Both the cultured cells and the mutant fish exhibited decreased cathepsin D enzymatic activity, fingerprint-like inclusion body formation, and neuronal death. These findings indicated that abnormalities in lysosome function are the primary defect in KRS/PARK9.

## 2. Results

### 2.1. Wild-type ATP13A2 localizes to the lysosome, but some disease-relevant variants localize to the endoplasmic reticulum

To investigate the subcellular localization of ATP13A2, we used an anti-V5 antibody and antibodies that recognize several markers of intracellular organelles to double label SH-SY5Y cells that stably expressed wild-type ATP13A2 fused to the V5 epitope (WT-V5). The WT-V5 signal largely colocalized with the signal of cathepsin D, a lysosomal aspartic protease (Fig. 1A and S1). Additionally, WT-V5 partially overlapped with makers of the Golgi apparatus (GM130), the late endosome (Rab7) and the autophagosome (LC3B), but not with the markers of the endoplasmic reticulum (ER) (GRP78), the mitochondria (Tom20), the early endosome (EEA1 and Rab5), or the exocytotic vesicles (Rab3 and 4) (Fig. 1A, B and S1). GFP signals in SH-SY5Y cells that stably expressed GFP-tagged wild-type ATP13A2 (GFP-WT) strongly colocalized with either of two lysosomal membrane proteins, Lamp2 (Fig. 1A and S1) or Lamp1 (data not shown). Furthermore, immunoelectron microscopy using ultrathin cryosections of GFP-WT SH-SY5Y cells confirmed that gold particles labeling GFP-WT (arrow head: 10 nm gold) and those labeling Lamp1 (arrow: 5 nm gold) co-localized on the lysosomal membrane (Fig. 1C). These data indicated that ATP13A2 is a resident protein of the lysosomal membrane.

Thus far, eight disease-associated mutations have been identified in the *ATP13A2* gene, including one that we described initially [3–8]. To determine whether the mutant proteins were mislocalized and whether any such mislocalization has etiological importance, we assessed the subcellular localization of five pathogenic protein variants (Fig. 2A). These mutants were each tagged with a V5 epitope and then transiently transfected into SH-SY5Y cells. To, first, characterize the subcellular distribution of KRS mutants, we separated cell homogenates by Percoll density gradient centrifugation. This Percoll gradient system separates dense lysosomes (near bottom) from lighter particles such as ER. We found that WT-V5 was mainly co-fractionated with lysosomal component (Lamp1 and cathepsin D; Fraction No. 13–16, Fig. 2B). However, an ATP13A2 variant that lacks exon 13 (1306 + 5G→A) but maintains the reading frame was co-fractionated not with lysosomal proteins but with GRP78 (Fraction No. 2–8, Fig. 2B): this finding indicated that the protein variant accumulated in the ER. Some other ATP13A2 pathogenic variants (F182L and G504R) also accumulated in the ER in a manner similar to the 1306 + 5G→A variant. In contrast, the other two pathogenic variants (T12M and G533R) were co-fractionated with lysosomal components, as was the WT protein. To, further, confirm the subcellular localization, we carried out double fluorescence immunocytochemistry of permeabilized SH-SY5Y cells by staining KRS mutants (V5), the ER marker (GRP78), the lysosomal marker (cathepsin D) and the mitochondria marker (Tom20). Along with immunoblotting of Percoll density gradient fractionation, immunofluorescence studies showed 1306 + 5G→A mutant, F182L and G504R colocalized with GRP78 and the other two variants (T12M and G533R) colocalized with cathepsin D as already reported [10–14] (Fig. 2C and S1). The expression levels of ATP13A2 mutant proteins that accumulated in the ER (F182L, G504R, and 1306 + 5G→A) were lower than that of the WT protein and MG132 inhibited degradation of the all ATP13A2 variants as well as that of the WT protein (Fig. 2D).

### 2.2. Stable knockdown of ATP13A2 induces cathepsin D deficiency and structures that resemble neuronal ceroid-lipofuscinosis deposits

To assess whether the loss of normal ATP13A2 functions has a causal role in PD pathogenesis, the expression of endogenous ATP13A2 was suppressed in SH-SY5Y cells by gene knockdown. By using antibodies that we generated (Fig. 3A), we showed that endogenous ATP13A2 protein levels in SH-SY5Y cells that stably expressed ATP13A2 shRNA (ATP13A2shRNA-1 or -2) were efficiently suppressed (Fig. 3B). Water soluble Tetrazolium salts (WST)-8 assay demonstrated that ATP13A2 knockdown caused a significant reduction of the cell growth in SH-SY5Y cells (Fig. 3B). Next, we determined the distribution and the morphology of lysosomes in SH-SY5Y cells subjected to ATP13A2 knockdown. Immunostaining of cathepsin D and Lamp2 indicated that ATP13A2 deficiency led to the assembly of lysosomes in the perinuclear region and decreased cathepsin D staining (Fig. S2). The protein amount of full length (52-kDa), immature (44-kDa) and mature (32-kDa) forms of cathepsin D of the cells expressing ATP13A2 shRNAs were all decreased together with the enzyme activity (Fig. 3C). To show the reduction of cathepsin D activity was indeed induced by the reduction of ATP13A2, not by the non-specific effect of shRNA, we generated shRNA-resistant species of ATP13A2. SH-SY5Y cells stably expressing ATP13A2 shRNAs that were transfected with shRNA resistant ATP13A2 showed comparable cathepsin D activity to the controls (Fig. S3A). The activity and amount of mature forms (25-kDa) of cathepsin B and L were also decreased in the knockdown cells line by ATP13A2 shRNA-1, but not shRNA-2, suggesting that the loss of ATP13A2 principally gives rise to the reduction of cathepsin D (Fig. 3C). Finally, analysis via transmission electron microscopy revealed that lysosome-like bodies became more numerous in ATP13A2 shRNAs expressing cells and that very immense high density structure appeared adjacent to the nucleus in these cells (Fig. 3D: a–h). High-magnification images revealed that the abnormal structures in these cells included fingerprint profiles-like structure (Fig. 3D: e), and these structures were very similar to the structure of the neurons in mice lacking cathepsin D [15]. The levels of LC3-II, autophagosome marker, in shRNA-1 transfected cells were increased (Fig. 3E) and p62 a substrate of autophagic degradation, was accumulated in both of the shRNA expressing lines (Fig. S3B), suggesting that an appearance of abnormal structure induced by ATP13A2 deficiency might be involved with impaired lysosomal proteolysis. Taken together, these findings indicated that loss of ATP13A2 led to lysosomal pathology and, more specifically, a reduction in cathepsin D activity.

### 2.3. Generation of an Atp13a2 mutant medaka fish

Next, we generated and evaluated medaka fish with an Atp13a2 mutation to investigate the mechanism of KRS/PARK9-associated neurodgeneration in vivo. The draft of the medaka genome contains only one identifiable ortholog of the human *ATP13A2* gene. We cloned the medaka *atp13a2* gene by RT-PCR and RACE, and it encoded a protein consisting of 1159 amino acids. The amino-acid sequence showed 51.3% homology to human ATP13A2 protein (Fig. S4A). To characterize medaka *atp13a2* expression, we used in situ hybridization to visualize medaka *atp13a2* mRNA. No signal was observed with the sense RNA probe. However, the anti-sense RNA probe resulted in diffuse signals in the gray matter of medaka brain (Fig. 4A). The telencephalon and diencephalon that contain the putative striatum and many dopaminergic neurons, respectively, were intensely labeled by the anti-sense probe. The optic tectum was also intensely labeled, but the hindbrain and spinal cord were scarcely labeled (Fig. 4A).

We then used TILLING (Targeting Induced Local Lesions In Genomes) method to generate an Atp13a2 mutant fish [16]. We

Please cite this article in press as: Matsui, H., et al. ATP13A2 deficiency induces a decrease in cathepsin D activity, fingerprint-like inclusion body formation, and selective degeneration of dopaminergic neurons. FEBS Lett. (2013), http://dx.doi.org/10.1016/j.febslet.2013.02.046

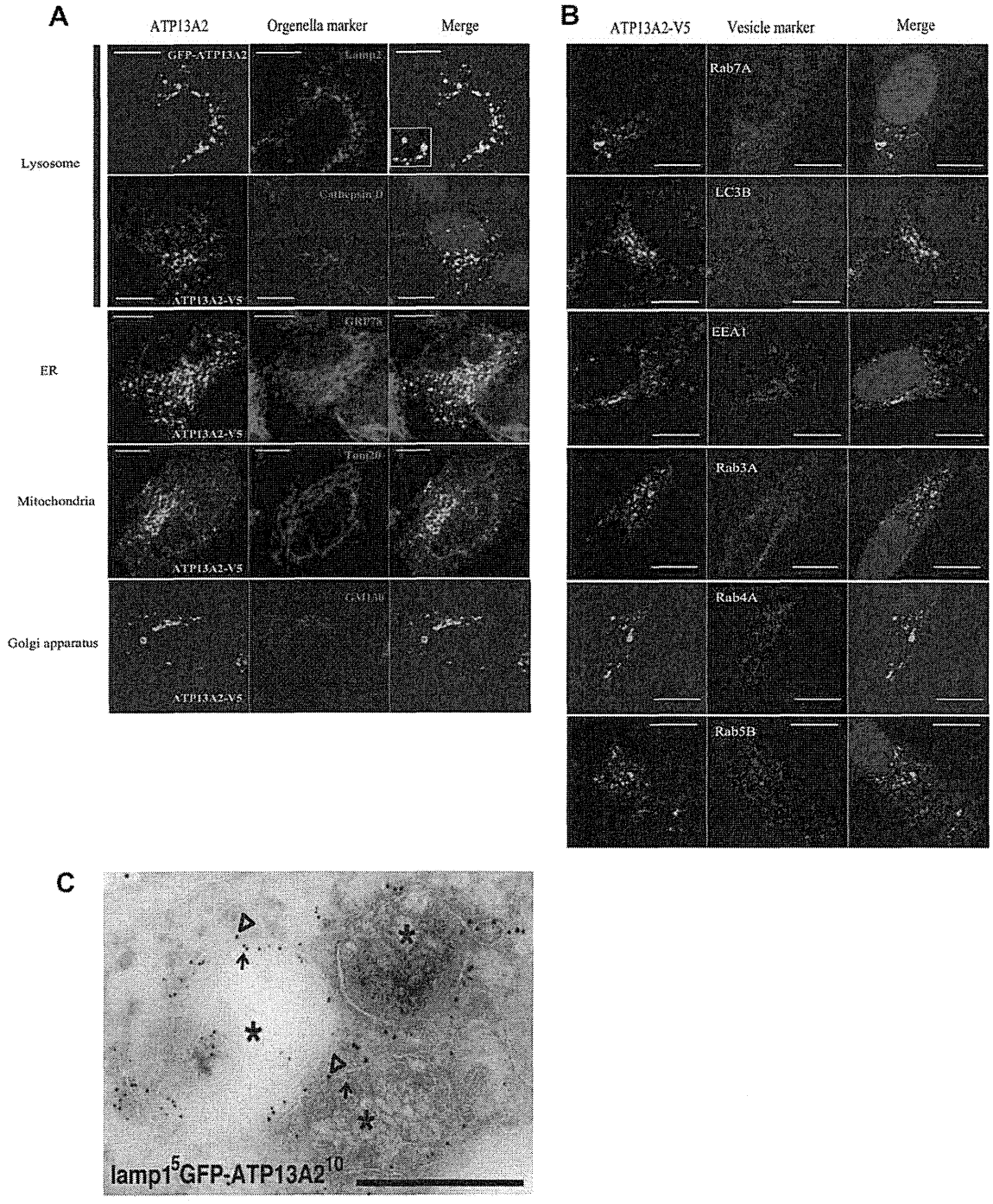



**Fig. 1.** Wild-type ATP13A2 localizes at lysosomal membranes. (A and B) WT-V5 and GFP-WT fusion proteins consistently co-localizes with Lamp2 and cathepsin D. Additionally, WT-V5 partially overlapped with GM130, Rab7 and LC3B. (Scale bar; 10 μm). (C) Immunoelectron microscopy using ultrathin cryosections. Double immunostaining of GFP-ATP13A2 (gold particles, 10 nm in diameter (open arrowheads)) and Lamp1 (gold particles, 5 nm in diameter (arrows)). Both types of immuno-gold label clearly localizes along the membranes around lysosomes (asterisks). (Scale bar; 0.5 μm).

sequenced the genomes of 5771 samples obtained from our ENU-mutagenized medaka library, and identified one mutation "IVS13, T-C, +2" that resulted in an aberrant splice donor site (Fig. 4B). The IVS13, T-C, +2 mutant from this strain was subjected to six sequential backcrosses to generate the mutant used in the following experiments. A cross between heterozygous "IVS13, T-C, +2" mutant pair resulted in wild-type fish (WT/WT), heterozygous mutants (WT/mt), and homozygous mutants (mt/mt) in Mendelian ratios. RT-PCR analysis revealed an abnormal splice variant in the WT/mt and mt/mt medaka (Fig. 4C), and the sequence of these PCR products indicated that exon 13 was skipped in the mutant mRNAs (Fig. 4D). Surprisingly, this abnormal splicing pattern was almost identical to that in the human KRS/PARK9 patient [3], in which the 111-bp exon 13 is skipped (Fig. S4B). Real-time PCR showed a marked reduction (17.8%) in the normal *atp13a2* mRNA in the mt/mt medaka brain (Fig. 4E). We therefore concluded that we had succeeded in identifying an Atp13a2 mutant in medaka, and this mutation was similar to a pathogenic KRS/PARK9 mutation in human.

Atp13a2 mutant medaka fish grew normally during early development without any obvious morphological abnormalities.

Please cite this article in press as: Matsui, H., et al. ATP13A2 deficiency induces a decrease in cathepsin D activity, fingerprint-like inclusion body formation, and selective degeneration of dopaminergic neurons. FEBS Lett. (2013), http://dx.doi.org/10.1016/j.febslet.2013.02.046

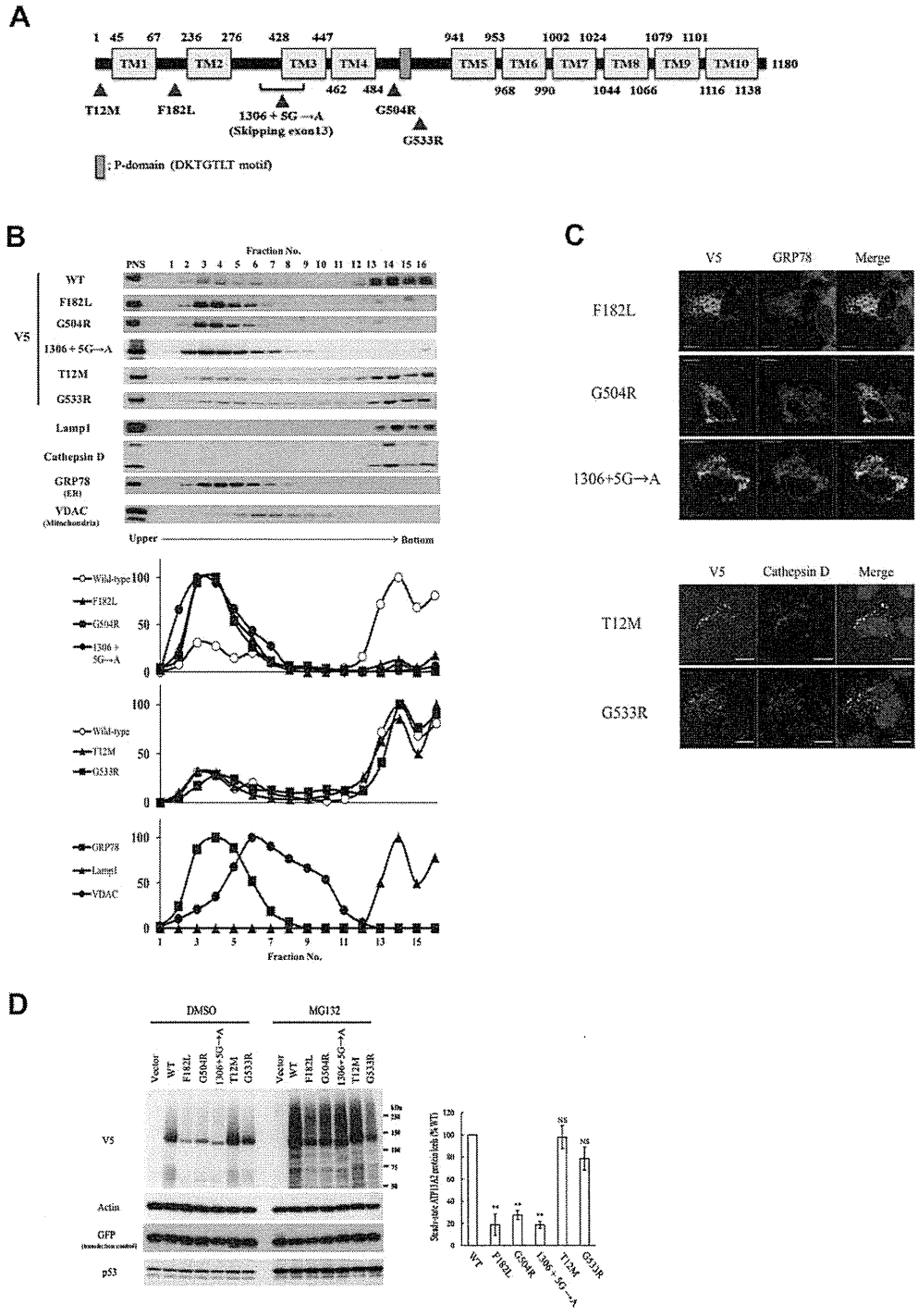



**Fig. 2.** Characterization of subcellular localization of KRS mutants. (A) Schematic diagram of disease-relevant mutants of ATP13A2 used in this study. (B) Percoll gradient fractionation of ATP13A2 protein. The graphs indicate densitometry of bands. This Percoll gradient system separates dense lysosomes (near bottom, Fraction 13–16) from lighter particles such as ER (Fraction 2–8). (C) Double immunofluorescence studies for KRS mutants. Colocalization of T12M or G533R with cathepsin D, a lysosomal protein, is observed. Other mutants localizes with GRP78, a resident ER chaperone protein. (D) Protein blot analysis of ATP13A2 WT and mutant V5-tagged constructs in transiently transfected SH-SY5Y cells. GFP was co-transfected with the ATP13A2s as a transfection control. The proteasome inhibitor MG132 (10 μM) stabilizes PARK9 mutants after 24 h. The antibody against p53, whose degradation is known to dependent to proteasome, is used as a control for MG132 treatment. Densitometry analysis indicates the steady state protein levels of each variant. Data are represented as percent of WT. Error bars, S.E.M. $n = 3$. $^{**}P < 0.01$ vs WT.

Remarkably, the mt/mt medaka fish showed a significant reduction of the life span relative to WT/WT and to WT/mt fish (Fig. 4F). The body weight of mt/mt fish was normal (Fig. 4G). We examined the internal organs including brains of the dead mt/mt fish, but we could not identify a specific reason for the shorter lifespan of mt/mt medaka fish. We next quantified spontaneous swimming movement in mt/mt medaka fish. At 4 months, mt/mt fish exhibited a mild locomotor increase with significant differences reported only in swimming duration whereas distance and velocity are normal. All the genotypes showed comparable movement at 12 months, irrespective of *atp13a2* genotype (Fig. 4H).

Collectively, we generated *atp13a2* mutant medaka fish carrying almost identical mutation to human KRS/PARK9 patient. The homozygous mutant fish grew normally, but relative to wild-type and heterozygous animals, these mutants exhibited more spontaneous swimming movement at 4 months and had a shorter life span.

Please cite this article in press as: Matsui, H., et al. ATP13A2 deficiency induces a decrease in cathepsin D activity, fingerprint-like inclusion body formation, and selective degeneration of dopaminergic neurons. FEBS Lett. (2013), http://dx.doi.org/10.1016/j.febslet.2013.02.046

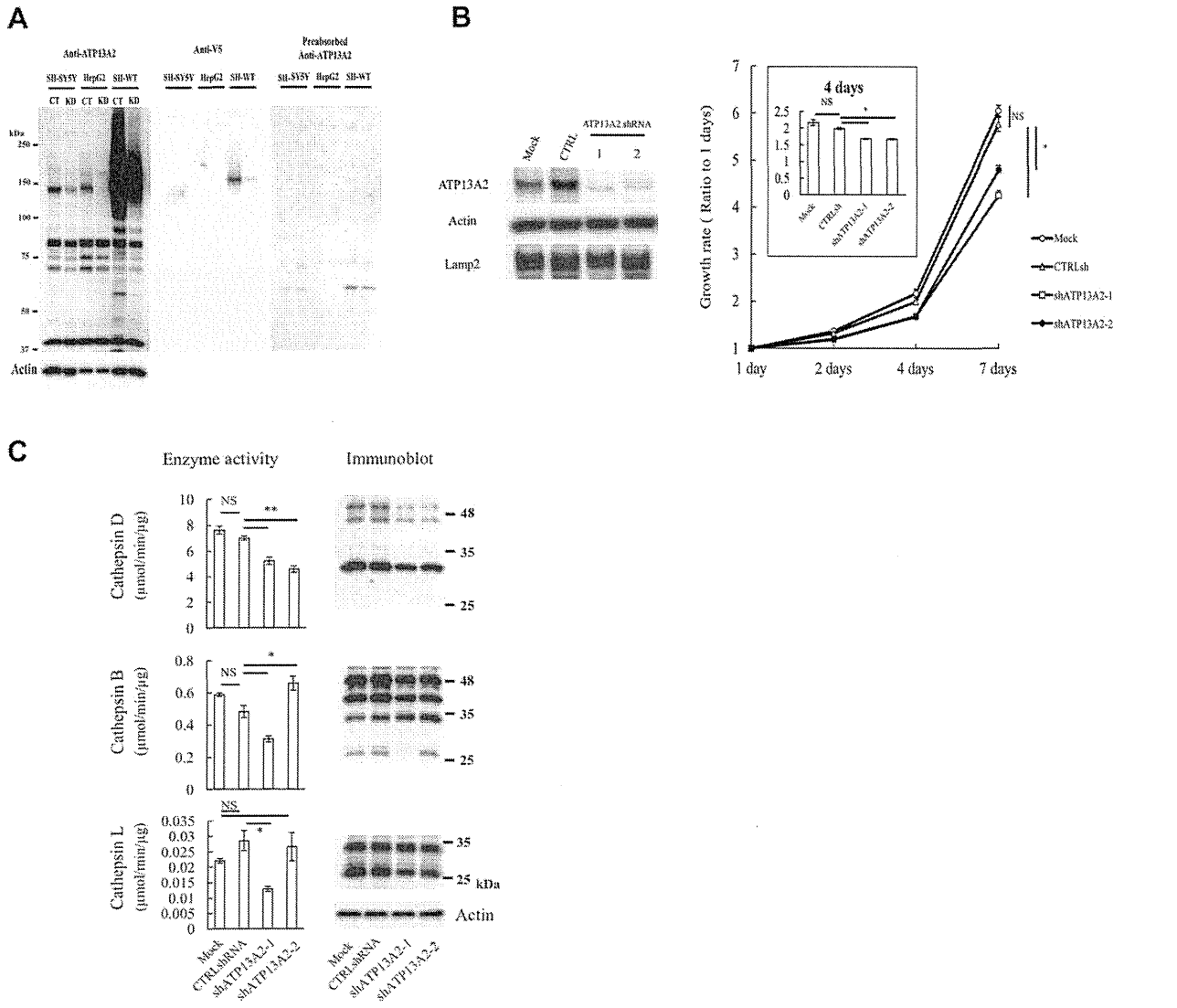
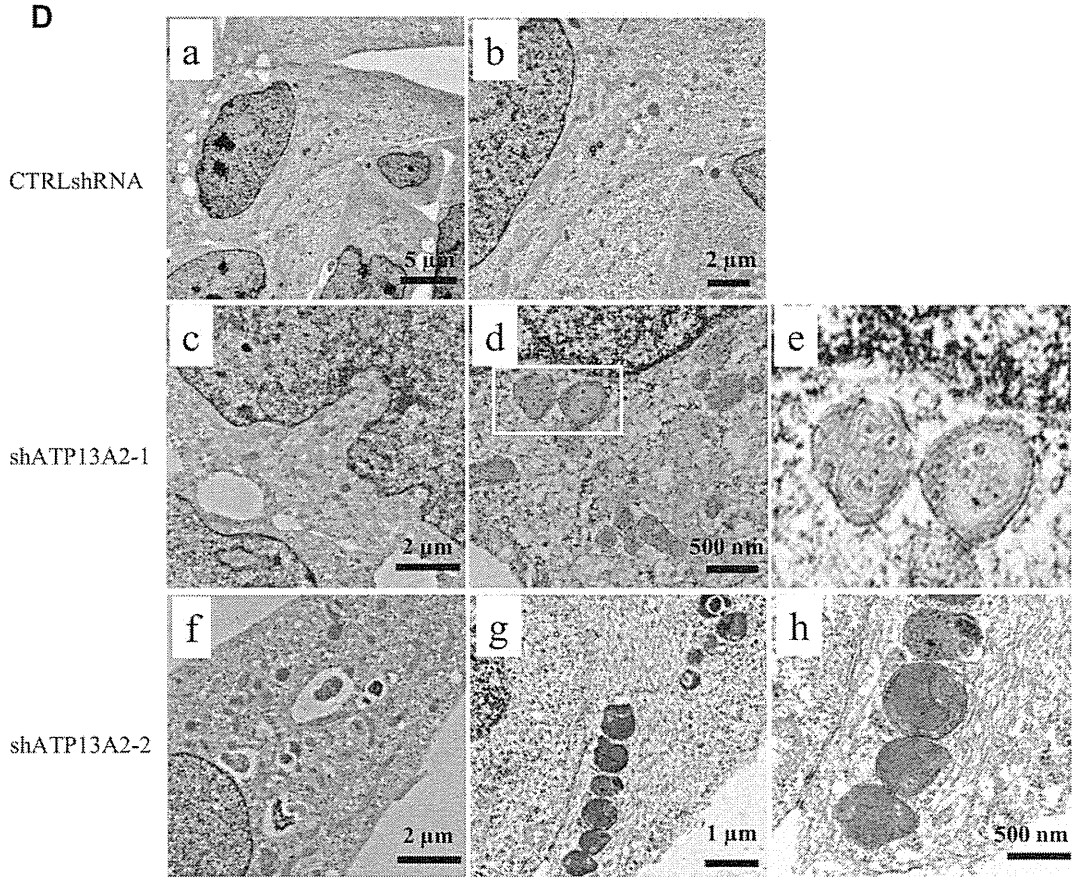



**Fig. 3.** Suppression of ATP13A2 leads to cathepsin D deficiency and accumulation of fingerprint-like structures in SH-SY5Y cells. (A) Immunoblotting with anti-ATP13A2 antibody (left panel), anti-V5 antibody (middle panel) and anti-ATP13A2 antibody (preabsorbed) (right panel) shows the ATP13A2 expression levels of SH-SY5Y cells, HepG2 cells and SH-SY5Y cells stably expressing WT (SH-WT). Cells are transfected with negative control (CT) or ATP13A2 siRNA (KD) for 72 h. (B) Immunoblotting using anti-ATP13A2 antibody and cell growth assay of SH-SY5Y cell lines stably expressing shRNA against human ATP13A2. The graph shows the growth rate of cells. Data are the means of triplicate experiments. Error bars, S.E.M. $^{*}P < 0.05$. (C) Measurement of cathepsin D, B and L activity and protein level in the extracts from ATP13A2-knockdown SH-SY5Y cells. Enzyme activity assays were performed in three independent experiments. Error bars, S.E.M. $^{*}P < 0.05$. $^{**}P < 0.01$. (D) Electron microscopic examination of SH-SY5Y cells that stably express shATP13A2-1 or shATP13A2-2. The diminished ATP13A2 expression induces lysosome-like bodies that contain granular deposits and fingerprint-like structure. Boxed areas (d) are shown enlarged in the left (e). (E) Evaluation of LC3 accumulation using anti-LC3B antibody in SH-SY5Y cell lines stably expressing shATP13A2 (each 3 clones). $^{**}P < 0.01$.

### 2.4. Neuropathology of *Atp13a2* mutant medaka fish

Selective and progressive loss of dopaminergic/noradrenergic cells constitutes the characteristic pathology of human PD patients. Having previously identified TH-positive (TH+) dopaminergic neurons and noradrenergic neurons in the medaka brain [17], we could examine histologically these TH+ neurons in the brain tissue of *atp13a2* mutants. At 4 months, the number of TH+ neurons did not differ significantly among mt/mt, mt/WT, and WT/WT fish. However at 8 and 12 months, the number of TH+ neurons in the middle diencephalon and the density of TH+ fibers in the telencephalon were lower in the mt/mt medaka than in mt/WT or in WT/WT fish (Fig. 5A). The mt/mt medaka fish at these stages also had fewer noradrenergic neurons in the medulla oblongata than did mt/WT or WT/WT fish (Fig. 5A). The reduction of TH+ neurons was not robust but age-dependent and progressive. Additionally, we examined tryptophan hydroxylase and serotonin levels via immunohistochemistry; neither the number of tryptophan hydroxylase positive neurons in the raphe nor the intensity of serotonin signals in the diencephalon differed significantly among all the genotypes (Fig. 5A). Although no TUNEL-positive dopaminergic neuron was observed in the WT/WT brains, a few TUNEL/TH double positive neurons were detected in the mt/mt brains (Fig. 5B). To exclude the possibility of developmental disorder of the dopaminergic neurons, we also counted the TH+ neurons in the middle diencephalon at 1 month. At this larval stage, mt/mt fish showed comparable number of dopaminergic neurons to WT/WT fish (Fig. S5) indicating the loss of dopaminergic neurons seen at 8 and 12 months was indeed late-onset phenotype. Next, we measured the amount of dopamine, noradrenaline, and serotonin in whole-brain samples from mutant fish at 4 and 12 months. The amount of dopamine in the brain samples from mt/mt medaka was comparable to that in the

Please cite this article in press as: Matsui, H., et al. ATP13A2 deficiency induces a decrease in cathepsin D activity, fingerprint-like inclusion body formation, and selective degeneration of dopaminergic neurons. FEBS Lett. (2013), http://dx.doi.org/10.1016/j.febslet.2013.02.046



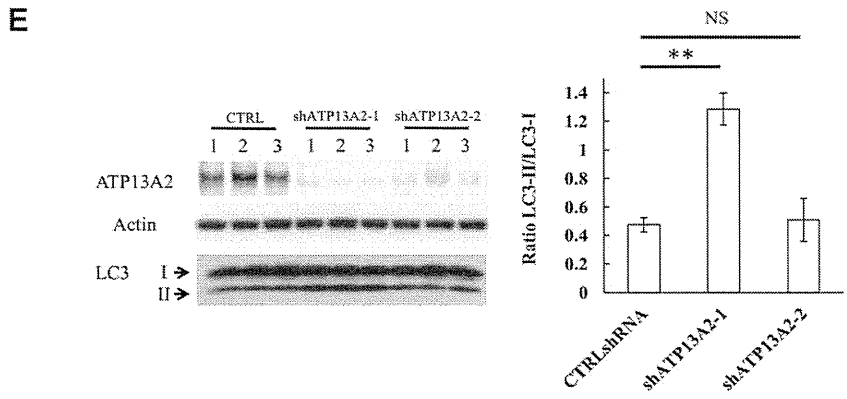




**Fig. 3.** (*continued*)

WT/WT or the WT/mt at 4 months, but lower than that in the WT/WT or the WT/mt at 12 months (Fig. 5C, upper). The noradrenaline in mt/mt medaka brain samples also tended to be lower at 12 months, but the differences were not statistically significant (Fig. 5C, middle). The amount of serotonin in whole-brain samples did not differ significantly among the genotypes (Fig. 5C, lower). Taken together, these findings indicated that *atp13a2* homozygous mutant medaka fish showed selective and progressive loss of dopaminergic/noradrenergic neurons, and such loss is a typical feature of human PD.

To examine in further detail the pathology associated with the *apt13a2* mutation, middle diencephalon samples from mt/mt, mt/WT, and WT/WT animals were analyzed via transmission electron microscopy as described previously [18]. Structures resembling fingerprints-profile, like those seen in ATP13A2-knockdown SH-SY5Y cells, were observed in thin sections taken from each mt/mt brain examined (Fig. 5D). However, these structures were not observed in the sections taken from WT/WT or WT/mt medaka brain samples. Fingerprint-profiles have been observed in cathepsin D-deficient mice [15] and in human patients with neuronal ceroid lipofuscinosis [19–21], and these structures are thought to indicate an autophagy/lysosome disorder. We used western blots to measure the amount of cathepsin D protein in brain tissue samples, and we found that mt/mt fish had less cathepsin D protein than did mt/WT or WT/WT fish (Fig. 5E and S6). We also showed that mt/mt medaka brain tissue, like the ATP13A2-knockdown cells, exhibited a significant reduction in cathepsin D activity (Fig. 5E). However, cathepsin K activity, cathepsin H activity, and proteasome activity were not affected by the *atp13a2* mutation (Fig. 5E), indicating the dysfunction of lysosomal enzymes was relatively specific to cathepsin D. Alpha-synuclein accumulation is one of the specific characters of idiopathic PD patients, Thus, we

Please cite this article in press as: Matsui, H., et al. ATP13A2 deficiency induces a decrease in cathepsin D activity, fingerprint-like inclusion body formation, and selective degeneration of dopaminergic neurons. FEBS Lett. (2013), http://dx.doi.org/10.1016/j.febslet.2013.02.046

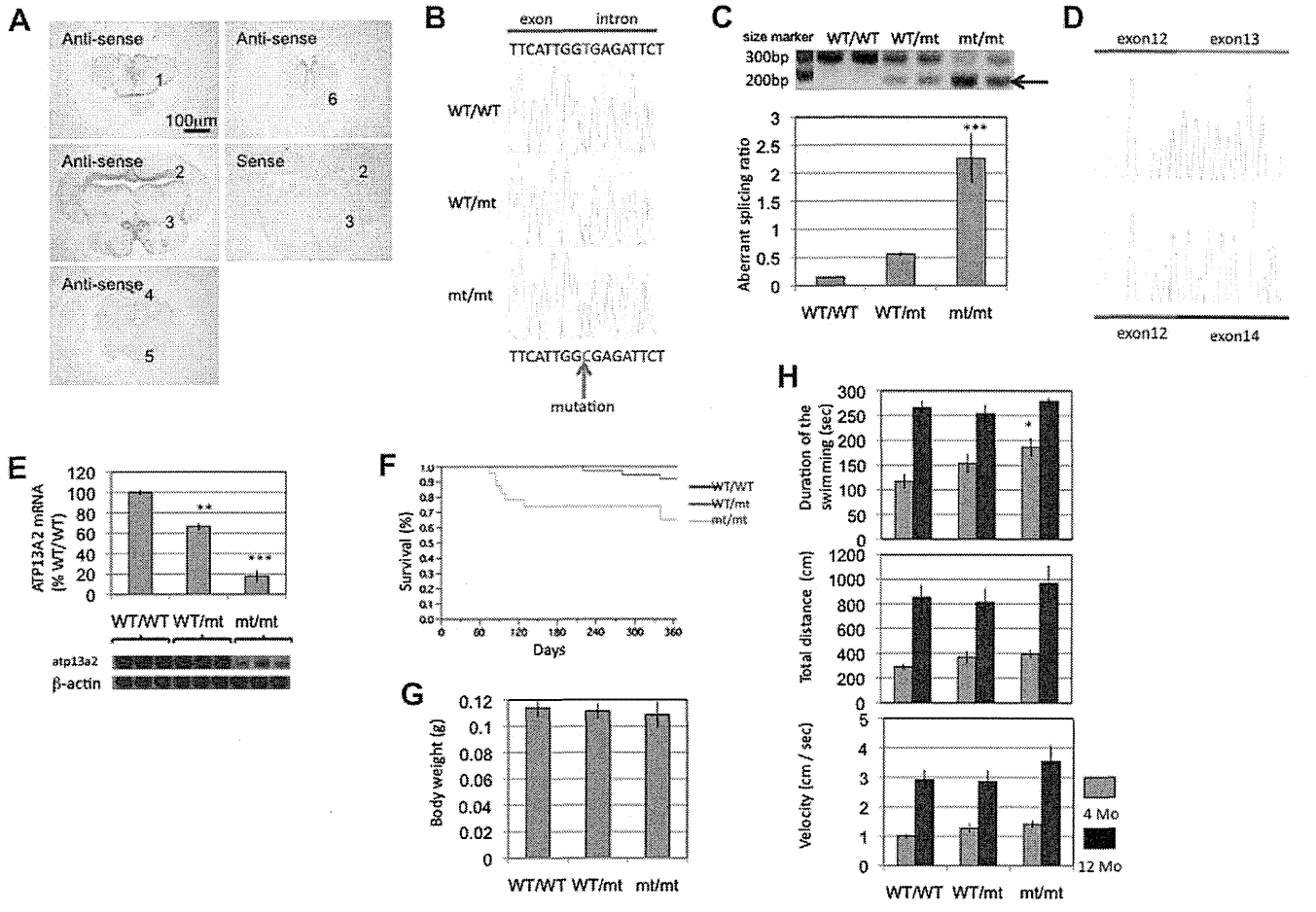



**Fig. 4.** Generation of Atp13a2 mutant medaka. (A) In situ hybridization of medaka *atp13a2* mRNA. Anti-sense signals and sense control of *Kyoto-Cab* medaka brain (12 months). 1: telencephalon, 2: optic tectum, 3: diencephalon, 4: cerebellum, 5: medulla oblongata, 6: spinal cord. (B) Sequence data for each genotype. A = green, T = red, G = black, and C = blue. The red arrow indicates the mutation site. This T to C mutation in the genomic sequence disrupts a splice donor site. (C) RT-PCR amplification of *atp13a2* mRNA from each genotype. WT/WT medaka show single band, whereas WT/mt and mt/mt medaka have an additional shorter band (arrow). bp: Base pairs. The graph indicates densitometric ratio of the shorter product/normal product. ***$P < 0.001$ vs. WT/WT and $P < 0.01$ vs. WT/mt. $n = 4$ for each genotype. Error bars, S.E.M. (D) Sequence of the PCR products. The upper band indicates the normal splicing product and the lower band is the abnormal splicing product. Exon 13 skipping occurrs in the Atp13a2 mutant medaka. (E) Real-time PCR of normal *atp13a2* mRNA. **: $P < 0.01$ vs. WT/WT, ***: $P < 0.001$ vs. WT/WT and WT/mt. $n = 3$ for each genotype. Error bars, S.E.M. (F) Survival curves of each genotype. The end point is the death of each medaka or day 365. The results show mild but significant shortening of the life span in mt/mt medaka ($P < 0.001$) ($n = 23$), relative to that in WT/WT ($n = 25$) or WT/mt ($n = 37$). Death before 1 month stage was not counted. (G) Body weight of Atp13a2 mutant medaka at 12 months. No significant differences were seen ($n = 20$ for each group). Error bars, S.E.M. (H) Duration of swimming, total swimming distance and swimming velocity during spontaneous swimming behavior of Atp13a2 mutant medaka ($n = 15$ for each group). *$P < 0.05$ vs. WT/WT. Error bars, S.E.M.

analyzed the alpha-synuclein status in our cell line and medaka models. However, we could not demonstrate consistent and significant differences between ATP13A2-deficient models and controls (Figs. S3 and S7), and these findings indicated that alpha-synuclein accumulation might not be the causative roles of KRS/PARK9.

In sum, *atp13a2* homozygous mutant medaka exhibited dopaminergic neurodegeneration, a deficiency of cathepsin D, and abnormal lysosome-related structures in the brain.

## 3. Discussion

Findings from previous studies clearly indicate that PARK gene products associate with each other via protein degradation pathways including the autophagy–lysosome system. Indeed, dysfunction of protein degradation has emerged as an important contributor to nigral neuronal death in PD. Presence of Lewy bodies is strong evidence of impaired protein degradation in PD. Lewy bodies consist of aggregated proteins, and alpha-synuclein is a major component of these structures [22]. Thus aggregation of alpha-synuclein has emerged as one of the most important processes in nigral degeneration in PD. Although soluble alpha-synuclein is degraded both during autophagy and by the proteasome, aggregated alpha-synuclein is degraded and cleared mainly via the autophagy–lysosome pathway [23]. Other PARK gene products, specifically Parkin and PINK1, work together to clear damaged mitochondria from cells via mitochondria-specific autophagy called mitophagy [24]. Furthermore, ATP13A2 mainly localizes to lysosomes, as we and other groups demonstrated [3,10–14,25]. Recently, mutations in the gene encoding glucocerebrosidase (GBA), a lysosomal enzyme, have been shown to be significant genetic risk factors for PD [26]. These observations indicate that lysosomal function is important for the maintenance of dopamine neurons, and they led us to investigate the function of ATP13A2 in the pathophysiology of KRS/PARK9.

ATP13A2 deficiency resulted in an abnormal aggregation of lysosomes at perinuclear site. Furthermore, these accumulated vesicles were enlarged, as previously reported [27]. Moreover, we found evidence of lysosomal dysfunction in that cathepsin D activity was, specifically, reduced in ATP13A2-knockdown cells. Cathepsin D is a ubiquitously expressed lysosomal protease that is involved in proteolytic degradation, cell invasion, and apoptosis. Cathepsin D deficiencies cause neuronal ceroid lipofuscinosis, a

Please cite this article in press as: Matsui, H., et al. ATP13A2 deficiency induces a decrease in cathepsin D activity, fingerprint-like inclusion body formation, and selective degeneration of dopaminergic neurons. FEBS Lett. (2013), http://dx.doi.org/10.1016/j.febslet.2013.02.046

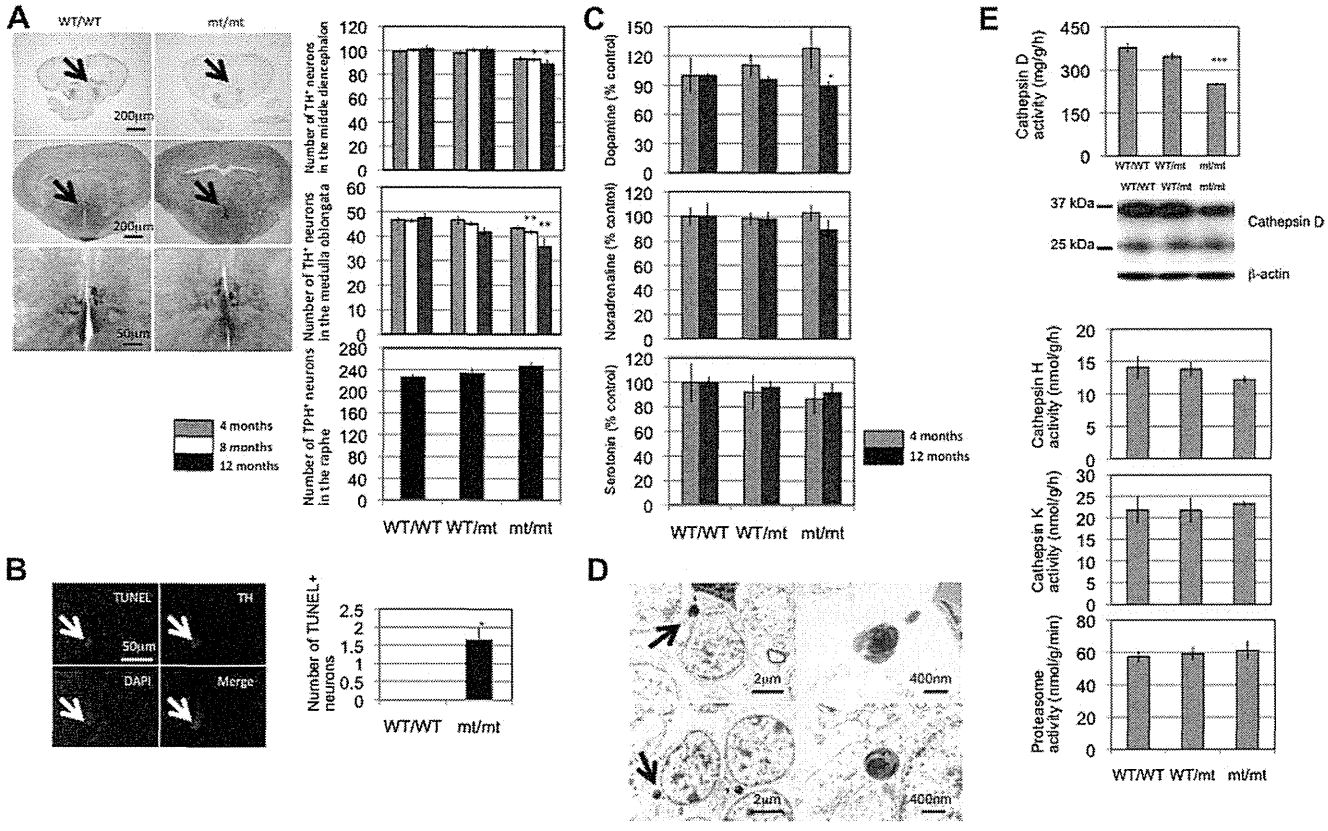



**Fig. 5.** Neuropathology of Atp13a2 mutant medaka. (A) Axial sections of telencephalon (upper) and middle diencephalon (middle and lower) of medaka brain at 12 months. Arrows indicate TH+ fibers and neurons. The lower figures are the enlarged image of middle figures. The graphs indicate the number of TH+ neurons in the middle diencephalon (upper) and medulla oblongata (middle) and the number of tryptophan-hydroxylase positive (TPH+) neurons in the raphe. $^{*}P < 0.05$ vs. WT/WT and WT/mt. $^{**}P < 0.01$ vs. WT/WT. ($n = 16$ for each group). Error bars, SEM. (B) TUNEL assay in medaka brain at 12 months. White arrow indicates one TUNEL/TH double positive neuron in the middle diencephalon of mt/mt fish. The graph shows the number of TH/TUNEL double positive neurons in the middle diencephalon ($n = 3$). $^{*}P < 0.05$ vs. WT/WT. (C) Amount of dopamine (upper), noradrenaline (middle), and serotonin (lower) in the brain of Atp13a2 mutant medaka. All values are expressed as a percentage of the amount (ng) per protein weight (mg) for WT/WT ($n = 8$ for each group). $^{*}P < 0.05$ vs. WT/WT. Error bars, S.E.M. (D) Fingerprint-like structures in Atp13a2 mutant medaka brain. Arrows indicate fingerprint-like structures in mt/mt brain. The right figures are the high magnification images of these structures. (E) Enzyme activity (Cathepsin D, H, K and proteasome activity) in the medaka brain. $^{***}P < 0.001$ vs. WT/T and WT/mt. Error bars, SEM. Image of a western blot shows cathepsin D protein in the medaka brain. Cross reactivity of the antibody against medaka cathepsin D is shown in the supplementary information (Fig. S5).

fatal neurogenerative disease in human and sheep [21,28–30]. Using electronmicroscopy, we also detected lysosome like body and granular deposits. Interestingly, we observed subcellular structures that resemble fingerprint-profile, and these structures resemble abnormal structures in the neurons of cathepsin D-deficient mice [15] and in human patients with neuronal ceroid lipofuscinosis [19–21] or with sphingolipidoses [31,32]. These results indicate that the primary cause of KRS/PARK9 is a lysosomal dysfunction and KRS/PARK9 could also be classified into a "lysosome disease".

We have recently used medaka fish to develop an animal model of PD [17,18,33,34]. Here, we found a mutation in our TILLING library that is almost identical to a PD-associated mutation in human patients. This medaka mutation results in the same abnormal splicing that is seen in the human patients with KRS/PARK9. Homozygous mutant fish exhibited selective loss of dopaminergic and noradrenergic neurons; this type of neuron loss is a pathology typically seen in human PD patients. Additionally, we found that tissues and cells in the brains from homozygous mutant medaka exhibited a specific reduction of cathepsin D protein and developed fingerprint-like subcellular structures. Both findings strongly indicate that the ATP13A2 mutation could lead to the dysfunction of lysosomes in medaka neurons.

Recently, Fonseca et al. injected Morpholinos against *atp13a2* into zebrafish embryo and showed that loss of Atp13a2 results in embryonic lethality [35]. As they also showed in zebrafish, *atp13a2* mRNA expressed not only in the brain but also in the entire body in medaka larvae (data not shown). This suggested that Atp13a2 is also important for some unknown function in other organs than the central nervous system. Our medaka model mimics the human mutation and showed pronounced reduction of *atp13a2* mRNA but not null expression. This might be helpful to study the long-term effect of Atp13a2 dysfunction.

As is the case of human patients with PD, cell death was specific to dopamine and noradrenaline neurons in our medaka model of KRS/ATP13A2. This cell-type specificity is also evident with our other medaka models of PD, including the models resulting from a lysosome inhibitor treatment [18,34]. The *atp13a2* mRNA, like other PD-related mRNAs, is expressed ubiquitously in the medaka brain [33]. Thus, the expression pattern of *atp13a2* could not explain the selective cell death in our model. Dopamine neurons contain toxic proteins derived from dopamine itself [36,37], and lysosomal function is essential in these neurons for preventing accumulation of the toxic proteins and other toxic metabolic products. Therefore, we speculate that dopamine neurons are especially vulnerable to lysosome dysfunction.

One negative finding is that mutant fish did not show slow locomotive movement, as do human PD patients. Based on our analysis, it seemed that the mutant fish swam the same amount, or more, than did control fish. In humans, loss of at least 80% of the dopaminergic neurons in substantia nigra seems necessary to evoke clear PD symptoms. The extent of the loss of dopaminergic

Please cite this article in press as: Matsui, H., et al. ATP13A2 deficiency induces a decrease in cathepsin D activity, fingerprint-like inclusion body formation, and selective degeneration of dopaminergic neurons. FEBS Lett. (2013), http://dx.doi.org/10.1016/j.febslet.2013.02.046

neurons in our medaka model might not be enough to evoke locomotive impairment. The mild increase of locomotion at 4 months seen in homozygous mutant might have a relation with the non-significant increase of dopamine at the same stage. Similar tentative increase of dopamine at younger stage is also observed in another PD model fish [33]. Such increased dopamine might harm the neurons, because the metabolism of dopamine is accompanied by the generation of oxidative radicals [38].

In conclusion, we demonstrated that reduction in ATP13A2 function in vitro or in vivo resulted in dysfunction of cathepsin D and the appearance of abnormal structures that are associated with lysosomal disorders. We used a teleost fish, medaka, to successfully generate an animal model suffered selective degeneration of dopaminergic neurons. Our findings indicate that lysosome-mediated autophagy may play a key role to protect dopaminergic neurons.

## Acknowledgments

We wish to thank Kondoh Differentiation Signaling Project, JST, for permission to use the Kyoto-cab strain. We are grateful to Ai Tanigaki, Rie Hikawa and Junji Ezaki, who were very supportive of our experiments. We are also grateful to Satoshi Fukui, Mitsutaka Yoshida, Kaori Moriya, and Hidetake Kurihara for excellent assistance with our electron microscopy studies. This work was supported by JST-CREST. A part of this research was supported by a Grant-in-Aid for Young Scientists (B) (F. Sato) and a Grant-in-Aid for Scientific Research on Innovative Areas (Comprehensive Brain Science Network) (F. Sato) from the Ministry of Education, Science, Sports and Culture of Japan.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.febslet.2013.02.046.

## References

[1] Gasser, T. (2007) Update on the genetics of Parkinson's disease. Mov. Disord. 22 (Suppl. 17), S343–S350.
[2] Najim al-Din, A.S., Wriekat, A., Mubaidin, A., Dasouki, M. and Hiari, M. (1994) Pallido-pyramidal degeneration, supranuclear upgaze paresis and dementia: Kufor-Rakeb syndrome. Acta Neurol. Scand. 89, 347–352.
[3] Ramirez, A., Heimbach, A., Gründemann, J., Stiller, B., Hampshire, D., Cid, L.P., Goebel, I., Mubaidin, A.F., Wriekat, A.L., Roeper, J., et al. (2006) Hereditary parkinsonism with dementia is caused by mutations in ATP13A2, encoding a lysosomal type 5 P-type ATPase. Nat. Genet. 38, 1184–1191.
[4] Di Fonzo, A., Chien, H.F., Socal, M., Giraudo, S., Tassorelli, C., Iliceto, G., Fabbrini, G., Marconi, R., Fincati, E., Abbruzzese, G., et al. (2007) ATP13A2 missense mutations in juvenile parkinsonism and young onset Parkinson disease. Neurology 68, 1557–1562.
[5] Ning, Y.P., Kanai, K., Tomiyama, H., Li, Y., Funayama, M., Yoshino, H., Sato, S., Asahina, M., Kuwabara, S., Takeda, A., et al. (2008) PARK9-linked parkinsonism in eastern Asia: mutation detection in ATP13A2 and clinical phenotype. Neurology 70, 1491–1493.
[6] Paisán-Ruiz, C., Guevara, R., Federoff, M., Hanagasi, H., Sina, F., Elahi, E., Schneider, S.A., Schwingenschuh, P., Bajaj, N., Emre, M., et al. (2010) Early-onset L-dopa-responsive parkinsonism with pyramidal signs due to ATP13A2, PLA2G6, FBXO7 and spatacsin mutations. Mov. Disord. 25, 1791–1800.
[7] Santoro, L., Breedveld, G.J., Manganelli, F., Iodice, R., Pisciotta, C., et al. (2011) Novel ATP13A2 (PARK9) homozygous mutation in a family with marked phenotype variability. Neurogenetics 12, 33–39.
[8] Crosiers, D., Ceulemans, B., Meeus, B., Nuytemans, K., Pals, P., Van Broeckhoven, C., Cras, P. and Theuns, J. (2011) Juvenile dystonia–parkinsonism and dementia caused by a novel ATP13A2 frame-shift mutation. Parkinsonism Relat. Disord. 17, 135–138.
[9] Lesage, S. and Brice, A. (2009) Parkinson's disease: from monogenic forms to genetic susceptibility factors. Hum. Mol. Genet. 18, R48–R59.
[10] Covy, J.P., Waxman, E.A. and Giasson, B.I. (2012) Characterization of cellular protective effects of ATP13A2/PARK9 expression and alterations resulting from pathogenic mutants. J. Neurosci. Res. 90, 2306–2316.
[11] Dehay, B., Ramirez, A., Martinez-Vicente, M., Perier, C., Canron, M.H., Doudnikoff, E., Vital, A., Vila, M., Klein, C. and Bezard, E.C. (2012) Loss of P-type ATPase ATP13A2/PARK9 function induces general lysosomal deficiency and leads to Parkinson disease neurodegeneration. Proc. Natl. Acad. Sci. USA 109, 9611–9616.
[12] Podhajska, A., Musso, A., Trancikova, A., Stafa, K., Moser, R., Sonnay, S., Glauser, L. and Moore, D.J. (2012) Common pathogenic effects of missense mutations in the P-type ATPase ATP13A2 (PARK9) associated with early-onset parkinsonism. PLoS ONE 7, e39942.
[13] Ramonet, D., Podhajska, A., Stafa, K., Sonnay, S., Trancikova, A., Tsika, E., Pletnikova, O., Troncoso, J.C., Glauser, L. and Moore, D.J. (2012) PARK9-associated ATP13A2 localizes to intracellular acidic vesicles and regulates cation homeostasis and neuronal integrity. Hum. Mol. Genet. 21, 1725–1743.
[14] Usenovic, M., Knight, A.L., Ray, A., Wong, V., Brown, K.R., Caldwell, G.A., Caldwell, K.A., Stagljar, I. and Krainc, D. (2012) Identification of novel ATP13A2 interactors and their role in α-synuclein misfolding and toxicity. Hum. Mol. Genet. 21, 3785–3794.
[15] Koike, M., Nakanishi, H., Saftig, P., Ezaki, J., Isahara, K., Ohsawa, Y., Schulz-Schaeffer, W., Watanabe, T., Waguri, S., Kametaka, S., et al. (2000) Cathepsin D deficiency induces lysosomal storage with ceroid lipofuscin in mouse CNS neurons. J. Neurosci. 20, 6898–6906.
[16] Taniguchi, Y., Takeda, S., Furutani-Seiki, M., Kamei, Y., Todo, T., Sasado, T., Deguchi, T., Kondoh, H., Mudde, J., Yamazoe, M., et al. (2006) Generation of medaka gene knockout models by target-selected mutagenesis. Genome Biol. 7, R116.
[17] Matsui, H., Taniguchi, Y., Inoue, H., Uemura, K., Takeda, S. and Takahashi, R. (2009) A chemical neurotoxin, MPTP induces Parkinson's disease like phenotype, movement disorders and persistent loss of dopamine neurons in medaka fish. Neurosci. Res. 65, 263–271.
[18] Matsui, H., Ito, H., Taniguchi, Y., Inoue, H., Takeda, S. and Takahashi, R. (2010) Proteasome inhibition in medaka brain induces the features of Parkinson's disease. J. Neurochem. 115, 178–187.
[19] Palmer, D.N., Fearnley, I.M., Walker, J.E., Hall, N.A., Lake, B.D., Wolfe, L.S., Haltia, M., Martinus, R.D. and Jolly, R.D. (1992) Mitochondrial ATP synthase subunit c storage in the ceroid-lipofuscinoses (Batten disease). Am. J. Med. Genet. 42, 561–567.
[20] Tsiakas, K., Steinfeld, R., Storch, S., Ezaki, J., Lukacs, Z., Kominami, E., Kohlschütter, A., Ullrich, K. and Braulke, T. (2004) Mutation of the glycosylated asparagine residue 286 in human CLN2 protein results in loss of enzymatic activity. Glycobiology 14, 1C–5C.
[21] Tyynelä, J., Palmer, D.N., Baumann, M. and Haltia, M. (1993) Storage of saposins A and D in infantile neuronal ceroid-lipofuscinosis. FEBS Lett. 330, 8–12.
[22] Spillantini, M.G., Schmidt, M.L., Lee, V.M., Trojanowski, J.Q., Jakes, R. and Goedert, M. (1997) Alpha-synuclein in Lewy bodies. Nature 388, 839–840.
[23] Webb, J.L., Ravikumar, B., Atkins, J., Skepper, J.N. and Rubinsztein, D.C. (2003) Alpha-synuclein is degraded by both autophagy and the proteasome. J. Biol. Chem. 278, 25009–25013.
[24] Narendra, D., Tanaka, A., Suen, D.F. and Youle, R.J. (2008) Parkin is recruited selectively to impaired mitochondria and promotes their autophagy. J. Cell Biol. 183, 795–803.
[25] Tan, J., Zhang, T., Jiang, L., Chi, J., Hu, D., Pan, Q., Wang, D. and Zhang, Z. (2011) Regulation of intracellular manganese homeostasis by Kufor-Rakeb syndrome-associated ATP13A2 protein. J. Biol. Chem. 286, 29654–29662.
[26] Sidransky, E., Nalls, M.A., Aasly, J.O., Aharon-Peretz, J., Annesi, G., Barbosa, E.R., Bar-Shira, A., Berg, D., Bras, J., Brice, A., et al. (2009) Multicenter analysis of glucocerebrosidase mutations in Parkinson's disease. N. Engl. J. Med. 361, 1651–1661.
[27] Usenovic, M., Tresse, E., Mazzulli, J.R., Taylor, J.P. and Krainc, D. (2012) Deficiency of ATP13A2 leads to lysosomal dysfunction, α-synuclein accumulation, and neurotoxicity. J. Neurosci. 32, 4240–4246.
[28] Siintola, E., Partanen, S., Strömme, P., Haapanen, A., Haltia, M., Maehlen, J., Lehesjoki, A.E. and Tyynelä, J. (2006) Cathepsin D deficiency underlies congenital human neuronal ceroid-lipofuscinosis. Brain 129, 1438–1445.
[29] Fritchie, K., Siintola, E., Armao, D., Lehesjoki, A.E., Marino, T., Powell, C., Tennison, M., Booker, J.M., Koch, S., Partanen, S., et al. (2009) Novel mutation and the first prenatal screening of cathepsin D deficiency (CLN10). Acta Neuropathol. 117, 201–208.
[30] Steinfeld, R., Reinhardt, K., Schreiber, K., Hillebrand, M., Kraetzner, R., Bruck, W., Saftig, P. and Gartner, J. (2006) Cathepsin D deficiency is associated with a human neurodegenerative disorder. Am. J. Hum. Genet. 78, 988–998.
[31] Jellinger, K., Anzil, A.P., Seemann, D. and Bernheimer, H. (1982) Adult GM2 gangliosidosis masquerading as slowly progressive muscular atrophy: motor neuron disease phenotype. Clin. Neuropathol. 1, 31–44.
[32] Idoate, M.A., Pardo-Mindan, F.J. and Gonzalez Alamillo, C. (1992) Fabry's disease without angiokeratomas showing unusual eccrine gland vacuolation. J. Pathol. 167, 65–68.
[33] Matsui, H., Taniguchi, Y., Inoue, H., Kobayashi, Y., Sakaki, Y., Toyoda, A., Uemura, K., Kobayashi, D., Takeda, S. and Takahashi, R. (2010) Loss of PINK1 in medaka fish (*Oryzias latipes*) causes late-onset decrease in spontaneous movement. Neurosci. Res. 66, 151–161.
[34] Matsui, H., Ito, H., Taniguchi, Y., Takeda, S. and Takahashi, R. (2010) Ammonium chloride and tunicamycin are novel toxins for dopaminergic neurons and induce Parkinson's disease-like phenotypes in medaka fish. J. Neurochem. 115, 1150–1160.

Please cite this article in press as: Matsui, H., et al. ATP13A2 deficiency induces a decrease in cathepsin D activity, fingerprint-like inclusion body formation, and selective degeneration of dopaminergic neurons. FEBS Lett. (2013), http://dx.doi.org/10.1016/j.febslet.2013.02.046

[35] Lopes da Fonseca, T., Correia, A., Hasselaar, W., van der Linde, H.C., Willemsen, R. and Outeiro, T.F. (2013) The zebrafish homologue of Parkinson's disease ATP13A2 is essential for embryonic survival. Brain Res. Bull. 90, 118–126.
[36] Barzilai, A., Daily, D., Zilkha-Falb, R., Ziv, I., Offen, D., Melamed, E. and Shirvan, A. (2003) The molecular mechanisms of dopamine toxicity. Adv. Neurol. 91, 73–82.
[37] Gandhi, S., Vaarmann, A., Yao, Z., Duchen, M.R., Wood, N.W. and Abramov, A.Y. (2012) Dopamine induced neurodegeneration in a PINK1 model of Parkinson's disease. PLoS ONE 7, e37564.
[38] Blum, D., Torch, S., Lambeng, N., Nissou, M., Benabid, A.L., Sadoul, R. and Verna, J.M. (2001) Molecular pathways involved in the neurotoxicity of 6-OHDA, dopamine and MPTP: contribution to the apoptotic theory in Parkinson's disease. Prog. Neurobiol. 65, 135–172.



Please cite this article in press as: Matsui, H., et al. ATP13A2 deficiency induces a decrease in cathepsin D activity, fingerprint-like inclusion body formation, and selective degeneration of dopaminergic neurons. FEBS Lett. (2013), http://dx.doi.org/10.1016/j.febslet.2013.02.046

JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

pubs.acs.org/JAFC

# Novel Isoflavone Glucosides in Groundnut (*Apios americana* Medik) and Their Antiandrogenic Activities

Marina Ichige,[†,∥] Erina Fukuda,[†,∥] Saki Miida,[†,∥] Jun-ichiro Hattan,[‡] Norihiko Misawa,[‡] Shun Saito,[§] Takahiro Fujimaki,[§] Masaya Imoto,[§] and Kazutoshi Shindo*[,†]

[†]Department of Food and Nutrition, Japan Women's University, 2-8-1, Mejirodai, Bunkyo-ku, Tokyo 112-8681, Japan

[‡]Research Institute for Bioresources and Biotechnology, Ishikawa Prefectural University, Suematsu, Nonoichi-shi, Ishikawa 921-8836, Japan

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

ABSTRACT: Isoflavone glucosides (2′-hydroxy,5-methoxy genistein-7-*O*-glucoside (**1**), 2′-hydroxy genistein-7-*O*-gentibioside (**2**), 5-methoxy genistein-7-*O*-glucoside (**3**), 3′,5-dimethoxy genistein-7-*O*-glucoside (**4**), 2′-hydroxy genistein-7-*O*-glucoside (**5**), genistein-7-*O*-gentibioside (**6**), 2′-hydroxy,5-methoxy genistein-4′,7-*O*-diglucoside (**7**), and 2′-hydroxy genistein-4′,7-*O*-diglucoside (**8**)) were isolated from the groundnut of *Apios americana* Medik. Their structures were elucidated on the basis of HR-ESI-MS and 1D- and 2D-NMR analyses. Compounds **1**, **2**, **4**, and **7** are new compounds presented here for the first time. Compounds **2** and **5** were proven to be androgen receptor antagonists due to their binding activities for androgen receptors ($IC_{50}$ 280 and 160 $\mu$M, respectively) and the inhibitory activity of androgen-induced expression of prostate-specific antigen (PSA) mRNA in LNCaP (prostate adenocarcinoma) cells ($IC_{50}$ 20 and 18 $\mu$M, respectively).

KEYWORDS: *Apios americana Medik, isoflavone glucosides, antiandrogenic activity, LNCaP human prostate adenocarcinoma cells*

## INTRODUCTION

Groundnut (*Apios americana* Medik) is a leguminous perennial vine native to North America that generates edible tubers, which were used as an important food by native Americans. This plant is called hodoimo or America-hodoimo in Japan, and its tubers have also been eaten in Japan, specifically in Aomori Prefecture, for more than one hundred years. This food is believed to be very nutritious, especially for women just after childbirth. Several studies have reported on the fatty acid,[1] amino acid,[2] and carbohydrate[3] compositions in groundnut, while other studies have reported on its secondary metabolites (saponin,[4] genistein,[5] and genistein glucosides[6]).

Since many isoflavones have been shown to possess interesting biological activities,[7−9] we analyzed the extract of the groundnut using PDA HPLC to obtain new bioactive isoflavones and found 7 isoflavone glucosides that have not been reported previously (Figure 1). In this study, we isolated all of these compounds using chromatographic methods and determined their structures (4 compounds were novel).

In addition, because some isoflavones such as genistein and daidzein were reported to show an estrogen agonist ability,[10] we tested the receptor binding activities of the isoflavone glucosides for the estrogen receptor (ER) and androgen receptor (AR) (androgen is structurally similar to estrogen). Two compounds were clarified to possess binding affinity for AR, while none of the compounds possessed binding affinity for ER. The activity of the 2 compounds for AR was proven to be antagonistic with an assay using LNCaP human prostate adenocarcinoma cells.

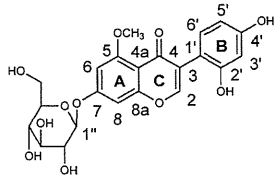
## MATERIALS AND METHODS

**Sample Materials.** Groundnut (*Apios americana* Medik) was purchased from e-yakusou.com (Niigata, Japan; http://www.e-yakusou.com/nae/apos.html) and planted in Ishikawa Prefectural Universiy in 2011, and the harvested tubers were used in this study. Gensitein, daidzein, and flutamide were purchased from Tokyo Chemical Industry (Tokyo, Japan).

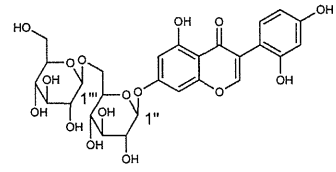
**Isolation of Isoflavone Glucosides from Groundnut.** Freeze-dried groundnuts (261.9 g) were pulverized in a mixer, and extracted with 2 L of MeOH by stirring for 30 min at room temperature (×2). The extract was filtered, and the filtrate was concentrated to dryness in vacuo to give a yellowish oil (1.02 g). This oil was added to 100 mL of hexane and sonicated for 5 min. The insoluble precipitate (1.02 g) was collected by centrifugation and applied on a silica gel column [30 i.d. × 200 mm, Silica Gel60 (Kanto Chemical Co. Inc., Tokyo, Japan)] prepared and developed with $CH_2Cl_2$:MeOH:$H_2O$ (3:1:0.1). Fractions 15−70 were collected and concentrated to dryness to give a pale yellowish oil (0.26 g). This oil was analyzed by RP-HPLC chromatography using Developsil C30 (20 i.d. × 250 mm; Nomura Chemical Co. Ltd., Aichi, Japan) developed with 17% $CH_3CN$ (flow rate: 7.0 mL/min). From this, pure compounds **1**, **2**, **3**, **4**, **5**, and **6** and crude compounds **7** and **8** were obtained. Crude compounds **7** and **8** were both further purified by ODS HPLC using CAPCELLPAK SG (10 i.d. × 250 mm; Shiseido, Tokyo, Japan) developed with 20% MeOH (flow rate: 3.0 mL/min) to give pure compounds.

**Spectroscopic Analysis.** NMR spectra were measured on AVANCE400 (Bruker BioSpin, Karlsruhe, Germany) in DMSO-$d_6$ with the residual solvent peak as an internal standard ($\delta_C$ 39.5, $\delta_H$ 2.50 ppm) for **1**, **3**, **4**, **5**, **7**, and **8** or $CD_3OD$ ($\delta_C$ 49.5, $\delta_H$ 3.30 ppm) for **2**,

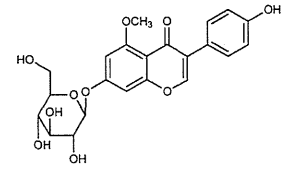
Received: December 7, 2012
Revised: February 6, 2013
Accepted: February 12, 2013
Published: February 12, 2013



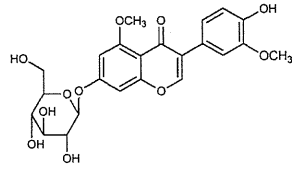
2'-hydroxy,5-methoxy genistein-7-O-glucoside (**1**)



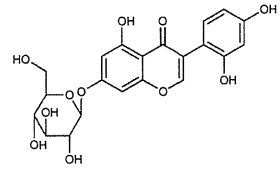
2'-hydroxy genistein-7-O-gentibioside (**2**)



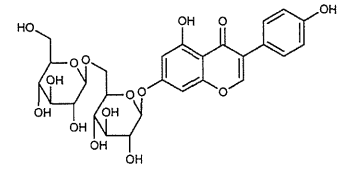
5-methoxy genistein-7-O-glucoside (**3**)



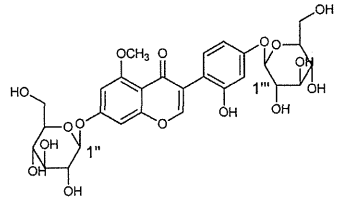
3',5-dimethoxy genistein-7-O-glucoside (**4**)



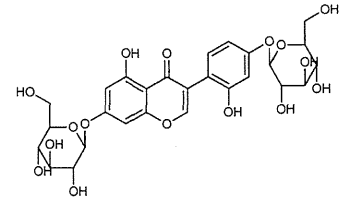
2'-hydroxy genistein-7-O-glucoside (**5**)



genistein-7-O-gentibioside (**6**)



2'-hydroxy,5-methoxy genistein-4',7-O-diglucoside (**7**)



2'-hydroxy genistein-4',7-O-diglucoside (**8**)

**Figure 1.** The structures of isoflavone glucosides.

pyridine-$d_5$ ($\delta_C$ 135.5, $\delta_H$ 7.58 ppm) for **6**, or $CDCl_3$ ($\delta_C$ 77.0, $\delta_H$ 7.25 ppm) for hexa-acetyl **1**. HR-ESI-MS was recorded with JMS-T100LP (Jeol, Tokyo, Japan) using reserpine as an external standard.

**Physico-Chemical Data for Compounds 1 – 8 and Hexa-acetyl 1.** *2′-Hydroxy,5-methoxy genistein-7-O-glucoside (1).* White solid; $[\alpha]_D^{22}$ −72.0° (*c* 0.1, 50%MeOH), UV(MeOH) $\lambda_{max}$ 256 nm. HR-ESI-MS (negative) *m/z* 461.10893 $[M - H]^-$ (calcd for $C_{22}H_{21}O_{11}$ 461.10839). $^1$H NMR (DMSO-$d_6$) $\delta$: 3.16 (m, 1H, H-4′′), 3.30 (m, 1H, H-2′′), 3.36 (m, 1H, H-3′′), 3.45 (2H, H-5′′ and H-6′′b), 3.72 (m, 1H, H-6′′a), 3.83 (s, 3H, H-5OCH$_3$), 5.08 (d, *J* = 7.2 Hz, 1H, H-1′′), 6.25 (dd, *J* = 2.0, 8.2 Hz, 1H, H-5′), 6.34 (d, *J* = 2.0 Hz, 1H, H-3′), 6.60 (d, *J* = 2.0 Hz, 1H, H-6), 6.72 (d, *J* = 2.0 Hz, 1H, H-8), 6.91 (d, *J* = 8.2 Hz, 1H, H-6′), 8.01 (s, 1H, H-2). $^{13}$C NMR (DMSO-$d_6$) $\delta$: 56.1 (C-5OCH$_3$), 60.7 (C-6′′), 69.8 (C-4′′), 73.1 (C-3′′), 76.6 (C-2′′), 77.3 (C-5′′), 95.6 (C-8), 97.1 (C-6), 99.1 (C-1′′), 102.8 (C-3′), 106.2 (C-5′), 109.5 (C-4a), 110.3 (C-1′), 123.3 (C-3), 132.1 (C-6′), 152.1 (C-2), 156.4 (C-2′), 158.3 (C-4′), 158.8 (C-8a), 160.6 (C-5), 161.3 (C-7), 174.4 (C-4).

*Hexacetyl 1.* White solid; HR-ESI-MS (negative) *m/z* 713.17195 $[M - H]^-$ (calcd. for $C_{34}H_{33}O_{17}$ 713.17177). $^1$H NMR ($CDCl_3$) $\delta$: 2.06 – 2.13 (CH$_3$COOH × 6), 3.92 (s, 3H, C-5OCH$_3$), 3.96 (ddd, *J* = 2.3, 5.6, 9.0 Hz, 1H, H-5′′), 4.22 (dd, *J* = 2.3, 12.3 Hz, 1H, H-6′′b), 4.28 (dd, *J* = 5.6, 12.3 Hz, 1H, H-6′′a), 5.18 (dd, *J* = 9.0, 9.3 Hz, 1H, H-4′′), 5.24 (d, *J* = 7.2 Hz, 1H, H-1′′), 5.31 (dd, *J* = 7.2, 9.0 Hz, 1H, H-2′′), 5.35 (dd, *J* = 9.0, 9.3 Hz, 1H, H-3′′), 6.45 (d, *J* = 2.0 Hz, 1H, H-6), 6.58 (d, *J* = 2.0 Hz, 1H, H-8), 7.03 (d, *J* = 2.2 Hz, 1H, H-3′), 7.06 (dd, *J* = 2.2, 8.3 Hz, H-5′), 7.35 (d, *J* = 8.3 Hz, H-6′), 7.70 (s, 1H, H-2). $^{13}$C NMR ($CDCl_3$) $\delta$: 20.6 – 20.9 (CH$_3$COOH × 6), 56.5 (C-5OCH$_3$), 62.0 (C-6′′), 68.1 (C-4′′), 71.0 (C-2′′), 72.5 (C-3′′), 72.5 (C-5′′), 95.8 (C-8), 97.6 (C-6), 98.0 (C-1′′), 111.1 (C-4a), 116.4 (C-3′), 119.0 (C-5′), 122.4 (C-1′), 123.0 (C-3), 132.1 (C-6′), 149.4 (C-2′), 151.1 (C-4′), 151.7 (C-2), 159.4 (C-8a), 160.6 (C-7), 161.7 (C-5), 168.8 – 170.4 (CH$_3$COOH × 6), 174.0 (C-4).

*2′-Hydroxy Genistein-7-O-gentibioside (2).* White solid; $[\alpha]_D^{22}$ −53.2° (*c* 0.5, MeOH), UV(MeOH) $\lambda_{max}$ 260 nm. HR-ESI-MS (negative) *m/z* 609.14474 $[M - H]^-$ (calcd for $C_{27}H_{29}O_{16}$ 609.14556). $^1$H NMR ($CD_3OD$) $\delta$: 3.23 (2H, H-2‴ and H-5‴), 3.27 (m, 1H, H-3‴), 3.33 (m, 1H, H-4′′), 3.47 (m, 1H, H-3′′), 3.49 (m, 1H, H-2′′), 3.67 (dd, *J* = 6.2, 11.8 Hz, 1H, H-6‴b), 3.84 (m, 1H, H-6′′b), 3.85 (m, 1H, H-5′′), 3.89 (d, *J* = 11.8 Hz, 1H, H-6‴a), 4.16 (d, *J* = 9.0 Hz, 1H, H-6′′a), 4.37 (d, *J* = 7.2 Hz, 1H, H-1‴), 5.05 (d, *J* = 7.2 Hz, 1H, H-1′′), 6.37 (dd, *J* = 2.0, 8.1 Hz, 1H, H-5′), 6.57 (d, *J* = 2.0 Hz, 1H, H-6), 6.57 (d, *J* = 2.0 Hz, 1H, H-3′), 6.87 (d, *J* = 2.0 Hz, H-8), 7.05 (d, *J* = 8.1 Hz, 1H, H-5′), 8.11 (s, 1H, H-2). $^{13}$C NMR ($CD_3OD$) $\delta$: 62.8 (C-6‴), 70.5 (C-6′′), 71.6 (C-4′′)[a], 71.7 (C-4‴)[a], 74.7 (C-3′′), 75.2 (C-3‴), 77.3 (C-5′′), 77.8 (C-2′′), 78.0 (C-5‴), 78.1 (C-2‴), 96.0 (C-8), 101.3 (C-6), 101.5 (C-1′′), 104.2 (C-3′), 105.2 (C-1‴), 108.1 (C-4a), 108.1 (C-5′), 110.6 (C-1′), 122.9 (C-3), 133.3 (C-6′), 157.3 (C-2), 157.8 (C-2′), 159.4 (C-8a), 160.3 (C-4′),163.3 (C-5), 164.7 (C-7), 182.9 (C-4). [a]: interchangeable.

*5-Methoxy Genistein-7-O-glucoside (3).* White solid; $[\alpha]_D^{22}$ −44.3° (*c* 0.5, MeOH), UV(MeOH) $\lambda_{max}$ 257 nm. HR-ESI-MS (negative) *m/z* 445.11528 $[M - H]^-$ (calcd for $C_{22}H_{21}O_{10}$ 445.11347). $^1$H NMR (DMSO-$d_6$) $\delta$: 3.16 (m, 1H, H-4′′), 3.30 (m, 1H, H-2′′), 3.32 (m, 1H, H-3′′), 3.45 (2H, H-5′′ and H-6′′b), 3.74 (m, 1H, H-6′′a), 3.84 (s, 3H, C-5OCH$_3$), 5.08 (d, *J* = 7.1 Hz, 1H, H-1′′), 6.60 (d, *J* = 2.0 Hz, 1H, H-6), 6.73 (d, *J* = 2.0 Hz, 1H, H-8), 6.79 (d, *J* = 8.4 Hz, 2H, H-3′ and H-5′), 7.31 (d, *J* = 8.4 Hz, 2H, H-2′ and H-6′), 8.16 (s, 1H, H-2). $^{13}$C NMR (DMSO-$d_6$) $\delta$: 56.1 (C-5OCH$_3$), 60.7 (C-6′′), 69.8 (C-4′′), 73.1 (C-2′′), 76.6 (C-3′′), 77.3 (C-5′′), 95.6 (C-8), 97.1 (C-6), 99.9 (C-1′′), 109.5 (C-4a), 114.8 (C-3′ and C-5′), 122.6 (C-1′), 124.9 (C-3), 131.5 (C-2′ and C-6′), 150.8 (C-2), 157.2 (C-4′), 158.8 (C-8a), 160.8 (C-5), 161.3 (C-7), 173.9 (C-4).

*3′,5-Dimethoxy Genistein-7-O-glucoside (4).* White solid; $[\alpha]_D^{22}$ −48.4° (*c* 0.5, MeOH), UV(MeOH) $\lambda_{max}$ 259 nm. HR-ESI-MS (negative) *m/z* 475.12413 $[M - H]^-$ (calcd for $C_{23}H_{23}O_{11}$ 475.12404). $^1$H NMR (DMSO-$d_6$) $\delta$: 3.16 (m, 1H, H-4′′), 3.30 (2H, H-2′′ and H-3′′), 3.45 (2H, H-5′′ and H-6′′b), 3.72 (m, 1H, H-6′′a), 3.78 (s, 3H, C-3′′OCH$_3$), 3.84 (s, 3H, C-5OCH$_3$), 5.08 (d, *J* = 7.2 Hz, 1H, H-1′′), 6.60 (s, 1H, H-6), 6.73 (s, 1H, H-8), 6.79 (d, *J* = 8.2 Hz, 1H, H-5′), 6.90 (d, *J* = 8.2 Hz, 1H, H-6′), 7.11 (s, 1H, H-2′), 8.19 (s, 1H, H-2). $^{13}$C NMR (DMSO-$d_6$) $\delta$: 55.6 (C-3′OCH$_3$), 56.1


dx.doi.org/10.1021/jf305233t | J. Agric. Food Chem. 2013, 61, 2183–2187

(C-5O$\underline{C}$H$_3$), 60.7 (C-6″), 69.7 (C-4″), 73.1 (C-3″), 76.6 (C-2″), 77.3 (C-5″), 95.6 (C-8), 97.1 (C-6), 99.8 (C-1″), 109.5 (C-4a), 113.4 (C-2′), 115.1 (C-5′), 121.6 (C-6′), 123.0 (C-1′), 124.9 (C-3), 146.4 (C-4′), 147.0 (C-3′), 151.1 (C-2), 158.7 (C-8a), 160.8 (C-5), 161.3 (C-7), 173.8 (C-4).

*2′-Hydroxy Genistein-7-O-glucoside (5).* White solid; $[\alpha]_D^{22}$ −58.8° (*c* 0.5, MeOH)UV(MeOH) $\lambda_{max}$ 260 nm. HR-ESI-MS (negative) *m/z* 447.09314 $[M - H]^-$ (calcd for $C_{21}H_{19}O_{11}$ 447.09274). $^1$H NMR (DMSO-$d_6$) $\delta$: 3.16 (m, 1H, H-4″), 3.29 (m, 1H, H-2″), 3.30 (m, 1H, H-3″), 3.44 (m, 1H, H-5″), 3.46 (m, 1H, H-6″b), 3.70 (m, 1H, H-6″a), 5.06 (d, *J* = 7.4 Hz, 1H, H-1″), 6.27 (d, *J* = 8.1 Hz, 1H, H-5′), 6.38 (s, 1H, H-3′), 6.46 (s, 1H, H-6), 6.70 (s, 1H, H-8), 6.99 (d, *J* = 8.1 Hz, 1H, H-6′), 8.26 (s, 1H, H-2). $^{13}$C NMR (DMSO-$d_6$) $\delta$: 60.6 (C-6″), 69.6 (C-4″), 73.1 (C-3″), 76.4 (C-2″), 77.2 (C-5″), 94.4 (C-8), 99.5 (C-6), 99.9 (C-1″), 102.6 (C-3′), 106.1 (C-4a), 106.2 (C-5′), 108.3 (C-1′), 120.8 (C-3), 132.2 (C-6′), 155.9 (C-2), 156.5 (C-2′), 157.2 (C-8a), 158.7 (C-4′), 161.6 (C-5), 162.9 (C-7), 180.7 (C-4).

*Genistein-7-O-gentibioside (6).* White solid; $[\alpha]_D^{22}$ −62.4° (*c* 0.5, MeOH)UV(MeOH) $\lambda_{max}$ 262 nm. HR-ESI-MS (negative) *m/z* 593.15152 $[M - H]^-$ (calcd. for $C_{27}H_{29}O_{15}$ 593.15064). $^1$H NMR (pyridine-$d_5$) $\delta$: 3.92 (m, 1H, H-5‴), 4.22 (m, 1H, H-4‴), 4.25 (2H, H-2‴ and H-3‴), 4.27 (2H, H-2″ and H-4″), 4.35 (m, 1H, H-3″), 4.36 (m, 1H, H-6‴b), 4.38 (m, 1H, H-6″b), 4.44 (m, 1H, H-5″), 4.56 (m, 1H, H-6‴a), 4.81 (m, 1H, H-6″a), 5.08 (d, *J* = 7.8 Hz, 1H, H-1‴), 1H, 5.75 (d, *J* = 7.3 Hz, 1H, H-1″), 6.87 (d, *J* = 1.9 Hz, 1H, H-6), 7.09 (d, *J* = 1.9 Hz, 1H, H-8), 7.26 (d, *J* = 8.4 Hz, 2H, H-3′ and H-5′), 7.63 (d, *J* = 8.4 Hz, 2H, H-2′ and H-6′), 8.08 (s, 1H, H-2). $^{13}$C NMR (pyridine-$d_5$) $\delta$: 62.1 (C-6‴), 69.6 (C-6″), 70.5 (C-4″), 71.2 (C-4‴), 74.1 (C-2″), 74.7 (C-2‴), 76.9 (C-5″), 77.7 (C-3″)[a], 77.8 (C-3‴)[a], 77.9 (C-5‴), 94.7 (C-8), 101.1 (C-6), 101.3 (C-1″), 105.1 (C-1‴), 106.8 (C-4a), 115.6 (C-3′ and C-5′), 121.4 (C-1′), 123.4 (C-3), 130.4 (C-2′ and C-6′), 153.2 (C-2), 157.5 (C-4′), 158.6 (C-8a), 162.4 (C-5), 163.4 (C-7), 180.7 (C-4). [a]: interchangeable.

*2′-Hydoxy,5-methoxy Genistein-4′,7-O-diglucoside (7).* White solid; $[\alpha]_D^{22}$ −40.6° (*c* 0.1, 50%MeOH), UV(MeOH) $\lambda_{max}$ 256 nm. HR-ESI-MS (negative) *m/z* 623.16312 $[M - H]^-$ (calcd for $C_{28}H_{31}O_{16}$ 623.16121). $^1$H NMR (DMSO-$d_6$) $\delta$: 3.18 (m, 1H, H-4″), 3.22 (m, 1H, H-2‴), 3.24 (m, 1H, H-3‴), 3.28 (m, 1H, H-2″), 3.29 (m, 1H, H-4‴), 3.31 (m, 1H, H-3″), 3.45 (4H, H-5″, H-5‴, H-6″b, and H-6‴b), 3.72 (2H, H-6″a and H-6‴a), 3.83 (s, 3H, C-5OC$\underline{H}_3$), 4.82 (d, *J* = 7.5 Hz, 1H, H-1‴), 5.08 (d, *J* = 7.1 Hz, 1H, H-1″), 6.62 (d, 1H, *J* = 8.5 Hz, H-5′), 6.53 (s, 1H, H-3′), 6.61 (s, 1H, H-6), 6.74 (s, 1H, H-8), 7.03 (d, 1H, *J* = 8.5 Hz, H-6′), 8.04 (s, 1H, H-2). $^{13}$C NMR (DMSO-$d_6$) $\delta$: 56.1 (C-5O$\underline{C}$H$_3$), 60.6 (C-6‴)[a], 60.7 (C-6″)[a], 69.6 (C-4‴)[b], 69.8 (C-4″)[b], 73.1 (C-3‴)[c], 73.2 (C-3″)[c], 76.6 (C-2″), 76.6 (C-2‴), 77.0 (C-5″)[d], 77.3 (C-5‴)[d], 95.6 (C-8), 97.2 (C-6), 99.9 (C-1″), 100.4 (C-1‴), 104.0 (C-3′), 106.6 (C-5′), 109.5 (C-4a), 113.3 (C-1′), 123.0 (C-3), 132.0 (C-6′), 152.3 (C-2), 156.4 (C-2′), 158.2 (C-4′), 158.8 (C-8a), 160.6 (C-5), 161.3 (C-7), 174.1 (C-4). [a, b, c, d]: interchangeable.

*2′-Hydroxy Genistein-4′,7-O-diglucoside (8).* White solid; $[\alpha]_D^{22}$ −61.6° (*c* 0.1, 50%MeOH), UV(MeOH) $\lambda_{max}$ 256 nm. HR-ESI-MS (negative) *m/z* 609.14469 $[M - H]^-$ (calcd for $C_{27}H_{29}O_{16}$ 609.14556). $^1$H NMR (DMSO-$d_6$) $\delta$: 3.18 (m, 1H, H-4″), 3.23 (m, 1H, H-2‴), 3.34 (m, 1H, H-3‴), 3.26 (m, 1H, H-2″), 3.29 (m, 1H, H-4‴), 3.31 (m, 1H, H-3″), 3.45 (4H, H-5″, H-6″b, H-5‴, and H-6‴b), 3.72 (2H, H-6″a and H-6‴a), 4.82 (d, *J* = 7.4 Hz, 1H, H-1‴), 5.06 (d, *J* = 7.3 Hz, 1H, H-1″), 6.47 (d, *J* = 1.6 Hz, 1H, H-6), 6.55 (d, *J* = 8.3 Hz, 1H, H-5′), 6.59 (s, 1H, H-3′), 6.71 (d, *J* = 1.6 Hz, 1H, H-8), 7.11 (d, *J* = 8.3 Hz, 1H, H-6′), 8.29 (s, 1H, H-2). $^{13}$C NMR (DMSO-$\delta_6$) $\delta$: 60.6 (C-6‴), 60.6 (C-6″), 69.6 (C-4‴), 69.6 (C-4″), 73.1 (C-3‴)[a], 73.2 (C-3″)[a], 76.4 (C-2‴)[b], 76.6 (C-2″)[b], 77.0 (C-5″)[c], 77.2(C-5‴)[c], 94.5 (C-8), 99.6 (C-6), 99.9 (C-1″), 100.4 (C-1‴), 103.9 (C-3′), 106.1 (C-4a), 106.6 (C-5′), 111.4 (C-1′), 120.5 (C-3), 132.1 (C-6′), 156.0 (C-2), 156.5 (C-2′), 157.3 (C-8a), 158.6 (C-4′), 161.6 (C-5), 162.9 (C-7), 180.5 (C-4). [a, b, c]: interchangeable

**[$^3$H]DHT-AR in Vitro Binding Assay.** This assay was performed according to the method described previously[11]. In brief, the gene sequence corresponding to the ligand-binding domain (AR-LBD, 609–919 a.a.) in the C-terminus of AR was expressed in the *E. coli* strain DH5$\alpha$ as a maltose-binding protein-fused protein (MBP-AR-LBD), followed by purification using amylose resin (BIO-RAD). Thus, the obtained recombinant MBP-AR-LBD (50 $\mu$g/mL), [$^3$H]-dihydrotestosterone (DHT, 2 nM), and test samples were incubated at 4 °C for 3 h. [$^3$H]DHT-bound MBP-AR-LBD was then precipitated with hydroxyapatite, and radioactivity was measured with a liquid scintillation counter. Values are the means of three independent determinations.

**Detection of PSA mRNA by Real-Time RT-PCR.** Prostate cancer (LNCaP) cells were incubated in RPMI 1640 medium supplemented with 2% charcoal-stripped serum for 24 h. Cells were then treated with DHT (0.1 nM) and test compounds. After 12 h, RNA from these cells was isolated, and the expression of PSA genes was determined by real-time quantitative reverse transcription-PCR (RT-PCR) and normalized to GAPDH mRNA. The primer sequences used were as follows: for PSA, 5′-AGG TCG GAG TCA ACG GAT TT-3′ (forward) and 5′-TAG TTG AGG TCA ATG AAG GG-3′ (reverse); for GAPDH, 5′-GGT CCT CAC AGC TGC CCA TC-3′ (forward) and 5′-CAG CCT GAG GCG TAG CAG GT-3′ (reverse). Values are the means of three independent determinations.

## RESULTS

**Isolation of Isoflavone Glucosides.** Freeze-dried groundnuts (261.9 g) were pulverized in a mixer and extracted with MeOH. The filtrate was concentrated to dryness and washed with hexane. The insoluble precipitate (1.02 g) was applied on a silica gel column prepared and developed with $CH_2Cl_2$:MeOH:$H_2O$ (3:1:0.1). From this chromatography, genistein was eluted at fraction 9–11. Fractions 15–70, which contain isoflavone glucosides as confirmed by TLC and HPLC analyses, were collected and further chromatographed with reversed phase HPLC using Developsil C30 developed with 17% $CH_3CN$. From this chromatography, pure compounds **1** (7.9 mg), **2** (5.4 mg), **3** (8.5 mg), **4** (3.2 mg), **5** (3.1 mg), and **6** (24.4 mg) were eluted at 18.3 min, 19.7 min, 28.1 min, 32.9 min, 34.4 min, and 40.3 min, respectively, and crude compounds **7** and **8** were eluted at 8.9 and 11.6 min, respectively. Crude compounds **7** and **8** were both purified by ODS HPLC developed with 20% MeOH (flow rate: 3.0 mL/min). The pure compounds were eluted at 15.0 min [**7** (3.2 mg)] and 16.8 min [**8** (3.6 mg)], respectively.

**Structural Elucidations of Compounds 1–8.** Compound **6** was identified as genistein-7-*O*-gentibioside by comparing MS, $^1$H, and $^{13}$C NMR data in pyridine-$d_5$ with those of reported data.[6]

The HR-ESI-MS(−) of **1** showed a $(M - H)^-$ peak at *m/z* 461.10949, and the molecular formula of **1** was determined as $C_{22}H_{22}O_{11}$ (calcd for $C_{22}H_{21}O_{11}$ 461.10839). $^1$H,$^{13}$C NMR, $^1$H–$^1$H DQF COSY, and HMQC spectral analyses of **1** clearly showed that the structure of **1** was constituted by O-substituted isoflavone (aglycone), one $\beta$-hexose ($J_{1,2}$ = 7.2 Hz), and one methoxy function ($\delta_H$ 3.83). In $^1$H NMR, two meta-coupled (*J* = 2.0 Hz) sp$^2$ mehines were observed at $\delta_H$ 6.60 (H-6) and $\delta_H$ 6.72 (H-8), which are characteristic for 5-OH and 7-OH substitutions in the A ring. In addition, an sp$^2$ methine network of $\delta_H$ 6.91 (*J* = 8.2 Hz, H-6′) − $\delta_H$ 6.25 (*J* = 2.0, 8.2 Hz, H-5′) − $\delta_H$ 6.34 (*J* = 2.0 Hz, H-3′) was also observed in $^1$H NMR, indicating a 2′,4′ or 3′,4′-*O*-substitution in the B ring. Furthermore, $^1$H–$^{13}$C long-range couplings were observed from methoxy ($\delta_H$ 3.83) and H-6 to $\delta_C$ 160.6 (C-5), and from H-6′ to $\delta_C$ 123.3 (C-3) in the HMBC experiment. From these observations, the methoxy function at C-5 and 2′,4′-*O*-substitution in the B ring were confirmed, and the aglycone in **1** was determined to be 2′-hydroxy,5-methoxy genistein. In


dx.doi.org/10.1021/jf305233t | J. Agric. Food Chem. 2013, 61, 2183–2187

$^1$H NMR, most signals derived from $\beta$-hexose were observed around the solvent peak ($\delta_H$ 3.30) under severe overwrapping, and we could not analyze their vicinal coupling constants. Thus, **1** was acetylated with $Ac_2O$/pyridine to give hexa-acetylated **1**. $^1$H NMR and $^1$H–$^1$H DQF COSY spectral analyses of hexa-acetylated **1** clearly proved the hexose to be $\beta$-glucose due to its large vicinal coupling constants [H-1′($\delta_H$ 5.24, d, $J$ = 7.2 Hz), H-2′($\delta_H$ 5.31, dd, $J$ = 7.2, 9.0 Hz), H-3′($\delta_H$ 5.35, dd, $J$ = 9.0, 9.3 Hz), H-4′ ($\delta_H$ 5.18, dd, $J$ = 9.0, 9.3 Hz), H-5′($\delta_H$ 3.96, ddd, $J$ = 2.3, 5.6, 9.0 Hz), H-6′($\delta_H$ 4.22, dd, $J$ = 2.3, 12.3 Hz and $\delta_H$ 4.28, dd, $J$ = 5.6, 12.3 Hz)]. The linkage of $\beta$-glucose at C-7 in aglycone was proven by the observations of $^1$H–$^{13}$C long-range couplings from H-1″ ($\delta_H$ 5.08), H-6, and H-8 to C-7 ($\delta_C$ 161.3) in the HMBC experiment. From the observations described above, **1** was identified as 2′-hydroxy,5-methoxy genistein-7-*O*-glucoside. Compound **1** was novel.

The HR-ESI-MS(−) of **2** showed a $(M - H)^-$ peak at $m/z$ 609.14469, and the molecular formula of **2** was determined to be $C_{27}H_{30}O_{16}$ (calcd for $C_{27}H_{29}O_{16}$ 609.14556). $^1$H,$^{13}$C NMR, $^1$H–$^1$H DQF COSY, and HMQC (in DMSO) spectral analyses of **2** showed that the structure of **2** was constituted by an O-substituted isoflavone (aglycone) and two $\beta$-hexoses (each $J_{1,2}$ = 7.2 Hz). The $^1$H and $^{13}$C NMR signals derived from aglycone were similar between **1** and **2**, while the methoxy signal in **1** disappeared in **2**. Thus, the aglycone in **2** was determined to be 2′-hydroxy genistein. The two $\beta$-hexoses were confirmed to possess a gentibioside structure because the $^1$H and $^{13}$C NMR chemical shifts of the two hexose signals were completely identical to those of **6** in DMSO-$d_6$. The attachment of gentibioside at C-7 in aglycone was proven by the observations of $^1$H–$^{13}$C long-range couplings from H-1″ ($\delta_H$ 5.05), H-6 ($\delta_H$ 6.57), and H-8 ($\delta_H$ 6.82) to C-7 ($\delta_C$ 164.7) in the HMBC experiment. Thus, **2** was identified as 2′-hydroxy genistein-7-*O*-gentibioside. Compound **2** was novel.

The HR-ESI-MS(−) of **3** showed a $(M - H)^-$ peak at $m/z$ 445.11528, and the molecular formula of **3** was determined to be $C_{22}H_{22}O_{10}$ (calcd for $C_{22}H_{21}O_{10}$ 445.11347). The $^1$H and $^{13}$C NMR spectra of **3** were similar to those of **1**. Differences were observed only on the signals in the B ring. In $^1$H NMR, the 2H equivalent methines of the B ring in **3** were observed as a doublet ($J$ = 8.4 Hz) at $\delta_H$ 6.79 and $\delta_H$ 7.31 suggesting a 4′-*O*-substitution in the B ring. Considering these observations and the molecular formula of **3** (= **1** − O), **3** was identified as 5-methoxy genistein-7-*O*-glucoside.[12]

The HR-ESI-MS(−) of **4** showed a $(M - H)^-$ peak at $m/z$ 475.12547, and the molecular formula of **4** was determined to be $C_{23}H_{24}O_{11}$ (calcd for $C_{23}H_{23}O_{11}$ 475.12404). The $^1$H- and $^{13}$C NMR spectra of **4** were similar to those of **1** and **3**. These signals derived from the B ring and the appearance of a methoxy signal ($\delta_H$ 3.78, $\delta_C$ 55.6) in **4** were different from **1** and **3**. The two doublet sp$^2$ methine signals, which coupled each other [H-5′ ($\delta_H$ 6.79, $J$ = 8.2 Hz) and H-6′ ($\delta_H$ 6.90, $J$ = 8.2 Hz)], and one singlet methine signal [H-2′ ($\delta_H$ 7.11)] were observed as B ring signals in the $^1$H NMR spectrum of **4**, and $^1$H–$^{13}$C long-range couplings were observed from H-2′ to C-6′ ($\delta_C$ 121.6) and C-3 ($\delta_C$ 124.9) in the HMBC spectrum of **4**. Therefore, a 3′,4′-O-substitution (B ring) in **4** was proven. In addition, $^1$H–$^{13}$C long-range couplings from the methoxy ($\delta_H$ 3.78) and H-5′ to C-3′ ($\delta_C$ 147.0) were also observed, and the attachment of the methoxy at C-3′ was proven. From these findings, **4** was identified as 3′,5-dimethoxy genistein-7-*O*-glucoside. Compound **4** was novel.

The HR-ESI-MS(−) of **5** showed a $(M - H)^-$ peak at $m/z$ 447.09314, and the molecular formula of **5** was determined to be $C_{21}H_{20}O_{11}$ (calcd for $C_{21}H_{19}O_{11}$ 447.09274). The $^1$H and $^{13}$C NMR signals of **5** derived from aglycone were identical to those of **2**, while the signals due to one $\beta$-glucose were observed in **5**. The attachment of $\beta$-glucose at C-7 was confirmed by the observations of $^1$H–$^{13}$C long-range couplings from H-1″ ($\delta_H$ 5.06) and H-6 ($\delta_H$ 6.46) to C-7 ($\delta_C$ 162.9). Thus, **5** was identified as 2′-hydroxy genistein-7-*O*-glucoside.[13]

The HR-ESI-MS(−) of **7** showed a $(M - H)^-$ peak at $m/z$ 623.16188, and the molecular formula of **7** was determined to be $C_{28}H_{32}O_{16}$ (calcd for $C_{28}H_{31}O_{16}$ 623.16121). The $^1$H signals of **7** derived from aglycone were closely similar to those of **1**, while low field shifts were observed at H-3′ and H-5′ ($\Delta$0.27 and $\Delta$0.19, respectively). In addition, the $^1$H and $^{13}$C signals due to two $\beta$-glucose units were observed in **7**. The attachment of the two $\beta$-glucoses was determined to be C-7 and C-4′ in aglycone because $^1$H–$^{13}$C long-range couplings from H-1″ ($\delta_H$ 5.08) and H-6 ($\delta_H$ 6.61) to C-7 ($\delta_C$ 161.3) and from H-1‴ ($\delta_H$ 4.82) to C-4′ ($\delta_C$ 158.2) and NOE between H-1‴ and H-5′ ($\delta_H$ 6.52) were observed in the HMBC and NOESY spectra of **7**, respectively. Thus, **7** was identified as 2′-hydroxy,5-methoxy genistein-4′,7-*O*-diglucoside. Compound **7** was novel.

The HR-ESI-MS(−) of **8** showed a $(M - H)^-$ peak at $m/z$ 609.14469, and the molecular formula of **8** was determined to be $C_{27}H_{30}O_{16}$ (calcd for $C_{27}H_{29}O_{16}$ 609.14556).

The $^1$H and $^{13}$C NMR of **8** were similar to those of **7**. The only difference between **7** and **8** was the disappearance of methoxy signals in **8**. Therefore, **8** was identified as 2′-hydroxy genistein-4′,7-*O*-diglucoside.[14]

**Effects of Compounds 1–8 on the Binding of DHT to AR.** Compounds **2** and **5** inhibited the binding of DHT to AR in a dose-dependent manner, the $IC_{50}$ values of **2** and **5** were 280 and 160 $\mu$M, respectively, and these inhibitory activities were about 4–7-fold less potent than that of flutamide,[15] which was clinically used for the treatment of prostatic diseases. However, compounds **1**, **3**, **4**, **6**, **7**, and **8** did not inhibit the binding of DHT to AR even at 400 $\mu$M. None of these compounds inhibited the binding of estradiol to the estrogen receptor up to 400 $\mu$M. (The binding inhibitory activities ($IC_{50}$) of genistein and daidzein for the estrogen receptor were 0.75 $\mu$M and 26 $\mu$M, respectively.) The above results are listed in Table 1.

**Effects of Compounds 1–8 on DHT-Induced PSA Expression.** Prostate-specific antigen (PSA) is a 33-kDa serine

**Table 1. Effects of Compounds 1–8 on Binding to the Androgen Receptor (AR) and Estrogen Receptor (ER)**

| compd | AR | ER |
|---|---|---|
| | $IC_{50}$ ($\mu$M) | $IC_{50}$ ($\mu$M) |
| 1 | >400 | >400 |
| 2 | 280 | >400 |
| 3 | >400 | >400 |
| 4 | >400 | >400 |
| 5 | 160 | >400 |
| 6 | >400 | >400 |
| 7 | >400 | >400 |
| 8 | >400 | >400 |
| flutamide | 39 | >400 |
| genistein | 270 | 0.75 |
| daidzein | >400 | 26 |


dx.doi.org/10.1021/jf305233t | J. Agric. Food Chem. 2013, 61, 2183–2187

protease, whose expression in the prostate is triggered by the androgen-mediated action of AR. Therefore, to determine whether compounds **1**–**8** showed AR agonistic or antagonistic activity, we examined the effects of these compounds on the DHT-induced expression of PSA mRNA in prostate cancer LNCaP cells. As shown in Table 2, compounds **2** and **5**, which

**Table 2. Effects of Compounds 1–8 on DHT-Induced PSA Expression**

| compd | $IC_{50}$ ($\mu$M) |
|---|---|
| 1 | >50 |
| 2 | 20 |
| 3 | >50 |
| 4 | >50 |
| 5 | 18 |
| 6 | >50 |
| 7 | >50 |
| 8 | >50 |
| flutamide | 2.3 |

inhibited DHT binding to AR *in vitro*, suppressed the DHT-induced expression of endogenous PSA mRNA in LNCaP cells with $IC_{50}$ values of 20 and 18 $\mu$M, respectively, whereas each compound alone failed to induce the expression of PSA mRNA (data not shown). These results showed that **2** and **5** possess AR antagonistic but not agonistic activities.

## DISCUSSION

In this study, we succeeded in isolating 8 isoflavone glucosides (**1**–**8**) from the tubers of groundnut. From these compounds, **1**, **2**, **4**, and **7** were novel. Compounds **3**, **5**, and **8** have been reported to be included in the legume family such as *Ulex* and *Lupinus*,[12−14] with compound **6** being included in groundnut (*Apios*).[6] We could not find genistin, which was reported to be present in groundnut.[6]

We tested the binding activities of **1**–**8** for ER and AR and found that **2** and **5** inhibited DHT binding to AR but not estrogen binding to ER, indicating that these two compounds can bind to AR selectively. Furthermore, the activity of these compounds for AR was proven to be antagonistic with an assay using cultured LNCaP cells. Compounds **2** and **5** have been shown to be the first isoflavone glucosides to possess AR antagonistic activity.

AR, a member of the nuclear receptor superfamily, is a critical mediator of prostate cancer and benign prostatic hyperplasia; therefore, treatment with AR antagonists is expected to be an effective prostate cancer and benign prostatic hyperplasia therapy. The AR antagonist flutamide has been clinically used for prostate cancer therapy Thus, the development of a new type of AR antagonist is an attractive strategy to overcome these diseases.

We examined the structure–activity relationship of compounds **1**–**8** for AR antagonistic activity and concluded that the coexistence of the 2′, 4′-OH structure in the B ring and 5-OH in the A ring was important. Further biological studies on **1**–**8** are in progress.

## AUTHOR INFORMATION

### Corresponding Author

*Tel: +81-3-5981-3433. Fax: +81-3-5981-3433. E-mail: kshindo@fc.jwu.ac.jp.

### Author Contributions

[‖]M.I., E.F., and S.M. equally contributed to this work.

### Notes

The authors declare no competing financial interest.

## REFERENCES

(1) Wilson, P. W.; Gorny, J. R.; Blackmon, W. J.; Thresher, W. C. Fatty acids on the American groundnut (*Apios americana*). *J. Food Sci.* **1986**, *51*, 1387–1388.

(2) Wilson, P. W.; Pichardo, F. J.; Liuzzo, J. A.; Blackmon, W. J.; Rwynolds, B. D. Amino acids on the American groundnut (*Apios americana*). *J. Food Sci.* **1987**, *52*, 224–225.

(3) Ogasawara, Y.; Hidano, Y.; Kato, Y. Study on carbohydrate composition of Apios (*Apios americana* Medikus) flower and tubers. *Nippon Shokuhin Kagaku Kogaku Kaishi* **2006**, *53*, 130–136.

(4) Okubo, K.; Yoshiki, Y.; Sugihara, T.; Tsukamoto, C.; Hoshikawa, K. DDMP-conjugated Saponin (soyasaponin $\beta$g) isolated from American groundnut (*Apios americana*). *Biosci. Biotechnol. Biosci.* **1994**, *58*, 2248–2250.

(5) Krishman, H. B. Identification of genistein, an anticarcinogenic compound, in the edible tubers of the American groundnut (*Apios americana* Medikus). *Crop Sci.* **1998**, *38*, 1952–1056.

(6) Nara, K.; Nihei, K.; Ogasawara, Y.; Koga, H.; Kato, Y. Novel isoflavone diglycoside in groundnut (*Apios americana* Medik). *Food Chem.* **2011**, *124*, 703–710.

(7) Aoki, T.; Akashi, T.; Ayabe, S. Flavonoids of leguminous plants: structure, biological activity, and biosynthesis. *J. Plant Res.* **2000**, *113*, 475–488.

(8) Naim, M.; Gestetner, B.; Bondi, A.; Birk, Y. Antioxidative anf antiheolytic activities of soybean isoflavones. *J. Agric. Food Chem.* **1976**, *24* (6), 1174–1177.

(9) Lee, H. C.; Yang, L.; Xu, Z. J.; Yeung, V. S.; Huang, Y.; Chen, Z. Relative antioxidant activity of soybean isoflavones and their glucosides. *Food Chem.* **2005**, *90* (4), 735–741.

(10) Hsieh, C.; Santell, C. R.; Haslam, Z. S.; Heiferich, G. W. Estrogenic effects of genistein on the growth of estrogen receptor-positive humanbrest cancer (MCF-7) cells *in vitro* and *in vivo*. *Cancer Res.* **1998**, *58*, 3833–3838.

(11) Kawamura, T.; Fujimaki, T.; Hamanaka, N.; Torii, K.; Kobayashi, H.; Takahashi, Y.; Igarashi, M.; Kinoshita, N.; Nishimura, Y.; Tashiro, E.; Imoto, M. Isolation and structure elucidation of a novel androgen antagonist, arabilin, produced by *Streptomyces* sp. MK756-CF1. *J. Antibiot.* **2010**, *63*, 601–605.

(12) De Rodriguez, D. J.; Chulia, J.; Simoes, C. M.; Amoros, M.; Mariotte, A. M.; Girre, L. Search for "in vivo" antiviral activity of a new isoflavonic glucoside from Ulex europaeus. *Planta Med.* **1990**, *56* (1), 59–62.

(13) Shibuya, Y.; Tahara, S.; Kimura, Y.; Mizutani, J. New isoflavone glucosides from white lupine (*Lupinus albus* L.). *J. Biosci.* **1991**, *6* (7–8), 513–518.

(14) Watanabe, K.; Kinjo, J.; 'Nohara, T. Leguminous plants. XXXIX. Three new isoflavone glucosides from *Lupinus luteus* and *L. polyphyllus x arboreus*. *Chem. Pharm. Bull.* **1993**, *41* (2), 394–396.

(15) Sogai, C. P.; Vagaiwala, R. M.; Whitmore, F. W., Jr. Experience with flutamide in patients with advanced prostate cancer without prior endocrine therapy. *Cancer* **1984**, *4*, 744–750.

**Running title:** A Detection System for IRE1α Dimerization Using a BiFC Assay

# Establishment of a New Detection System for the Dimerization of IRE1α Using a BiFC Assay

**Satoko SHINJO, Etsu TASHIRO†, Masaya IMOTO**

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

Received February 12, 2013; Accepted March 18, 2013

**†To whom correspondence should be addressed:** Fax: +81 45 566 1557; e-mail: tashiro@bio.keio.ac.jp

**Abbreviations:** ATF6, activating transcription factor 6; BiFC, bimolecular fluorescent complementation; BiP, immunoglobulin-binding protein; CHOP, C/EBP homologous protein; DTT, dithiothreitol; ER, endoplasmic reticulum; ERAD, ER-associated degradation; GRP78, glucose-regulated proteins 78; IRE1α, inositol requiring kinase 1 α; PERK, protein kinase regulated by RNA-like ER kinase; RT-PCR, reverse transcription PCR; UPR, unfolded protein response; XBP1, x-box binding protein 1

**Abstract (within 60 words)**

We developed a new detection system for the activation of an ER stress sensor, IRE1α, by evaluating its dimerization using a bimolecular fluorescence complementation (BiFC) assay. By detecting the fluorescence derived from the reconstituted cerulean, this assay system enables us to distinguish the activation behaviors of IRE1α toward ER stress-inducing compounds.

**Key words**

IRE1α; imaging; BiFC assay; UPR; small molecule

The endoplasmic reticulum (ER) is the organelle within which newly synthesized transmembrane or secreted proteins are folded. The accumulation of unfolded proteins in the ER, a condition referred to as "ER stress", activates the unfolded protein response (UPR) in order to restore homeostasis in the ER. In mammalian cells, UPR is triggered by the activation of three ER stress sensors: activating transcription factor 6 (ATF6), inositol-requiring kinase 1 α (IRE1α), and protein kinase regulated by RNA-like ER kinase (PERK) (reviewed in[1, 2]). However, these three ER stress sensors have been reported to display distinct activation behaviors toward different forms of ER stress that are induced by pharmacological agents[3]. Therefore, monitoring the activation behavior of ER stress sensors may provide a better understanding of UPR functions in physiological and pharmacological conditions.

Among the three ER stress sensors, IRE1α is the most evolutionarily conserved, having been detected in yeast through to mammalian cells[4]. Upon activation, the luminal domain of IRE1α undergoes dimerization, triggered by the dissociation of glucose-regulated proteins 78 (GRP78, also known as BiP, immunoglobulin-binding protein)[5, 6]. The dimerization allows their transautophosphorylation and cofactor binding[7, 8], which activates RNase domains and initiates the unconventional mRNA splicing of x-box binding protein 1 (XBP1)[9, 10]. The translation of the spliced form of XBP1 mRNA produces a potent transcription factor, which can up-regulate several UPR target genes, such as ER chaperones[11]. In this study, we aimed to develop a new detection system for the activation behavior of IRE1α.

Recently, the bimolecular fluorescent complementation (BiFC) assay has been developed to evaluate protein interactions in cultured cells[12, 13]. The BiFC assay is based on the formation of a fluorescent complex, which results when two proteins fused to non-fluorescent fragments of a fluorescent protein interact with each other. Because the dimerization of IRE1α has been reported to be a sufficient and efficient step for activation[4, 14-17], the activation of IRE1α is expected to be monitored by evaluating its dimerization by using a BiFC assay.

We first developed expression vectors for the BiFC system with the intention of detecting IRE1α dimerization in mammalian cells (illustrated in Fig. 1A). Since the luminal domain of IRE1α, which is located in its N-terminal, has been shown to be responsible for IRE1α dimerization[14], the N-terminal half of IRE1α (IRE1αΔKR; residues 1-570) was fused to the N-terminal half of cerulean (cerN; residues 1-154) or the C-terminal half of cerulean (cerC; residues 155-238). Furthermore, we established IRE1αΔKR-cerN and -cerC expressing stable clones derived from HeLa cells (named HeLa/IRE1α-BiFC cells). The sequences encoding IRE1αΔKR-cerN and -cerC were linked with the 2A peptide from the insect Thosea asigna virus, which works as a self-cleaving signal and enables expression of two separate proteins from a single transcript[18, 19] (Fig. 1A). We transfected plasmid-encoding IRE1αΔKR-cerN-2A- and IRE1αΔKR-cerC to HeLa cells and selected neomycin-resistant cells. Two independently isolated HeLa transfectants, designated "HeLa/IRE1α-BiFC#5" and "HeLa/IRE1α-BiFC#10", were shown to express both IRE1αΔKR-cerN and -cerC by western blotting using anti-FLAG antibody (Fig. 1B), because both IRE1αΔKR-cerN and –cerC have FLAG-tag at their C-terminus. Both IRE1αΔKR-cerN and –cerC were detected as double bands (Fig. 1B). Since previous studies have reported that a cytoplasmic domain deletion mutant of IRE1α was detected as double bands in Western blot analysis due to the glycosylation at Asn-176[15, 20], it is likely that the faster and slower migrating form of both IRE1αΔKR-cerN and –cerC are non-glycosylated and glycosylated proteins, respectively.

Next, we examined whether fluorescence derived from cerulean was observed in HeLa/IRE1α-BiFC cells in response to ER stress inducer dithiothreitol (DTT), because it has been shown that DTT induced dimerization of IRE1α[6]. HeLa/IRE1α-BiFC#5 and #10 cells were treated with 3 mM DTT for 1 hour, stained with anti-FLAG antibody and observed under a confocal microscope. The fluorescence derived from cerulean was hardly detected in HeLa/IRE1α-BiFC#5 and #10 cells under normal condition; however, it could be observed in DTT-treated HeLa/IRE1α-BiFC#5

and #10 cells, which showed a dot-like distribution, referred to as "foci" (Fig. 1C). Furthermore, IRE1αΔKR-cerN and –cerC, which were stained by anti-FLAG, also formed foci and the staining was overlapped with the fluorescence derived from cerulean in DTT-treated HeLa/IRE1α-BiFC#5 and #10 cells, whereas IRE1αΔKR-cerN and –cerC were perinuclear staining under normal condition (Fig. 1C). Foci formation of IRE1α has been reported to be detected after treatment with tunicamycin or DTT, which is caused by the oligomerization subsequent to the dimerization of IRE1α, because the foci formation is impaired by the dimerization-disrupted mutant of IRE1α[16]. Therefore, the cerulean fluorescence observed in HeLa/IRE1α-BiFC#5 and #10 cells is considered to be the consequence of IRE1α dimerization, as judged from its overlapping with foci. Furthermore, the fluorescence derived from the reconstitution of cerulean was also observed after the treatment of HeLa/IRE1α-BiFC#5 cells with other ER stress inducers, i.e. thapsigargin, tunicamycin or 2-deoxyglucose. Similar to DTT, the fluorescence induced by thapsigargin, tunicamycin, or 2-deoxyglucose also resulted in the formation of foci, which overlapped with the staining by anti-FLAG (Fig. 2). On the other hand, non-ER stress inducers[21, 22], i.e. camptothecin, an inhibitor of topoisomerase I, and staurosporine, a kinase inhibitor, induced neither foci formation nor the reconstitution of cerulean (Fig. 2) even at the concentrations that inhibited the growth of HeLa cells[23, 24]. These results indicated that IRE1α dimerization could be easily detected by the fluorescence derived from the reconstituted cerulean in HeLa/IRE1α-BiFC cells.

In order to quantify the reconstitution of cerulean, we next calculated the ratio of the cells that exhibited fluorescence (hereafter designated as BiFC-positive cells) as follows: the number of cells stained with the anti-FLAG antibody was divided by the number of BiFC-positive cells. As a result, DTT and thapsigargin increased the ratio of BiFC-positive cells, which was approximately two-fold and constant when measured at time points from 1 to 8 hours after treatment, with statistical significance, while tunicamycin and 2-deoxyglucose time-dependently and significantly increased the ratio

of BiFC-positive cells after treatment for 4 hours (Fig. 3A). To examine when endogenous IRE1α was activated upon treatment with these compounds, we performed RT-PCR to detect splicing of XBP1 mRNA, a downstream event of IRE1α dimerization, using the same method as previously described[25]. As a result, the ratio of BiFC-positive cells was increased at the same time as or prior to the splicing of XBP1 mRNA (Fig. 3B). Therefore, the ratio of BiFC-positive cells is suggested to reflect the activation of endogenous IRE1α. Moreover, these observations were consistent with the previous report by DuRose et al., which indicated that the activation of IRE1α induced by tunicamycin was slower than that induced by DTT and thapsigargin[3]. On the other hand, camptothecin and staurosporine, non-ER stress inducers, increased neither the ratio of BiFC-positive cells nor the splicing of XBP1 mRNA at any time point measured between 1 and 8 hours (data not shown). These results further confirmed that the fluorescence derived from the reconstituted cerulean in HeLa/IRE1α-BiFC cells would reflect endogenous IRE1α dimerization.

In conclusion, we established a new assay system for the detection of IRE1α dimerization using a BiFC assay. In this assay system, the fluorescence derived from the reconstituted cerulean was successfully observed in response to treatment with ER stress-inducing compounds. Moreover, this assay system enables us to distinguish between the activation behaviors of IRE1α induced by tunicamycin from those induced by DTT and thapsigargin, which was consistent with a previous report[3]. Therefore, while IRE1α dimerization could be evaluated by detecting foci formation of IRE1α-GFP[16], our method might be also useful for the quantitative detection of the dimerization of IRE1α.

**Acknowledgement**

We would like to thank Dr. Atsushi Miyawaki for kindly providing cDNA of Cerulean. This work is supported by KAKENHI (grant numbers 23510283 to E.T.) from MEXT,