

國立清華大學

碩士論文

人類粒線體第一蛋白質複合體中
NDUFS7 次單元擁有粒線體與細胞核兩
種不同的分佈

**Human mitochondrial complex I NDUFS7
subunit has a dual distribution both in
mitochondria and nuclei**

所別: 分子醫學研究所

學號: 9680534

研究生: 周定芳 Ting-Fang Chou

指導教授: 高茂傑 Mou-Chieh Kao

中華民國九十八年七月

July 2009

Acknowledgments

經過兩年的奮鬥，終於完成了這本碩士論文。首先要誠摯的感謝我的指導教授高茂傑老師，從論文題目開始的選擇就給予我很大的自由度，讓我嘗試許多的方向。另外從高老師身上也學到很多研究觀念以及實驗設計的邏輯，往往激發出我許多新的想法，尤其老師對於文獻探討所下的工夫，更另我佩服。而老師對學問的嚴謹態度以及做事的認真態度都是我們所應該學習的對象。除此之外，這本論文能以英文的形式完成，一切都得感謝老師幫忙。

另外本論文的完成還需感謝周美智老師以及張壯榮老師，感謝兩位老師能在百忙之中抽空擔任我的論文指導老師，並且給與許多寶貴意見與看法，以及發現一些實驗設計上的盲點，將使得這本論文更加完備，也規劃出本實驗將來所能延伸的方向。

感謝韻涵學姐及建師、凱惇、義心與俊仁學長所建立的實驗室，有了學長姐的努力，才使得我有良好的實驗操作環境。也要感謝與我今年一起畢業的同學們，認真溫柔且唯一能管得動我的欣瑜、平時陪我聊棒球的博淳、常麻煩借東西的婉慈與瑩兒，有了這些夥伴的相陪與互相鼓勵，才能順利畢業。還要感謝仍在努力奮鬥的學弟妹們：于廷、綢鈺、卉盈、艾凌、富茗、忠凱，有了你們大家的R313，隨時都充滿快樂的氣氛，也因為有你們才能讓平淡無奇的研究生活添加一些樂趣，祝你們未來也能保持這愉悅的心情做實驗，順順利利的畢業。

最後也要感謝我的父母、爺爺、奶奶、妹妹、弟弟、叔叔與三位姑姑，有你們在背後對我默默的支持，才有寫這本論文的動力。此外本實驗還要感謝財團法人罕見疾病基金會對於科學研究的支持與肯定，未來我們也會繼續努力。

在此，將這本小小的碩士論文，獻給這一路上所有幫助或關心我的人，謝謝你們！

周定芳 謹上

僅在此感謝財團法人罕見疾病基金會提供之
博碩士論文之獎助



中文摘要

人類粒線體第一蛋白質複合體中的 NDUFS7 (NADH dehydrogenase ubiquinone Fe-S protein 7) 在演化上具有高度保留的特性，並在氧化磷酸化系統中 (oxidative phosphorylation system, OXPHOS) 扮演重要的角色。NDUFS7 含有一個[4Fe-4S]的鐵硫中心 (iron-sulfur cluster N2, tetranuclear)，其功能為擔任第一蛋白質複合體中電子傳遞的最後電子接受者。於臨床研究中發現三種 NDUFS7 的突變與 Leigh 症候群 (Leigh syndrome, LS) 有關。NDUFS7 是由細胞核基因組所轉錄，於細胞質轉譯成蛋白質後再進入粒線體，並組裝於第一蛋白質複合體的親水區塊。在 T-REx-293 細胞內，利用核糖核酸干擾技術(RNAi)來抑制 NDUFS7 的表現，而 NDUFS7 表現下降會導致細胞在含有乳糖的培養基中生長緩慢以及細胞中 H₂O₂ 上升等表型。此結果顯示 NDUFS7 在粒線體能量代謝及細胞的生長扮演重要角色。多數表現在粒線體基質的蛋白質是在細胞質中合成，且藉由粒線體膜上的 TIM/TOM 蛋白質複合體辨認本身攜帶的粒線體標的訊號 (mitochondrial targeting signal, MTS) 來運輸。經預測軟體分析，NDUFS7 的 N 端前 29 個胺基酸可能為 MTS 的所在。在本次實驗中，我們建構出刪除不同片段大小的 NDUFS7 與綠螢光蛋白 (EGFP) 融合的質體，來觀察其在 T-REx-293 細胞內的表現位置。我們發現 NDUFS7 前 60 個胺基酸的片段即具有良好運送 EGFP 至粒線體的能力，顯示 MTS 位於此蛋白的 N 端。接著我們藉由定點突變的方法來研究鹼性與疏水性胺基酸對於此 MTS 的功能影響，發現這兩類胺基酸對於 NDUFS7 MTS 的運輸能力皆很重要。另外，我們也定義出 NDUFS7 的 C 端區域含有一段有效的細胞核輸入訊號 (nuclear localization signal, NLS) 以及細胞核運出訊號 (nuclear export signal, NES)。由這些結果可證明 NDUFS7 在細胞中擁有粒線體與細胞核兩種不同的分佈。

Abstract

Human NADH dehydrogenase (ubiquinone) Fe-S protein 7 (NDUFS7), one of the most conserved core subunits of mitochondrial complex I, plays an important role in the oxidative phosphorylation system (OXPHOS). This protein binds one iron-sulfur cluster N2 (tetranuclear) which is the terminal redox center in the electron transport process of complex I. Three types of mutations in this subunit have been associated with Leigh syndrome (LS). NDUFS7 protein is encoded by nuclear genome and incorporated in the peripheral segment of complex I facing the mitochondrial matrix. The NDUFS7 was suppressed in human T-REx-293 cells using the RNA interference (RNAi) technology. The reduction in the NDUFS7 expression caused a slow growth rate in galactose containing medium and increased H₂O₂ generation. These results indicated that NDUFS7 may play an important role in cell energy production and survival. Most mitochondrial matrix proteins are synthesized in the cytoplasm and imported into mitochondria by TIM/TOM complexes recognizing the mitochondrial targeting sequences (MTSs). Using prediction softwares, the N-terminal fragment was suggested to be the MTS of NDUFS7. In this study, we fused NDUFS7 containing N-terminal deletions, C-terminal deletions or different portions of the protein to enhanced green fluorescent protein (EGFP), and used these obtained constructs to study their intracellular localization in T-REx-293 cells. We found that the chimeric NDUFS7₁₋₆₀-EGFP was colocalized with mitochondria, demonstrating that this N-terminal fragment contained an effective MTS. We then used site-directed mutagenesis analyses to study the role of basic and hydrophobic residues in the first 60 amino acids of NDUFS7. These results suggested that both two types of amino acids are important for the mitochondrial import of the MTS in NDUFS7. In addition, we also demonstrated that there is a nuclear

localization signal (NLS) and a nuclear export signal (NES) located in the C-terminus of NDUFS7. These results suggested that NDUFS7 has a dual distribution both in mitochondria and nuclei.



Table of contents

Acknowledgments	i
中文摘要	iii
Abstract	iv
Abbreviations	ix
Introduction	1
The background of mitochondrion	1
The relationship between mitochondrial evolution and mitochondrial genome	1
Mitochondrial complex I and the electron transfer chain	2
The background of NDUFS7	3
The association between NDUFS7 defect and Leigh syndrome	3
The function of NDUFS7	4
Mitochondrial targeting sequence	5
Nuclear localization signal	6
Nuclear export signal	7
Dual localization of protein in eukaryotes	8
Materials and methods	10
Results	17
The function of NDUFS7 subunit in mammalian cell	17
Mitochondrial targeting signal of NDUFS7	18
A competent NLS of NDUFS7 was located at the C-terminal region	21
The C-terminal region of NDUFS7 may contain a NES sequence	23
NDUFS7 has a dual distribution both in mitochondria and nuclei	25
Discussion	26

Tables	30
Figures	36
References	74
Appendixes	79



Figures

Figure 1. Overview of the five mitochondrial complexes of the OXPHOS	36
Figure 2. The RNA interference knock down effect on the mRNA and protein expression levels of NDUFV2	38
Figure 3. The Growth capacity of T-REx293 and NDUFV2 knock down 1F4 and 1F6 cell lines	39
Figure 5. ROS assay detecting the H ₂ O ₂ levels in T-REx293, 1F4 and 1F6 cells.....	40
Figure 6. Mitochondrial targeting signals (MTSs)	41
Figure 7. The secondary structure prediction of NDUFV2	42
Figure 8. α -helical wheel diagram of (A) the first 32 and (B) the last 27 amino acids of NDUFV2	43
Figure 9. The MTS of NDUFV2 was located at N-terminus	46
Figure 10. The effective MTS region in NDUFV2 was located at amino acids 1 to 60	50
Figure 11. The characteristic of MTS in NDUFV2 was identified by site-directed mutagenesis.....	55
Figure 12. A competent NLS of NDUFV2 is located at the C-terminal region	58
Figure 13. A competent NLS of NDUFV2 is located at the C-terminal region	63
Figure 14. The NLS of NDUFV2 (from <i>H. sapiens</i>) and NuoB (from <i>N. crassa</i>) can carry LacZ into nuclei.....	65
Figure 15. The C-terminal region of NDUFV2 may contain a NES	68
Figure 16. NDUFV2 has a dual distribution both in mitochondria and nuclei	71
Figure 17. Overview the signal peptides of NDUFV2	72
Figure 18. The SUMOylation may be the key regulation for the dual distribution of NDUFV2	73

Abbreviations

NDUFS7: NADH dehydrogenase ubiquinone Fe-S protein 7

OXPHOS: oxidative phosphorylation system

Leigh syndrome: LS

MTS: mitochondrial targeting signal

NLS: nuclear localization signal

NES: nuclear export signal

RNAi: RNA interference

EGFP: enhanced green fluorescent protein

ATP: adenosine triphosphate

Complex I: NADH-ubiquinone oxidoreductase

Complex II: succinate-ubiquinone oxidoreductase

Complex III: ubiquinone-cytochrome c oxidoreductase

Complex IV: cytochrome c oxidase

Complex V: ATP synthase

mtDNA: mitochondrial DNA

nDNA: nuclear DNA

rRNA: ribosomal RNA

tRNA: transfer RNA

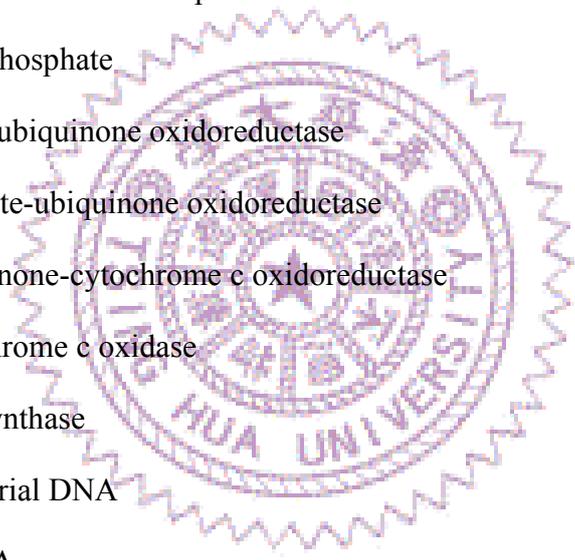
FMN: noncovalently bound flavin mononucleotide

EPR: electron paramagnetic resonance

BN-PAGE: blue native PAGE

TIM: translocase of the outer-membrane

TOM: translocase of the inner-membrane



MPPL mitochondrial-processing peptidases

IMP: inner membrane peptidase

MIP: mitochondrial intermediate peptidase

pALDH: aldehyde dehydrogenase

DHFR: dihydrofolate reductase

2-D: 2-dimensional

LMB: Leptomycin B

PTS: peroxisomal targeting sequence

ERTS: endoplasmic reticulum targeting sequence

TS: targeting sequences

FPGS: Folylpoly-gamma-glutamate synthetase

UNG: human uracil-DNA glycosylase

GR: glutathione reductase

DMEM: Dulbeccos modified Eagle medium

3' UTR: 3' unstranlation region

GAPDH: glyceraldehyde 3-phosphate dehydrogenase

carboxy-H₂DCFDA: 5-(and-6)-carboxy-2',7'- dichlorodihydrofluorescein diacetate

PBS: phosphate-buffered saline

PFA: paraformaldedyde

ECL: chemiluminescence

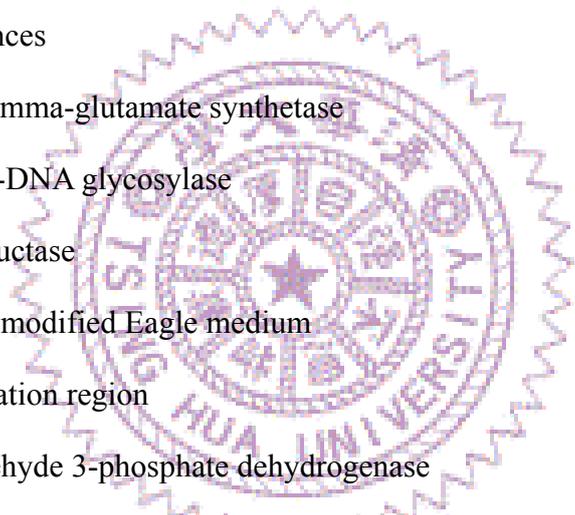
shRNA: short hairpin RNAs

LacZ: β-galactosidase

ROS: reactive oxygen species

H₂O₂: hydrogen peroxide

SUMO: small ubiquitin-like modifier



Introduction

The background of mitochondrion

Mitochondrion, stemming from the Greek “mito” and “chondrial” two words which mean “line” and “pellet” respectively, is a special organelle for energy metabolism in eukaryotic cells. This organelle is formed by two phospholipid membranes, outer- and inner-membrane, to separate three compartments which from outside to inside are called cytosol, inter-membrane space and matrix. There are many protein complexes on the inner-membrane which participate in various metabolic processes such as oxidative phosphorylation (OXPHOS), fatty acid oxidation and citric acid cycle. The inner-membrane is often protruded to form cristae which will increase the area of membrane surface and enhance overall enzyme efficiency. In OXPHOS, five enzyme complexes are involved in electron transport and adenosine triphosphate (ATP) production which are important for cell survival. These complexes include NADH-ubiquinone oxidoreductase (complex I), succinate-ubiquinone oxidoreductase (complex II), ubiquinone-cytochrome c oxidoreductase (complex III), cytochrome c oxidase (complex IV) and ATP synthase (complex V). The numbers of mitochondria are not constant but depend on the energy requirement of cells. In human cells, more than 90~98% of the ATP is generated by mitochondria, and the glycolytic source of ATP is less than 10% [1].

The relationship between mitochondrial evolution and mitochondrial genome

According to endosymbiotic theory, mitochondrion is likely the remnant of an oxidative prote-bacterium which is captured by endocytosis to be the partner in a symbiotic relationship with a proto-eukaryotic cell. This relationship has been

established for over a billion years. During the evolution process, the proto-bacterium has gradually lost most of its genes and these genes are transferred to the eukaryotic nucleus. Nowadays the human mitochondrial DNA (mtDNA) is about 16,569 base pairs in size and can conduct replication, transcription and translation by itself. It only contains 37 genes, including genes for 7 subunits (ND1, 2, 3, 4, 4L, 5 and 6) of complex I, 1 gene for subunit (cytochrome b) of complex III, 3 genes for subunits (COX1, 2 and 3) of complex IV, 2 genes for subunits (ATP6 and 8) of complex V, 2 genes for rRNAs (12S and 16S rRNA) and 22 genes for tRNAs. Therefore, except 13 mitochondria encoded subunits just mentioned, the vast majority of the over 1000 proteins of mitochondria are encoded by nuclear DNA (nDNA) [1-3].

Mitochondrial complex I and the electron transfer chain

Mitochondria complex I, the biggest and most complicated enzyme complex of OXPHOS is located on the inner-membrane and composed of 45 subunits in bovine. Among them 7 subunits are encoded by mtDNA and 38 genes including (NDUFV1-3, NDUFS1-8, NDUFA1-13, NDUFB1-11, NDUFAB1, NDUFC1-2 and 10.6 kDa) are encoded by nDNA. Complex I together with complex II are the two electron entries of the OXPHOS. Mammalian complex I oxidizes NADH and contains 8 iron sulfur (Fe-S) clusters for electron transfer. They include two binuclear Fe-S clusters (N1a and N1b) and 6 tetranuclear Fe-S clusters (N2, N3, N4, N5, N6a and N6b). The electrons are finally accepted by ubiquinone (Q). The likely electron transfer pathway from FMN (noncovalently bound flavin mononucleotide) to ubiquinone is FMN-N3-N1b-N4-N5-N6a-N6b-N2-Q. This electron transfer process also accompanies the pumping of protons across the inner-membrane from matrix to inner-membrane space thus generates proton concentration gradients. Finally,

complex V can use this proton-motive force to produce ATP [4].

The background of NDUFS7

Human NADH dehydrogenase (ubiquinone) Fe-S protein 7 (NDUFS7) has various names for homologues in different species, including PPST or 20 kDa in *Bos Taurus*, NuoB or 19.3 kDa in *Neurospora crassa*, NUKM in *Yarrowia lipolytica*, NQO6 in *Paracoccus denitrificans* and NuoB in *Thermus thermophilus* and *Escherichia coli* [5].

The first cDNA sequence of the NDUFS7 has been found in 1996 [NM_024407]. The coding sequence of NDUFS7 has 93% identity with that of the bovine homologous protein (PSST, 20 kDa). The NDUFS7 protein has 213 amino acids and 86.3% homology with PSST. This gene was identified to be localized on the short arm of chromosome 19 (19p13) by *in situ* hybridization [6].

The association between NDUFS7 defect and Leigh syndrome

Leigh syndrome (LS) is a genetic progressive neurodegenerative disorder caused by deficiency in the oxidative phosphorylation system. This disease was discovered and named by Denis Leigh in 1951, and patients suffering from this syndrome developed several phenotypes which included subacute necrotizing encephalomyelopathy, demyelination, spongiosis, gliosis, capillary proliferation and necrosis [7]. Mutations in both nuclear- and mitochondrial-encoded proteins which are related with mitochondrial energy metabolism could cause LS. Three types of mutations are found in LS: The first is mutation in mtDNA, like the gene encoded ATP6 subunit of ATP synthase. The second is mutation in nuclear DNA, like the gene encoded NDUFS8 (R102H). The last is deletion or insertion in both mtDNA and nuclear DNA [8, 9].

Three mutations in the NDUFS7 gene was found in patients with LS. Among them, two cases were missense mutations (V122M and R145H), and the other mutation (c.17-1167 C>G) could lead to activation of a cryptic exon. This abnormal splice site was predicted to produce a truncated protein of 41 residues which only the first five amino acids possessed the correct sequence [10-12]. In yeast model, it was shown that the mitochondrial membrane from a *Yarrowia lipolytica* strain carrying the mutation of NDUFS7 (V122M) caused complex I defects with decreased the V_{\max} value of complex I activity by about 50% [13].

The function of NDUFS7

NDUFS7, a subunit containing the last Fe-S cluster for electron transfer in complex I, can bind a tetranuclear N₂ which could be detected by mean of electron paramagnetic resonance (EPR) spectroscopy. It was demonstrated that the N₂ has the highest redox potential of -160 mV in complex I in *N. crassa* [14, 15]. The conserved motif for binding the [4Fe-4S] N₂ cluster is CCXXE(X)₆₀C(X)₃₀CP based on multiple sequence alignment of the NDUFS7 protein with various homologues [5]. The mutation on V122M in *Y. lipolytica* NUKM decreased the function of mitochondrial complex I activity and implied that this mutation might affect the formation of iron sulfur cluster [13].

NDUFS7 is a core subunit incorporated in the peripheral segment of complex I and faces the mitochondrial matrix. It locates at a critical position which connects the hydrophobic and hydrophilic parts of complex I [16]. The result from knockout studies of the 19.3 kDa (NuoB) in *N. crassa* showed that the loss of this subunit caused the defect in complex I assembly. The hydrophilic part of the complex was missing from the mutated mitochondria when samples were extracted by detergent. This observation suggested that 19.3 kDa is essential for the hydrophilic arm

assembly in complex I in *N. crassa* [15]. Mitoplasts extracted from skins fibroblasts of patients with the Leigh syndrome showed a truncated NDUFS7 protein of 41 amino acid residues when samples were separated by blue native PAGE (BN-PAGE) or SDS-PAGE. Similarly, fully assembled mitochondrial complex I was decreased in the BN-PAGE analysis with anti-GRIM19, NDUFA9, or NDUFS3 antibodies, but the assembly of complex II, III and IV was not affected [12]. In summary, all available data suggested that NDUFS7 is an iron sulfur (Fe-S) protein in the electron transfer chain and is important for the assembly of mitochondrial complex I.

Mitochondrial targeting sequence

Mostly cytosolic precursors (usually unfolded or loosely folded) of mitochondrial proteins need a mitochondrial targeting signal (MTS) to help their translocation into mitochondria. This signal will be recognized by the translocase of the outer-membrane (TOM) and the translocase of the inner-membrane (TIM) complex. Following of translocation, the majority of MTSs of these proteins were usually cleaved by mitochondrial-processing peptidases (MPPs) including inner membrane peptidase (IMP) and mitochondrial intermediate peptidase (MIP) to create the mature form of proteins [17-19]. For example, the mature form of yeast aldehyde dehydrogenase (pALDH) is located in mitochondrial matrix. It was found that the precursor form of this protein would be readily degraded when it was not correctly processed by MPP in the yeast system [20].

This signal peptide is majorly present at the N-terminus in most matrix and some inner membrane proteins. The other mitochondrial proteins which lack the N-terminal signals contain internal targeting signals for mitochondrial targeting. The internal signals are often present in outer-membrane proteins such as bcl-2, Tom5 or Fis1, and can not be cleaved. Infrequently, the MTS can be found to locate at the C-terminus.

For example, a yeast protein called Hm1p is a DNA helicase which has its MTS located at the C-terminus [17-19, 21].

The typical length of the N-terminal signal is about 10-80 amino acids. Both basic and hydrophobic amino acids are often enriched in the N-terminal targeting sequences to form amphipathic α -helices [18, 19]. In the case of mouse dihydrofolate reductase (DHFR) protein, it was demonstrated that amphiphilicity was necessary for mitochondrial import but forming α -helices may not be essential [22]. Artificial mitochondrial presequences have been synthesized and contained only arginine, serine and leucine, in addition to the first methionine to confirm the nature of MTS rules. These artificial signals had the ability of mitochondrial import when they were added to the N-terminus of the COX4 in yeast [23]. Interestingly, an artificial MTS fused to the C-terminus of the protein could still translocate this protein into mitochondria, but in the C to N direction [24].

In previous studies, complex I was purified from bovine heart mitochondria and PSST was separated and identified by 2-dimensional (2-D) gel. The amino acid sequence of mature form of this protein (identified by N-terminal sequence analysis) was Pro-Ser-Ser-Thr (PSST). Its molecular weight was estimated to be 20,077.5 Da [25]. The 1-38 amino acids of the precursor form was suggested to function as the signal peptide which could carry this protein to mitochondria. Similarly, the first 56 amino acids of NuoB (a homologue from *N. crassa*) was predicted to contain a putative MTS, and the molecular weight of the mature form was about 19,337 Da [26]. However, the location of signal peptide in human NDUFS7 has not been identified.

Nuclear localization signal

Some proteins with the molecular weight less than approximately 60 kDa can enter the nuclei without their nuclear localization signals (NLS) by diffusing passively.

In other words, proteins larger than this size would definitely need a NLS to help their entries into nuclei [27, 28]. Generally, the NLS has a small size and contains a rich number of basic amino acids including arginine (R) and lysine (K). In addition, some NLSs may also contain proline (P) or histidine (H). It was demonstrated that a NLS from SV40 T antigen fused with a cytoplasmic protein pyruvate kinase could translocate the passenger protein into nuclei [29].

Recently, based on the features of NLS, nuclear leader signals were categorized into three types. The first one like the classical type of the SV40 large T antigen contains a defined motif P-K-K-K-R-L-V [30-32]. This type of NLS contains 4 basic residues (K or R), or composes 3 basic residues (K or R) and one more residue which could either be P or H. The next one is the bipartite NLS, and the classical example is first found in the *Xenopus* protein nucleoplasmin. This type of NLS is composed of a fragment containing 2 basic residues as the starting and a segment containing 4 basic residues as the end. These two basic segments are separated by a spacer of 10 amino acids and making up a bipartite signal K-R-(X)₁₀-K-K-K-L [33].

The third class of NLS is identified according to the yeast protein Mat α 2 which contains a N-terminal signal, L-I-P-I-K [34]. This type of NLS does not have a NLS motif similar to those just mentioned, but contains basic residues and one or more hydrophobic residues in the NLS. The detailed characteristic of this group is still not clear. A family of transport receptors called karyopherins which can recognize the NLS of nuclear proteins and form the transport complex with the nuclear pore complex (NPC), which is a large protein complex and spans the nuclear envelope. Importin α/β is a heterodimer which belongs to a common family of karyopherins and can recognize the NLS of nuclear proteins in human [32, 35-38].

Nuclear export signal

In eukaryotic cells, nucleocytoplasmic transport of proteins plays an important role in regulating cell cycle, proliferation, transformation and tumorigenesis. Some proteins depend on a nuclear export signal (NES) to be exported from nuclei to cytoplasm. NES has a loosely conserved motif containing three or more hydrophobic leucine-rich residues, for example ⁷⁵L-P-P-L-E-R-L-T-L⁸³ in protein HIV Rev [39].

More than 75 NESs have been identified. Among them, p53, cyclin D1, HIV Rev and protein kinase A inhibitor are the most well known examples which contain a functional NES [39-41]. These NES are recognized specifically by chromosomal region maintenance (CRM1), a nuclear exporter protein participating in forming the transport complex [42]. Leptomycin B, a CRM1 inhibitor, can bind CRM1 to block the formation of transport complex [43]. Some proteins like Pho4 and Mig1 lack a hydrophobic NES but can still be transported from nuclei to cytoplasm by regulation through specific exportins (e.g. Msn5) with phosphorylation on the NES site [44-46].

Dual localization of protein in eukaryotes

After finishing human genome sequencing projects, the numbers of predicted genes are apparently smaller than scientists' original anticipation in various eukaryotes. One possible explanation for this finding is that one protein or a single gene product may have having different (or the same) functions in dual or multiple locations in cells. In recent reports, many cases that a single gene of which the derived products have dual or more subcellular localizations have been reported. The protein transports of these multiple targeting proteins to different intracellular organelles are executed by using mitochondrial targeting sequence (MTS), nuclear localization signal (NLS), nuclear export signal (NES), peroxisomal targeting sequence (PTS) or endoplasmic reticulum targeting sequence (ERTS).

Dual subcellular localization can be carried out by several mechanisms. A single

gene can get two translation products which may contain a different targeting sequence (TS) in each protein or only one has the TS but the other does not (distributed in cytosol) by dual transcription initiations (e.g. Folylpoly-gamma-glutamate synthetase, FPGS) [47], mRNA alternative splicing (e.g. human uracil-DNA glycosylase, UNG) [48] and multiple translation start sites (e.g. NifS) [49]. The other type is one product from one gene. The available examples include two different TSs on the same polypeptide (e.g. P4502B1), post-translational modification (e.g. WND) [50], inefficient/chimeric targeting (e.g. HMGCL and AMACR) [51, 52], protein-protein interactions (e.g. Apn1) [53], ambiguous signal (e.g. glutathione reductase from pea, GR) [54] and shuttle or export from mitochondrion (e.g. GOT2) [55].

According to MITOP2 program analyses, about 92% of the mitochondrial proteins are specifically localized in mitochondria without dual localization [56]. It means that most enzymes of energetic metabolism are restricted in mitochondria for specific functions. However, more than 31 human mitochondrial proteins that are derived from a single gene but have multiple distributions have been identified. For example, proteins participated in lipid or fatty acid β -oxidation in peroxisomes and mitochondria (e.g. HMGCL and AMACR) [51, 52], protein synthesis in cytosol and mitochondria (e.g. KARS) [57], DNA repair in nuclei and mitochondria (e.g. UNG and APE2) [48, 58-62].

Materials and methods

1. Cell culture

T-REx-293 cells (Invitrogen, Eugene, USA), human embryonic kidney cells with tetracycline-regulated expression system, were maintained in Dulbeccos modified Eagle medium (DMEM) with 10% fetal bovine serum (FBS), 100 U/ml penicillin and 100 µg/ml streptomycin at 37 °C.

2. Protein sequence alignment and analysis

The following protein sequences of NDUFS7 homologues were obtained from the NCBI database: *Homo sapiens* NDUFS7 (NP_077718), *Bos taurus* PSST (NP_001033111), *Neurospora crassa* 19.3 kDa (CAF06152), *Yarrowia lipolytica* NUKM (CAB65525), *Paracoccus denitrificans* NQO6 (P29918), *Thermus aquaticus thermophilus* NuoB (AAA97939) and *Escherichia coli* NuoB (YP_001731224). The retrieved multiple protein sequences of NDUFS7 homologues were aligned by EMBL-EBI ClustalW2 (<http://www.ebi.ac.uk/Tools/clustalw2/>), displayed by BOXSHADE server (http://www.ch.embnet.org/software/BOX_form.tml) and the percentages of sequence identity and similarity comparing to those collected sequences were analyzed by FASTA (http://fasta.bioch.virginia.edu/asta_www2/fasta_www.cgi?rm=lalign&pgm=lal).

3. Suppressed expression of NDUFS7 by RNA interference (RNAi)

Two knock down RNAi target sites designed by Clontech and invitrogen were located at 3' untranslated region (3' UTR) and the coding region of NDUFS7 gene. The oligonucleotide sequences of these two target sites were: CCG TGA GGT TGT

CAA TAA A (3' UTR) and GAG GAG GCT ACT ACC ACT A (coding region). This designed short hairpin oligonucleotides of *NDUFS7* were cloned into a pSUPER.neo+EGFP vector (Appendix 1) by *Hind* III (NEB, Ipswich, USA) and *Xho* I (NEB), and transfected into T-REx-293 cells with Lipofectamine 2000 (Invitrogen). The stable clone cells were selected in a medium containing 500 $\mu\text{g/ml}$ G418.

4. Real-time reverse-transcription polymerase chain reaction analysis

The total cellular RNA of cells was isolated by TRI reagent (molecular research center, Cincinnati, USA) and converted to cDNA by reverse-transcriptase from ABI High Capacity cDNA RT kit (Applied biosystems, Lincoln, USA). Specific primers were designed and synthesized to detect the expression of *NDUFS7* and glyceraldehyde 3-phosphate dehydrogenase (GAPDH) by real-time PCR. The primers used were: *NDUFS7*-F3: 5'-CTA CGA CAT GGA CCG CTT TG -3', *NDUFS7*-R3: 5'-CC GAG TAG GAA TAG TGG TAG -3', *GAPDH*-F2: 5'-GTT CGT CAT GGG TGT GAA CC -3' and *GAPDH*-R1: 5'-GCA TGG ACT GTG GTC ATG AG-3').

5. Growth measurements

2×10^4 cells were seeded on 6-well plates in the glucose containing medium (10% FBS, DMEM, 4.5 mg/ml glucose and 0.11 mg/ml pyruvate) and galactose containing medium (10% FBS, 0.9 mg/ml galactose and 0.5 mg/ml pyruvate). The number of cells was counted by trypan blue dye exclusion method for 8 consecutive days.

6. Reactive oxygen species (ROS) assay

Cells were seeded in cover glasses and incubated with 25 μM

5-(and-6)-carboxy-2'7'- dichlorodihydrofluorescein diacetate (carboxy-H₂DCFDA, Invitrogen) for 30 min. After washing once in phosphate-buffered saline (PBS), the cells were mounted by the VECTASHIELD Mounting Medium (Vector Laboratories, Burlingame, USA). Fluorescence was visualized by LSM510 confocal microscope.

7. Constructs

(1) Construction of plasmids expressing full-length and truncated NDUFS7 proteins

The pcDNA4/TO/*myc*-His A vector (Invitrogen) (Appendix 2) was used as the expression plasmid, and *Eco*R I (NEB) and *Xho* I (NEB) were employed as restriction enzyme sites for cloning. The pcDNA4-NDUFS7 vector constructed in a previous study in Dr. Kao's laboratory was used as the template for its various deletion constructs (NDUFS7 $_{\Delta 1-29}$ -Myc-His, NDUFS7 $_{\Delta 1-38}$ -Myc-His, NDUFS7 $_{\Delta 1-60}$ -Myc-His, NDUFS7 $_{\Delta 1-96}$ -Myc-His, NDUFS7 $_{\Delta 199-213}$ -Myc-His, NDUFS7 $_{\Delta 1-60/187-213}$ -Myc-His and NDUFS7 $_{\Delta 1-119/187-213}$ -Myc-His). The sequences of primers used are shown in Table 2.

(2) Construction of NDUFS7-EGFP and EGFP-NDUFS7

Various DNA fragments coding for N-terminal, internal or C-terminal peptides of NDUFS7 were designed and generated to fuse with *EGFP* gene in the pEGFP-N3 expression vector (Clontech Laboratories, Terra Bella Ave, USA) (Appendix 3). The restriction enzyme sites used for this purpose are *Xho* I and *Eco*R I (NEB). The pcDNA4-NDUFS7-*myc*-His was used as the template to generate the various different N-terminal proteins of NDUFS7. These constructs included: NDUFS7 $_{1-213}$ -EGFP (full-length NDUFS7), NDUFS7 $_{1-27}$ -EGFP,

NDUFS7₁₋₂₉-EGFP, NDUFS7₁₋₃₈-EGFP, NDUFS7₁₋₅₀-EGFP, NDUFS7₁₋₆₀-EGFP, NDUFS7₃₀₋₆₀-EGFP, NDUFS7₃₀₋₉₆-EGFP, NDUFS7₃₉₋₁₈₉-EGFP, NDUFS7₁₈₇₋₂₁₃-EGFP, NDUFS7₁₈₉₋₂₁₃-EGFP, EGFP-NDUFS7₁₋₂₁₃, EGFP-NDUFS7₁₋₆₀, EGFP-NDUFS7₁₈₇₋₂₁₃, EGFP-NDUFS7₁₉₉₋₂₁₃, EGFP-NDUFS7₁₉₉₋₂₀₆, NDUFS7_{Δ1-29}-EGFP, NDUFS7_{Δ1-38}-EGFP and NDUFS7_{Δ1-60}-EGFP. All of fragments in NDUFS7 were amplified using the pfu turbo DNA polymerase (Stratagene, USA). The sequences of primers used are shown in Table 2.

(3) Construction of NDUFS7 missense mutations

The pEGFP-N3-NDUFS7 was used as the template for mutation of basic, acidic and hydrophobic residues in the first 1-60 and 187-213 amino acids of NDUFS7 using site-directed mutagenesis methodology. These constructs included: NDUFS7_{1-60(L4Q)}-EGFP, NDUFS7_{1-60(L9Q)}-EGFP, NDUFS7_{1-60(I14T)}-EGFP, NDUFS7_{1-60(L15Q)}-EGFP, NDUFS7_{1-60(L17Q)}-EGFP, NDUFS7_{1-60(I14T/L15Q)}-EGFP, NDUFS7_{1-60(I14T/L17Q)}-EGFP, NDUFS7_{1-60(L15Q/L17Q)}-EGFP, NDUFS7_{1-60(L4Q/I14T/L15Q)}-EGFP, NDUFS7_{1-60(R10G)}-EGFP, NDUFS7_{1-60(R13G)}-EGFP, NDUFS7_{1-60(R18G)}-EGFP, NDUFS7_{1-60(R10G/R13G)}-EGFP, NDUFS7_{1-60(R10G/R18G)}-EGFP, NDUFS7_{1-60(R13G/R18G)}-EGFP, NDUFS7_{1-27(R10G)}-EGFP, NDUFS7_{1-27(L4Q)}-EGFP, NDUFS7_{1-27(L15Q)}-EGFP, EGFP-NDUFS7_{199-213(R199G)}, EGFP-NDUFS7_{199-213(K200G)}, EGFP-NDUFS7_{199-213(R199G/K200G)}, EGFP-NDUFS7_{199-213(K202G)}, EGFP-NDUFS7_{199-213(R203G)}, EGFP-NDUFS7_{199-213(K202G/R203G)}, EGFP-NDUFS7_{199-213(R205G)}, EGFP-NDUFS7_{199-213(R206G)}, EGFP-NDUFS7_{199-213(R205G/R206G)}, EGFP-NDUFS7_{199-213(R212G)}, EGFP-NDUFS7_{199-213(R213G)}, EGFP-NDUFS7_{199-213(R212G/R213G)}, EGFP-NDUFS7_{199-213(K200G/R206G)}, EGFP-NDUFS7_{199-213(R203G/R206G)}, EGFP-NDUFS7_{187-213(E204T)}, NDUFS7_{187-213(E204T)}-EGFP, NDUFS7_{189-213(E204T)}-EGFP, NDUFS7_{189-213(E204A)}-EGFP, NDUFS7_{189-213(E204L)}-EGFP, NDUFS7_{189-213(E204D)}-EGFP,

NDUFS7_{189-213(E204Q)}-EGFP, NDUFS7_{188-213(E188A/E204A)}-EGFP, EGFP-NDUFS7_{187-213(L195A/L197A)} and EGFP-NDUFS7_{187-213(L190A/L191A/L195A/L197A)}. All of mutations in NDUFS7 were made with the pfu turbo DNA polymerase (Stratagene) and the used primers are shown in Table 2.

(4) Construction of *N. crasa* NuoB

The fragment containing 212-226 residues of *N. crasa* NuoB was designed and generated by annealing two primers which contained BsrG I (NEB) and Not I (NEB) restriction sites at the two ends. This fragment was fused with *EGFP* gene in pEGFP-N3 expression vector (Clontech) and the resultant construct was named EGFP-NuoB₂₁₂₋₂₂₆. The used primers are the forward primer 5'- GTG TAC AAG AGG AAG ATG AGG AAT ACG AAG ATC ACG AGG ATG-3' and the reverse primer 5'- AGC GGC CGC CTA CTT GCG GTA CCA CAT CCT CGT GAT CTT CGT ATT-3'.

(5) Construction of LacZ-NDUFS7 and LacZ-NuoB

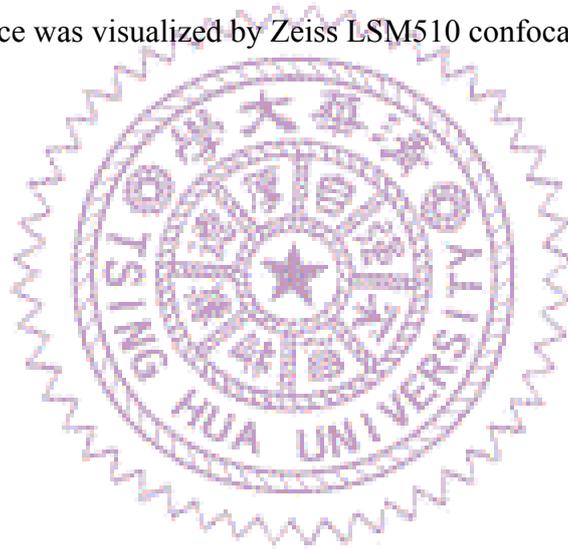
The pcDNA4/TO/*myc*-His/LacZ vector (Invitrogen) (Appendix 4) was used as the expression plasmid, and *Not* I (NEB) and *Xho* I (NEB) were employed as restriction enzyme sites for cloning. The EGFP-NDUFS7₁₉₉₋₂₁₃ and EGFP-NuoB₂₁₂₋₂₂₆ was used as the template to generate the fragment 199-213 residues of NDUFS7 and the fragment 212-226 residues of *N. crasa* NuoB, respectively, by primers containing *Not* I (NEB) and *Xho* I (NEB) restriction sites. The used primers are shown in Table 2.

8. Immunofluorescence staining with antibodies

T-REx-293 cells were trypsinized and seeded at approximately 50~70% confluency in 24-well plates containing cover glasses, and then transfected with TransIT-LT1 transfection Reagent (Mirus, Madison, USA). Twenty-four hours later, cells were induced with tetracycline at the concentration of 0.5 $\mu\text{g/ml}$ for 24 hours. After tetracycline induction, they were incubated for 30 min at 37 °C with 0.5 ml growth medium containing 100 nM Mito Tracker Red (CMX-Ros; Molecular probe, Eugene, USA) and then washed once in 0.5 ml PBS (phosphate-buffered saline). After this step, all samples should be protected from light. Cells were permeabilized and fixed with 0.5 ml acetone and methanol mixture (acetone/methanol = 3/1) for 5 min on ice. After fixation, specimens were moved to 6-well plates and blocked with 2 ml growth medium at room temperature (RT) for 2 hours. Samples were then incubated for 1 hours at RT with diluted primary antibody (1:50~1:100). After 5 times of PBS washing for 5 min each, cells were incubated for another 1 hour at RT with diluted secondary antibody (1:100), and washed again by PBS for 5 times. Finally, the cover glass was mounted with VECTASHIELD Mounting Medium (Vector Laboratories) at RT for 10 min. Fluorescence was visualized by Zeiss LSM510 confocal microscope. The primary detection was performed using monoclonal mouse anti-c-*myc* antibody (Calbiochem, San Diego, USA), monoclonal mouse anti-ATPase alpha subunit antibody (Santa Cruz Biotechnology, Delaware, USA) and rabbit polyclonal affinity purified anti-c-Myc antibody (Santa Cruz). The secondary detection was performed using goat anti-mouse IgG (H+L) FITC conjugate antibody (Calbiochem, San Diego, USA), goat anti-mouse IgG (H+L) rhodamine conjugate antibody (Molecular Probe) and anti-rabbit IgG (H+L) FITC conjugate antibody (Zymed, Eugene, USA).

9. Immunofluorescence staining with EGFP

T-REx-293 cells were trypsinized and seeded at approximately 50~70% confluency in 24-well plates containing cover glasses and transfected with TransIT-LT1 transfection Reagent (Mirus). Twenty-four hours later, the cells were incubated for 30 min at 37 °C with 0.5 ml growth medium containing 100 nM Mito Tracker Red (Molecular probe) and then washed once in 0.5 ml PBS. After this step, all samples should protect from light. Cells were fixed with 4% paraformaldehyde (PFA) in PBS for 15 min at RT, and cells were then permeabilized with 0.5 ml methanol for 5 min on ice. After washed with PBS for 5 min, the cover glass was mounted with VECTASHIELD Mounting Medium (Vector Laboratories) at RT for 10 min. The fluorescence was visualized by Zeiss LSM510 confocal microscope.



Results

The function of NDUF7 subunit in mammalian cell

NDUF7, one of the most conserved core subunits of mitochondrial complex I, is expected to play an important role during the oxidative phosphorylation (OXPHOS) process (Fig. 1). It has very high homologies with subunits from *Bos Taurus*, *Neurospora crassa*, *Yarrowia lipolytica*, *Paracoccus denitrificans* and, slightly less with *Thermus thermophilus* and *Escherichia coli* (Fig. 2 and Table 1). Therefore, it is a rational speculation that NDUF7 in human mitochondria has similar functions as those discoveries made in other organisms. However, the subunit compositions of mammalian complex I are much complicated than those of lower eukaryotic organisms or prokaryotes, and the results derived from lower order model organisms may not be totally applicable to higher order organisms or even human being. For this reason, we targeted at human NDUF7 in this study and would like to know more about its functions and mitochondrial targeting.

RNA interference (RNAi) methodology was first applied to knock down NDUF7 gene expression. Two different short hairpin RNAs (shRNA) targeting at the coding region and 3' UTR respectively, were designed and cloned into pSUPER plasmid (Fig. 3A). The derived vector was transfected into T-REx-293 cells of the tetracycline on system. After screening, several stable clone cell lines with mRNA levels of NDUF7 been knocked down were selected for farther studies. According to data obtained from real-time PCR studies, the mRNA expression levels of NDUF7 were suppressed efficiently by 81% and 99.1% in U19 and U21 cell lines respectively (Fig. 3B).

ATP can be produced from glycolysis in cytoplasm or OXPHOS in mitochondria in mammalian cell. However, between these two sources, the majority of ATP is

produced from mitochondria. If the cell has a functional defect in mitochondria, it will grow slowly and increase the ATP production in the glycolysis pathway for compensation. We measured the growth rate of cells cultured in DMEM medium containing high glucose only or galactose only to understand the effect of NDUFS7 knockdown on cell growth. The growth rates of both U19 and U21 cells were slower than those of control cells in glucose medium (Fig. 4A), and the differences were much more significant in galactose containing medium (Fig. 4B). Furthermore, knock-down cells did not survive in galactose containing medium when TTFA (complex II inhibitor) was added at concentration of 200 μ M to block the complex II source of electron entry (Fig. 4F).

The complex I not only can transfer electrons from NADH to ubiquinone, but is also a major source of production of reactive oxygen species (ROS). The generation of H_2O_2 , which is one of the ROS, in knock-down cells was detected by the addition of DCFH-DA. The image analyses from cofocal microscopy studies indicated that U21 cells produced a large number of H_2O_2 when they were compared with T-REx-293 cells (Fig. 5). These results revealed that the function of NDUFS7 is important in mammalian mitochondria for cell survival and defense.

Mitochondrial targeting signal of NDUFS7

Mitochondrial genome only contains 37 genes and most mitochondrial proteins are encoded by the nuclear genome. The mitochondrial targeting sequence (MTS) can lead these nuclear-encoded mitochondrial proteins into the mitochondria (Fig. 6). In previous studies, NDUFS7, a nuclear-encoded protein, was thought to be located only in the mitochondrial matrix part of complex I. However, the location and characteristics of MTS in NDUFS7 have not been detailedly analyzed in human or yeast. According sequence analyses, both N- and C-terminus of NDUFS7 had many

positive charged and hydrophobic residues capable of forming a potential amphipathic α -helix thus become an effective MTS (Fig. 2, 7 and 8). In order to investigate whether these two regions contain a MTS sequence, N- or C-terminal deletions of NDUF7 were fused with enhanced green fluorescent protein (EGFP), and the obtained constructs were used to study the intracellular localization of the resultant hybrid proteins.

Four different deleted constructs, named NDUF Δ_{1-38} -EGFP, NDUF Δ_{1-60} -EGFP and NDUF $\Delta_{199-213}$ -EGFP, contained deletions of the N-terminal 38, 60, or C-terminal 15 amino acid residues respectively, and the full length of NDUF $_{71-213}$ -EGFP were generated (Fig. 9A). These constructs were then transfected into T-REx-293 cells for 24 hours, and the subcellular localization of expressed fusion proteins and mitochondria was visualized by EGFP and Mito Tracker Red (dye for mitochondria) respectively. The full-length of NDUF7 could take EGFP to the mitochondria but the construct with EGFP itself only could not (Fig. 9B, a and b). Similar results were also observed in HeLa cell (Fig. 9B, f). This result indicated that NDUF7 may carry a signal peptide for mitochondrial targeting. When construct with deletion of the C-terminal 15 (Fig. 9B, e) amino acid residues in NDUF7 was fused with EGFP, this deletion did not affect the ability of mitochondria targeting of EGFP. However, the proteins expressed from two other deletion constructs, NDUF Δ_{1-38} -EGFP (Fig. 9B, c) and NDUF Δ_{1-60} -EGFP (Fig. 9B, d), were diffused throughout the cells. These observations indicated that the MTS of NDUF7 should be located at N-terminus of the protein.

To further investigate the minimum covering region of MTS in NDUF7, EGFP was also used as the report protein and fused with a NDUF7 fragment at its N-terminus or C-terminus. The constructs and the result of image analyses were outlined in Figure 10A. In summary, N-terminal 60 residues showed a similar

mitochondrial targeting efficiency to that of the full-length NDUFS7 (Fig 10A, f). Based on the results from the following constructs: NDUFS₁₋₂₇-EGFP (Fig. 10B, b), NDUFS₁₋₂₉-EGFP (Fig. 10B, c), NDUFS₁₋₃₈-EGFP (Fig. 10B, d), NDUFS₁₋₅₀-EGFP (Fig. 10B, e) and NDUFS₁₋₆₀-EGFP (Fig. 10B, f) it was concluded that within the range of N-terminal 60 amino acids residues, the longer sequence selected for MTS had a stronger mitochondrial targeting efficiency. All of the aforementioned constructs showed at least partial mitochondrial localization, whereas the construct with the N-terminal 19 residues of NDUFS7 (Fig. 10B, a) did not. Interestingly, all of the constructs expressing residues 30-60 (Fig. 10B, g), 30-96 (Fig. 10B, f) and 30-189 (Fig. 10B, i) in front of the N-terminus of EGFP were diffused throughout the cells. These results implied that the residues 1-27 were the minimum sequence required for mitochondrial targeting of NDUFS7, but for more efficient mitochondrial import, a longer MTS such the one constructed from the N-terminal 60 residues might be required.

The results also showed that the constructs with the N-terminal 60 residues (Fig. 10B, k) or full-length NDUFS7 (Fig. 10B, j) fusing at the C-terminus of EGFP could not be imported into mitochondria, even though these constructs contained the full-length MTS. These observations indicated that the correct direction of the NDUFS7 MTS located at the N- or C-terminus of the protein is important for mitochondrial import of this protein.

To explore the critical residues in the MTS of NDUFS7, site-directed mutagenesis method was used to replace leucine, isoleucine and arginine to glutamine, threonine and glycine, respectively, with the NDUFS₁₋₆₀-EGFP construct as the template. All residues selected for mutations was located within the first 27 residues, because we have demonstrated in previous studies that the N-terminal 27 residues is the minimum NDUFS7 MTS required for mitochondrial targeting (Fig. 11A). The

results showed that only one point mutation (L4Q) lost the ability to target to mitochondria completely, and others (e.g. L9Q, I14T, L15Q, L17Q, R10G, R13G or R18G) had milder effects on mitochondrial targeting (Fig. 11B, a-e and j-l). In constructs with double or triple mutations, proteins expressed from I14T/L15Q, L15Q/L17Q, L4Q/I14T/L15Q, R10G/R13G or R10G/R18G were not colocalized with Mito Tracker Red (Fig. 11B, f, h-i and m-n), but proteins expressed from I14T/L17Q or R13G/R18G could have partial colocalization (Fig. 11B, g and o). Similar results were obtained when three critical amino acids just identified were individually mutated (R10G, L4Q and L15Q) in the NDUFS₁₋₂₇-EGFP construct (Fig. 11B, p-r). These results clearly indicated that these three amino acids (L4, R10 and L15) in the NDUFS7 MTS are critical for mitochondrial import of this protein. In addition, we concluded that both positive charged and hydrophobic residues which may form amphipathic α -helices in the MTS are essential for mitochondrial import of NDUFS7.

A competent NLS of NDUFS7 was located at the C-terminal region

When EGFP was fused with the full-length NDUFS7 in the N-terminus, it showed that the EGFP signals diffused throughout the cells and had prominent accumulation in nuclei (Fig. 10B, j). Based on this result, NDUFS7 may have a nuclear localization signal (NLS) for nuclear targeting. However, no typical NLSs which contained three or more basic residues (R or K) were recognized at the N-terminus using PSORT II sequence analysis programs (<http://psort.ims.u-tokyo.ac.jp/form2.html>). In contrast, the C-terminal 187-213 residues of NDUFS7 contained 8 basic residues which formed 4 sets of 2 connected basic residues that might be an effective NLS. In order to study this region to see if it contains a competent NLS, various C-terminal fragments from NDUFS7 were fused

with EGFP in the N-terminus and the resultant constructs were transfected into T-REx-293 cells (Fig, 12A).

When the peptides of amino acids 187-213 were fused with EGFP (EGFP-NDUFS7₁₈₇₋₂₁₃), two significantly different distribution patterns were observed: one had the fusion protein distributed majorly in the cytoplasm (Fig. 12B, a), the other had the most fusion protein accumulated in nuclei (Fig. 12B, b). To narrow down the region of the putative NLS, two peptides with residues 199-213 and 199-206 were chosen for further NLS studies using similar strategies. Both fusion proteins derived from these two peptides were able to accumulate in nuclei (Fig. 12B, c and d). Especially, the distribution of the EGFP-NDUFS7₁₉₉₋₂₁₃ construct was exclusively in nuclei. These observations indicated that the peptide converting the C-terminal 199-213 amino acids of NDUFS7 is a competent NLS which can carry EGFP into nuclei but does not have a typical NLS sequence. Excitingly, the corresponding sequence in the C-terminus of NuoB (a NDUFS7 homologous from *N. crassa*) also had a similar nuclear import ability (Fig. 12B, e). This observation suggested that the possibility of nuclear localization of NDUFS7 not only exists in human but may also in other eukaryotic species like *B. Taurus*, *N. crassa* or even *Y. lipolytica* (Fig. 2).

To define the important residues in the NLS of NDUFS7, site-directed mutagenesis was conducted in previous identified NLS region to change arginine or lysine to glycine, using the EGFP-NDUFS7₁₉₉₋₂₁₃ construct as the template (Fig. 13A). The results were summarized in Figure 13. The efficiency of nuclear import was slightly decreased when individual basic residue was mutated to glycine. These mutative constructs included R199G, K200G, K201G, R202G, R205G, R206G, R212G and R213G (Fig. 13B, b-c, e-f, h-i and k-l). However, when both the connected basic amino acids were mutated (e.g. R199G/K200G, K201G/R202G,

R205G/R206G and R212G/R213G), the EGFP signals were diffused throughout the cells (Fig. 13B, d, g, j and m). A similar trend was also observed when double mutations were conducted two non-connected basic residues (e.g. K200G/R206G and R203G/R206G, Fig. 13B, n and o). Meaningfully, these 8 basic residues are all related to the nuclear import ability of NDUF7 NLS and the effect is synergistic.

Recently, it was demonstrated that the EGFP could enter nuclei by itself in some occasions, even though it did not have any putative NLS sequence [63]. It was also well recognized that some nuclear proteins with molecular weights less than approximately 60 kDa could enter into nuclei by passive diffusion. However, for nuclear proteins larger than this size would need a NLS for successful nuclear entry [27, 28]. To avoid the possible influence of EGFP on the nuclear targeting of the tested NLS constructs, LacZ (β -galactosidase), which is a large protein with 120 kDa size, was fused in front of the NLS sequences of NDUF7 and NuoB as the passenger protein by using the pcDNA4/TO/LacZ/myc/His vector as the template. With this design, it was clearly observed that these two peptides selected were able to carry most of the LacZ into nuclei of T-REx-293 cells (Fig. 14B, b and c) when they were compared with the LacZ only construct (Fig. 14B, a). A similar result also appeared in a test using HeLa cells as the experimental system (Fig. 14B, d) indicating, that NDUF7 has the nuclear import ability because of its C-terminal 199-213 residues.

The C-terminal region of NDUF7 may contain a NES sequence

In previous results, it showed that the nuclear targeting efficiency of EGFP-NDUF7₋₁₉₉₋₂₁₃ is higher than that of the longer EGFP-NDUF7₋₁₈₇₋₂₁₃ construct. In other words, the EGFP-NDUF7₋₁₈₇₋₂₁₃ construct has a more cytoplasmic distribution compared with the EGFP-NDUF7₋₁₉₉₋₂₁₃ construct (Fig. 12B, a and c). Therefore, it is a reasonable speculation that the region covering NDUF7₋₁₈₇₋₂₁₃ may

contain a nuclear export signal (NES) sequence. By using NetNES 1.1 as a NES prediction program (<http://www.cbs.dtu.dk/services/NetNES/>) (Fig. 15A), it could be found that the region of NDUFS7 C-terminus may contain a NES. When this putative NES sequence was compared with well known NES sequences, including IκBα, PRRSV N protein, p53 and p73, the putative NES of NDUFS7 was partially fitted in the classical NES motif as L-(X)₂₋₃-L-(X)₂₋₃-L-X-L (Fig. 15B). To investigate the involvement of the conserved residues in the NES motif, several leucines in this region was mutated to alanines by site-directed mutagenesis with EGFP-NDUFS7₁₈₇₋₂₁₃ as the template. Interestingly, when both L195 and L197 were mutated at the same time (L195A/L197A), the nuclear distribution was sharply increased (Fig. 15D, b) as compared with the control (Fig. 15D, a). Even more significantly, the fusion proteins were exclusively accumulated in nuclei when all four leucine residues were mutated to alanines simultaneously in the L190A/L191A/L195A/L197A construct (Fig. 15D, c). These results suggested that the C-terminal 187-213 amino acids of NDUFS7 contain a competent NES, and the leucine residues in this region are important for nuclear export.

To confirm the aforementioned finding, a nuclear export inhibitor leptomycin B (LMB) was used to identify the role of NES in the nuclear export of NDUFS7. After transfected with the EGFP-NDUFS7₁₈₇₋₂₁₃ construct for 2 days, T-REx-293 cells were incubated in absence or presence of LMB at concentration of 20 ng/ml for 4 more hours. In the absence of this drug, most of the EGFP fusion proteins were diffused in cytoplasm (Fig. 15D, a). However, a striking accumulation in nuclei was observed when LMB was added to the cell culture, suggesting the loss of export ability of the fusion proteins (Fig. 15D, d). These results demonstrated that the nuclear export ability of the NDUFS7 NES was affected by LMB which has the function of inhibiting formation of the nuclear CRM1 export complex.

NDUFS7 has a dual distribution both in mitochondria and nuclei

In previous section, the EGFP fusion system was applied to locate the covering region and essential amino acid residues of the MTS, NLS and NES in NDUFS7. In order to conform that NDUFS7 has a dual localization (mitochondria and nuclei), full-length NDUFS7 fusing with both c-Myc and His tags in the pcDNA4/TO/myc-His plasmid was designed and constructed (Fig. 16A). The subcellular localization of these constructs and mitochondria were probed with anti-His and anti-ATPase alpha subunit antibodies, respectively, in T-REx-293 cells. In most cases, the subcellular distribution of NDUFS7₁₋₂₁₃-c-Myc-His was mostly in mitochondria (Fig. 16B, a), but some cells had targeted proteins colocalized both in mitochondria and nuclei (Fig. 16B, b), in T-REx-293 cells. To further confirm this finding, the confocal microscopy (LSM510, ZEISS, Welwyn Garden City, Germany) was used to scan each Z section from 0 to 8 (Fig. 16B, d-1). The nucleus and mitochondrial patterns were recognized in Z0-Z6 and Z1-Z8, respectively. From these data it was apparent that NDUFS7 had a dual distribution both in mitochondria and nuclei, and the result was corresponded to previous EGFP data.

Discussion

In our present studies, the NDUF57 was suppressed in human T-REx-293 cells using the RNAi technology. The reduction in the NDUF57 expression caused a slow growth rate in galactose containing medium without glucose and increased ROS generation. These results indicated that NDUF57 may play an important role in cell energy production and survival. In addition, we also discovered the signal sequences including MTS, NLS and NES in the NDUF57 and investigated the critical amino acid residues involved in the targeting by using the EGFP fusion system. Although the region of N-terminal 1-27 residues in NDUF57 is capable of mitochondria targeting, the N-terminal 1-60 residues contains a more sufficient ability to carry EGFP into mitochondria. Both NLS and NES are localized at the C-terminus of NDUF57, and they can carry passenger proteins efficiently in or out, respectively, through the nuclei. In addition, to prevent the possible influence caused by the use of EGFP fusion techniques, a small tag fusion system such as c-Myc and His tags were also applied in this study. The results indicated that the construct of wild type NDUF57 fused with c-Myc and His tags also has a dual distribution both in mitochondria and nuclei, and suggested that this phenomenon may also exist in the native situation for NSUF57.

The NDUF57 studies of mitochondrial targeting

In previous studies, the N-terminal 38 amino acids of bovine PSST were predicted to be the MTS [4, 25]. In this study, we identified that the N-terminal residues 1-27 was the minimum region required for mitochondrial targeting. Although this region could carry EGFP to mitochondria, the transport efficiency of this peptide has its limitation. Our data suggested that a longer N-terminal peptides has better mitochondrial targeting ability, and the first 60 amino acids already have the same

mitochondrial targeting efficiency as that of the full-length NDUF57.

Because NDUF57₃₀₋₆₀-EGFP, NDUF57₃₀₋₉₆-EGFP and NDUF57₃₉₋₁₈₉-EGFP constructs were unable to be transported to mitochondria, these results implied that the key region of mitochondrial targeting ability is located at first 1-27 amino acids, and the residues 30-60 play a supporting role in the MTS of NDUF57. In order to confirm this hypothesis, we used Target P 1.1 server (<http://www.cbs.dtu.dk/services/TargetP/>) to predict the mitochondrial targeting ability in different NDUF57 constructs. The results showed that the full-length NDUF57 scored 0.948 which is a strong indication of a mitochondrial protein and the N-terminal 1-27 residues of NDUF57 scored 0.617, whereas the N-terminal 28-60 residues of NDUF57 scored only 0.153 which has a very little mitochondrial targeting potential. These bioinformatic results may support our hypothesis that the first 1-27 amino acids play an important role for mitochondrial targeting and the residues 28-60 has a supporting role to increase the MTS efficiency.

From the PSIPRED secondary structure prediction server (<http://bioinf.cs.ucl.ac.uk/psipred/psiform.html>), it was showed that the NDUF57 N-terminal 1-32 residues contains a α -helix-loop- α -helix motif which is a common structure found in the MTS. However, when the first 1-32 amino acids of NDUF57 were arranged by Helical Wheel Projection (<http://rzlab.ucr.edu/scripts/wheel/wheel.cgi>), the results showed that it does not contain a very classical amphipathic pattern. It indicated that the MTS of NDUF57 is not belonging to the common MTS of mitochondrial matrix proteins.

Moreover, we used the site-directed mutagenesis analysis to investigate the critical amino acid residues in the MTS of NDUF57. We found that the L9Q, I14T, L15Q and L17Q point mutations in the NDUF57₁₋₆₀ constructs would partially decrease the ability of mitochondrial targeting, but when the L4 of NDUF57₁₋₆₀ was

mutated, the MTS capability would be mostly destroyed. The double mutations of I14 and L15 in NDUFS7₁₋₆₀ constructs would cause a synergistic loss of the mitochondrial targeting ability. In contrast, any one of basic residues (R10G, R13G and R18G) mutated in NDUFS7₁₋₆₀ constructs only caused a slightly decrease in the MTS capability, but double mutations in any two basic residues would completely lose the mitochondrial targeting ability except the R13G/R18G pair. These results demonstrated that both basic and hydrophobic residues are very important for the mitochondrial targeting ability of the MTS in NDUFS7, and their effects are synergistic. Intriguingly, the MTS ability was destroyed completely when one basic (R10) or hydrophobic (L15) residue in NDUFS7₁₋₂₇ was mutated. This result confirmed that the supporting role of the residues 28-60 of NDUFS7 is important for efficient mitochondrial import.

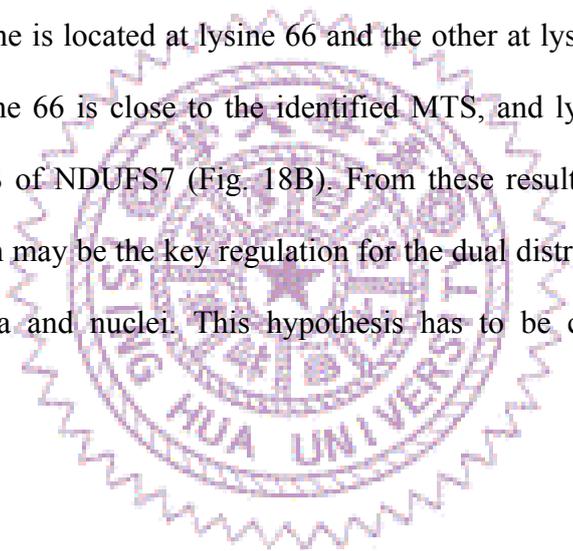
The NLS and NES studies of NDUFS7

In this study, we also identified that NDUFS7 has other signal peptides, including non-classical NLS and NES, which are both located at the C-terminus of NDUFS7. These results strongly suggested that NDUFS7 has a dual distribution both in mitochondria and nuclei in the native state (Fig. 17). However, what factors determine its distribution in these two organelles remain unknown. In addition, there is still no clue to the function of the NDUFS7 in the nucleus. Further experiments are required to answer these two questions.

The perspective of SUMOylation and the dual distribution of NDUFS7

In recently studies, scientists have found that the distribution of some nuclear proteins are related to a novel host cell posttranslational modification system, termed

sumoylation. Many SUMOylated proteins are localized on nuclei, and SUMOylation is known as the key to regulate the distribution of nucleocytoplasmic transport [64, 65]. For example, the level of nuclear localization in KLF5 was enhanced by SUMOylation, and its SUMOylation site was close to the NES of this protein [66]. To evaluate the possibility that the SUMOylation is involved in the NDUFS7 distribution, the SUMOplot analysis program (http://www.abgent.com/tools/sumoplot_login) was used to predict the SUMOylation sites on the basis of a consensus motif Ψ KXE (Ψ , a large hydrophobic residue; K, lysine; X, any residue; and E, an acidic residue which is primarily glutamate) [64, 65]. Interestingly, two putative SUMOylation sites were found in NDUFS7: one is located at lysine 66 and the other at lysine 202 (Fig. 18A). Surprisingly, the lysine 66 is close to the identified MTS, and lysine 202 is located within NLS and NES of NDUFS7 (Fig. 18B). From these results, we hypothesized that the SUMOylation may be the key regulation for the dual distribution of NDUFS7 between mitochondria and nuclei. This hypothesis has to be confirmed by more experiments.



Tables

Table 1. Sequence identity (%) and similarity (%) of NDUFS7 to its homologous proteins from various species

Species	Identity (%)	Similarity (%)
<i>Bos taurus</i>	85.8	95.4
<i>Neurospora crassa</i>	62.3	88.9
<i>Yarrowia lipolytica</i>	62.4	84.7
<i>Paracoccus denitrificans</i>	81.3	94.2
<i>Thermus thermophilus</i>	58.4	86.9
<i>Escherichia coli</i>	46.9	77.1

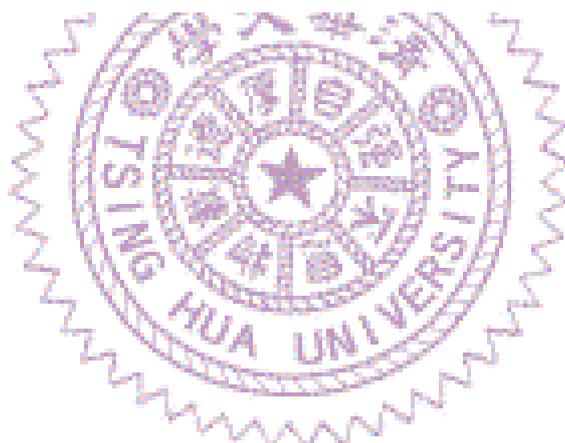


Table 2. Primers used in this study

Applications/Primer name	Nucleotide sequence (5'→3')	Comment
For shRNA		
NDUFS7-shRNA-ORF-F1	AGCTTCCGAGGAGGCTACTACCACTATTCAAGAGATAGTGGTAGTAGCCTCCTCTTTTC	the coding region
NDUFS7-shRNA-ORF-R1	TCGAGAAAAAGAGGAGGCTACTACCACTATCTCTTGAATAGTGGTAGTAGCCTCCTCGGA	the coding region
NDUFS7-shRNA-3'UTR-F1	AGCTTCCCCGTGAGGTTGTCAATAAAATTCAAGAGATTTATTGACAACCTCACGGTTTTTC	3'UTR
NDUFS7-shRNA-3'UTR-R1	TCGAGAAAAACCGTGAGGTTGTCAATAAAATCTCTTGAATTTATTGACAACCTCACGGGGA	3'UTR
pSUPER-F1	ATGAGACCACAGATCTAAGC	For sequencing
pSUPER-R1	TTAGCTCACTCATTAGGCAC	For sequencing
pSUPER-F2	CGAACGCTGACGTCATCAAC	For sequencing
Reverse Transcription-PCR		
NDUFS7-F1	GCACGAGGTGTCCATCAGAG	Identification the coding region of NDUFS7
NDUFS7-R1	CAACCTCACGGCACACAAGC	Identification the coding region of NDUFS7
NDUFS7-F3	CTACGACATGGACCGCTTTG	Identification the coding region of NDUFS7
NDUFS7-R3	CCGAGTAGGAATAGTGGTAG	Identification the coding region of NDUFS7
GADPH-F2	GTCGTCATGGGTGTGAACC	Identification the coding region of NDUFS7
GADPH-R1	GCATGGACTGTGCTCATGAG	Identification the coding region of NDUFS7
Cloning of NDUFS7 in pcDNA4/TO/myc-His		
CMVF-F1	CGCAAATGGGCGGTAGGCGTGT	For sequencing
Vector Reverse	TAGAAGGCACAGTCGAGG	For sequencing
myc-R	GATCCTCTTCTGAGATGAG	For sequencing
NDUFS7-EcoRI-F	CGAATTCGCCAAGATGGCGGTGCTGTC	Identification the full length coding region of NDUFS7
NDUFS7-XhoI-R	GCTCGAGTCTGCGGTACCAGATCTGCAG	Identification the full length coding region of NDUFS7
LacZ-F	GATTGGTGGCGACGACTC	Identification the coding region of LacZ

Cloning of NuoB in pcdNA4/TO/myc-IIs

NuoB-212-226-NotI-F

TTGCGGCCGCAAGGAAGATGAGGAATACG

Identification the fragment of NuoB

NuoB-212-226-XhoI-R

CCTCGAGACTTCCGGTACCACATCCT

Identification the fragment of NuoB

Cloning of NDUFS7 in pEGFP-N3

pEGFP-N3-NDUFS7-XhoI-F

GCTCGAGATGGCGGTGCTGTCTCAGCTCC

Identification the full length coding region of NDUFS7

pEGFP-N3-NDUFS7-EcoRI-R

GGAATTCGCCCTGCGGTACCAGATCTGC

Identification the full length coding region of NDUFS7

pEGFP-N3-NDUFS7-BsrGI-F

CTGTACAAGATGCCGGTGTCTCAGCTC

Identification the full length coding region of NDUFS7

pEGFP-N3-NDUFS7-NotI-R

CGCGGCCCGCTACCTGCGGTACCAGATCTG

Identification the full length coding region of NDUFS7

pEGFP-N3m-NDUFS7-EcoRI-F1

CGCCACCGAATTCATGGCGGTGCTGTCTCAGC

Identification the full length coding region of NDUFS7

pEGFP-N3m-NDUFS7-XhoI-R1

CCTCGAGCCCTGCGGTACCAGATCTG

Identification the fragment of NDUFS7

pEGFP-N3-NDUFS7-1-19-EcoRI-R

GGAATTCGCCGAGCGCAGACCAAGGATC

Identification the fragment of NDUFS7

pEGFP-N3-NDUFS7-1-29-EcoRI-R

GGAATTCGCACCTCGTGCCATGCACAGC

Identification the fragment of NDUFS7

pEGFP-N3-NDUFS7-1-27-EcoRI-R

GGAATTCGCTGCCCTGCACAGCCAGGCC

Identification the fragment of NDUFS7

pEGFP-N3-NDUFS7-1-38-EcoRI-R

GGAATTCGCCCATCGGTGGCCACGGCTC

Identification the fragment of NDUFS7

pEGFP-N3-NDUFS7-1-50-EcoRI-R

GGAATTCGCTCTGGCCCTTGGCAGGGC

Identification the fragment of NDUFS7

pEGFP-N3-NDUFS7-1-60-EcoRI-R

GGAATTCGCGCCCGGGCTGCTGGGTTT

Identification the fragment of NDUFS7

pEGFP-N3-NDUFS7-30-60-XhoI-F

GCTCGAGATGGTCCATCAGAGCGTG

Identification the fragment of NDUFS7

pEGFP-N3-NDUFS7-30-96-EcoRI-R

GGAATTCGCCATGTGCATCATCTCCAC

Identification the fragment of NDUFS7

pEGFP N3 NDUFS7 39 189 XhoI F

GCTCGAGATGCCAAGCAGCAGCCAGCCTG

Identification the fragment of NDUFS7

pEGFP-N3-NDUFS7-39-189-EcoRI-R

GGAATTCGCGCCCTCGGCCGTAGGTG

Identification the fragment of NDUFS7

pEGFP-N3-NDUFS7-199-213-EcoRI-F

CGAATTCAGGAAGATCAACCGGGAG

Identification the fragment of NDUFS7

pEGFP-N3-NDUFS7-199-213-XhoI-R1

CCTCGAGACCTGCGGTACCAGATCTGC

Identification the fragment of NDUFS7

pEGFP-N3-NDUFS7-189-213-XhoI-F

GCTCGAGATGGCCCTGCTCTACGGCATC

Identification the fragment of NDUFS7

pEGFP-N3-NDUFS7-187-213-XhoI-F

GCTCGAGATGGCCGAGGCCCTGCTCTAC

Identification the fragment of NDUFS7

pEGFP-N3-NDUFS7-187-213-BsrGI-F

GTGTACAAGATGCCCGAGGCCCTGCTCTAC

Identification the fragment of NDUFS7

pEGFP-N3-NDUFS7-199-213-BsrGI-F

GTGTACAAGAGGAAGATCAAGCGGGAG

Identification the fragment of NDUFS7

NDUFS7-1-60-NotI-R

CGCGGCCCGCTAGCCCCGGCTGCTGGGTTG

Identification the fragment of NDUFS7

EGFP-F1

CTACCTGAGCACCCAGTC

For sequencing

EGFP-R1

CGCGGCCCGCTTACTTGTAC

For sequencing

Cloning of NuoB in pEGFP-N3

NuoB-212-226-BsrGI-F	GTGTACAAGAGGAAGAATGAGGAATACGAAGATCAAGGAGGATG	Identification the fragment of NuoB
NuoB-212-226-NotI-R	AGCGGCCCGCCTACTTGCGGTACCACATCCTCGTGATCTTCGTATT	Identification the fragment of NuoB
Site-directed mutagenesis		
L4Q-F	CGAGATGGCGGTGCAGTCAGCTCCTGGCC	Leu4 to Gln
L4Q-R	GGCCAGGAGCTGACTGCACCGCCATCTCG	Leu4 to Gln
L9Q-F	GTCAGCTCCTGGCCAGCGCGGCTTCCGGA	Leu9 to Gln
L9Q-R	TCCGGAAGCCCGCTGGCCAGGAGCTGAC	Leu9 to Gln
I14T-F	GCGCGGCTTCCGGACCCTTGGTCTGCGCTC	Ile14 to Thr
I14T-R	GAGCGCAGACCAAGGGTCCGGAAGCCGCGC	Ile14 to Thr
L15Q-F	CGGCTTCCGGATCCAGGCTCGCGCTCCAG	Leu15 to Gln
L15Q-R	CTGGAGCGCAGACCCTGGATCCGGAAGCCG	Leu15 to Gln
L17Q-F	CCGGATCCTTGGTCAGCGCTCCAGCGTGG	Leu17 to Gln
L17Q-R	CCACGCTGGAGCGCTGACCAAGGATCCGG	Leu17 to Gln
I14T+L15Q-F	GCGCGGCTTCCGGACCAGGGTCTGCGCTCCAG	Ile14 to Thr and Leu15 to Gln
I14T+L15Q-R	CTGGAGCGCAGACCCTGGTCCGGAAGCCCGCC	Ile14 to Thr and Leu15 to Gln
I14T+L17Q-F	CCGGACCCTTGGTCAGCGCTCCAGCGTGG	Ile14 to Thr and Leu17 to Gln
I14T+L17Q-R	CCACGCTGGAGCGCTGACCAAGGGTCCGG	Ile14 to Thr and Leu17 to Gln
L15Q+L17Q-F	CCGGATCCAGGGTCCAGCGCTCCAGCGTGG	Leu15 to Gln and Leu17 to Gln
L15Q+L17Q-R	CCACGCTGGAGCGCTGACCCTGGATCCGG	Leu15 to Gln and Leu17 to Gln
R10G-F	CAGCTCCTGGCCTGGCGGGCTTCCGGATC	Arg10 to Gly
R10G-R	GATCCGGAAGCCGCCAGGCCAGGAGCTG	Arg10 to Gly
R13G-F	GCCTGCGCGGCTTCGGGATCCTTGGTCTG	Arg13 to Gly
R13G-R	CAGACCAAGGATCCCGAAGCCGCCAGGC	Arg13 to Gly
R18G-F	CGGATCCTTGGTCTGGGCTCCAGCGTGGGC	Arg18 to Gly
R18G-R	GCCCACGCTGGAGCCCAGACCAAGGATCCG	Arg18 to Gly
R10G+R13G-F	GCCTGGGCGGCTTCGGGATCCTTGGTCTG	Arg10 to Gly and Arg13 to Gly
R10G+R13G-R	CAGACCAAGGATCCCGAAGCCGCCAGGC	Arg10 to Gly and Arg13 to Gly

F12V-F	TGGCCTGGCGGGCTACCGGATCCTTGGTC	Phe12 to Tyr
F12V-R	GACCAAGGATCCGGTAGCCGCGCAGGCCA	Phe12 to Tyr
E204I-F	GGAAAGATCAAGCGGACGCGGAGGCTGCAG	Glu204 to Thr
E204T-R	CTGCAGCCTCCGCGTCCGCTTGATCTTCC	Glu204 to Thr
E204A-F	GAAGATCAAGCGGGCGGGAGGCTGCAG	Glu204 to Ala
E204A-R	CTGCAGCCTCCGCGCCGCTTGATCTTC	Glu204 to Ala
E204D-F	AAGATCAAGCGGGATCGGAGGCTGCAGATC	Glu204 to Asp
E204D-R	GATCTGCAGCCTCCGATCCGCTTGATCTT	Glu204 to Asp
E204Q-F	GGAAAGATCAAGCGGCAGCGGAGGCTGCAG	Glu204 to Gln
E204Q-R	CTGCAGCCTCCGCTGCCGCTTGATCTTCC	Glu204 to Gln
E204L-F	GGAAAGATCAAGCGGCTGCGGAGGCTGCAGA	Glu204 to Leu
E204L-R	TCTGCAGCCTCCGCAGCCGCTTGATCTTCC	Glu204 to Leu
E188A-F	ATCTCGAGATGGCCGCGGCCCTGCTCTAC	Glu188 to Ala
E188A-R	GIAGAGCAGGGCCGCGGCCAATCTCGAGAT	Glu188 to Ala
R199G-F	ACGAGCTGTACAAGGGGAAGATCAAGCGG	Arg199 to Gly
R199G-R	CCGCTTGATCTTCCCCTTGACAGCTCGT	Arg199 to Gly
K200G-F	GAGCTGTACAAGAGGGGTATCAAGCGGGAGCG	Lys200 to Gly
K200G-R	CGCTCCCGCTTGATACCCCTCTTGACAGCTC	Lys200 to Gly
R199G+K200G-F	ACGAGCTGTACAAGGGTGGTATCAAGCGGGA	Arg199 to Gly and Lys200 to Gly
R199G+K200G-R	TCCCGCTTGATACCACCCTTGACAGCTCGT	Arg199 to Gly and Lys200 to Gly
K202G-F	ACAAGAGGAAGATCGGACGGGAGCGGAGGCT	Lys202 to Gly
K202G-R	AGCCTCCGCTCCCGTCCGAICTTCCICTTGT	Lys202 to Gly
R203G-F	ACAGGAAGATCAAGGGACAGCCGAGGCTGCA	Arg203 to Gly
R203G-R	TGCAGCCICCGCTCTCCCTTGATCTTCCICT	Arg203 to Gly
K202G+R203G-F	ACAAGAGGAAGATCGGTGGAGAGCGGAGGCT	Lys202 to Gly and Arg203 to Gly
K202G+R203G-R	AGCCTCCGCTCTCCACCGATCTTCTCTTGT	Lys202 to Gly and Arg203 to Gly

R205G-F	GATCAAGCCGGAGGGAAGGCTGCAGATCTG	Arg205 to Gly
R205G-R	CAGATCTGCAGCCTTCCCTCCCGCTTGATC	Arg205 to Gly
R206G-F	TCAAGCCGGAGCGGGACTGCAGATCTGGTA	Arg206 to Gly
R206G-R	IACCAGATCTGCAGTCCCCGCTCCCCGCTTGA	Arg206 to Gly
R205G+R206G-F	GATCAAGCCGGAGGGTGGACTGCAGATCTG	Arg205 to Gly and Arg206 to Gly
R205G+R206G-R	CAGATCTGCAGTCCACCCTCCCGCTTGATC	Arg205 to Gly and Arg206 to Gly
R212G-F	TGCAGATCTGGTACGGCAGGIAGGCGGC	Arg212 to Gly
R212G-R	GCCGCCTACCTGCCGTACCAGATCTGCA	Arg212 to Gly
R213G-F	AGATCTGGTACCGCGGATAGGCGGCCGCGACT	Arg213 to Gly
R213G-R	AGTCGCGGCCGCTATCCGCGGTACCAGATCT	Arg213 to Gly
R212G+R213G-F	TGCAGATCTGGTACGGCGGATAGGCGGCC	Arg212 to Gly and Arg213 to Gly
R212G+R213G-R	GGCCGCCTATCCGCGGTACCAGATCTGCA	Arg212 to Gly and Arg213 to Gly
R203G+R206G-F	AGAGGAAGATCAAGGGAGAGCGGGACTGCA	Arg203 to Gly and Arg206 to Gly
R203G+R206G-R	TGCAGTCCCCGCTCTCCCTTGATCTTCCTCT	Arg203 to Gly and Arg206 to Gly
L195A+L197A-F	CTGCTCTACGGCATCGCGCAGGCGCAGAGGAAGATCAA	Leu195 to Ala and Leu197 to Ala
L195A+L197A-R	TTGATCTTCTCTGCGCCTGCGCGATGCCGTAGACGAG	Leu195 to Ala and Leu197 to Ala
L190A+L191A-F	CAAGATGGCCGAGGCCGCGGCCTACGGCATCGCGCAG	Leu190 to Ala and Leu191 to Ala
L190A+L191A-R	CTGCGCGATGCCGTAGGCCGCGGCCTCGGCCATCTTG	Leu190 to Ala and Leu191 to Ala



Figures

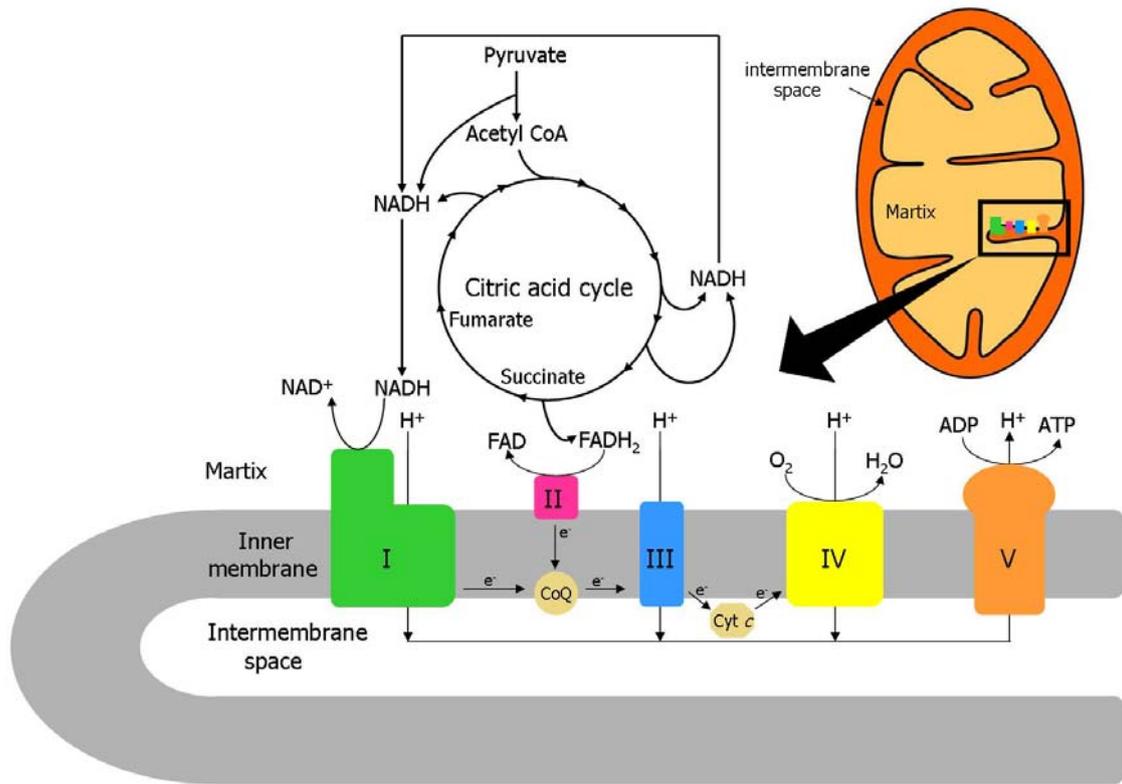


Figure 1. Overview of the five mitochondrial complexes of the OXPHOS.



```

H. sapiens      ---MAVLSAPGLRGRFRILGLRSSVGPVQARGVHQSD-----VATDGPSSSTQPALPKA 49
B. Taurus      ---MAALAALRLL-HPILAVRSGVGAALQVRGVHSS-----MAADSPSSSTQPAVSQA 48
N. crassa      MMSVVRTGASMLRARPQAIVPFRAAAVASISSSSRKDATGAVAPAGAQHGIARRERRE 60
Y. lipolytica  MLRSQIG--RLALRP---TLVP--ATVIPQTRAYS-----APAGTPRVSSSSMPTD 44
P. denitrificans -----MTGLNTAGADRD-----LATAELN----- 19
T. thermophilus -----MALKDLFER-----DV 11
E. coli        -----MDYTLTRIDPNCENDR-----YPLQKQEIVT 26
              1      1      g  av      g

H. sapiens      RAVAPK---PSSRGEYVVAKLDDLDVNWARRSSSLWPMTFGLACCAVEMMHMAAPRYDMDR 105
B. Taurus      RAVVPKPAALPSSRGEYVVAKLDDLDVNWARRSSSLWPMTFGLACCAVEMMHMAAPRYDMDR 108
N. crassa      VPPLPSQEG--TKGAVQYALITLDSIVNWARRSSSLWPMTFGLACCAVEMMHLSSTPRYDQDR 118
Y. lipolytica  FPLPSQOK--PNSAVDYTLTLLDAVANWARQGSFVWVTFGLACCAVEMMHVSAPRYDQDR 102
P. denitrificans RELQDK-----GFLLTTEEDIINWARRNGSLHWMTFGLACCAVEMMQTSMPRYDLER 70
T. thermophilus QELERE-----GILFTTLEKLVAVGRSNSLWPAITFGLACCAVEMMASTDARNDLAR 62
E. coli        DPLEQEVN-----KNVFMGKLNDMVNWGRKNSLWVYVTFGLSCCAVEMVTSFTAVIDHVAR 80
              I              yvvttlddlvnrWAr      slwpmTFGLACCAVEMHhms prydmdr

H. sapiens      FG-VVFRASPRQSDVMIVAGTLTNKMAPALRKVYDQMPPEPRYVISMGSCANGGGYYHYSY 164
B. Taurus      FG-VVFRASPRQSDVMIVAGTLTNKMAPALRKVYDQMPPEPRYVISMGSCANGGGYYHYSY 167
N. crassa      LG-IIFRASPRQSDLMIVAGTLTNKMAPALRQVYDQMPDPRVVISMGSCANGGGYYHYSY 177
Y. lipolytica  LG-IIFRASPRQSDLMIVAGTLTNKMAPVLRQVYDQMPPEPRVVISMGSCANGGGYYHYSY 161
P. denitrificans FG-TAPRASPRQSDLMIVAGTLTNKMAPALRKVYDQMPPEPRYVISMGSCANGGGYYHYSY 129
T. thermophilus FGSEVFRASPRQADVMIVAGRLSKKMAPVMRRVWVWQMPDPRVVISMGACASSGGMFN-NY 121
E. coli        FGAEVLFRASPRQADLMVIVAGTCTFTKMAPVIQRVYDQMPPEPRVVISMGACANSGGMYD-IY 139
              FG vvfRASPRQsDvMiVAGTltnkMAPAlrkvydQmpEprwvISMGScAngGGyyhysY

H. sapiens      SVVRGCDRIVPVDIYVPGCPPTAEALLYGILQLQKIK-----RERRLQ 208
B. Taurus      SVVRGCDRIVPVDIYVPGCPPTAEALLYGILQLQKIK-----REKRLR 211
N. crassa      SVVRGCDRIVPVDIYVPGCPPTSEALMYGIFQLQKMR-----NTKITR 221
Y. lipolytica  SVVRGCDRIVPVDIYVPGCPPTSEALMYGVFQLQKMR-----NTKITR 205
P. denitrificans SVVRGCDRIVPVDIYVPGCPPTAEALLYGILQLQRAS-----GAP--A 171
T. thermophilus AIVQNVDSVVPVDVYVPGCPPRPEALIIYAVMQLQKVRGQAY-----NERGERLPPVA 174
E. coli        SVVQGVDKFIPVDVYVPGCPPRPEAYMQALMLQESIGKERRPLSVVVGDQGVYRANMQS 199
              svvrgcdRivPVDiYvPGCPpt EAllygilqLQrkik              r r

H. sapiens      IWYRR----- 213
B. Taurus      IWYRR----- 216
N. crassa      MWYRK----- 226
Y. lipolytica  MWYRK----- 210
P. denitrificans RW----- 173
T. thermophilus AWKRTRG----- 181
E. coli        ERERKRGERIAVTNLRTPDEI 220
              iwyr

```

Figure 2. Multiple sequence alignment of NDUFS7 protein from various homologues. Sequence alignment was generated by EMBL-EBI ClustalW2, and displayed by BOXSHADE server. Red boxes are cysteine residues of [4Fe-4S] cluster motif: CCXXE(X)₆₀C(X)₃₀CP.

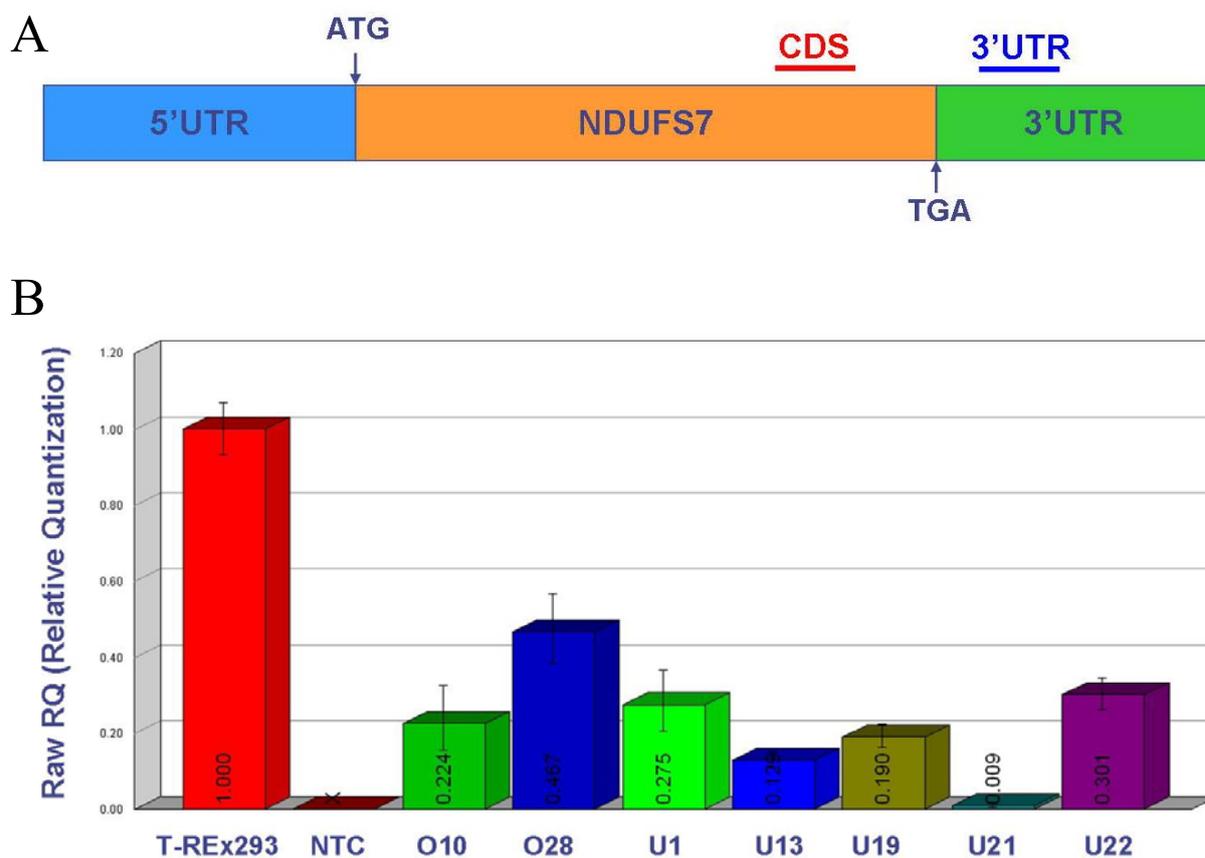


Figure 3. The knock down effect of RNA interference on the mRNA level of NDUFS7. (A) Two different targeting regions (CDS and 3' UTR) of shRNA were selected for NDUFS7 suppression. (B) Quantification of the relative amounts of NDUFS7 was determined by real-time PCR in T-REx-293, O10, O28, U1, U13, U19, U21 and U22 cells.

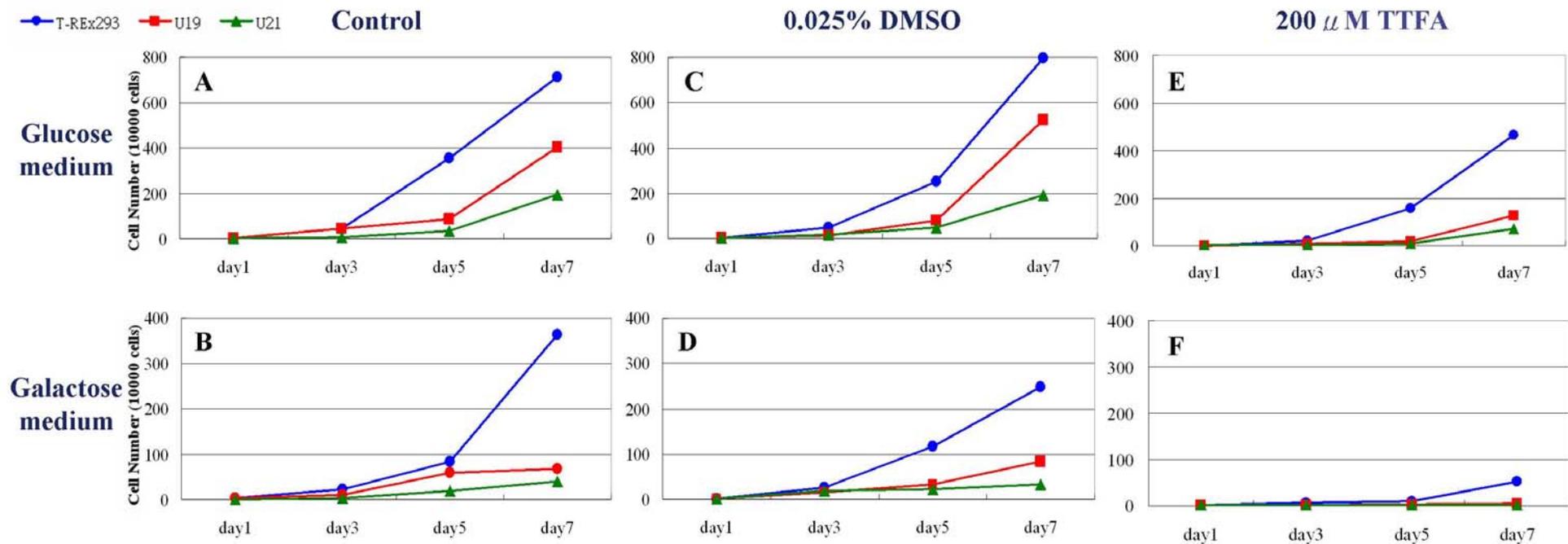


Figure 4. The growth rate of T-REx293 and NDUFS7 knock down U19 and U21 cell lines were measured in glucose (A, C and E) and galactose (B, D and F) containing medium. Cells were seeded at 2×10^4 per well at first day 0 and counted for 7 days. DMSO (0.025%) was added in mediums (C and D), and the 200 μ M TTFA (a complex II inhibitor) was added in mediums (E and F).

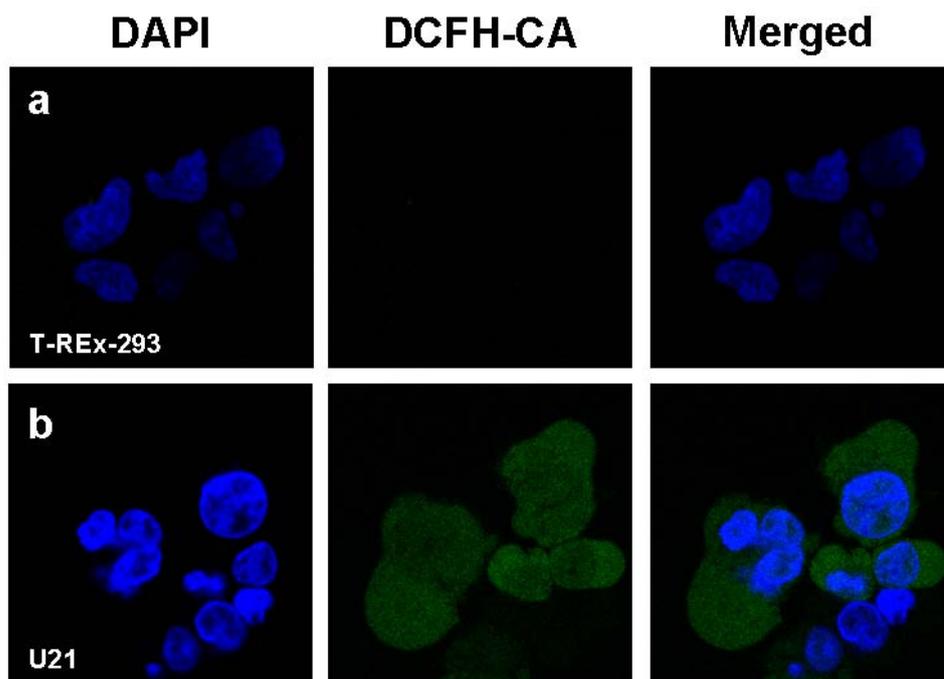
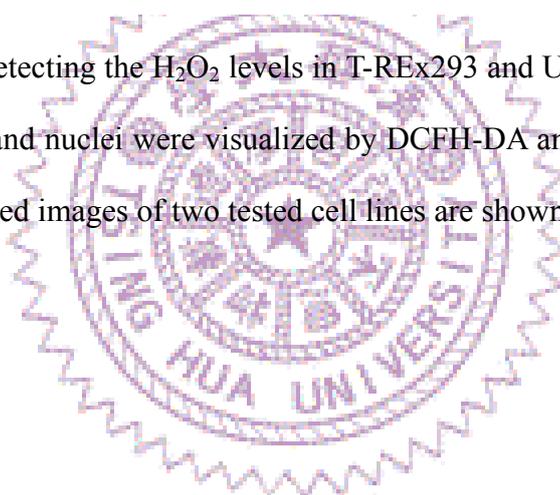


Figure 5. ROS assay detecting the H₂O₂ levels in T-REx293 and U21 cells. The subcellular localizations of H₂O₂ and nuclei were visualized by DCFH-DA and DAPI, respectively, in T-REx-293 cells. Merged images of two tested cell lines are shown in the right panel.



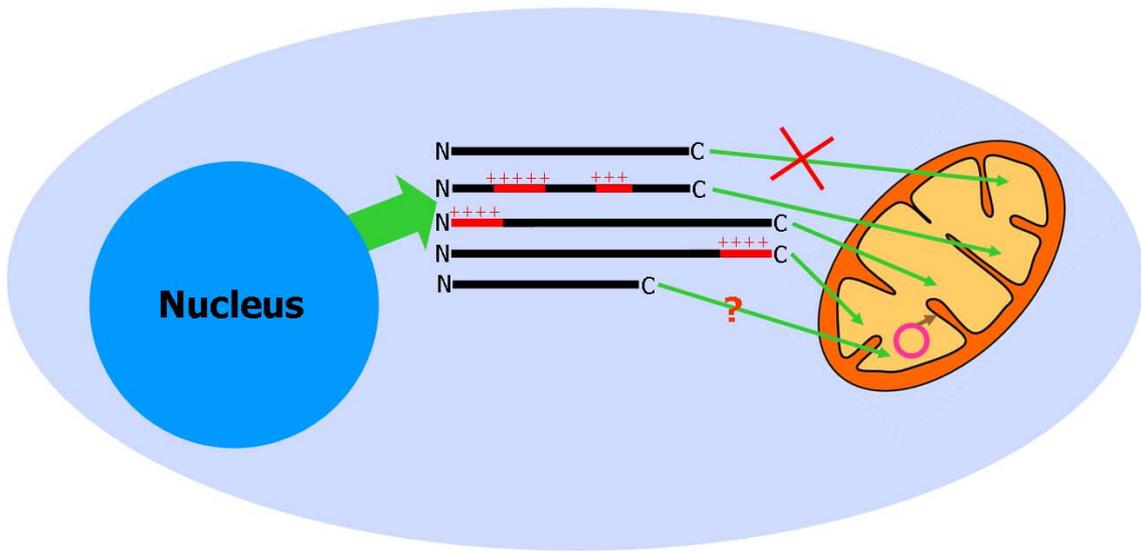
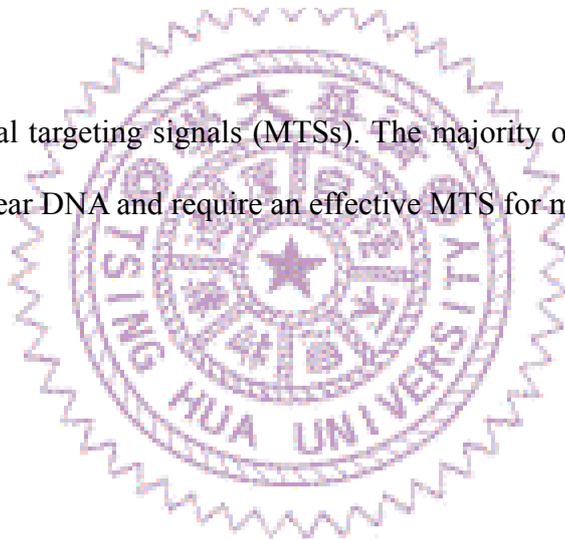


Figure 6. Mitochondrial targeting signals (MTSSs). The majority of mitochondrial proteins are encoded from nuclear DNA and require an effective MTS for mitochondrial import.



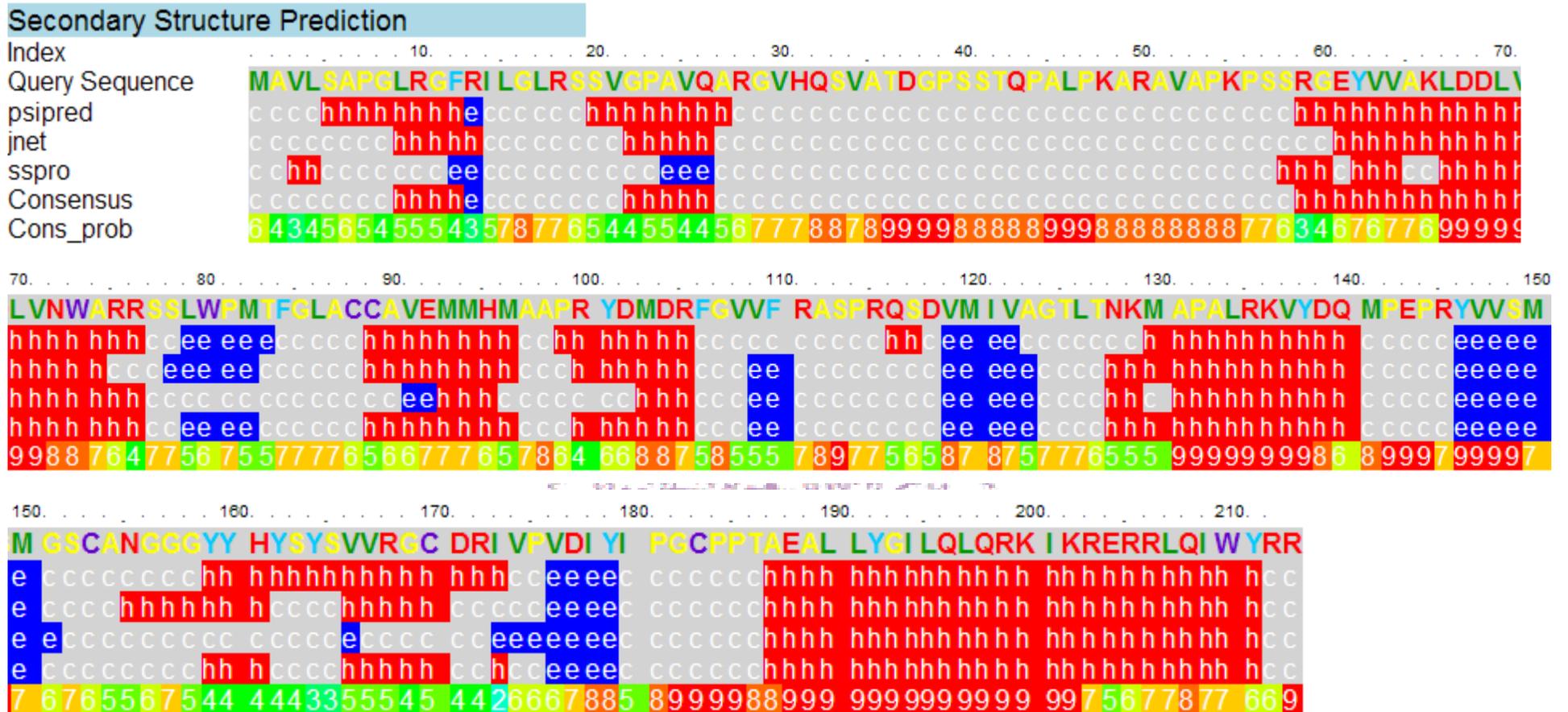
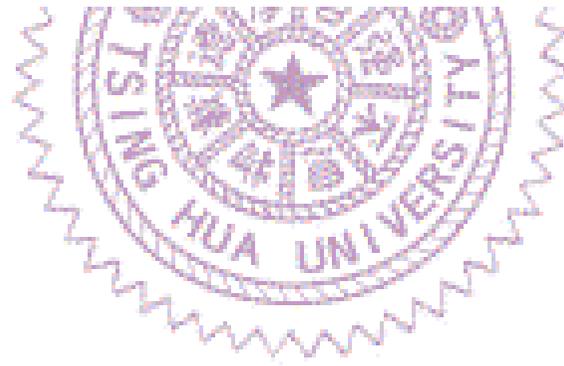
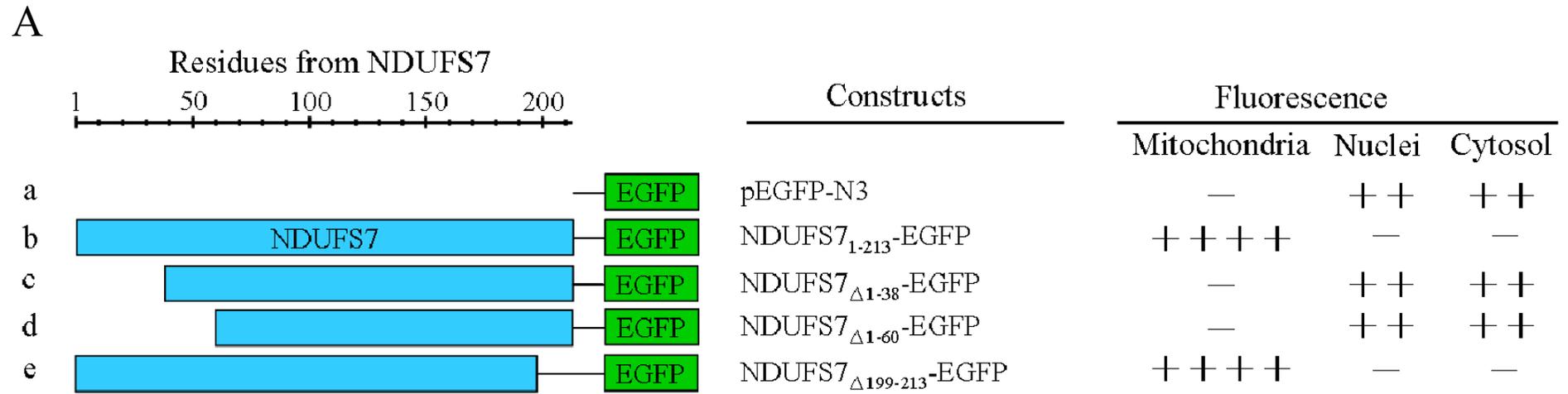


Figure 7. The secondary structure prediction of NDUFS7. The Phyre program was used as a prediction program for secondary structure prediction (<http://www.sbg.bio.ic.ac.uk/phyre/index.cgi>). Secondary structures such as α -helix, β -sheet and non-special were marked with h, e and c, respectively.



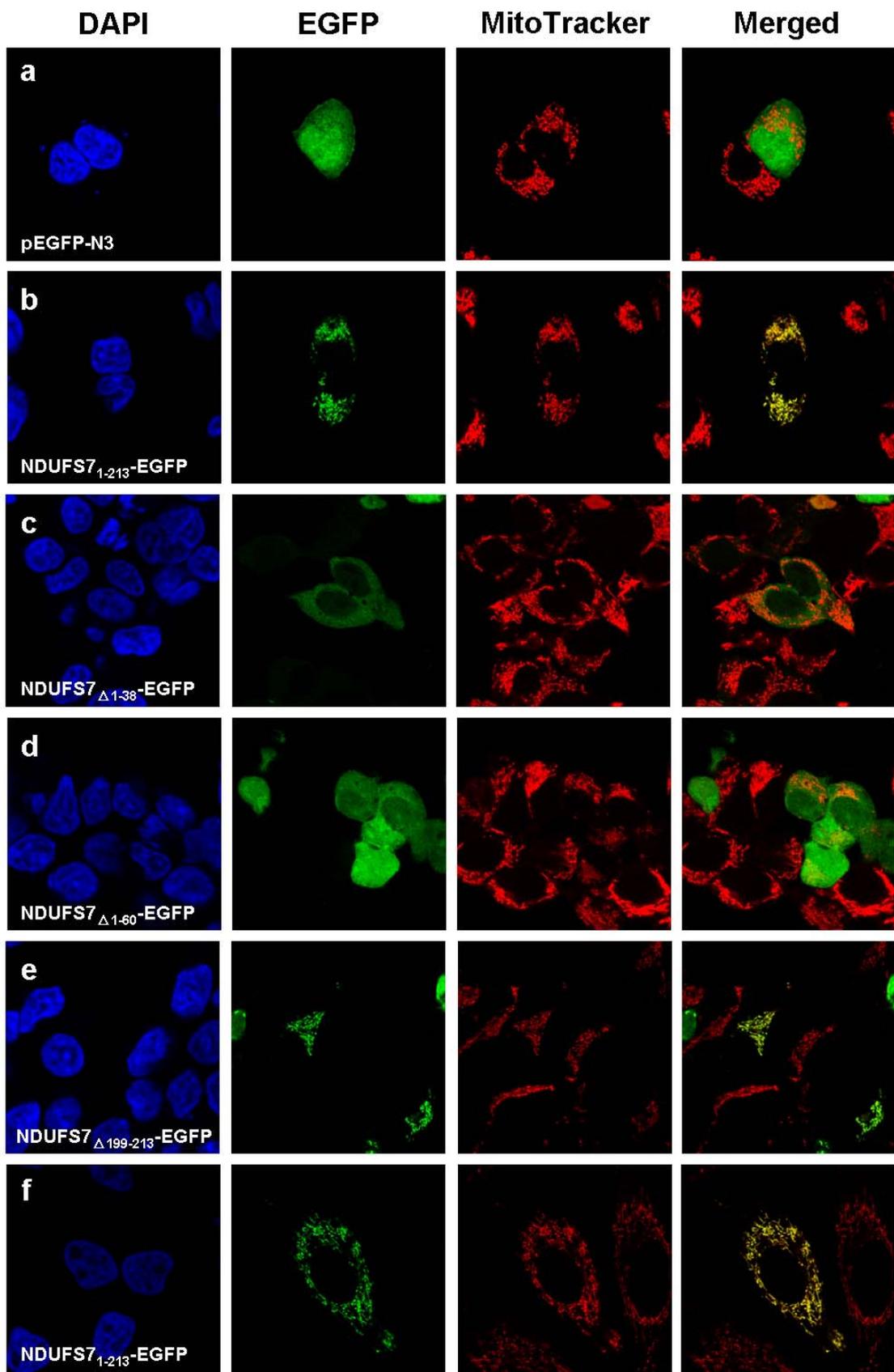
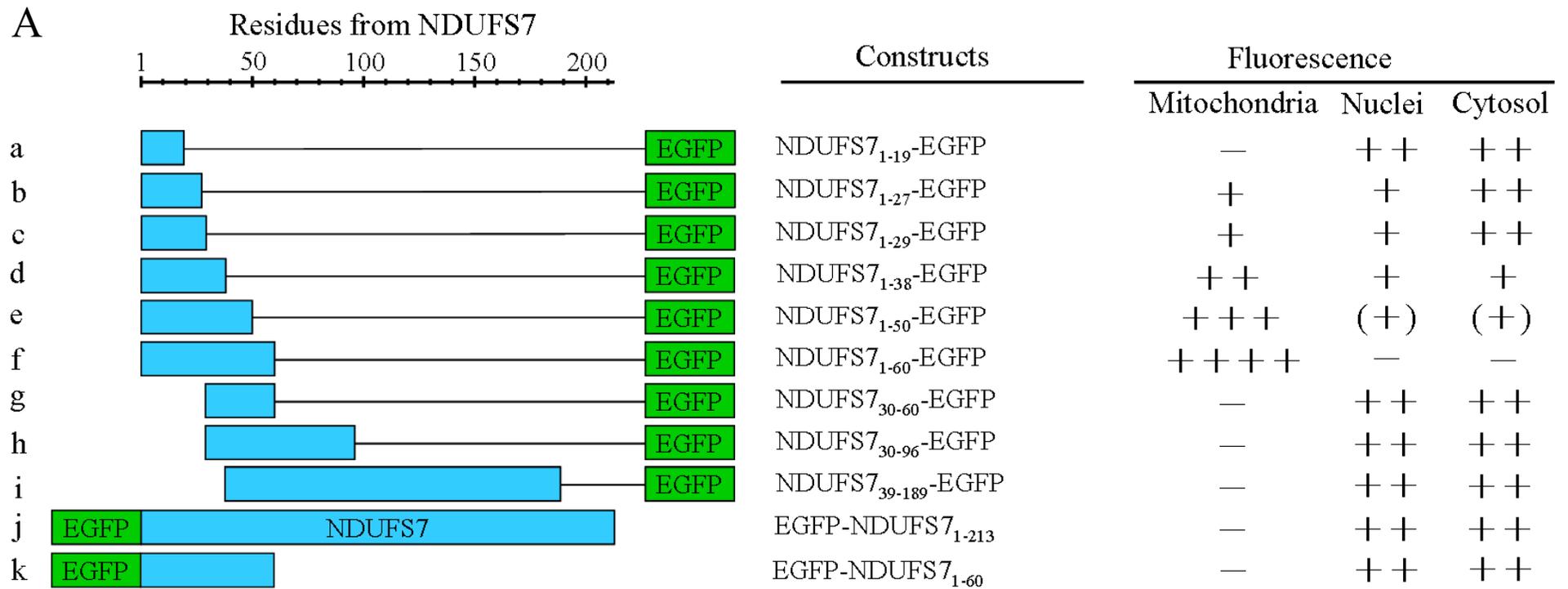
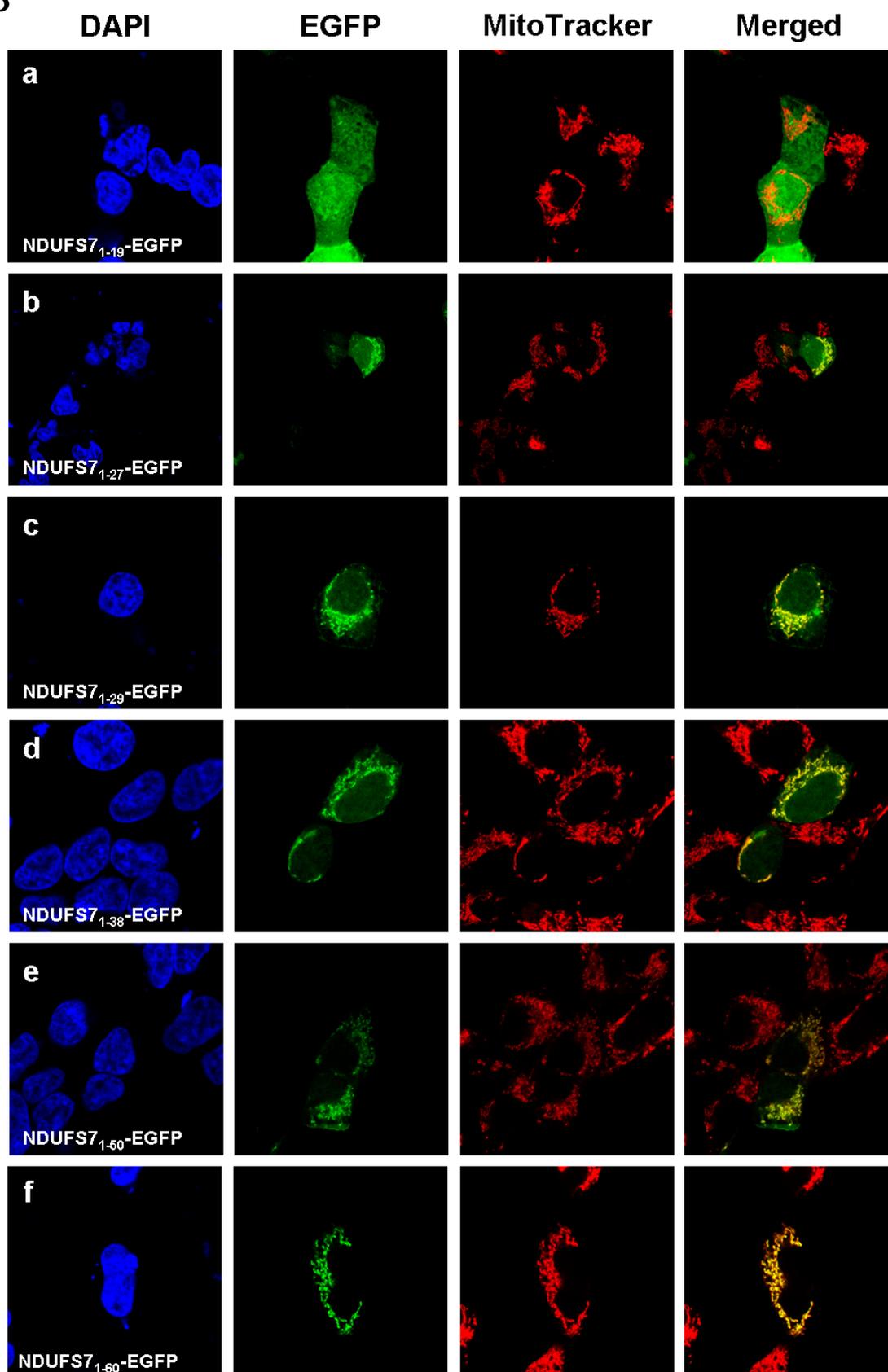
B

Figure 9. The MTS of NDUFS7 was located at N-terminus. (A) Overview of the various deletions of NDUFS7 fusing with EGFP and their mitochondrial targeting abilities. (B) The subcellular localizations of expressed fusion proteins and mitochondria were visualized by EGFP and Mito Tracker Red, respectively, in T-REx-293 (a-e) or HeLa (f) cells. Merged images of two results (EGFP and Mito Tracker Red) are shown in the right panel.





B

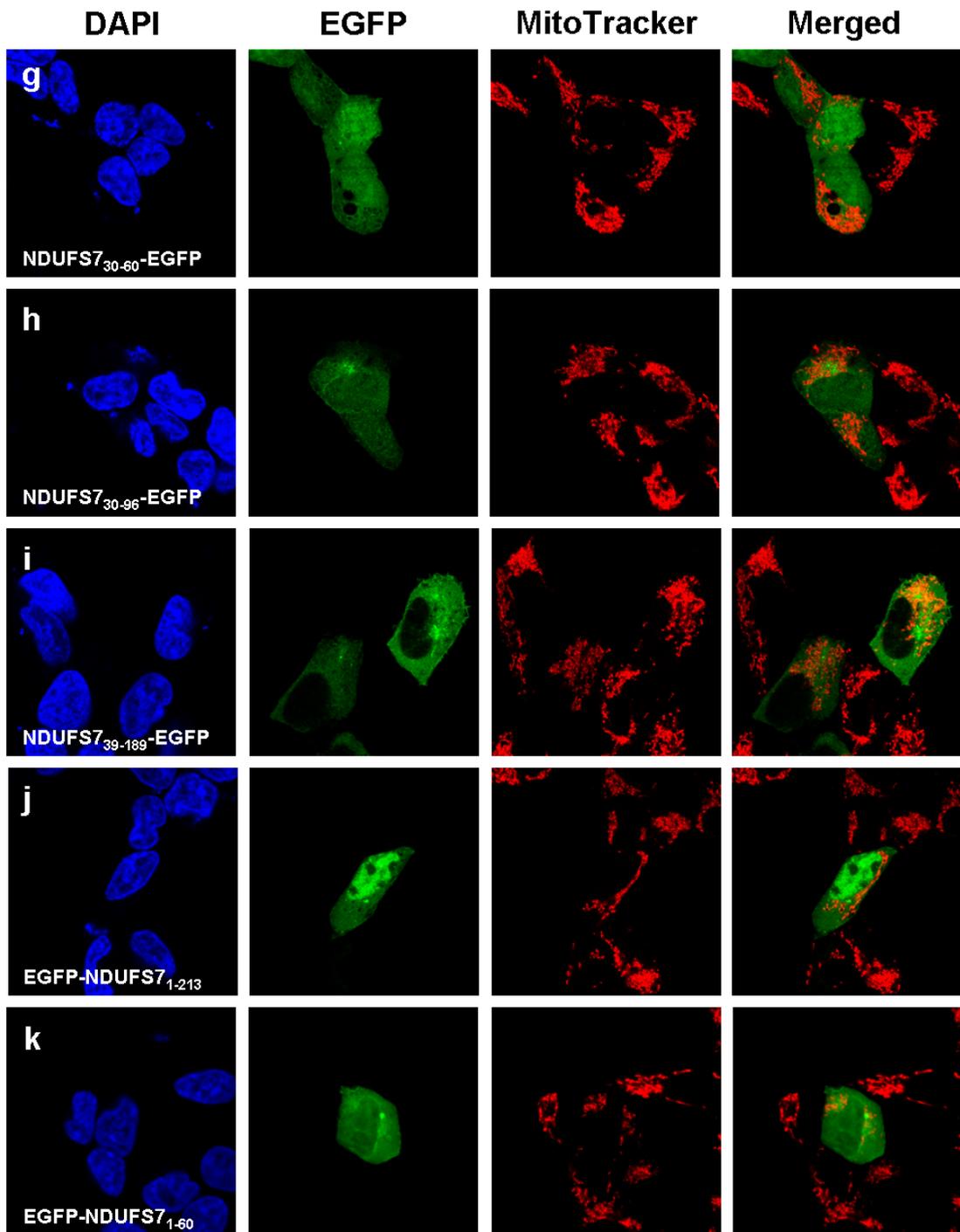


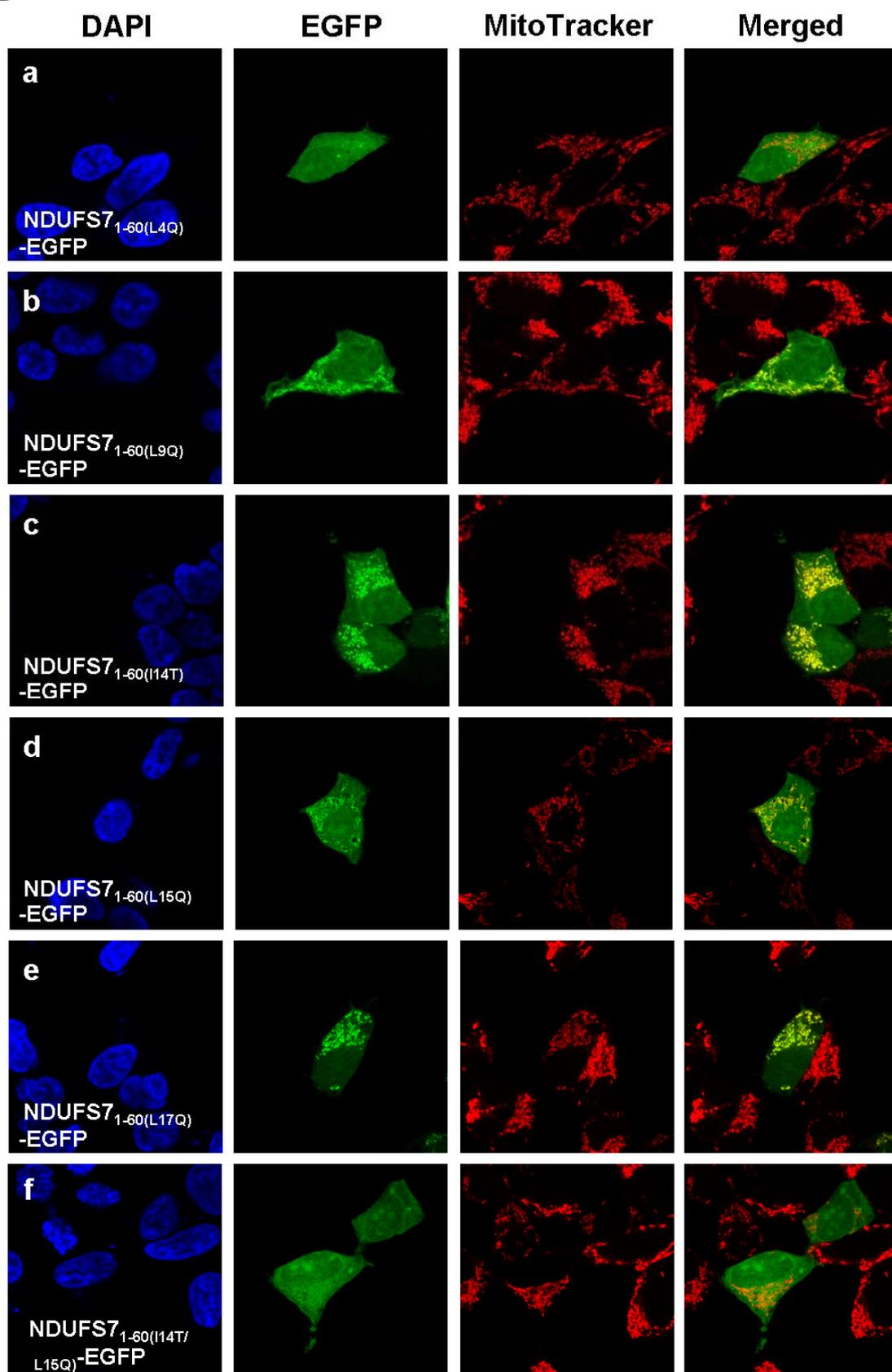
Figure 10. The effective MTS region in NDUF57 was located at amino acids 1 to 60. (A) Overview of the different N-terminal NDUF57 fragments fusing with EGFP and their mitochondrial targeting abilities. (B) The subcellular localizations of expressed fusion proteins and mitochondria were visualized by EGFP and Mito Tracker Red, respectively, in T-REx-293 cells. Merged images of two results (EGFP and Mito Tracker Red) are shown in the right panel.

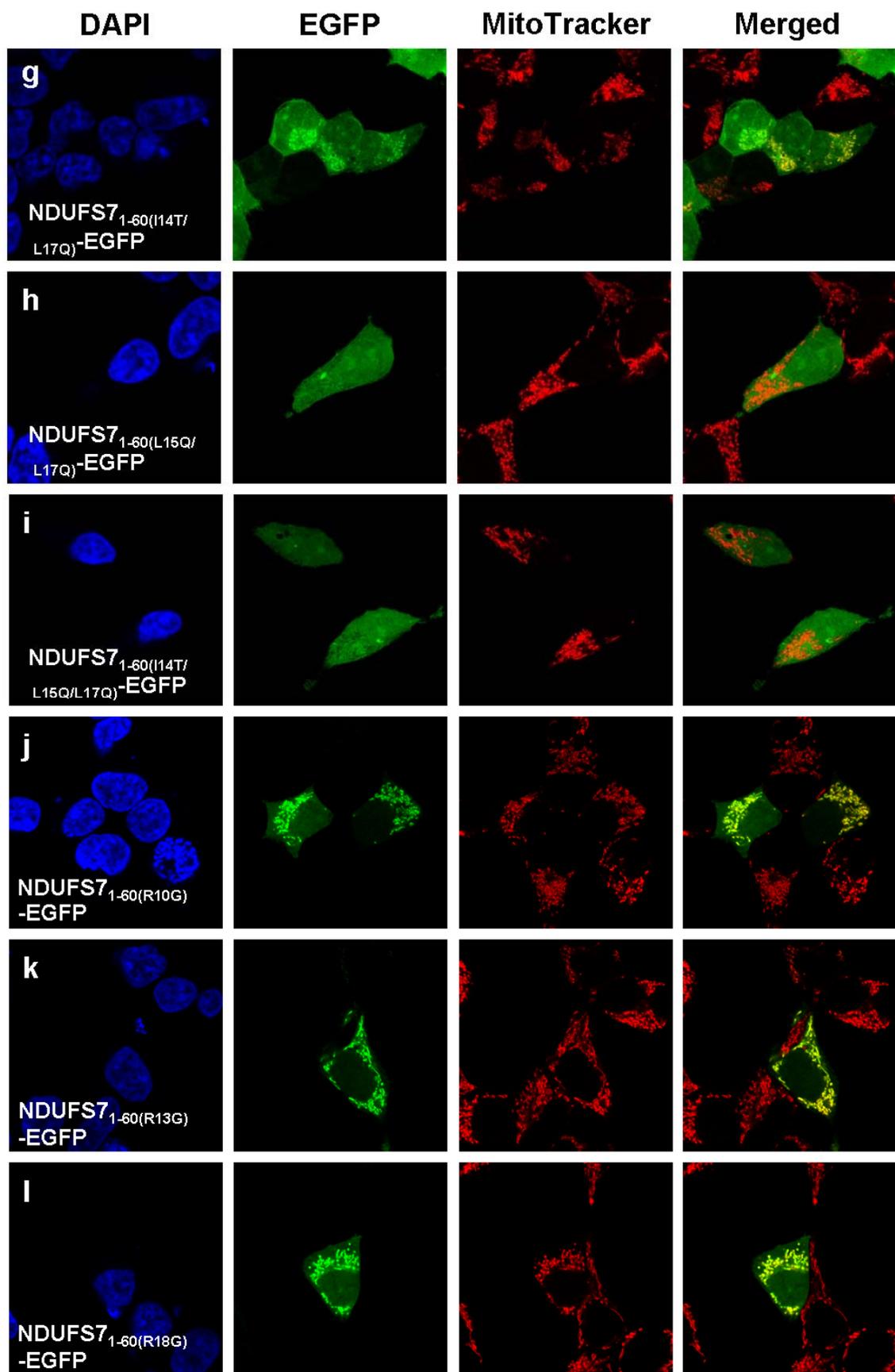


A

Residues from NDUF7		Constructs	Fluorescence		
			Mitochondria	Nuclei	Cytosol
	1 4 10 14 18 28 60 MAVLSAPGLRGFRILGLRSSVGPVQARGVH-EGFP	NDUFS7 ₁₋₆₀ -EGFP	++++	—	—
a	MAVLSAPGLRGFRILGLRSSVGPVQARGVH-EGFP	NDUFS7 _{1-60(L4Q)} -EGFP	—	++	++
b	MAVLSAPGLRGFRILGLRSSVGPVQARGVH-EGFP	NDUFS7 _{1-60(L9Q)} -EGFP	++	+	+
c	MAVLSAPGLRGFRILGLRSSVGPVQARGVH-EGFP	NDUFS7 _{1-60(I14T)} -EGFP	++	+	+
d	MAVLSAPGLRGFRILGLRSSVGPVQARGVH-EGFP	NDUFS7 _{1-60(L15Q)} -EGFP	+	+(+)	+(+)
e	MAVLSAPGLRGFRILGLRSSVGPVQARGVH-EGFP	NDUFS7 _{1-60(L17Q)} -EGFP	++	+	+
f	MAVLSAPGLRGFRILGLRSSVGPVQARGVH-EGFP	NDUFS7 _{1-60(I14T/L15Q)} -EGFP	—	++	++
g	MAVLSAPGLRGFRILGLRSSVGPVQARGVH-EGFP	NDUFS7 _{1-60(I14T/L17Q)} -EGFP	+	+(+)	+(+)
h	MAVLSAPGLRGFRILGLRSSVGPVQARGVH-EGFP	NDUFS7 _{1-60(L15Q/L17Q)} -EGFP	—	++	++
i	MAVLSAPGLRGFRILGLRSSVGPVQARGVH-EGFP	NDUFS7 _{1-60(L4Q/I14T/L15Q)} -EGFP	—	++	++
j	MAVLSAPGLRGFRILGLRSSVGPVQARGVH-EGFP	NDUFS7 _{1-60(R10G)} -EGFP	++	+	+
k	MAVLSAPGLRGFRILGLRSSVGPVQARGVH-EGFP	NDUFS7 _{1-60(R13G)} -EGFP	++	+	+
l	MAVLSAPGLRGFRILGLRSSVGPVQARGVH-EGFP	NDUFS7 _{1-60(R18G)} -EGFP	++	+	+
m	MAVLSAPGLRGFRILGLRSSVGPVQARGVH-EGFP	NDUFS7 _{1-60(R10G/R13G)} -EGFP	—	++	++
n	MAVLSAPGLRGFRILGLRSSVGPVQARGVH-EGFP	NDUFS7 _{1-60(R10G/R18G)} -EGFP	—	++	++
o	MAVLSAPGLRGFRILGLRSSVGPVQARGVH-EGFP	NDUFS7 _{1-60(R13G/R18G)} -EGFP	++	+	+
	1 4 10 14 18 27 MAVLSAPGLRGFRILGLRSSVGPVQA-EGFP	NDUFS7 ₁₋₂₇ -EGFP	+	+	++
p	MAVLSAPGLRGFRILGLRSSVGPVQA-EGFP	NDUFS7 _{1-27(R10G)} -EGFP	—	++	++
q	MAVLSAPGLRGFRILGLRSSVGPVQA-EGFP	NDUFS7 _{1-27(L4Q)} -EGFP	—	++	++
r	MAVLSAPGLRGFRILGLRSSVGPVQA-EGFP	NDUFS7 _{1-27(L15Q)} -EGFP	—	++	++

B





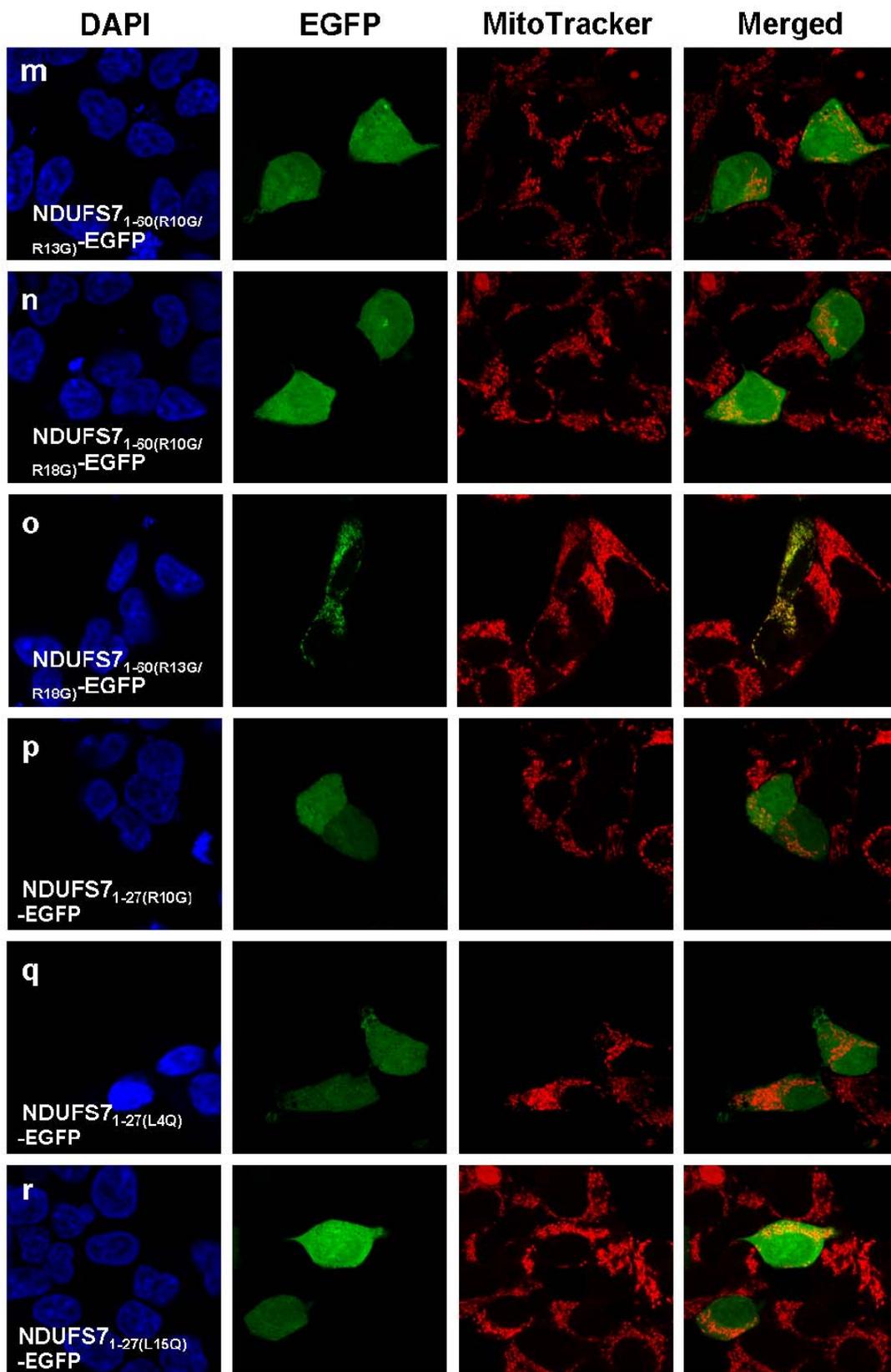
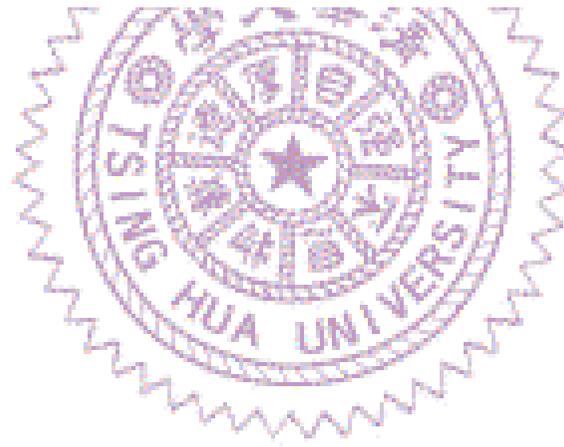


Figure 11. The characteristic of MTS in NDUFS7 was identified by site-directed mutagenesis. (A) Overview of the various site-directed mutations of NDUFS7 fusing with EGFP and their mitochondrial targeting abilities. Blue, green and yellow letters represent the basic, hydrophobic and mutated amino acids, respectively. (B) The subcellular localizations of expressed fusion proteins and mitochondria were visualized by EGFP and Mito Tracker Red, respectively, in T-REx-293 cells. Merged images of two results (EGFP and Mito Tracker Red) are shown in the right panel.



A

	Residues from NDUFS7				Constructs	Fluorescence	
	187	199	204	213		Nuclei	Cytosol
a, b	EGFP	A E A L L Y G I L Q L Q	R K I K R E R R L Q I W Y R R		EGFP-NDUFS7 ₁₈₇₋₂₁₃	++	++
c	EGFP		R K I K R E R R L Q I W Y R R		EGFP-NDUFS7 ₁₉₉₋₂₁₃	+++ (+)	(+)
d	EGFP		R K I K R E R R		EGFP-NDUFS7 ₁₉₉₋₂₀₆	++	++
e	EGFP		R K M R N T K I T R M W Y R K		EGFP-NuoB ₂₁₂₋₂₂₆	+++ (+)	(+)



B

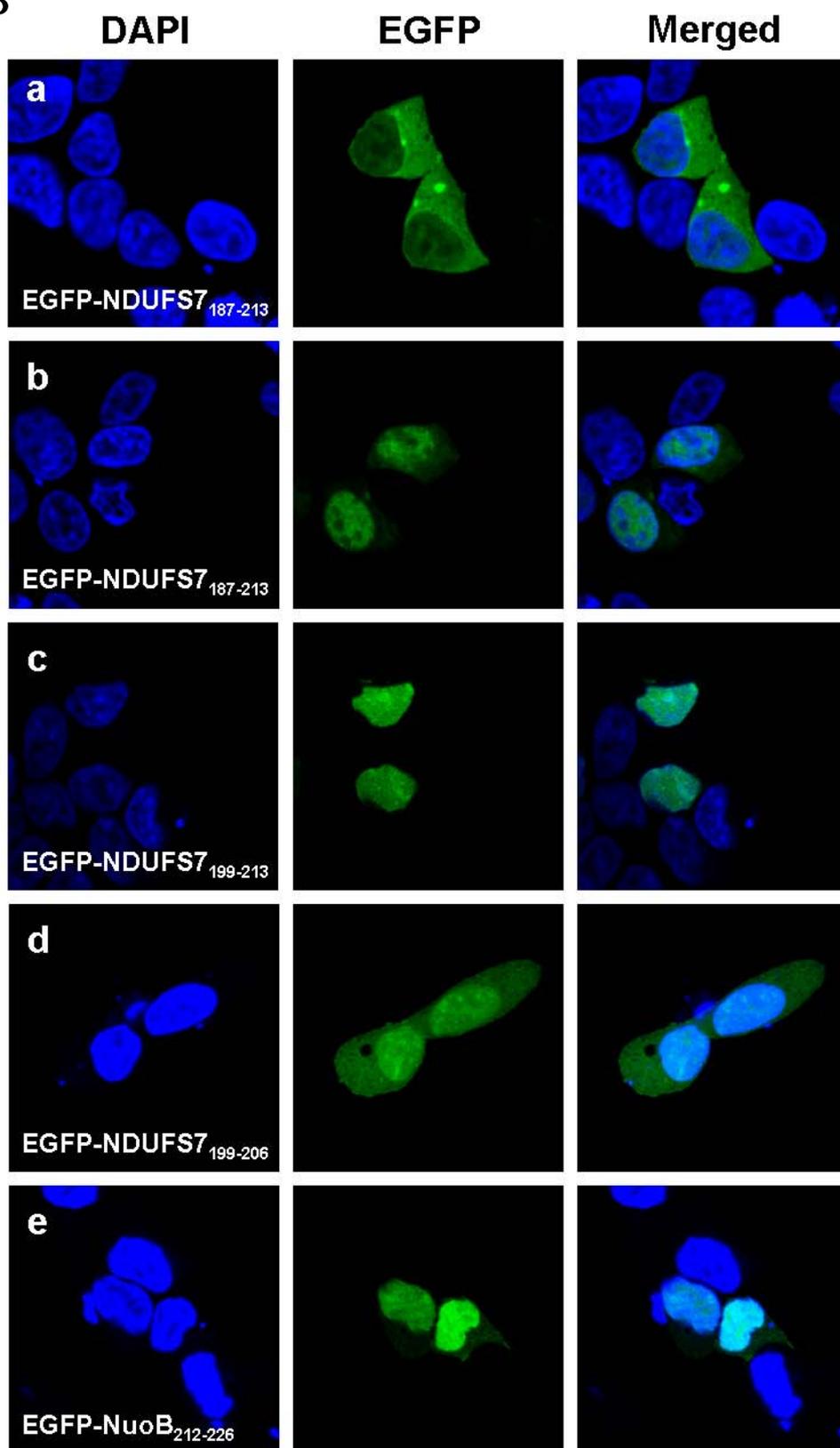
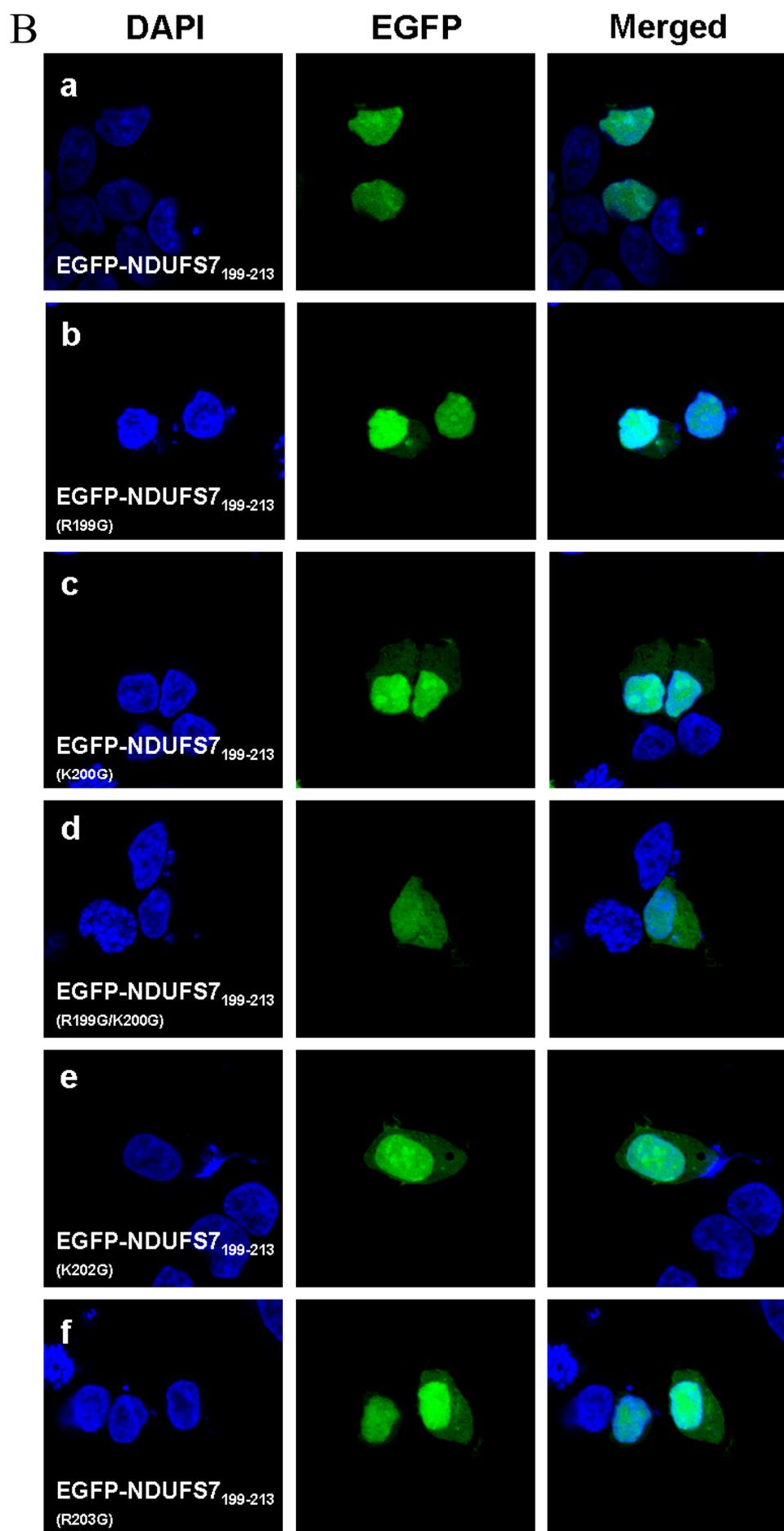
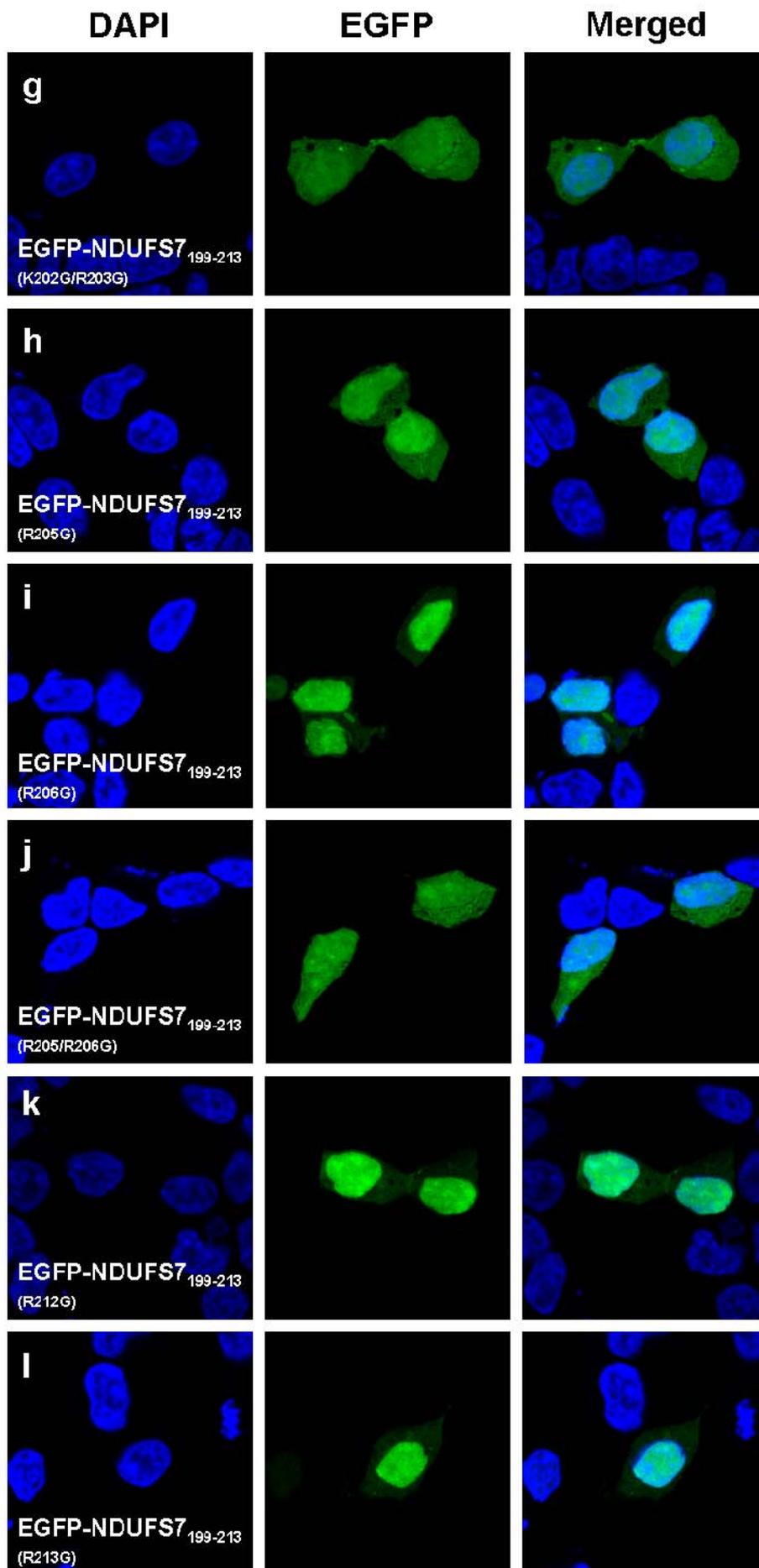


Figure 12. A competent NLS of NDUFS7 is located at the C-terminal region. (A) Overview of the various deletions and site-directed mutations of NDUFS7 fusing with EGFP and their nuclear targeting abilities. Blue, green, red and yellow letters identify the basic, hydrophobic, negative charge and mutated amino acids, respectively. (B) The subcellular localizations of expressed fusion proteins and nuclei were visualized by EGFP and DAPI, respectively, in T-REx-293 cells. Merged images of two results (DAPI and EGFP) are shown in the right panel.



A	Residues from NDUFS7			Constructs	Fluorescence				
	199	204	213		Nuclei	Cytosol			
a	EGFP	RKIKRE	RRLQI	WYRR	EGFP-NDUFS7 ₁₉₉₋₂₁₃	+++(+)	(+)		
b	EGFP	GKIKRE	RRLQI	WYRR	EGFP-NDUFS7 ₁₉₉₋₂₁₃ (R199G)	+++(+)	(+)		
c	EGFP	RGIKRE	RRLQI	WYRR	EGFP-NDUFS7 ₁₉₉₋₂₁₃ (K200G)	+++	+		
d	EGFP	GGIKRE	RRLQI	WYRR	EGFP-NDUFS7 ₁₉₉₋₂₁₃ (R199G/K200G)	++	++		
e	EGFP	RKI	GRE	RRLQI	WYRR	EGFP-NDUFS7 ₁₉₉₋₂₁₃ (K202G)	+++	+	
f	EGFP	RKIK	GER	RRLQI	WYRR	EGFP-NDUFS7 ₁₉₉₋₂₁₃ (R203G)	+++	+	
g	EGFP	RKI	GGER	RRLQI	WYRR	EGFP-NDUFS7 ₁₉₉₋₂₁₃ (K202G/R203G)	++	++	
h	EGFP	RKIKRE	GRL	QI	WYRR	EGFP-NDUFS7 ₁₉₉₋₂₁₃ (R205G)	+++	+	
i	EGFP	RKIKRE	RGL	QI	WYRR	EGFP-NDUFS7 ₁₉₉₋₂₁₃ (R206G)	++(+)	+(+)	
j	EGFP	RKIKRE	GGL	QI	WYRR	EGFP-NDUFS7 ₁₉₉₋₂₁₃ (R205G/R206G)	++	++	
k	EGFP	RKIKRE	RRLQI	WY	GR	EGFP-NDUFS7 ₁₉₉₋₂₁₃ (R212G)	+++	+	
l	EGFP	RKIKRE	RRLQI	WY	RG	EGFP-NDUFS7 ₁₉₉₋₂₁₃ (R213G)	+++	+	
m	EGFP	RKIKRE	RRLQI	WY	GG	EGFP-NDUFS7 ₁₉₉₋₂₁₃ (R212G/R213G)	+++	+	
n	EGFP	RGIKRE	RGL	QI	WYRR	EGFP-NDUFS7 ₁₉₉₋₂₁₃ (K200G/R206G)	++	++	
o	EGFP	RKIK	GER	GRL	QI	WYRR	EGFP-NDUFS7 ₁₉₉₋₂₁₃ (R203G/R206G)	++	++





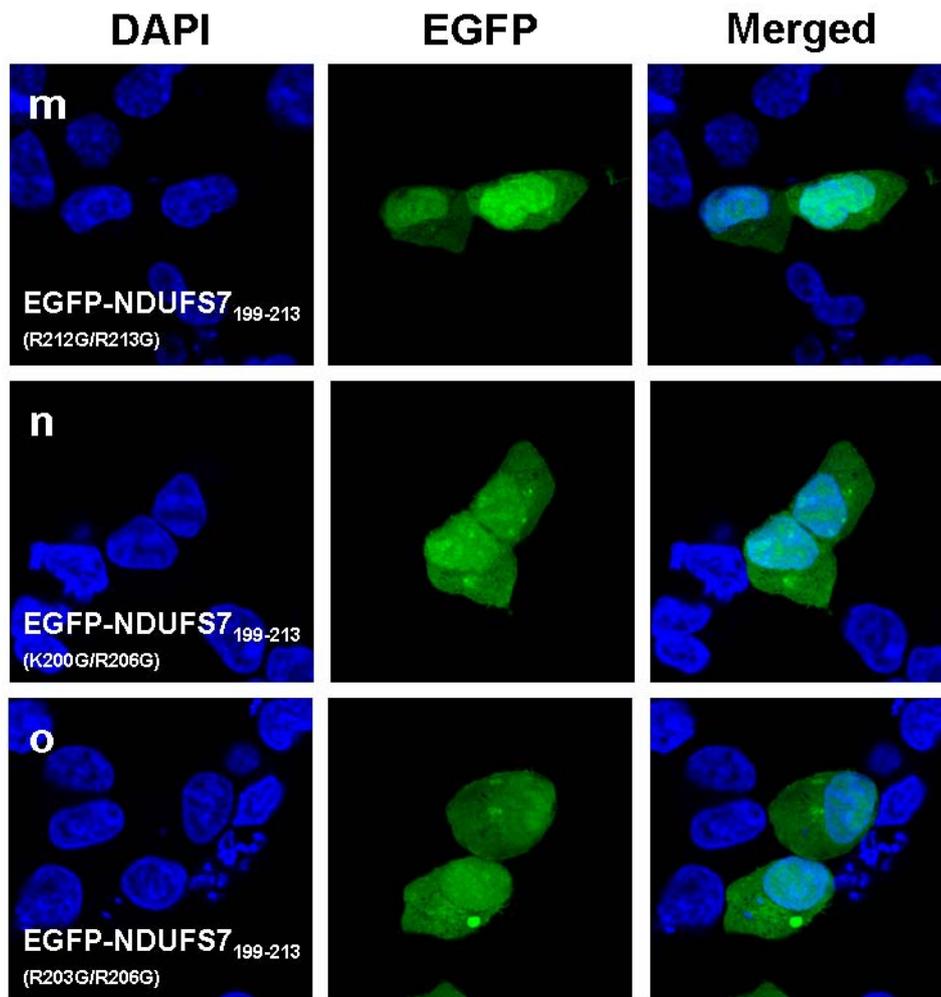
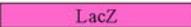
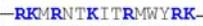


Figure 13. A competent NLS of NDUFS7 is located at the C-terminal region. (A) Overview of the various deletions and site-directed mutations of NDUFS7 fusing with EGFP and their nuclear targeting abilities. Blue, green, red and yellow letters identify the basic, hydrophobic, negative charge and mutated amino acids, respectively. (B) The subcellular localizations of expressed fusion proteins and nuclei were visualized by EGFP and DAPI, respectively, in T-REx-293 cells. Merged images of two results (DAPI and EGFP) are shown in the right panel.



A

	Constructs			Fluorescence	
				Nuclei	Cytosol
a			LacZ-c-Myc-His	—	++++
b	 - 		LacZ-NDUFS7 ₁₉₉₋₂₁₃ -c-Myc-His	++++	—
c	 - 		LacZ-NuoB ₂₁₂₋₂₂₆ -c-Myc-His	++++	—

B

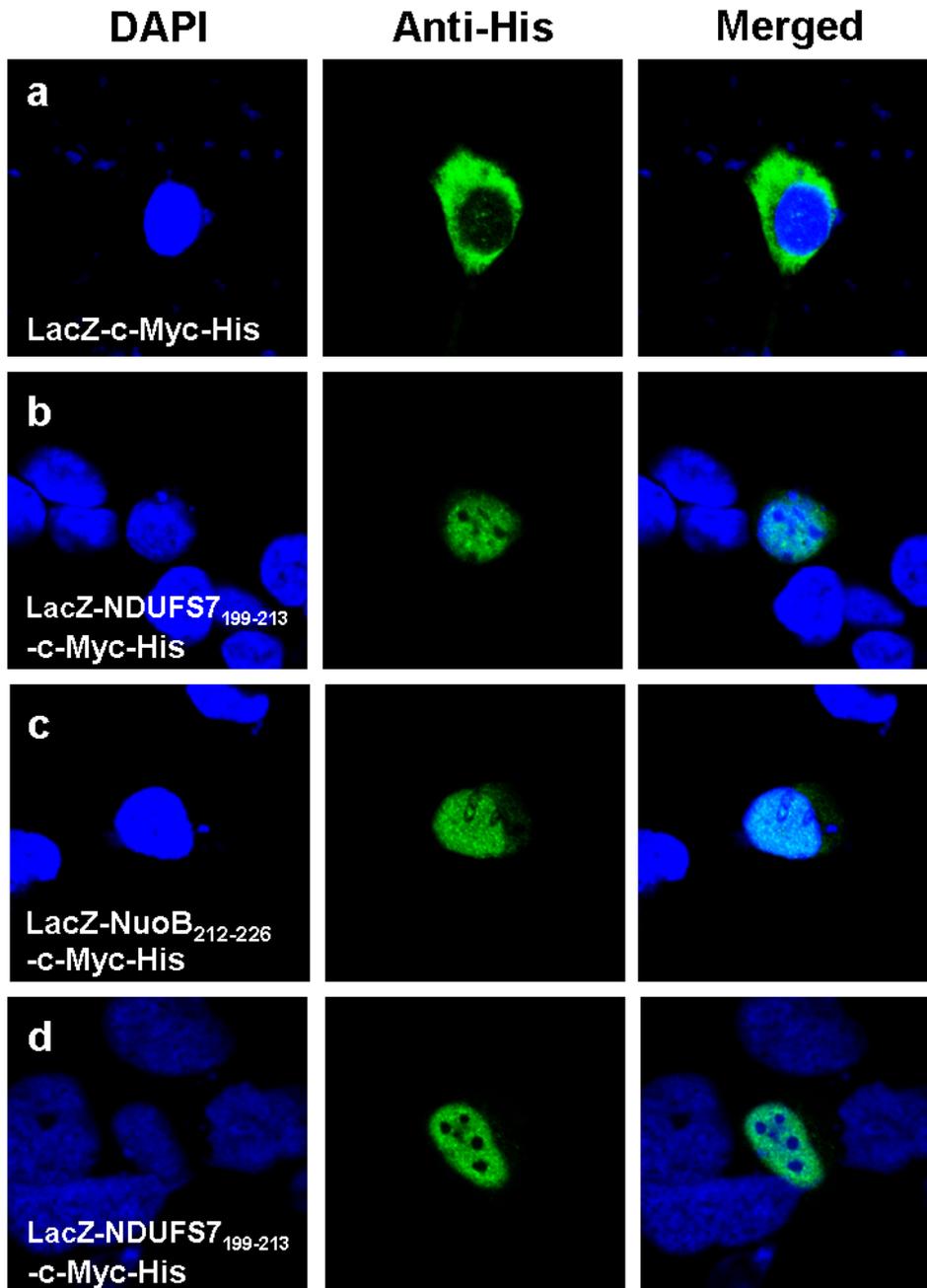
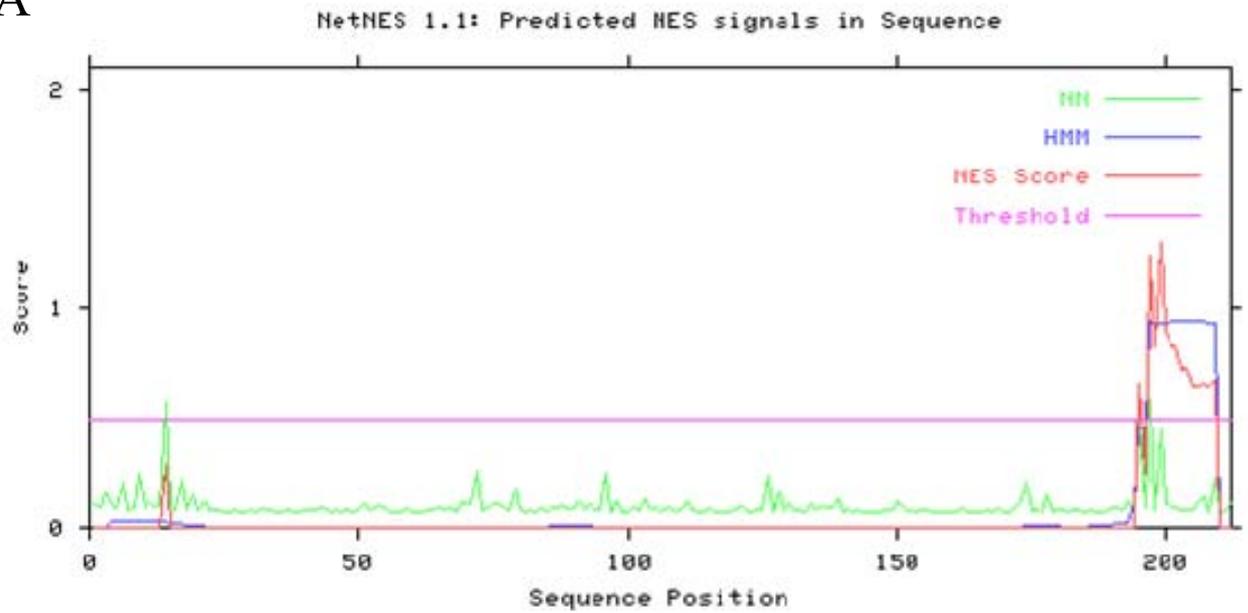


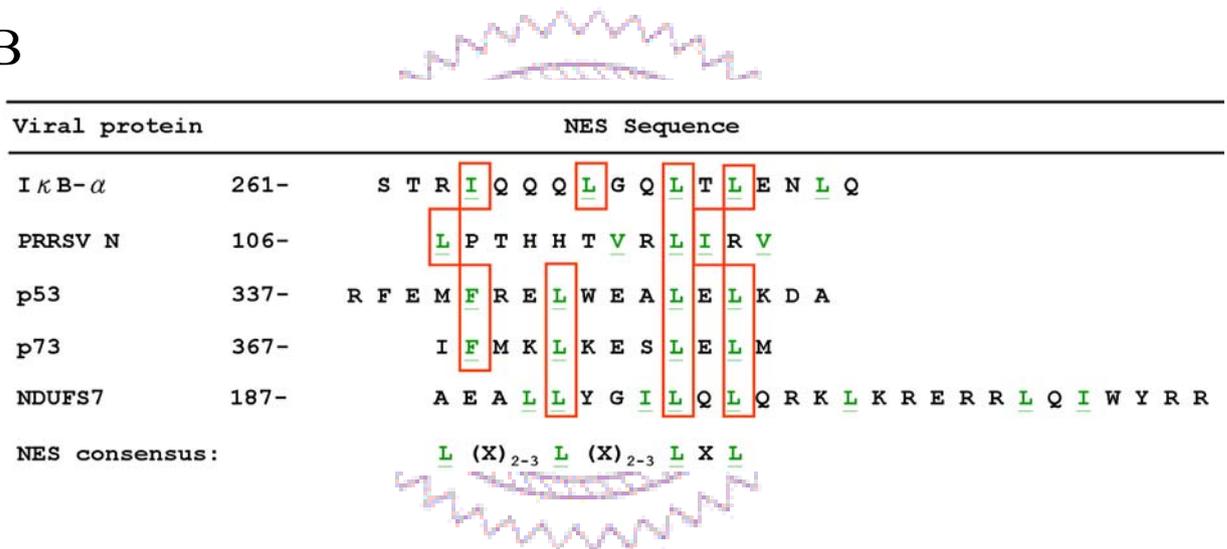
Figure 14. The NLS of NDUFS7 (from *H. sapiens*) and NuoB (from *N. crassa*) can carry LacZ into nuclei. (A) Overview of the NLS of NDUFS7 and NuoB fusing with LacZ and their nuclear targeting abilities. Blue letters identify the basic amino acids. (B) The subcellular localizations of nuclei and expressed fusion proteins were visualized by DAPI and probed with anti-His, in T-REx-293 (a-c) or HeLa (d) cells. Merged images of two results (DAPI and EGFP) are shown in the right panel.



A



B



C

	Residues from NDUFS7				Constructs	Fluorescence	
	187	199	204	213		Nuclei	Cytosol
a	EGFP	A E A L L Y G I L Q L Q R K I K R E R R L Q I W Y R R			EGFP-NDUFS7 ₁₈₇₋₂₁₃	++	++
b	EGFP	A E A L L Y G I A Q A Q R K I K R E R R L Q I W Y R R			EGFP-NDUFS7 _{187-213(L195A/L197A)}	+++	+
c	EGFP	A E A A A Y G I A Q A Q R K I K R E R R L Q I W Y R R			EGFP-NDUFS7 _{187-213(L190/191/195/197A)}	+++(+)	(+)

D

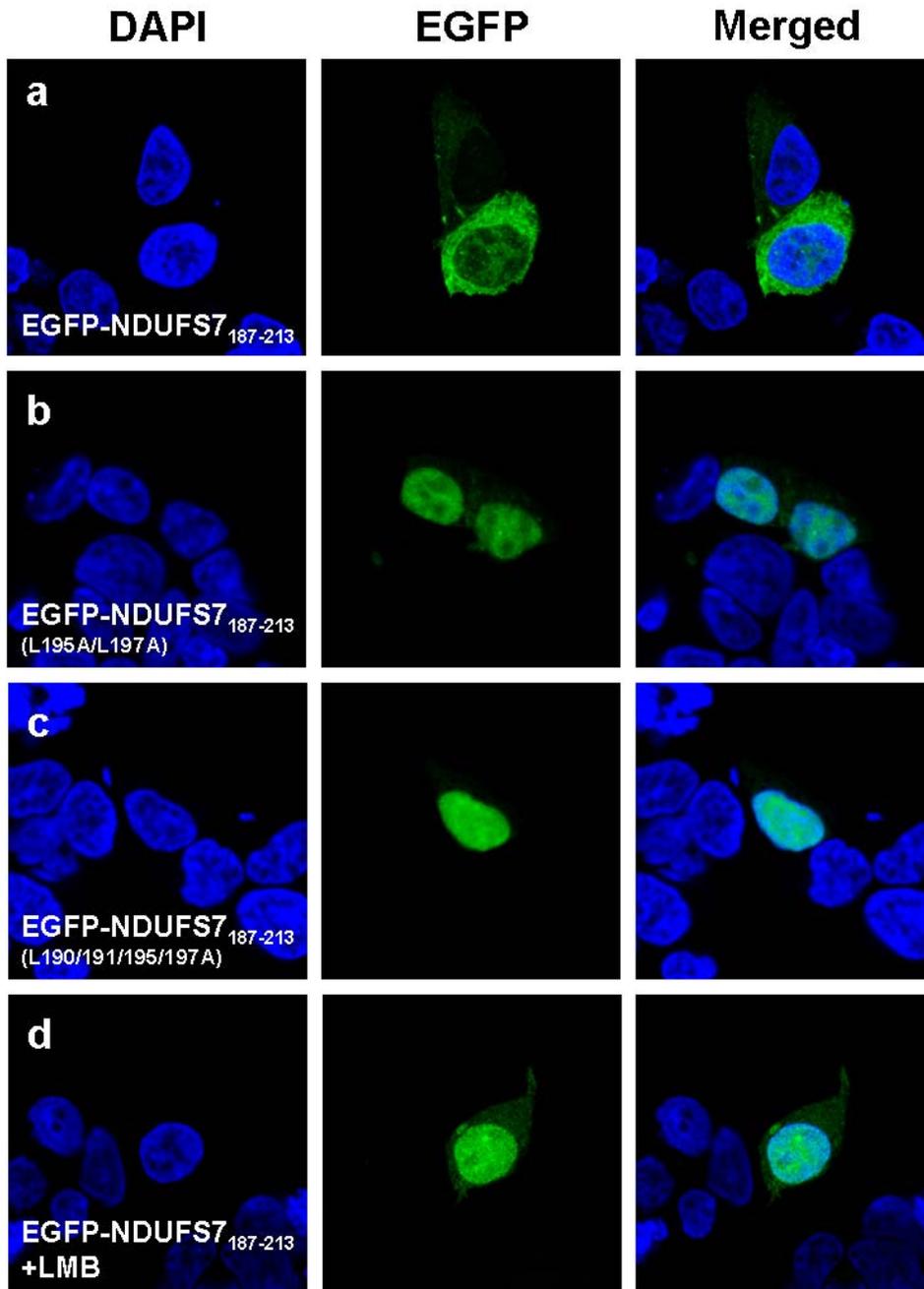
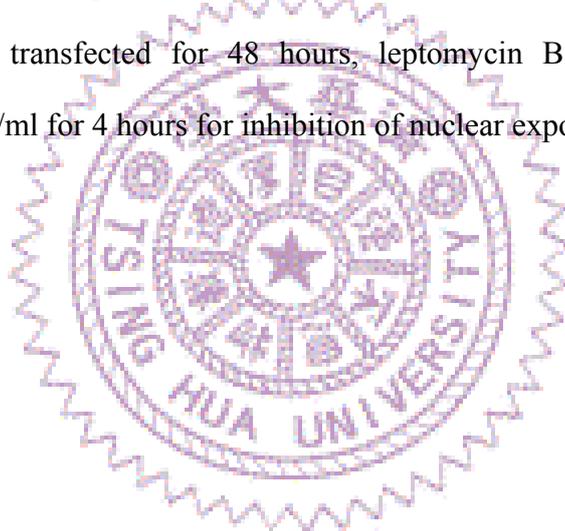


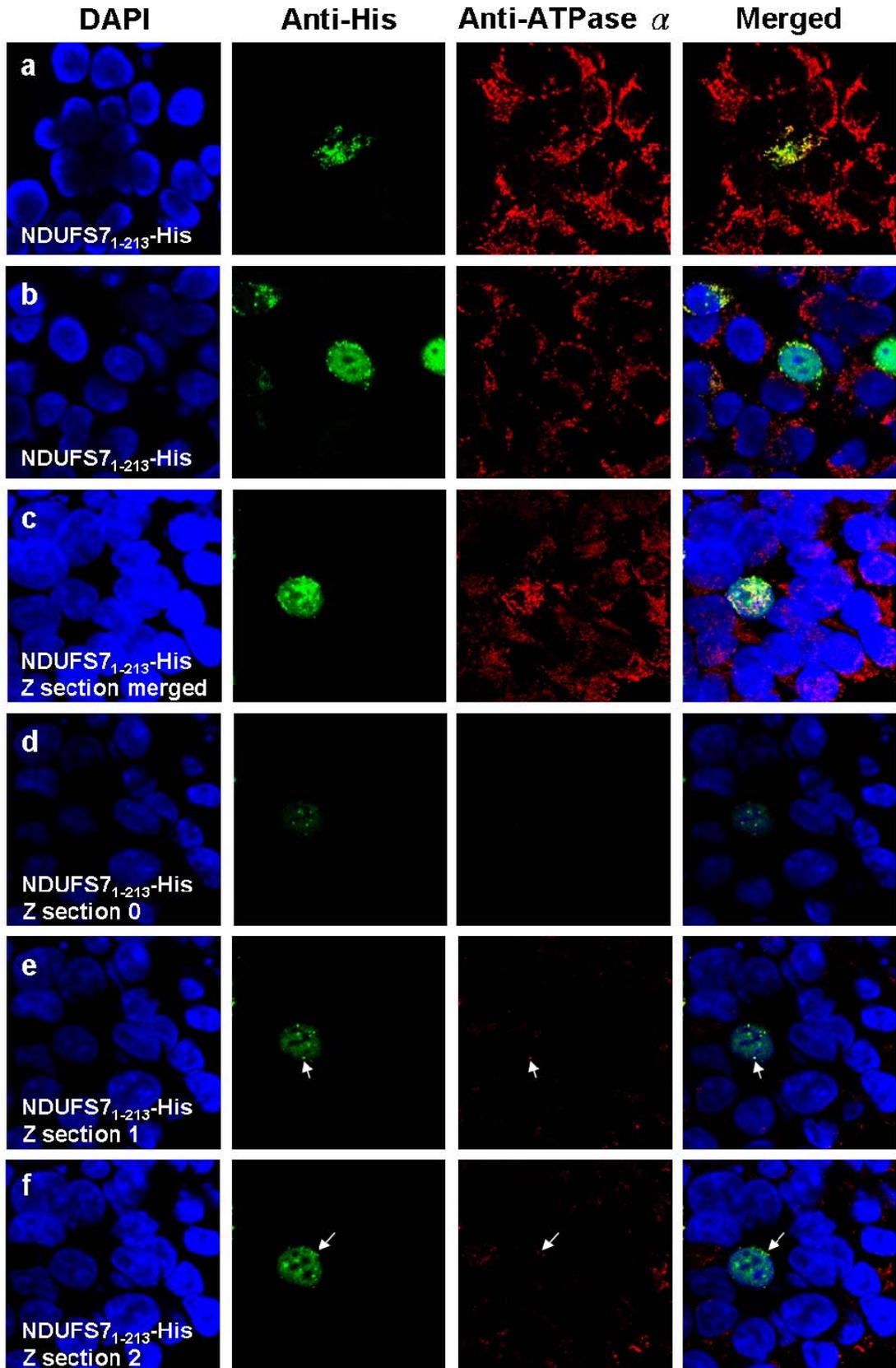
Figure 15. The C-terminal region of NDUFS7 may contain a NES. (A) The nuclear export signal (NES) prediction was carried out using NetNES 1.1 (<http://www.cbs.dtu.dk/services/NetNES/>). Blue, green, red and yellow letters identify the basic, hydrophobic, negative charge and mutated amino acids, respectively. (B) The C-terminal of NDUFS7 was compared with several known NES sequences. Red frames are the classical NES sequences and green letters are the hydrophobic amino acids with underlined. (C) Overview of the various site-directed mutations of NDUFS7 C-terminal peptide fusing with EGFP and their nuclear targeting abilities. (D) The subcellular localizations of expressed fusion proteins and nuclei were visualized by EGFP and DAPI, respectively, in T-REx-293 cells. Merged images of two results (DAPI and EGFP) are shown in right panel. (d) After cells were transfected for 48 hours, leptomycin B (LMB) was added at concentration of 20 ng/ml for 4 hours for inhibition of nuclear export.



A



B



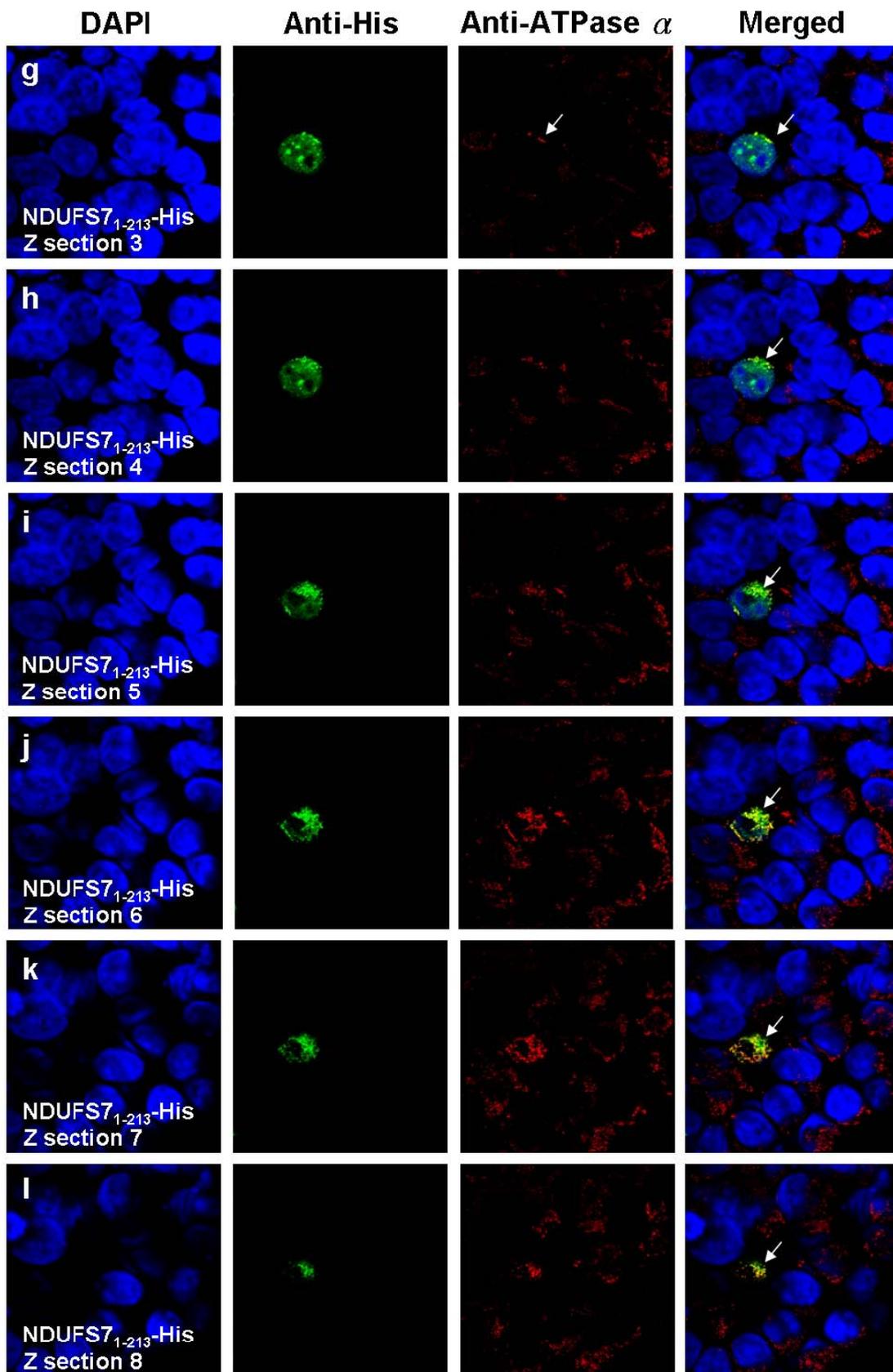


Figure 16. NDUF57 has a dual distribution both in mitochondria and nuclei. (A) Overview of NDUF57 fusing with c-Myc/His and their mitochondrial or nuclear targeting abilities. (B) The subcellular localizations of NDUF57 were probed with anti-His and anti-ATPase alpha subunit antibodies, in T-REx-293 cells. Merged images of two results (anti-His and anti-ATPase alpha) are shown in the right panel. (c) The confocal images were merged from all Z sections data. (d-l) The results of each Z section data. White arrows identify the distribution of the mitochondria.



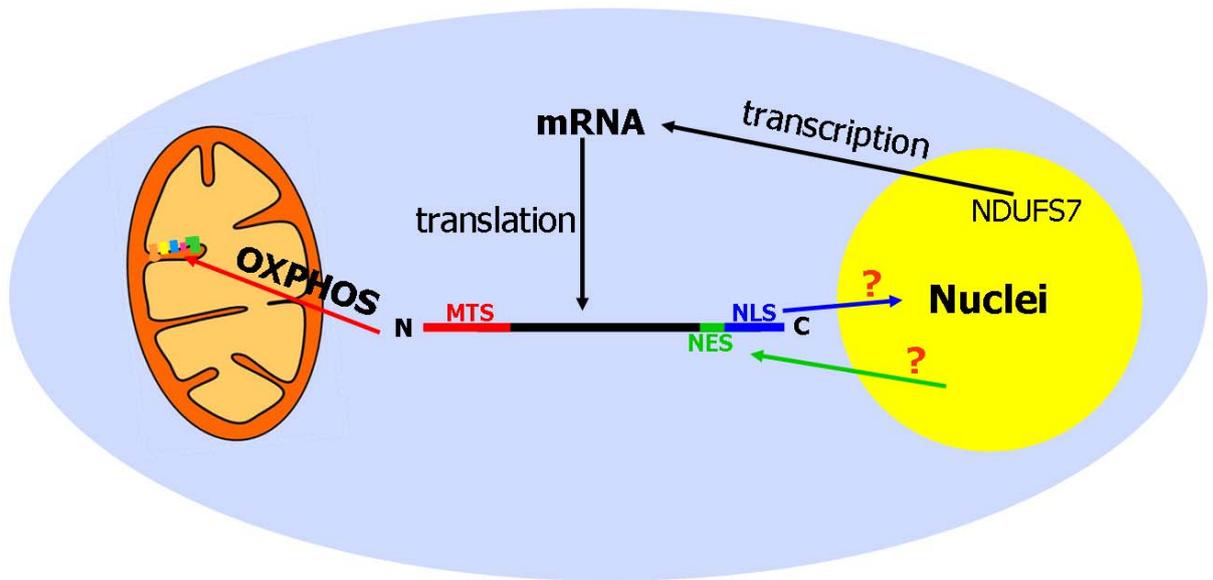
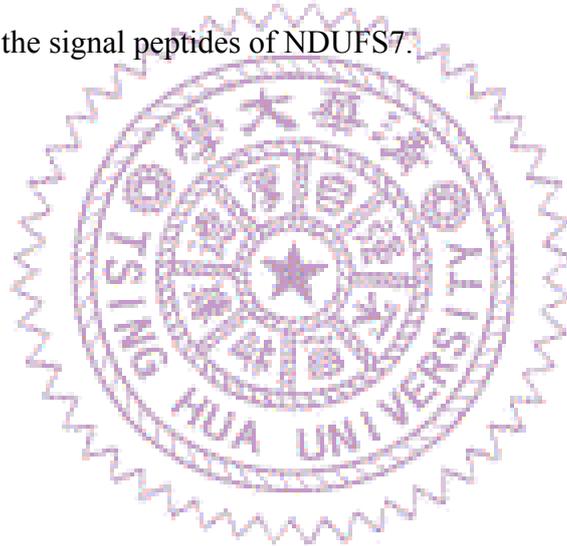


Figure 17. Overview the signal peptides of NDUFS7.



A

SUMOplot™ Prediction

Developed by Abgent, copyright 2003-2004

Protein ID:	NDUFS7
Defintion:	N/A
Length:	213 aa

■ Motifs with high probability

■ Motifs with low probability

■ Overlapping Motifs

```

1 MAVLSAPGLR GFRILGLRSS VGPAVQARGV HQSVATDGPS STQPALPKAR
51 AVAPKSSRG EYVVAKLDDL VNWARRSSLW PMTFGLACCA VEMMHMAAPR
101 YDMDFGVVF RASPRQSDVM IVAGTLINKM APALRKVYDQ MPEPRYVISM
151 GSCANGGGYY HYSYSVVRGC DRIVFVDIYI PGCPPTAEAL LYGILQLQRK
201 IKRERLQIW YRR
  
```

No.	Pos.	Group	Score
1	K202	QLQRK <u>IK</u> RE <u>RL</u> QI	0.94

No.	Pos.	Group	Score
2	K66	GEYVV <u>AK</u> LD DLVNW	0.79

B

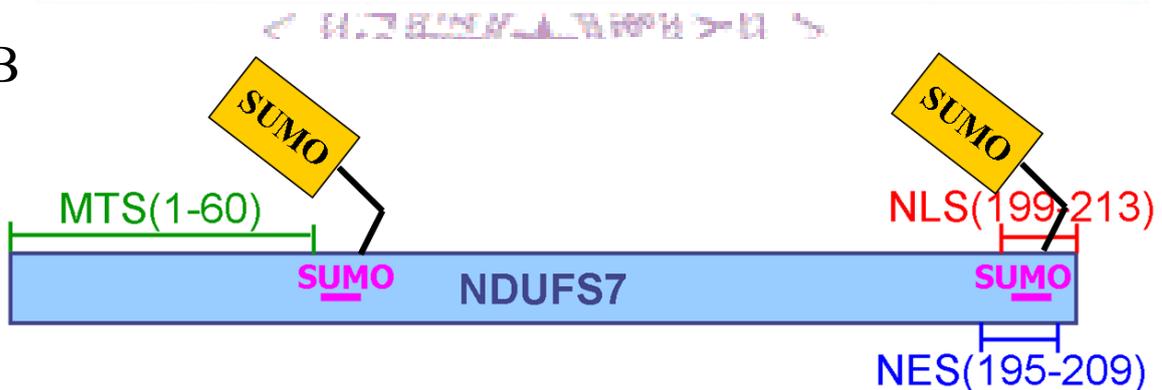


Figure 18. The SUMOylation may be the key regulation for the dual distribution of NDUFS7. (A) The SUMOylation site prediction was carried out using SUMOplot analysis program (http://www.abgent.com/tools/sumoplot_login). (B) Regions of three signal peptides and two SUMOylation sites in NDUFS7 are shown in the relative localizations.

References

1. Scheffler IE: **Mitochondria make a come back.** *Adv Drug Deliv Rev* 2001, **49**(1-2):3-26.
2. Taylor RW, Turnbull DM: **Mitochondrial DNA mutations in human disease.** *Nat Rev Genet* 2005, **6**(5):389-402.
3. Taanman JW: **The mitochondrial genome: structure, transcription, translation and replication.** *Biochim Biophys Acta* 1999, **1410**(2):103-123.
4. Janssen RJ, Nijtmans LG, van den Heuvel LP, Smeitink JA: **Mitochondrial complex I: structure, function and pathology.** *J Inherit Metab Dis* 2006, **29**(4):499-515.
5. Ohnishi T: **Iron-sulfur clusters/semiquinones in complex I.** *Biochim Biophys Acta* 1998, **1364**(2):186-206.
6. Hyslop SJ, Duncan AM, Pitkanen S, Robinson BH: **Assignment of the PSST subunit gene of human mitochondrial complex I to chromosome 19p13.** *Genomics* 1996, **37**(3):375-380.
7. Leigh D: **Subacute necrotizing encephalomyelopathy in an infant.** *J Neurol Neurosurg Psychiatry* 1951, **14**(3):216-221.
8. Dahl HH: **Getting to the nucleus of mitochondrial disorders: identification of respiratory chain-enzyme genes causing Leigh syndrome.** *Am J Hum Genet* 1998, **63**(6):1594-1597.
9. Smeitink J, van den Heuvel L: **Human mitochondrial complex I in health and disease.** *Am J Hum Genet* 1999, **64**(6):1505-1510.
10. Triepels RH, van den Heuvel LP, Loeffen JL, Buskens CA, Smeets RJ, Rubio Gozalbo ME, Budde SM, Mariman EC, Wijburg FA, Barth PG *et al*: **Leigh syndrome associated with a mutation in the NDUFS7 (PSST) nuclear encoded subunit of complex I.** *Ann Neurol* 1999, **45**(6):787-790.
11. Lebon S, Rodriguez D, Bridoux D, Zerrad A, Rotig A, Munnich A, Legrand A, Slama A: **A novel mutation in the human complex I NDUFS7 subunit associated with Leigh syndrome.** *Mol Genet Metab* 2007, **90**(4):379-382.
12. Lebon S, Minai L, Chretien D, Corcos J, Serre V, Kadhom N, Steffann J, Pauchard JY, Munnich A, Bonnefont JP *et al*: **A novel mutation of the NDUFS7 gene leads to activation of a cryptic exon and impaired assembly of mitochondrial complex I in a patient with Leigh syndrome.** *Mol Genet Metab* 2007, **92**(1-2):104-108.
13. Ahlers PM, Garofano A, Kerscher SJ, Brandt U: **Application of the obligate aerobic yeast *Yarrowia lipolytica* as a eucaryotic model to analyse Leigh syndrome mutations in the complex I core subunits PSST and TYKY.**

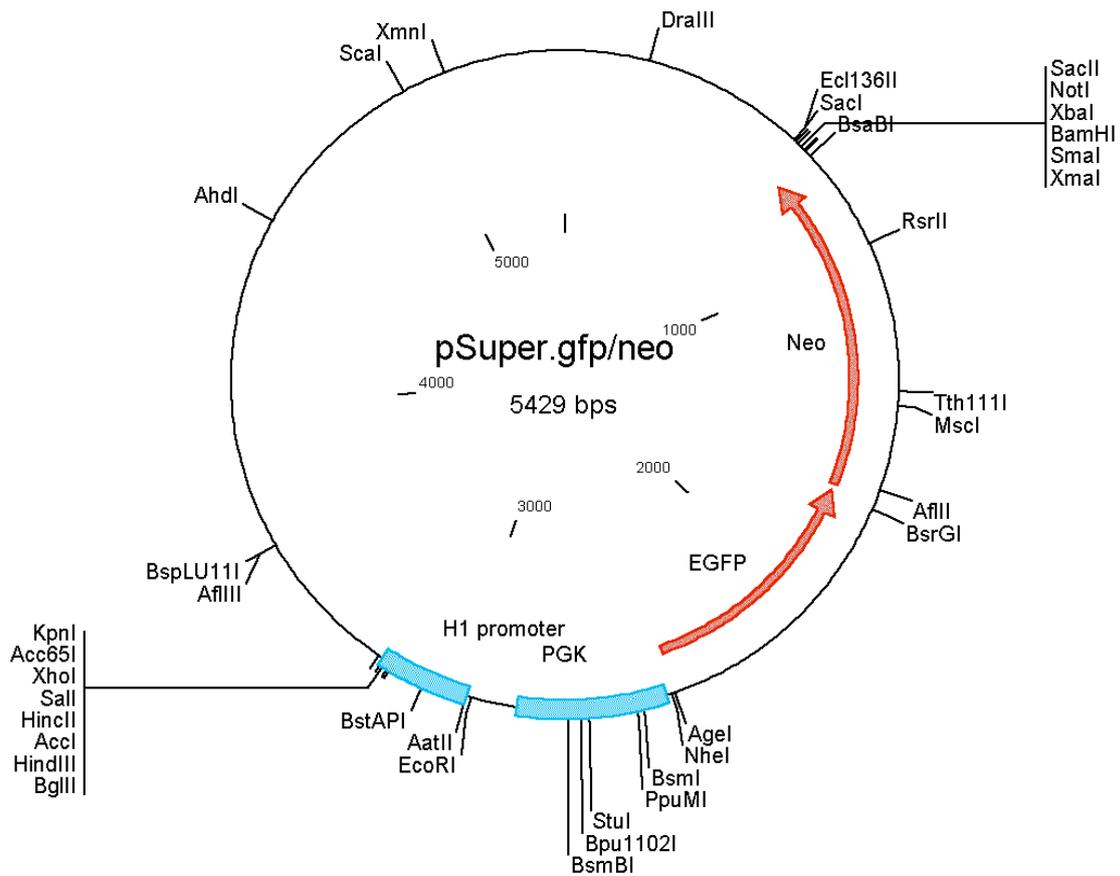
- Biochim Biophys Acta* 2000, **1459**(2-3):258-265.
14. Wang DC, Meinhardt SW, Sackmann U, Weiss H, Ohnishi T: **The iron-sulfur clusters in the two related forms of mitochondrial NADH: ubiquinone oxidoreductase made by *Neurospora crassa***. *Eur J Biochem* 1991, **197**(1):257-264.
 15. Duarte M, Populo H, Videira A, Friedrich T, Schulte U: **Disruption of iron-sulphur cluster N2 from NADH: ubiquinone oxidoreductase by site-directed mutagenesis**. *Biochem J* 2002, **364**(Pt 3):833-839.
 16. Friedrich T, Steinmuller K, Weiss H: **The proton-pumping respiratory complex I of bacteria and mitochondria and its homologue in chloroplasts**. *FEBS Lett* 1995, **367**(2):107-111.
 17. Gakh O, Cavadini P, Isaya G: **Mitochondrial processing peptidases**. *Biochim Biophys Acta* 2002, **1592**(1):63-77.
 18. Neupert W, Herrmann JM: **Translocation of proteins into mitochondria**. *Annu Rev Biochem* 2007, **76**:723-749.
 19. Pfanner N, Geissler A: **Versatility of the mitochondrial protein import machinery**. *Nat Rev Mol Cell Biol* 2001, **2**(5):339-349.
 20. Mukhopadhyay A, Yang CS, Wei B, Weiner H: **Precursor protein is readily degraded in mitochondrial matrix space if the leader is not processed by mitochondrial processing peptidase**. *J Biol Chem* 2007, **282**(51):37266-37275.
 21. Lee CM, Sedman J, Neupert W, Stuart RA: **The DNA helicase, Hmi1p, is transported into mitochondria by a C-terminal cleavable targeting signal**. *J Biol Chem* 1999, **274**(30):20937-20942.
 22. Roise D, Theiler F, Horvath SJ, Tomich JM, Richards JH, Allison DS, Schatz G: **Amphiphilicity is essential for mitochondrial presequence function**. *EMBO J* 1988, **7**(3):649-653.
 23. Allison DS, Schatz G: **Artificial mitochondrial presequences**. *Proc Natl Acad Sci U S A* 1986, **83**(23):9011-9015.
 24. Folsch H, Gaume B, Brunner M, Neupert W, Stuart RA: **C- to N-terminal translocation of preproteins into mitochondria**. *EMBO J* 1998, **17**(22):6508-6515.
 25. Arizmendi JM, Runswick MJ, Skehel JM, Walker JE: **NADH: ubiquinone oxidoreductase from bovine heart mitochondria. A fourth nuclear encoded subunit with a homologue encoded in chloroplast genomes**. *FEBS Lett* 1992, **301**(3):237-242.
 26. Duarte M, Finel M, Videira A: **Primary structure of a ferredoxin-like iron-sulfur subunit of complex I from *Neurospora crassa***. *Biochim Biophys*

- Acta* 1996, **1275**(3):151-153.
27. Schmidt-Zachmann MS, Nigg EA: **Protein localization to the nucleolus: a search for targeting domains in nucleolin.** *J Cell Sci* 1993, **105** (Pt 3):799-806.
 28. Nigg EA: **Nucleocytoplasmic transport: signals, mechanisms and regulation.** *Nature* 1997, **386**(6627):779-787.
 29. Kalderon D, Roberts BL, Richardson WD, Smith AE: **A short amino acid sequence able to specify nuclear location.** *Cell* 1984, **39**(3 Pt 2):499-509.
 30. Kalderon D, Smith AE: **In vitro mutagenesis of a putative DNA binding domain of SV40 large-T.** *Virology* 1984, **139**(1):109-137.
 31. Lanford RE, Butel JS: **Construction and characterization of an SV40 mutant defective in nuclear transport of T antigen.** *Cell* 1984, **37**(3):801-813.
 32. Jans DA, Xiao CY, Lam MH: **Nuclear targeting signal recognition: a key control point in nuclear transport?** *Bioessays* 2000, **22**(6):532-544.
 33. Robbins J, Dilworth SM, Laskey RA, Dingwall C: **Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence.** *Cell* 1991, **64**(3):615-623.
 34. Hall MN, Hereford L, Herskowitz I: **Targeting of E. coli beta-galactosidase to the nucleus in yeast.** *Cell* 1984, **36**(4):1057-1065.
 35. Hicks GR, Raikhel NV: **Protein import into the nucleus: an integrated view.** *Annu Rev Cell Dev Biol* 1995, **11**:155-188.
 36. Chan CK, Jans DA: **Using nuclear targeting signals to enhance non-viral gene transfer.** *Immunol Cell Biol* 2002, **80**(2):119-130.
 37. Ketha KM, Atreya CD: **Application of bioinformatics-coupled experimental analysis reveals a new transport-competent nuclear localization signal in the nucleoprotein of influenza A virus strain.** *BMC Cell Biol* 2008, **9**:22.
 38. Sorokin AV, Kim ER, Ovchinnikov LP: **Nucleocytoplasmic transport of proteins.** *Biochemistry (Mosc)* 2007, **72**(13):1439-1457.
 39. Fischer U, Huber J, Boelens WC, Mattaj JW, Luhrmann R: **The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs.** *Cell* 1995, **82**(3):475-483.
 40. la Cour T, Gupta R, Rapacki K, Skriver K, Poulsen FM, Brunak S: **NESbase version 1.0: a database of nuclear export signals.** *Nucleic Acids Res* 2003, **31**(1):393-396.
 41. Wen W, Taylor SS, Meinkoth JL: **The expression and intracellular distribution of the heat-stable protein kinase inhibitor is cell cycle**

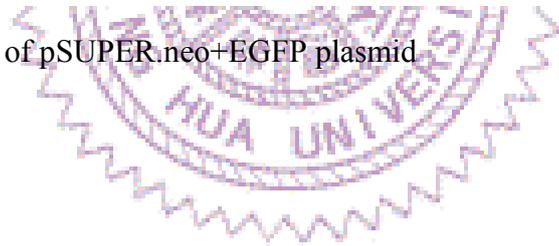
- regulated.** *J Biol Chem* 1995, **270**(5):2041-2046.
42. Johnson AW, Lund E, Dahlberg J: **Nuclear export of ribosomal subunits.** *Trends Biochem Sci* 2002, **27**(11):580-585.
43. Engel K, Kotlyarov A, Gaestel M: **Leptomycin B-sensitive nuclear export of MAPKAP kinase 2 is regulated by phosphorylation.** *EMBO J* 1998, **17**(12):3363-3371.
44. DeVit MJ, Johnston M: **The nuclear exportin Msn5 is required for nuclear export of the Mig1 glucose repressor of *Saccharomyces cerevisiae*.** *Curr Biol* 1999, **9**(21):1231-1241.
45. Kaffman A, Rank NM, O'Neill EM, Huang LS, O'Shea EK: **The receptor Msn5 exports the phosphorylated transcription factor Pho4 out of the nucleus.** *Nature* 1998, **396**(6710):482-486.
46. Kaffman A, O'Shea EK: **Regulation of nuclear localization: a key to a door.** *Annu Rev Cell Dev Biol* 1999, **15**:291-339.
47. Freemantle SJ, Taylor SM, Krystal G, Moran RG: **Upstream organization of and multiple transcripts from the human folylpoly-gamma-glutamate synthetase gene.** *J Biol Chem* 1995, **270**(16):9579-9584.
48. Otterlei M, Haug T, Nagelhus TA, Slupphaug G, Lindmo T, Krokan HE: **Nuclear and mitochondrial splice forms of human uracil-DNA glycosylase contain a complex nuclear localisation signal and a strong classical mitochondrial localisation signal, respectively.** *Nucleic Acids Res* 1998, **26**(20):4611-4617.
49. Land T, Rouault TA: **Targeting of a human iron-sulfur cluster assembly enzyme, nifs, to different subcellular compartments is regulated through alternative AUG utilization.** *Mol Cell* 1998, **2**(6):807-815.
50. Lutsenko S, Cooper MJ: **Localization of the Wilson's disease protein product to mitochondria.** *Proc Natl Acad Sci U S A* 1998, **95**(11):6004-6009.
51. Ashmarina LI, Pshezhetsky AV, Branda SS, Isaya G, Mitchell GA: **3-Hydroxy-3-methylglutaryl coenzyme A lyase: targeting and processing in peroxisomes and mitochondria.** *J Lipid Res* 1999, **40**(1):70-75.
52. Amery L, Fransen M, De Nys K, Mannaerts GP, Van Veldhoven PP: **Mitochondrial and peroxisomal targeting of 2-methylacyl-CoA racemase in humans.** *J Lipid Res* 2000, **41**(11):1752-1759.
53. Ramotar D, Kim C, Lillis R, Dempfle B: **Intracellular localization of the Apn1 DNA repair enzyme of *Saccharomyces cerevisiae*. Nuclear transport signals and biological role.** *J Biol Chem* 1993, **268**(27):20533-20539.
54. Chew O, Rudhe C, Glaser E, Whelan J: **Characterization of the targeting signal of dual-targeted pea glutathione reductase.** *Plant Mol Biol* 2003,

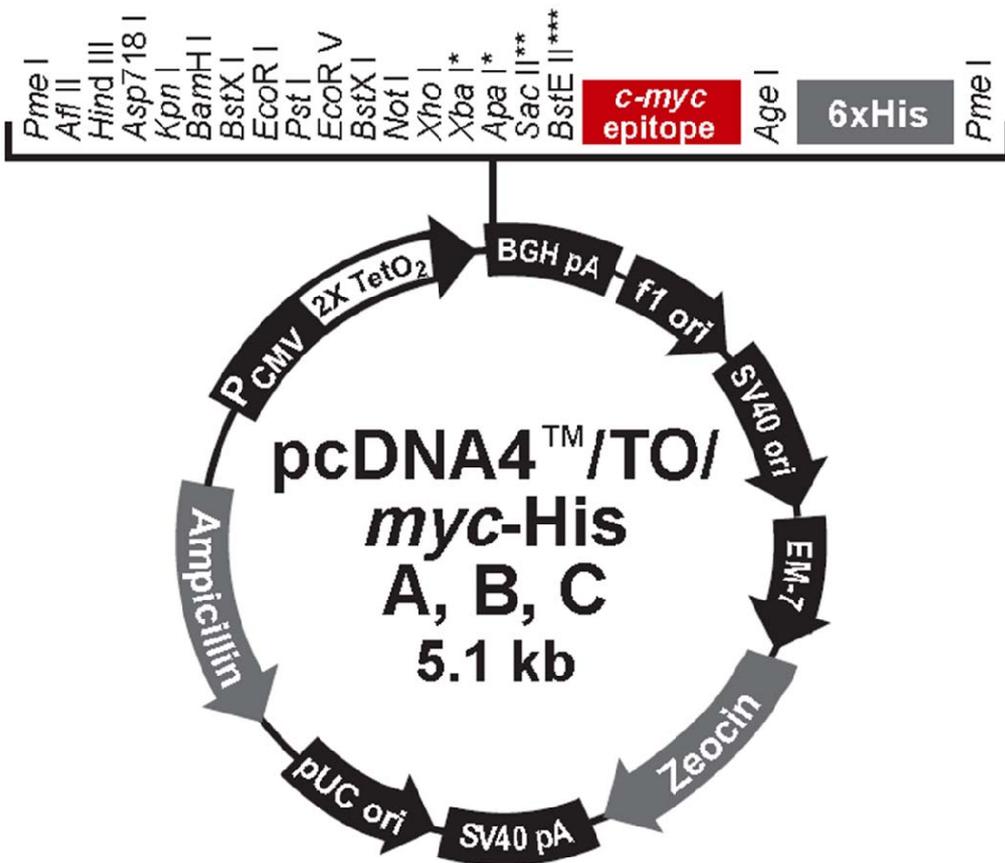
- 53(3):341-356.
55. Bradbury MW, Berk PD: **Mitochondrial aspartate aminotransferase: direction of a single protein with two distinct functions to two subcellular sites does not require alternative splicing of the mRNA.** *Biochem J* 2000, **345 Pt 3**:423-427.
56. Prokisch H, Scharfe C, Camp DG, 2nd, Xiao W, David L, Andreoli C, Monroe ME, Moore RJ, Gritsenko MA, Kozany C *et al*: **Integrative analysis of the mitochondrial proteome in yeast.** *PLoS Biol* 2004, **2(6)**:e160.
57. Tolkunova E, Park H, Xia J, King MP, Davidson E: **The human lysyl-tRNA synthetase gene encodes both the cytoplasmic and mitochondrial enzymes by means of an unusual alternative splicing of the primary transcript.** *J Biol Chem* 2000, **275(45)**:35063-35069.
58. Tsuchimoto D, Sakai Y, Sakumi K, Nishioka K, Sasaki M, Fujiwara T, Nakabeppu Y: **Human APE2 protein is mostly localized in the nuclei and to some extent in the mitochondria, while nuclear APE2 is partly associated with proliferating cell nuclear antigen.** *Nucleic Acids Res* 2001, **29(11)**:2349-2360.
59. Mueller JC, Andreoli C, Prokisch H, Meitinger T: **Mechanisms for multiple intracellular localization of human mitochondrial proteins.** *Mitochondrion* 2004, **3(6)**:315-325.
60. Karniely S, Pines O: **Single translation--dual destination: mechanisms of dual protein targeting in eukaryotes.** *EMBO Rep* 2005, **6(5)**:420-425.
61. Danpure CJ: **How can the products of a single gene be localized to more than one intracellular compartment?** *Trends Cell Biol* 1995, **5(6)**:230-238.
62. Regev-Rudzki N, Pines O: **Eclipsed distribution: a phenomenon of dual targeting of protein and its significance.** *Bioessays* 2007, **29(8)**:772-782.
63. Seibel NM, Eljouni J, Nalaskowski MM, Hampe W: **Nuclear localization of enhanced green fluorescent protein homomultimers.** *Anal Biochem* 2007, **368(1)**:95-99.
64. Johnson ES: **Protein modification by SUMO.** *Annu Rev Biochem* 2004, **73**:355-382.
65. Wilson VG, Rangasamy D: **Intracellular targeting of proteins by sumoylation.** *Exp Cell Res* 2001, **271(1)**:57-65.
66. Du JX, Bialkowska AB, McConnell BB, Yang VW: **SUMOylation regulates nuclear localization of Kruppel-like factor 5.** *J Biol Chem* 2008, **283(46)**:31991-32002.

Appendixes

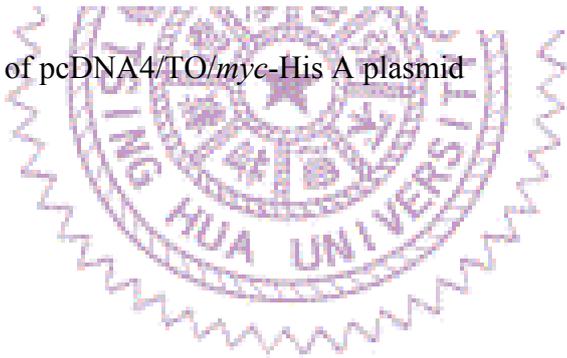


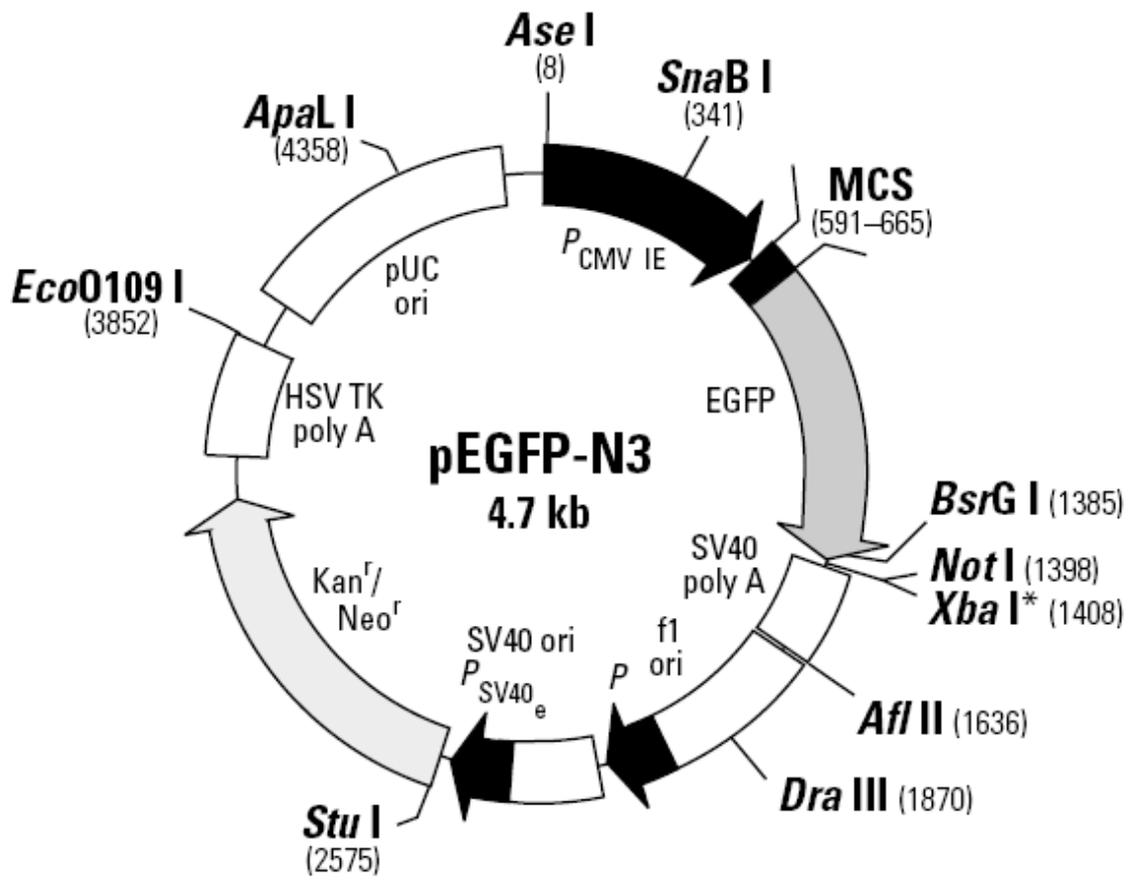
Appendix 1. The map of pSUPER.neo+EGFP plasmid



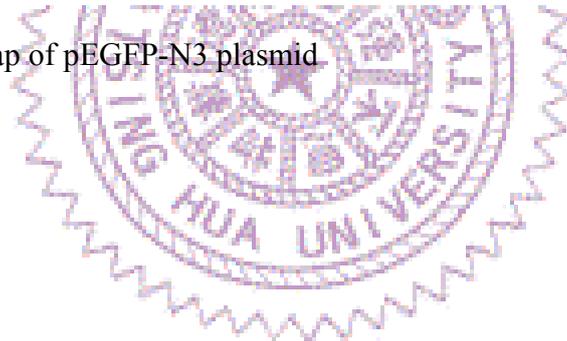


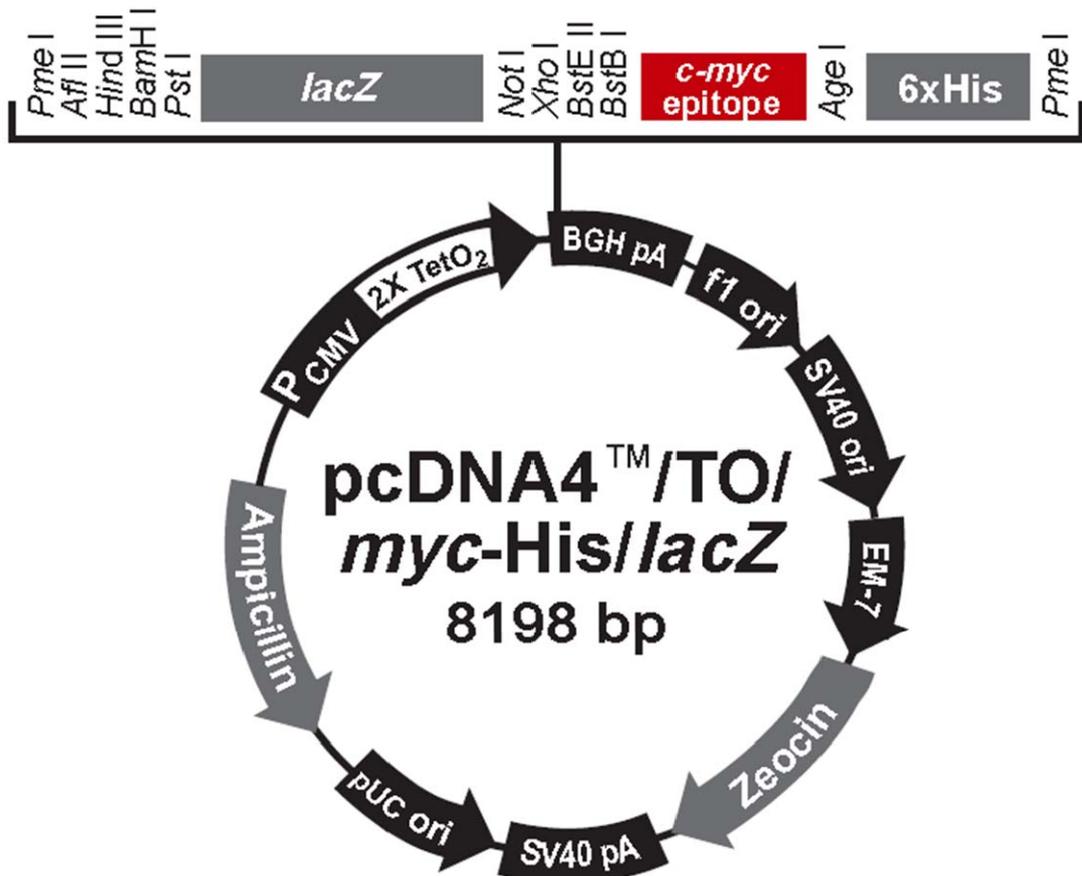
Appendix 2. The map of pcDNA4/TO/myc-His A plasmid





Appendix 3. The map of pEGFP-N3 plasmid





Appendix 4. The map of pcDNA4/TO/LacZ/myc-His A plasmid