BMC Cell Biology

Methodology article

Protein-trap version 2.1: screening for expressed proteins in mammalian cells based on their localizations.

Olga O Sineshchekova, Toshimitsu Kawate, Oleksandr V Vdovychenko and Thomas N Sato*

Address: The University of Texas Southwestern Medical Center at Dallas, Dallas, Texas 75390, USA

Email: Olga O Sineshchekova - Olga.Sineshchekova@UTSouthwestern.edu; Toshimitsu Kawate - Toshimitsu.Kawate@UTSouthwestern.edu; Oleksandr V Vdovychenko - Oleksandr.Vdovychenko@UTSouthwestern.edu; Thomas N Sato* - island1005@aol.com

* Corresponding author

Published: 02 February 2004

BMC Cell Biology 2004, **5**:8

Received: 04 September 2003

Accepted: 02 February 2004

This article is available from: http://www.biomedcentral.com/1471-2121/5/8

Abstract

Background: "Protein-trap" is a method that allows epitope-tagging of endogenous proteins. This method allows for the identification of endogenously expressed proteins that exhibit specific localization of interest. This method has been recently reported for its application in the study of *Drosophila* development by using a relatively large epitope, green-fluorescent-protein (GFP).

Result: Herein, we report a new "protein-trap" vector for mammalian cells. This new method utilizes a much smaller epitope-tag and also allows for drug-selection prior to the epitope-tagging. Pre-selection by drug resulted in the highly efficient protein-trapping frequency.

Conclusion: The "protein-trap" method based on this new vector is expected to serve as a complimentary approach to the previously reported GFP-based method.

Background

Protein-trap is a method that allows for the identification of proteins of interest based on their unique subcellular localization without the use of specific antibodies to each protein. Protein-trap is a complimentary approach to currently available gene screening methods that are all based on the detection of global changes in gene expression at a genome level [1-3].

In another class of gene-expression based approach is gene-trap. This method uses reporters such as lacZ and GFP to tag, mutate and identify insertions into endogenous genes [4-6]. In addition to the regulation of gene expression, protein localization also plays a critical role in many biological processes. Despite such importance, a method was lacking that allows for the detection of a family of proteins based on their unique cellular localization in a relatively global manner. The modification of original gene trap approaches was used to identify specifically secreted or transmembrane proteins [7]. However, more universal method to screen proteins based on their subcellular localization would be useful. Such a method would require a large panels of specific antibodies to each protein. This is theoretically possible but is not a trivial problem in practice. Protein-trap was invented to overcome this problem.

The concept of protein-trap was originally proposed by Smith and Jarvik et al [8,9]. The principle is to tag proteins by an epitope and localize epitope-tagged proteins by using specific antibodies against the epitope. This concept was later adapted using GFP as a tag, thus eliminating the use of antibodies [10-12]. However, these original versions (Version 1.x; v1.x) of protein-trap schemes use

cDNA-GFP fusion library or genomic fragemnts fused to GFP. Therefore, they do not offer useful information about the regulation of endogenous protein expression. Most recently, the second version (Version 2.0; v2.0) of protein-trap strategy was reported by introducing a GFP based protein-trap vector into the genome of *Drosophila* [13]. This strategy allowed the GFP epitope-tagging of endogenous proteins in vivo. Thus, the protein-trap v2.0 can be used to identify proteins of interest based on their subcellular localization in *Drosophila* without the use of specific antibodies to each protein.

In this report, we describe a modified version (Version 2.1; v2.1) of a protein-trap scheme that utilizes smaller epitope-tag (43 amino acids length) and allows for the trapping of endogenously expressed proteins in mammalian cells. This new epitope-tag (43 amino acids) is considerably smaller than GFP (approximatley 239 amino acids) used in the protein-trap v2.0. Herein, we report the protein-trap v2.1 scheme and its unique applications to the mammalian cells.

Results

Principle of protein-trap v2.1

Our protein-trap v2.1 method is schematically described in Fig. 1A. The principle of this method is based on epitope-tagging each protein by a small identifiable epitope sequence (43 amino acids) at a genome level. The usefulness of this principle was previously proposed, but has never been developed for the use for endogenously expressed proteins at a genome level [9].

The myc-epitope that is commonly used for tagging proteins is inserted as a single copy "mini-exon" into the genome (Fig. 1A). This mini-exon behaves like a normal exon as the epitope sequence is flanked by splice-acceptor and -donor sequences. When the mini-exon is inserted into an intron of an expressed gene, the protein product is expressed as a myc-tagged form. Therefore, both intracellular and extracellular localization of the protein can be identified by a commonly used monoclonal antibody against the myc-epitope, 9E10 (Fig. 1A).

The protein-trap v2.1 is designed for its specific application to mammalian cells. For this specific application, we have devised a vector called "EpiTag" (Fig. 1B). This vector uses smaller epitope sequences (43 amino acids) in the mini-exon than GFP-tag (approximately 239 amino acids), as the smaller epitope-tag is potentially less detrimental to the normal localization of the proteins [14].

The efficient epitope-tagging was accomplished by two modifications. First, the mini-exon used in EpiTag vector reads the identical myc-epitope-tag sequences in all three reading-frame, thus eliminating the use of three independent mini-exon constructs. The efficiency was further improved by incorporating a puromycin-resistant cassette in the vector. This vector ("EpiTag" vector) can be delivered into the genome of cells in culture. The cells that incorporated this EpiTag vector into expressed genes become puromycin-resistant (Fig. 1B). Therefore, each puromycin-resistant colony represents a single independent expressed gene. Approximately 600 puromycin-resistant colonies were typically obtained from the infection of 5×10^5 cells. Upon excision of the puromycin-resistant gene by Cre-recombinase, each gene becomes epitope-tagged by the myc-sequence (Fig. 1B).

Retrovirus vector was used to enhance the probability of integration as a single-copy. In order to provide an estimate on the percentage of successful single-site integration events, Southern blot analysis was performed (Fig. 1C). This Southern blot analysis indicates that approximately 80% of the clones represent single-site integration events.

Identification of epitope-tagged proteins by protein-trap v2.1

Using the EpiTag vector, we successfully established a library of puromycin-resistant clones. Cre-mediated excision of the puromycin-cassette and the staining of total 5,670 puromycin-resistant colonies with 9E10 antibody identified 596 epitope-tagged clones. This is approximately 10% efficiency. The identity of proteins that are epitope-tagged from a total of 24 myc-antibody positive puromycin-resistant clones were determined by 3'-RACE (Table 1). Some of the proteins were trapped multiple times (e.g., three independent clones each representing the trapping of the same protein, 40S ribosomal protein S12).

Intracellular proteins exhibited the known localization (Fig. 2). Examples of three known nuclear localized proteins are shown in Fig. 2A. The nuclear localization of the myc epitope-tagged proteins was confirmed by co-staining with DAPI (Fig. 2A). One cytoskeletal protein was trapped (beta-actin:ACTB) (Fig. 2B). The co-staining with FITC-phalloidin confirmed the cytoskeletal localization of the myc epitope-tagged actin protein (Fig. 2B). Examples of mitochondrial proteins are shown in Fig. 2C. Colocalization of epitope-tagged F1-beta-ATPase (ATP5B) was confirmed by staining with both anti-myc antibody and commercially available anti-ATP5B antibody (Fig. 2C). Another mitochondrial protein, phosphate carrier precursor isoform 1b (PC) also exhibited the known mitochondrial localization (Fig. 2C).

In addition to intracellular proteins, one secreted protein, plasminogen activator inhibitor, was also identified by the ELISA assay among 5,670 puromycin-resistant

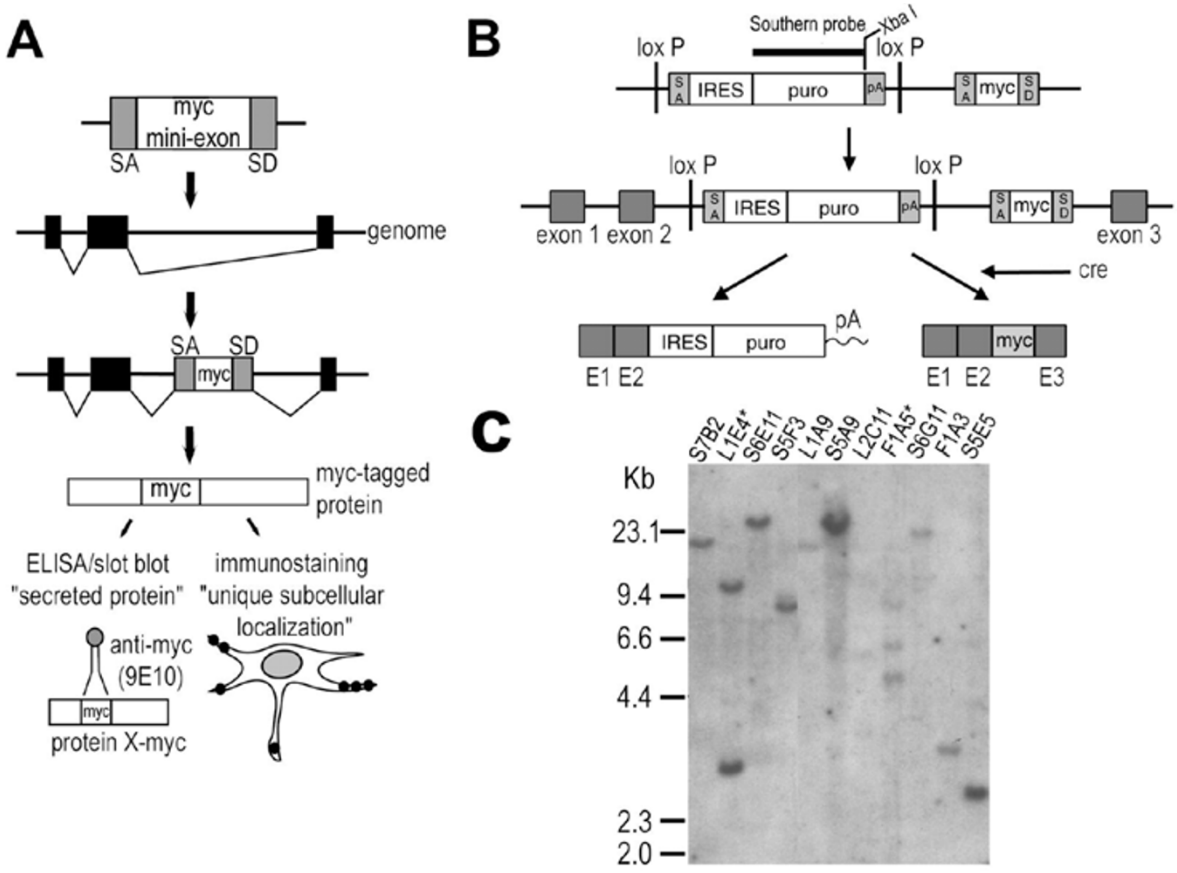

Figure 1

Principle of in vivo epitope-tagging method. **(A)** The concept of the method is schematically shown. The mini-exon consisting of a synthetic myc-epitope sequence flanked by splice-acceptor and -donor sequences is integrated into an intron of a gene as a single copy. This permits the incorporation of an epitope-tag sequence within a protein coding sequences. Secreted proteins can be detected by ELISA using 9E10 antibody, a commonly used antibody against the myc-epitope sequence [26]. Intracellular protein localization can be detected by immunostaining method using 9E10 antibody. **(B)** "EpiTag" vector and its mode of operation. EpiTag vector consists of a puromycin-resistant gene cassette flanked by a pair of loxP sequences that preceeds the mini-exon cassette. The puromycin-resistant cassette consists of a synthetic splice acceptor sequence followed by internal ribosomal entry sequence (IRES) fused to the puromycin-resistant sequence that is followed by poly-adenylation sequences. When this vector integrates into an intron, the endogenous protein is expressed as a truncated form but the puromycin-resistant marker is also expressed using the second translation start site in IRES. As a consequence, the cells survive in puromycin containing media. When the puromycin-resistant cassette is excised by Cre-recombinase, the protein is expressed as a full length and as a myc-epitope tagged form. Unique restriction enzyme site, XbaI, in the construct is indicated. The probe used for Southern blot analysis is also indicated. **(C)** Southern blot analysis to determine the number of integration events of EpiTag transgene per single cell (clone). Results from ten independent clones are shown. The XbaI-digested DNA was analyzed. Single-band indicates a single-site integration event. Multiple-bands indicate a single-copy but a multiple-site integration event. The clone numbers are indicated at the top and the ones that show a multiple-site integration event are indicated with (*). This result indicates approximately 80% of the clones represent single-site integration events of the transgene. S7B2: beta-actin; L1E4:beta ATP synthase; S6E11:CD44 antigen; S5F3:dual specificity protein phosphatase 5; L1A9:enolase alpha; S5A9:integrin alpha6; L2C11:phosphate carrier precursor isoform 1b; F1A5:similar to prokaryotic-type class I peptide chain release factors; S6G11:ribosomal protein L3; F1A3:40S ribosomal protein S12; S5E5:40S ribosomal protein S19.

Table 1: Examples of trapped proteins.

Protein name	Localization	AC#	INCIDENCE	MYC insertion
40S ribosomal protein S12	cytoplasm	X53505	3	E2/E3, E3/E4
40S ribosomal protein S19	cytoplasm	AF092907	1	E1/E2
Beta actin	cytoskeleton	M10277	1	E1/E2
Beta ATP synthase	mitochondrion	M27132	1	E1/E2
CD44 antigen	membrane	AJ251595	2	E1/E2
Dual specificity protein phosphatase 5	nucleus	U15932	1	E1/E2
Enolase alpha	cytoplasm	X16287	1	E2/E3
Eukaryotic translation initiation factor 1A	cytoplasm	L18960	1	E1/E2
Heterogeneous nuclear ribonucleoprotein A1	nucleus	X12671	3	E1/E2, E2/E3
Integrin α6	membrane	X53586	1	***
Metallothionein 2A	cytoplasm	J00271	2	E2/E3
Phosphate carrier precursor isoform 1b	mitochondrion	X60036	2	E3/E4
Plasminogen activator inhibitor, type 1	secreted	M16006	2	E2/E3
Regulator of chromosome condensation 1	nucleus	D00591	1	E5/E6
Ribosomal protein L3	cytoplasm	BC008492	1	E2/E3
Similar to prokaryotic-type class I peptide chain release factors	mitochondrion	BC011873	1	E6/E7

*** Insertion was in intron 1 and a cryptic translation of the intron sequence occurred possibly due to a splicing to a cryptic splice acceptor sequence in the intron.

colonies (Fig. 3). These results confirm that the protein-trap v2.1 scheme using EpiTag vector is effective in epitope-tagging both intracellular and extracellular proteins.

Sequence identification by RACE reactions shows the insertion of the myc mini-exon epitope in the expected exon/exon boundaries in proteins (Figure 4). For heterogeneous ribonucleoprotein, two independent insertion events exibit the identical localization pattern (Fig. 2A). This suggested that epitope-location within the full-length heterogeneous ribonucleoprotein A1 does not affect its normal localization.

Epitope-tagged proteins are synthesized as full-length proteins

The translation products of epitope-tagged ATP5B and ACTB was studied by Western blot analysis (Fig. 5). These two proteins were selected for the Western blot analysis since specific antibodies for these proteins that are suited for Western blot analysis are readily available. The Western blot analysis of ATP5B and ACTB confirmed that the epitope-tagged proteins are synthesized as expected full-length sizes (Fig. 5). Furthermore, the translation of epitope-tagged proteins is specifically dependent on Cre-mediated excision of puromycin-cassette (Fig. 5). The synthesis of full-length proteins is also reflected on the transcription of full-length mRNA based on the RACE reaction. Only exception to this fidelity of the system is the integrin α6 (Table 1). In the case of integrin α6, the epitope-encoding mini-exon was inserted into the intron 1. However, mini-exon is spliced into the cryptic splice-donor site present at the downstream part of the intron 1. This resulted in the truncation of normal integrin α6 protein.

Application of protein-trap v2.1 screening to identify protein translocations upon extracellular signals

Commonly used gene expression based screening identifies genes that are turned on and off in response to a variety of signals such as cell proliferation and differentiation, and exposure to chemical and physiological signal [1-3]. However, these differential gene expression based screening methods fail to identify proteins that are constitutively expressed but alter their localizations in response to such signals.

Protein translocation is known to play a critical role in a variety of biological processes. Identification of proteins that translocate in response to specific signals has been conventionally achieved by examining the translocation of each protein in response to specific signals. This approach requires targeted examination of specific proteins with specific prior hypotheses and predictions. It also requires the availability of useful antibodies to localize such proteins. To bypass these limitations, it would be extremely useful if a family of proteins that may play important biological functions by translocation can be identified without prior hypotheses and predictions. The success in this type of "reverse-genetic" approach provides more opportunities to study a variety of biological processes from different angles.

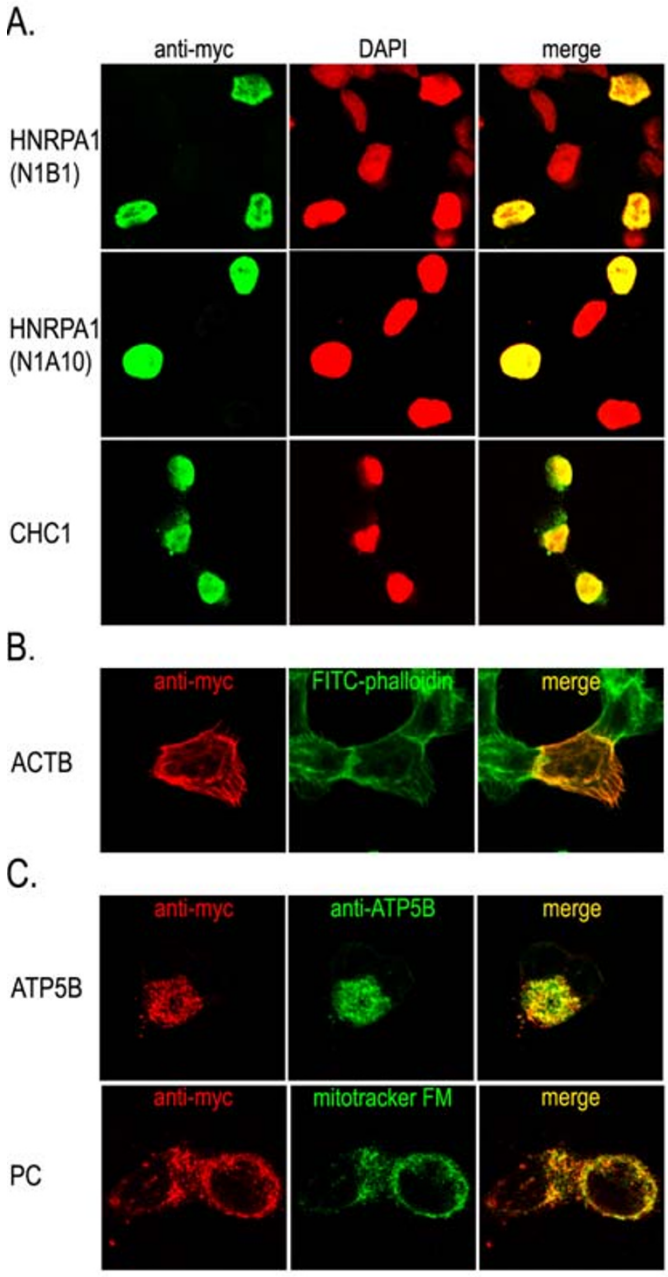

Figure 2
Examples of intracellular localized epitope-tagged proteins. **(A)** Nuclear localized proteins (Heterogeneous ribonucleoprotein A1: HRNPA1, chromosome condensation1: CHC1). Their nuclear localization was confirmed by co-staining with DAPI. In the case of HRNPA1, two clones that incorporated the epitope-tag in the two distinct part of the proteins were isolated. However, both of these epitope-tagged HNRPA1 (N1B1 and N1A10) exhibited the identical localization pattern. **(B)** Cytoskeletal protein. Beta-actin was epitoped tagged. The co-staining of the cells with anti-myc antibody and FITC-phalloidin confirmed the colocalization. **(C)** Mitochondrial proteins (F1-beta-ATPase: ATP5B, phosphate carrier precursor isoform 1b: PC). The epitope-tagged ATB5B and endogenous ATB5B proteins exhibited the identical localization as shown by co-staining the cells with anti-myc and anti-ATP5B antibodies. Since no useful antibodies against PC are available, the cells were co-stained with anti-myc and Mitotracker Green FM (Molecular Probes). Both anti-myc and Mitotracker FM showed identical staining, confirming the mitochondrial localization of epitope-tagged PC protein. All of the colocalization studies were preformed with confocal microscopy. Only the cells that underwent the successful Cre-mediated excision express the myc-tagged proteins and were stained with 9E10 antibody. The efficiency of Cre-mediated excision varied greatly depending on the target gene (<0.5% – 2%).

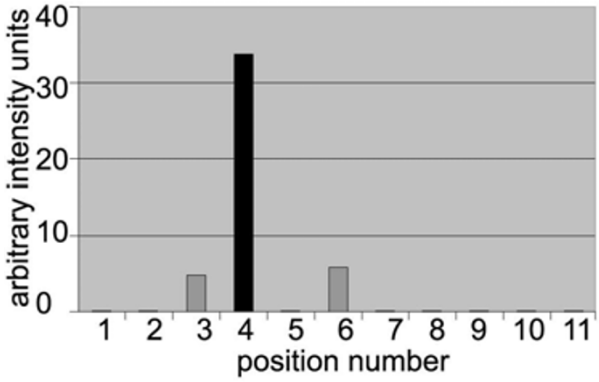

Figure 3
Trapping of secreted proteins. ELISA was used to detect secreted epitope-tagged protein in the media. Position number 4 in this case identifies a significantly high signal. Following the screening of total 5,670 clones, two clones gave significantly high ELISA signals. One was identified as plasminogen activator I (PAII). However, the other clone failed to yield consistent sequencing result following the 3'-RACE reaction. Thus the protein identity of this second clone remains unknown.

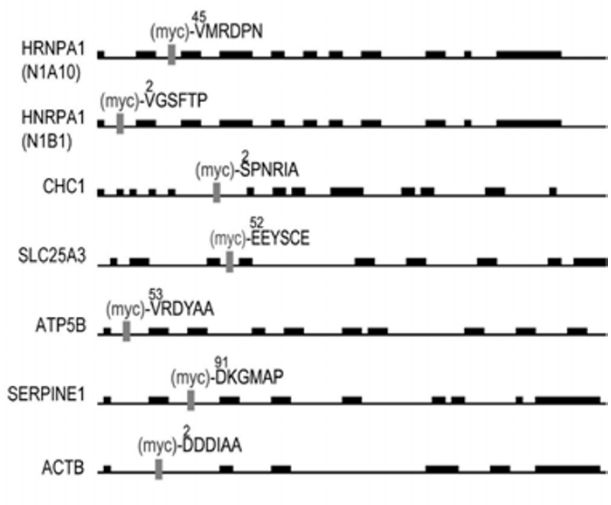

Figure 4
Epitope insertion site in each gene. RACE analysis confirmed that the epitope is inserted precisely at the known exon/exon boundaries in analyzed proteins. The mini-exon is indicated as a grey box in each gene. The position of the myc-epitope relative to the full-length protein for each gene is indicated by an amino acid position (indicated as a number). HRNPA1: heterogeneous nuclear ribonucleoprotein A1; CHC1: regulator of chromosome condensation 1; SLC25A3: phosphate carrier precursor isoform 1b; ATP5B: beta-ATP synthase; SERPINE1: type 1 plasminogen activator inhibitor; ACTB: beta-actin.

To test the applicability of our protein-trap v2.1 scheme in accomplishing this goal, we attempted to identify proteins that are constitutively expressed but alter their subcellular localizations in response to hypoxia. Hypoxia is one of the most critical biological signals for nearly all cells [15-17]. A large number of genes have been previously identified for their differential expression in cells in response to hypoxia [18-20]. However, very little is known about putative pathways that may not be directly linked to gene expression but yet play critical roles in cells' response to hypoxia. Therefore, we tested whether there are proteins that change their cellular localizations in response to hypoxia by using the protein-trap v2.1 (Fig. 6A).

Each puromycin-resistant colony following the insertion of the EpiTag vector was picked individually and arrayed onto 96-well culture plates (Fig. 6A). Master plates were stored and duplicate tester plates were subjected to Cre-recombinase mediated excision of the puromycin-resistant gene (Fig. 6A). One of the duplicate plates was cultured in normoxic and the other in hypoxic conditions (Fig. 6A). The myc-epitope tagged proteins were detected with the 9E10 antibody and immunofluorescence staining patterns were compared (Fig. 6A).

Using this screening scheme, we have identified a protein that changes subcellular localization upon hypoxia treatment (Fig. 6B). It is eukaryotic translation initiation factor 1A (EIF1A). In normoxia, its localization is patchy but ubiquitously distributed throughout the cytoplasm (Fig. 6B). In hypoxia, the EIF1A localization was specifically detected in a fibrous pattern similar to β-actin localization (Fig. 2).

To confirm that this staining pattern change is due to the protein translocation of the same protein, Western blot analysis was performed (Fig. 6C). Unavailability of an useful antibody for EIF1A forced us to resort to this indirect method. The Western blot analysis clearly shows that the protein of identical molecular weight is detected in the lysates prepared from the cells cultured in both normoxia and hypoxia conditions (Fig. 6C). Although further studies are necessary, this result provides evidence that EIF1A translocates upon hypoxia.

The identification of EIF1A as a protein that translocates in response to hypoxic environment was achieved without prior prediction. Therefore, this result implicates that our protein-trap v2.1 scheme can be adapted to discover

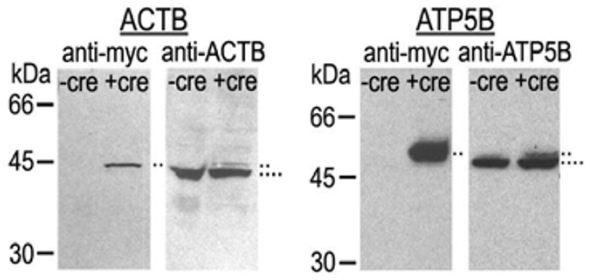

Figure 5
Synthesis of full-length epitope-tagged proteins. Epitope-tagged ATP5B and ACTB proteins were analyzed by Western blot to confirm the translation of full-length proteins. The Western blots were first probed with anti-myc antibody and secondly with anti-ATP5B or anti-ACTB following the stripping of the anti-myc antibody. Cell lysates before (-Cre) and after (+Cre) the Cre-mediated excision of the SA-IRES-puromycin cassette were analyzed. Specific myc-epitope signal (••) was detected from the (+Cre) cell lysates. No myc-epitope signal was detected from the (-Cre) cell lysates. Anti-ATP5B and anti-ACTB detected both endogenous ATP5B and ACTB proteins (••••), respectively, in addition to the epitope-tagged corresponding proteins (••). Slight size difference (approximately 3–4 kDa) between the epitope-tagged protein (••) and the endogenous counterpart of each protein (••••) reflects the addition of the mini-exon encoded epitope-tag. The "+cre" cell lysates were prepared from the cells that were transiently infected with the cre virus. Thus, they are mixtures of cells that express the tagged proteins and the cells that do not. As the cre-mediated excision efficiency is in the range of 0.5%-2%, the myc-tagged protein bands are always stained less intense. Molecular weight (in kDa) standards are also indicated.

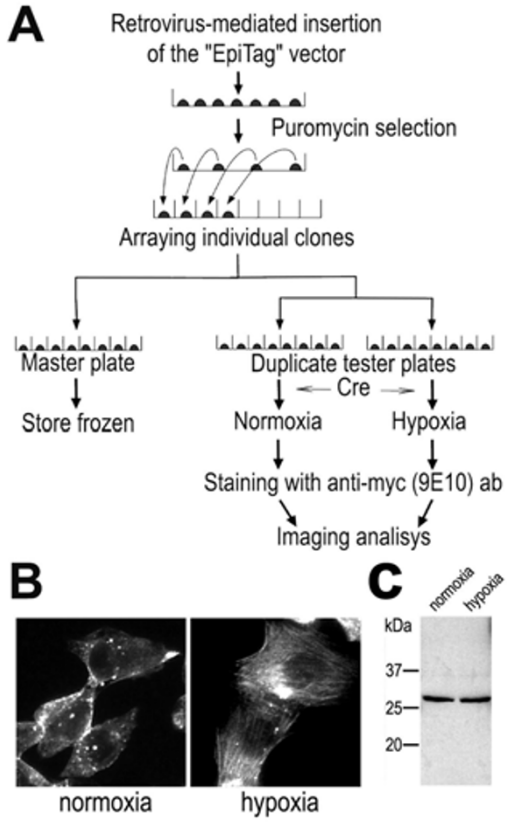

Figure 6
The screening for a family of proteins that translocates in hypoxia. **(A)** Screening scheme. A single copy of EpiTag vector is inserted into the genome by replication-defective retrovirus. Following integration of the vector, cells were selected by puromycin and each puromycin-resistant colony was picked and replated to 96-well cell culture plates individually. The master plates were stored frozen and the duplicate tester plates were generated. Both tester plates were subjected to the infection by another replication-defected retrovirus expressing Cre-recombinase. One of the tester plates was cultured in normoxia and the other tester plate was cultured in hypoxia. Following incubation in each environment for 18 hrs, cells were fixed and stained with 9E10 antibody and staining patterns were analyzed using Axiovert 200 M microscope (Carl Zeiss) equipped with a motorized automated stage and fluorescence filters using Open Lab Image analysis software (Improvision). **(B)** Translocation of epitope-tagged EIF1 in response to hypoxia. In normoxia, epitope-tagged EIF1 is localized ubiquitously in cytoplasm in a patchy manner. In hypoxia, its localization shifts to fibrous pattern resembling actin cytoskeleton. (**C**) Western blot analysis of the cell lysates. The blot was probed with anti-myc antibody. The protein with the same molecular weight was detected in the lysates prepared from the cells cultured in normoxia and hypoxia conditions. The size of the protein band is in accordance with the known molecular weight of EIF1A protein.

protein translocation events in response to a specific biological signal without any prior hypotheses or predictions.

Discussion

New feature of protein-trap v2.1

In this report, we described a modified protein-trap scheme for endogenous proteins. There are a few unique features with our protein-trap v2.1. First, our epitope-tag encoded by a mini-exon is considerably smaller (43 amino acids) as compared to GFP (239 amino acids). Although GFP is frequently used to tag proteins, a smaller tag such as myc may potentially have less effect on the normal protein localization. Furthermore, the myc-tag is commonly used as a tag for a variety of proteins.

The second unique feature is that our mini-exon encodes 43 amino acid peptides that contain the myc-epitope in all three reading frames without any interrupting stop

codons [9]. This reading-frame independent epitope-tagging eliminates the use of three protein-trap vectors.

The third feature is the incorporation of the puromycin-selection cassette into the vector. The incorporation of a puromycin-resistant cassette enables the cells to survive in puromycin-containing media when the cells incorporated the protein-trap vector in the expressed gene locus. This puromycin-cassette can be removed by a simple Cre recombinase transduction, which subsequently results in the myc-epitope tagging of the trapped protein (Fig. 1B). By this pre-drug selection, we were able to achieve relatively efficient protein-trapping frequency (approximately 10% of the puromycin-resistant clones expressed detectable levels of tagged-proteins).

The last unique feature is the ability to trap both intracellular and extracellular proteins with our scheme (Fig. 2 &3). The myc epitope-tagged secreted proteins were detected through the use of two different antibodies for mini-exon encoded epitopes in the sandwich ELISA format (Fig. 3).

Potential applications of protein-trap v2.1 in mammalian cells

Our protein-trap v2.1 method can be adapted to discover proteins of interest based on a number of criteria in mammalian cell culture system. The most straightforward approach is to identify proteins simply by their unique cellular localizations. For example, one could use this method to identify proteins that exhibit interesting distributions in nucleus during cell division or at the cell-cell junction.

Our scheme can be also used to identify proteins that translocate in response to biological, physiological or chemical signals. As an example of this type of application, we showed the scheme to identify proteins that translocate in response to hypoxic condition (Fig. 5). In this report, we tested our system only with one signal, hypoxia. However, our system can be adapted to identify proteins that translocate in response to a large number of different signals.

Another application is to screen for agonists and antagonists causing the specific cellular localization of the protein by using an identified clone of cells that express a unique epitope-tagged protein. This would be useful especially when the regulated specific cellular localization of the protein is critical for its function.

In this paper, we have established proof-of-principle of a new protein-trap method with human lung carcinoma cell line, NCI-H1299. However, this method is expected to work in other cells. One of the most potentially useful cells is embryonic stem (ES) cells. ES cells can be maintained as undifferentiated and subsequently induced to differentiate to a variety of cell types in vitro. Therefore, it is possible to identify translocatable proteins upon ES cell differentiation.

Potential further improvements of protein-trap v2.1

Although our protein-trap v2.1 is useful for a number of reasons discussed above, there is room for further improvement in the future. Our currently described EpiTag vector is based on a retrovirus-mediated integration. However, the LTRs that remain in the genome could affect the expression of other genes near the integration site. Furthermore, the retroviral vector could integrate into "hot spots" in the genome. These preferential integration events could result in biased representation of the trapped proteins. Therefore, it would be useful to have an option to be able to use non-viral integration method such as transposon system in mammalian cells in the future [21].

The frequency of the protein-trapping events in our system is currently 1/10. This suggests that 9 out of 10 puromycin-resistant clones remain unstained for the epitope-tag. Although there could be some leakiness in the puromycin-selection, it is also possible that 9E10 antibody staining method to detect the epitope-tagged proteins may not be sensitive enough to detect all of the epitope-tagged proteins. When the abundance of the epitope-tagged proteins is low in the cell, the cells become puromycin-resistant but the amount of the tagged protein could be insufficient to be detected by 9E10 antibody. Therefore, it would be useful to invent a way to enhance the sensitivity of the immuno-detection method in the future.

The major advantage of the use of GFP as a protein-trapping tag is its autofluorescent nature of the molecule. This eliminates the use of antibodies to detect the trapped proteins, thus allowing visualization in living organisms. Our protein-trap v2.1 uses smaller epitope-tag but requires the use of antibody to detect the epitope-tagged protein. Therefore, it would be further useful to invent a mini-exon cassette that encodes a small amino acid sequences of which expression can be detected without a specific antibody.

Conclusions

Herein, we present a new protein-trap system that is specifically applicable to mammalian cells. This system allows for the discovery of proteins based on their unique localization. The protein-trap system is expected to compliment other genomic and proteomic approaches in search of proteins that play critical roles in a variety of biological and/or pathological processes.

Methods

Constructs and virus

The EpiTag vector was generated by fusing the splice acceptor (SA) originating from the intron/exon2 boundary of adenovirus (kindly provided by Phil Soriano) to the IRES::Puromycin^r::pA cassette of the pIRESpuro vector (Clonetech). This SA::IRES::puromycin^r::pA cassette was flanked by a pair of loxP sequences. The mini-exon cassette was kindly provided by Desmond Smith and is previously described [9]. The final construct containing puromycin and mini-exon cassettes was inserted into the XhoI site of the replication-defective retrovirus vector, pGen$^-$ (kindly provided by Phil Soriano), in reverse orientation.

The retrovirus vector expressing Cre-recombinase was generated by replacing the lacZ sequence of the LZR-SpBMN-Z (kindly provided by Gary Nolan) with the Cre-NLS (kindly provided by Stephen O'Gorman) [22,23]. Replication-defective retrovirus was packaged using the Phoenix cell line (kindly provided by Gary Nolan) as described before [24].

"EpiTag virus stock" was prepared by collecting the media from the Phoenix cells transiently transfected with the constructs as described above. "Cre virus stock" was prepared by collecting the media from the Phoenix cells that were stably transfected with the retrovirus vector containing the Cre-NLS described above.

Cell culture

NCI-H1299 cells (human lung carcinoma cell line) were cultured in 10%FCS, RPMI1640, penicillin, streptomycin and infected with EpiTag vector virus. NCI-H1299 cells (5×10^5 cells in a 10 cm plate) were infected with 2 ml of "EpiTag virus stock". Two days post infection, the media was replaced with the addition of 1.75 μg/ml puromycin. The cells grown in this selection media for 7–10 days and the puromycin-resistant colonies were picked and transferred to 96-well cell culture plates. The master plates were stored at -80°C and duplicate tester plates were subjected to infection by Cre virus to remove the SA::IRES::puromycin^r::pA cassette. Cells in each well of 96-well plates received 50 μl of "Cre virus stock". Following Cre virus infection, cells were cultured in normoxic or hypoxic conditions, and stained with 9E10 antibody to localize the myc-epitope tagged proteins. Concurrently, the culture media was assayed for the potential presence of secreted myc-epitope tagged protein by ELISA.

Hypoxia was achieved by infusing a gas mixture (95% N_2/5% CO_2) into an air chamber (Billups-Rothenberg) as described before [25]. The plates in the hypoxia chamber and the ones in normoxia (95% air/5% CO_2) were always cultured and assayed at the same time.

Immunofluorescence staining and ELISA

Cells were fixed and permeablized by 1% TritonX-100, 3.6% paraformaldehyde in PBS at 4°C for 10 min. Non-specific binding sites were blocked by 2% BSA, 0.1%Triton X-100. PBS and the myc-epitope tagged proteins were detected by using 9E10 antibody (conditioned media prepared from the cultured hybridoma cells purchased from Developmental Studies Hybridoma Bank) and Cy3-conjugated donkey anti-mouse IgG (Jackson ImmunoResearch).

For the detection of secreted myc-epitope tagged proteins, the conditioned media from each well was incubated in a 96-well ELISA plate (Nunc) coated with 5 μg/ml affinity purified rabbit anti-Ep2 epitope IgG. Ep2 is an epitope (EWSRSSSPRRTST) just outside of the myc-epitope generated from the mini-exon coding sequence [9]. Rabbit polyclonal anti-Ep2 antibody was prepared by immunizing rabbits with Ep2 peptide conjugated with carrier proteins. The bound myc-epitope tagged proteins were subsequently detected by 9E10 antibody and HRP-conjugated goat anti-mouse IgG (Jackson ImmunoResearch).

Western Blot Analysis

Cells were lysed in the SDS-PAGE sample buffer (2% SDS, 10% glycerol, 20 mM DTT, 65 mM Tris-Hcl, pH6.8, 0.05% bromophenol blue). The cell lysates were loaded onto SDS-PAGE and the proteins were transferred to Hybond-C Extra membrane (Amersham). The membrane-transferred proteins were probed with rabbit anti-myc polyclonal antibodies (Cell Signaling Technology) or corresponding antibodies against the specific endogenous proteins. The signal was detected by using ECL system (Amersham).

Southern blot analysis

Each genomic DNA was digested with XbaI. The probe was labeled with ^{32}P-dCTP using Prime-It II kit (Stratagene). The blot was hybridized and washed (final with 0.1XSSC, 0.1%SDS at 67°C) at the highest stringency.

Gene identification

The myc-epitope tagged gene was identified by 3'-RACE using outer and inner primers derived from the 3'-end of the mini-exon (outer primer TS862: 5'-GAAGCTCATCTC-CGAGGAGG-3', inner primer TS863: 5'-AGCTCATCTC-CGAGGAGGAC-3') and RLM-Race kit (Ambion).

Author's contributions

O.O.S. and T.K. carried out establishing experimental system described in this manuscript. O.V.V. carried out a part of the cloning and expansion of protein-trap clones. T.N.S. conceived the idea of "protein-trap version 2.1" in 1997 and designed the vector and experimental strategy.

Acknowledgements

We thank Eric Biesterveld for setting up the ELISA system, and Junko Kuno for assisting the construction of vectors, and Drs. D. Smith, P. Soriano, G. Nolan, and S. O'Gorman for reagents. TNS would also like to acknowledge lively discussion with Emery Kalinov (SURF student from UT Austin) during the early phase of this project. We would also like to thank Dr. Steve McKnight for his critical reading of the manuscript. This work was supported by NIH.

References

1. Wan JS, Sharp SJ, Poirier GM, Wagaman PC, Chambers J, Pyati J, Hom YL, Galindo JE, Huvar A, Peterson PA, Jackson MR, Erlander MG: **Cloning differentially expressed mRNAs.** *Nat Biotechnol* 1996, **14:**1685-1691.
2. Claverie JM: **Computational methods for the identification of differential and coordinated gene expression.** *Hum Mol Genet* 1999, **8:**1821-1832.
3. Diatchenko L, Lau YF, Campbell AP, Chenchik A, Moqadam F, Huang B, Lukyanov S, Lukyanov K, Gurskaya N, Sverdlov ED, Siebert PD: **Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries.** *Proc Natl Acad Sci U S A* 1996, **93:**6025-6030.
4. Stanford WL, Cohn JB, Cordes SP: **Gene-trap mutagenesis: past, present and beyond.** *Nat Rev Genet* 2001, **2:**756-768.
5. Joyner AL: **Gene targeting and gene trap screens using embryonic stem cells: new approaches to mammalian development.** *Bioessays* 1991, **13:**649-656.
6. Lai Z, Han I, Park M, Brady RO: **Design of an HIV-1 lentiviral-based gene-trap vector to detect developmentally regulated genes in mammalian cells.** *Proc Natl Acad Sci U S A* 2002, **99:**3651-3656.
7. Skarnes WC, Moss JE, Hurtley SM, Beddington RS: **Capturing genes encoding membrane and secreted proteins important for mouse development.** *Proc Natl Acad Sci U S A* 1995, **92:**6592-6596.
8. Jarvik JW, Adler SA, Telmer CA, Subramanian V, Lopez AJ: **CD-tagging: A new approach to gene and protein discovery and analysis.** *BioTechniques* 1996, **20:**896-904.
9. Smith DJ: **Mini-exon epitope tagging for analysis of the protein coding protential of genomic sequence.** *BioTechniques* 1997, **23:**116-120.
10. Cutler SR, Ehrhardt DW, Griffitts JS, Somerville CR: **Random GFP::cDNA fusions enable visualization of subcellular structures in cells of Arabidopsis at a high frequency.** *Proc Natl Acad Sci U S A* 2000, **97:**3718-3723.
11. Misawa K, Nosaka T, Morita S, Kaneko A, Nakahata T, Asano S, Kitamura T: **A method to identify cDNAs based on localization of green fluorescent protein fusion products.** *Proc Natl Acad Sci U S A* 2000, **97:**3062-3066.
12. Ding DQ, Tomita Y, Yamamoto A, Chikashige Y, Haraguchi T, Hiraoka Y: **Large-scale screening of intracellular protein localization in living fission yeast cells by the use of a GFP-fusion genomic DNA library.** *Genes Cells* 2000, **5:**169-190.
13. Morin X, Daneman R, Zavortink M, Chia W: **A protein trap strategy to detect GFP-tagged proteins expressed from their endogenous loci in Drosophila.** *Proc Natl Acad Sci U S A* 2001, **98:**15050-15055.
14. Griffin BA, Adams SR, Tsien RY: **Specific covalent labeling of recombinant protein molecules inside live cells.** *Science* 1998, **281:**269-272.
15. Semenza GL: **Perspectives on oxygen sensing.** *Cell* 1999, **98:**281-284.
16. Semenza GL: **Hypoxia-inducible factor 1: master regulator of O2 homeostasis.** *Curr Opin Genet Dev* 1998, **8:**588-594.
17. Semenza GL: **Expression of hypoxia-inducible factor 1: mechanisms and consequences.** *Biochem Pharmacol* 2000, **59:**47-53.
18. Bruick RK: **Expression of the gene encoding the proapoptotic Nip3 protein is induced by hypoxia.** *Proc Natl Acad Sci U S A* 2000, **97:**9082-9087.
19. Seta KA, Kim R, Kim HW, Millhorn DE, Beitner-Johnson D: **Hypoxia-induced regulation of MAPK phosphatase-1 as identified by subtractive suppression hybridization and cDNA microarray analysis.** *J Biol Chem* 2001, **276:**44405-44412.
20. Scandurro AB, Weldon CW, Figueroa YG, Alam J, Beckman BS: **Gene microarray analysis reveals a novel hypoxia signal transduction pathway in human hepatocellular carcinoma cells.** *Int J Oncol* 2001, **19:**129-135.
21. Ivics Z, Hackett PB, Plasterk RH, Izsvak Z: **Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells.** *Cell* 1997, **91:**501-510.
22. Kinsella TM, Nolan GP: **Episomal vectors rapidly and stably produce high-titer recombinant retrovirus.** *Hum Gene Ther* 1996, **7:**1405-1413.
23. O'Gorman S, Dagenais NA, Qian M, Marchuk Y: **Protamine-Cre recombinase transgenes efficiently recombine target sequences in the male germ line of mice, but not in embryonic stem cells.** *Proc Natl Acad Sci U S A* 1997, **94:**14602-14607.
24. Grignani F, Kinsella T, Mencarelli A, Valtieri M, Riganelli D, Lanfrancone L, Peschle C, Nolan GP, Pelicci PG: **High-efficiency gene transfer and selection of human hematopoietic progenitor cells with a hybrid EBV/retroviral vector expressing the green fluorescence protein.** *Cancer Res* 1998, **58:**14-19.
25. Brogi E, Schatteman G, Wu T, Kim EA, Varticovski L, Keyt B, Isner JM: **Hypoxia-induced paracrine regulation of vascular endothelial growth factor receptor expression.** *J Clin Invest* 1996, **97:**469-476.
26. Squinto SP, Aldrich TH, Lindsay RM, Morrissey DM, Panayotatos N, Bianco SM, Furth ME, Yancopoulos GD: **Identification of functional receptors for ciliary neurotrophic factor on neuronal cell lines and primary neurons.** *Neuron* 1990, **5:**757-766.