

局在解析から蛋白質－蛋白質相互作用解析へ Localization to protein-protein interactions

局在解析グループ 長田 直樹
Naoki Nagata

The localization of a protein is related to its function. Proteins showing the same localization may function by interacting each other. Thus, the information on their localizations is useful not only to predict the function of the protein but also to select target proteins for searching novel protein-protein interactions. Using the green fluorescent protein (GFP) fusion technique, we analyzed the localization of proteins encoded by novel cDNAs whose sequences had already been determined, resulting in discovery of many clones showing an interesting localization pattern. As an example, we showed the analysis of a cDNA encoding a nuclear protein that exist in a speckle pattern identical to that of spliceosome. Furthermore we developed a novel technique, named "two-hybrid localization method", for detecting protein-protein interaction, based on an observation of the localization change of GFP-fused protein induced by its interaction with a protein possessing a localization signal after coexpressing both proteins in the cells. This method was successfully applied to detect the interaction between the WW domain and the PGR motif.

はじめに

機能未知の蛋白質が与えられた場合、局在がわかれば、その蛋白質の機能についてある程度の目処を付けることができる。例えば、核に局在していれば、DNAの複製・転写・スプライシングなどに、またミトコンドリアに局在していれば、ATP合成などに関与している可能性が高いであろう。また、蛋白質－蛋白質相互作用を調べる際にも、局在が同じもの同士の中から検索したほうが、見つかる可能性が高い。このようなわけで、塩基配列が決定された完全長 cDNA のクローンについて、緑色蛍光蛋白質 (GFP) を融合させた蛋白質を動物細胞内で発現させて、その局在を網羅的に調べることにした。その過程で、核内に斑点上の特徴的な局在を示したクローンについて、さらに相互作用する蛋白質の検索を行い、このクローンがスプライセオソームの一構成成分であることを明らかにした。また、網羅的局在解析のために作製した GFP 融合蛋白質並びに得られた局在情報を有効利用するため、蛋白質同士の結合によって局在が変わることを指標にして蛋白質間相互作用を検出する方法を考案した。

完全長 cDNA クローンの網羅的局在解析

全長配列が決定された新規の完全長 cDNA クローンについては、オープンリーディングフレーム(ORF)がコードしている蛋白質の推定アミノ酸配列の情報を基に、C 末端に緑色蛍光蛋白質(GFP)を融合させたものを発現させるベクターを構築した。この発現ベクターをリポフェクション法を用いて、サル腎臓細胞株 COS-7 に導入した後、融合蛋白質の発現を蛍光顕微鏡で調べた。もし、細胞内に GFP の蛍光が観測できれば、ORF がコードしている蛋白質と GFP の融合蛋白質が細胞内で確かに発現していることを意味する。したがって、推定した ORF が正しいかどうかのチェックを行っていることにもなる。また、蛍光を発している場所から、融合蛋白質の細胞内局在部位を決定できる。ただ、局在化シグナルが C 末端にある場合には、GFP の融合によってシグナルがマスクされることがありうるので、正しい局在を示さなくなる可能性がある。

これまでに約 150 種類の蛋白質について、GFP 融合蛋白質の発現を行い、その局在部位を決定した。表 1 にその結果の一部を示す。GFP 融合蛋白質の発現が全く認められないものもあったが、ほとんどのクローンは蛍光の強さに差はあるものの、それぞれに特徴的な局在パターンが観測された(本報告書付録プロテオグラム参照)。ミトコンドリア、核、小胞体などへの局在は、一目で判定できるが、細胞全体が光るものについては、その局在を既知の構造体に対応付けることが困難なものが多い。また、凝集塊を形成するものも多く認められた。現在用いている発現系は、SV40 のプロモーターを用いた強制発現系なので、大量発現の結果発現産物が凝集してしまう場合もあると思われる。いずれにしても、融合蛋白質という人工産物を強制発現させているので、局在を正しく反映しているとは限らない。したがって、この方法によって得られる局在は、あくまでも参考データと考え、本当の局在を知るためには、内在性蛋白質の局在を免疫染色法などによって確認する必要がある。免疫染色に必要な抗体は、完全長 cDNA 発現ベクターを直接動物に接種する cDNA 免疫法で作製することを試みた。これについては、伊藤の報告書を参照していただきたい。

新規スプライセオソーム構成成分 Nps20 の発見

上記局在解析の結果、核内に斑点上の局在パターンを示すクローンが得られたので、これに着目しさらに解析を進めることにした。本クローンは、157 アミノ酸残基からなる蛋白質をコードしており、インビトロ翻訳産物を SDS-PAGE 電気泳動により解析したところ、分子量約 20kDa の産物が得られた。また、GFP 融合蛋白質を一過性に発現させた HeLa 細胞を共焦点レーザー顕微鏡で観察したところ、9~12 時間後には、核内に局在する強い GFP 由来のシグナルが観察された(図 1A)。このシグナルは、核内で核小体以外の部位に斑点状に広がる分布を示した。同様の細胞内局在パターンは、COS-7 細胞を用いた場合、もしくは本蛋白質の N 末端に Flag タグを融合した蛋白質を免疫染色で検出した際にも確認された。これらのユニークな局在パターン

は、細胞内において発現された本蛋白質が、核内の mRNA プロセッシング因子の集合体として知られるスプライセオソームへ移行することを示唆した。そこで、代表的なスプライセオソーム構成成分である SC35 を分子マーカーとして用い細胞内での局在を比較した。GFP 融合蛋白質を発現した Hela 細胞を抗 SC35 抗体により免疫染色し、共焦点レーザー顕微鏡で観察したところ、両者に由来する蛍光シグナルは、細胞核内においてよく一致した (図 1B, 1C)。以上の結果から、本蛋白質は核内に移行後、スプライセオソームに局在し得ることが示された。そこで、本蛋白質を Nps20 (Nuclear protein co-localizing with spliceosome with a molecular mass of 20 kDa) と命名した。なお、Nps20 を一過性に発現した細胞は、GFP のみを発現したものと明らかに異なり、多くが約 8~24 時間後に球形に変形し、培養器から脱落、死滅に至った。従って、Nps20 の過剰発現は、なんらかのメカニズムにより細胞に生育阻害を示すと考えられる。

Nps20 は、N 末端に塩基性アミノ酸に富む核移行シグナル (NLS) 様の配列を有していた (図 2A)。また C 末端には疎水性アミノ酸であるロイシンの繰り返し領域が認められることから、核外移行シグナル (NES) 様の配列を併せ持つと推定された。まず Nps20 の N 末端側の NLS 様塩基性アミノ酸クラスターに着目し、このクラスター内の塩基性アミノ酸に関してアラニンとの置換体を作製したところ、6 位のリジン残基の置換体 (K6A 置換体) については核への移行能が失われ、細胞質への局在が観察された (図 2B, 2C)。以上の結果から、Nps20 の核移行において 6 位のリジンが必須であることが判明した。なお、K6A 置換体を強制発現させた細胞について SC35 抗体を用いて免疫染色したところ、内在性の SC35 は野生型を強制発現させた場合と同様の斑点状パターンの局在が観察され、スプライセオソーム形成初期にみられる SC35 などの SR 蛋白質の集合体形成に影響はなかった (図 2B)。また、この K6A 置換体の核からの排除は、Nps20 が NLS 以外に NES をも併せ持つことを示唆した。現在、Nps20 の C 末端側のロイシンに富む配列をもとに NES 活性部位の特定を進めている。

次に、Nps20 と相互作用する蛋白質を検索する目的で、酵母 2-ハイブリッド系によるスクリーニングを行った。この結果、スプライセオソーム構成成分として、U2 snRNP と複合体を形成し、pre-mRNA のスプライシングに関わることが知られている SF3a60/SAP61 との相互作用が見い出された。スクリーニングで得られた複数のクローンは、いずれも SF3a60 の C 末端部分をコードしていた。この部分は RNA 結合に関与すると考えられているジンクフィンガードメインを含む領域である。よって、Nps20 との相互作用に関与するドメインは、SF3a60 の C 末端部分に存在するものと考えられた。さらに Nps20 と SF3a60 の相互作用を GST ブルダウンアッセイにより検証した。大腸菌で発現させた GST タグ付き Nps20 と SF3a60cDNA のインビトロ翻訳産物との結合実験をおこなったところ、GST-Nps20 と SF3a60 が共沈したことから、両者はインビトロでも結合することが示された。このことから、Nps20 は動物細胞内でスプライセオソーム構成成分である SF3a60 と直接の相互作用を示しながら pre-mRNA のスプライシングに何らかの調節因子として関与するものと考えられる。

従来、スプライセオソーム構成成分研究は、核から分離したスプライセオソーム中の構成分子群を蛋白質レベルで個々に単離、同定する手法がとられてきた。一方 Nps20 は本プロジェクトの出発点である完全長 cDNA バンクの網羅的局在解析によって発見されたものである。また、Nps20 は N 末端側に核移行シグナルを、また C 末端側に核外移行シグナルを併せもつことから、核と細胞質間でシャトルしうる分子である事が示唆された。すなわち、Nps20 はその細胞内でのユニークな動態ともあいまって、従来のアプローチ法では見い出すことが難しかった分子であると考えられる。以上、この新たなスプライセオソーム構成成分の発見は、われわれのアプローチの有効性を実証するとともに、いままで知られていない pre-mRNA スプライシングの分子レベルでの制御機構の解明に発展し得るものと期待される。

細胞内局在解析を利用した蛋白質間相互作用の検出法

局在データを基に新しい蛋白質相互作用を見出した例を紹介してきたが、これをさらに発展させる方法を考案した。網羅的局在解析の結果、我々は、GFP タグのついた蛋白質発現ベクターとそれらの局在データを有している。同一ハターンの局在を示す蛋白質は、お互いに結合して複合体を形成している可能性がある。このように局在パターンが同じもの同士の間相互作用を簡便に検出する方法を考案した。まず、GFP を融合させた目的蛋白質 X から局在化シグナル配列(LS)を除いておく (X(Δ LS)-GFP)。これを細胞内で発現させても、GFP 融合蛋白質の特定の局在は認められない。一方、完全長 cDNA バンクの中から X と同じ局在ハターンを示す候補分子 Y を選択する。これら両者を動物細胞内で同時に発現させる。もし、両者間に相互作用があれば、X(Δ LS)-GFP と Y は結合し、X(Δ LS)-GFP は Y の局在する場所へ移動すると考えられる。一方、相互作用しない場合には、X(Δ LS)-GFP の局在ハターンの変化は観測されない。このように 2 つの蛋白質のハイブリッド形成に伴う局在変化を指標として両者の結合能を検出することから、この方法を「2 ハイブリッド局在化法」と命名した (本報告書研究成果集参照)。

本法が蛋白質間結合の検出法として有効かどうかを確かめるため、我々が新たに見出した Npw38 と NpwBP の結合 (本報告書小室の要旨参照) をモデルとして実験を行った。いずれも核に局在する蛋白質である。あらかじめ核局在化シグナル(NLS)を欠損させた Npw38 に GFP を融合させたもの(Npw38(Δ NLS)- GFP)と NpwBP を COS-7 細胞内で発現させたのち、蛍光顕微鏡下で観察した。Npw38(Δ NLS)- GFP を単独で発現させた場合は、細胞質、核いずれにも融合蛋白質が認められたが、NpwBP を同時発現させたものでは、予想どおり核への局在化が認められた (図 3)。また、組み合わせを逆にしても、同様の結果が得られた。すなわち NpwBP の PGR モチーフ (Npw38 の WW ドメインと結合する配列) -GFP を単独で発現させた場合、細胞全体に融合蛋白質の発現が認められるが、Npw38 を同時発現させると、核への局在化が起こった (表 2)。また、この場合に、Npw38 の WW ドメイン変異体を同時発現させても、核への局在化は起こらなかったことから、このような局在変化は、

WW ドメインと PGR モチーフの結合を介して起こっていることを意味している。

上記の例では、核への局在化を指標としたが、他の局在化シグナルが使えるかどうかを検討した。NpwBP の PGR モチーフの下流にミトコンドリアへの局在化シグナル (Bcl-X_L) を融合させたものを作製した。これを Npw38(Δ NLS)-GFP と同時に発現させたところ、ミトコンドリア (抗ミトコンドリア抗体染色により確認) にも GFP の発光が認められた (図3)。すなわち、付加する局在化シグナルを種々替えることによって、任意の局在化を指標とする系を構築できることになる。以上の二つのモデル実験の結果、本法が動物細胞内で蛋白質間相互作用を直接モニターする方法として有効であることが示された。

おわりに

我々はすでに完全長 cDNA クローンを有しているので、これを用いれば細胞内で cDNA がコードしている蛋白質を発現させることができる。この際、GFP などの目印をつけてやれば、その細胞内局在も簡単に知ることができる。このようにして局在によって cDNA クローンを分類した後、ある特定の局在を示す蛋白質群の中から、相互作用するものの組み合わせを網羅的に調べていけば、より効率良く蛋白質間相互作用ネットワークを見つけだすことができると思われる。そのためには細胞を構成している蛋白質の完全長 cDNA クローンをできるだけ多くそろえることと、2 ハイブリッド局在化法を多くの種類の蛋白質間相互作用に適用できるように改良することが望まれる。

表1 GFP融合蛋白質の局在解析

No.	HP No.	Encoded protein	Localization	No.	HP No.	Encoded protein	Localization
1	hP10552s		Aggregate	61	HP10137s	Hypothetical protein DKFZp564D177	Particle & aggre
2	hP10659s		Cytoskeleton-like	62	hP10601s		Particle & nucle
3	HP10512s		Cytosol	63	hP01495s	E. coli BclA protein-like	Particle & nucle
4	hP10515s	Hypothetical protein DMC63B12-like	Cytosol (aggregate)	64	HP10496s	Ubiquitously-expressed transcript (HSP)	Particle or nucle
5	hP01017s	short-chain-type dehydrogenase	Cytosol (mesh)	65	hP03314s	DNA repair protein (mei-1)-like	Whole
6	hP10582s	Hypothetical protein 108.7 kDa-like	Cytosol (mesh)	66	hP03090s	Hypothetical 32.0 kDa protein-like (glyo)	Whole
7	HP01927s	Mitochondrial glutathione S-transferase-	Cytosol (mesh) & r	67	hP03264s	ER1 (potential transcription factor) -like	Whole
8	hP10098s		Cytosol (particle)	68	hP10052s		Whole
9	hP10650s		Cytosol (particle)	69	hP10090s	Hypothetical protein HSPC014	Whole
10	hP10162s	Hypothetical protein-like	Cytosol (particle)	70	hP10106s	beta-transducin family WD repeat?	Whole
11	hP01640s	Plasmodium falciparum surface antigen	Cytosol (whole)	71	hP10111s		Whole
12	HP03248s	Hypothetical protein CGI-44 (Flavocyto	Cytosol & aggrega	72	HP10149s	Hypothetical protein CEW02A11-like	Whole
13	HP02409s	zinc finger protein-like	Cytosol & Nuclear	73	hP10160s	Hypothetical protein CELZK1248-like	Whole
14	hP10410s		Cytosol & nuclear	74	hP10173s	Hypothetical protein CEC04H5-like	Whole
15	HP10018s	Hypothetical protein 15.4 kDa-like	Cytosol & particle	75	HP10177s	Adenosine 5'-diphosphosugar pyrophos	Whole
16	hP10456s	Developmental protein DG1118-like (Ra	Cytosol aggregate	76	hP10200s		Whole
17	hP10151s		Golgi	77	hP10211s		Whole
18	hP10565s		Golgi	78	hP10327s		Whole
19	hP10658s		Golgi	79	HP10350s	G8 protein	Whole
20	hP03145s	Mitochondrial para-hydroxybenzoate=pc	Golgi & ER	80	HP10400s		Whole
21	HP10534s	DKFZp586A0722	Golgi & ER	81	hP10525s	Hypothetical protein SPAC8C9.11-like	Whole
22	hP03233s	Putative ubiquinone biosynthesis methyl	Golgi & ER	82	HP10546s	HSPC007	Whole
23	hP10570s		Golgi & ER	83	hP10641s		Whole
24	hP10649s		Membrane?	84	hP10645s	Hypothetical protein SPBC1709.14-like	Whole
25	hP00698s	3-hydroxyisobutyrate dehydrogenase-lik	Mitochondria	85	hP10654s		Whole
26	HP00758s	Mitochondrial ribosomal protein L2-like	Mitochondria	86	hP10657s	Ubiquitin-like	Whole
27	HP02206s	Mitochondrial ribosomal protein BL22-lik	Mitochondria	87	hP10334s	SH3 domain-binding glutamic acid-rich-l	Whole
28	hP02612s	Ribosomal protein BL9-like	Mitochondria	88	HP10382s	WW-domain binding protein 2	Whole
29	HP03324s	Hypothetical protein CGI-22 (mitochond	Mitochondria	89	hP10417s		Whole
30	hP10117s	Ribosome recycling factor-like	Mitochondria	90	hP10482s	Profilaggrin (antisense)	Whole
31	HP10352s	Hypothetical protein CGI-139 (Gangliosic	Mitochondria	91	hP10498s	Hypothetical protein C24D10.6-like	Whole
32	hP10421s		Mitochondria	92	hP10499s		Whole
33	hP10522s	Mitochondria ribosomal protein?	Mitochondria	93	hP10501s		Whole
34	HP01748s	Mitochondria ribosomal protein S2-like	Mitochondria & ag	94	hP10511s		Whole
35	HP02051s	Mitochondria ribosomal protein-like	Mitochondria & cyt	95	hP10553s	Hypothetical protein YDR031w-like	Whole
36	hP10505s	Hypothetical protein HSPC139 (Postsyn	Mitochondria & nu	96	hP10560s		Whole
37	HP02907s	Hypothetical protein PTD017 (rp BS18)	Mitochondria or M	97	hP10564s		Whole
38	hP10660s		no expression	98	hP10569s		Whole
39	hP10653s		no PCR product	99	hP02573s	GTP-binding protein CgpA-like	whole
40	HP10030s	Hypothetical 6.3 kDa protein ZK652.2-lik	no signal	100	hP10021s	Hypothetical protein-MTH1307-like	whole
41	hP10434s		no signal	101	hP10120s	Hypothetical protein CEF45G2-like	whole
42	hP10500s		no signal	102	hP10321s		Whole
43	HP00726s	Transcription factor HES-1-like (Hairy hc	Nuclear	103	hP10384s		Whole (aggreg
44	HP10121s	MORF-related gene 15	Nuclear	104	HP10342s		Whole (aggreg
45	hP10356s		Nuclear	105	hP10532s		Whole (aggreg
46	hP10504s		Nuclear	106	HP10116s	T-cell activation protein (PGR1)	Whole (large ce
47	hP10648s	Hypothetical protein Y40B1B.7-like	Nuclear	107	hP10633s		Whole (weak)
48	hP10239s	Hypothetical protein x 0001	Nuclear	108	hP10503s		Whole (weak)
49	HP10509s	Postsynaptic protein CRIPT-like	Nuclear	109	hP10637s		Whole & aggreg
50	hP10559s	Hypothetical protein KIAA0276-like	Nuclear	110	hP10437s	pp21 homolog-like	Whole & nuclea
51	hP10662s		Nuclear	111	hP10077s		Whole & nuclea
52	hP10626s		Nuclear & cytosol	112	hP10681s		Whole & particl
53	hP03185s	Histone macroH2A1-like	Nuclear & nucleol	113	hP02901s	Hypothetical protein-like (rp BL18)	Whole & particl
54	hP10416s		Nuclear & particle	114	hP10562s	Transcription factor LZIP-like	Whole & particl
55	HP00718s	High mobility group-like nuclear protein	Nuclear?	115	hP10431s		Whole and aggr
56	hP02644s	Probable ATP-dependent RNA helicase	Nucleolar	116	HP10458s	Sorting nexin 5 (SNX5)	Whole and aggr
57	HP10275s	E2IG3 (contains ATP/GTP binding dom	Nucleolar	117	hP10533s		Whole and aggr
58	hP10558s		Nucleolar	118	hP10178s	Small nuclear ribonucleoprotein associa	Whole and nucl
59	hP10561s		Nucleolar	119	hP10332s		Whole or Golgi
60	hP03115s	Midline 1 cerebellar isoform 1-like (Ring	Particle	120	hP10543s	Leucine-rich-domain interacting protein	Whole or Nucle

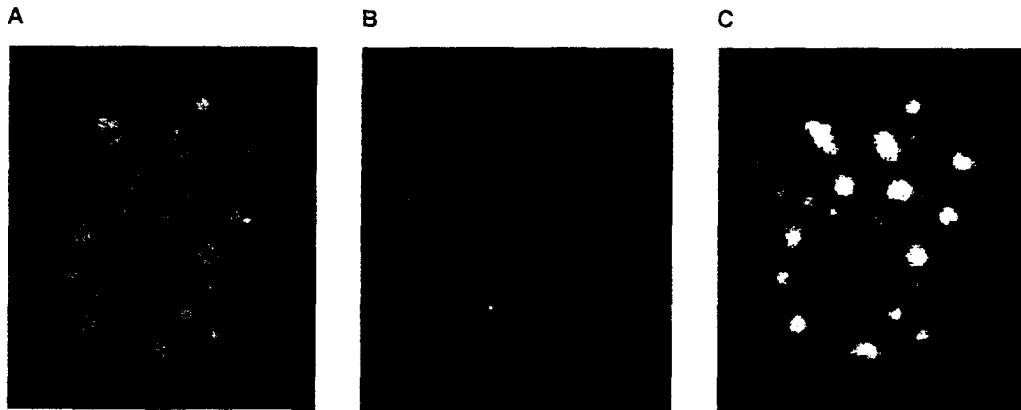


Fig. 1. Transiently expressed EGFP-Nps20 fusion protein is co-localized with SC35 at spliceosome.

A: EGFP-tagged Nps20 (Green). B: Anti-SC35/rhodamine (Red). C: Merged image of A and B. Cell: COS-7.

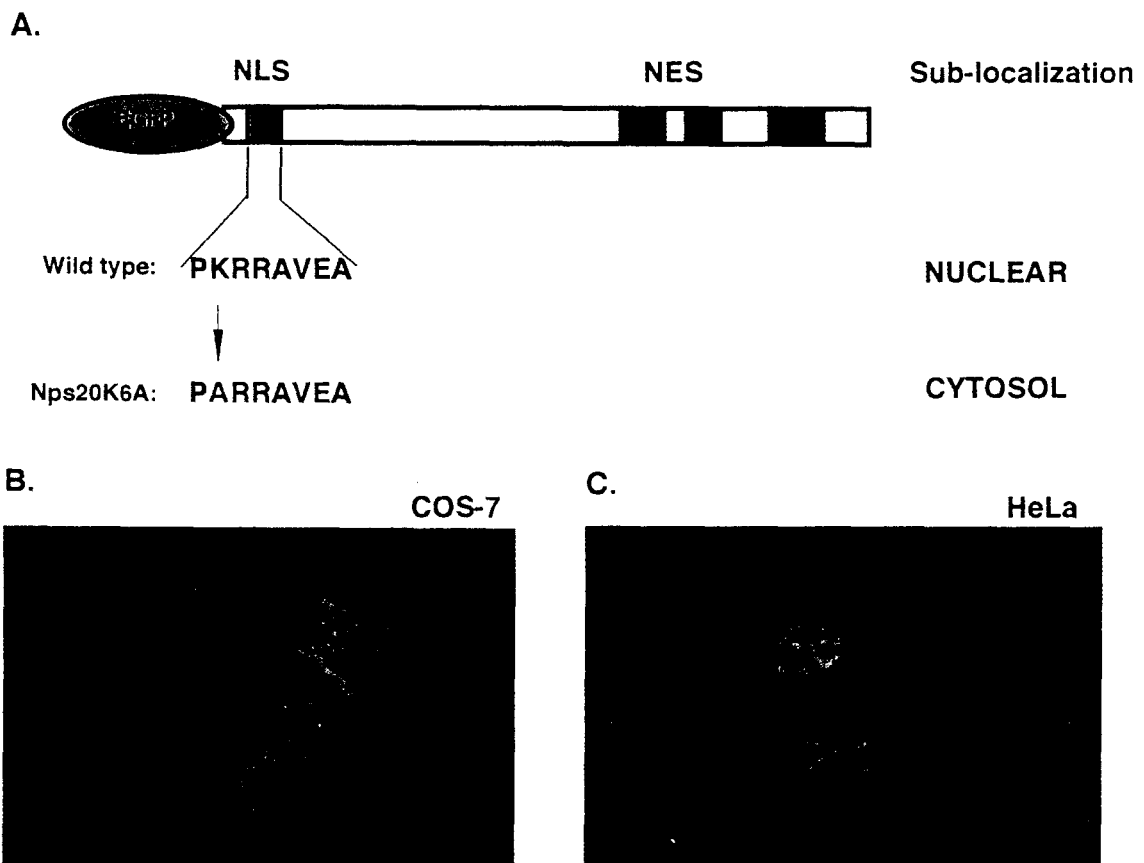


Fig.2 The N-terminal lysine at position 6 is essential for nuclear sub-localization of Nps20.

Transiently expressed EGFP-Nps20K6A is located in the cytosol. Green: EGFP-Nps20K6A. Red: SC35/rhodamine.

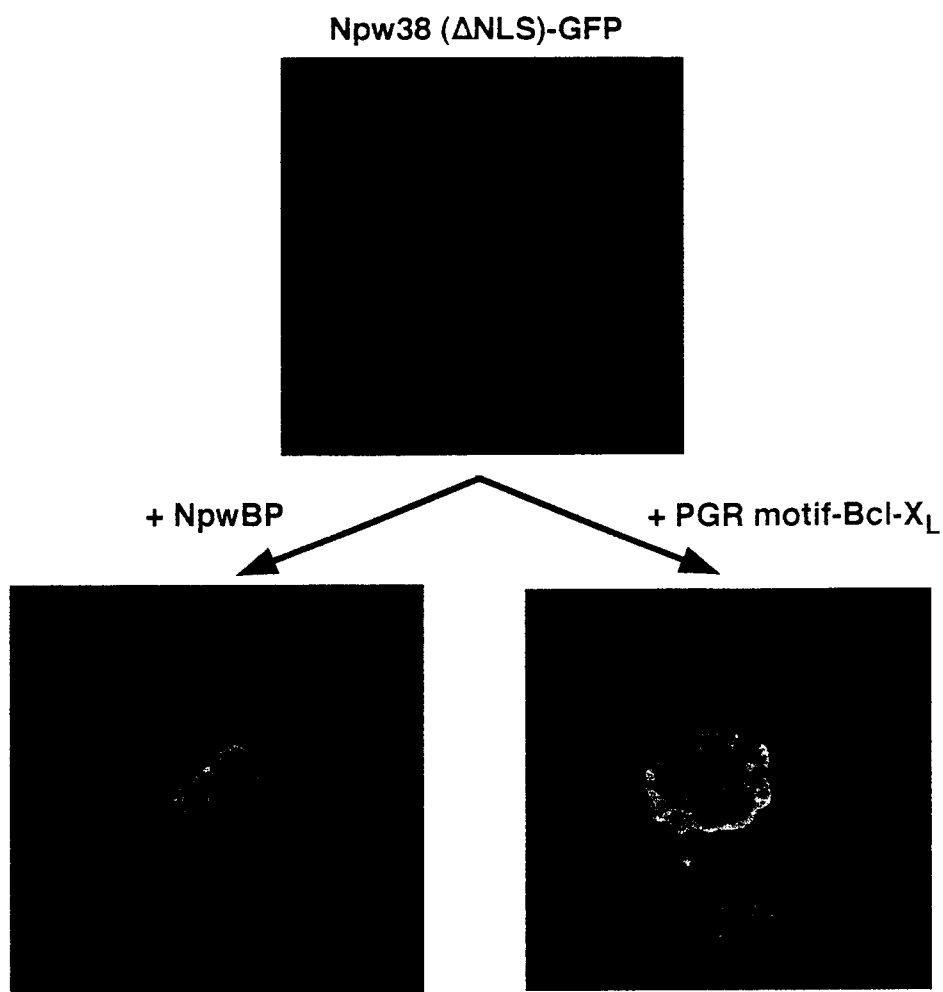


Fig. 3. Localization change induced by protein-protein interaction.

表 2 共発現の組み合わせと局在部位

X-GFP	Y	Localization
Npw38 (Δ NLS)	-	Whole
Npw38 (Δ NLS)	NpwBP	Nuclear
Npw38 (Δ NLS)	PGR motif-Bcl-X _L	Mitochondria
PGR motif	-	Whole
PGR motif	Npw38	Nuclear
PGR motif	Npw38 (W52A)	Whole