyethyleneglycol (PEG) spacer (16) can enter the ribosomal A site to bind covalently to the C-terminal end of the protein that it encodes. (**B**) Schematic representation of iterative selection for protein–protein interactions using IVV. [1] A cDNA library encoding various proteins is PCR-amplified, transcribed and ligated with a PEG spacer having puromycin. [2] The IVV template RNA library and bait template RNA are co-translated in a cell-free translation system. [3] The complex of bait protein and IVV library are subjected to affinity selection. [4] The RNA portion of the bound IVVs is reverse-transcribed and PCR-amplified. [5] The RT–PCR product is subjected to the next round of selection or [6] identified by cloning and sequencing. (**C**) Construction of the bait Jun for the selection. Jun protein has three conserved domains, delta domain, transactivation domain and bZIP domain. As a bait, the fragment containing the bZIP domain of Jun was fused with a T7-tag for confirmation of expression of the bait protein and with the TAP affinity selection tag, which contains the IgG binding domain of protein A, TEV protease cleavage site and calmodulin binding peptide (1).

proteins. Of the 143 clones, 7 were further removed, because these clones corresponded to 5′-untranslated region (5′-UTR) or 3′-UTR in mRNA sequences. Consequently, a total of 81 clones (37% of total clones) were eliminated from the obtained 217 clones as false positives. Surviving clones from the above examinations were further characterized.

The remaining 136 clones were clustered into 20 distinct sequence groups by the CLUSTALW algorithm (Table 1). Ten of the clusters consist of siblings, and the others consist of single clones (Table 1). BLASTN search revealed that 16 of the clusters involved known proteins. The other four proteins, 4732436F15Rik, 9130229H14Rik, 1200008A14Rik and B130050I23Rik, are hypothetical proteins, which have been reported in the full-length cDNA sequencing and functional annotation project 'FANTOM2' (23). Characterization of the amino acid sequences revealed that 14 of the 20 proteins contain leucine heptad repeats, which have the potential to form a leucine zipper motif. Four of the 20 proteins, Fos, Jun, Atf4 and Jdp2, have already been reported to interact with Jun directly (24), but the other 16 have not. Thus, we further examined their specific interactions with Jun by means of real-time PCR, *in vitro* pull-down assays and co-immunoprecipitation assays.

Table 1. A total of 20 selected proteins from IVV selection and sequence analyses

Gene symbol	Accession no.	Number of clones	Locus on mRNA sequence (base)	Previous report	Leucine heptad repeats	*In vitro* pull-down assay[a]
SNAP19	NM_183316.1	78	1...285	Unknown	Y	++
Kif5C (region C)	NM_008449.1	17	2473...2672	Unknown	Y	+
Kif5A (region C)	NM_008447.2	5	2654...2851	Unknown	Y	+
Eef1d	NM_023240.1	5	149...522	Unknown	Y	ND[b]
Jdp2	NM_030887.2	5	481...717	Known	Y	+++
Kif5C (region N)	NM_008449.1	4	907...1115	Unknown	N	—
Nef3	NM_008691.1	4	1086...1251	Unknown	N	—
4732436F15Rik	XM_143418.3	3	2087...2287	Unknown	Y	++
Fos	NM_010234.2	3	493...740	Known	Y	+++
9130229H14Rik	XM_135706.3	2	96...267	Unknown	Y	++
Atf4	NM_009716.1	1	1091...1305	Known	Y	+
Mapre3	NM_133350.1	1	724...980	Unknown	N	+
Cspg6	NM_007790.2	1	2474...2689	Unknown	Y	++
Mapk8ip3	NM_013931.1	1	1413...1624	Unknown	Y	+
Jun	NM_010591.1	1	904...1036	Known	Y	++
1200008A14Rik	NM_028915.1	1	1522...1677	Unknown	Y	+
GFAP	K01347.1	1	892...1025	Unknown	N	—
B130050I23Rik	NM_153536.2	1	1151...1424	Unknown	Y	++
Kif5A (region N)	NM_008447.2	1	1427...1463	Unknown	N	—
Kif5B (region N)	NM_008448.1	1	1229...1362	Unknown	N	—

[a]Interaction level of selected proteins with bait Jun based on the result of pull-down assay: —, none; +, weak; ++, strong; +++, very strong; and ND, no data.
[b]Eef1d contains leucine heptad repeats, and significant interaction was observed in the presence of bait protein, but similar behavior was also observed in the absence of bait protein.

Quantitative analysis of selected protein clones using real-time PCR

The finally selected clones may merely contain RNA that is abundant in the initial library, such as β-actin. Such negative clones can be distinguished from positive clones that are expected to be enriched in the bait (+) selection, but not enriched in the bait (−) selection. Thus, we used real-time PCR analysis to determine the amounts of the DNA molecules encoding the selected proteins in the DNA libraries from each round of the selection. Of the 20 candidates, 19 selected proteins were enriched in each round in the presence of Jun bait. The enrichment rates were between 80- and 2.0×10^4-fold, while β-actin (negative control) was not enriched. In contrast, none of the selected proteins, nor β-actin, was enriched in the bait (−) selection. These results support the conclusion that the 19 selected proteins were specifically enriched by interacting directly or indirectly with Jun bait protein. In the initial library and in the 5th bait (+), library the 19 proteins accounted for less than 0.1% and over 50% of the total, respectively. The region N of Kif5A with a chain length of 38 bp was not analyzed, because the clone was too short to perform real-time PCR analysis.

Verification of protein–protein interactions *in vitro* by pull-down assay

To determine whether the selected proteins have the ability to interact directly with Jun, *in vitro* pull-down assay was performed. A C-terminal-specific fluorescence labeling technique, which is a simple and convenient method (25), was employed for the pull-down assay. As shown in Table 1, 14 proteins including 4 known positives, Jun, Fos, Atf4 and Jdp2, exhibited direct interactions with Jun bait protein *in vitro*. All of the 14 proteins, except for Mapre3, contain leucine heptad repeats. The other proteins, Kif5A (region N), Kif5B (region N), Kif5C (region N), Nef3 and GFAP, except for Eef1d (see Table 1) neither interacted nor contained leucine heptad repeats. We considered that these proteins might interact with Jun indirectly via other Jun-associated proteins, because there are some findings indicating interactions between selected proteins. For example, regions C and N of Kif5 family (Kif5s) proteins (Table 1) are known to interact in a single molecule, generating a compact structure to control the motor activity of the Kif5s (26–28). The region C clones have leucine heptad repeats and interacted directly with Jun *in vitro*; thus, region N fragments of Kif5s might interact with bait Jun via region C fragments of Kif5s in this selection. Also, Nef3 are known to interact with Kif5A (29), and selected regions of Nef3 are highly homologous to parts of GFAP. Consequently, all five proteins for which direct interactions with Jun were not confirmed by pull-down assay, in spite of the specific enrichment confirmed by real-time PCR, may interact indirectly with Jun through other positive clones.

Verification of protein–protein interactions in cultured cells

Transfected Jun protein was assayed by co-immunoprecipitation with 10 GFP- and FLAG-tagged selected proteins which exhibited interaction with Jun *in vitro* (Figure 2). All 10 selected proteins were immunoprecipitated (Figure 2B, upper panels), while Jun was co-immunoprecipitated with only 6 of the 10 proteins, SNAP19, Cspg6, 9130229H14Rik, 1200008A14Rik, B130050I23Rik and 4732436F15Rik (Figure 2B, lower panels). The other four proteins, Kif5A (region C), Kif5C (region C), Mapk8ip3 and Mapre3, apparently did not interact with Jun.

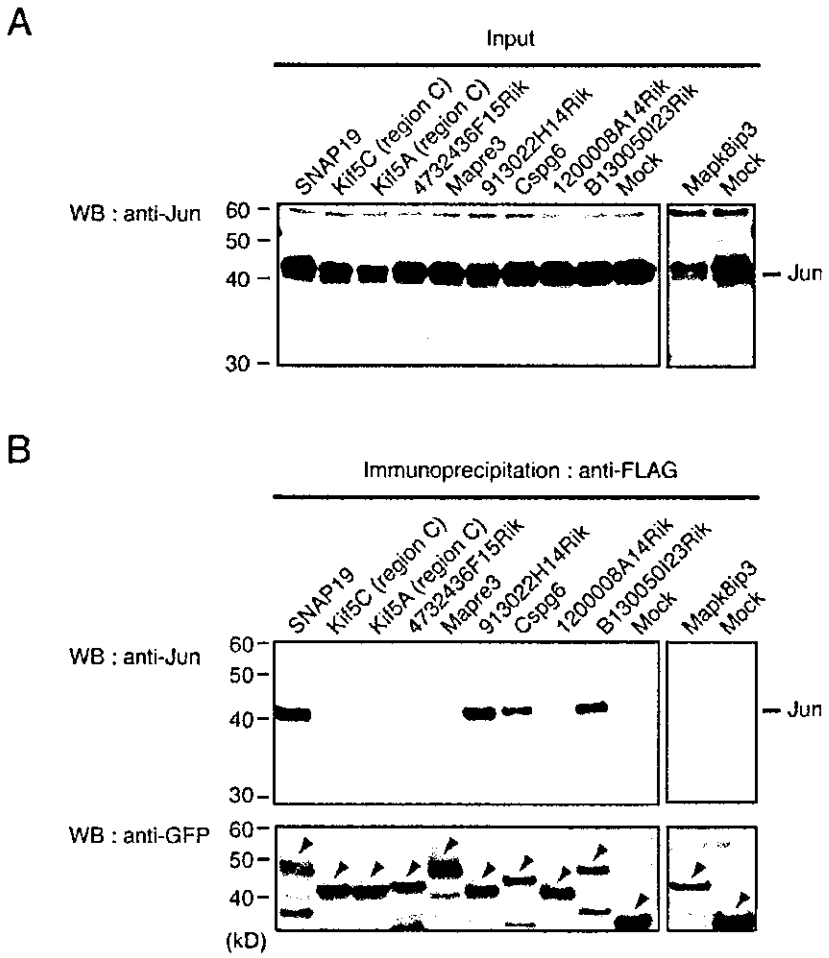

Figure 2. Co-immunoprecipitation assay between Jun and selected Jun-associated protein candidates. (A) Western-blot analysis, showing the expression levels of Jun after co-transfection with Jun and each of 10 selected proteins. 'Mock' was co-transfected with Jun and GFP-HL4-FLAG. (B) Extracts of COS7 cells which were transfected with Jun and each of the ten selected proteins were immunoprecipitated with anti-FLAG antibody. This was followed by western-blot analysis using the anti-Jun antibody to detect co-immunoprecipitated Jun protein (top) and anti-GFP antibody to detect GFP-fused selected proteins (arrowhead on the bottom).

Determination of subcellular localization of selected protein fragments

To elucidate why the interaction of some proteins that interacted with Jun *in vitro* could not be confirmed by co-immunoprecipitation assay, we observed the subcellular localization of 10 GFP-tagged selected protein fragments in COS7 cells (Figure 3). As a control, a GFP-tagged selected Fos protein fragment and GFP-HL4-FLAG protein alone (mock) were also transfected. All of the protein fragments that were co-immunoprecipitated with Jun, SNAP19, 4732436F15Rik, 913022H14Rik, Cspg6 and B130050I23Rik, except for 1200008A14Rik, were located mostly in the nucleus, like Fos. On the other hand, the proteins that did not co-immunoprecipitate Jun, Kif5C (region C), Kif5A (region C), Mapre3 and Mapk8ip3, were located mostly in cytoplasm. Mock protein was located ubiquitously in the cells. These results imply that differences in the subcellular localization of the transfected protein fragments could explain the result of the co-immunoprecipitation assay.

Prediction of cellular function

In order to elucidate the cellular roles of the selected proteins, functional annotations of these proteins in public databases were searched from the entries in GO and Refseq. As shown in Figure 4, the 13 unreported proteins and 4 known positives were clustered into five functional groups, microtubule association (Kif5A, Kif5B, Kif5C and Mapre3), kinesin complex (GFAP, Nef3, Kif5A, Kif5B, Kif5C, Mapk8ip3 and Cspg6), chromosome segregation (Cspg6, 9130229H14Rik and 1200008A14Rik), DNA repair (Mapk8ip3, Cspg6 and 9130229H14Rik), and transcriptional regulation (SNAP19, Fos, Jdp2, Atf4 and Jun). Two hypothetical proteins, B130050I23Rik and 4732436F15Rik, have no clear functional annotations (black-bordered boxes in Figure 4).

DISCUSSION

Jun protein is a eukaryotic transcription factor, which plays an important role in a variety of cellular functions, including

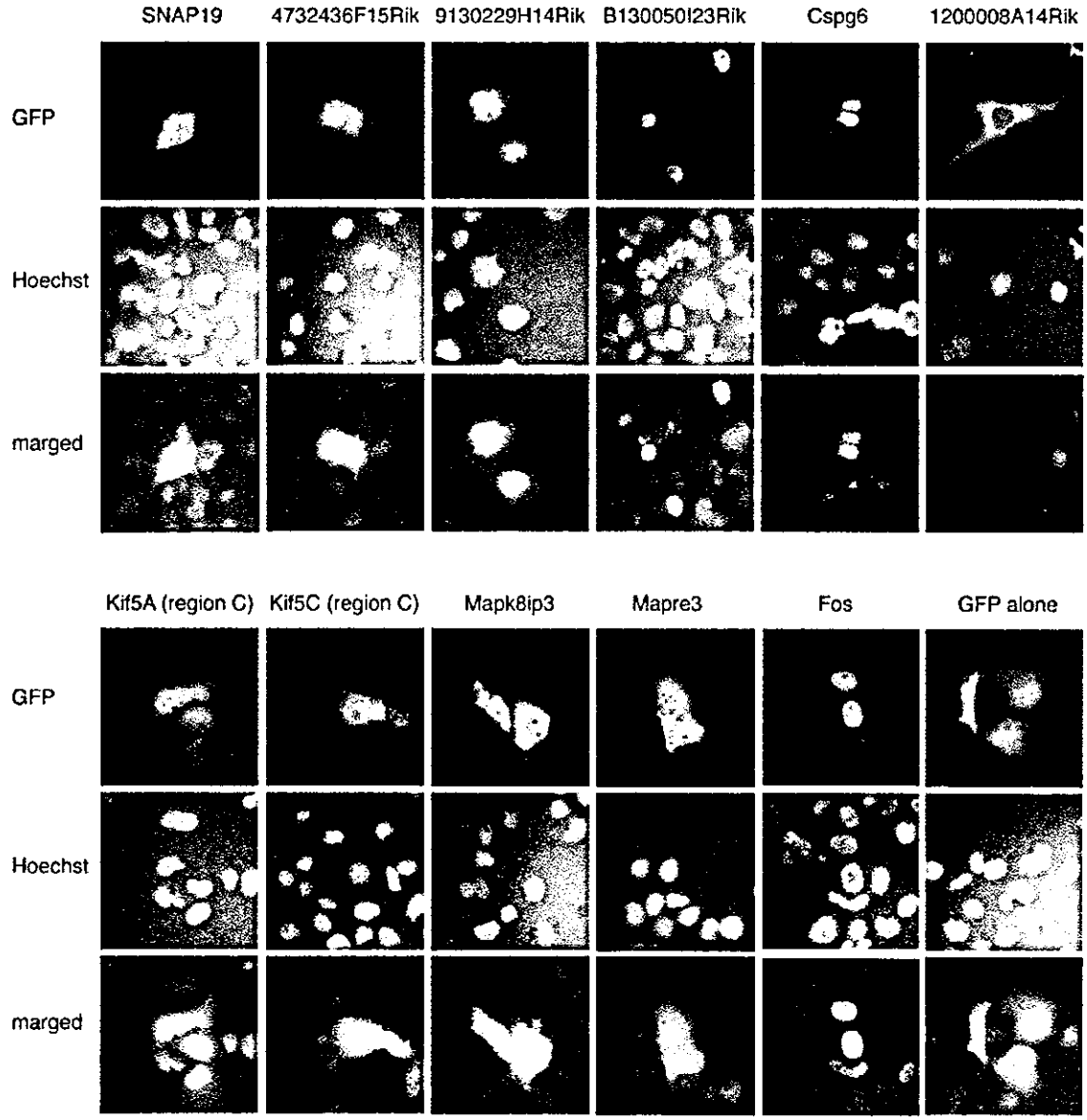

Figure 3. Subcellular localization of the 10 selected protein fragments. COS7 cells were transfected with GFP-tagged selected protein fragments that interacted with Jun *in vitro*, GFP-tagged selected Fos fragment, and GFP-HL4-FLAG alone (negative control). The Fos fragment was used as a positive control of co-localization with Jun, because it was efficiently co-immunoprecipitated with Jun. Upper panels show GFP-tagged proteins. Middle panels show the nucleus stained with Hoechst33342. Lower panels show merged images. GFP and Hoechst33342 appear as green and blue, respectively.

proliferation, differentiation and tumorigenesis (30). The cellular functions of Jun vary according to the interacting partners (24). So far, over 50 Jun-associated proteins have been found in various tissues by using biochemical methods, Y2H assay, and other techniques (24). Jun forms a homodimer and heterodimers with Fos/Jun family proteins such as Fos, Fra1, Fra2, FosB, Jun, JunB and JunD (31–34), and with other bZIP family proteins (24). Jun also interacts directly with many proteins, such as transcriptional co-activators, structurally unrelated DNA binding proteins and nuclear structural components (24). Most of the known Jun-associated proteins are transcriptional regulators.

In this study, we were able to select 20 candidate Jun-associated proteins from a mouse brain cDNA library by using the IVV selection system. Of the 20 candidates, 16 are previously unreported interactions. Of the 16 selected proteins, 10 proteins were confirmed to interact directly with Jun bait protein *in vitro* (Table 1). All 10 proteins, except for Mapre3, contain leucine heptad repeats. This result seems reasonable, because almost all Jun-associated proteins that interact with the bZIP domain of Jun have a leucine zipper motif, which is essential to form heterodimers with Jun (24). Furthermore, 6 of the 10 candidates, SNAP19, Cspg6, 9130229H14Rik, 1200008A14Rik, B130050I23Rik and

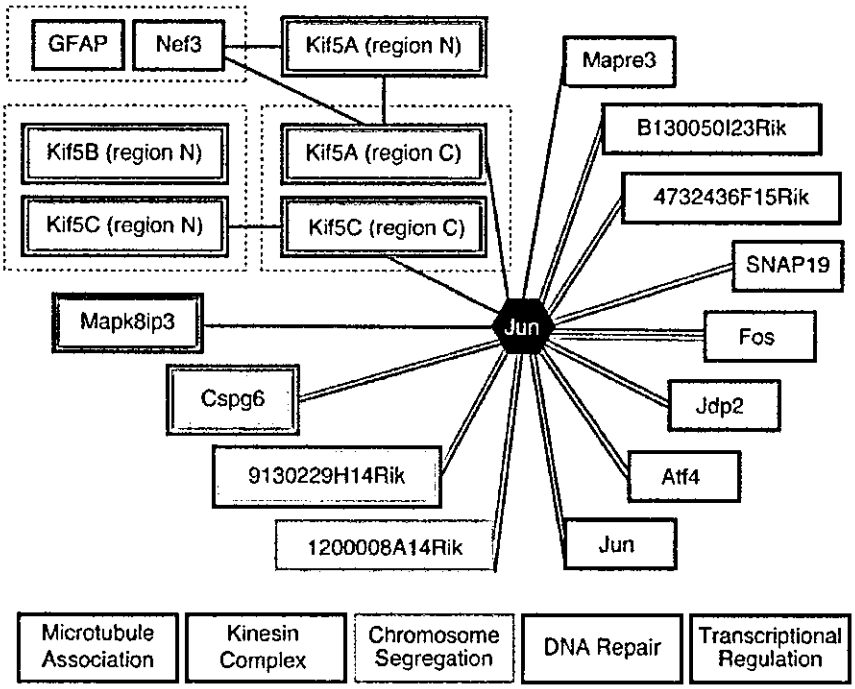

Figure 4. Protein–protein interaction mapping of the selected proteins with functional annotations. Protein–protein interactions confirmed by co-immunoprecipitation and pull-down assay are represented as magenta and blue lines, respectively. Black lines indicate previously reported interactions. Color-bordered boxes represent the functional annotations of the proteins as indicated in the lower panels. Two hypothetical proteins (black-bordered boxes) have no clear functional annotations. Broken-lined squares represent highly homologous protein fragment pairs.

4732436F15Rik, were confirmed to interact with Jun in a cultured cell line by using co-immunoprecipitation assay (Figure 2). We predicted the cellular functions of these proteins by means of functional clustering using annotations.

SNAP19, the most abundant protein (Table 1; 57% of the positives) in this selection, was clustered into the functional group of transcriptional regulation (Figure 4). SNAP19 is known as a 19 kDa subunit of a small nuclear RNA-activating protein complex (35). Previous reports indicate that SNAP19 works in the nucleus as a subunit of the transcriptional regulator protein complex (35). Although no functional or physical relationship between SNAP19 and Jun has been reported previously, our present findings strongly suggest an interaction between SNAP19 and Jun in living cells.

Cspg6, 9130229H14Rik and 1200008A14Rik were clustered into two functional groups, DNA repair and chromosome segregation. This suggests that Jun plays unreported roles in these functions (Figure 4). There are some examples of transcription factors with other non-transcriptional cellular functions; for example, a bZIP transcription factor controls the cell cycle by interacting directly with Cdk2 protein without the involvement of any transcriptional event (36), and an unexpected interaction of Jun with cytoskeletal materials has been reported recently (37). Otherwise our selected proteins may work cooperatively with Jun as transcriptional regulators. For example, Mmip1 protein, an isoform of Cspg6, was reported as a transcriptional suppressor interacting with Mad proteins, a bHLH-ZIP transcriptional regulator family (38,39), implying that Cspg6 protein may also suppress the transcriptional activity of Jun.

Although B130050I23Rik and 4732436F15Rik have no clear annotation, their nuclear localization predicted by the PSORTII program (http://psort.nibb.ac.jp) implies interaction with Jun in the nucleus *in vivo*, and possible functions related to transcriptional regulation. Further *in vivo* experiments are necessary to clarify these proteins' cellular functions.

We could not confirm the interactions between Jun and the other four candidates, Kif5A (region C), Kif5C (region C), Mapk8ip3 and Mapre3, by means of co-immunoprecipitation assay (Figure 2), in spite of the *in vitro* interactions. A possible reason for this would be a difference of subcellular localization between these proteins and Jun. Indeed, all of the proteins that were confirmed to interact with Jun in cultured cells, except for 1200008A14Rik, were located mostly in the nucleus, like Fos protein, a well-known Jun-interactor, while the above-mentioned four proteins were located mostly in cytoplasm (Figure 3), and are known to interact with the cytoskeleton in living cells (40–42). Although the *in vitro* interactions of these four proteins may be biological false positives, biologically significant interaction with Jun cannot be ruled out, because subcellular localizations of some proteins are known to be tightly restricted to a specific phase *in vivo*. For example, Kif17, a closely related paralogous protein of Kif5 family proteins, is located in the nucleus and interacts with ACT protein, a transcriptional co-activator, only at the specific stage of spermatogenesis *in vivo*, and the interaction was only confirmed by immunostaining analysis of mouse testis tissue (43).

So far, genome-wide analyses of protein–protein interactions have been performed by only Y2H and biochemical

methods using MS. The IVV selection system exemplified in this study has several advantages over previous techniques, for example, convenient removability of false positives arising from the selection system itself, availability of a wider range of interacting conditions and availability for analysis of the interactions of toxic proteins. Every current technique for screening of protein–protein interactions has some disadvantages as well as advantages, and therefore the use of a range of different techniques is important to obtain as complete a map as possible of protein–protein interactions in various organisms.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

ACKNOWLEDGEMENTS

We thank Dr Takashi Takeuchi for experimental advice. This work was supported in part by a Grant-in-Aid for the 21st Century Center of Excellence (COE) Program entitled 'Understanding and Control of Life's Function via Systems Biology (Keio University)' and a Special Coordination Fund of the Ministry of Education, Culture, Sports, Science and Technology, Japan.

REFERENCES

1. Puig,O., Caspary,F., Rigaut,G., Rutz,B., Bouveret,E., Bragado-Nilsson,E., Wilm,M. and Seraphin,B. (2001) The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods*, **24**, 218–229.
2. Fields,S. and Song,O. (1989) A novel genetic system to detect protein–protein interactions. *Nature*, **340**, 245–246.
3. Vidalain,P.O., Boxem,M., Ge,H., Li,S. and Vidal,M. (2004) Increasing specificity in high-throughput yeast two-hybrid experiments. *Methods*, **32**, 363–370.
4. Smith,G.P. (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science*, **228**, 1315–1317.
5. Bradbury,A.R. and Marks,J.D. (2004) Antibodies from phage antibody libraries. *J. Immunol. Methods*, **290**, 29–49.
6. Hanes,J. and Pluckthun,A. (1997) *In vitro* selection and evolution of functional proteins by using ribosome display. *Proc. Natl Acad. Sci. USA*, **94**, 4937–4942.
7. Mattheakis,L.C., Bhatt,R.R. and Dower,W.J. (1994) An *in vitro* polysome display system for identifying ligands from very large peptide libraries. *Proc. Natl Acad. Sci. USA*, **91**, 9022–9026.
8. Nemoto,N., Miyamoto-Sato,E., Husimi,Y. and Yanagawa,H. (1997) *In vitro* virus: bonding of mRNA bearing puromycin at the 3′-terminal end to the C-terminal end of its encoded protein on the ribosome *in vitro*. *FEBS Lett.*, **414**, 405–408.
9. Miyamoto-Sato,E., Nemoto,N., Kobayashi,K. and Yanagawa,H. (2000) Specific bonding of puromycin to full-length protein at the C-terminus. *Nucleic Acids Res.*, **28**, 1176–1182.
10. Roberts,R.W. and Szostak,J.W. (1997) RNA–peptide fusions for the *in vitro* selection of peptides and proteins. *Proc. Natl Acad. Sci. USA*, **94**, 12297–12302.
11. Yonezawa,M., Doi,N., Kawahashi,Y., Higashinakagawa,T. and Yanagawa,H. (2003) DNA display for *in vitro* selection of diverse peptide libraries. *Nucleic Acids Res.*, **31**, e118.
12. Takahashi,T.T., Austin,R.J. and Roberts,R.W. (2003) mRNA display: ligand discovery, interaction analysis and beyond. *Trends Biochem. Sci.*, **28**, 159–165.
13. Lipovsek,D. and Pluckthun,A. (2004) *In-vitro* protein evolution by ribosome display and mRNA display. *J. Immunol. Methods*, **290**, 51–67.
14. Hammond,P.W., Alpin,J., Rise,C.E., Wright,M. and Kreider,B.L. (2001) *In vitro* selection and characterization of Bcl-X(L)-binding proteins from a mix of tissue-specific mRNA display libraries. *J. Biol. Chem.*, **276**, 20898–20906.
15. Cujec,T.P., Medeiros,P.F., Hammond,P., Rise,C. and Kreider,B.L. (2002) Selection of v-abl tyrosine kinase substrate sequences from randomized peptide and cellular proteomic libraries using mRNA display. *Chem. Biol.*, **9**, 253–264.
16. Miyamoto-Sato,E., Takashima,H., Fuse,S., Suc,K., Ishizaka,M., Tateyama,S., Horisawa,K., Sawasaki,T., Endo,Y. and Yanagawa,H. (2003) Highly stable and efficient mRNA templates for mRNA–protein fusions and C-terminally labeled proteins. *Nucleic Acids Res.*, **31**, e78.
17. Sawasaki,T., Ogasawara,T., Morishita,R. and Endo,Y. (2002) A cell-free protein synthesis system for high-throughput proteomics. *Proc. Natl Acad. Sci. USA*, **99**, 14652–14657.
18. Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403–410.
19. Thompson,J.D., Higgins,D.G. and Gibson,T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**, 4673–4680.
20. Arai,R., Ueda,H., Kitayama,A., Kamiya,N. and Nagamune,T. (2001) Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Eng.*, **14**, 529–532.
21. Kanagawa,T. (2003) Bias and artifacts in multitemplate polymerase chain reactions (PCR). *J. Biosci. Bioeng.*, **96**, 317–323.
22. Hertveldt,K., Dechassa,M.L., Robben,J. and Volckaert,G. (2003) Identification of Gal80p-interacting proteins by *Saccharomyces cerevisiae* whole genome phage display. *Gene*, **307**, 141–149.
23. Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Nikaido,I., Osato,N., Saito,R., Suzuki,H. *et al.* (2002) Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature*, **420**, 563–573.
24. Chinenov,Y. and Kerppola,T.K. (2001) Close encounters of many kinds: Fos–Jun interactions that mediate transcription regulatory specificity. *Oncogene*, **20**, 2438–2452.
25. Doi,N., Takashima,H., Kinjo,M., Sakata,K., Kawahashi,Y., Oishi,Y., Oyama,R., Miyamoto-Sato,E., Sawasaki,T., Endo,Y. *et al.* (2002) Novel fluorescence labeling and high-throughput assay technologies for *in vitro* analysis of protein interactions. *Genome Res.*, **12**, 487–492.
26. Stock,M.F., Guerrero,J., Cobb,B., Eggers,C.T., Huang,T.G., Li,X. and Hackney,D.D. (1999) Formation of the compact confomer of kinesin requires a COOH-terminal heavy chain domain and inhibits microtubule-stimulated ATPase activity. *J. Biol. Chem.*, **274**, 14617–14623.
27. Friedman,D.S. and Vale,R.D. (1999) Single-molecule analysis of kinesin motility reveals regulation by the cargo-binding tail domain. *Nature Cell Biol.*, **1**, 293–297.
28. Seiler,S., Kirchner,J., Horn,C., Kallipolitou,A., Woehlke,G. and Schliwa,M. (2000) Cargo binding and regulatory sites in the tail of fungal conventional kinesin. *Nature Cell Biol.*, **2**, 333–338.
29. Xia,C.H., Roberts,E.A., Her,L.S., Liu,X., Williams,D.S., Cleveland,D.W. and Goldstein,L.S. (2003) Abnormal neurofilament transport caused by targeted disruption of neuronal kinesin heavy chain KIF5A. *J. Cell Biol.*, **161**, 55–66.
30. Shaulian,E. and Karin,M. (2002) AP-1 as a regulator of cell life and death. *Nature Cell Biol.*, **4**, E131–136.
31. Kovary,K. and Bravo,R. (1991) Expression of different Jun and Fos proteins during the G_0-to-G_1 transition in mouse fibroblasts: *in vitro* and *in vivo* associations. *Mol. Cell. Biol.*, **11**, 2451–2459.
32. Lallemand,D., Spyrou,G., Yaniv,M. and Pfarr,C.M. (1997) Variations in Jun and Fos protein expression and AP-1 activity in cycling, resting and stimulated fibroblasts. *Oncogene*, **14**, 819–830.
33. McCabe,L.R., Kockx,M., Lian,J., Stein,J. and Stein,G. (1995) Selective expression of fos- and jun-related genes during osteoblast proliferation and differentiation. *Exp. Cell Res.*, **218**, 255–262.
34. Sonnenberg,J.L., Macgregor-Leon,P.F., Curran,T. and Morgan,J.I. (1989) Dynamic alterations occur in the levels and composition of transcription factor AP-1 complexes after seizure. *Neuron*, **3**, 359–365.
35. Henry,R.W., Mittal,V., Ma,B., Kobayashi,R. and Hernandez,N. (1998) SNAP19 mediates the assembly of a functional core promoter complex (SNAPc) shared by RNA polymerases II and III. *Genes Dev.*, **12**, 2664–2672.
36. Izumiya,Y., Lin,S.F., Ellison,T.J., Levy,A.M., Mayeur,G.L., Izumiya,C. and Kung,H.J. (2003) Cell cycle regulation by Kaposi's

sarcoma-associated herpesvirus K-bZIP: direct interaction with cyclin-CDK2 and induction of G1 growth arrest. *J. Virol.*, **77**, 9652–9661.
37. Luettich,K. and Schmidt,C. (2003) TGFbeta1 activates c-Jun and Erk1 via alphaVbeta6 integrin. *Mol. Cancer*, **2**, 33.
38. Gupta,K., Anand,G., Yin,X., Grove,L. and Prochownik,E.V. (1998) Mmip1: a novel leucine zipper protein that reverses the suppressive effects of Mad family members on c-myc. *Oncogene*, **16**, 1149–1159.
39. Jones,S. and Sgouros,J. (2001) The cohesin complex: sequence homologies, interaction networks and shared motifs. *Genome Biol.*, **2**, RESEARCH0009.
40. Kanai,Y., Okada,Y., Tanaka,Y., Harada,A., Terada,S. and Hirokawa,N. (2000) KIF5C, a novel neuronal kinesin enriched in motor neurons. *J. Neurosci.*, **20**, 6374–6384.
41. Verhey,K.J., Meyer,D., Deehan,R., Blenis,J., Schnapp,B.J., Rapoport,T.A. and Margolis,B. (2001) Cargo of kinesin identified as JIP scaffolding proteins and associated signaling molecules. *J. Cell Biol.*, **152**, 959–970.
42. Nakagawa,H., Koyama,K., Murata,Y., Morito,M., Akiyama,T. and Nakamura,Y. (2000) EB3, a novel member of the EB1 family preferentially expressed in the central nervous system, binds to a CNS-specific APC homologue. *Oncogene*, **19**, 210–216.
43. Macho,B., Brancorsini,S., Fimia,G.M., Setou,M., Hirokawa,N. and Sassone-Corsi,P. (2002) CREM-dependent transcription in male germ cells controlled by a kinesin. *Science*, **298**, 2388–2390.

Rapid Communication

J. Biochem. 137, 121–124 (2005)
DOI: 10.1093/jb/mvi023

Application of Quantitative Real-Time PCR for Monitoring the Process of Enrichment of Clones on *In Vitro* Protein Selection

Kenichi Horisawa, Nobuhide Doi, Hideaki Takashima and Hiroshi Yanagawa*

Department of Biosciences and Informatics, Faculty of Science and Technology, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama 223-8522

Received November 29, 2004; accepted December 24, 2004

***In vitro* selection of proteins from cDNA libraries using display technologies, such as the *in vitro* virus method, is a powerful means for the discovery of novel protein interactions. After iterative screening, selected proteins are usually identified and evaluated by cloning and sequencing analysis. Previously we applied real-time PCR for evaluation of the sequences obtained on *in vitro* virus screening. Here, we have presented additional data regarding monitoring of the process of enrichment of selected clones in each round of selection and elimination of false positives by real-time PCR, and have also discussed the utility of the novel method. This approach should also be applicable to other display technologies.**

Key Words: display technology, false positive, *in vitro* virus, mRNA display, real-time PCR.

Abbreviations: IVV, *in vitro* virus; Nrbf2, nuclear receptor binding factor 2; Aes, amino-terminal enhancer of split; Gas5, growth arrest–specific 5.

Display technologies (*1–3*) that link genotype (DNA or RNA) and phenotype (protein) molecules are powerful tools for the discovery of interaction partner proteins for various targets (*e.g.*, proteins, nucleic acids, small-molecular compounds, and drug candidates), using cDNA libraries or artificial random-sequence libraries. Phage display (*4–8*) is the most widely used display technology, and has uncovered many novel functional proteins, protein–protein interactions, and DNA–protein interactions. Furthermore, totally *in vitro* display technologies involving cell-free translation systems, such as ribosome display (*9–11*), mRNA display (*12, 13*), and DNA display (*14–16*), have also been applied for the discovery of novel functional proteins, screening of drug targets (*17*), and protein–protein interaction analysis (*18*).

Figure 1 is a schematic representation of a typical screening procedure involving totally *in vitro* display technologies. Initial libraries of genotype-phenotype linking molecules are affinity-screened and amplified iteratively. The resulting DNA library is cloned into some kind of cloning vector and then sequences are determined by commonly used methods. However, we cannot easily determine much about the abundance ratio or the process of enrichment of the selected clones in each round from only the cloned numbers of the sequences, although such information would be quite useful for evaluating the clones and for optimizing the selection conditions. In addition, the resulting clones often include false positives that are merely abundant in the initial library or that are accidentally picked up in spite of having no binding activity. Therefore, we planned to apply a quantitative real-time PCR technique to monitoring of the process of enrichment of clones on *in vitro* protein selection and elimination of false positives easily by accurately determining the numbers of molecules of each selected DNA in the resulting libraries to confirm the specific enrichment of the sequences.

We recently performed affinity screening of Jun-associated proteins from a mouse brain cDNA library (*19*) using our *in vitro* virus (IVV) method (*12, 13*), one of the mRNA display technologies. In this study, we further analyzed 451 clones including 217 previously analyzed clones which had been picked up from the library on five rounds of iterative screening (*19*). Of the 451 clones, 271 (about 60% of the total analyzed clones) were confirmed to represent intact RNA-protein conjugated molecules without any frame-shift or stop codon on sequence analysis. These 271 clones were clustered into 22 independent sequence groups that were considered to be credible candidates for Jun-associated proteins (Table 1); they included three unreported candidates, nuclear receptor binding factor 2 (Nrbf2) (*20*), amino-terminal enhancer of split (Aes) (*21*), and growth arrest–specific 5 (Gas5) (*22*), together with 19 that we had previously reported (*19*).

We performed real-time PCR analysis to determine the numbers of DNA molecules of the 22 selected sequences (Table 1) in the initial cDNA library, and the libraries obtained on each round of IVV screening in the presence [bait (+)] and absence [bait (–)] of the Jun bait protein (Fig. 2). We found that all of the selected sequences except for those of Aes and Gas5 were enriched in each round of bait (+) selection (Fig. 2A), whereas β-actin (negative control) in the bait (+) selection (Fig. 2A), and all sequences in the bait (–) selection (Fig. 2B) were found not to be enriched. These results indicate that 20 of the 22 sequences had been selected specifically on the basis of affinity for the Jun bait protein, whereas Aes and Gas5 are false positives. Thus, of the three new candidates discovered here, only Nrbf2 was concluded to be a true positive. The interaction between Jun and Nrbf2 was

*To whom correspondence should be addressed. Tel: +81 45 566 1775, Fax: +81 45 566 1440, E-mail: hyana@bio.keio.ac.jp

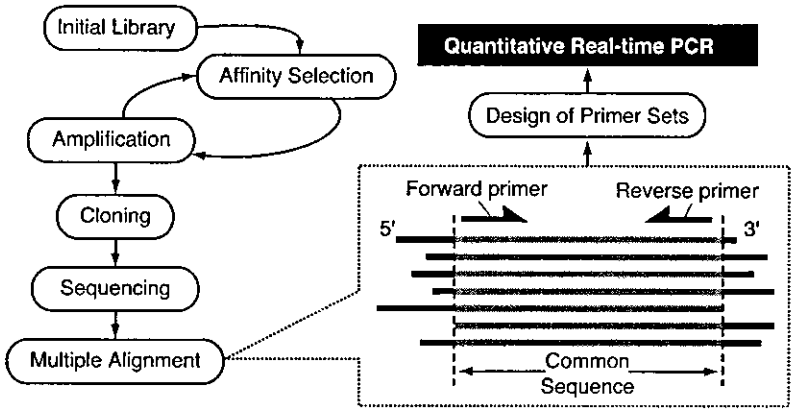

Fig. 1. **A schematic representation of screening with totally *in vitro* display technologies and quantitative real-time PCR analysis.** Initial DNA libraries were constructed by reverse-transcription from poly A (+) RNAs (cDNA library) or chemical synthesis (random-sequence library). Affinity screening and amplification of each library were performed iteratively. After the screening, the resulting library was cloned into a cloning vector and the DNA sequences were determined. The selected DNA sequences were analyzed as described previously (*19*) and were clustered by multiple alignment using the CLUSTALW program (*24*). Primer sets for real-time PCR were designed for each of the clustered sequences. In the case of the sequences of siblings, specific primers were designed based on the common region of the clone encoding each selected sequence. Real-time PCR was performed with a Lightcycler FastStart DNA master SYBR green I kit (Roche) and protein-specific primer sets (the amplicon size was 56 to 200 bp) with a LightCycler (Roche). The standard template DNA was PCR-amplified from each selected sequence on a pDrive vector (Qiagen) using primers 5'M13F (5'-GTTTTCCCAGTCACGACGTTG-3') and 3'M13R (5'-GAAACAGCTATGACCATGATTACG -3'). The numbers of DNA molecules of the selected sequences in 5 ng aliquots (~10^{10} molecules) of DNA libraries were determined.

confirmed by *in vitro* pull-down assays (data not shown), as had been done with other previously confirmed sequences (*19*). Interestingly, the process of enrichment of each selected clone varied. The highest and lowest enrichment ratios in this selection were about 2.0×10^4 and 80-fold, respectively. Furthermore, the process of enrichment of the clones based on the results of real-time PCR is informative for refining the selection protocol, in particular, the selection conditions.

Figure 3 shows the correlation between the abundance ratio of the 22 selected sequences analyzed by real-time PCR, and that found on cloning and sequencing analysis of the library after the fifth round of screening. The two abundance ratios were well correlated in most cases, but not for the Gas5 and Aes sequences (arrows in Fig. 3). Considering that the two sequences were each discovered as a single clone (Table 1), they are certainly false positives.

Real-time PCR is a PCR application that allows accurate quantification of DNA molecules. This technique is based on the PCR kinetics (*23*); the quantification of amplified DNA molecules is performed during the amplification by using intercalators such as a SYBR Green reagent, and the absolute number of DNA molecules is calculated from the results obtained with standards that are observed at the same time. The real-time PCR technique has been used for gene expression analysis and the detection of mutations. In our previous study, we introduced real-time PCR analysis to evaluate the results of *in vitro* selection. In this paper, we have presented additional data regarding monitoring of the process of enrichment of the selected sequences and elimination of false positives by means of real-time PCR. The use of real-time PCR for this purpose has several advantages over previous methods. For example, binding assays, such as *in vitro* pull-down assays and coimmunoprecipitation, are relatively

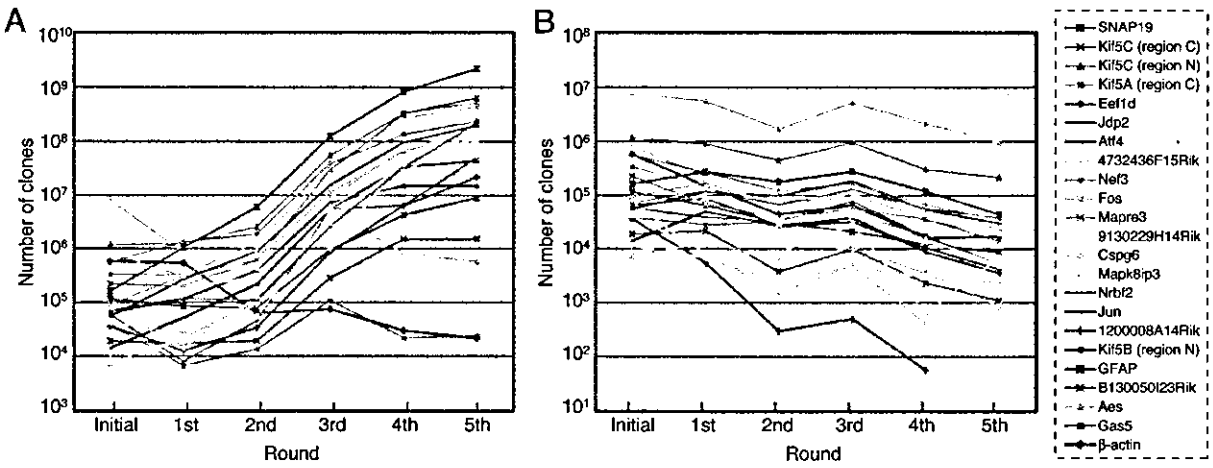

Fig. 2. **Quantitative real-time PCR analysis of the selected sequences in the library obtained on each round.** The numbers of DNA molecules of the 22 specifically selected clones in 5 ng aliquots (~10^{10} molecules) of the libraries, (A) bait (+) and (B) bait (–), obtained on each round were plotted logarithmically. β-Actin was chosen as a negative control.

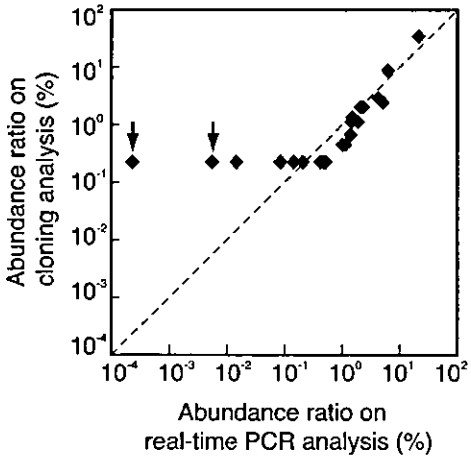

Fig. 3. **Correlation of the abundance ratio on cloning analysis and that on quantitative real-time PCR analysis.** On a logarithmic scale, the vertical axis indicates the abundance ratio on cloning analysis [(number of obtained clones/total number of analyzed clones) × 100 (%)], and the horizontal axis indicates the abundance ratio on real-time PCR analysis [(number of DNA molecules in 5 ng of a DNA library/1.0×10^{10} molecules) × 100 (%)]. The broken line indicates complete correspondence of the two abundance ratios. The arrows indicate Gas5 and Aes from left to right, respectively.

Table 1. **Candidates for Jun-associated proteins selected on IVV screening.**

Gene symbol	Number of clones
SNAP19	155
Kif5C (region C)	39
Kif5A (region C)	13
Kif5C (region N)	11
Eef1d	9
Jdp2	9
Nef3	6
4732436F15Rik	6
Fos	5
Mapre3	3
9130229H14Rik	2
Atf4	2
Cspg6	2
Mapk8ip3	1
Nrbf2	1
Jun	1
1200008A14Rik	1
Kif5B (region N)	1
GFAP	1
B130050I23Rik	1
Aes	1
Gas5	1

tedious to perform. In addition, binding assays involving binary interaction analysis cannot detect indirect interactions, while *in vitro* display technologies can, in principle, detect indirect interactors that form complexes with the bait via direct interactors (*19*).

In summary, we applied real-time PCR to evaluate the sequences of 22 candidate Jun-associated proteins, including three novel candidates picked up on IVV screening. The results show that real-time PCR analysis is a useful tool for monitoring the process of enrichment of specifically selected sequences and for eliminating false positives. This approach should be useful not only for IVV screening, but also for other display technologies, such as ribosome display, DNA display, and phage display.

We wish to thank Dr. S. Tateyama, Dr. E. Miyamoto-Sato and Dr. M. Yonezawa for the useful suggestions during the preparation of the manuscript. This work was supported in part by a Grant-in-Aid for the 21st Century Center of Excellence (COE) Program entitled "Understanding and Control of Life's Function *via* Systems Biology (Keio University)," and special coordination funds from the Ministry of Education, Culture, Sports, Science and Technology of Japan, funds from the New Energy and Industrial Technology Development Organization (NEDO) of Japan, and a Grant-in-Aid for Research on Advanced Medical Technology, Health and Labor Sciences Research Grants from the Ministry of Health, Labor and Welfare of Japan.

REFERENCES

1. Doi, N. and Yanagawa, H. (2001) Genotype-phenotype linkage for directed evolution and screening of combinatorial protein libraries. *Comb. Chem. High Throughput Screen* **4**, 497–509
2. Lipovsek, D. and Plückthun, A. (2004) *In-vitro* protein evolution by ribosome display and mRNA display. *J. Immunol. Methods* **290**, 51–67
3. Kondo, A. and Ueda, M. (2004) Yeast cell-surface display–applications of molecular display. *Appl. Microbiol. Biotechnol.* **64**, 28–40
4. Smith, G.P. (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**, 1315–1317
5. Fujii, I., Fukuyama, S., Iwabuchi, Y., and Tanimura, R. (1998) Evolving catalytic antibodies in a phage-displayed combinatorial library. *Nat. Biotechnol.* **16**, 463–467
6. Faix, P.H., Burg, M.A., Gonzales, M., Ravey, E.P., Baird, A., and Larocca, D. (2004) Phage display of cDNA libraries: enrichment of cDNA expression using open reading frame selection. *Biotechniques* **36**, 1018–1022, 1024, 1026–1029
7. Hagiwara, H., Kunihiro, S., Nakajima, K., Sano, M., Masaki, H., Yamamoto, M., Pak, J.W., Zhang, Y., Takase, K., Kuwabara, I. *et al.* (2002) Affinity selection of DNA-binding proteins from yeast genomic DNA libraries by improved lambda phage display vector. *J. Biochem.* **132**, 975–982
8. Yamauchi, A., Nakashima, T., Tokuriki, N., Hosokawa, M., Nogami, H., Arioka, S., Urabe, I., and Yomo, T. (2002) Evolvability of random polypeptides through functional selection within a small library. *Protein Eng.* **15**, 619–626
9. Mattheakis, L.C., Bhatt, R.R., and Dower, W.J. (1994) An in vitro polysome display system for identifying ligands from very large peptide libraries. *Proc. Natl Acad. Sci. USA* **91**, 9022–9026
10. Hanes, J. and Pluckthun, A. (1997) *In vitro* selection and evolution of functional proteins by using ribosome display. *Proc. Natl Acad. Sci. USA* **94**, 4937–4942
11. Zhou, J.M., Fujita, S., Warashina, M., Baba, T. and Taira, K. (2002) A novel strategy by the action of ricin that connects phenotype and genotype without loss of the diversity of libraries. *J. Amer. Chem. Soc.* **124**, 538–543
12. Nemoto, N., Miyamoto-Sato, E., Husimi, Y., and Yanagawa, H. (1997) *In vitro* virus: bonding of mRNA bearing puromycin at the 3'-terminal end to the C-terminal end of its encoded protein on the ribosome in vitro. *FEBS Lett.* **414**, 405–408

13. Miyamoto-Sato, E., Takashima, H., Fuse, S., Sue, K., Ishizaka, M., Tateyama, S., Horisawa, K., Sawasaki, T., Endo, Y., and Yanagawa, H. (2003) Highly stable and efficient mRNA templates for mRNA-protein fusions and C-terminally labeled proteins. *Nucleic Acids Res.* **31**, e78
14. Doi, N. and Yanagawa, H. (1999) STABLE: protein-DNA fusion system for screening of combinatorial protein libraries *in vitro*. *FEBS Lett.* **457**, 227–230
15. Yonezawa, M., Doi, N., Kawahashi, Y., Higashinakagawa, T., and Yanagawa, H. (2003) DNA display for *in vitro* selection of diverse peptide libraries. *Nucleic Acids Res.* **31**, e118
16. Odegrip, R., Coomber, D., Eldridge, B., Hederer, R., Kuhlman, P.A., Ullman, C., FitzGerald, K., and McGregor, D. (2004) CIS display: *In vitro* selection of peptides from libraries of protein-DNA complexes. *Proc. Natl Acad. Sci. USA* **101**, 2806–2810
17. McPherson, M., Yang, Y., Hammond, P.W., and Kreider, B.L. (2002) Drug receptor identification from multiple tissues using cellular-derived mRNA display libraries. *Chem. Biol.* **9**, 691–698
18. Hammond, P.W., Alpin, J., Rise, C.E., Wright, M., and Kreider, B.L. (2001) *In vitro* selection and characterization of Bcl-X(L)-binding proteins from a mix of tissue-specific mRNA display libraries. *J. Biol. Chem.* **276**, 20898–20906
19. Horisawa, K., Tateyama, S., Ishizaka, M., Matsumura, N., Takashima, H., Miyamoto-Sato, E., Doi, N., and Yanagawa, H. (2004) *In vitro* selection of Jun-associated proteins using mRNA display. *Nucleic Acids Res.* **32**, e169
20. Yasumo, H., Masuda, N., Furusawa, T., Tsukamoto, T., Sadano, H., and Osumi, T. (2000) Nuclear receptor binding factor-2 (NRBF-2), a possible gene activator protein interacting with nuclear hormone receptors. *Biochim. Biophys. Acta* **1490**, 189–197
21. Miyasaka, H., Choudhury, B.K., Hou, E.W., and Li, S.S. (1993) Molecular cloning and expression of mouse and human cDNA encoding AES and ESG proteins with strong similarity to Drosophila enhancer of split groucho protein. *Eur. J. Biochem.* **216**, 343–352
22. Smith, C.M. and Steitz, J.A. (1998) Classification of gas5 as a multi-small-nucleolar-RNA (snoRNA) host gene and a member of the 5'-terminal oligopyrimidine gene family reveals common features of snoRNA host genes. *Mol. Cell. Biol.* **18**, 6897–6909
23. Higuchi, R., Fockler, C., Dollinger, G., and Watson, R. (1993) Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. *Biotechnology* **11**, 1026–1030
24. Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680

生物機能解明に向けた*in vitro* virus（IVV）法によるゲノムネットワーク解析

宮本悦子・柳川弘志

ゲノムネットワーク解析における蛋白質間相互作用（PPI）解析法として，筆者らはこれまで世界に先駆けて，独自性の高い *in vitro* virus（IVV）法とC末端ラベル化法の2つの技術（ピューロマイシン・テクノロジーと総称）を開発し，これらの試験管内で実現できる技術を応用したハイスループットで偽陽性の少ないPPI解析法を構築してきた．本稿では，IVV法による大規模解析システムの構想を述べ，IVV法の特徴として，塩基配列解析による相互作用配列（モチーフやドメイン）の抽出と1：多分子解析による複合体解析について解説し，IVV法の生物機能解析やゲノム創薬への可能性について言及する．

Key Words　*in vitro* virus（IVV）　C末端ラベル化　蛋白質間相互作用　PPI　ゲノムネットワーク

●はじめに

ヒトとチンパンジーの塩基配列解読により，その塩基配列の差は約1％であるが，配列変化は遺伝子領域に集中しており，遺伝子の表現型である蛋白質で比較するとその差は80%であることが判明した[1]．このことは，ゲノムプロジェクトの塩基配列解析完了は決してゴールではなく通過点であり，ヒトがヒトであるゆえんに関しての知見を得るためには，ポストゲノムプロジェクトこそが重要な意味をもつ研究であることを示唆している．2003年4月，ヒトゲノムの全塩基配列解読が基本的に完了したことを契機として，米国でENCODE（Encyclopedia of DNA Elements）計画が発表され，本格的なゲノム機能解析の時代に入った．日本では，ヒトをはじめ，マウス，線虫，シロイヌナズナ，チンパンジー，イネなどの世界一を誇るcDNAコレクションの資源を今こそ有効利用し，世界に貢献できるゲノムネットワークの基盤データを創出し，生命現象の理解や創薬などゲノム産業の基盤を育てるべきときが到来している．

本稿では，ゲノムネットワーク解析の蛋白質間相互作用（PPI）解析法の1つとして，*in vitro* virus（IVV）法[2,3]を紹介する．本手法は最初，蛋白質の進化分子工学として，蛋白質の試験管内進化（*in vitro* selection）のためのツールとして誕生した．1997年に筆者らが世界に先駆けて発表し[2]，次いで米国のSzostakらのグループが発表した[4]．その目的は，蛋白質の進化の仕組みの解明と，その工学的応用として，分子レベルのダーウィン進化の仕組みを利用して，所望の機能をもつ蛋白質を試験管内で創出しようとするものであった．進化分子工学において，遺伝子型と表現型の対応づけは重要な要素技術であり，IVVはピューロマイシンを介した遺伝子型（mRNA）と表現型（蛋白質）の対応づけ分子である[5]．

IVV法がゲノムネットワーク解析へ応用可能である理由は，進化分子工学の要素技術の1つである対応づけ分子を利用することで，ゲノムネットワークを形成している多数の遺伝子の遺伝子型（mRNA）に対応した表現型（蛋白質）の機能を一度に検出できるためである[6]．

本稿では，ゲノムネットワーク解析におけるIVV法の大規模解析システムについて概説し，IVV法により得られる大量のコンテンツを利用した生物機能解析やゲノム創薬への応用などの展開について述べる．

Etsuko Miyamoto-Sato, Hiroshi Yanagawa, 慶應義塾大学理工学部生命情報学科　E-mail：nekoneko@educ.cc.keio.ac.jp, hyana@bio.keio.ac.jp

Analysis of genome network toward understanding biological function using *in vitro* virus

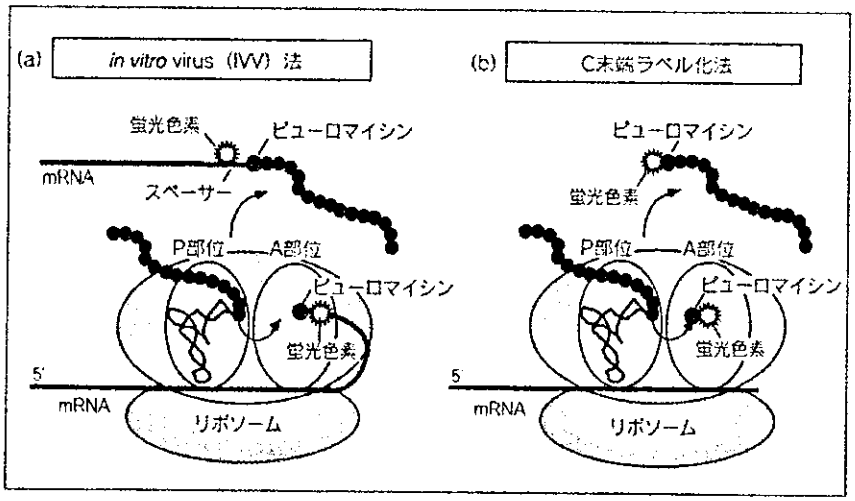

図1 ピューロマイシン・テクノロジー
(a) 無細胞翻訳系のリボソーム上で合成された表現型である蛋白質とその遺伝子型であるmRNAがピューロマイシンを介して共有結合した対応づけ分子を*in vitro* virus (IVV) と名づけた[3]．
(b) ピューロマイシン誘導体の濃度を調整することにより，無細胞翻訳系のリボソーム上で合成された全長蛋白質のC末端にピューロマイシン誘導体が結合できる特性を利用したものがC末端ラベル化法である[7]．この図の場合，ピューロマイシン誘導体はピューロマイシンと蛍光色素の複合体である．
P部位：ペプチジル部位，A部位：アミノアシル部位．

I. IVV法の原理とゲノムネットワーク解析

遺伝子型と表現型の対応づけ手法であるIVV法[2,3]は，図1aに示すように，mRNAの3′末端にスペーサーを介して抗生物質の一種のピューロマイシンを結合し，それを鋳型として無細胞翻訳反応を行なうことにより，蛋白質とmRNAがピューロマイシンを介して共有結合した単純なmRNA-蛋白質連結分子IVVが構築される．また，IVV法の構築過程で，図1bのように，低濃度のピューロマイシン誘導体を無細胞翻訳系に投入すると，合成された蛋白質のC末端にピューロマイシン誘導体が特異的に結合することを見いだした[7]．この原理を応用して，蛍光色素をピューロマイシンに連結させた化合物を用いることにより，蛋白質のC末端を蛍光色素でラベルすることが可能になった[8]．筆者らは，IVV法とC末端ラベル化法の2つの技術を，ピューロマイシン・テクノロジーと名づけた[3]．その後，筆者らは，平成11～15年度「科学技術振興調整費によるゲノムフロンティア開拓推進制度」プロジェクト研究において，ゲノム機能解析の基盤技術として，ハイスループット解析に適したIVV共翻訳セレクション法による複合体解析法[9]，およびC末端ラベル化法を用いたプロテインチップあるいは蛍光相互相関分光法（fluorescence cross correlation spectroscopy：FCCS）によるPPI解析法[10,11]による網羅的解析システムを構築してきた．

図2に示すように，本システムは，IVVライブラリーのなかからベイト蛋白質と結合する蛋白質を含むIVVを試験管内で釣り上げたのち，そこに連結している情報タグ（mRNA）を逆転写・PCRで増幅し，塩基配列を解読することによって，相互作用する蛋白質群を容易に同定することができるシステムである．

ゲノムネットワーク解析のための大規模PPI解析ツールとしてよく知られている方法は，酵母ツーハイブリッド法（Y2H）[12]とタグ精製法であるTAP（tandem affinity purification）-MS（質量分析）法[13]である．しかしながら，表1に示すように，これらの方法では生きた細胞を使用するために，毒性やライブラリーサイズの制限などの問題があった．筆者らは，これらの欠点を克服する手法としてIVV法を実現した．IVV法はツーハイブリッド法と比べて，間接的な相互作用を含めた複合体の解析が可能であり，転写因子など複合体を形成する分子群などの解析に有利である．また，IVV法は蛋白質ではなく核酸による検出であるため，TAP-MS法と比べて感度が百万倍も高く，ごく微量の蛋白質でも検出が可能な革新的な手法であり，かつ相互作用配列（モチーフ，機能エレメント）を得ることが可能である．さらに，IVV法は，スクリーニングに濃縮をくり返すセレクションを採用するため偽陽性が少なく，信頼性の高いデータが得られる．また，ネットワークの全体像を知るには，1つで完璧な方法というものはなく，いくつかの方法で補うことが必要であることから，IVV法はゲノムネットワーク解析の基盤データ創出において，重要な役割を担うツールの1つとして期待できる．

これまで*in vitro*で蛋白質間相互作用を解析する他の手法として，ファージディスプレイ法[14]やリボソーム

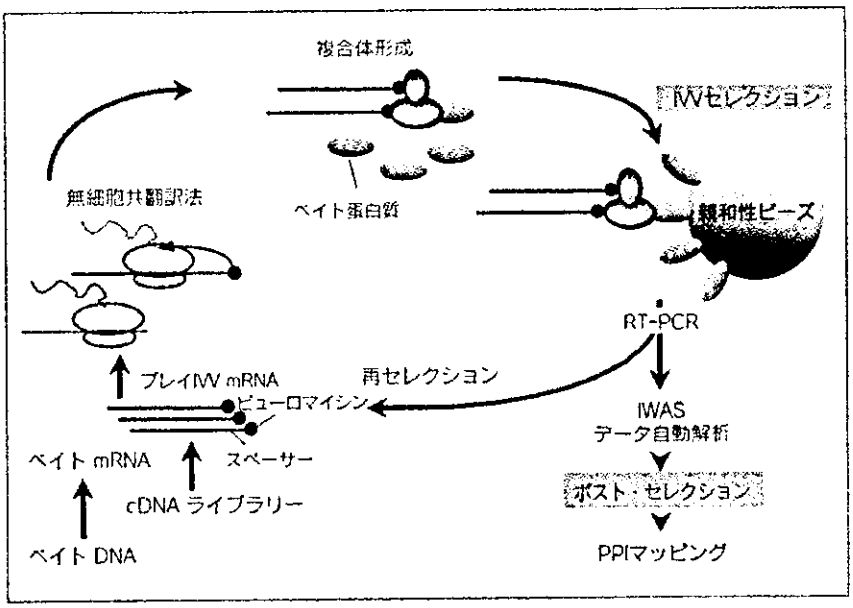

図2 IVV法による複合体とPPI解析
本システムは，セレクション，ポストセレクション（セレクションの検証工程），シークエンス解析と遺伝子解析（IWASデータ解析），PPIマッピングの工程から構成される．セレクションは，ベイト蛋白質とプレイIVVライブラリーを無細胞翻訳系で共翻訳することにより，ベイトとプレイの複合体を形成し，ベイトのタグでプレイをセレクションし，プレイIVVのmRNAを利用して，RT-PCR（逆転写PCR）によって検出されたプレイの遺伝子群を解析する[9]．ポストセレクションは，C末端ラベル化法を用いたプロテインチップあるいは蛍光相互相関分光法（FCCS）によるPPI解析法[10,11]によりセレクションの検証を行なう．IWASデータ解析，PPIマッピングの工程については，図3を参照．

表1 IVV法と他の方法の特徴の比較

特 徴	Y2H法	IVV法	TAP-MS法
相互作用解析	1：1分子	1：多分子	1：多分子
検出方法	核酸	核酸	蛋白質
in vitro/in vivo	*in vivo*	*in vitro*	*in vivo*

ディスプレイ法[15]が提案されている．しかし，ファージディスプレイ法は大腸菌を用いるため，ライブラリーのサイズが制限され，大腸菌に対して毒性をもつ蛋白質は選択されないなどの欠点がある．一方，スイス・チューリッヒ大のグループが開発したリボソームディスプレイ法は，IVV法と同様，無細胞翻訳系を用いた完全に*in vitro*の実験系であるため，細胞を使うことによる欠点はないが，mRNA-蛋白質-リボソーム三者複合体が不安定なため，スクリーニング中にmRNAと蛋白質が離れてしまう恐れがあり，さらに遺伝子型と表現型の対応づけ効率が0.1%と非常に低いという致命的な問題がある．

これに対して，IVV法は共有結合を介して対応づけされているため，リボソームディスプレイ法よりもはるかに安定性が高く，また対応づけ効率も約70%[3]であり，リボソームディスプレイ法に比べて数百倍も高い．したがって，現時点でIVV法が蛋白質間相互作用の*in vitro*スクリーニングに最も有効な方法であるといえる．ヒトのゲノムワイドなネットワーク解析が，IVV法によって達成できれば，日本発の独自技術による初めての大規模解析として，ポストゲノム時代における国際的な優位性を示せる．

II. IVV法による大規模解析システム

図2に示したIVV法によるPPI解析システムは，セレクション，ポストセレクション（セレクションの検証工程），シークエンス解析と遺伝子解析，PPIマッピングとマップ解析の工程から構成される．IVV法によるゲノムネットワーク解析のための大規模解析を念頭におくとき，最も時間がかかる工程はセレクション後のシークエンス解析と遺伝子解析である．シークエンス解析は，GENETIXのコロニー・ピッカー（Qpix）とQiagenのラボロボット（BioRobot 8000）の導入により，一晩で数百個のシークエンスデータが得られるようになった．次に，これら大量のシークエンスデータを入力すると自動的に解析されて，遺伝子リストが得られるIVV自動解析システム（IWAS）を，富士通の協力を得て開発した（図3a）．このシステムでは，一晩で数百シークエンスの遺伝子解析が可能で，自動作成された遺伝子リストは，遺伝子解析ソフト（富士通 Genesphere；以前の名称 Xminer）にリンクし，PPIマップが自動的に作成できる（図3b）．さらに，Genesphereは，LocusLink, UniGene, PubMedなどのデータベースを利用したマッピング解析を提供してくれる．このシステムによって，これまで数

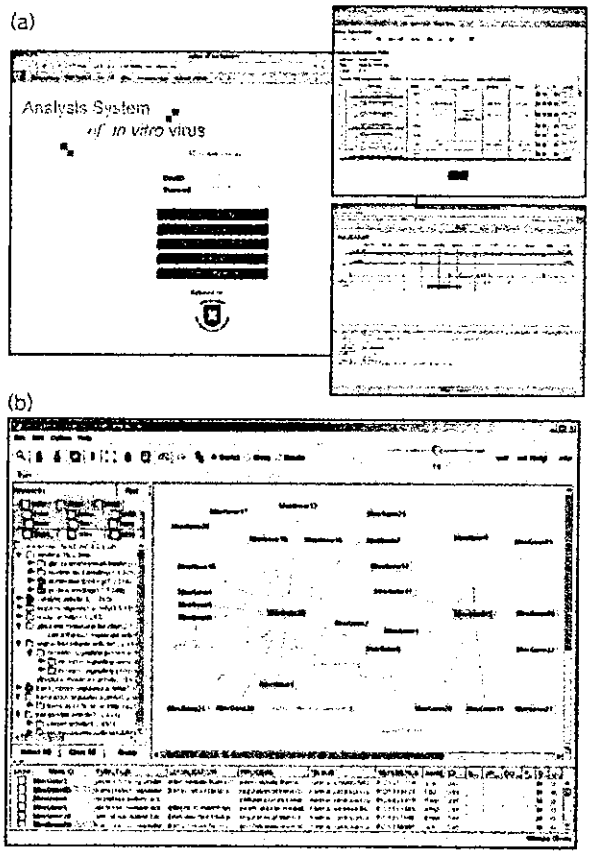

図3　IVVデータ自動解析システム
(a) 図2によって検出されたプレイの遺伝子群のシークエンス解析データを自動的に遺伝子検索にかけ，遺伝子リストやゲノムビュアーとして表示できるIVV解析システム（IWAS）[17]，(b) IWASによって解析された遺伝子群のリストは，遺伝子解析ソフト（富士通Genesphere）とリンクしており，IVV法で検出された複合体やPPIマップが自動的に表示でき，機能解析や疾患解析などが可能である[17]．

カ月かかっていたシークエンスと遺伝子解析，およびPPIマッピングとマップ解析の工程が数日で可能になり，ゲノムワイドなネットワーク解析のための土台が出来上がった．

IVV法による大規模解析を実現するために，図4に示すようなシステムの構築を目指している．日本のゲノムプロジェクトの資産であるヒトやマウスのcDNAコレクションを利用して数百単位のベイトを設計し，ラボロボット（BioRobot 8000）をシステム化したIVVセレクション・ロボットによる並列セレクション解析，さらにクローニング・シークエンス解析による数万に及ぶ大量データの高速処理システムとして，IWAS（一晩で数千シークエンス処理），および遺伝子機能解析が可能な相互作用マッピングツールとしてGenesphereなどを利用した大規模解析を構想している．IVV法では，1つのベイトを用いたセレクションで平均数十の相互作用が得られるので，数百ベイトのセレクションで数千の相互作用が解析されると予想される．

ここで，IVV法の相互作用は，先に述べたように，1：1分子解析としてのPPI解析のみならず，1：多分子解析としての複合体解析が実現できる．一方，C末端ラベル化法のプロテインチップによる解析は，1：1分子解析としてのPPI解析データである．IVV法では，偽陽性は少ないが偽陰性が多い．反対に，プロテインチップでは，偽陰性は少ないが偽陽性が多いので，相補しあえる関係にある．よって，同じベイトを用いたIVV法によるデータとプロテインチップによるデータをマージすることにより，バイオインフォマティクスによる複合

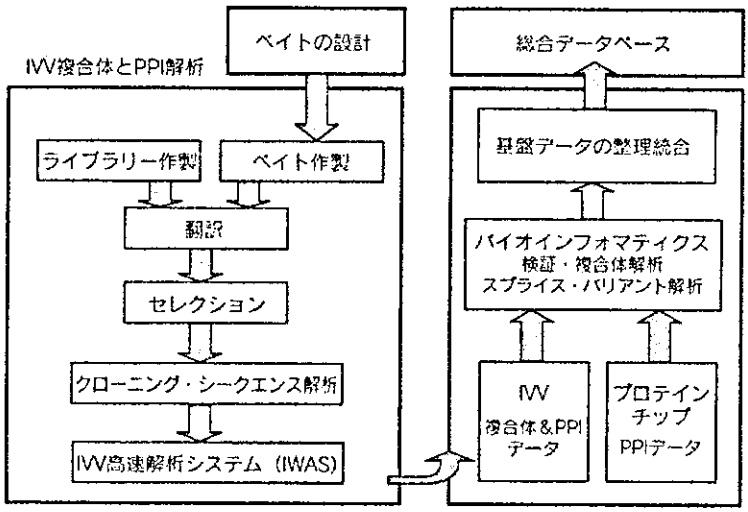

図4　IVV法による大規模解析システム
IVV法で得た複合体データと，プロテインチップでのPPIデータを統合して，バイオインフォマティクスで信頼性の高い複合体情報を抽出し，さらにスプライス・バリアント解析などによるアノテーション情報を備えた基盤データを創出し，データベース化することを目指している．

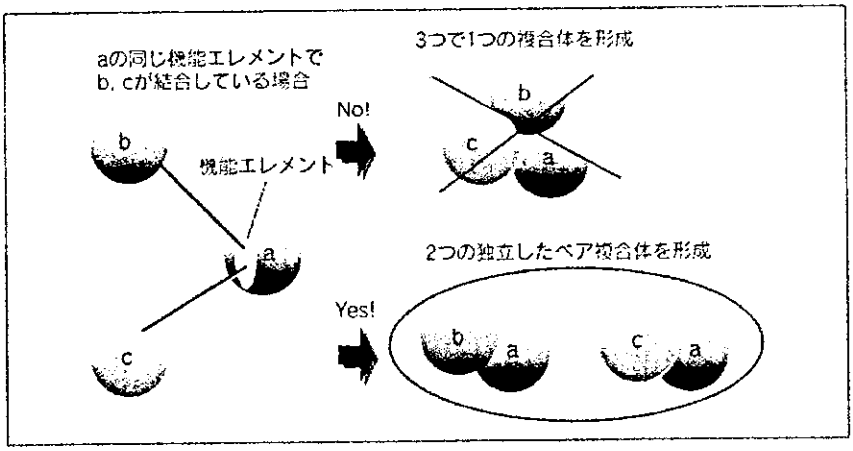

図5　モチーフあるいは機能エレメントに基づく複合体予測
蛋白質aの同じ機能エレメントによって，蛋白質aが蛋白質bおよびcと相互作用していれば，複合体の形成は3つの蛋白質による複合体ではなくて，2つの独立したペア複合体を形成していると推測できる．

体の抽出や，データの信頼性が向上することが可能と考えられる．さらに，IVV法が塩基配列の解析法であるため，疾患関連情報として重要なスプライス・バリアント解析[16]が可能なことから，複合体情報とモチーフ情報のみならず，スプライス・バリアント情報などもアノテーションとしてもつIVVデータベースを得ることを目指している．

III. モチーフ抽出に基づくIVV法による複合体解析

図4に示したゲノムネットワーク解析のための大規模解析システムの特徴として，他のPPI解析法と区別できる点を強調しておきたい．まず，① 1：多分子解析により複合体解析が可能であること，② 相互作用に必要な最低限の結合配列(モチーフあるいは機能エレメント)領域を同時に解析可能であることがあげられる．複合体の解析はTAP-MS法で可能であるが，解析と同時に機能エレメントの情報を得ることはできない．すなわち，同じ複合体の解析方法でも，TAP-MS法の場合とIVV法の場合では，解析後のバイオインフォマティクスのレベルでの複合体解析のための情報量が格段に違ってくるのである．

たとえば，図5に示すように，3つの蛋白質がどのような複合体を形成して機能しているかを予測する場合，モチーフが抽出されていない場合は3つで複合体を形成して機能しているのか，あるいは2つのペアで複合体を形成して機能しているのかを予想することは困難である．モチーフが抽出されている場合，aの蛋白質のもつ同じモチーフがbの蛋白質とcの蛋白質との相互作用に使われていれば，(a, b, c) の3つで複合体を形成しているのではなく，(a, b)，(a, c) の2つのペアで複合体を形成して機能していることが予想される．このことはさらに，今まで単に遺伝子同士を結ぶだけのマッピングであったものを，マップからその複合体のあり方を具体的に抽出することによって，カスケードの予測に繋げていける可能性も期待できる．

IV. IVV法による生物機能解析やカスケード解析へ向けて

IVV法による解析には，あらゆる生物の種や組織への適用の壁がない．すなわち，あらゆる生物の種や組織由来の少量のmRNAがあれば，自由に実験に用いられる点も大きな特徴である．このことから，容易に組織間やマウスとヒトなどの種間で[17]，あるいは正常系と疾患系の組織間などでネットワークパターンの違いを比較していくことが可能な実験系である(図6)．

これまでmRNAプロファイルなどで量的な変動の差は計測されていたが，PPI解析の相互作用パターンを比較することができれば，具体的な相互作用の変化などをとおして，どのような機能に影響するかを解析できる．これらの解析をとおして，IVV法による生物機能解明やゲノム創薬を実現していくためには，ゲノムネットワーク解析によって，カスケード解析に繋がるようなPPI解析を戦略的に行なっていくことが必要である．そのためには，先に説明したIVV法によって1：多分子解析が可能なことと，相互作用に必要な配列としてモチーフあるいは機能エレメントが解析できる点を利用して，具

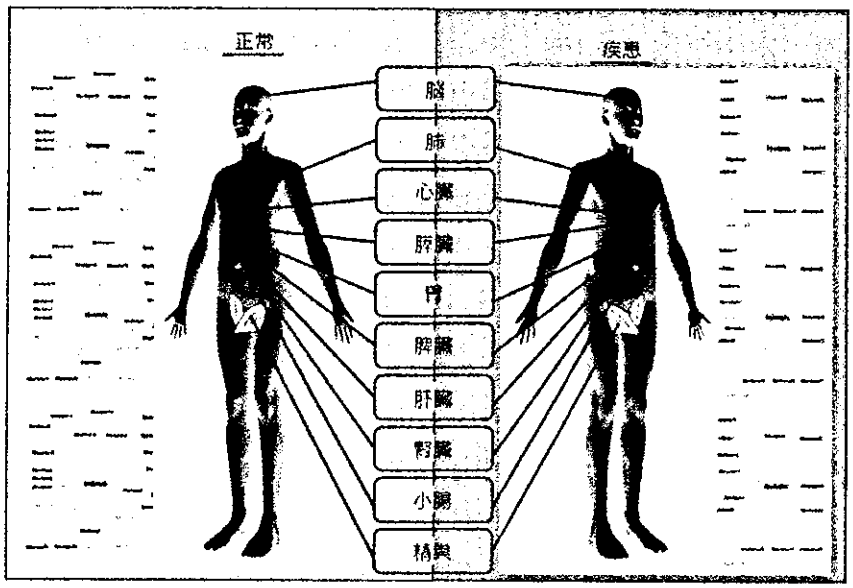

図6　ヒト正常・疾患組織のネットワーク・パターン比較
IVV法では，市販の疾患ライブラリーを用いて，あるいは疾患組織からmRNAを少量抽出できればライブラリーを構築できるため，ヒト正常・疾患組織のネットワーク・パターンの比較が可能である．

体的な複合体予想やその複合体の機能予想からカスケード解析に近づけること，および生物現象にとって最も重要なイベントとして，ゲノムネットワークの要である転写制御について着目し，転写因子複合体とDNA結合部位のプロモーター配列を1つの組として解析していくことが重要であると考えている．なぜなら，多くの生物機能で転写因子複合体の組換え，あるいは転写因子複合体へのコファクターの結合によりその転写因子複合体と結合するプロモーターをもつ遺伝子が変わることによって，さまざまな機能や疾患が発生している例が見られるためである[16]．

これらのことから，ゲノムネットワーク解析によるカスケード解析のための情報として，先に示したスプライス・バリアント情報，複合体情報，モチーフ情報のみならず，プロモーター情報などもアノテーションとしてもつIVVデータベースを得ることを目指している．

●おわりに

ゲノムネットワーク解析におけるIVV法の大規模解析の構想を述べた．今後，あらゆるツールによるあらゆる角度からのゲノムネットワーク解析が進み，大量のコンテンツがデータベースとして公開されることになるであろう．それらのコンテンツをいかに利用し，どのようにしてそこから個別の機能や疾患を解明していくかが重要になってくる．しかしながら，*in vitro*のPPI解析結果から個別機能や疾患解明を実現するには，*in silico*の解析により生物機能に重要な遺伝子や創薬ターゲット分子を絞り込み，さらに*in vivo*解析などで検証していく必要がある．*in silico*の解析では，バイオインフォマティクスの活躍が期待され，*in vivo*解析では，IVV法を本来の進化分子工学的手法として利用することで，ターゲット分子の抗体やアンタゴニストを得ることができるので[17]，培養細胞，組織，モデル動物実験で検証していくことが可能になる．このように，ゲノムネットワーク解析で大量のコンテンツを創出するだけでなく，生物機能解明への道筋をきちんと検証していくことが今後重要であり，これは，IVV法の信頼性の確立にとっても避けて通ることのできない道だと考えている．

IVV自動解析システム（IWAS）は，富士通ライフサイエンスシステム事業部とバイオIT事業開発本部の協力を得て開発された．また，本研究の一部は，文部科学省の科学技術振興調整費による「細胞内でネットワークを構成しているタンパク質の相互作用を試験管内で解析するための新しいツールの開発」の一環として行なわれた．

文　献

1) Sakaki, Y. *et al.* : *Nature*, 429, 382-388 (2004)
2) Nemoto, N., Miyamoto-Sato, E., Husimi, Y., Yanagawa, H. : *FEBS Lett.*, 414, 405-408 (1997)
3) Miyamoto-Sato, E. *et al.* : *Nucleic Acids Res.*, 31, e78 (2003)
4) Roberts, R. W., Szostak, J. W. : *Proc. Natl. Acad. Sci. USA*, 94, 12297-12302 (1997)
5) 宮本悦子・柳川弘志：蛋白質 核酸 酵素, 46, 138-147 (2001)

6) 宮本悦子・柳川弘志：プロテオミクス（伊藤隆司・谷口寿章 編），pp.136-145，中山書店（2000）
7) Miyamoto-Sato, E., Nemoto, N., Kobayashi, K., Yanagawa, H.: *Nucleic Acids Res.*, 28, 1176-1182 (2000)
8) Nemoto, N., Miyamoto-Sato, E., Yanagawa, H.: *FEBS Lett.*, 462, 43-46 (1999)
9) 宮本悦子・柳川弘志：蛋白質 核酸 酵素 増刊，化学と生物学の接点がつくるNewバイオロジー，48, 1474-1480 (2003)
10) Doi, N. *et al.*: *Genome Res.*, 12, 487-492 (2002)
11) Kawahashi, Y. *et al.*: *Proteomics*, 3, 1236-1243 (2003)
12) Ito, T. *et al.*: *Proc. Natl. Acad. Sci. USA*, 98, 4569-4574 (2001)
13) Uetz, P. *et al.*: *Nature*, 403, 623-627 (2000)
14) Tong, A. H. Y. *et al.*: *Science*, 295, 321-324 (2002)
15) Hanes, J. *et al.*: *Nat. Biotechnol.*, 12, 1287-1292 (2000)
16) Resch, A. *et al.*: *J. Proteome Res.*, 3, 76-83 (2004)
17) 宮本悦子・柳川弘志：Bioベンチャー，7-8, 61-64 (2004)