4. Results and discussion

4.1. Construction of human chromosome 21 cell arrays

4.1.1. Working flow

Each open reading frame (ORF) of these genes was cloned into a mammalian expression vector containing an amino-terminal 6×-Histidine tag. Purified plasmids were diluted with gelatin solution to the final concentration of plasmid at 50~100ng/µl. Prior to printing, the normal microscope glass slides were treated with poly-lysine (see Appendis for preparation protocol) for better retention of DNA and cells. Depending on the number of sample needed for printing, plasmid-gelatin solutions were spotted manually or by conventional arrayer. HEK293T cells were then reverse transfected and cultured for less than 72 hours. Finally, Chr21 proteins expressed on the cell arrays were detected using anti-His antibody and additional assays were applied for organelles colocalization or apoptosis identification (see following sections). The efficiency of protein expression and labeling was monitored using a BioCCD laser scanning system with different filters. For single-cell analysis, fluorescent and confocal microscopies were used. The scheme of working procedures was presented in Fig. 4.1, and the detailed protocols can be found in Material and Methods.
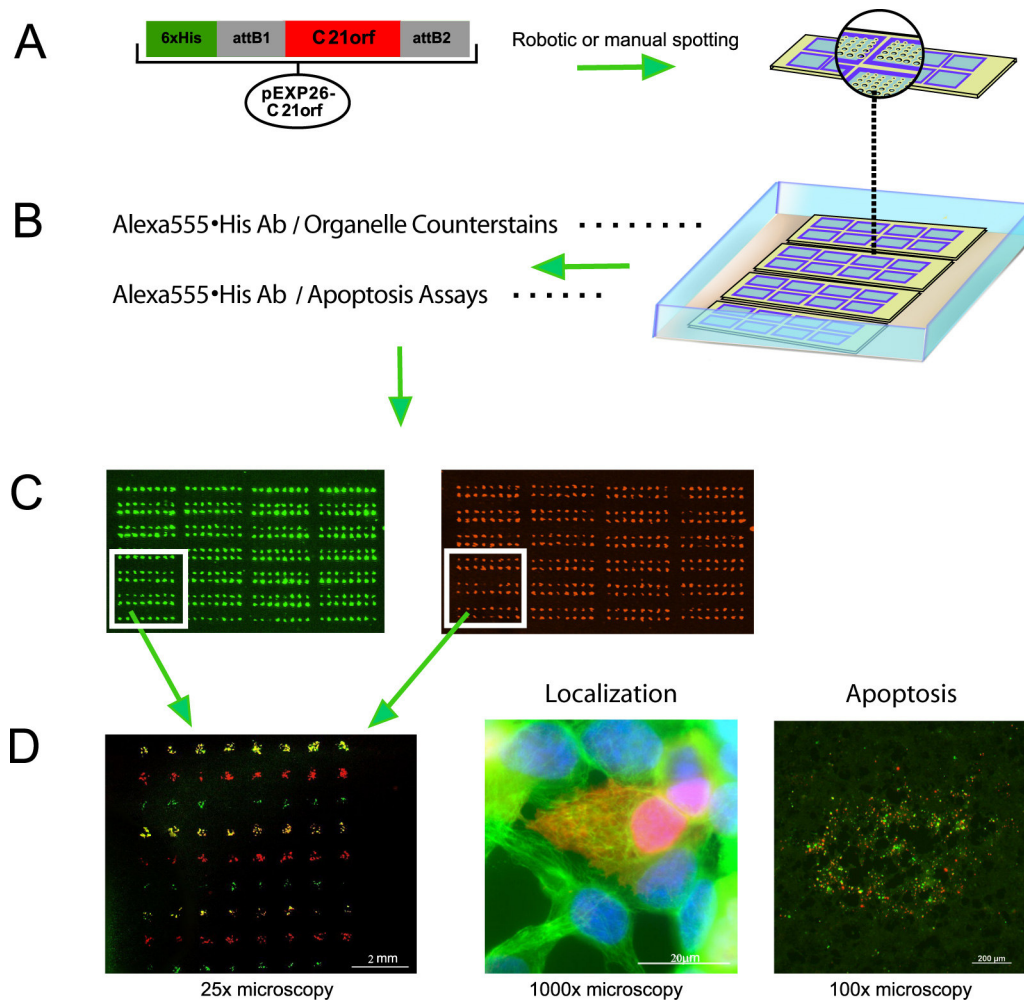
Fig. 4.1 Schematic presentation of human chrosome 21 (Chr21) cell array preparation and functional detection

(A) Plasmid construction of Chr21 ORFs and array preparation. (B) Reverse transfection into HEK293T cell line, followed by labeling of Chr21 proteins (using anti-His antibodies) and functional detections (using organelle counterstaining for protein localization or apoptosis detection). (C) Image capture—BioCCD laser scanning. The green signal indicates the expression of enhanced green fluorescent protein. The red signal indicates His_6-tagged protein. (D) Image capture—fluorescent and confocal microscopy. Left: merged microscope image of one block of spotting frame. Middle: subcellular localization analysis under 1000× magnification (using confocal system). Red signal in the image represents expressed C21orf25 protein, whereas the green and blue signals represent labeled microtubules and nucleus, respectively. From this image, this Chr21 protein localizes in the nucleus and the cytoplasm, which is not overlapped with the microtubules. Right: apoptosis detection of one cell cluster under 100× magnification. Red signal represents the cell cluster transfected by one plasmid spot overexpressing Bax, a pro-apoptotic protein, while green signal indicates the dying cells detected by TUNEL reaction. Due to overlap of red and green, most of the cells in this cluster appear as yellow, indicating the apoptosis induction effect induced by Bax overexpression.

4.1.2.   Selection of anti-His antibodies

Anti-His antibodies were applied to detect the recombinant proteins expressed on the cell arrays. Performance of the immunocytochemical detection greatly depends on the quality of antibodies. There are a number of anti-His antibodies commercially available to label the His_6-tagged proteins with the fluorescence of interest, through "one-step" (fluorochrome-conjugated anti-His antibody) or "two-step" (anti-His antibody plus fluorochrome-conjugated secondary antibody) reaction. "One-step" labeling is time-efficient, while "two-step" labeling can better amplify the signal. To assess the performance of antibodies from different manufactures, His_6-tagged eGFP (enhanced GFP) were expressed on cell arrays, followed by anti-His antibody labeling with a red fluorochrome. The labeling performance of antibodies was evaluated based on the overlap between red signal and the GFP signal. In total, 8 antibodies were screened (Table 4.1) and Penta•His Alexa Fluor antibodies gave the most sensitive and specific labeling (Fig. 4.2).

Table 4.1 Immunofluorescent (IF) labeling performance of seven anti-His antibodies

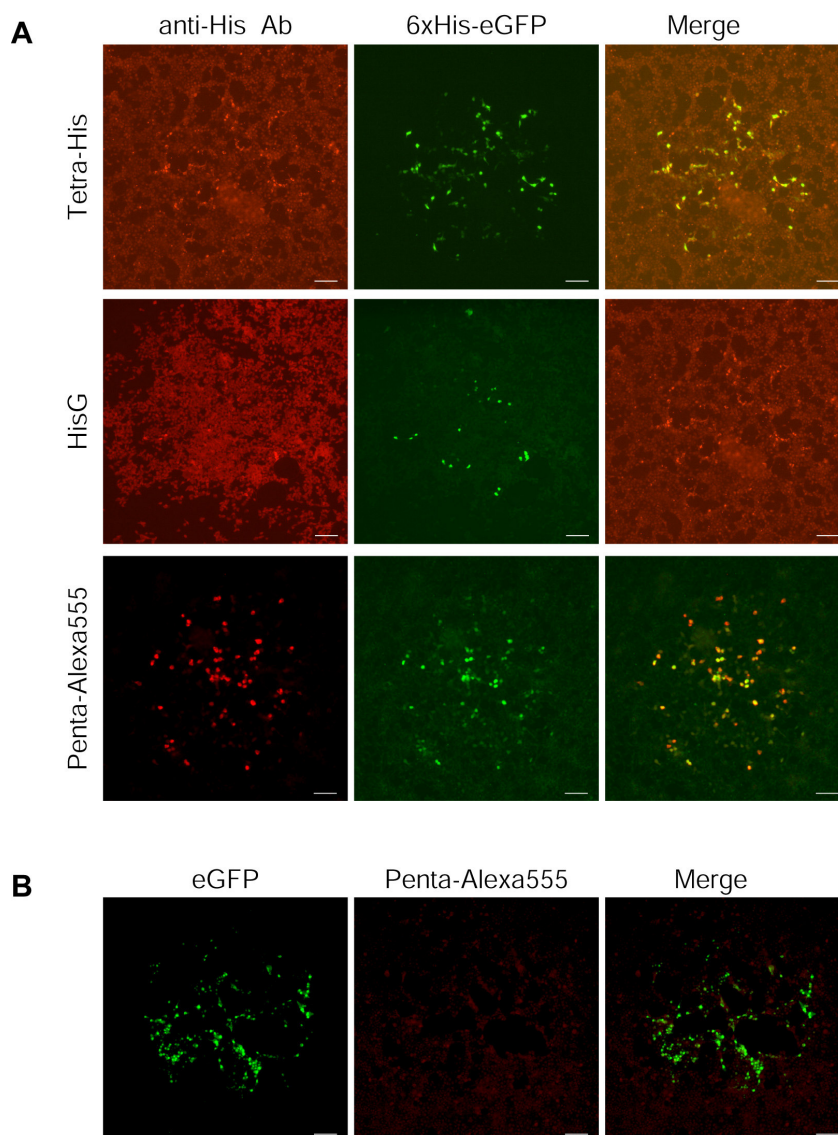| Antibody Name	Tetra•His	Penta•His	RGS•His	Penta•His Alexa555 /Alexa488	anti-HisG	anti-His tag	anti-His tag
Resource	Qiagen	Qiagen	Qiagen	Qiagen	Invitrogen	DPC Biermann	Abcam
IF Performance	unspecific	unspecific	low intensity	specific; sensitive	unspecific	unspecific	low intensity

IF: Immunofluorescence labeling

Fig. 4.2 Differential labeling performance of anti-His antibodies from different resources

(A) Immunofluorescent (IF) labeling of 6xHis-eGFP on cell array showed different detection performance by different antibodies. Left panel: IF detection of the expressed 6xHis-eGFP spots by Tetra•His (Qiagen), HisG (Invitrogen) or Penta•His Alexa555 (Qiagen) antibody, as indicated. Middle panel: GFP autofluorescence of the same eGFP spot as that shown in left panel. Right panel: merge of IF detection and antofluorescent images. The most sensitive staining was obtained using Penta•His Alexa555 (Qiagen). (B) High specificity of Penta•His Alexa555 labeling. Cell cluster expressing untagged eGFP (left) were not stained by Penta•His Alexa555 antibody (middle). Scale bars, 100μm.

4.1.3. Construction of control proteins and organelles-colocalization

In order to guarantee the coverage of all DNA/RNA spots, cells growing on cell arrays have to reach complete confluence, which ocaasionally results in condensed

and overlapping cellular structures. In this study, for instance, the ER in small cytoplasm was found very hard to be distinguished from the Golgi apparatus. Therefore, it is essential to identify the subcellular compartments for any protein localization study. To this end, two approaches (control proteins and organelles counterstaining) were applied in this study.

4.1.3.1. Control Proteins

Proteins with known localization pattern can be used as controls of proper compartmentalization. When fluorescent proteins (e.g. GFP and its derivatives) are fused with the control proteins, the cellular compartments can be "highlighted" by the fluorescence. Thus, co-expression of these control proteins with studied protein in a same cell can help determine its localization.

For large-scale localization study, it is however unlikely to use the co-expression approach of control proteins with candidate proteins due to several reasons. The actual localization of candidate proteins are hard to predict before experiment, hence, co-transfection of each candidate protein along with all control proteins is difficult to perform due to too large number of experiments. It is also uncertain if co-expression of control proteins will influence the distribution of test protein or not. This question has to be addressed especially when the control and test proteins are functionally related.

The control proteins are thus often used to evaluate the experimental platform (e.g. cloning and transfection strategy), rather than be co-expressed for organelle identification. In this study, several control proteins encoded by gene *KDELR1*, *LMNA*, *Pex11a*, *TGN38*, *CDH8*, and *GFP* were constructed following the established cloning, reverse transfection and immunostaining strategies (see methods and Fig. 4.1), and were found to localize in the predicted compartments (Fig. 4.3 (Hu et al., 2005)). Thus, the cell array platform has been proved reliable for protein subcellular localization.
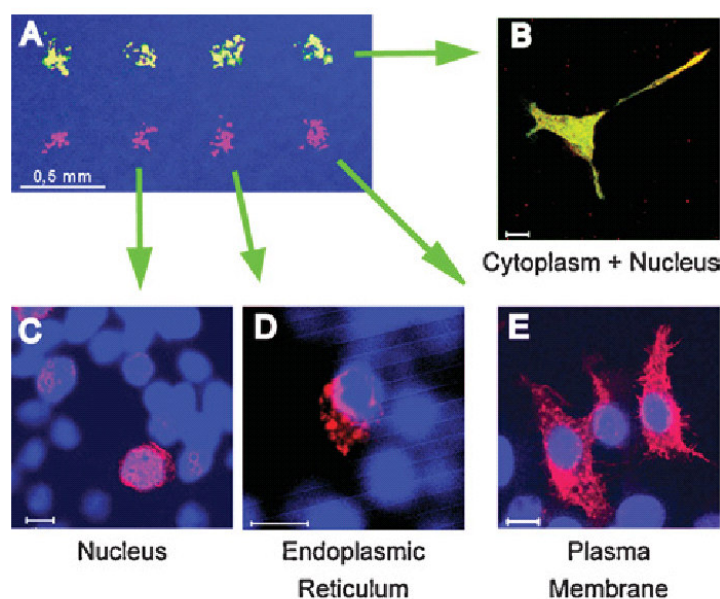
(Hu et al., 2005)

Fig. 4.3 Consistent localization of control proteins

A. Microscopy image of the cell clusters transfected by a 4×2 grid of plasmid spots. Scale bar, 0.5 mm. Image **B**, **C**, **D** and **E** show localization of control proteins including GFP proteins, Cadherin-8, the KDEL (Lys-Asp-Glu-Leu) receptor 1, and lamin A, respectively. Scale bars, 10 µm.

4.1.3.2. Organelles-colocalization

Instead of ectopic expression of control proteins, the *in situ* components of organelle structures can be used as localization markers. Antibodies and/or dyes that can specifically interact with an organelle-specific component are generally introduced into each test to label the organelles. This has been shown to be effective in accurate identification of the subcellular localization of single proteins. However, it has not been applied so far for large-scale characterizations.

For the first time, a high-throughput colocalization approach was developed in this study for a large set of organelles on cell arrays. Organelle-specific antibodies or dyes from different suppliers were screened and their performing conditions were optimized. In total, 9 reagents were found to provide sensitive and specific organelle labeling (Table 4.2). They are then subsequently counterstained with Penta•His AlexaFluor antibodies for colocalization study of Chr21 proteins.

Table 4.2 Organelle counterstains used for colocalization study

Organelle	Reagent	Supplier
Nucleus	DAPI	Sigma
F-Actin	Phalloidin (Rhodamine-conjugated)	Molecular Probes
Mitochondria	α-prohibitin antibody (rabbit)	Abcam
Peroxisome	α-catalase antibody (rabbit)	Abcam
ER	PDI (mouse)	Stressgen
Golgi	Adaptin- γ (mouse)	BD Biosciences
Lysosome	LAMP2 (mouse)	H4B4, Developmental Studies Hybridoma Bank
Microtubule	α-tubulin (mouse)	Sigma
Intermediate filament	Vimentin (mouse)	Affinity BioReagents

4.2. Subcellular localization reveals novel functional features of Chr21 proteins

4.2.1. Subcellular localization of Chr21 proteins

The working flow described in section 4.1 was then used for localization screening of 89 Chr21ORF constructs. These genes are distributed over the majority of 21q, ranging from RBM11 (14.5 Mb) to MCM3AP (46.6 Mb), with part of them aberrantly expressed in the Down's syndrome mouse model. Most of the genes have an OMIM entry, and have been known or suspected to lead to disease susceptibility or actual phenotypes in Down syndrome or congenital diseases (see Table 4.3 for details).

In total, 52 out of 89 proteins were detected on Hek293T cell arrays using optimized Penta•His AlexaFluor antibodies and their subcellular localization patterns were determined after counterstaining with each compartment marker. The rest of the tested proteins could not be detected as their expression was below the immunnofluorescence detection threshold of the antibodies on cell arrays. For 34 of the detected proteins, this study describes, for the first time, their subcellular distribution in mammalian cells. A summary of the localization data is presented in Table 4.3, together with the protein accession numbers and information on gene functions and expression patterns. Localization images not shown in this dissertation

can be found at the web site http://www.molgen.mpg.de/~ag_onf/chr21ORF-SubLoc/index.html.
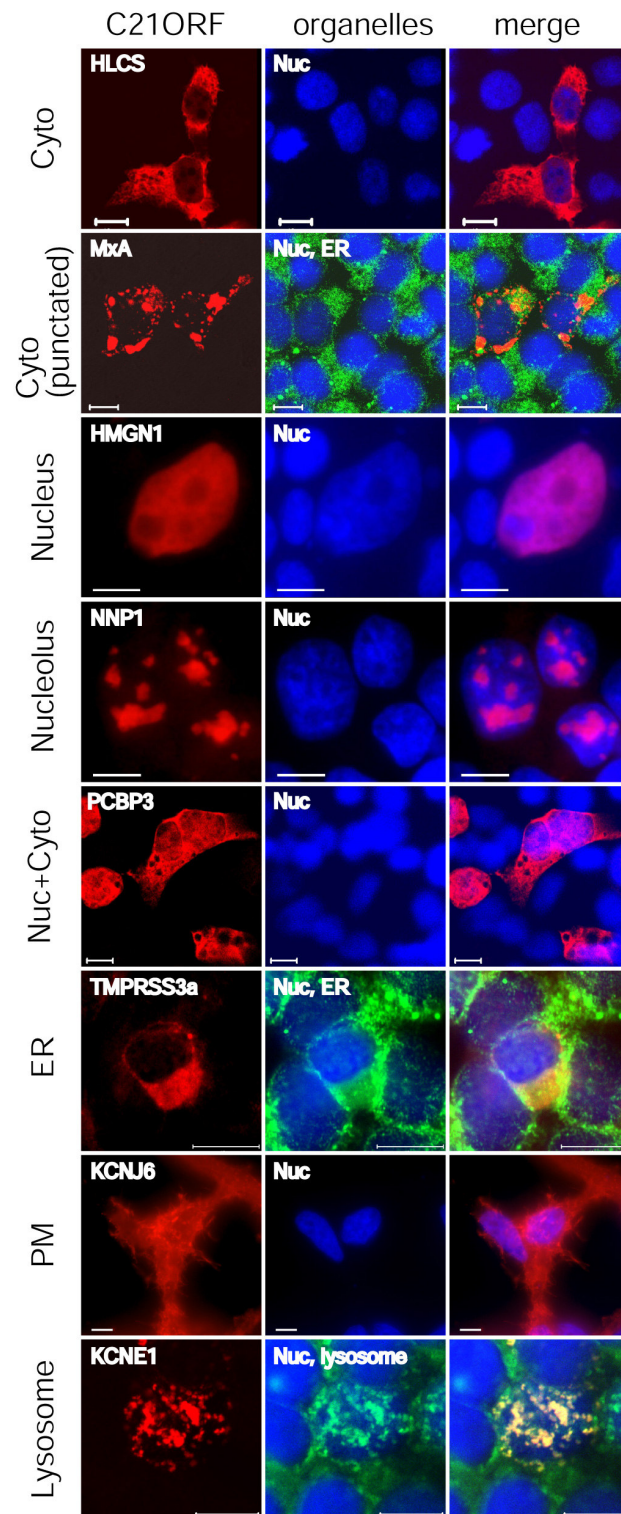

Fig. 4.4 Localization categories of detected Chr21 proteins

Chromosome 21 (Chr21) proteins were localized in a variety of cellular compartments. The images in the left panel represent anti-His staining of Chr21 proteins, with the gene names indicated in the upper left corner of each image. The images in the middle panel show applied counterstaining, with organelles indicated in the upper left corner. The images in the right panel present a merged picture from detection of Ch21 proteins and counterstaining. Scale bar, 10 μm.

Table 4.3 Subcellular localization of 52 human chromosome 21 proteins

Gene Symbol	Function class	Localization in HEK293T[a]	Localization in Swiss-Prot	Transcript alteration in Ts65Dn mice[b]	OMIM ID[c]	GenBank Protein Acc.No.
21 genes located in Down Syndrome Critical Region						
B3GALT5	galactosyl-transferase	Golgi/ER	Golgi		604066	NP_006048.1
C21orf19	unknown	Nuc/Cyto	—			AAL34462.1
C21orf4	unknown	PM	—			AAC05974.2
C21orf96	unknown	Cyto (punct)	—			NP_079419.1
CHAF1B	chromatin assembly factor	Nucleoplasm Cyto-M phase	Nuc; Cyto-M phase		601245	NP_005432.1
CLDN14	tight junction	ER/PM(less)	—		605608 ALLELIC VARIANTS	AAG60052.1
CRYZL1	oxidoreductase	Cyto	—		603920	BAA91605.1
DSCR3	unknown	Nuc	—	20% cohort to DSCAM	605298	NP_006043.1
ETS2	transcription factor	Nuc	—		164740	NP_005230.1
HLCS	protein ligase	Cyto	Cyto/Mito		253270 CLINICAL SYNOPSIS	NP_000402.2
HMGN1	DNA binding	Nuc	Nuc		163920	AAA52676.1
IFNGR2	receptor	PM/ER	—		209950 CLINICAL SYNOPSIS	AAH03624.1
KCNE1	K-channel	Lyso/PM	—		176261 ALLELIC VARIANTS / CLINICAL SYNOPSIS	AAH36452.1
KCNE2	K-channel	Lyso/PM	—	3.39 in Ts65Dn midbrain	603796 ALLELIC VARIANTS	NP_005127.1
KCNJ15	K-channel	PM/Golgi	—		602106	NP_002234.2
KCNJ6	K-channel	PM/Golgi	—	20% cohort to DSCAM	600877	NP_002231.1
MxA	dynamin & large GTPases	Cyto(punct)	Cyto	3.70 in Ts65Dn skeletal muscle; 2.70 in Ts65Dn heart	147150	NP_002453.1
PCP4	unknown	Nuc/Cyto	—	20% cohort to DSCAM	601629	CAA63724.1
RPS5L	unknown	Cyto	—			pseudogene, 81% identity to BAB79493.1
SH3BGR	SH3 adaptor	Cyto	—	0.55 in Ts65Dn midbrain; 2.61 in Ts65Dn kidney	602230	AAH06371.1
WDR9_3'	unknown	Nuc	—			BAA92123.1

31 genes not located in Down Syndrome Critical Region						
ABCG1	ATPase	PM/Golgi	ER/Golgi	20% cohort to DSCAM	603076	CAA62631.1
AGPAT3	acyltransferase	ER/PM(less)	—			AAH11971.1
BACH1	transcription regulation	Cyto(punct) Nuc-M-phase	Nuc (predict)		602751	BAA24932.1
C21orf103	unclear	Cytoplasm	—			NP_853633.1
C21orf25	unknown	Nuc/Cyto	—			XP_032945.2
C21orf30	unknown	Nuc	—			CAB56001.2
C21orf59	unknown	Nuc/Cyto	—			AAG00496.1
C21orf69	unknown	ER	—			AAK60445.1
C21orf7 (TAK1L)	transcription factor like	Nuc/Cyto	—	0.97 in Ts65Dn kidney; 1.15 in Ts65Dn testis		AAF81754.1
CBS	cystathionine-beta-synthase	Cyto	Cyto		236200 ALLELIC VARIANTS / CLINICAL SYNOPSIS	NP_000062.1 (splicing isoform)
CCT8	chaperonin	Cyto	Cyto			BAA02792.1
CLDN17	tight junction	PM/Golgi	—			CAB60616.1
CLDN8	tight junction	ER/PM(less)	—			NP_036264.1
CXADR	receptor	PM	—		602621	AAH10536.1
DNMT3L	methyl-transferase like	Nuc/Cyto	Nuc (predict)		606588	AAH02560.1
GCFC	transcriptional repressor	cyto	Nuc (predict)	20% cohort to DSCAM		AAD34617.1
HSF2BP	transcription factor binding	Cyto	—		604554	NP_008962.1
KIAA0179	unknown	Nuc/Cyto(punct)	—			XP_035973.4
MCM3AP	DNA binding	Cyto/Nuc	Nuc/Cyto (predict)		603294	BAA25170.1
NNP1	RNA processing	Nucleolus	Nuc; Nucleolar			AAH00380.1
PCBP3	RNA binding	Cyto/Nuc	Cyto		608502	AAH12061.1
PDE9A2	phosphodiesterase	Cyto (accum)	—		602973	AAH09047.1
PDXK	kinase	Cyto	Cyto		179020	AAH00123.1

PFKL	kinase	Cyto (accum)	—	20% cohort to DSCAM	171860	AAH09919.1
PKNOX1	transcription factor	Nuc/Cyto	Nuc (predict)		602100	AAH07746.1
PPIA3L	peptidylprolyl isomerase A-like	Nuc/Cyto	Cyto		123840	CAA37039.1
TMPRSS3a	protease	ER	ER		605511 ALLELIC VARIANTS	NP_076927.1
TSGA2	chromosome-associated	Cyto /Nuc	Cyto/Nuc(metaphase)			NP_543136.1
UBASH3A	catalytic activity	Cyto	Nuc (predict)		605736	NP_061834.1
UBE2G2	ligase	Cyto	—		603124	AAC32312.1
WDR4	unknown	Nucleoplasm	—		605924	AAH06341.1

Subcellular localization of 52 Chr21 proteins was determined, among which 34 proteins have been newly reported. Up to 21 genes were located in the Down syndrome critical region as defined by Delabar et al. 1993. In total, the localization patterns were highly consistent with described gene functions. [a] Accum, accumulated; Cyto, cytosol; ER, endoplasmic reticulum; Lyso, lysosome and endosome; Mem: membrane; Mito: mitochondrion; Nuc: nucleus; PM: plasma membrane; Punct: punctate; Predict: 3 non-experimental localization qualifiers including "potential", "probable" and "by similarity"; [b] 4 genes escaped the 1.5 overexpression rule in the tissues of Ts65Dn mouse model for trisomy 21; 6 genes were in the 20% cohort for similar expression pattern as DSCAM, a key regulator for neurogenesis. This co-expression with DSCAM may indicate a role of these proteins in neurogenesis (Kahlem et al. 2004). [c] 33 genes had an OMIM entry, among which 7 had known allelic variants associated with diseases, and 4 had a clinical synopsis.

A significant proportion of the proteins were found either in the cytosol (31%) or in the nucleus (17%). Twenty-three percent of the proteins localized in both the cytoplasm and the nucleus, or were found to transport between these two compartments. Localization of 29% of the proteins was associated with components of the ER secretory pathway, including the ER (10%), plasma membrane (PM) (13%), Golgi (2%), and endosome/lysosome compartment (4%). No protein was found in the peroxisome, mitochondria or cytoskeletal structures, probably due to the limited number of proteins detected. Figure 4.4 shows localization examples of Chr21 proteins at different cellular compartments. In the following paragraphs, several features of protein distribution were described and discussed with protein functions that were previously reported or inferred.

Cytosol

Two distinct localization features were observed in this study for cytosolic proteins. Most cytosolic proteins were evenly distributed throughout the cytoplasm (Fig. 4.4), while a few (4 out of 16 proteins) formed a punctate pattern of localization. The punctate pattern could render a morphology similarity to some certain cell organelles such as lysosomes and peroxisome. However, the application of organelles counterstaining in this study allowed for unambiguous distinction of this accumulated cytosolic feature from surrounding organelles. For example, myxovirus resistance protein 1, which is encoded by the *MxA* gene — revealed a punctate pattern in the cytosol, which does not overlap with any organelle staining. The accumulated localization pattern of MxA supports previous reports that self-assembled MxA proteins wrap around or interact with the incoming viral nucleocapsids and thereby prevent their replication (Andersson et al., 2004). Accola *et al.* found that the MxA protein localizes to the smooth ER and suggested that it might inhibit viral replication by altering membrane organizations of host cells (Accola et al., 2002). In this study, however, MxA was not observed to co-localize with the ER marker protein disulphide isomerase (PDI) (Fig. 4.4).

Nucleus

The non-histone chromosomal protein HMG-14, which is encoded by the *HMGN1* gene, was found in the nucleus but outside of the nucleolus (Fig. 4.4). This localization correlated with its suggested function as a modulator of the interaction between DNA and the histone octamer through its binding to nucleosomal DNA. In contrast, the novel nuclear protein 1 (NNP-1) localized exclusively in the nucleolus

and the nucleolus location is in accordance with its role in the synthesis of 28S rRNA (Savino et al., 1999) (Fig. 4.4).

Nucleus and cytoplasm

Twenty-three percent of the 52 detected proteins were found to localize in both the nucleus and the cytoplasm. The nucleus/cytoplasm distribution ratio for a particular protein, however, varied from cell to cell. In some cells, the proteins with dual localization could be found either only in the nucleus or only in the cytoplasm, suggesting a continuous translocation activity. In contrast to a previous claim (Simpson et al., 2000) that dual protein localization delivers little information, we think that a dual localization pattern probably reflects the complexity in the protein function. For example, the *PCBP3* gene encodes poly(rC)-binding protein 3, which belongs to the K homology (KH) domain protein subfamily. The KH domain binds to RNA, and therefore has an important role in post-transcriptional activities. It has been found that PCBP3 proteins, when fused to a *myc* epitope and a pyruvate kinase tag, were only observed in the cytoplasm, whereas other members in this subfamily mainly localize in the nucleus (Chkheidze and Liebhaber, 2003). Interestingly, by using small His_6 tag, the distribution of PCBP3 protein were found remarkable versatile, ranging from the cytoplasm or nucleus only, to both nucleus and cytoplasm at variable distribution ratios (Fig. 4.4). The nucleus and/or cytoplasm localization pattern of PCBP3 is probably associated with its diverse functions in mRNA stabilization and translational regulation via its RNA-binding properties (Chkheidze and Liebhaber, 2003).

In another example, the variable distribution of the MCM3AP protein (minichromosome maintenance protein 3 associated protein) in the nucleus and cytoplasm (Fig. 4.5 C) is in accordance with its role in the translocation of MCM3 proteins from the cytosol into the nucleus. The MCM3 protein was known as an essential factor that allows DNA to undergo a single round of replication per cell cycle. The observed variable distribution of MCM3AP between the nucleus and the cytosol probably reflects the cell cycle-dependent nuclear import of MCM3 protein.

ER-associated secretory pathway

A series of organelles including ER, Golgi, lysosome/endosome, are involved in ER-associated secretory/localization pathway, which targets a large class of proteins to its proper cellular compartments. Plasma membrane proteins, lysosome proteins and secretory proteins are first targeted to the ER and then to the Golgi apparatus for

modifications before being transported elsewhere. In contrast, proteins that are resident in the ER and the Golgi are retrieved and retained in these two organelles. Counterstaining with specific antibodies for ER, Golgi and lysosome/endosome allowed for precise localization of the proteins in these compartments. For example, the TMPRSS3 protein (transmembrane protease serine 3), whose defect results in autosomal-recessive neurosensory deafness 8 and 10, was found to locate exclusively in the ER, as previously described (Guipponi et al., 2002) (Fig. 4.4).

For some proteins, the entire protein-processing pathway could even be traced throughout different organelles. Good examples are two integral membrane proteins that are encoded by the *KCNE1* and *KCNE2* genes, both of which belong to the potassium channel, voltage-gated, ISK-related subfamily (subfamily E). Figure 4.5 (A) shows consecutive steps of KCNE2 protein maturation, including its glycosylation site in the Golgi complex and its final destination at the PM. We report here, for the first time, that KCNE1 and KCNE2 proteins also localize to the lysosomes (Fig. 4.4 and Fig. 4.5 A). The co-localization of the two KCNE proteins with the lysosome marker LAMP2 (lysosomal-associated membrane glycoprotein 2) could reflect the ability of the PM proteins to shuttle between the PM and the endosome/lysosome compartment (Pollard and Earnshaw, 2004). Alternatively, the lysosomal location of KCNE1 and KCNE2 proteins could result from their direct transport from the Golgi to the lysosome. So far, it is not clear to us which of the two mechanisms accounts for this lysosomal localization pattern. Interestingly, the KCNJ6 potassium channel protein, which is a member of subfamily J, never showed the lysosomal resident pattern and only localized at the PM in the same cell array experiment (Fig. 4.4). It has been known that subfamily E and subfamily J, associated with their respective functional partners, regulate potassium current under different biological circumstances. Whether the differential localization between these two subfamilies could contribute to the functional difference still need further investigation.
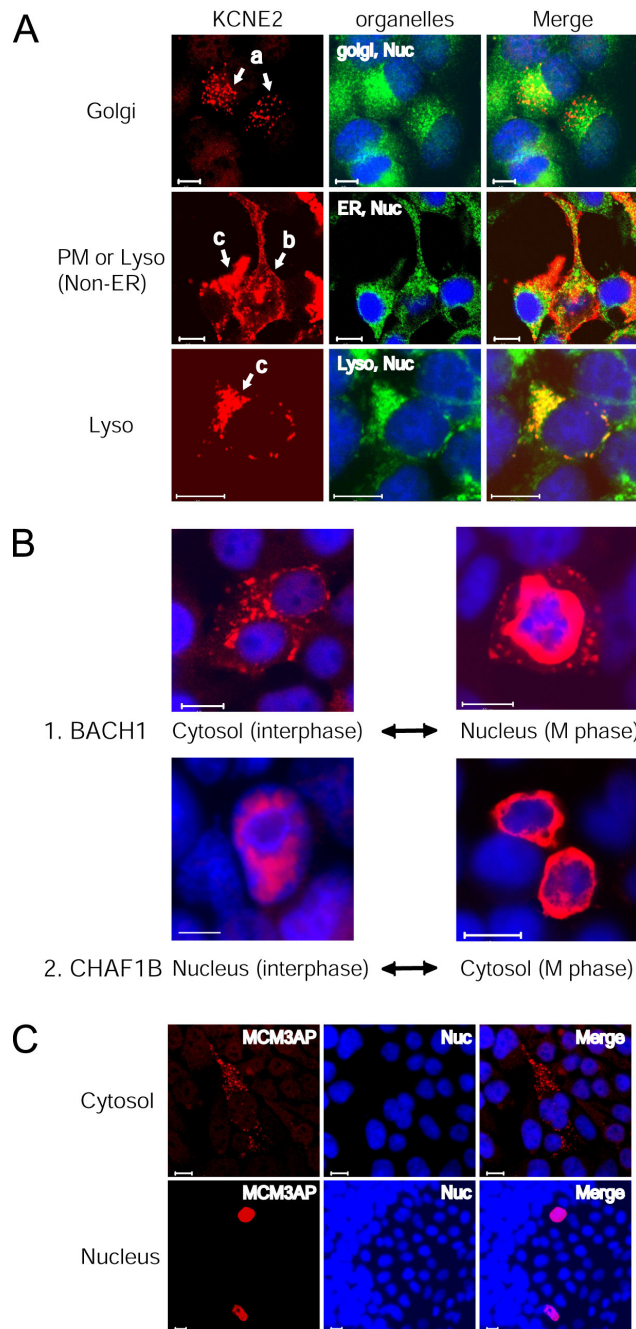
Fig. 4.5 Protein intracellular translocation observed on cell arrays

(A) Stages of post-translational trafficking of KCNE2, a plasma membrane (PM) protein, including the early modification within the Golgi complex (a) and final localization at the PM (b) or in the lysosomes (c), without retention in the ER (middle row of the images). (B) Translocation of BACH1 and CHAF1B proteins during cell interphase and mitosis. 1. BACH1 proteins distributed in the entire cytosol at interphase. In early mitosis, most proteins accumulated in the nucleus and wrapped condensed chromosome. 2. During the interphase, CHAF1B protein localized in the nucleus, whereas after cell division it was found in the cytoplasm of two daughter cells. (C) Differential localization pattern of MCM3AP protein in the cytoplasm and nucleus. Scale bar, 10 μm.

4.2.2. Cell-cycle-dependent protein translocation

A cell-cycle-dependent localization pattern was detected in the case of BACH1 and CHAF1B proteins. BACH1 proteins contained a BTB (broad complex-tramtrack-bric-a-brac) or POZ (poxvirus and zinc finger) protein interaction domain and a CNC (Cap'n'collar)-type bZIP domain. It has been demonstrated that BACH1 can function as transcriptional activators or repressors when expressed in mammalian cells (Oyake et al., 1996). In this study, a complex cell-division-dependent localization pattern was observed for this protein. At interphase, BACH1 was distributed in the entire cytosol. During the early phase of mitosis, however, the protein accumulated in the vicinity of the condensing chromatin (Fig. 4.5 B). This suggested that the nuclear transportation of BACH1 protein is highly associated with cell division. Hence, BACH1, acting as a transcription activator, probably regulate the genes that are involved in cell mitosis.

For CHAF1B (chromatin assembly factor I p60 subunit), we observed the translocation from the nucleus into the cytoplasm during cell division (Fig. 4.5 B). Consistent with the previous report (Marheineke and Krude, 1998), we observed that at interphase most of the CHAF1B protein localizing in the nucleus, where it is involved in chromatin assembly and DNA replication. After cytokinesis had occurred, we detected the presence of the CHAF1B protein in the cytoplasm of two daughter cells, probably in its inactive form after dissociation from chromatin (Marheineke and Krude, 1998).

4.2.3. Localization of claudins and the accompanying cell morphological changes

Claudins comprise one of the two major integral-membrane protein families that are found in the tight junctions of epithelial- and endothelial-cell sheets. There are more than 20 members of the human claudin gene family, which are characterized by heterogeneous tissue localization patterns (Rahner et al., 2001; Ruffer and Gerke, 2004). Daugherty *et al*. found that translocation of claudin-5 from the PM to an intracellular compartment helps regulate tight junction permeability during the differentiation of human fetal lung alveolar epithelial cells to a type II cell phenotype (Daugherty et al., 2004). Loss of human claudin-14 is known to cause autosomal-recessive deafness (Ben-Yosef et al., 2003). Furthermore, loss of claudin-3 may

initiate cilia degeneration of the retinal pigment epithelium (Nishiyama et al., 2002) and correlates with experimental autoimmune encephalomyelitis (Wolburg et al., 2003). Despite the clinical relevance of the claudin proteins, their structure within tight junctions and the regulatory mechanisms of protein expression is not entirely understood.

Here, the subcellular localization data was reported for the first time for claudin-8 and -17. Claudin-17 localized at the PM, whereas claudin-8 and -14 predominantly distributed in the ER (Fig. 4.6). Moreover, morphological changes were observed for HEK293 cells overexpressing Claudin-8 and claudin-14, but not claudin-17.
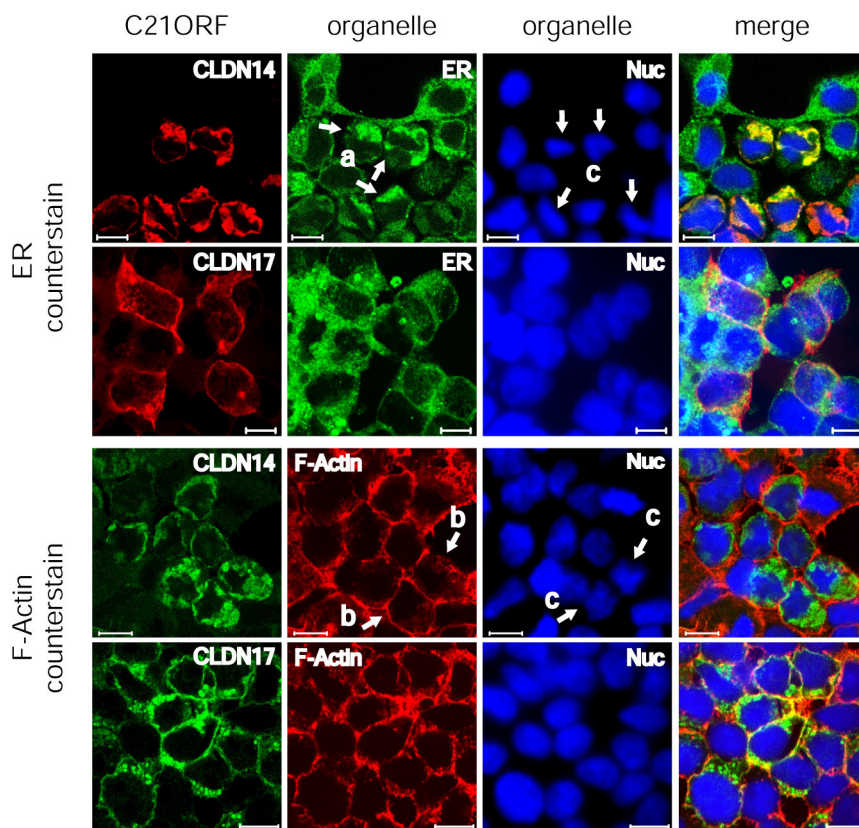

Fig. 4.6 Influence of Chr21 protein overexpression on cellular physiology

Overexpressed claudin-14 was retained in the retracted endoplasmic reticulum (a), which resulted in loss of cell membrane protrusions (b). Claudin-17 localized in the plasma membrane and its expression did not influence cell morphology. A distinct change in nuclear shape (c) was observed in cells expressing claudin-14 but not in those expressing claudin-17. Scale bar, 10 μm.

Hek293T cells, when used on cell arrays, retained their epithelial cellular morphology. Normally, the cells attached to the surface of glass slides, stretched out and connected with each other, to perform a variety of cellular functions including

intercellular communication. However, clusters of cells transfected with gene *CLDN8* or *CLDN14,* the genes coding for claudin-8 and claudin-14 respectively, were found to undergo cell-morphological changes. Expression of claudin-8 and -14, but not claudin-17, resulted in loss of cell-membrane protrusions, and the cells retracted. In addition, a distinct change in nuclear shape could be observed in the cells overexpressing claudin-8 and –14. In some cases, the nuclear DNA was found to degrade, as indicated by the lower intensity to DNA DAPI staining. Again, all these nuclear features were not shown in the cells overexpressing claudin-17 (Fig. 4.6).

4.2.4. Proteins relevant to the pathology of Down's syndrome

Twenty-one proteins described in this study located to the Down's syndrome critical regions (DSCR) of human Chr21 (see Table 4.3), based on the genetic mapping done by Delabar et al., which until recently has been considered to be critical for the etiology of Down syndrome (Delabar et al., 1993). The localization of 16 of these proteins has been reported here for the first time. Figure 4.7 shows the localization patterns of four selected DSCR proteins that are encoded by the *C21orf4, DSCR3, PCP4* and *SH3BGR* genes. The C21orf4 protein was predicted to be an integral membrane protein. Indeed, we determined its localization in the plasma membrane. The *DSCR3* gene has significant homology to the *Hbeta58* mouse gene, which is active in embryogenesis (Nakamura et al., 1997). In this study, the DSCR3 protein was found in the nucleus, suggesting that it may function in the regulation of transcription or in the maintenance of nuclear structure. The *PCP4* gene encodes brain-specific polypeptide PEP-19, which has a key role in brain development (Slemmon et al., 2000). However, the precise function of PEP-19 in this process has not been determined. For the first time, a dual localization of PEP-19 in the cytosol and in the nucleus was reported here (Fig. 4.7).

In Chr21 trisomic cells, genes are usually overexpressed by 1.5-fold in comparison to diploid cells. There is, however, a number of exceptions, as described recently (Kahlem et al., 2004). The *SH3BGR* and *KCNE2* genes showed a 2.61- and 3.39-fold increase of expression in the kidney and midbrain, respectively, in the Down's syndrome mouse model. In this study, the SH3BGR protein revealed a uniform localization pattern in the cytosol (Fig. 4.7). As described in previous paragraphs of this dissertation (see section 4.2.1), the KCNE2 protein showed a

variable distribution pattern, probably due to the differential level of expression in particular cells (Fig. 4.5A). The abundance of the KCNE2 protein in the PM was lower than that in the lysosome/endosome (Fig. 4.5A). This abundance-dependent protein localization pattern might be particularly relevant to the case of Chr21 trisomic cells, where *KCNE2* gene are overexpressed. This hypothesis, however, needs further functional evaluation.
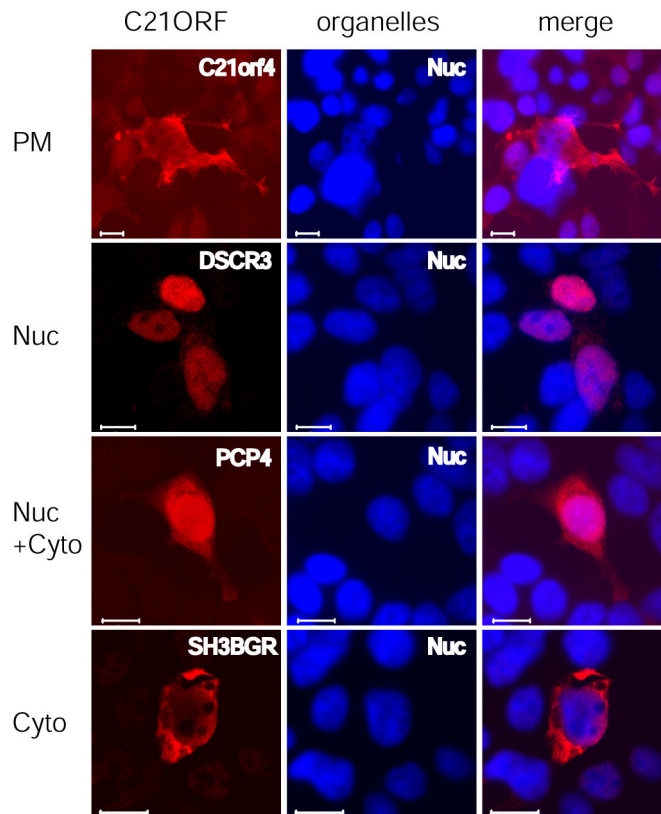

Fig. 4.7 Subcellular localization of the proteins relevant to Down's syndrome pathogenesis

The localization pattern of 21 chromosome 21 proteins related to the Down's syndrome critical region was described in this study. The figure shows the localization pattern for four of these proteins. Gene names are shown in the left panel, together with an indication of protein localization. The middle panel shows nucleus DAPI staining. Scale bar, 10 μm.

4.2.5. Advantages and limitations of His$_6$ tag

In order to minimize the tag-caused localization artifacts we took advantage of His$_6$ tag, which is a much smaller epitope than other commonly used fusion tags such as GFP. In this study, 22 out of 24 evaluated proteins (including 22 Chr21 proteins and 6 control proteins) localized to the cellular compartments in accordance to previously

described sites. This suggested that application of His_6 tag does not influence protein's native compartmentalization. Furthermore, the well-established metal-chelate chromatography technology, e.g. Ni-NTA, makes His_6 applicable for the high-throughput purification of recombinant proteins for subsequential applications, such as antibody screening or protein structure analysis.

Detection of His_6-tagged proteins, however, requires the process of immunofluorescence labeling. An efficient labeling greatly relies on the quality of His-antibodies and sometimes also on the staining protocol used. Moreover, during the entire labeling process, extra attention should be paid to avoid the loss of cells. This is of particular importance to maintain the cell monolayer intact when carry out an immunofluorescent labeling on cell arrays. On the contrary, GFP and its spectral derivants are popularly used as fusion tags to monitor gene expression and protein localization in living cells thanks to their autofluorescence (Chalfie et al., 1994). However, usage of the large (ca. 28 kDa for GFP alone) fluorescent proteins as fusion tags should be taken with caution since their presence may imfluence the native localization of test proteins. For example, GFP fusion was found to result in protein mislocalization through altering the protein folding or masking the targeting signal of the protein (Brosius et al., 2002; Pouli et al., 1998).

4.2.6. Confirmation of protein localization using C-terminal tag

C-terminus-tagged Chr21 proteins were next constructed to answer the question whether the involvement of fusion tag at the N- and C-terminus could lead to different localizations. The ORFs of seventeen Chr21 proteins were re-cloned into the expression vector containing a C-terminal Myc epitope. The localizations of these proteins represent all localization categories that were previously classified using His_6 tag. Following the same workflow, the Myc-tagged proteins were expressed on Hek293T cell arrays and were detected with anti-Myc antibody. The localization of C-terminus-tagged proteins did not substantially differ from the localization of N-terminus-tagged proteins. The only exception came from the plasma membrane protein KCNJ6. This potassium channel protein, when tagged by C-terminal Myc tag, was only found in the intracellular vesicles rather than being transported to the plasma membranes with N-terminal His_6 tag (Fig. 4.8). From the data, the small epitope such

as His_6 (6-aa) and Myc (11-aa) appeared to a safe tag with little effect on proteins localization.
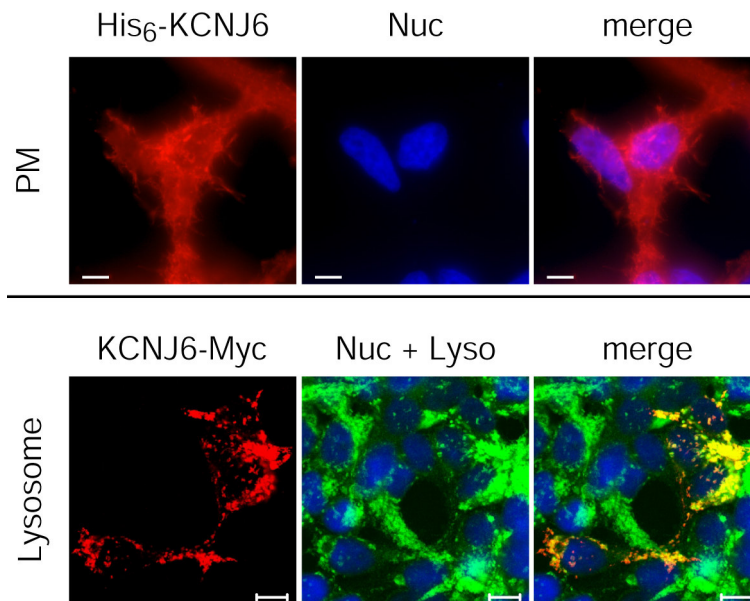

Fig. 4.8 Altered localization of KCNJ6 protein due to different fusion tag and origination

KCNJ6 protein, the potassium channel protein, when fused with His_6 tag to the N-terminus, localized predominantly to the plasma membrane (upper panel). When used with Myc tag to the C-terminus, it distributed only in the intracellular vesicles, which highly overlapping with the cell lysosomes/endosomes (lower panel). Scale bars, 10 µm.

The fusion tag should be put at which end of the recombinant protein has always been argued in protein localization study. A common preference was given to C-terminal tag or, if possible, the use of both N- and C-terminal tag, respectively. This is because the N-terminal fusion of GFP was found, in some cases, to mask the targeting signals of mitochondrial and plasma membrane proteins (Wiemann et al., 2004). However, misdirection of the proteins localizing to the ER and peroxisomes were also described due to the presence of C-terminal GFP (Huh et al., 2003; Wiemann et al., 2003). In both cases, the localization artifacts should be also considered as a result of GFP tag in addition to the terminal influence. GFP alone has preferential localization in the nucleus and cytoplasm and its big size may interfere with the protein targeting processing, which might happens at both the N-terminus (e.g. signal peptide cleavage) and the C-terminus (e.g. palmitoylation or farnesylation) (Davis, 2004; Ogawa et al., 1995). In this study, we used small epitopes to retain the physiological function of tagged proteins, and the localization results from N-terminal His_6 tag were almost

completely confirmed by C-terminal Myc tag. However, for those clones that were failed to be detected using N-terminal tag, change of fusion terminus may be helpful to reveal their localization behaviors.

4.2.7. Influences of cells on protein localization

To answer the question if cellular environment influence the localization of overexpressed proteins, the workflow established for HEK cells was applied in Hela cell arrays. Compared with Hek293T cells, transfection efficiency of Hela cell arrays decreased dramatically and resulted in low detection rate. Nevertheless, all detected proteins were found in the same localizations as in Hek cells, suggesting similar sorting mechanism between these two human epithelial cell types. However, to what extent the cellular variations between epithelial and non-epithelial cells would influence the protein localization remains unclear.

Differences in cellular environments between prokaryotes and eukaryotes showed a strong impact on protein synthesis and sorting. The same Chr21 ORFs were cloned into an *E. coli* expression vector for protein synthesis. Among the 52 proteins detected in Hek293T cell arrays, 21 proteins could not be expressed in *E. coli*. More interestingly, up to 14 proteins not expressed in *E. coli* were plasma membrane and ER proteins, representing 90% of the proteins found in these two locations in Hek cells. This data strongly argued the fact that prokaryocytes are not able to perform an efficient synthesis of membrane and secretory mammalian proteins. Hence, mammalian cells remain a better choice when perform functional studies of ER-associated secretory and membrane proteins.

4.3. Comparison of experimental localization results with *in silico* prediction

Four computational programs (PSORTII (Horton and Nakai, 1997), ProtComp v. 5 (Softberry), PENCE proteome Analyst (Lu et al., 2004) and SubLoc (Hua and Sun, 2001)), which based their predication on different strategies (see section 1.2.2.1 for details) were used to assign 52 proteins detected in Hek cell arrays into multiple

intracellular compartments. Given the fact that the proteins with different subcellular localization usually have characteristic amino acid composition (Hua and Sun, 2001). SubLoc utilizes this feature for protein localization prediction. PENCE predicts localization solely upon sequence homology and searches database for homologous proteins with known localization. Besides homologue search, ProtComp5 also looks for certain targeting motifs and functional peptides that are relevant to protein localization. PSORTII generates its prediction through the combination of the overall amino acid composition, N-terminal targeting sequence information and motifs. The prediction results for the 52 proteins were summarized in Table 4.4 and were compared with the localization pattern from Hek cell arrays.

The percentage of the positive predications, which were consistent with the experimental findings, was calculated for each of the four programs. It was found to vary among different programs. In total, PENCE and PrompComp presented a significantly better performance with 77% and 75% of the prediction in consistency, whereas only 58% and 57% of predication were positive using SubLoc and PSORTII, respectively (Fig. 4.9). The fact that both PENCE and PrompComp rely on homology comparison indicates a better performance of this approach for protein localization predication. However, if there is no homologous protein available with annotated localization, the prediction results can be quite ambiguous. This is exactly what I observed in this study. For all 89 Chr21 proteins, PENCE did not give any prediction for 14 proteins. This brings a big challenge when using homolog-based predictor such as PENCE for genome-wide protein localization prediction, even though it can provide better accuracy in case the protein homologs with known localization are present.

To evaluate whether the prediction performance is associated with specific localization site, the prediction results were grouped into different categories based on the experimental localization results. The number of predictions consistent with experiments was counted for each localization category (Fig. 4.9). As shown in Fig. 4.9, the prediction accuracy was tightly correlated with localization sites. For example, PSORT II could correctly predict most of nuclear and cytoplasmic proteins (11 out of 12), but very few of plasma membrane proteins (only 2 out of 11). The predications of ProtComp5 were correct for most of proteins in all localization sites except for nucleus (only 7 out of 15). The reasons accounting for the localization site-dependent performances among different programs could originate, on one hand, from the

Table 4.4 Comparison of localization experimental results of 52 Chr21 proteins with in silico prediction

Gene name	GenBank Protein Acc.No.	Funtion class	Localization in HeK293T	Bioinformatic Localization Prediction			
				PENCE	ProtComp5	PSORTII	SubLoc
ABCG1	CAA62631.1	ATPase	PM/Golgi	ER	PM	PM	Extracell
AGPAT3	AAH11971.1	acyltransferase	ER/PM(less)	ER	PM	ER	Cyto
B3GALT5	NP_006048.1	galactosyl-transferase	Golgi/ER	Golgi	Golgi	extracell	Mito
BACH1	BAA24932.1	transcription regulation	Cyto(punct) Nuc-M-phase	Nuc	Nuc	Nuc	Nuc
C21orf103	NP_853633.1	unclear	Cyto	Cyto	Extracell	Cyto	Nuc
C21orf19	AAL34462.1	unknown	Nuc/Cyto	Cyto	Mem-bound-Extracell	Mito	Nuc
C21orf25	XP_032945.2	unknown	Nuc/Cyto	extracellular	Nuc	Nuc	Nuc
C21orf30	CAB56001.2	unknown	Nuc	__	Extracell	Nuc	Nuc
C21orf4	AAC05974.2	unknown	PM	PM	PM	PM	Cyto
C21orf59	AAG00496.1	unknown	Nuc/Cyto	__	Cyto	Cyto	Cyto
C21orf69	AAK60445.1	unknown	ER	__	Extracell	Nuc	Nuc
C21orf96	NP_079419.1	unknown	Cyto (punct)	__	Nuc	Nuc	Mito
CBS	NP_000062.1 (splicing isoform)	lyase	Cyto	Cyto	Cyto	ER	Cyto
CCT8	BAA02792.1	chaperonin	Cyto	Cyto	Cyto	ER	Cyto
CHAF1B	NP_005432.1	chromatin assembly factor	Nucleoplasm Cyto-M phase	Nuc	Cyto	Nuc	Cyto
CLDN14	AAG60052.1	tight junction	ER/PM(less)	PM	PM	ER	Extracell
CLDN17	CAB60616.1	tight junction	PM/Golgi	PM	PM	ER	Extracell
CLDN8	NP_036264.1	tight junction	ER/PM(less)	PM	PM	ER	Extracell
CRYZL1	BAA91605.1	oxidoreductase	Cyto	Cyto	Cyto	Cyto	Cyto
CXADR	AAH10536.1	receptor	PM	PM	PM	extracell	Nuc
DNMT3L	AAH02560.1	methyltransferase like	Nuc/Cyto	Nuc	Nuc	Cyto	Cyto
DSCR3	NP_006043.1	unknown	Nuc	Cyto	Cyto	Cyto	Cyto
ETS2	NP_005230.1	transcription factor	Nuc	Nuc	Nuc	Nuc	Nuc
*GCFC**	AAD34617.1	transcriptional repressor	cyto	Nuc	Nuc	Cyto	Nuc
HLCS	NP_000402.2	protein ligase	Cyto	Cyto	Mem-bound-Cyto	Cyto	Cyto
HMGN1	AAA52676.1	DNA binding	Nuc	Nuc	Nuc	Nuc	Nuc
HSF2BP	NP_008962.1	transcription factor binding	Cyto	Cyto	Extracell	Cyto	Cyto

IFNGR2	AAH03624.1	receptor	ER/PM(less)	extracell	PM	ER	Extracell
KCNE1	AAH36452.1	K-channel	Lyso/PM	PM	PM	ER	Extracell
KCNE2	NP_005127.1	K-channel	Lyso/PM	PM	PM	Cyto	Extracell
KCNJ15	NP_002234.2	K-channel	PM/Golgi	PM	PM	ER	Cyto
KCNJ6	NP_002231.1	K-channel	PM/Golgi	PM	PM	ER	Cyto
KIAA0179	XP_035973.4	unknown	Nuc/Cyto(punct)-M phase	Nuc	Nuc	Nuc	Nuc
MCM3AP	BAA25170.1	DNA binding	Cyto/Nuc	Nuc	Cyto	Cyto	Cyto
MX1	NP_002453.1	dynamin & large GTPases	Cyto(punct)	Cyto	Cyto	Cyto	Cyto
NNP1	AAH00380.1	RNA processing	Nucleolus	Nuc	Nuc	Nuc	Cyto
PCBP3	AAH12061.1	RNA binding	Cyto/Nuc	Nuc	Mem-bound-Nuc	Nuc	Cyto
PCP4	CAA63724.1	unknown	Nuc/Cyto	Cyto	Nuc	Nuc	Mito
PDE9A	AAH09047.1	phosphodiesterase	Cyto (accum)	Cyto	Nuc	Nuc	Nuc
PDXK	AAH00123.1	kinase	Cyto	Cyto	Cyto	Cyto	Extracell
PFKL	AAH09919.1	kinase	Cyto (accum)	Cyto	Mem-bound-Mito	Cyto	Mito
PKNOX1	AAH07746.1	transcription factor	Nuc/Cyto	Nuc	Nuc	Nuc	Nuc
PPIA3L	CAA37039.1	peptidylprolyl isomerase A	Nuc/Cyto	Cyto	Cyto	Cyto	Cyto
RPS5L	pseudogene, 81% identity to BAB79493.1	unknown	Cyto	Cyto	Nuc	Nuc	Extracell
SH3BGR	AAH06371.1	SH3 adaptor	Cyto	Cyto	Nuc	Cyto	Nuc
TAK1L	AAF81754.1	transcription factor like	Nuc/Cyto	Cyto	Cyto	Nuc	Extracell
TMPRSS3a	NP_076927.1	protease	ER	extracell	PM-	Cyto	Extracell
TSGA2	NP_543136.1	chromosome-associated	Cyto /Nuc	Cyto	Cyto	Cyto	Cyto
UBASH3A*	NP_061834.1	catalytic activity	Cyto	Nuc	Nuc	Nuc	Cyto
UBE2G2	AAC32312.1	ligase	Cyto	Cyto	Extracell or Cyto	Cyto	Nuc
WDR4	AAH06341.1	unknown	Nucleoplasm	Cyto	Extracell	Cyto	Cyto
WDR9_3'	BAA92123.1	unknown	Nuc	Nuc	Nuc	Nuc	Nuc

The localization property of 52 Chr21 proteins obtained in Hek293T cells was compared with prediction results given by 4 computational programs. Accum, accumulated; Cyto, cytosol; ER, endoplasmic reticulum; Extracell, extracellular secreted protein; Lyso, lysosome and endosome; Mem-bround, membrane-bound; Mito, mitochondria; Nuc, Nucleus; PM, plasma membrane; Punct, punctated.

different prediction methods the programs use, on the other hand, from the different amount of knowledge on the protein targeting mechanisms for different localization sites. For example, the sequence and structure of signal peptide (SP), a motif directing proteins to the ER membrane, is more well-studied compared to the nuclear localization signals (NLS). Therefore, further discovery of protein targeting motifs and their underlying mechanism will help improve the accuracy for protein localization prediction.

Computational prediction *vs* Experiment results
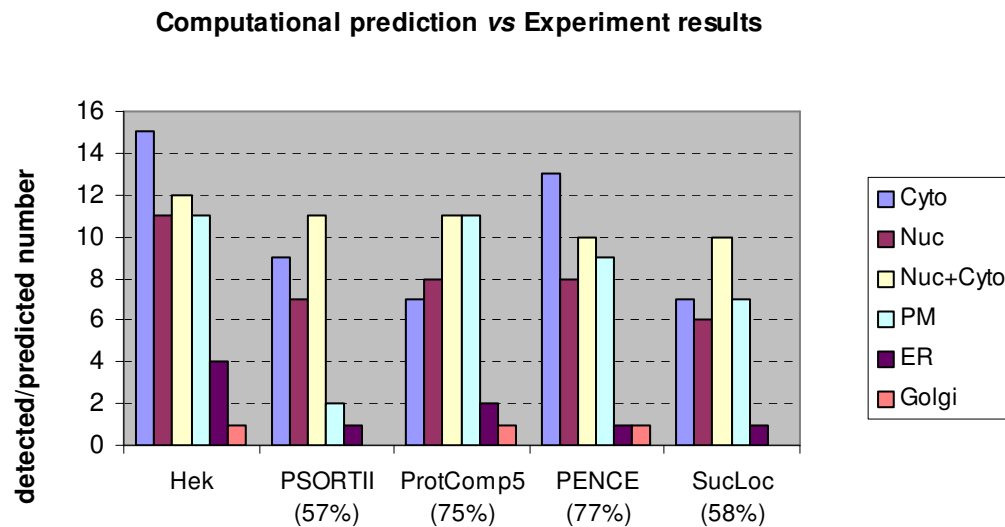
Fig. 4.9 Comparison of prediction performances of different computation programs

The prediction performance was found to vary among different programs. PENCE and PrompComp presented a significantly better performance with 77% and 75% of the prediction in consistency with experimental results, whereas SubLoc and PSORTII provided only 58% and 57% consistent predication, respectively. Beside, the prediction accuracy was found to associate with specific localization site.

The results showed the prediction performance varied largely among different programs and different localization categories. As a result, it is always wise to apply more than one localization predictor, especially those using different prediction methods. Moreover, when interpretating the predication results, a special attention should be paid to the relative confidence scores assigned to different localization sites. Generally, a big difference between the second best score and the best one implies a reliable prediction whereas similar scores obtained for different locations may reflect either unreliability of the prediction or the protein's multiple localization patterns. A good example provides protein C21orf7 in our study. C21orf7 (TAK1-like) gene

shares a homology with human TAK1 (TGF-beta activated kinase), which play a critical role in the TGF-beta signaling transduction pathway. For this protein, ProtComp5 gave the two best localization scores 2.6 and 1.95 to the cytoplasm and the nucleus, while PSORT II predicted the protein localizing in the nucleus and the cytoplasm with 35% and 30% probability, respectively. In our experiments, the actual localization of this protein was quite dynamic as we found it in both the cytoplasm and nucleus.

In some cases, the predictions might be false due to many unknown mechanisms of protein sorting, even though more than one predictor were used. In this study, the actual localization of several proteins was found in disagreement with all predictors. For example, gene WDR4 encodes a member of the WD repeat protein family and is a candidate for some disorders mapped to 21q22.3 as well as for Down syndrome phenotypes (Michaud et al., 2000). ProtComp5 predicted WDR4 to be an extracellular protein whereas the other three programs predicted it in the cytoplasm. In actual experiment, WDR4 proteins were found to distribute in the nucleus, concentrating in the nucleoplasm. The yeast homologue of WDR4, Trm82, is known to be required for 7-methylguanosine modification of tRNA (Alexandrov et al., 2002). Because such pre-tRNA processing has been known to occur in the nucleoplasm before the resulted mature tRNAs are transported out to the cytoplasm (Lodish et al., 2000), Trm82 was expected to locate in the nucleus, especially in the nucleoplasm as we observed for WDR4. Although the functional role of WDR4 in human cells has not been experimentally verified, Alexandrov et al. have found that WDR4, when forming a complex with METTL1, are required for 7-methylguanosine modification of yeast tRNA (Alexandrov et al., 2002). Supported by our localization results, it is very likely that WDR4 performs similar function in human cells on tRNA processing as its yeast homologue.

4.4. Apoptosis screening of Chr21 reveals novel cell death regulators

4.4.1. Construction of multiple death detection assays

As discussed in the introduction (section 1.3.4.1), a number of biochemical assays are commercially available for apoptosis identification, but not all of them can be applied to the cell array. Although the usefulness of a few of these assays (e.g. TUNEL) on the cell array has been suggested in previous reviews (Vanhecke and Janitz, 2004; Wheeler et al., 2005), none of the assays has been so far experimentally demonstrated. Therefore the fist step in this study was to select a suitable apoptosis detection assay for the cell array. For this purpose, a construct containing the full-length sequence of Bax protein, whose overexpression induces cell apoptosis, was used as a positive control for assay optimization. Vectors overexpressing autofluorescent protein eGFP (green fluorescence) and HcRed (red fluorescence), respectively, were included for the assessment of transfection efficiency. These two constructs were also used as negative control to evaluate the specificity of apoptosis assay.

4.4.1.1. TUNEL

The reverse transfection was initially carried out in Hek293T cells for 48 hours and the transfected cell array was subjected to TUNEL (FITC-conjugated) reaction (see section 3.7.1). The simultaneous use of anti-Bax antibody (together with AlexaFluor 568-conjugated secondary antibody) allowed distinguishing the apoptotic cells induced by overexpression of Bax from those undergoing spontaneous cell death. On the same array, the spontaneous apoptosis background (TUNEL-positive and Bax-negative cells) was randomly distributed throughout the entire array slide, whereas the Bax-induced apoptotic cells (both TUNEL and Bax positive) were highly concentrated in the clusters of Bax spots, and could be easily distinguished from the non-transfected surrounding cells (Fig. 4.10). The expression of eGFP and HcRed did not increase the cell death detected by TUNEL.
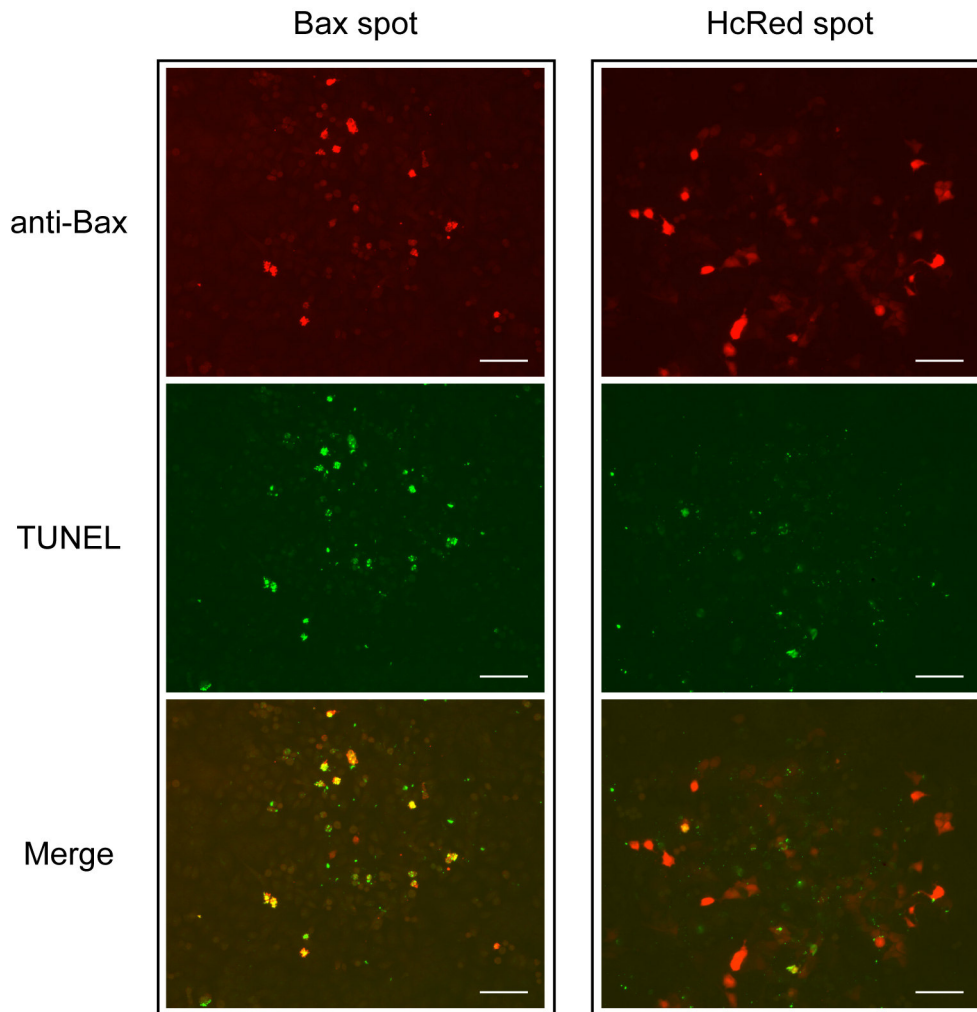
Fig. 4.10 Apoptotic cell clusters detected by TUNEL reaction

Overexpressing of Bax (left panel) but not HcRed (right panel) resulted in an apoptotic cell cluster detected by TUNEL. Anti-Bax: immunofluorescence labeling of cell clusters using anti-Bax antibody. Red signal in Bax spot indicates overexpression of Bax proteins, while the red signal in HcRed spot indicates overexpression of HcRed proteins. TUNEL: dying cells were labeled with green fluorochrome by using FITC-conjugated TUNEL assay. Merge: combination of anti-Bax and TUNEL labeling. Overlap of red and green signal results in yellow spot, indicating the apoptotic cells induced by Bax overexpression. Scale bars, 100μm.

To study the dynamics of spontaneous cell death and Bax induced apoptosis, the same tests were carried out at 24, 48 and 72 hours after transfection. The number of spontaneous apoptotic cells went up continuously and reached the highest level at 72 hours, when cells appeared to overgrow. The expression of Bax slightly increased from 24 hours to 48 hours and became stable afterwards, so was the number of apoptotic cells induced by Bax overexpression. The similar trend was observed in HeLa cells. Therefore, 48 hours after the transfection was chosen as the optimal time point for the subsequent TUNEL detection experiments. Taken together, TUNEL

reaction, when used on cell arrays, proved to be a simple, quick and specific assay for apoptosis detection.

Bax is a key regulatory protein in the mitochondrial apoptotic pathway. Overexpression of Bax has been known to induce apoptosis. In our experiments, however, not all of the cells overexpressing recombinant Bax were positive to TUNEL reaction. Cells at early stage or latent phase (the first phase of apoptosis where the cell looks still morphologically normal) of apoptosis were unable to be detected (Fig. 4.11 b, c). Only the cells at the late stage of apoptosis (cell cytoplasm and nucleus shrank; nuclear DNA underwent severe fragmentation; cells started to detach from the glass surface, resulting in raised focus under microscopy) could be detected (Fig. 4.11 a). In addition, the abundance of exogenous Bax was found to vary from cell to cell, probably due to the fact that different cells took up different quantities of vectors over the time in Bax transfection. The variation in Bax expression might lead to differential apoptotic stages among cells at the time of detection and thus accounted for the TUNEL-negative cells at the early stage of apoptosis. Therefore multiple cell death assays were developed for the detection of apoptosis at different stages.

4.4.1.2. Triple death assay

Biotin or FITC conjugated Annexin V, antibodies against activated caspases (caspase-3, -7, -9) and cleaved PARP, and FLICAs were tested on on cell arrays and the staining protocols were optimized respectively (see section 3.6). Annexin V-Biotin and cleaved caspase-3 antibody were found to give the most sensitive and specific staining, and thus used for the construction of "triple death assay" in subsequent tests. The rest of the assays did not perform satisfactory staining when used on cell arrays. Annexin V- FITC and the cleaved PARP antibody stained too weakly to be capable for large-scale screening, whereas the detection of cleaved caspase-7 and -9 with the antibodies used appeared to be weak and unspecific. FLICAs, the inhibitors of activated caspase, proved to be functional with limited applications, even though the signals weakened after the cell fixation and severe treatments of immunofluorescent staining.
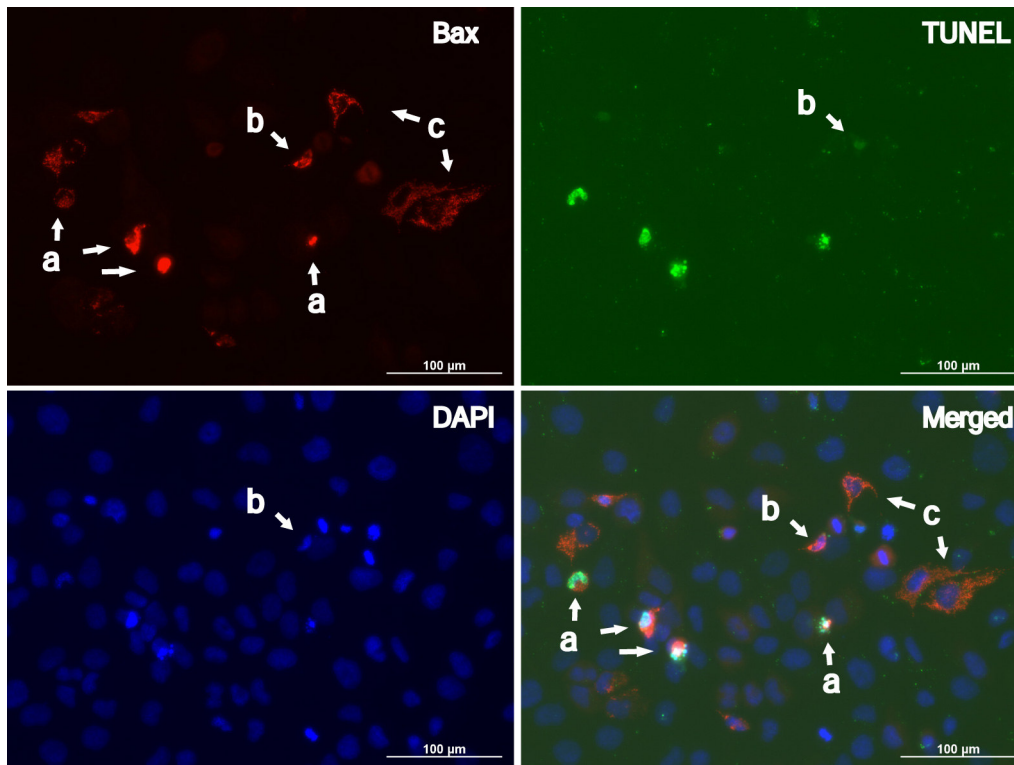
Fig. 4.11 Cells at different stages of apoptosis were differentially detected by TUNEL

The same cell cluster transfected by Bax gene was labeled with anti-Bax antibody, TUNEL and DAPI nucleus dye, as indicated. Cells overexpressing Bax underwent apoptosis at different time. Late stage of apoptosis could be well detected by TUNEL (a); early stage of apoptosis was weakly detected (b); (c) cells in latent phase (the first phase of apoptosis where the cell looks still morphologically normal) of apoptosis were negative to TUNEL. Scale bars, 100μm.

To test whether Annexin V-Biotin and anti-cleaved Caspase3 detect the same apoptotic stage as TUNEL, three assays were combinatorially applied to the same apoptotic cell clusters induced by Bax overexpression. The assays were found not always targeting at the same apoptotic cells. Some cells that were positive to Annexin V were not detected by Caspase3 and TUNEL assays (Fig. 4.12 A). The staining difference might originate from the biological bases of the assays. On cell arrays, Annexin V was found dominantly binding to the cell plasma membranes where the asymmetry of the membrane structure has been lost. Because cell membranes were more abundant at early stage of apoptosis, Annexin V binding of these cells was much more intensive compared to the cells at later stage of apoptosis with fragmented membranes. In contrast, most of the TUNEL-sensitive cells were often characterized with collapse in the cytoplasm and nucleus, and most frequently with cell blebbing or

fragmentation—the features of very late stage of apoptosis. Activated caspsase3 were often detected in the cells between middle and early late stage, whose cytoplasm shrank and the nucleus lost its integrity but without fragmentation. Thanks to their differential performances, the three assays (Annexin V-Biotin, cleaved Caspase3 antibody, and TUNEL) when used with separate color fluorochromes allowed distinguishing the cells at different stages of apoptosis. On the other hand, all assays can be used with a same color fluorochrome, termed as "triple death assay" in this study, to collect the nonsynchronous apoptotic signals ranging from early to late stage (Fig. 4.12 B). This is of particular use in order to eliminate the false-negatives in large-scale apoptosis screening where cells may respond to the death inducers at various rates.

4.4.2. Over-expression of two Chr21 proteins (claudin-14 and -8) induces the cell death undergoing non-apoptosis pathway

4.4.2.1. "Triple death assay" detected the claudin-induced cell death

The apoptosis assays were then applied to study chromosome 21 proteins for apoptosis induction on Hek293T cell arrays. The combined "triple death assay" (Annexin V-Biotin, cleaved Caspase3 antibody, and TUNEL) was performed to collect the nonsynchronous death signal. Again, anti-His antibody conjugated with another colored fluorochrome was simultaneously used to indicate the overexpression of Chr21 proteins. Three tight-junction proteins (claudin-8, -14 and –17) were investigated since the overexpression of claudin-8 and claudin-14 was found in this study to result in severe cell shrinkage, indicating possible processing of apoptosis. In "triple death detection", claudin-8 and -14 were found to be positive, whereas claudin-17 showed negative staining. As shown in Fig.4.13, 48 hours after the reverse transfection, over 70% of the cells overexpressing claudin-14 were positively stained by the triple assay, whereas none of the cells overexpressing claudin-17 was stained. In the subsequent time-course experiments performed in both Hek23T and HeLa cells at 24, 48 and 72 hours after transfection, at least 70% positive staining was obtained for claudin-14-transfected cells irrespective to the transfection time and cell lines. To answer the question if these results are cell array-dependent, normal transfections of
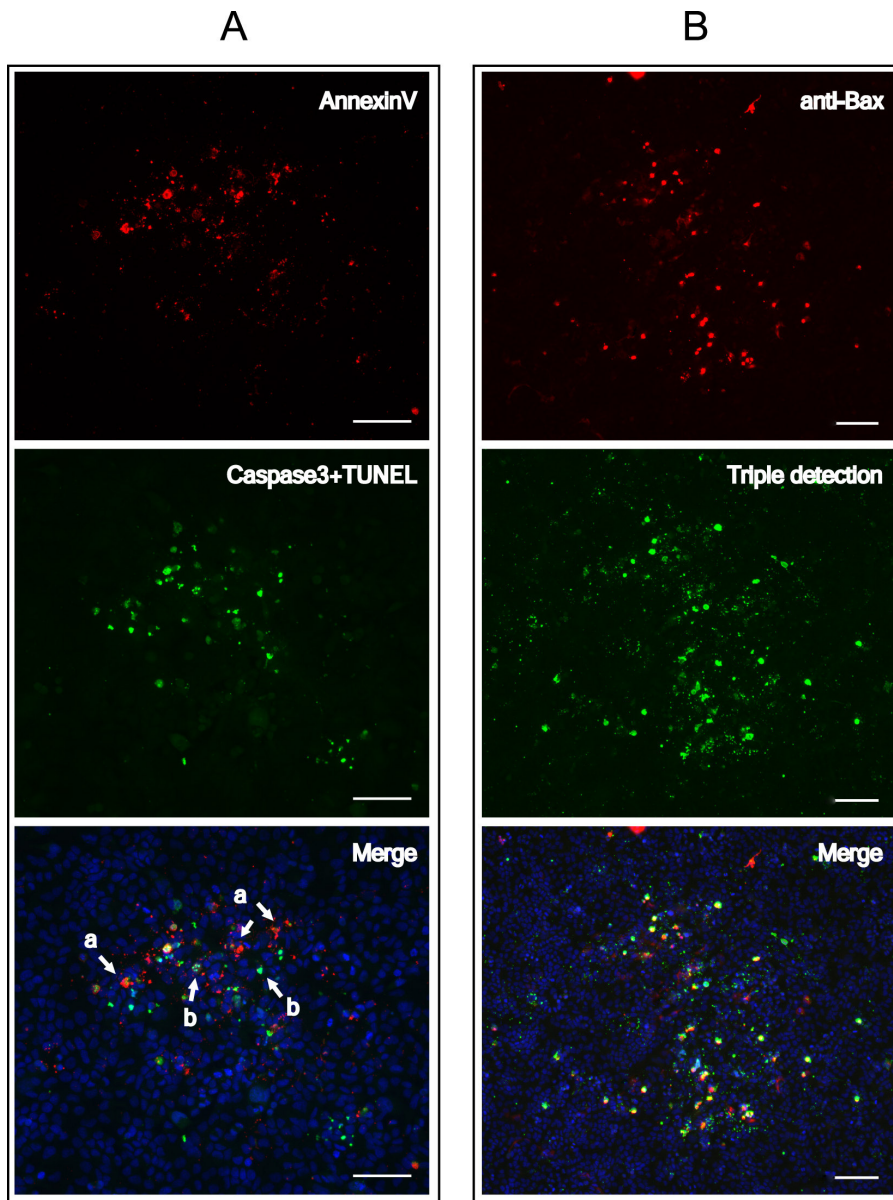
A   B

Fig. 4.12 Multiple apoptosis detection on cell arrays

(A) Different assay (as indicated) targets the cells at different stage of apoptosis. Annexin V, anti-activated Caspase3 antibody and TUNEL were used to label the same apoptotic cell cluster induced by Bax overexpression with different fluorochromes (as indicated). Non-overlapped signals in the merged image (bottom) indicate different apoptosis stages of cells. Cells in red (a) were stained by Annexin V only, a marker for early stage of apoptosis; green cells (b) underwent middle to late stage of apoptosis detected by TUNEL and antibody against activated capase3. (B) Combined apoptosis assays can collect the nonsynchronous death signals induced by Bax overexpression. Anti-Bax: immunofluorescence labeling using anti-Bax antibody, indicating a cell cluster overexpressing Bax. Triple detection: combined Annexin V, activated Caspase3 Ab and TUNEL assays. Merge: combination of anti-Bax staining and triple death detection. Scale bars, 100μm.

the three claudins were performed in 6-well plates and the "triple death assay" was applied to the cells fixed on glass coverslip. Again, most of the cells overexpressing claudin-14 or claudin-8 were positive to death detection, whereas only the cells undergoing spontaneous apoptosis were found in the case of claudin-17 transfection.

The findings from both the cell array and normal transfection experiments support the thesis that the overexpression of claudin-14 and -8 but not claudin-17 induces cell death. Interestingly, in the protein subcellular localization study, claudin-14 and -8 had been found predominantly in the ER instead of the plasma membrane, where tight-junction proteins (including claudin-17) localize. Hence, the ER retention of claudin-14 and –8 likely correlates to the cell death induced by their overexpression.
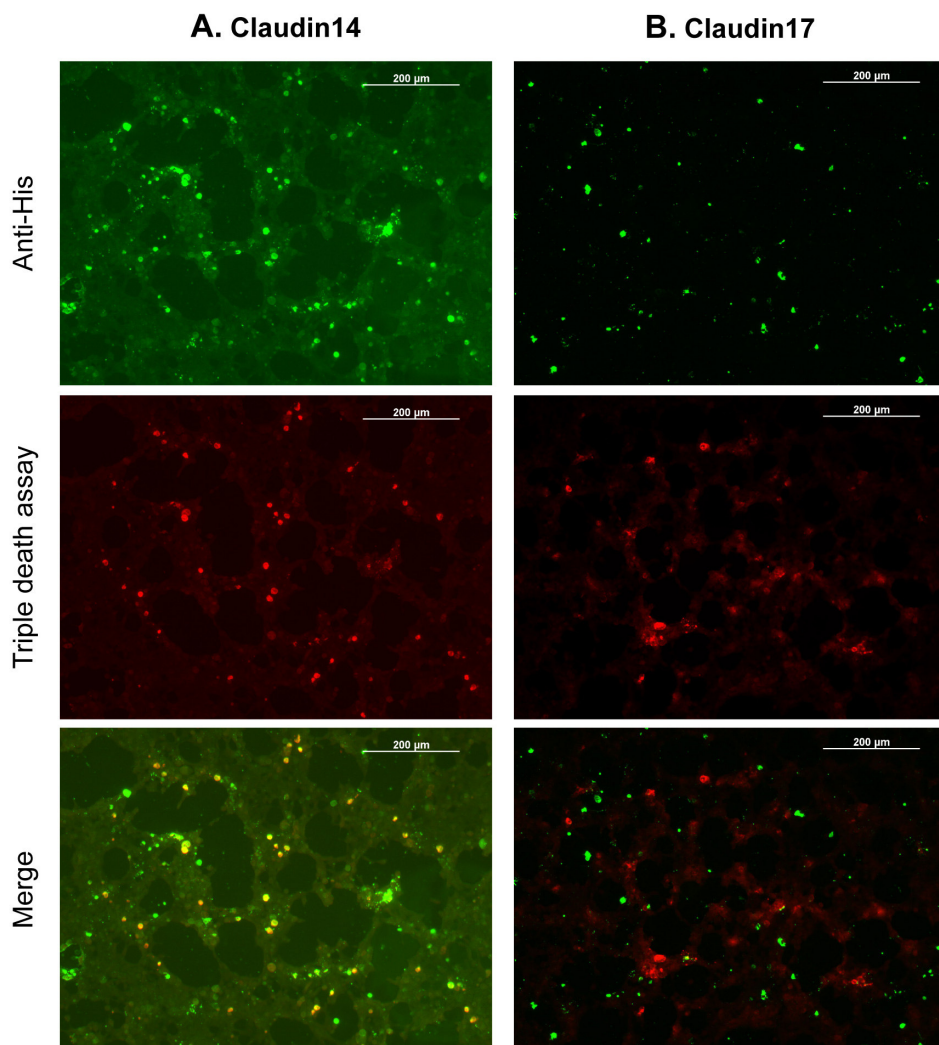

Fig. 4.13 Over-expression of claudins lead to programmed cell death

Overexpression of claudin 14 (A), but not claudin 17 (B), resulted in positive detection by "triple death assay". Upper images, the cell clusters transfected by the spots of plasmid coding for claudin 14 (upper left) or claudin 17 (upper right), respectively. The expressed proteins

were labeled with anti-His antibody that was conjugated with green fluorescence. Middle images, simultaneous detection of the same cell clusters by Annexin V, activated Caspase3 antibody and TUNEL with red fluorescence; Merged images: combination of anti-His staining and death detection, the yellow cells in left image indicating the dying cells as a result of claudin 14 overexpression. Scale bars, 200μm.

4.4.2.2. Separate use of the assays revealed a non-apoptotic form of cell death

Despite cell shrinkage and nuclear DNA degradation (see section 4.2.3), the additional typical apoptosis features such as cell blebbing were not observed in the dying cells following claudin-14 or -8 overexpression. The question whether the cell death induced by the two claudins is a classic caspase-dependent apoptosis was subsequently addressed.

Annexin V, activated Caspase3 antibody, and TUNEL assays were separately applied in order to find out the origination of positive staining and thus to elucidate the type and possible mechanism of the induced cell death. Bax, GFP, HcRed were introduced in each array as positive (Bax) and negative (GFP and HcRed) controls for apoptosis induction. In the cells transfected by claudin 14 or 8, Annexin V rather than the other two assays contributed to the positive labeling within the "triple death assay". Under higher magnification, Annexin V was found to bind specifically to the outer leaflets of the cell membrane that wrapped the condensed ER where claudin-14 localized (Fig. 4.14 A). Moreover, the nuclear DNA of the dying cells appeared to degrade as the DAPI staining greatly decreased. For the cells at late death stage (the cytoplasm extremely condensed), the nuclear DNAs were barely stained by DAPI (Fig. 4.14 B). TUNEL-positive staining, however, was only observed in few cells overexpressing claudins (Fig. 4.15), suggesting the DNA fragmentation is not compulsory for the claudin-induced cell death. Surprisingly, caspase-3, an essential processor for classic apoptosis, were not actively cleaved in the dying cells induced by the two claudins (Fig. 4.16). Meanwhile, the typical apoptosis phenotypes such as cell blebbing and the formation of apoptotic bodies were not found in the case of claudin-induced cell death. These results were confirmed by normal transfections subjected to separate death assay. Thus, it might be hypothesized that the cell death, as a result of claudin-14 and -8 overexpression, undergos an alterlative pathway that is independent to classic caspases-mediated apoptosis.
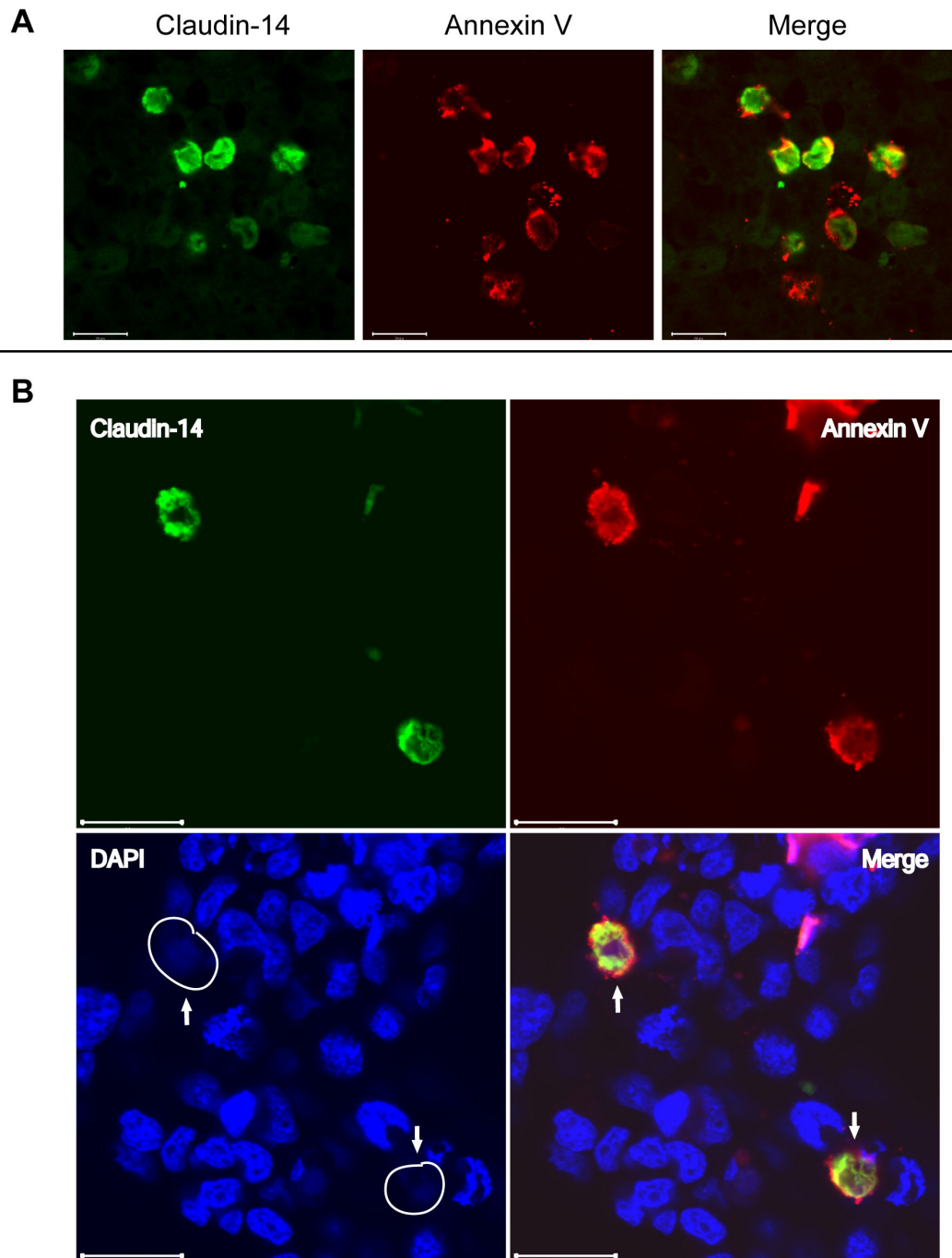
Fig. 4.14 Cell death induced by claudin-14 overexpression is positive to Annexin V

Annexin V was found to bind specifically to the outer leaflets of the cell membrane where the membrane asymmetry was lost due to claudin-14 overexpression (A and B). In the cells overexpressing claudin-14, the nuclear DNA degraded and was barely stained by DAPI (arrow). Green fluorescence: overexpressed claudin-14 labeled with Alexa488-conjugated antibody against the His_6 epitope fused to claudin-14. Red fluorescence: biotin-conjugated Annexin V bound to Alexa594-conjugated streptavidin. Merge: combination of claudin-14, Annexin V, and DAPI staining. Scale bars, 20µm.
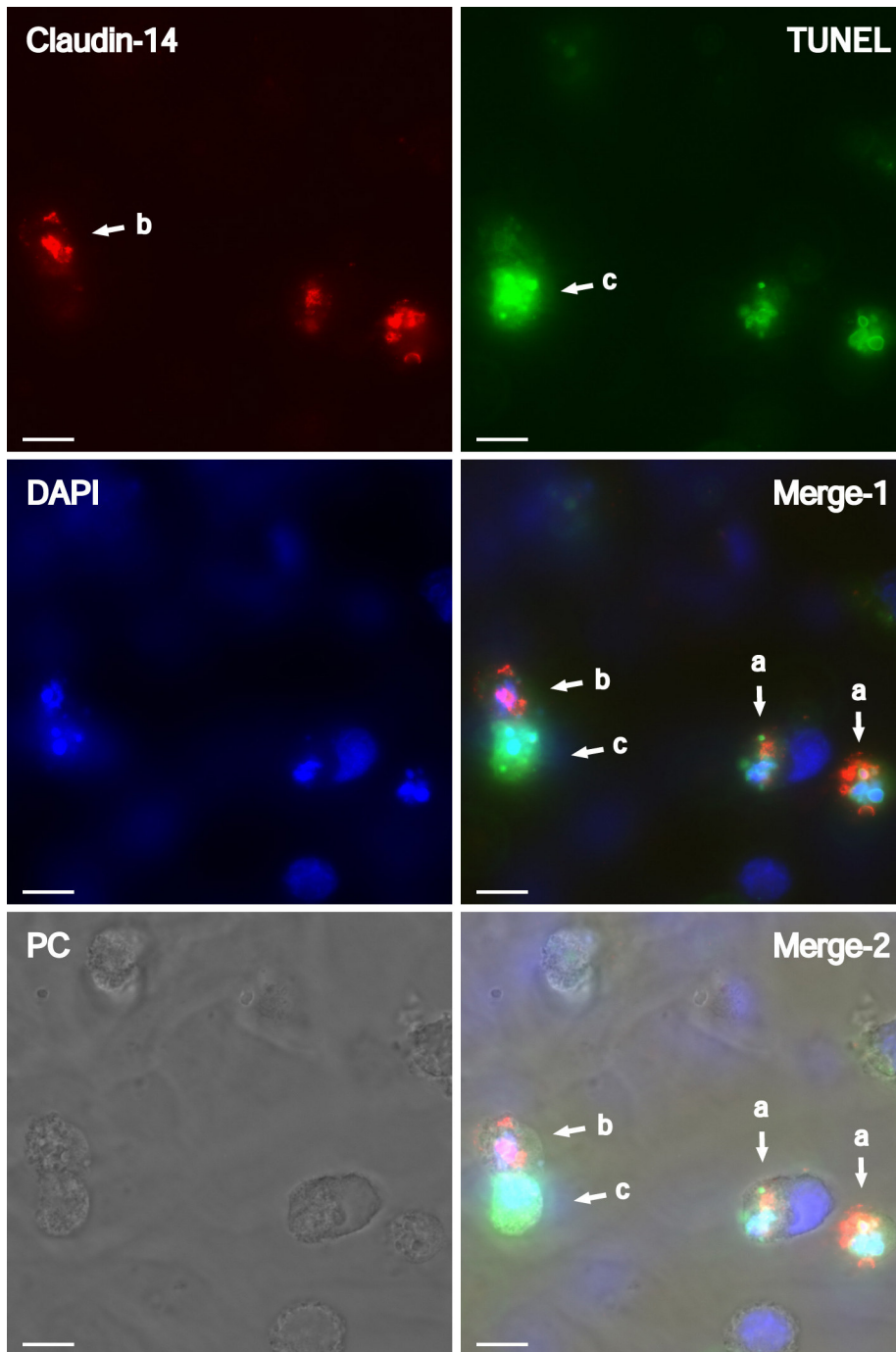
Fig. 4.15 Cell death induced by claudin-14 overexpression is only partially positive to TUNEL

DNA fragmentation detected by TUNEL reaction was observed in some (a) but not all of the cells overexpressing claudins (b). (c), a spontaneous apoptotic cell positive to TUNEL reaction. Red fluorescence: overexpressed claudin-14 labeled with Alexa555-conjugated antibody against the His$_6$ epitope fused to claudin-14. Green fluorescence: FITC-conjugated TUNEL reaction. PC: Phase contrast. Merge-1: combination of claudin-14, TUNEL, and DAPI staining; Mege-2: combination of merge-1 and phase contrast. Scale bars, 20µm.
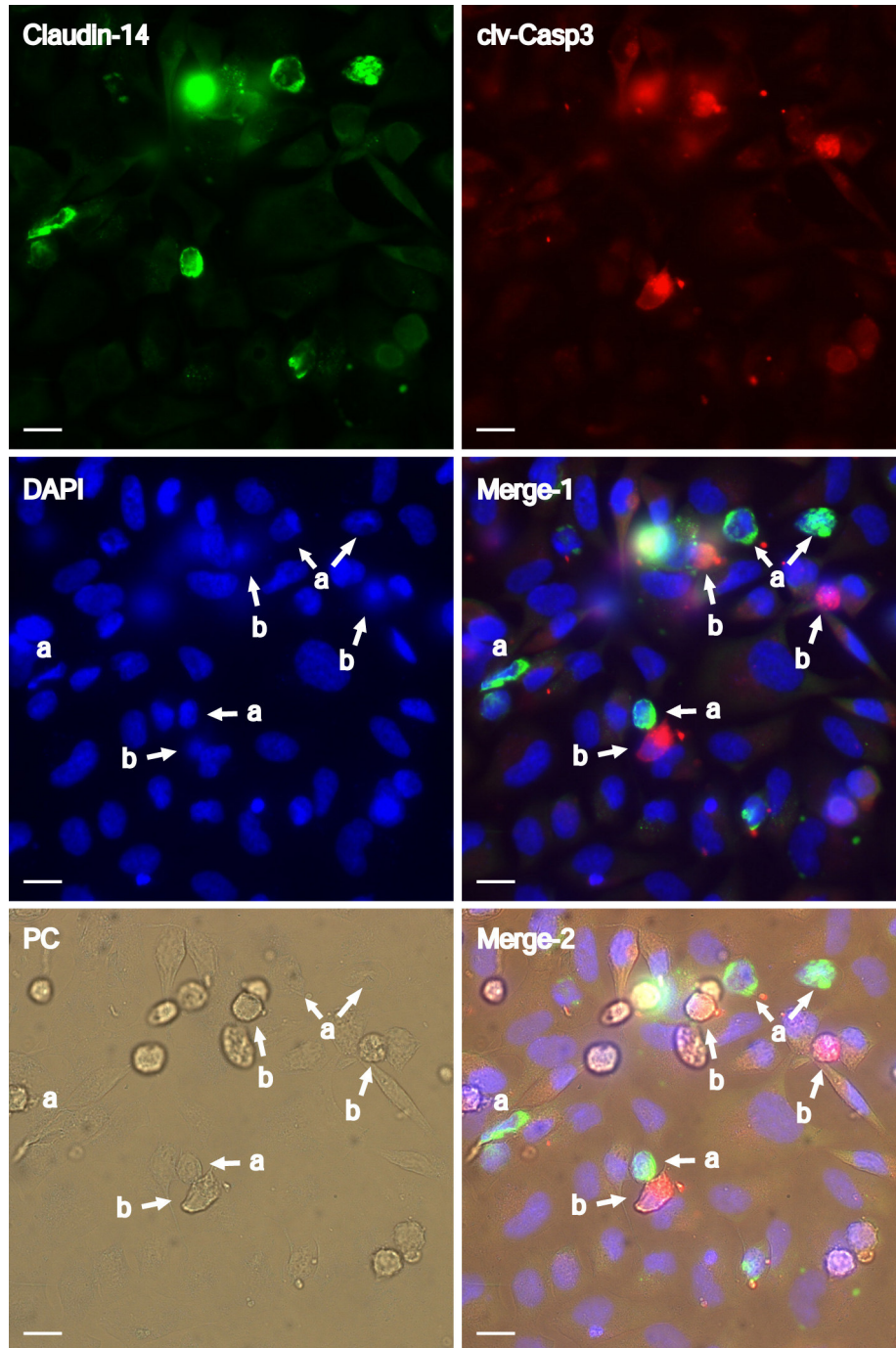
Fig. 4.16 Cell death induced by claudin-14 overexpression is caspase-3 independent

Cleaved caspase 3, an essential processor for classic apoptosis, was not found in the dying cells induced by the claudin overexpression (a), but only found in the spontaneous apoptotic cells (b). Green fluorescence: overexpressed claudin-14 labeled by using Alexa488-conjugated antibody against the His_6 epitope fused to claudin-14. Red fluorescence (clv-Casp3): spontaneous apoptotic cells labeled by the antibody against cleaved capsase-3 (Asp293). PC: Phase contrast. Merge-1: combination of claudin-14, cleaved capsase-3, and DAPI staining; Mege-2: combination of merge-1 and phase contrast. Scale bars, 20μm.

Claudin-correlated non-classic cell death was previously described in epidermal development at early embryogenesis of vertebrates. Claudin-1 was found to localize in the periderm and the subperiderm of chicken embryo, and its function in tight-junctions was involved in the formation of the embryonic diffusion barrier, which is lost by desquamation at the end of embryonic period. Interestingly, a special form of programmed cell death was found to account for the desquamation of the embryonic diffusion barrier. The cell death occurring at the periderm and the subperiderm was partially positive to TUNEL, whereas the activity of caspase-3, -6, and -7 was found to be absent (Saathoff et al., 2004). This type of ell death was very similar to the one induced by claudin-14 and –8 overexpression as observed in this study. All together, these results suggested a non-classic type of programmed cell death in close correlation to certain claudins.

4.5. RNAi-induced apoptosis on cell arrays

In previous sections it has been demonstrated that the multiple death assays, i.e. "triple death assay", when used on cell arrays were sensitive enough to identify the cell death as a result of protein overexpression. This strategy can be applied for large-scale gene screening with the aim to identify novel genes that function as cell death inducers. In principle, in combination with RNA interference (RNAi) technology, the same platform can also be used to identify the genes that function as cell death suppressors, if only the down-regulation of the genes induces cell death. Hence, a proof-of-principle experiment combining RNAi, cell array technique and apoptosis assays was carried out in this study.

4.5.1. Development of siRNA-cell arrays for endogenous gene silencing

A few groups had reported the application of cell array technique in RNAi silencing (Silva et al., 2004; Wheeler et al., 2005). Traditionally, chemically synthesized siRNA and the DNA constructs expressing shRNA in transfected cells were used as the resources for gene silencing. In this project, the chemically synthesized siRNAs that are commercially available were used.

To evaluate the knockdown efficiency of siRNA on cell arrays, two siRNAs targeting human gene lamin A/C and gene hnRNP-A3, respectively, were initially analyzed. Alexa488 or Rhodamine-conjugated siRNA whose sequence does not target any human gene was used as negative control in order to assess silencing specificity and transfection efficiency. Prior to spotting, all siRNAs were treated with the transfection reagents for siRNA transfection (see section 3.2.1.2). The arrayed siRNA slides were air-dried and transfected into HeLa cells. The process of the reverse transfection was monitered with fluorecent microscope on fluorochrome-labeled control siRNA. The control siRNAs were found to enter the cells a few hours after the transfection and retained in the cytoplasm up to 72 hours after transfection. At the 24, 48 and 72 hours after transfection, the cell arrays were fixed for immunofluorescent labeling with the antibodies against human protein Lamin A/C and hnRNP-A3, respectively. The silencing efficiency of the endogenous genes could therefore be evaluated through the reduction in the immunofluorescence intensity. For example, cells on top of a siRNA spot targeting lamin A/C would present a "dark spot" (decrease in immunofluorescence) in anti-lamin A/C staining, but not in anti-hnRNP-A3 staining, and vice versa. GenePix 6.0 software (Axon laboratory) was used to quantify the signal reduction at each "dark spot" to assess the silencing efficiency. The results of time-course experiments showed that gene down-regulation was already achieved 24 hours after the transfection, equal to half of the total silencing detected at 72hours. Till the 48 hourse after transfection, the gene silencing effect has almost reached a plateau, as observed that only 5% further knockdown was achieved from 48 to 72 hours. Moreover, two different transfection reagents from Qiagen (HiperFect and RNAiFect) were ever compared in parallel. RNAiFect provided a better assistance for siRNA transfection on cell arrays, and thus was used in further studies (Fig. 4.17).
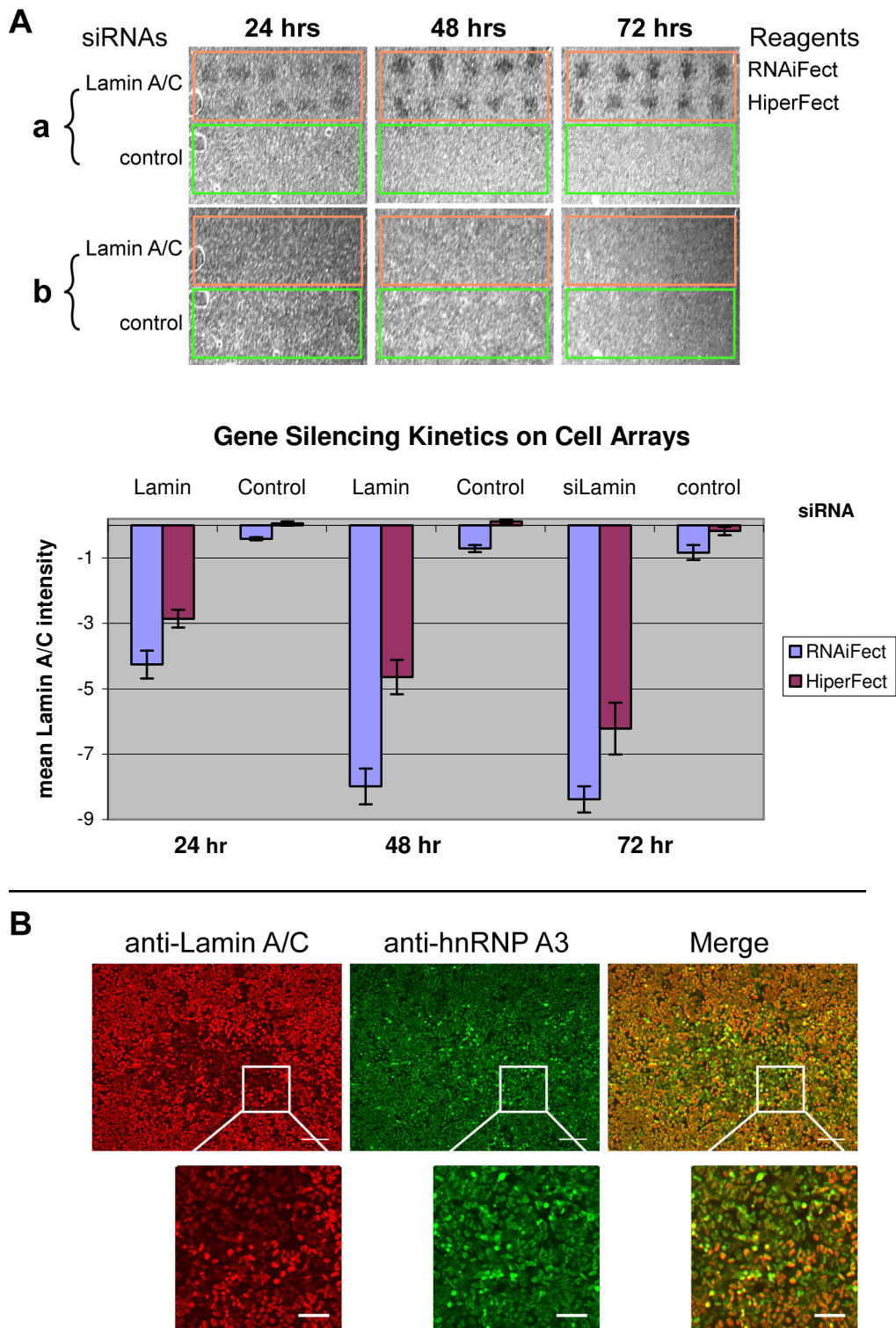
Fig. 4.17 siRNA-induced gene silencing on cell arrays

(A) Silencing kinetics of endogenous Lamin A/C. Upper panel shows three arrays scanned by BioCCD-scanner at 24, 48 and 72 hours after transfection. At the time of detection, siRNA cell arrays were fixed and labeled using antibodies against endogenous Lamin A/C (**a**) and hnRNP A3 (**b**). Cells on top of lamin A/C siRNA spots that were treated by RNAiFect or HiperFect (5 × 2, as indicated) presented a "dark spot" (diminishment of immunofluorescence) in anti-lamin A/C staining but not in anti- hnRNP A3 staining, indicating a specific silencing of lamin A/C. Control siRNA spots (5 × 2, as indicated) did not lead to the reduction in either

lamin A/C or hnRNP A3 expression. Quantification of lamin A/C knockdown using GenePix 6.0 software was performed. For the quantification, each siRNA spot was treated as one data set, and the lamin A/C expression in each spot was measured by the mean intensity inside the spot subtracted by its local background intensity. Therefore, knockdown of lamin A/C expression was presented as a minus. The measurements were carried out for all lamin A/C and control siRNA spots at different transfection time, and the averaged scores were shown in the columns. The silencing rate at 48 hours was almost at highest level as increase of the cell culture time to 72 hours only slightly (5%) improved the silencing effect. Moreover, two different transfection reagents from Qiagen (HiperFect and RNAiFect) were compared in parallel and RNAiFect revealed to be a better transfection reagent. **(B)** Microscopy images of the cell cluster on top of single lamin A/C siRNA spot (upper panel, low magnification). Lower panel, the same spot with higher magnification. Red, anti-Lamin A/C immunofluorescence (Alexa568); Green, anti- hnRNP A3 immunofluorescence (Alexa488); Blue, DAPI staining. Scale bars in low and high magnification indicate 150 μm and 50 μm, respctively.

4.5.2. RNAi-mediated apoptosis induction on cell arrays

The cell array-based multiple apoptosis detection was applied for large-scale siRNA screening towards the identification of anti-apoptotic genes. In the proof-of-principle experiments, a siRNA library known to specifically target 18 human genes was analyzed (see Table 2.1). Some of the genes are known to be essential for cell proliferation and survival, while the functions of a few genes are not fully understood. All siRNAs are commercially synthesized (Qiagen, Inc.), and the knockdown effect of each siRNA sequence had been reported in previous publication and/or confirmed by the manufacturer. All siRNAs, together with Bax-overexpressing plasmid, were printed on a batch of slides and transfected simultaneously in HeLa cells (see section 3.2). Cell death detections were carried out at 24, 48 and 72 hours after transfection. At each time-point, Annexin V, activated caspase-3 antibody, and TUNEL detection were applied all together with a same fluorochrome and separately with different fluorochromes.

Two siRNAs that target human gene SGTA (small glutamine-rich tetratricopeptide repeat-containing, alpha) and CCNB1 (CyclinB1), respectively, were found to provoke cell death. In the example shown in Fig. 4.18, siRNAs against SGTA, CyclinB1 and Lamin A/C, control siRNAs, and Bax-overexpressing plasmid were repetitively arrayed in a 7 × 2 format surrounded by GFP frame. At 48 hours after transfection, the "triple death assay" (Annexin V, activated caspase-3 antibody, and TUNEL) labeled with Alexa488 fluorochrome (green) were carried out for the detection of dying cells. Cells growing on top of SGTA or CyclinB1 siRNA spots

presented clusters of green signal, a feature similar to that produced by Bax overexpression. In contrast, cells on Lamin A/C and control siRNA spots did not manifest a green death signal (Fig. 4.18 A). Quantitative analysis suggested 2 to 3 times increase of death signal induced by CyclinB1 and SGTA siRNAs, compared to Lamin A/C and control siRNAs, respectively (Fig. 4.18 C). Separate use of three assays revealed that the dying cells induced by CyclinB1 and SGTA siRNAs contained activated Caspase-3 and were positive to Annexin V-binding and TUNEL reaction. Under light microscope, the dying cells were also found to undergo the morphological changes of classic apoptosis—cell shrinking, blebbing and fragmentized into small apoptotic bodies.
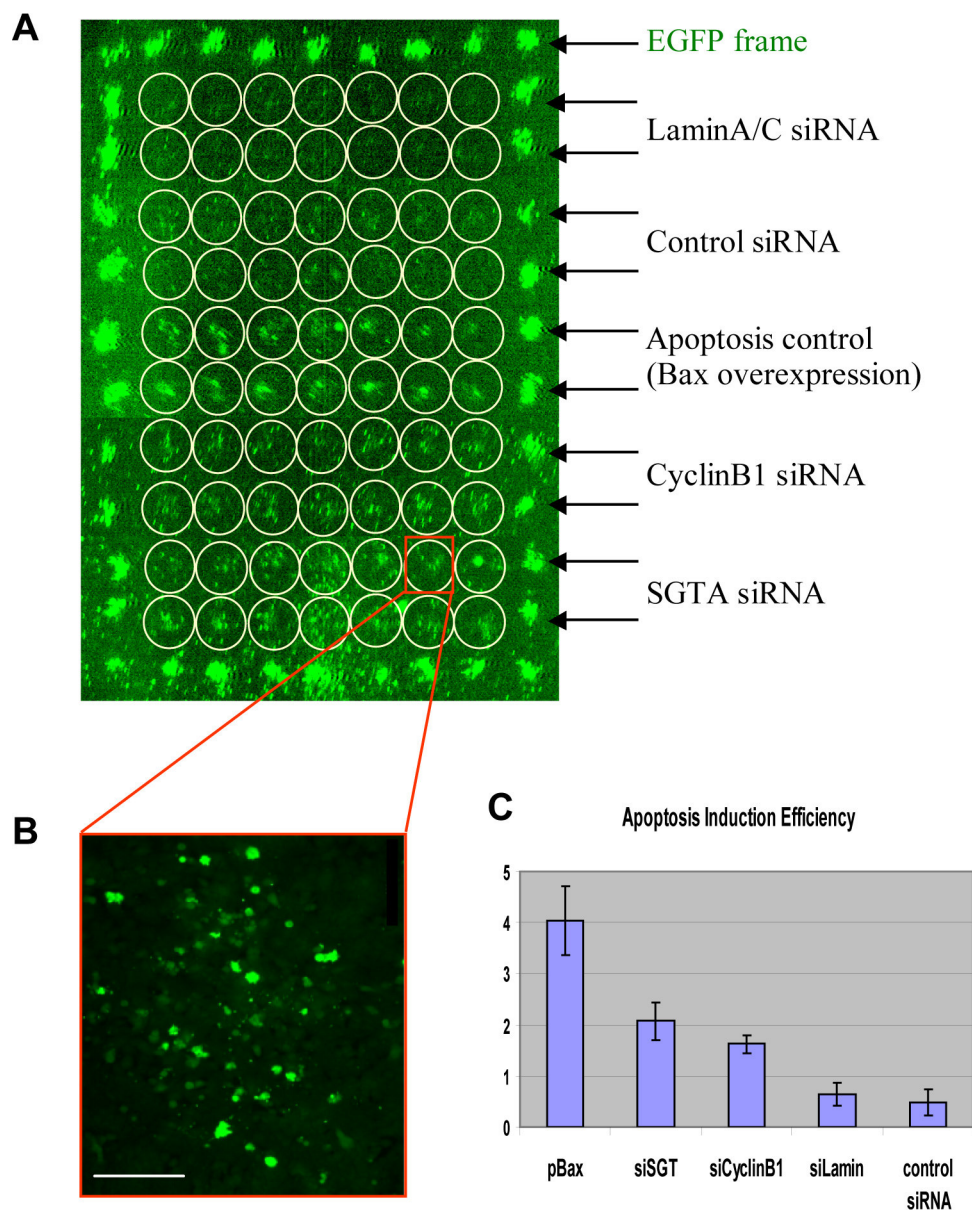

Fig. 4.18 siRNA-mediated apoptosis induction on cell arrays

(A) Scanned image of a spotting grid on apoptosis cell array. SGTA, CyclinB1, Lamin A/C and control siRNAs, and Bax-overexpressing plasmid (as indicated) were repetitively arrayed in a 7 × 2 format (as indicated by cycles) surrounded by GFP frame. 48 hours after transfection, "triple death assay" labeled with Alexa488 fluorochrome (green) was carried out to detect the dying cells. SGTA or CyclinB1 siRNA not Lamin A/C and control siRNA presented localized green apoptotic clusters, whose feature was similar to that produced by Bax overexpression. **(B)** Microscopic view on one apoptotic cluster induced by SGTA siRNA spot. Scale bar, 200 μm. **(C)** Quantification of apoptosis induction efficiency using GenePix 6.0 software. For quantification, each apoptotic cluster was treated as one data set, and the apoptosis induction efficiency was measured by the mean intensity inside the spot subtracted by its local background intensity. Quantitative analyses indicated 2 to 3 times increase of death signal induced by CyclinB1 and SGTA siRNA, compared to Lamin A/C and control siRNAs.

Cyclin B1 is a regulatory protein involved in mitosis and DNA damage-induced apoptosis (Porter et al., 2003). Down-regulation of cyclin B1 by RNAi have been reported to induce apoptosis in several cell lines including HeLa and can be detected by DNA ladder and Annexin V assays (Xie et al., 2005; Yuan et al., 2004). In this study, apoptosis induction effect of cyclin B1 knockdown was better demonstrated on living cells that could be detected by Annexin V, TUNEL and activated Caspase-3 antibody assays.

The protein encoded by SGTA is capable of interacting with the major nonstructural protein of parvovirus H-1 and 70-kDa heat shock cognate proteins. However, its function is still not clear. The expression of this gene has been found in various tissues as well as a number of human cell lines (Winnefeld et al., 2004). Interestingly, two previous reports have shown that SGTA plays a role in apoptosis regulation, but the results were completely opposite. In one report, the anti-apoptotic role of SGTA was revealed by the refrainment of NBE (SV40-transformed newborn human kidney epithelial) cells from the completion of cell division, which occured following siRNA-mediated depletion of SGTA. Consequently, the SGTA-depletive cells died of apoptosis (Winnefeld et al., 2004). Conversely, Wang et al. reported that overexpression of SGTA protein promotes apoptosis in hepatocarcinoma cells (Wang et al., 2005). The authors interpreted the inconsistency with previous study through a hypothesis that SGTA might have dual functions in apoptosis. They proposed that SGTA probably exerts its anti-apoptotic function or pro-apoptotic function through two different pathways in different cell types. This dissertation supported the anti-apoptotic function of SGTA in HeLa cells. On HeLa cell arrays, knockdown of SGTA

protein by siRNA treatment led to apoptosis, which was very likely via a classic caspase-dependent pathway. However, this work did not disprove Wang's hypothesis since I used different cell line from theirs. Further elucidation of SGTA's role in apoptosis regulation awaits more works. Nevertheless, cell array-based apoptosis detection established in this study was demonstrated to be a high-throughput and reliable platform that can be combined with large-scale siRNA screening for the identification of novel cell death regulators.