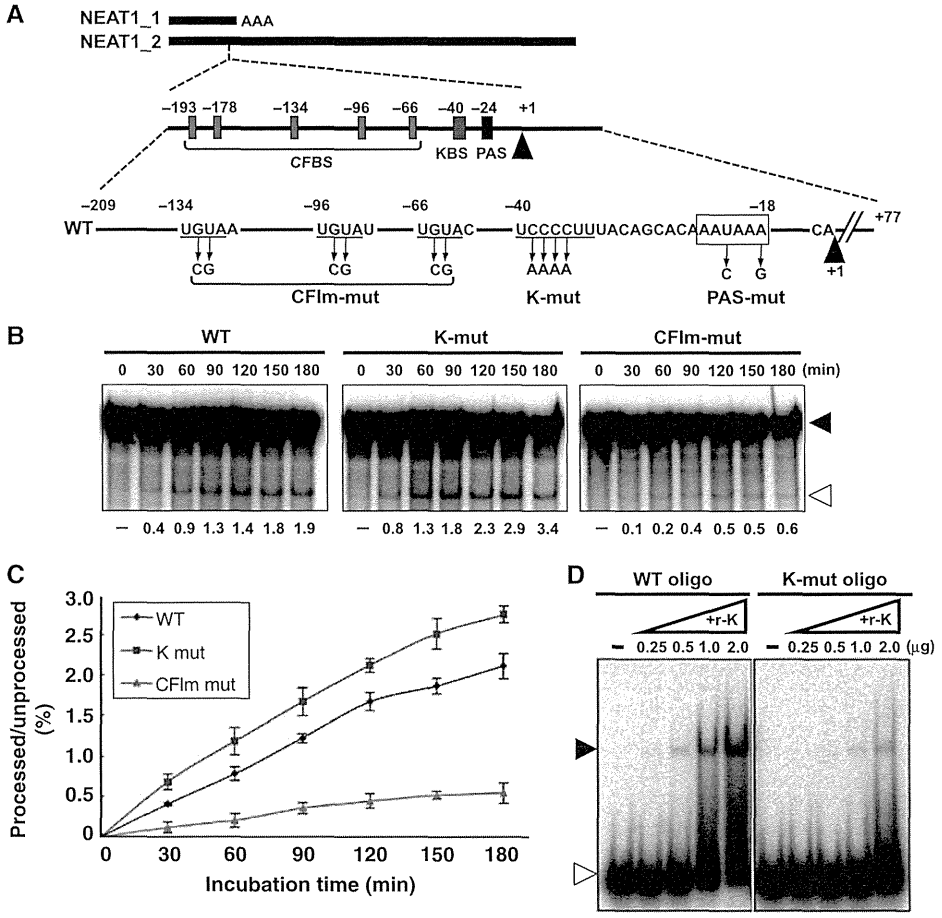

Figure 6 Roles of CFIm and HNRNPK in 3′-end processing of NEAT1_1. (**A**) Schematic representation of the substrate RNAs for the *in vitro* processing reaction that contains the region spanning the 3′-end of NEAT1_1. Numbers indicate distance from the polyadenylation site of NEAT1_1. Middle scheme represents putative sequences around the NEAT1_1 3′-end processing site that are recognized by the CFIm complex (CFBS: red boxes) or HNRNPK (KBS: blue box). Mutated positions on the mutant substrates (CFIm-mut, K-mut, and PAS-mut) are indicated. (**B**) Recapitulation of CFIm-dependent 3′-end processing of NEAT1_1 *in vitro*. Incubation time is shown above each panel. Substrate RNAs are represented on the top. Unprocessed and processed bands are shown with closed and open triangles, respectively, on the right. Processing efficiencies (%) are shown below each panel. (**C**) Average values of the processing efficiencies obtained from three independent experiments. (**D**) Detection of sequence-specific RNA binding of HNRNPK. Gel mobility shift assay to detect binding of recombinant HNRNPK protein (r-K) with RNA fragments (30 nt) derived from WT and K-mut, WT oligo, and K-mut oligo, respectively, are shown. The RNA–protein complex and free RNA are shown with closed and open triangles, respectively. Amounts of supplemented r-K (μg) are shown above each panel.

concentration-dependent manner (Figure 7B, lanes 1–4). Immunoprecipitation of UV-crosslinked CPSF6 and NUDT21 clearly revealed that the RNA bindings of CPSF6 and NUDT21 were diminished (~50%) in the presence of r-K (Figure 7C–E). The mutation of KBS (K-mut) substantially elevated the UV-crosslinking of CPSF6 and NUDT21 (Figure 7B, lanes 2 and 6). The interfering effect of r-K on CPSF6 binding was milder on K-mut compared with WT (Figure 7B, lanes 6–8). These data indicate that HNRNPK binding to KBS results in the arrest of CFIm binding.

To obtain further mechanistic insights into the HNRNPK-dependent arrest of the 3′-end processing of NEAT1_1, the interaction between HNRNPK and CFIm was investigated. Endogenous HNRNPK was coimmunoprecipitated with NUDT21 in the presence of RNase A (Figure 8A, lanes 5 and 6) but not with CPSF6 (Figure 8A, lanes 7 and 8). This interaction was confirmed during *in vitro* processing, in which the supplemented r-K prominently coprecipitated with αNUDT21 (Figure 8B). This result was supported by protein-interaction data in the public database, in which NUDT21 but not CPSF6 was listed as an HNRNPK interactor (http://www.genecards.org). Importantly, the NUDT21–CPSF6 interaction was markedly weakened (~50%) in the presence of excess r-K (Figure 8B and C), suggesting that HNRNPK competes for the binding of NUDT21 with CPSF6.

To verify the competition between HNRNPK and CPSF6 for binding with NUDT21, the recombinant CFIm complex (NUDT21–CPSF6) was purified from HEK293T cells that were contransfected with two expression plasmids for the streptavidin-binding peptide (SBP)-tagged NUDT21 and CPSF6 (Figure 8D). Purified CFIm complex was immobilized on streptavidin-conjugated beads through the SBP-tag on NUDT21 (Figure 8E). The beads were mixed with r-K, and the binding of r-K to the beads was detected simultaneously with CPSF6 dissociation from the beads (Figure 8E).

As shown in Figure 8F, r-K specifically associated with the beads only when the CFIm complex was conjugated to the beads (compare lanes 4 and 8 in the bottom panel of

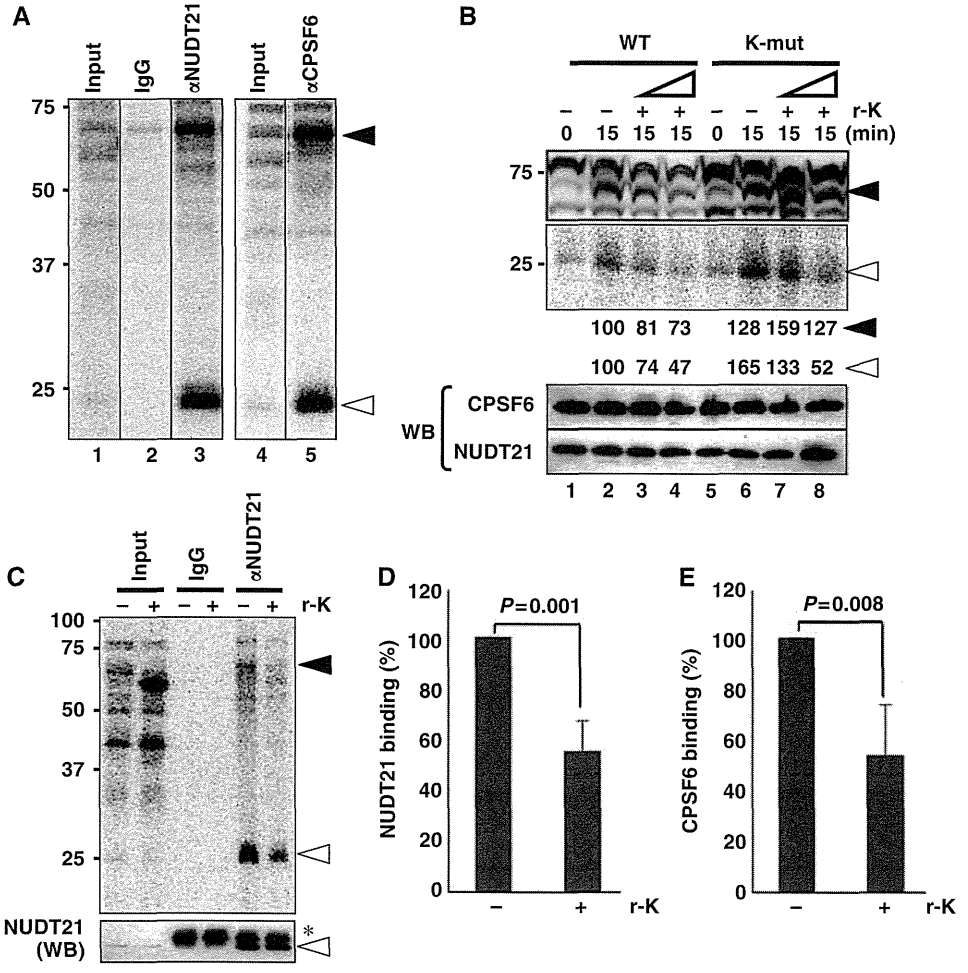

Figure 7 Molecular mechanism of the alternative 3′-end processing of NEAT1. (**A**) Detection of RNA binding of CFIm during *in vitro* processing by UV-crosslinking. UV-crosslinked WT substrate RNA-binding proteins were detected as ^{32}P-labelled proteins on SDS–PAGE. The 68- and 25-kDa UV-crosslinked RNA-binding proteins (closed and open arrows, respectively) were immunoprecipitated with antibodies against CPSF6 and NUDT21. (**B**) RNA bindings of CPSF6 and NUDT21 are affected by the addition of r-K. The two RNA substrates (WT and K-mut) employed are shown on the top. Addition of r-K (+) at two concentrations (10 × and 30 × excess of endogenous HNRNPK in HNE) and the incubation time (min) are shown above. Closed and open triangles are as in (**A**). Intensities of the 68- and 25-kDa bands were quantified and normalized by the levels of CPSF6 and NUDT21, respectively, which were detected by the western blot (WB) shown below. Molecular weight marker is shown on the left. (**C**) Confirmation of r-K-dependent inhibition of RNA binding of CPSF6 and NUDT21. Presence (+) or absence (−) of r-K (30 × excess) is indicated above the panel. UV-crosslinking of r-K (~55 kDa) was detected in the Input lane (+r-K) and in Supplementary Figure S6B. Closed and open triangles are as in (**A**). Amounts of NUDT21 in the input samples and immunoprecipitated samples were detected by the WB shown below the panel. Asterisk represents IgG light chain. (**D**) Quantification of immunoprecipitated, UV-crosslinked NUDT21 in the presence (+) or absence (−) of r-K (open triangle in the upper panel in **C**). Data were normalized with the total amounts of NUDT21 in each immunoprecipitation sample (lower panel in **C**). Graph shows the average (with s.d.) of three independent experiments. *P*-value was calculated by Student's *t*-test. (**E**) Quantification of immunoprecipitated CPSF6, as in (**D**). Antibodies are shown in Supplementary Table S4. Figure source data can be found with the Supplementary data.

Figure 8F). The addition of r-K caused the marked dissociation of CPSF6 from the beads, whereas the addition of BSA caused much less release of CPSF6 (compare lanes 8 and 3 in the top panel in Figure 8F). These results strongly support our argument that HNRNPK competes for the binding of NUDT21 with CPSF6, which is a possible underlying mechanism for the arrest of CFIm binding by HNRNPK (Figure 9A).

Discussion

Expansion of paraspeckle components

In the current study, the FLJ cDNA-based localization screening revealed 34 new PSPs. Many of the new PSPs are likely present throughout the nucleoplasm, and subsets may be concentrated in paraspeckles. An analysis of the compilation of all of the PSPs (Figure 4; Table I) indicated that most possess canonical RNA-binding domains. NUDT21, which was found to possess no canonical RNA-binding motif, has a NUDIX hydrolase domain that acts like an authentic RNA-binding protein (Yang *et al*, 2010). Some of the paraspeckle-localized RNA-binding proteins (e.g., NONO, SFPQ, RBM14, EWSR1, FUS, TAF15, and TARDBP) mediate both transcription and RNA processing (Auboeuf *et al*, 2005). Several of the new PSPs (e.g., AHDC1, DLX3, and ZNF335) are likely to be DNA-binding proteins that are involved in transcriptional control. This finding raises the possibility that paraspeckles may integrate tightly coupled transcription and post-transcriptional events.

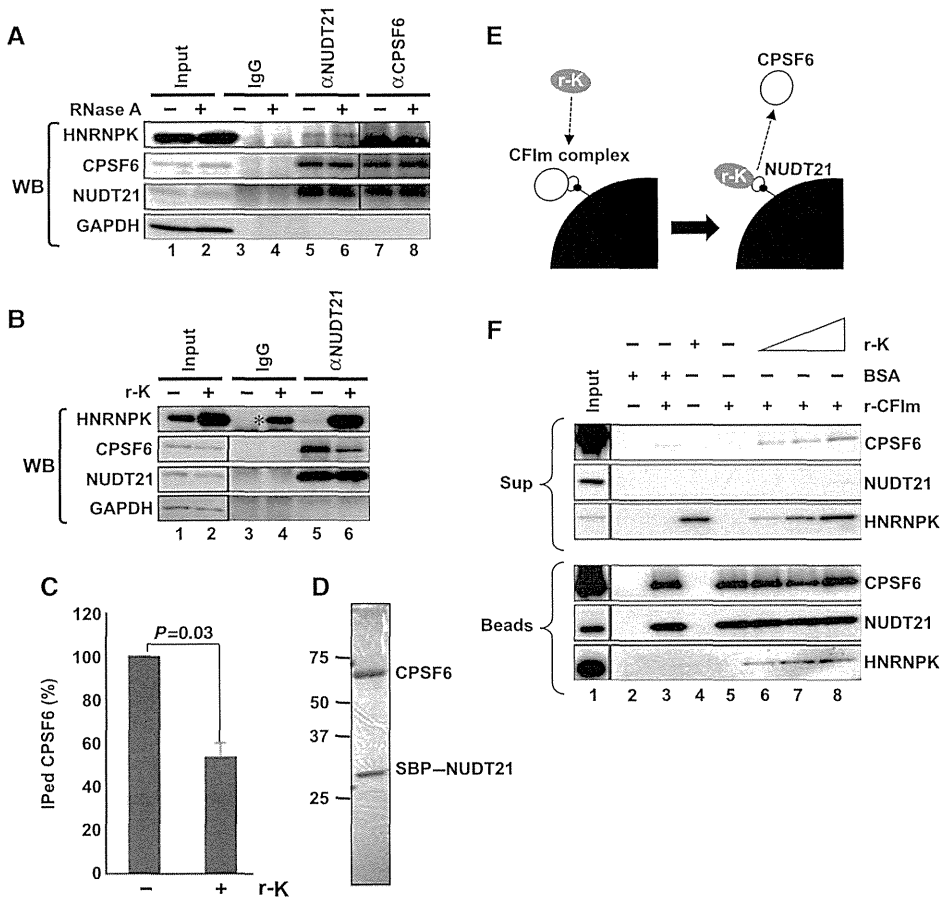

Figure 8 HNRNPK captures NUDT21 from the functional CFIm complex. (**A**) Interaction between HNRNPK and NUDT21 *in vivo*. Immunoprecipitations with αNUDT21 and αCPSF6 were performed in the presence (+) or absence (−) of RNase A. CPSF6, NUDT21, and HNRNPK were detected by WB. (**B**) The CPSF6–NUDT21 interaction is affected by HNRNPK. HNE was incubated with r-K in the presence of WT RNA. Immunoprecipitation with αNUDT21 was performed in the presence (+) or absence (−) of r-K. CPSF6, NUDT21, HNRNPK, and GAPDH were detected by WB. Immunoprecipitation of CPSF6 was diminished in the presence of r-K. Asterisk represents nonspecific binding of excess r-K to IgG. (**C**) Quantification of immunoprecipitated CPSF6 in the presence or absence of r-K. Relative amount (%) of CPSF6 was normalized by the immunoprecipitated NUDT21 level (**B**). Graph shows the average (with SD) of three independent experiments. *P*-value was calculated by Student's *t*-test. (**D**) Purified CFIm complex. The SDS–PAGE gel was stained with Coomassie Brilliant Blue. Both CPSF6 and SBP–NUDT21 possess Flag- and HA-tags. Size maker is shown on the left. (**E**) Schematics of the competitive binding assay carried out in **F**. The CFIm complex (30 pmol NUDT21–CPSF6 complex, white circles) was immobilized on streptavidin-conjugated beads (black quarter circle) through SBP-tag (small black circle). Recombinant HNRNPK (0–90 pmol r-K, grey circle) was mixed with the beads. Binding of r-K with NUDT21 is expected to lead to dissociation of CPSF6 from the beads. BSA (90 pmol) was used as a control. (**F**) Detection of CPSF6, NUDT21, and HNRNPK in bead (Beads) and supernatant (Sup) fractions. Proteins in the binding reaction are shown by + on the top of the panels. Each protein was detected by western blot, as shown on the right of the panels. Input lanes were loaded with 1/30 of the proteins used. Antibodies used are shown in Supplementary Table S4. Figure source data can be found with the Supplementary data.

Yang *et al* (2011) reported that nuclear body-localized ncRNAs mediate the attachment of the specific chromosomal locus to the nuclear bodies that control the epigenetic status of the gene loci. NEAT1 may play a similar role in the attachment of the specific chromosomal locus to the paraspeckles, where multiple regulatory PSPs are enriched for modulating gene expression from the specific locus. Alternatively, paraspeckles may be involved in RNA-dependent epigenetic regulation. Indeed, two of the identified PSPs, FUS and HNRNPK, participate in epigenetic regulation through their interactions with long ncRNAs (Wang *et al*, 2008; Huarte *et al*, 2010).

Several PSPs are disease-related. The genes for nine PSPs (NONO, SFPQ, CPSF6, EWSR1, FUS, TAF15, DAZAP1, RBM3, and SS18L1) are the breakpoints of chromosomal translocation that result in the production of abnormal fusion proteins responsible for various cancers (Kim *et al*, 2006). Four of them (SFPQ, NONO, DAZAP1, and FUS) belong to category 1, suggesting that the paraspeckle structure is altered in tumour cells in which the genes have undergone translocations. FUS and TARDBP are commonly associated with a neurodegenerative disease, familial amyotrophic lateral sclerosis (ALS) (Lagier-Tourenne and Cleveland, 2009). TARDBP was found to associate prominently with NEAT1 in the brain of patients with FTLD-TDP, which is an ALS-related neurodegenerative disease with TARDBP inclusions (Tollervey *et al*, 2011). These evidences suggest that TARDBP associated with NEAT1 is sequestered in paraspeckles, where it is functionally modulated.

Mechanism of alternative RNA processing of NEAT1 ncRNA

NEAT1_2 was found to be essential for paraspeckle formation. Therefore, the alternative 3′-end processing event that

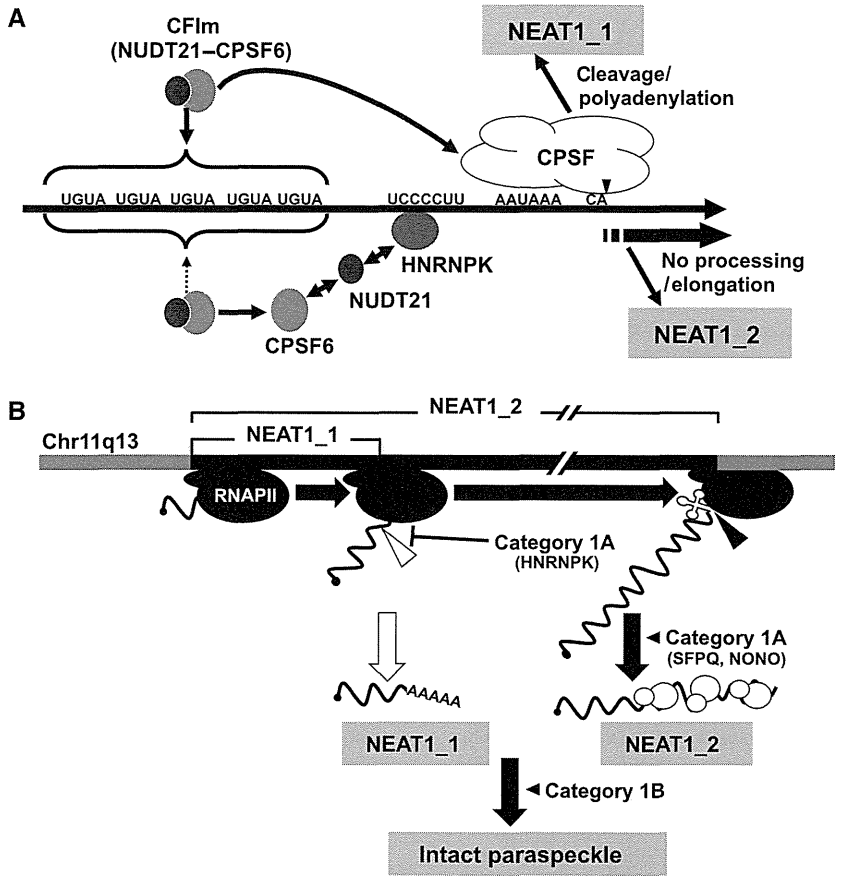

Figure 9 (**A**) Models of NEAT1 isoform synthesis. NEAT1 border region is shown with the critical sequence elements. Pathways for NEAT1_1 and NEAT1_2 synthesis are shown above and below the NEAT1 scheme. For NEAT1_2 synthesis, HNRNPK binds with the UCCCCUU sequence, captures NUDT21 from the functional CFIm (CPSF6–NUDT21), and arrests CFIm binding to upstream UGUA sequences. (**B**) Current model of intact paraspeckle formation. The essential steps, including 1) ongoing transcription of NEAT1 by RNA polymerase II (RNAPII), 2) NEAT1_2 synthesis by alternative 3′-end processing, 3) NEAT1_2 stabilization by category 1A proteins, such as SFPQ and NONO, and 4) subsequent assembly step(s), are schematized and represented with bold black arrows. Category 1B proteins act in an essential step other than NEAT1_2 accumulation. NEAT1_1 synthesis is dispensable; therefore, it is shown with a white arrow. The 3′-ends of NEAT1_1 and NEAT1_2 are formed by distinct mechanisms: canonical polyadenylation (open triangle) and RNase P cleavage (closed triangle). The significance of the noncanonical 3′-end processing of NEAT1_2 remains uncertain.

leads to NEAT1_2 accumulation is a fundamental molecular event for paraspeckle formation. Alternative 3′-end processing, which produces various mRNA isoforms with different 3′-UTR lengths, is utilized mainly for situations in which the produced mRNA isoforms are subjected to differential regulation by 3′-UTR-interacting factors (Lutz, 2008). In the case of NEAT1 ncRNA, this mechanism diversifies the ncRNA functions.

The alternative 3′-end processing of NEAT1 comprises two distinct 3′-end processing mechanisms: canonical polyadenylation for NEAT1_1 and RNase P-mediated cleavage for NEAT1_2. CFIm binds UGUA sequences and facilitates processing and polyadenylation at adjacent sites (Venkataraman *et al*, 2005). Our RNAi and *in vitro* analyses indicated a corresponding mechanism for the 3′-end processing of NEAT1_1.

PSP7/CPSF7 was reported to form a heterodimer with NUDT21 to facilitate 3′-end processing (Kim *et al*, 2010). However, we observed that CPSF7 RNAi markedly decreased the NEAT1_2 level, which was the opposite effect of NUDT21 RNAi (see Supplementary Table S3). This finding suggests that CPSF7 has an additional role in NEAT1_2 processing or stabilization. Alternatively, CPSF7 may play a counteracting role to that of CPSF6–NUDT21 in 3′-end processing under a specific (e.g., paraspeckle-localized) condition.

The multifunctional HNRNPK protein is involved in transcriptional regulation, pre-mRNA splicing, mRNA stability, and translation (Bomsztyk *et al*, 2004). Our data add a new function to this list: the regulation of 3′-end processing. The results of our RNAi and immunoprecipitation experiments suggest that HNRNPK is required for NEAT1_2 accumulation through its arrest of the 3′-end processing of NEAT1_1. The results of our *in vitro* processing and UV-crosslinking experiments reveal that HNRNPK interferes with the RNA binding of the CFIm complex through its binding to KBS. The possibility that HNRNPK additionally participates in the stabilization or noncanonical 3′-end processing of the NEAT1_2 isoform cannot be ruled out. Our preliminary results show that HNRNPK is required to maintain the NEAT1_2 level, even in the absence of CPSF6, suggesting additional role(s) of HNRNPK in the accumulation of the NEAT1_2 isoform. Mapping of the HNRNPK-binding sites may uncover additional function(s) of HNRNPK in NEAT1 expression.

Our coimmunoprecipitation results indicated that NUDT21 but not CPSF6 interacted with HNRNPK *in vivo* and *in vitro*. The supplemented r-K interacted with NUDT21, which resulted in a diminished interaction with CPSF6 and suggested the underlying mechanism. The binding of HNRNPK to KBS proximal to CFBS would provide an environment in which HNRNPK and CPSF6 effectively compete for association with NUDT21, which eventually determines the NEAT1 isoform ratio. The dissection of HNRNPK and NUDT21 to identify the interacting domain(s) and further interaction studies would solidify this model.

Several RNA-binding proteins (e.g., PTBP1, NONO, and ELAVL1) reportedly bind to upstream sequences implicated in the regulation of mRNA 3′-end processing; however, their detailed mechanisms of action remain to be investigated (Millevoi and Vagner, 2010). It would be intriguing to pursue the generality of the HNRNPK-dependent regulatory mechanism in the 3′-end processing of mRNAs and other ncRNAs. The NEAT1 ratio is controllable because it is variable in different mouse tissues (Nakagawa *et al*, 2011). The HNRNPK protein is expressed ubiquitously, but its activity is controlled by post-translational modifications, such as phosphorylation, methylation, and ubiquitination, under various conditions, including DNA damage (Chen *et al*, 2008). This fact raises the possibility that the NEAT1 ratio is controlled through the modification status of HNRNPK. Our experimental system for the functional rescue of HNRNPK (used in Figure 5C and D) would be useful to identify the important modification site(s) of HNRNPK required for the regulation of NEAT1 alternative processing.

CFIm complexes may be involved in the 3′-end processing of other RNAs in paraspeckles. Electron microscopic observation has revealed the localization of CPSF6 in the interior area of the paraspeckle (Cardinale *et al*, 2007), where the NEAT1_1 isoform is not present (Souquere *et al*, 2010). Alternatively, the paraspeckle interior area may serve as the storage site of CFIm complexes.

Steps required for paraspeckle formation

Our results provide several important insights into paraspeckle formation. The plasmid rescue experiment clarified that NEAT1_2 but not NEAT1_1 is a necessary RNA component for *de novo* paraspeckle formation. This evidence supports our previous observations that: (1) SFPQ or NONO RNAi leads to paraspeckle disintegration as a consequence of NEAT1_2 destabilization (Sasaki *et al*, 2009) and (2) paraspeckles are observable solely in the NEAT1_2-expressing cells of mouse tissues (Nakagawa *et al*, 2011). By contrast, Shevtsov and Dundr (2011) reported that tethering NEAT1_1 at the specific chromosomal site triggers on-site paraspeckle formation. Clemson *et al* (2009) reported that NEAT1_1 overexpression in a stable cell line increased the number of nuclear paraspeckles. We observed a similar effect with NEAT1_1 overexpression, although NEAT1_2 overexpression increased the paraspeckle numbers more effectively. Because these experiments were performed in cells possessing intact paraspeckles with endogenous NEAT1_2, it is likely that locally concentrated NEAT1_1 captured the preexisting paraspeckles or their subparticles containing NEAT1_2, which resulted in the formation of paraspeckles containing exogenous NEAT1_1.

The two isoforms of NEAT1 are differentially localized within the paraspeckle: NEAT1_2 is present in the interior core of the paraspeckle, whereas NEAT1_1 is located at the peripheral area (Souquere *et al*, 2010). This observation is consistent with our hypothesis that the overexpressed NEAT1_1 is efficiently incorporated into the peripheral area of paraspeckles, whose core is constructed around endogenous NEAT1_2.

The existence of category 1B proteins argues that NEAT1_2 accumulation alone is insufficient for paraspeckle formation. An additional step involving category 1B proteins is required for intact paraspeckle formation subsequent to assembly of the primary NEAT1_2 subcomplex with category 1A proteins. Category 1B proteins may be involved in the assembly of a higher-order paraspeckle structure that is built with multiple copies of the NEAT1_2 subcomplex, as well as with the NEAT1_1 subcomplex (Figure 9B). Indeed, DAZAP1 in category 1B has been shown to interact with SFPQ in category 1A (Yang *et al*, 2009). We cannot rule out the possibility that category 1B proteins bind to unidentified essential RNA component(s) of the paraspeckle, because all category 1B proteins possess RRMs.

The role of NEAT1_1 remains obscure, despite its higher abundance compared with NEAT1_2. The RNAi results indicate that category 3A and 1B proteins contribute to NEAT1_1 accumulation, suggesting that NEAT1_1 forms subcomplexes with these proteins. Paraspeckles presumably are involved in the nuclear retention of specific mRNAs, which raises an interesting possibility: the more conserved NEAT1_1 RNA may serve as a functional unit for paraspeckle-conducting events (such as nuclear mRNA retention), rather than for its structural maintenance. In this way, NEAT1_1 synthesis could account for the amplification of the functional units at the paraspeckle periphery. Our trials to identify PSPs that mediate the nuclear retention of mRNAs by RNAi were unsuccessful, which suggests functional redundancy within the PSPs or the presence of additional unidentified factors.

On the basis of the data presented in this manuscript, we constructed a model of paraspeckle formation (Figure 9B). To understand the details of each process, it is important to map the RNA–protein and protein–protein interactions in this structure. Further studies will identify additional PSPs and RNAs, some of which may be critical for the paraspeckle structure. Indeed, additional paraspeckle-localized proteins that are not included in our list were recently reported (Bond and Fox, 2009). It will be important to investigate the connections among chromatin structure, transcription machinery, and paraspeckle formation. The ongoing transcription of NEAT1 was recently found to be a prerequisite for paraspeckle formation (Mao *et al*, 2011). Therefore, the initial step of paraspeckle formation may occur cotranscriptionally (Figure 9B). Further mechanistic investigations of NEAT1 ncRNA and PSPs should elicit a novel view of the formation of these tremendously large ribonucleoprotein particles and their linkage to function.

Materials and methods

Cell cultures and transfection

HeLa, HEK293T, MEF, and NIH3T3 cells were grown in DMEM (10% FBS). Some cells were treated with actinomycin D (0.3 μg/ml, 4 h). Transfection of MEF cells was performed with the Nucleofector MEF starter kit and the Nucleofector device (Ronza) or FuGene HD (Promega). Expression from the NEAT1_1 or NEAT1_2 construct in MEF was confirmed by RT–qPCR. Transfection of other cell lines was performed with Lipofectamine 2000 or Lipofectamine LTX (Invitrogen).

In vitro 3′-end processing assay

The ^{32}P-labelled RNA substrate was synthesized with SP6 RNA polymerase (TaKaRa). HNEs were prepared according to Dignam *et al* (1983). The *in vitro* RNA processing reaction was performed as described, with minor modifications (Takagaki *et al*, 1988). Briefly, ^{32}P-labelled RNA (2×10^4 c.p.m., ~5 fmol) was incubated in a 12.5-μl reaction mixture containing 8 mM HEPES (pH 7.9), 8% glycerol, 40 mM KCl, 0.2 mM PMSF, 0.4 mM DTT, 2.08 mM EDTA, 40 mM creatine phosphate, 40 μg/ml *Escherichia coli* tRNA, 0.25 U of RNasin (Promega), 2.5% polyvinyl alcohol, and 4 μl of HNE (32%). After the solution was reacted at 30°C for the indicated time, RNA was extracted and separated by 6% PAGE containing 7 M urea.

UV-crosslinking

UV-crosslinking was performed as described (Hirose *et al*, 2006) after incubating the samples under conditions of *in vitro* 3′-end processing. UV light (1.8 J/cm^2) was applied to an open-top reaction tube on ice with a UV-crosslinking device (CL-1000, UVP). RNase A and RNase T1 were added and incubated for 15 min at 37°C, and precipitation with 50% acetone was performed. Precipitated proteins were fractionated by SDS–PAGE. Substrate RNA was pre-incubated with r-K for 15 min at 30°C, after which HNE was added for an additional 15 min incubation. The r-K was expressed in *E. coli* cells [BL21(DE3)-CodonPlus RILP (Stratagene)] and purified with Ni affinity chromatography.

Gel mobility shift assay

WT and K-mut RNA oligonucleotides were chemically synthesized by Hokkaido System Science Co., Ltd. The oligonucleotide sequences used were as follows: WT: 5′-AAUCACUUUUCUUCCCCUUUACAGC ACAAA-3′ and K-mut: 5′-AAUCACUUUUCUUAAAAUUUACAGCAC AAA-3′. ^{32}P-labelled RNA probes were prepared with [γ-^{32}P]ATP and T4 polynucleotide kinase (TaKaRa). ^{32}P-labelled RNA (0.5×10^4 c.p.m., ~150 fmol) was incubated in a 12.5-μl reaction mixture containing 8 mM HEPES (pH 7.9), 8% glycerol, 40 mM KCl, 2.08 mM EDTA, 0.2 mM PMSF, 0.4 mM DTT, 40 mM creatine phosphate, 40 μg/ml *E. coli* tRNA, 0.25 U of RNasin (Promega), 2.5% polyvinyl alcohol, and 0.25–2.0 μg of r-K. The mixture was incubated at 30°C for 15 min, and the RNA–protein complexes were separated on 5% native PAGE.

Immunoprecipitation

For protein immunoprecipitation, HeLa cells (1×10^7 cells) were lysed with lysis buffer (50 mM Tris–HCl [pH 7.5], 150 mM NaCl, 50 mM NaF, 1 mM Na_3VO_4, and 0.5% NP40) for 30 min on ice, and the supernatant was recovered by centrifugation at 10 000 *g* for 10 min. For each immunoprecipitation experiment, cell extract (1 mg protein) was incubated overnight at 4°C with antibody-Dynabead conjugates (25 μl) in the presence of 10 μg/ml RNase A. The beads were washed five times with lysis buffer, and bound proteins were eluted by directly adding SDS loading buffer to the beads. For RNA immunoprecipitation, the cell extract prepared as described above was incubated with antibody-Dynabead conjugates without RNase A for 3 h at 4°C. Bound RNAs were extracted by directly adding Trizol reagent (Invitrogen) to the beads.

Immunoprecipitations from the *in vitro* processing samples were performed as described (Ideue *et al*, 2007). Briefly, *in vitro* processing was performed at a 10-fold scale (125 μl). The reaction mixture was diluted eight-fold with NET2 buffer (20 mM Tris–HCl [pH 7.5], 150 mM NaCl, and 0.5% NP40) immediately after RNase treatment and incubated with antibody-Dynabead conjugates for 16 h at 4°C. The beads were washed five times with NET2 buffer, and bound proteins were eluted by directly adding SDS loading buffer to the beads. Information about the antibodies used is shown in Supplementary Table S4.

Protein-binding assay with purified recombinant proteins

The recombinant CFIm complex was purified basically as described (Arias-Palomo *et al*, 2011). Briefly, 4×10^7 HEK293T cells were cotransfected with pEF_Flag–HA–SBP–NUDT21 and pEF_Flag–HA–CPSF6 plasmids at a 1:3 ratio using Lipofectamine LTX (Invitrogen). After 48 h, the cells were lysed by sonication in F-lysis buffer (20 mM Tris–HCl [pH 7.5], 150 mM NaCl, 10% [w/v] sucrose, 1% Triton X-100, 0.5% NP40, 1 mM DTT, protease inhibitor cocktail [Roche], and phosphatase inhibitor cocktail [Roche]). They were incubated for 30 min on ice, and the soluble fraction was recovered by centrifugation at 15 000 *g* for 30 min. The soluble fraction was precleared with Sepharose 4B (Sigma), incubated with streptavidin Sepharose (GE Healthcare) for 2 h at 4°C with gentle rotation, and washed with F-lysis buffer. The affinity-purified SBP-tagged CFIm complex was eluted by incubation with F-lysis buffer containing 2 mM biotin (Sigma) for 30 min at 4°C.

The SBP-tagged CFIm complexes were immobilized to Dynabeads MyOne Streptavidin beads (Invitrogen) for 2 h at 4°C with gentle rotation and washed with NET2 buffer. Beads suspended in 30 μl of NET2 buffer were mixed with 10, 30, or 90 pmol of recombinant His-tagged HNRNPK for 15 min at RT. The bead suspension was further incubated for 2 h at 4°C, and the supernatant was collected. Beads were washed with NET2 buffer five times. Bound proteins were eluted by directly adding SDS loading buffer to the beads.

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (http://www.embojournal.org).

Acknowledgements

We thank Y Kisu and H Mochizuki for their help with data mining of the intracellular localization of Venus-fusion proteins, A Yamashita for kindly providing the plasmids and host cell lines for the preparation of recombinant proteins, and Y Hirose for his feedback on the *in vitro* RNA processing reaction. JA Steitz and K Tycowski are acknowledged for their critical reading of the manuscript. We thank M Nagai and the members of the Hirose laboratory for valuable discussions. This research was supported by the Funding Program for Next Generation World-Leading Researchers (NEXT Program) of the Japan Society for the Promotion of Science (JSPS), and by grants from the New Energy and Industrial Technology Development Organization (NEDO), Ministry of Education, Culture, Sports, Science, and Technology of Japan (MEXT), the Astellas Foundation for Research on Metabolic Disorders, and Takeda Science Foundation.

Author contributions: The study was conceived and supervised by TN and TH; TN performed most of the experiments. SN performed the experiment in Figure 1A. AT and YFS contributed with biochemical and microscope analyses. TN and TH wrote the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Arias-Palomo E, Yamashita A, Fernández IS, Núñez-Ramírez R, Bamba Y, Izumi N, Ohno S, Llorca O (2011) The nonsense-mediated mRNA decay SMG-1 kinase is regulated by large-scale conformational changes controlled by SMG-8. *Genes Dev* **25:** 153–164

Auboeuf D, Dowhan DH, Dutertre M, Martin N, Berget SM, O'Malley BW (2005) A subset of nuclear receptor coregulators act as coupling proteins during synthesis and maturation of RNA transcripts. *Mol Cell Biol* **25:** 5307–5316

Bomsztyk K, Denisenko O, Ostrowski J (2004) hnRNP K: one protein multiple processes. *Bioessays* **26:** 629–638

Bond CS, Fox AH (2009) Paraspeckles: nuclear bodies built on long noncoding RNA. *J Cell Biol* **186:** 637–644

Brown RS (2005) Zinc finger proteins: getting a grip on RNA. *Curr Opin Struct Biol* **15:** 94–98

Burd CG, Dreyfuss G (1994) Conserved structures and diversity of functions of RNA-binding proteins. *Science* **265:** 615–621

Cardinale S, Cisterna B, Bonetti P, Aringhieri C, Biggiogera M, Barabino SM (2007) Subnuclear localization and dynamics of the Pre-mRNA 3′ end processing factor mammalian cleavage factor I 68-kDa subunit. *Mol Biol Cell* **18:** 1282–1292

Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, Oyama R, Ravasi T, Lenhard B, Wells C, Kodzius R, Shimokawa K, Bajic VB, Brenner SE, Batalov S, Forrest AR, Zavolan M, Davis MJ, Wilming LG, Aidinis V *et al* (2005)

The transcriptional landscape of the mammalian genome. *Science* **309:** 1559–1563

Chen LL, Carmichael GG (2009) Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. *Mol Cell* **35:** 467–478

Chen LL, DeCerbo JN, Carmichael GG (2008) Alu element-mediated gene silencing. *EMBO J* **27:** 1694–1705

Chen Y, Zhou X, Liu N, Wang C, Zhang L, Mo W, Hu G (2008) Arginine methylation of hnRNP K enhances p53 transcriptional activity. *FEBS Lett* **582:** 1761–1765

Clemson CM, Hutchinson JN, Sara SA, Ensminger AW, Fox AH, Chess A, Lawrence JB (2009) An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol Cell* **33:** 717–726

Clemson CM, McNeil JA, Willard HF, Lawrence JB (1996) XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. *J Cell Biol* **132:** 259–275

Dettwiler S, Aringhieri C, Cardinale S, Keller W, Barabino SM (2004) Distinct sequence motifs within the 68-kDa subunit of cleavage factor Im mediate RNA binding, protein-protein interactions, and subcellular localization. *J Biol Chem* **279:** 35788–35797

Dignam JD, Lebovitz RM, Roeder RG (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res* **11:** 1475–1489

Fox AH, Lam YW, Leung AK, Lyon CE, Andersen J, Mann M, Lamond AI (2002) Paraspeckles: a novel nuclear domain. *Curr Biol* **12:** 13–25

Guru SC, Agarwal SK, Manickam P, Olufemi SE, Crabtree JS, Weisemann JM, Kester MB, Kim YS, Wang Y, Emmert-Buck MR, Liotta LA, Spiegel AM, Boguski MS, Roe BA, Collins FS, Marx SJ, Burns L, Chandrasekharappa SC (1997) A transcript map for the 2.8-Mb region containing the multiple endocrine neoplasia type 1 locus. *Genome Res* **7:** 725–735

Hirose T, Ideue T, Nagai M, Hagiwara M, Shu MD, Steitz JA (2006) A spliceosomal intron binding protein, IBP160, links position-dependent assembly of intron-encoded box C/D snoRNP to pre-mRNA splicing. *Mol Cell* **23:** 673–684

Huarte M, Guttman M, Feldser D, Garber M, Koziol MJ, Kenzelmann-Broz D, Khalil AM, Zuk O, Amit I, Rabani M, Attardi LD, Regev A, Lander ES, Jacks T, Rinn JL (2010) A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* **142:** 409–419

Hutchinson JN, Ensminger AW, Clemson CM, Lynch CR, Lawrence JB, Chess A (2007) A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics* **8:** 39

Ideue T, Sasaki YT, Hagiwara M, Hirose T (2007) Introns play an essential role in splicing-dependent formation of the exon junction complex. *Genes Dev* **21:** 1993–1998

Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, Willingham AT, Stadler PF, Hertel J, Hackermüller J, Hofacker IL, Bell I, Cheung E, Drenkow J, Dumais E, Patel S, Helt G, Ganesh M, Ghosh S, Piccolboni A, Sementchenko V *et al* (2007) RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* **316:** 1484–1488

Kim N, Kim P, Nam S, Shin S, Lee S (2006) ChimerDB—a knowledgebase for fusion sequences. *Nucleic Acids Res* **34:** D21–D24

Kim S, Yamamoto J, Chen Y, Aida M, Wada T, Handa H, Yamaguchi Y (2010) Evidence that cleavage factor Im is a heterotetrameric protein complex controlling alternative polyadenylation. *Genes Cells* **15:** 1003–1013

Lagier-Tourenne C, Cleveland DW (2009) Rethinking ALS: the FUS about TDP-43. *Cell* **136:** 1001–1004

Law WJ, Cann KL, Hicks GG (2006) TLS, EWS and TAF15: a model for transcriptional integration of gene expression. *Brief Funct Genomic Proteomic* **5:** 8–14

Lutz CS (2008) Alternative polyadenylation: a twist on mRNA 3′-end formation. *ACS Chem Biol* **3:** 609–617

Mao YS, Sunwoo H, Zhang B, Spector DL (2011) Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. *Nat Cell Biol* **13:** 95–101

Maruyama Y, Kawamura Y, Nishikawa T, Isogai T, Nomura N, Goshima N (2012) HGPD: Human Gene and Protein Database, 2012 update. *Nucleic Acids Res* **40:** D924–D929

Mercer TR, Dinger ME, Mattick JS (2009) Long non-coding RNAs: insights into functions. *Nat Rev Genet* **10:** 155–159

Millevoi S, Vagner S (2010) Molecular mechanisms of eukaryotic pre-mRNA 3′-end processing regulation. *Nucleic Acids Res* **38:** 2757–2774

Nakagawa S, Naganuma T, Shioi G, Hirose T (2011) Paraspeckles are subpopulation-specific nuclear bodies that are not essential in mice. *J Cell Biol* **193:** 31–39

Prasanth KV, Prasanth SG, Xuan Z, Hearn S, Freier SM, Bennett CF, Zhang MQ, Spector DL (2005) Regulating gene expression through RNA nuclear retention. *Cell* **123:** 249–263

Prasanth KV, Spector DL (2007) Eukaryotic regulatory RNAs: an answer to the 'genome complexity' conundrum. *Genes Dev* **21:** 11–42

Sasaki YT, Ideue T, Sano M, Mituyama T, Hirose T (2009) MENepsilon/beta noncoding RNAs are essential for structural integrity of nuclear paraspeckles. *Proc Natl Acad Sci USA* **106:** 2525–2530

Shav-Tal Y, Blechman J, Darzacq X, Montagna C, Dye BT, Patton JG, Singer RH, Zipori D (2005) Dynamic sorting of nuclear components into distinct nucleolar caps during transcriptional inhibition. *Mol Biol Cell* **16:** 2395–2413

Shevtsov SP, Dundr M (2011) Nucleation of nuclear bodies by RNA. *Nat Cell Biol* **13:** 167–173

Sone M, Hayashi T, Tarui H, Agata K, Takeichi M, Nakagawa S (2007) The mRNA-like noncoding RNA Gomafu constitutes a novel nuclear domain in a subset of neurons. *J Cell Sci* **120:** 2498–2506

Souquere S, Beauclair G, Harper F, Fox A, Pierron G (2010) Highly ordered spatial organization of the structural long noncoding NEAT1 RNAs within paraspeckle nuclear bodies. *Mol Biol Cell* **21:** 4020–4027

Spector DL (2006) SnapShot: cellular bodies. *Cell* **127:** 1071

Sunwoo H, Dinger ME, Wilusz JE, Amaral PP, Mattick JS, Spector DL (2009) MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res* **19:** 347–359

Takagaki Y, Ryner LC, Manley JL (1988) Separation and characterization of a poly(A) polymerase and a cleavage/specificity factor required for pre-mRNA polyadenylation. *Cell* **52:** 731–742

Thisted T, Lyakhov DL, Liebhaber SA (2001) Optimized RNA targets of two closely related triple KH domain proteins, heterogeneous nuclear ribonucleoprotein K and alphaCP-2KL, suggest Distinct modes of RNA recognition. *J Biol Chem* **276:** 17484–17496

Tollervey JR, Curk T, Rogelj B, Briese M, Cereda M, Kayikci M, König J, Hortobágyi T, Nishimura AL, Zupunski V, Patani R, Chandran S, Rot G, Zupan B, Shaw CE, Ule J (2011) Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat Neurosci* **14:** 452–458

Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, Freier SM, Bennett CF, Sharma A, Bubulya PA, Blencowe BJ, Prasanth SG, Prasanth KV (2010) The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell* **39:** 925–938

Venkataraman K, Brown KM, Gilmartin GM (2005) Analysis of a noncanonical poly(A) site reveals a tripartite mechanism for vertebrate poly(A) site recognition. *Genes Dev* **19:** 1315–1327

Wang KC, Chang HY (2011) Molecular mechanisms of long noncoding RNAs. *Mol Cell* **43:** 904–914

Wang X, Arai S, Song X, Reichart D, Du K, Pascual G, Tempst P, Rosenfeld MG, Glass CK, Kurokawa R (2008) Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. *Nature* **454:** 126–130

Yang HT, Peggie M, Cohen P, Rousseau S (2009) DAZAP1 interacts via its RNA-recognition motifs with the C-termini of other RNA-binding proteins. *Biochem Biophys Res Commun* **380:** 705–709

Yang L, Lin C, Liu W, Zhang J, Ohgi KA, Grinstein JD, Dorrestein PC, Rosenfeld MG (2011) ncRNA- and Pc2 methylation-dependent gene relocation between nuclear structures mediates gene activation programs. *Cell* **147:** 773–788

Yang Q, Gilmartin GM, Doublié S (2010) Structural basis of UGUA recognition by the Nudix protein CFI(m)25 and implications for a regulatory role in mRNA 3′ processing. *Proc Natl Acad Sci USA* **107:** 10062–10067

Zheng R, Shen Z, Tripathi V, Xuan Z, Freier SM, Bennett CF, Prasanth SG, Prasanth KV (2010) Polypurine-repeat-containing RNAs: a novel class of long non-coding RNA in mammalian cells. *J Cell Sci* **123:** 3734–3744

FEBS Letters xxx (2012) xxx–xxx

ELSEVIER

FEBS Letters

journal homepage: www.FEBSLetters.org

Expression screening of 17q12–21 amplicon reveals GRB7 as an ERBB2-dependent oncogene

Makoto Saito [a], Yukiko Kato [a], Emi Ito [b], Jiro Fujimoto [a], Kosuke Ishikawa [a,c], Ayano Doi [a], Kentaro Kumazawa [a], Atsuka Matsui [a], Shiori Takebe [a], Takaomi Ishida [a], Sakura Azuma [a], Hiromi Mochizuki [c], Yoshifumi Kawamura [c], Yuka Yanagisawa [b,d], Reiko Honma [b,d], Jun-ichi Imai [b], Hirokazu Ohbayashi [e], Naoki Goshima [f], Kentaro Semba [a,*], Shinya Watanabe [b]

[a] Department of Life Science and Medical Bioscience, School of Advanced Science and Engineering, Waseda University, 2-2 Wakamatsu-cho, Shinjuku-ku, Tokyo 162-8480, Japan
[b] Department of Clinical Genomics, Translational Research Center (Tokyo Branch), Fukushima Medical University, Shibuya-ku, Tokyo 151-0051, Japan
[c] Japan Biological Informatics Consortium (JBiC), 2-45 Aomi, Koto-ku, Tokyo 135-8073, Japan
[d] Nippon Gene, Co., Ltd., Kandanishiki-cho, Chiyoda-ku, Tokyo 101-0054, Japan
[e] Nichirei Biosciences Inc., Biosciences Research & Development Center, 1-52-14 Kumegawa-cho, Higashimurayama-shi, Tokyo 189-0003, Japan
[f] Biomedicinal Information Research Center (BIRC), National Institute of Advanced Industrial Science and Technology (AIST), 2-42 Aomi, Koto-ku, Tokyo 135-0064, Japan

ARTICLE INFO

Article history:
Received 10 April 2012
Revised 30 April 2012
Accepted 2 May 2012
Available online xxxx

Edited by Takashi Gojobori

Keywords:
Gene amplification
Expression screening
Tumorigenesis
GRB7
ERBB2

ABSTRACT

Gene amplification is a major genetic alteration in human cancers. Amplicons, amplified genomic regions, are believed to contain "driver" genes responsible for tumorigenesis. However, the significance of co-amplified genes has not been extensively studied. We have established an integrated analysis system of amplicons using retrovirus-mediated gene transfer coupled with a human full-length cDNA set. Applying this system to 17q12–21 amplicon observed in breast cancer, we identified GRB7 as a context-dependent oncogene, which modulates the ERBB2 signaling pathway through enhanced phosphorylation of ERBB2 and Akt. Our work provides an insight into the biological significance of gene amplification in human cancers.

1. Introduction

DNA amplification is a major genetic alteration contributing to oncogenesis [1]. Historically, some proto-oncogenes identified as cellular counterparts of retroviral oncogenes were found to be amplified in human cancers. Thus, it is thought that unidentified proto-oncogenes exist in amplified genomic regions called "amplicons", and amplified proto-oncogenes express large amount of proteins, leading to oncogenesis. We previously constructed gene expression maps of chromosomes in human breast cancer cell lines and extracted six novel amplicons [2]. Nevertheless, it is not easy to identify such proto-oncogenes because amplification events often include multiple genes, and because more information is required, including precise mapping of amplified regions in multiple cancers and deduced function of each gene. A few findings provided the significance of co-amplified genes except specific oncogenes in the amplicon in terms of cancer cell phenotypes. In non-small-cell lung cancer, for instance, co-amplification of *TTF-1* and *NKX2-8* in the 14q13.3 amplicon renders cancer cells resistance to cisplatin [3]. However, the biological significance of co-amplification for oncogenesis has not been validated extensively.

In this study, we focused on the functions of co-amplified genes localized in the 17q12–21 amplicon containing *ERBB2* as a driver gene [4]. The *ERBB2* amplicon is observed in 25% of breast cancers, and also in ovarian, gastric and esophagus cancers [5]. Clinicopathological data indicate that ERBB2 expression is a poor prognostic factor [6]. Furthermore, an active *ErbB2* mutant called *neu* oncogene causes cellular transformation of NIH3T3 cells [7] and breast cancer in trans-

Abbreviations: GRB7, growth factor receptor-bound protein 7; ERBB2, v-erb-b2 erythroblastic leukemia viral oncogene homolog; MMTV, mouse mammary tumor virus; CMV, cytomegalovirus; MAPK, mitogen-activated protein kinase; MEK, MAPK extracellular signal-regulated kinase; ERK, extracellular signal-regulated kinase; IGF1, insulin-like growth factor-1

* Corresponding author. Fax: +81 3 5369 7320.
E-mail address: ksemba@waseda.jp (K. Semba).

0014-5793/$36.00 © 2012 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.febslet.2012.05.003

Please cite this article in press as: Saito, M., et al. Expression screening of 17q12–21 amplicon reveals GRB7 as an ERBB2-dependent oncogene. FEBS Lett. (2012), http://dx.doi.org/10.1016/j.febslet.2012.05.003

genic mice expressing *neu* oncogene under the control of MMTV promoter [8]. These results support the idea that *ERBB2* functions as a driver gene for oncogenesis when mutated or overexpressed.

To examine the function of co-amplified genes in the *ERBB2* amplicon, we tested these genes for enhancement of ERBB2 cellular transforming activity. For this purpose, we introduced a human wild-type (WT) *ERBB2* expression vector into NIH3T3 cells and established "non-transformed" cells moderately expressing ERBB2 under the CMV promoter. Then, human full-length cDNAs of co-amplified genes were retrovirally introduced into ERBB2-expressing NIH3T3 cells, and the transforming activity was assessed by focus formation assays. Our screening system ensures multiple expression of complete proteins in cells, thereby enabling the examination of combinations of co-amplified genes in the *ERBB2* amplicon.

Here, we show *GRB7* gene, which is located about 10 kb from *ERBB2* locus and is frequently co-amplified with *ERBB2*, caused efficient transformation of NIH3T3 in concert with *ERBB2*; this ERBB2-dependent transforming activity was specific for GRB7 among the GRB7 family proteins. This was consistent with the observation that phosphorylation of ERBB2 was increased when GRB7 was expressed, but not other GRB7 family proteins. Importantly, phosphorylation of Akt at Thr308 and Ser473 was upregulated when both ERBB2 and GRB7 were expressed. We further examined the transforming activity of a series of GRB7 mutants, and showed that a BPS region deletion mutant of GRB7 is a potent activator of ERBB2 and Akt, while the transforming activities of other domain deletion mutants are severely impaired.

Our model of collaborative transformation by ERBB2 and GRB7 proposes that GRB7 is a cytoplasmic activator and adaptor of ERBB2, enhances ERBB2 phosphorylation, and connects ERBB2 to Akt. Our analysis also highlights the biological significance of gene amplification in terms of simultaneous overexpression of a driver gene and "supporter" genes.

2. Materials and methods

2.1. DNA constructs and antibodies

To construct an ERBB2 expression vector, human ERBB2 (RefSeq: NM_004448) cDNA was inserted in pQCXIN retroviral vector (Clontech, Mountain View, CA). Human full-length cDNAs were obtained from the human proteome expression resource (HuPEX) [9], and cloned into pMXs retroviral vector [10] using the Gateway Cloning system (Life Technologies, Carlsbad, CA) with or without N-terminal FLAG epitope tag. Venus fluorescent protein [11] was used as a control. Primary antibodies were as follows: anti-α-tubulin DM1A (Sigma, St. Louis, MO), anti-FLAG M2 (Sigma), anti-ERBB2 (SV2-61γ, Nichirei Bioscience, Japan), anti-HER2/ErbB2 (#2242, Cell Signaling Technology (CST), Danvers, MA), anti-Phospho-HER2/ErbB2 (Tyr877) (CST#2241), anti-Phospho-HER2/ErbB2 (Tyr1221/1222) (CST#2243), anti-Phospho-HER2/ErbB2 (Tyr1248) (CST#2247), anti-Erk1/2 (CST#4695), anti-Phospho-Erk1/2 (#4370), anti-Phospho-MEK1/2 (CST#9154), anti-Akt (pan) (#4691), anti-Phospho-Akt (Thr308) (CST#2965), anti-Phospho-Akt (Ser473) (CST#4060), and anti-Phosphotyrosine 4G10 (Millipore, Billerica, MA). Secondary antibodies for western blotting were purchased from GE Healthcare (Piscataway, NJ).

2.2. Cell culture

NIH3T3 cells were obtained from RIKEN Cell Bank (Tsukuba, Japan) and cultured in DMEM supplemented with 5% heat-inactivated calf serum, 100 U/ml penicillin, and 100 μg/ml streptomycin at 37 °C and 5% CO_2. 3T3-ERBB2 cells were established by retroviral infection of pQCXIN-ERBB2 and selection with 1 mg/ml G418. Plat-E packaging cells were obtained from T. Kitamura (Institute of Medical Science, University of Tokyo), and cultured

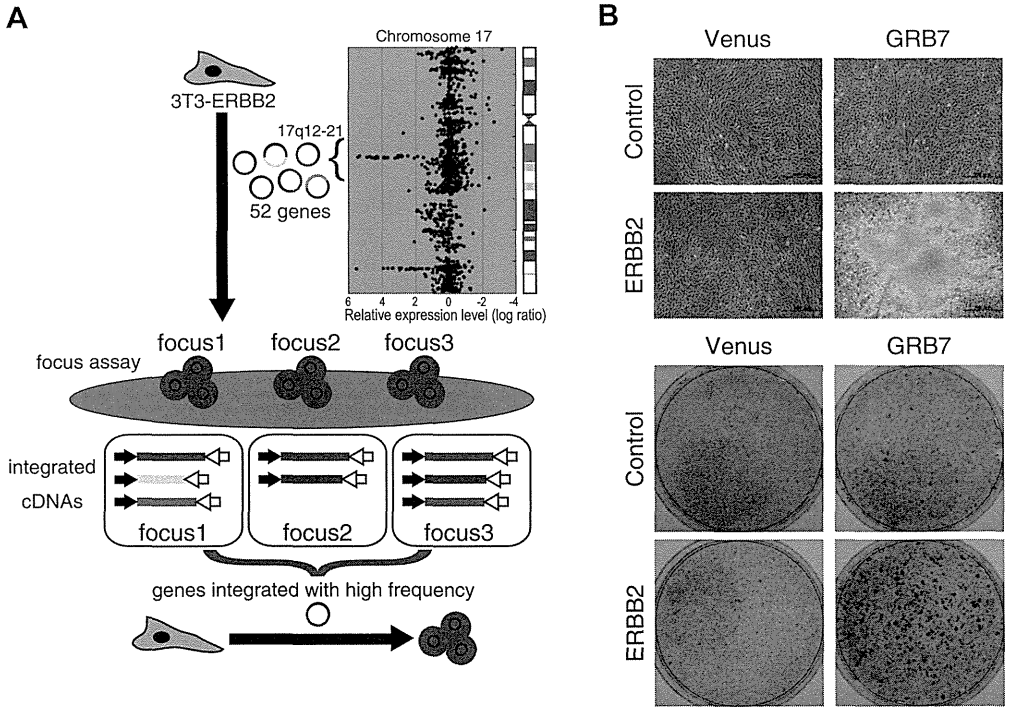

Fig. 1. ERBB2-dependent transforming activity of GRB7. (A) A scheme of oncogene-screening of ERBB2 amplicon. (B) Focus formation assays with ERBB2 and GRB7. 3T3-ERBB2 were infected with GRB7, and cultured for 17 days (upper). Scale bar, 500 μm. Cells were fixed and stained with crystal violet (lower). Venus fluorescent protein was used as a control. Volume of virus stock used in this experiment was 125 μl.

Please cite this article in press as: Saito, M., et al. Expression screening of 17q12–21 amplicon reveals GRB7 as an ERBB2-dependent oncogene. FEBS Lett. (2012), http://dx.doi.org/10.1016/j.febslet.2012.05.003

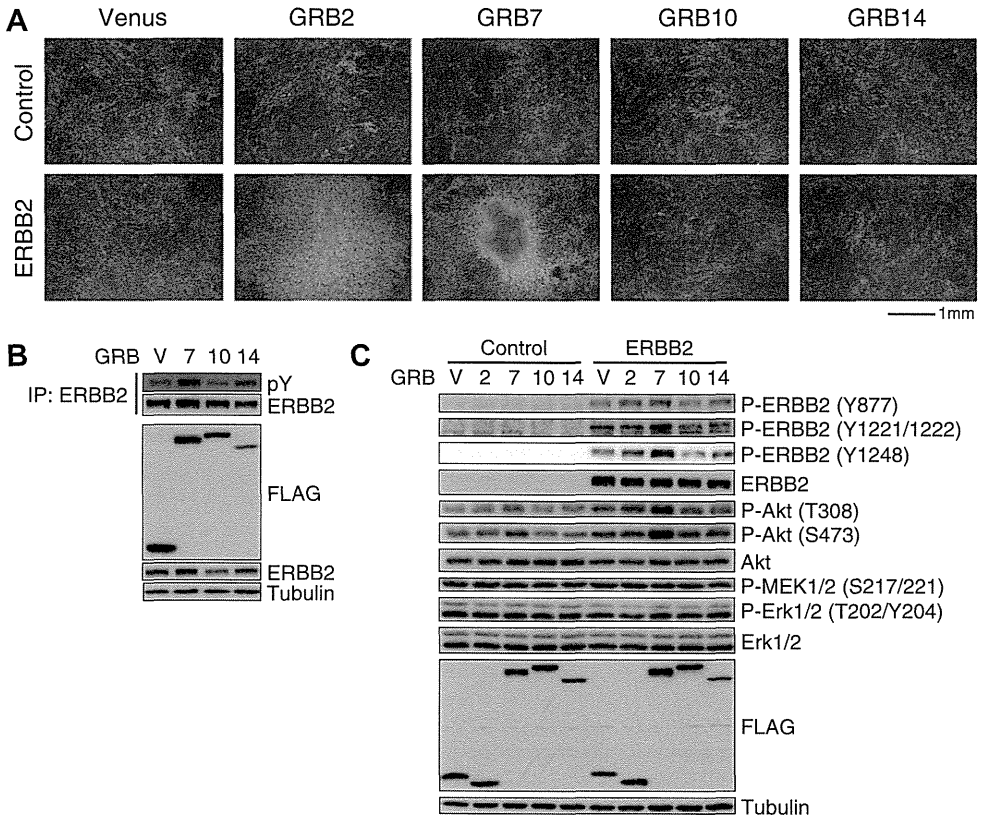

Fig. 2. ERBB2-dependent Akt activation and transformation by GRB7. (A) Focus formation assays with GRB proteins and ERBB2. 3T3-ERBB2 were infected with each GRB7 family protein or GRB2, and cultured for 20 days. Scale bar, 1 mm. (B) Phosphorylation status of ERBB2 in GRB7 family-expressing cells. ERBB2 immunoprecipitates prepared from GRB7 family-expressing cells were analyzed by anti-phosphotyrosine antibody. (C) Effect of expression of ERBB2 with GRB proteins. Phosphorylation status of ERBB2 and components of MAPK and Akt pathways were analyzed by phosphorylation-specific antibodies. Distinct amount of virus stock was used as follows: Venus (30 μl), GRB2 (480 μl), GRB7 (90 μl), GRB10 (360 μl), GRB14 (480 μl) for control NIH3T3, and Venus (20 μl), GRB2 (360 μl), GRB7 (60 μl), GRB10 (240 μl), GRB14 (360 μl) for 3T3-ERBB2.

in DMEM supplemented with 10% heat-inactivated FBS, penicillin, and streptomycin, as above.

2.3. Retroviral packaging

Plat-E cells were seeded into 6-cm culture dishes at a density of 1.0×10^6 and transfected with 4 μg of retroviral plasmids mixed with 10 μl of Lipofectamine 2000 reagent (Life Technologies). One day after transfection, the culture supernatant was replaced with NIH3T3 media. The following day, culture supernatant was harvested and centrifuged to remove cell debris, then stored at −80 °C as virus stock.

2.4. Focus formation assays

NIH3T3 cells were seeded into 12-well culture plates at a concentration of 5.0×10^4 and the following day, the cells were infected with 1 ml of appropriately diluted virus stock containing 8 μg/ml polybrene. To carefully adjust the expression level of each protein, distinct amount of virus stock were used. The next day, cells were seeded into 10-cm culture dishes and cultured for several days. Then, infected cells (1.0×10^5) were seeded again into 10-cm culture dishes and maintained for 2–3 weeks with medium change every 2 days. Cell foci were stained with 0.05% crystal violet and counted.

2.5. Recovery of cDNAs from NIH3T3 transformants

Transformed NIH3T3 cells were picked from each focus and cultured to isolate genomic DNA. cDNAs inserted into genome were amplified with a pair of primers for pMX vector (pMXs-s1811: GACGGCATCGCAGCTTGGATA and pMXs-AS3200: TTATCGTCGACCACTGTGCTG) and KOD FX polymerase (Toyobo, Japan) for 30 cycles of 98 °C for 10 s, 55 °C for 15 s, and 68 °C for 2 min, and subcloned into pBluescript SK(−) (Agilent Technologies, Santa Clara, CA). Two hundred and thirteen informative sequences were obtained by an Applied Biosystems 3130 genetic analyzer (Life Technologies).

2.6. Constructs of GRB7 mutants

To construct expression vectors for GRB7 mutants, pMXs-FLAG-GRB7 plasmid was amplified with an appropriate set of primers, then the PCR product was self-ligated to obtain a mutated plasmid [12].

2.7. Western blotting and immunoprecipitation

Cells were rinsed in ice-cold PBS and lysed in RIPA buffer (10 mM Tris–HCl, pH8.0, 150 mM NaCl, 1 mM EDTA, 1% NP-40, 0.1% sodium deoxycholate, 0.1% SDS, 1 mM Na_3VO_4, 10 mM NaF, 17.5 mM β-glycerophosphate, and 1 mM PMSF) for analysis of total cell lysates, or Triton lysis buffer (50 mM Tris–HCl, pH 8.0, 135 mM NaCl, 1 mM EDTA, 1% TritonX-100, 10% glycerol, 1 mM Na_3VO_4, 10 mM NaF, and 1 mM PMSF) for immunoprecipitation. Protein amounts used in Figs. 2B and 4 were 225 and 500 μg, respectively. Protein lysates were incubated with 2 μg anti-ERBB2 (SV2-61γ) for 1 h at 4 °C, and immunoprecipitated with 20 μl of protein A-Se-

Please cite this article in press as: Saito, M., et al. Expression screening of 17q12–21 amplicon reveals GRB7 as an ERBB2-dependent oncogene. FEBS Lett. (2012), http://dx.doi.org/10.1016/j.febslet.2012.05.003

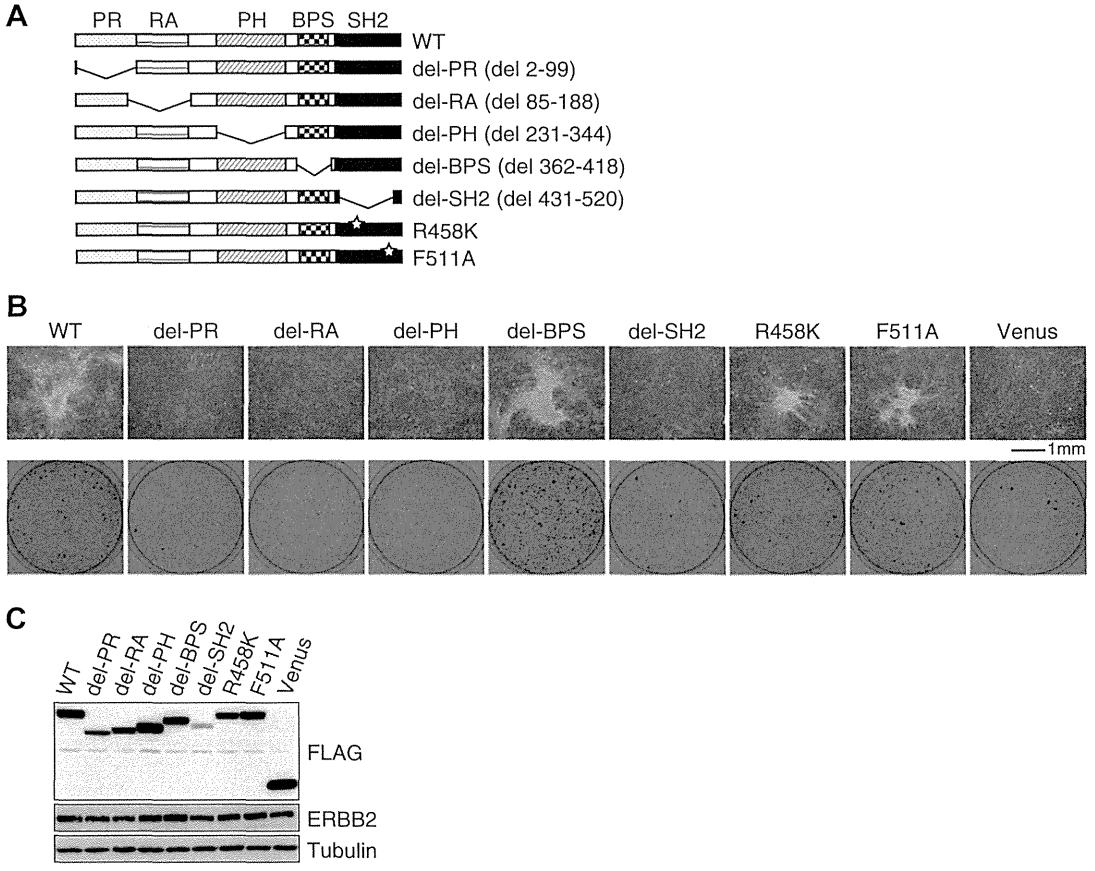

Fig. 3. Transforming activity of GRB7 domain deletion and point mutants. (A) Structures of the GRB7 mutants used in the experiments. (B) Focus formation assays with GRB7 mutants. 3T3-ERBB2 were infected with WT or mutant GRB7, and then cultured for 16 days (upper). Scale bar, 1 mm. Cells were fixed and stained with crystal violet (lower). (C) Expression level of ERBB2 and WT or mutant GRB7. Distinct amount of virus stock was used as follows: GRB7 WT (250 μl), del-PR (500 μl), del-RA (500 μl), del-PH (250 μl), del-BPS (250 μl), del-SH2 (500 μl), R458 K (500 μl), F511A (250 μl), Venus (250 μl).

pharose (GE Healthcare) for 1 h at 4 °C, then washed. Samples were boiled for 5 min in SDS–PAGE sample buffer, and separated with 7.5% or 10% acrylamide gels. Proteins were transferred onto PVDF membranes (Immobilon-P, Millipore), rinsed in TBS (20 mM Tris–HCl (pH 7.5), and 150 mM NaCl) and incubated in blocking buffer (TBS containing 3–5% non-fat dry milk or 5% bovine serum albumin) for 1 h at RT or overnight at 4 °C. The membranes were incubated for 1 h at RT or overnight at 4 °C with primary antibodies. Proteins were labeled using HRP-conjugated secondary antibodies diluted at 1:2000 for 1 h at RT, then visualized by the enhanced chemiluminescence method using Immobilon Western reagent (Millipore).

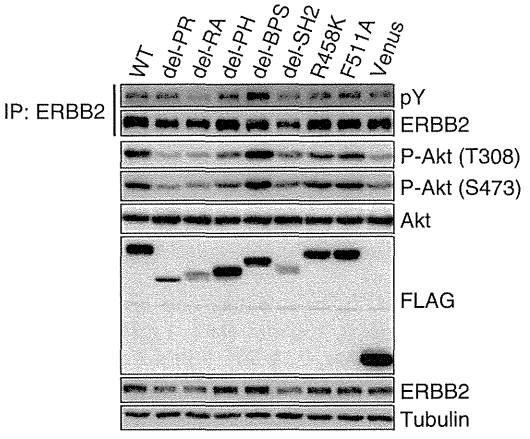

Fig. 4. Effect of GRB7 mutants on Akt phosphorylation. Lysates from 3T3-ERBB2, WT and mutant GRB7 were immunoblotted using anti-phospho-Akt-specific antibodies. ERBB2 immunoprecipitates were also analyzed to assess total phosphorylation by anti-phosphotyrosine antibody. Amounts of virus stock used are shown in Fig. 3.

3. Results

To screen for ERBB2-dependent transforming genes, we established NIH3T3 derivative (3T3-ERBB2) moderately expressing human WT ERBB2 under the control of CMV promoter. Mixtures of full-length human cDNA clones corresponding to the 52 genes (Table 1) within the *ERBB2* amplicon were cloned into pMX retroviral vectors, and then retroviral expression vectors were introduced into Plat-E packaging cells. 3T3-ERBB2 was infected with the resultant retrovirus mixture. Two to three weeks after infection, 30 foci were identified, whereas no foci were observed for 3T3-ERBB2 without retroviral infection. Each focus was isolated, expanded, and lysed for preparation of genomic DNA to analyze integrated cDNAs (Fig. 1A, Supplementary Table 1).

To test the transforming activity of most frequently recovered cDNAs, 12 representative cDNAs (Table 1) were individually introduced into 3T3-ERBB2. As a result, only GRB7 (RefSeq: NM_001030002) reproducibly induced foci. Transformed cells were smaller than normal cells and piled up on one another. The

Please cite this article in press as: Saito, M., et al. Expression screening of 17q12–21 amplicon reveals GRB7 as an ERBB2-dependent oncogene. FEBS Lett. (2012), http://dx.doi.org/10.1016/j.febslet.2012.05.003

Table 1
Fifty-two genes analyzed in this study.

Gene symbol	RefSeq ID	Frequency
C17orf78	**NM_173625**	**9**
TADA2A	NM_001166105	5
DUSP14	**NM_007026**	**9**
SYNRG	NM_001163546	0
DDX52	NM_007010	2
HNF1B	NM_000458	4
MRPL45	NM_032351	0
SOCS7	NM_014598	1
SRCIN1	NM_025248	0
MLLT6	NM_005937	0
PCGF2	NM_007144	5
PSMB3	NM_002795	3
PIP4K2B	NM_003559	0
CWC25	NM_017748	2
RPL23	NM_000978	6
LASP1	NM_006148	0
PLXDC1	NM_020405	4
CACNB1	NM_000723	3
RPL19	NM_000981	5
FBXL20	NM_032875	0
MED1	NM_004774	0
PPP1R1B	NM_032192	2
STARD3	**NM_001165937**	**11**
TCAP	**NM_003673**	**21**
PNMT	**NM_002686**	**11**
PGAP3	**NM_033419**	**16**
ERBB2	NM_004448	0
C17orf37	**NM_032339**	**7**
GRB7	**NM_001030002**	**11**
IKZF3	NM_012481	2
GSDMB	NM_001165959	0
ORMDL3	NM_139280	5
GSDMA	NM_178171	0
PSMD3	NM_002809	5
CSF3	**NM_172219**	**10**
MED24	NM_014815	1
THRA	**NM_001190919**	**7**
NR1D1	NM_021724	0
MSL1	NM_001012241	2
RAPGEFL1	NM_016339	2
WIPF2	NM_133264	0
CDC6	NM_001254	0
RARA	**NM_000964**	**9**
LOC100131821	AK123052	4
IGFBP4	NM_001552	6
TNS4	**NM_032865**	**13**
CCR7	NM_001838	2
SMARCE1	NM_003079	0
KRT24	NM_019016	2
KRT25	NM_181534	3
KRT28	NM_181535	3
TMEM99	NM_145274	0

Genes are aligned in order of chromosome location. Frequency of genes appeared in 30 foci are shown. 12 cDNAs individually introduced into 3T3-ERBB2 are shown in bold.

Table 2
Focus formation assay with GRB proteins and ERBB2.

	Number of foci after 20 days (10 cm-culture dish)				
	Venus	GRB2	GRB7	GRB10	GRB14
Experiment 1					
Control	0	0	0	0	0
ERBB2	0	2	6	0	0
Experiment 2					
Control	1	0	0	0	0
ERBB2	0	5	13	0	0

Distinct amount of virus stock were used as follows: Venus (30 μl), GRB2 (480 μl), GRB7 (90 μl), GRB10 (360 μl), GRB14 (480 μl) for control NIH3T3, and Venus (20 μl), GRB2 (360 μl), GRB7 (60 μl), GRB10 (240 μl), GRB14 (360 μl) for 3T3-ERBB2. Expression level of each protein was analyzed in Fig. 2C.

appearance of GRB7-transformed foci was smaller and higher than that of H-Ras(G12V)-transformed foci (data not shown). GRB7 induced transformation of 3T3-ERBB2, but not of parental NIH3T3 (Fig. 1B), indicating that transformation of NIH3T3 by GRB7 is dependent on ERBB2 expression.

GRB7 is one of the GRB7 family proteins, which are known to bind receptor tyrosine kinases and mediate signal to downstream effectors [13]. We carried out focus formation assays to determine whether ERBB2-dependent transforming activity is only limited to GRB7 or conserved in all GRB7 family proteins. We also tested GRB2, a well-known adaptor protein with distinct structure from GRB7, in addition to GRB7 family proteins. No foci were observed when each GRB protein (GRB2, GRB7, GRB10 and GRB14) was expressed individually in NIH3T3 cells; however, with the expression of ERBB2, a number of foci were reproducibly observed when either GRB7 or GRB2 was expressed (Fig. 2A, Table 2). This result indicates that ERBB2-dependent transforming activity was specific to GRB7 among the GRB7 family proteins.

To assess whether the transforming activity of GRB7 is based on activation of ERBB2, we evaluated the phosphorylation status of ERBB2 by expression of GRB7 compared with other GRB7 family proteins. To analyze total phosphorylation of ERBB2, whole cell lysates of 3T3-ERBB2 expressing one of the GRB7 family proteins were subjected to immunoprecipitation with an ERBB2 antibody. Then, the phosphorylation status of ERBB2 was analyzed using phosphotyrosine-specific antibody 4G10. As shown in Fig. 2B, ERBB2 phosphorylation was considerably upregulated in GRB7-expressing cells. GRB14 also increased ERBB2 phosphorylation; however, GRB10 had no effect. To determine which tyrosine residues of ERBB2 are phosphorylated when GRB7 is expressed, whole cell lysates of 3T3-ERBB2 expressing one of the GRB proteins were immunoblotted using a series of phospho-specific antibodies against ERBB2. As shown in Fig. 2C, phosphorylation of Tyr1221/1222 and Tyr1248 was upregulated in GRB7-expressing cells, while that of a Src phosphorylation site, Tyr877, was unaltered, suggesting that GRB7 promotes or retains ERBB2 autophosphorylation in particular. We therefore evaluated the effects of GRB7 and ERBB2 on signaling pathways that are frequently deregulated in cancer, that is, MAPK and PI3K-Akt pathways. We analyzed activation of components of these pathways by activation state-specific antibodies. These data showed that compared with other GRB proteins, co-expression of ERBB2 and GRB7 had no effect on the phosphorylation status of MEK or Erk. In contrast to components of the MAPK pathway, phosphorylation of Akt at Thr308 and Ser473 was marginally upregulated by expression of either ERBB2 or GRB7, but greatly upregulated when both proteins were expressed (Fig. 2C). These results suggest that cooperative transformation of ERBB2 and GRB7 may require Akt activation.

To dissect functional domains required for transforming activity of GRB7, we constructed five corresponding deletion mutants, PR (proline-rich), RA (Ras-associating), PH (Pleckstrin homology), BPS (between PH and SH2), and SH2 (Src-homology 2), and two point mutants for SH2 domain (Fig. 3A). R458K has a defect in the phosphotyrosine binding capacity of its SH2 domain, while F511A has a dimerization defect [14]. We assessed their transforming activity on 3T3-ERBB2 by focus formation assays. As shown in Fig. 3B and C, four GRB7 mutants, del-PR, del-RA, del-PH, and del-SH2, were defective in the transforming activity of 3T3-ERBB2 cells. R458K was also impaired in transforming activity, and the number of foci was reduced. In contrast, F511A retained comparable transforming activity with WT GRB7. However, del-BPS promoted foci formation, although its transforming activity was still dependent on ERBB2 (data not shown). Taken together, all four domains are required for ERBB2-dependent transforming activity of GRB7, except for its BPS region.

Please cite this article in press as: Saito, M., et al. Expression screening of 17q12–21 amplicon reveals GRB7 as an ERBB2-dependent oncogene. FEBS Lett. (2012), http://dx.doi.org/10.1016/j.febslet.2012.05.003

The BPS region of GRB7 family proteins is known to act as a pseudo substrate of the insulin and IGF1 receptors, and directly inhibits their catalytic activities [15]. Therefore, we hypothesized that BPS region of GRB7 downregulates ERBB2 kinase activity in spite of its activator function of ERBB2. To address this hypothesis, we compared WT and mutant GRB7 in terms of their effect on ERBB2 and Akt phosphorylation. As shown in Fig. 4, ERBB2 phosphorylation was increased by deletion of the BPS region. Moreover, Akt phosphorylation was concomitantly increased. These results indicate that the BPS region of GRB7 potentially represses ERBB2 activity in a similar manner to the inhibition of insulin and IGF1 receptors by GRB7 family proteins. Of the other transformation-deficient mutants, del-RA had no effect on ERBB2 and Akt phosphorylation, while del-PR, del-PH, and del-SH2 increased ERBB2 phosphorylation as well as WT but had no effect on Akt phosphorylation. Two point mutants with transforming activity retained ERBB2 and Akt phosphorylation. These results indicate that the transformed phenotype correlated with Akt activation status.

4. Discussion

Gene amplification is one of the major genetic alterations in cancer, which leads to overexpression of several genes. Amplicons recurrently observed in human cancers are likely to be positively selected owing to their contribution to oncogenesis, and indeed, amplicons have been shown to include cancer driver genes. But information on the significance of other co-amplified genes has been limited.

Accumulating clinical evidence indicate that *ERBB2* is a driver gene of several types of cancers. In fact, transforming activity of ERBB2 depends on its conformational status and expression level. An *ErbB2* mutant observed in rat neuroblastoma called *neu* oncogene, which tends to form dimers [16], has transforming activity [7]. Furthermore, strong expression of WT ERBB2 driven by retroviral LTR showed transforming activity [17]. In contrast, moderate expression of WT ERBB2 driven by SV40 promoter did not cause transformation [17]. We hypothesized that some of the genes in the 17q12–21 amplicon may cooperate with ERBB2 to cause tumorigenesis in the case ERBB2 expression is relatively low. To test this hypothesis, we assessed the transforming activity of all genes in the amplicon under the moderate expression of ERBB2. Our 3T3-ERBB2 cells formed neither foci nor colonies, enabling us to identify GRB7 as a gene that enhances the transforming activity of ERBB2.

In the screening of 52 genes in the amplicon, we detected integration of GRB7 in 11 out of 30 foci analyzed (Table 1, Supplementary Table 1). This result suggested the existence of oncogenes other than GRB7 in the amplicon. In fact, we identified another novel transforming gene in the amplicon with the same strategy using NMuMG-ERBB2, which we established by introducing ERBB2 expression vector into NMuMG, a mouse mammary epithelial cell line (A.M., manuscript in preparation). Some foci did not contain either of them (Supplementary Table 1). One possibility is that combination of multiple genes, which do not have independent transforming activity, causes transformation. Alternatively, recovering integrated cDNAs from genomic DNA of transformants may be biased in PCR amplification and cloning and thus some of the cDNAs might have been missed. Nevertheless, the fact that the remaining 50 genes (except for GRB7 and the novel transforming gene) in the *ERBB2* amplicon did not induce any foci strongly suggests that either of the two genes is necessary for cellular transformation.

GRB7 has a unique feature in the activation of ERBB2. ERBB2-dependent transforming activity is observed only in GRB7 among the GRB7 family proteins (Fig. 2A). Previous studies demonstrated GRB7 binds to ERBB family receptors, especially ERBB2 and ERBB3, through its SH2 domain [18]. Increased ERBB2 phosphorylation was observed in response to expression of GRB7, suggesting that this binding specificity contributes to its effect on ERBB2. Also in our transformation model, Akt phosphorylation correlated with the transformation phenotype of NIH3T3 (Figs. 3B and 4). Akt is deregulated in a wide spectrum of human cancers, such as breast, ovarian and thyroid cancers [19], and transforming activity of its truncated form in NIH3T3 has been reported [20]. Our results suggest activation of ERBB2-GRB7-Akt axis is sufficient for cellular transformation.

Our study with GRB7 mutants provides insights into the mechanism by which GRB7 activates ERBB2 and Akt. GRB7 family proteins consist of an RA domain, PH domain, BPS region and SH2 domain [13]. Among them, deletion of the BPS region in GRB7 increases ERBB2 phosphorylation and enhances transformation, suggesting two possibilities: first, the BPS region of GRB7 may inhibit ERBB2 kinase activities in a similar way that the BPS regions of GRB10 and GRB14 interfere with IGF1 receptors as pseudo-substrates [15]. Second, as it was recently shown that IGF1 receptor and ERBB2 were co-immunoprecipitated in SK-BR-3 breast cancer cells [21], GRB7 may enhance complex formation of ERBB2 and IGF1 receptor. If this is the case, lack of BPS region may facilitate IGF1 receptor to phosphorylate ERBB2. The RA domain, defined by sequence homology between the Ras effectors, is also involved in the phosphorylation of both ERBB2 and Akt. However, the molecular function of RA domain still remains to be elucidated.

Previous studies showed that GRB7 was required for SK-BR-3 cell proliferation and ERK1/2 and Akt phosphorylation [22,23]. Conversely, overexpression of GRB7 and ERBB2 in MCF-7 breast cancer cells enhance ERBB2 and Akt phosphorylation and tumor xenograft growth [24]. These studies suggested a possible involvement of GRB7 in the neoplastic phenotype of breast cancer cells. Beyond these studies, we further reveal GRB7 as a unique transforming gene, which does not show transforming activity by itself but cooperatively transforms NIH3T3 cells with ERBB2. Furthermore, we provide evidence for the significance of its BPS and RA domains in the ERBB2-GRB7-Akt signaling axis for cellular transformation.

Acknowledgments

We thank Kumiko Semba for her secretarial assistance. This research was partially supported by JSPS KAKENHI 23241064 and a grant for translational research programs from New Energy and Industrial Technology Development Organization (NEDO).

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.2012.05.003.

References

[1] Santarius, T., Shipley, J., Brewer, D., Stratton, M.R. and Cooper, C.S. (2010) A census of amplified and overexpressed human cancer genes. Nat. Rev. Cancer 10, 59–64.
[2] Ito, E. et al. (2007) Novel clusters of highly expressed genes accompany genomic amplification in breast cancers. FEBS Lett. 581, 3909–3914.
[3] Hsu, D.S. et al. (2009) Characterizing the developmental pathways TTF-1, NKX2-8, and PAX9 in lung cancer. Proc. Natl. Acad. Sci. U S A 106, 5312–5317.
[4] Kauraniemi, P. and Kallioniemi, A. (2006) Activation of multiple cancer-associated genes at the ERBB2 amplicon in breast cancer. Endocr. Relat. Cancer 13, 39–49.

Please cite this article in press as: Saito, M., et al. Expression screening of 17q12–21 amplicon reveals GRB7 as an ERBB2-dependent oncogene. FEBS Lett. (2012), http://dx.doi.org/10.1016/j.febslet.2012.05.003

[5] Yamamoto, T. et al. (2011) ErbB2/HER2: its Contribution to Basic Cancer Biology and the Development of Molecular Targeted Therapy in: Breast Cancer – Carcinogenesis, Cell Growth and Signalling Pathways (Gunduz, M., Ed.), InTech, Rijeka, Croatia.
[6] Slamon, D.J., Clark, G.M., Wong, S.G., Levin, W.J., Ullrich, A. and McGuire, W.L. (1987) Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science 235, 177–182.
[7] Bargmann, C.I., Hung, M.C. and Weinberg, R.A. (1986) Multiple independent activations of the neu oncogene by a point mutation altering the transmembrane domain of p185. Cell 45, 649–657.
[8] Muller, W.J., Sinn, E., Pattengale, P.K., Wallace, R. and Leder, P. (1988) Single-step induction of mammary adenocarcinoma in transgenic mice bearing the activated c-neu oncogene. Cell 54, 105–115.
[9] Goshima, N. et al. (2008) Human protein factory for converting the transcriptome into an in vitro-expressed proteome. Nat. Methods 5, 1011–1017.
[10] Kitamura, T., Koshino, Y., Shibata, F., Oki, T., Nakajima, H., Nosaka, T. and Kumagai, H. (2003) Retrovirus-mediated gene transfer and expression cloning: powerful tools in functional genomics. Exp. Hematol. 31, 1007–1014.
[11] Nagai, T., Ibata, K., Park, E.S., Kubota, M., Mikoshiba, K. and Miyawaki, A. (2002) A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. Nat. Biotechnol. 20, 87–90.
[12] Rabhi, I., Guedel, N., Chouk, I., Zerria, K., Barbouche, M.R., Dellagi, K. and Fathallah, D.M. (2004) A novel simple and rapid PCR-based site-directed mutagenesis method. Mol. Biotechnol. 26, 27–34.
[13] Han, D.C., Shen, T.L. and Guan, J.L. (2001) The Grb7 family proteins: structure, interactions with other signaling molecules and potential cellular functions. Oncogene 20, 6315–6321.
[14] Porter, C.J., Wilce, M.C., Mackay, J.P., Leedman, P. and Wilce, J.A. (2005) Grb7-SH2 domain dimerisation is affected by a single point mutation. Eur. Biophys. J. 34, 454–460.
[15] Stein, E.G., Gustafson, T.A. and Hubbard, S.R. (2001) The BPS domain of Grb10 inhibits the catalytic activity of the insulin and IGF1 receptors. FEBS Lett. 493, 106–111.
[16] Weiner, D.B., Liu, J., Cohen, J.A., Williams, W.V. and Greene, M.I. (1989) A point mutation in the neu oncogene mimics ligand induction of receptor aggregation. Nature 339, 230–231.
[17] Di Fiore, P.P., Pierce, J.H., Kraus, M.H., Segatto, O., King, C.R. and Aaronson, S.A. (1987) ErbB-2 is a potent oncogene when overexpressed in NIH/3T3 cells. Science 237, 178–182.
[18] Stein, D. et al. (1994) The SH2 domain protein GRB-7 is co-amplified, overexpressed and in a tight complex with HER2 in breast cancer. EMBO J. 13, 1331–1340.
[19] Vivanco, I. and Sawyers, C.L. (2002) The phosphatidylinositol 3-Kinase AKT pathway in human cancer. Nat. Rev. Cancer 2, 489–501.
[20] Cheng, J.Q., Altomare, D.A., Klein, M.A., Lee, W.C., Kruh, G.D., Lissy, N.A. and Testa, J.R. (1997) Transforming activity and mitosis-related expression of the AKT2 oncogene: evidence suggesting a link between cell cycle regulation and oncogenesis. Oncogene 14, 2793–2801.
[21] Nahta, R., Yuan, L.X., Zhang, B., Kobayashi, R. and Esteva, F.J. (2005) Insulin-like growth factor-I receptor/human epidermal growth factor receptor 2 heterodimerization contributes to trastuzumab resistance of breast cancer cells. Cancer Res. 65, 11118–11128.
[22] Chu, P.Y., Li, T.K., Ding, S.T., Lai, I.R. and Shen, T.L. (2010) EGF-induced Grb7 recruits and promotes Ras activity essential for the tumorigenicity of Sk-Br3 breast cancer cells. J. Biol. Chem. 285, 29279–29285.
[23] Nencioni, A. et al. (2010) Grb7 upregulation is a molecular adaptation to HER2 signaling inhibition due to removal of Akt-mediated gene repression. PLoS One 5, e9024.
[24] Bai, T. and Luoh, S.W. (2008) GRB-7 facilitates HER-2/Neu-mediated signal transduction and tumor formation. Carcinogenesis 29, 473–479.

Please cite this article in press as: Saito, M., et al. Expression screening of 17q12–21 amplicon reveals GRB7 as an ERBB2-dependent oncogene. FEBS Lett. (2012), http://dx.doi.org/10.1016/j.febslet.2012.05.003

CAXII Is a Sero-Diagnostic Marker for Lung Cancer

Makoto Kobayashi[1,2], Toshihide Matsumoto[1,4], Shinichiro Ryuge[6], Kengo Yanagita[1,2], Ryo Nagashio[1,2], Yoshitaka Kawakami[3], Naoki Goshima[3], Shi-Xu Jiang[4], Makoto Saegusa[4], Akira Iyoda[5], Yukitoshi Satoh[5], Noriyuki Masuda[6], Yuichi Sato[1,2]*

1 Department of Applied Tumor Pathology, Graduate School of Medical Sciences, Kitasato University, Kanagawa, Japan, **2** Department of Molecular Diagnostics, School of Allied Health Sciences, Kitasato University, Kanagawa, Japan, **3** Biomedicinal Information Research Center, National Institute of Advanced Industrial Science and Technology, Tokyo, Japan, **4** Department of Pathology, School of Medicine, Kitasato University, Kanagawa, Japan, **5** Department of Thoracic and Cardiovascular Surgery, School of Medicine, Kitasato University, Kanagawa, Japan, **6** Department of Respiratory Medicine, School of Medicine, Kitasato University, Kanagawa, Japan

Abstract

To develop sero-diagnostic markers for lung cancer, we generated monoclonal antibodies using pulmonary adenocarcinoma (AD)-derived A549 cells as antigens by employing the random immunization method. Hybridoma supernatants were immunohistochemically screened for antibodies with AMeX-fixed and paraffin-embedded A549 cell preparations. Positive clones were monocloned twice through limiting dilutions. From the obtained monoclonal antibodies, we selected an antibody designated as KU-Lu-5 which showed intense membrane staining of A549 cells. Based on immunoprecipitation and MADLI TOF/TOF-MS analysis, this antibody was recognized as carbonic anhydrase XII (CAXII). To evaluate the utility of this antibody as a sero-diagnostic marker for lung cancer, we performed dot blot analysis with a training set consisting of sera from 70 lung cancer patients and 30 healthy controls. The CAXII expression levels were significantly higher in lung cancer patients than in healthy controls in the training set ($P<0.0001$), and the area under the curve of ROC was 0.794, with 70.0% specificity and 82.9% sensitivity. In lung cancers, expression levels of CAXII were significantly higher in patients with squamous cell carcinoma (SCC) than with AD ($P=0.035$). Furthermore, CAXII was significantly higher in well- and moderately differentiated SCCs than in poorly differentiated ones ($P=0.027$). To further confirm the utility of serum CAXII levels as a sero-diagnostic marker, an additional set consisting of sera from 26 lung cancer patients and 30 healthy controls was also investigated by dot blot analysis as a validation study. Serum CAXII levels were also significantly higher in lung cancer patients than in healthy controls in the validation set ($P=0.030$). Thus, the serum CAXII levels should be applicable markers discriminating lung cancer patients from healthy controls. To our knowledge, this is the first report providing evidence that CAXII may be a novel sero-diagnostic marker for lung cancer.

Citation: Kobayashi M, Matsumoto T, Ryuge S, Yanagita K, Nagashio R, et al. (2012) CAXII Is a Sero-Diagnostic Marker for Lung Cancer. PLoS ONE 7(3): e33952. doi:10.1371/journal.pone.0033952

Editor: Vladimir Brusic, Dana-Farber Cancer Institute, United States of America

Received September 23, 2011; **Accepted** February 20, 2012; **Published** March 16, 2012

Funding: This study was supported in part by a Grant-in-Aid for Scientific Research Council (23590414) from the Japan Society for the Promotion of Science, grants from the Third Term Comprehensive Control Research for Cancer conducted by the Ministry of Health, Labour and Welfare of Japan, as well as from the Research Project (No. 2011-1006) of the School of Allied Health Sciences, Kitasato University. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. No additional external funding received for this study.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: yuichi@med.kitasato-u.ac.jp

Introduction

Lung cancer is the leading cause of cancer death, comprising 13% (1.6 million) of the total cancer cases and 18% (1.4 million) of the cancer deaths in the world in 2008 [1,2].

Tumor markers have been detected in sera, urine, and tissues from patients with malignant tumors, and can be used for an exact diagnosis, discrimination of benign or malignant tumors, follow-up after therapies, and prediction of the patient's outcome. At present, some sero-diagnostic markers are used for lung cancer, such as carcinoembryonic antigen (CEA) and sialyl Lewis X antigen (SLX) for adenocarcinoma (AD), and cytokeratin 19 fragment (CYFRA) and squamous cell carcinoma antigen (SCCa) for squamous cell carcinoma (SCC) [3]. The positive rates of CEA, SLX, CYFRA, and SCCa are reportedly 57, 40~50, 50~60, and 60~80%, respectively. However, it has been reported that these markers do not show sufficient tumor or organ specificities; for example, SLX can show false-positive results in the presence of pulmonary tuberculosis and pulmonary fibrosis, and CYFRA can elevate with interstitial pneumonia and renal failure.

Antibodies are usually developed using purified proteins or synthetic peptides. We have exhaustively generated monoclonal antibodies (MoAbs) against various tumor-associated proteins using the pulmonary AD-derived A549 cell as an antigen with the random immunization method [4], and over 1,000 MoAbs have been obtained [5]. This method is expected to generate antibodies against proteins with tumor-specific post-translational modifications, which are difficult to obtain by conventional immunization methods.

Carbonic anhydrase XII is a transmembrane zinc metalloenzyme that catalyzes the reversible hydration of carbon dioxide to form bicarbonate ($H_2O+CO_2 \Leftrightarrow H^+ + HCO_3^-$), and is a member of the alpha carbonic anhydrase (CA) family. CAXII has been proposed to be involved in the acidification of the extracellular microenvironment, which is suitable for rapid tumor growth. CAXII overexpression was initially detected in renal cell carcinoma, and subsequent studies confirmed its expression in

various human cancers, such as diffuse astrocytoma, breast, pancreatic, and ovarian carcinoma, as well as in non-small cell lung cancer (NSCLC) [6–11]. Its expression was influenced both by factors related to differentiation and hypoxia in breast cancer *in vivo*, and was associated with a more favorable prognosis in invasive breast carcinoma patients [12]. Higher CAXII expression was also correlated with a better overall and disease-specific survival in patients with resectable NSCLC [13]. However, no study has clarified CAXII in sera and its clinical utility as a sero-diagnostic marker for patients with malignant tumors.

In this study, the specificity of the obtained anti-CAXII antibody was confirmed by immunohistochemistry (IHC) and immunoblotting with lung cancer cell lines and lung cancer tissues. To further confirm its utility as a sero-diagnostic marker, CAXII levels in sera from patients with lung cancer were studied by dot blot analysis.

Materials and Methods

1. Cell lines

The A549 and LC-2/ad cells derived from lung AD were purchased from the Japanese Cancer Research Resources Bank (Tokyo, Japan) and RIKEN BioResource Center (Ibaraki, Japan), respectively. The RERF-LC-AI cells derived from lung SCC was purchased from the RIKEN BioResource Center. The N231 cells derived from SCLC were purchased from the American Type Culture Collection (Rockville, MD, USA). LCN1, a large cell neuroendocrine carcinoma (LCNEC) line, was established in our laboratory [14]. These cells were grown in RPMI-1640 medium (SIGMA, Steinheim, Germany) supplemented with 10% fetal bovine serum (FBS; Biowest, Miami, FL, USA), 100 units/ml of penicillin, and 100 μg/ml of streptomycin (GIBCO, Auckland, New Zealand). After harvesting and washing twice with phosphate-buffered saline without divalent ions (PBS-), sub-confluent cells were stored at −80°C for proteomics analysis or fixed in 10% formalin and embedded in paraffin for immunohistochemistry. A549 cells were also AMeX-fixed [15] for immunohistochemical screening. The SP2/O-Ag14 cells derived from a mouse myeloma were purchased from the RIKEN BioResource Center, and were grown in RPMI-1640 medium supplemented with 1× 8-azaguanine (50× Hybri-Max, SIGMA), 10% FBS, penicillin, and streptomycin.

2. Ethics statement

All samples were collected in accordance with the ethical guidelines and written consent mandated, and this study was approved by the Ethics Committee of Kitasato University School of Medicine. All patients and healthy controls were approached based on approved ethical guidelines, and those who agreed to participate in this study were required to sign consent forms. Patients could refuse entry and discontinue participation at any time. All participants provided written consent.

2.1. Sera. Sera from 70 patients with lung cancer (AD: 29, SCC: 21, SCLC: 17, and LCNEC: 3) and 30 healthy controls were used in the training set. In addition, a validation set consisting of sera from 26 patients with lung cancer (AD: 20, SCLC: 5, and LCNEC: 1) and 30 healthy controls was also studied. The clinicopathological characteristics of the patients data are summarized in Table 1.

Patient sera were collected at Kitasato University Hospital, and healthy control sera were provided by Kyowa Medex Co., Ltd. (Tokyo, Japan) and kept at −80°C until use.

3. Generation of monoclonal antibodies

A549 cell lysate was prepared with PBS(-) using an ultra-sonic homogenizer (UH-50; SMT Company, Tokyo, Japan). Five-week-old female BALB/c mice were immunized intra-peritoneally with 50 mg wet- weight of A549 cell lysate in 500 μl of PBS(-) 3 times with a two-week interval. The antibody titer was tested by IHC using 100-times diluted sera from the immunized mice as the first antibody on AMeX-fixed A549 cells. Three days prior to cell fusion, the animal with the highest titer was intra-peritoneally boosted by the same amount of A549 lysate. Hybridoma preparation and IHC screening with AMeX-fixed A549 cells were previously described [4,5].

Table 1. Clinicopathological characteristics of the patients.

Characteristics		Training set (N = 70)	Validation set (N = 26)
Age	<70	40 (57.1%)	19 (73.1%)
	≥70	30 (42.9%)	7 (26.9%)
Gender	Male	52 (74.3%)	16 (61.5%)
	Female	18 (25.7%)	10 (38.5%)
Stage	I	19 (27.2%)	17 (65.4%)
	II	11 (15.7%)	2 (7.7%)
	III	26 (37.1%)	4 (15.4%)
	IV	14 (20.0%)	3 (11.5%)
Tumor differentiation	Well	7 (13.2%)	11 (52.4%)
(NSCLC)	Moderate	10 (18.9%)	5 (23.8%)
	Poor	18 (34.0%)	4 (19.0%)
	Unknown	18 (34.0%)	1 (4.8%)
Tumor size	<3 cm	24 (34.3%)	15 (57.7%)
	≥3 cm	45 (64.3%)	6 (23.1)
	Unknown	1 (1.4%)	5 (19.2)
Nodal status	N0	23 (32.9%)	18 (69.3%)
	N1	12 (17.1%)	1 (3.8%)
	N2	23 (32.9%)	5 (19.2%)
	N3	12 (17.1%)	2 (7.7%)
Distant metastasis	M0	56 (80.0%)	23 (88.5%)
	M1	14 (20.0%)	3 (11.5%)
Histological type	AD[a]	29 (41.4%)	20 (77.0%)
	SCC[b]	21 (30.0%)	0 (0.0%)
	SCLC[c]	17 (24.3%)	5 (19.2%)
	LCNEC[d]	3 (4.3%)	1 (3.8%)

[a]Adenocarcinoma.
[b]Squamous cell carcinoma.
[c]Small cell lung carcinoma.
[d]Large cell neuroendocrine carcinoma.
doi:10.1371/journal.pone.0033952.t001

4. Proteomics analysis

4.1. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Proteins were extracted from each of A549, LC-2/ad, RERF-LC-AI, N231, and LCN1 cells with detergent lysis buffer [16] using an ultra-sonic homogenizer. Ten μg each of extracted proteins were boiled and separated by SDS-PAGE with 10% polyacrylamide gel at a constant current of 20 mA. After SDS-PAGE, proteins in gels were transferred to a polyvinylidene difluoride (PVDF) membrane (Millipore Corp., Billerica, MA, USA) for immunoblotting.

4.2. Immunoblotting. Blotting membranes were blocked with 0.5% casein from bovine milk (Sigma, St. Louis, MO, USA) for 30 min at RT. The membranes were then reacted with non-

diluted hybridoma supernatant for 1 hr at RT, followed by incubation with 1,000-times diluted horseradish peroxidase-conjugated rabbit anti-mouse IgG polyclonal antibody (Dako, Glostrup, Denmark) with 0.025% Casein for 45 min at RT. Finally, signals were developed using Immobilon Western HRP reagent (Millipore Corp.).

4.3. Determination of antibody isotype. To determine the isotype of the established KU-Lu-5 antibody, we used the IsoStrip^{TM} Mouse Monoclonal Antibody Isotyping Kit (Roche Diagnostics, Mannheim, Germany) according to the manufacturer's instructions.

4.4. Immunoprecipitation. The immunoprecipitation method used in this study was previously described [17]. In brief, A549 cells were washed with PBS (-) and treated with radio-immunoprecipitation assay (RIPA) buffer containing Complete-mini EDTA-free (Roche Diagnostics) on ice for 30 min. After centrifugation at 15,000 rpm for 30 min at 4°C, the supernatant was collected and precleared with protein G sepharose (50% slurry) (GE Healthcare Bio-Sciences Corp., Piscataway, NJ, USA) at 4°C overnight. To conjugate the primary antibody, 250 μL of primary antibody (KU-Lu-5 hybridoma supernatant) and 20 μL of protein G sepharose beads suspended in RIPA buffer were incubated with mixing at 4°C overnight. After centrifugation, the antibody-sepharose conjugate and 500 μg of total cellular protein from the precleared supernatant were incubated with mixing at 4°C for 4 hrs. The immunoprecipitates were collected by centrifugation at 15,000 rpm for 5 min at 4°C. After washing four times with RIPA buffer, the supernatant was carefully removed and the pellets were resuspended in 15 μL of 1×Laemmili's buffer. Then, 15 μL of samples were boiled and separated by SDS-PAGE with 10% polyacrylamide gel. After SDS-PAGE, gels were Zn-stained with the Negative Gel Stain MS kit (Wako Pure Chemical, Tokyo, Japan) according to the manufacturer's instructions.

4.5. Identification of antigen protein. *4.5.1. In-gel digestion.* The protein spot was excised from the SDS-PAGE gel and minced to 1 mm^3, destained with destaining solution (Wako Pure Chemical), dehydrated with 100% (v/v) ACN, and dried under vacuum conditions. Tryptic digestion was performed with a minimal volume of digestion solution which contained 20 ng/μl of trypsin (Trypsin Gold, Mass Spectrometry Grade, Promega, Madison, WI, USA) and 25 mM NH_4HCO_3 for 24 hrs at 37°C. After incubation, digested protein fragments eluted in solution were collected, and gels were washed once in 5% (v/v) trifuloroacetic acid /50% (v/v) ACN and collected in the same tube.

4.5.2. Protein identification. The collected peptide fragments were analyzed using autoflex III matrix-associated laser desorption/ionization-time of flight/time of flight mass spectrometry (MALDI-TOF/TOF MS; Bruker Daltonik, Bremen, Germany). A disposable plate, spotted α-cyano-4-hydroxycinnamic acid matrix for samples, and PAC Peptide Calibstandard for calibration (Prespotted AnchorChip 96 set for Proteomics, Bruker Daltonik) were used. Peptide mass fingerprints (PMF) were measured, and then a few peaks obtained from PMF were further measured for their tandem mass spectra as parent masses. MASCOT (http://www.matrixscience.com) using the IPI Human database (93,289 sequences; 36,994,704 residues), released on 3 May, 2011 (http://www.matrixscience.com), was used to determine proteins from PMF and tandem mass data.

5. Immunoblot analysis with recombinant CAXII protein

Recombinant CAXII protein and Venus protein as a negative control with GST-tag were prepared using a wheat germ cell-free system [18]. Fourteen μg each of recombinant CAXII and Venus proteins were boiled and separated with SDS-PAGE, followed by immunoblotting with KU-Lu-5 antibody, as mentioned in 2.4.1.

6. Immunohistochemical staining

Three-μm-thick sections, made from 10% formalin-fixed and paraffin-embedded lung cancer cell lines and 37 surgically resected lung cancers (AD: 28, SCC: 9) were deparaffinized in xylene, rehydrated in a descending ethanol series, and then treated with 3% hydrogen peroxide for 20 min. After the antigen was retrieved by autoclaving in 0.01 mol/L citrate buffer (pH 6.0) with 0.1% Tween 20 at 121°C for 10 min, the sections were reacted with non-diluted KU-Lu-5 hybridoma supernatant for 16–18 hrs at room temperature (RT). After rinsing in TBS three times for 5 min each, the sections were reacted with ChemMate Envision reagent (Dako) for 30 min at RT. Finally, the sections were visualized with Stable DAB solution (Invitrogen Corp.) and counterstained with Mayer's hematoxylin.

7. Dot blot analysis

7.1. Sample preparation. *7.1.1. Removal of albumin and IgG from serum samples.* The removal of albumin and IgG from sera was performed using a ProteoExtract Albumin/IgG Removal kit (Merck, Darmstadt, Germany) according to the manufacturer's instructions. A 60-μL sample of each sera was diluted with 540 μL of binding buffer, and allowed to pass the column by gravity flow. The flow-through fraction was collected in a collection tube. To wash the column, binding buffer was allowed to pass the column by gravity flow. The flow-through fraction was collected in the same collection tube.

7.1.2. Desalting and concentration by ultrafiltration. The albumin- and IgG-depleted samples were buffer-exchanged and concentrated using 10-kDa molecular-weight cut-off ultra-filtration VIVASPIN 2 (Sartorius, Gottingen, Germany). The samples were centrifuged at 6,000×g at 4°C until less than 100 μL, and then the buffer was exchanged for PBS (-) with concentration at 6,000×g at 4°C until concentrated to less than 50 μL. The concentrated samples were adjusted to a final volume of 60 μL with PBS (-).

7.2. Dot blot analysis. One μl each of albumin- and IgG-depleted samples diluted to 1:20 with PBS(-) and mouse IgG (purified in our laboratory) for a positive control were spotted on a PVDF membrane (Millipore Corp.) using the automatic dot blot system with a 256-solid pin configuration (Kakengeneqs Inc., Chiba, Japan). Two sheets of membrane were prepared for one set of experiment. Spotted membranes were washed in TBS for 10 min, and blocked with 0.5% casein (Sigma) for 1 hr at RT. One membrane was then reacted with non-diluted KU-Lu-5 hybridoma supernatant, and the other membrane was reacted with antibody diluting solution [20-times diluted 0.5% casein with 0.1% Tween 20 added TBS (TBS-T)] for 30 min at RT. After rinsing in TBS-T 3 times for 5 min each, membranes were incubated with 1,000-times diluted horseradish peroxidase-conjugated rabbit anti-mouse IgG polyclonal antibody (Dako) for 30 min at RT. Finally, signals were developed with Immobilon Western reagent (Millipore Corp.). The data were analyzed using DotBlotChipSystem Ver. 4.0 (Dynacom Co., Ltd., Chiba, Japan). Each normalized signal was presented as the ratio of the positive intensity versus the negative background intensity.

8. Statistical analysis

Serum CAXII levels in patients with lung cancer and healthy controls were statistically analyzed using the Mann-Whitney *U*-test. Sensitivity, specificity, and predictive values were calculated with the SPBS software package (Ver. 9.42 for Windows) for each variable at a corresponding cut-off. Discriminant function analysis

was performed to classify patients in the "lung cancer" vs. "healthy control" group, according to the status of the biomarkers, using the SPBS software package. The area under the curve (AUC) and best cut-off point were calculated employing receiver operating characteristic (ROC) analysis. Results were considered significant when $P<0.05$.

Results

1. Confirmation of antibody titer in mouse sera

The antibody titer was tested by IHC with 1,000-times diluted sera of immunized mice as the first antibody on AMeX-fixed A549 cells. As a result, the sera from immunized mice contained antibodies that reacted with various components of A549 cells.

Using AMeX-fixed A549 cell preparations for the immunohistochemical screening of hybridomas, we finally established 188 MoAbs in total and a further study was performed with the KU-Lu-5 clone, which showed intense staining in A549 cells (Fig. 1 A).

2. Identification of antigen protein

In order to identify the antigen protein recognized by the KU-Lu-5 antibody, we performed IP with lysate from A549 cells. The results of IP are shown in Fig. 1 B, C. The antigenic protein was observed at roughly 40 kDa. To determine the antigenic protein recognized by KU-Lu-5 antibody, we excised and collected the spot from the Zn-stained gel, and proceeded with in-gel digestion. After analysis employing a MALDI-TOF/TOF MS and a MASCOT search, the protein was determined as isoform 2 of carbonic anhydrase XII (CAXII, accession: IPI00221392), which is composed of 343 amino acids with a predicted M.W. of 38,384 Da. The result was confirmed by immunoblot analysis with recombinant CAXII protein using KU-Lu-5 hybridoma supernatant as the first antibody (Fig. 1 D). The immunoglobulin isotype of KU-Lu-5 antibody was determined as IgG_1, κ.

3. Immunoblot analysis

Expression of CAXII was detected only in A549 cells as a roughly 40-kDa protein, and no clear band was detected in other cells used in this study (Fig. 2 A).

4. Immunohistochemical staining for CAXII

Immunohistochemically, membranous expression of CAXII was observed only in A549 cells (Fig. 2 B). Membranous staining was detected in 2 of the 28 ADs (7.1%) and in 2 of the 9 SCCs (22.2%) (Fig. 2 C, D).

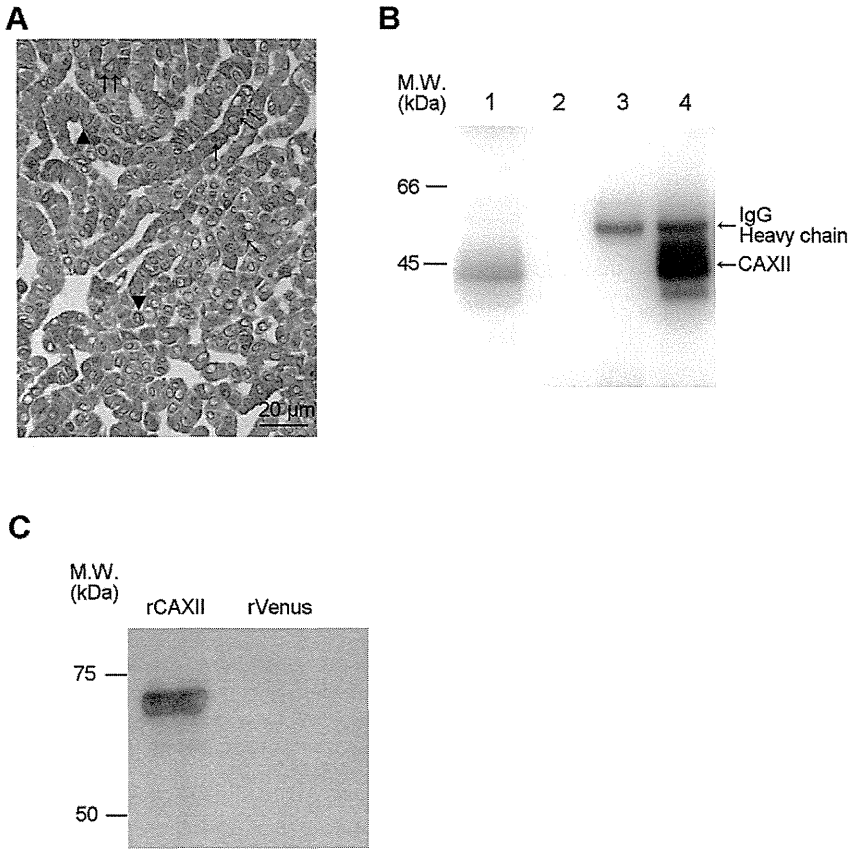

Figure 1. Production of anti-CAXII monoclonal antibody and its antigen identification. (A) The antibody titer was tested immunohistochemically using 1,000-times diluted sera of immunized mice as the first antibody on AMeX-fixed A549 cells, which were used as an immunogen. The sera of immunized mice contained antibodies that reacted with various cell components, such as the nucleus (↑), plasma membrane (▲), and cytoplasm (↑ ↑). (B) Immunoprecipitation with KU-Lu-5 antibody. Immunoblot analysis using KU-Lu-5 hybridoma supernatant as the first antibody [lane 1: A549 lysate, lane 2: A549 lysate combined with protein G, lane 3: KU-Lu-5 antibody combined with protein G, lane 4: A549 lysate combined with KU-Lu-5 antibody]. Lanes 2 to 3 are negative controls, and immunoprecipitated product with KU-Lu-5 antibody was detected in lane 4 (↑). (C) Confirmation of identified antigen protein. KU-Lu-5 antibody reacted with recombinant CAXII protein (64 kDa), but not with recombinant Venus protein.
doi:10.1371/journal.pone.0033952.g001

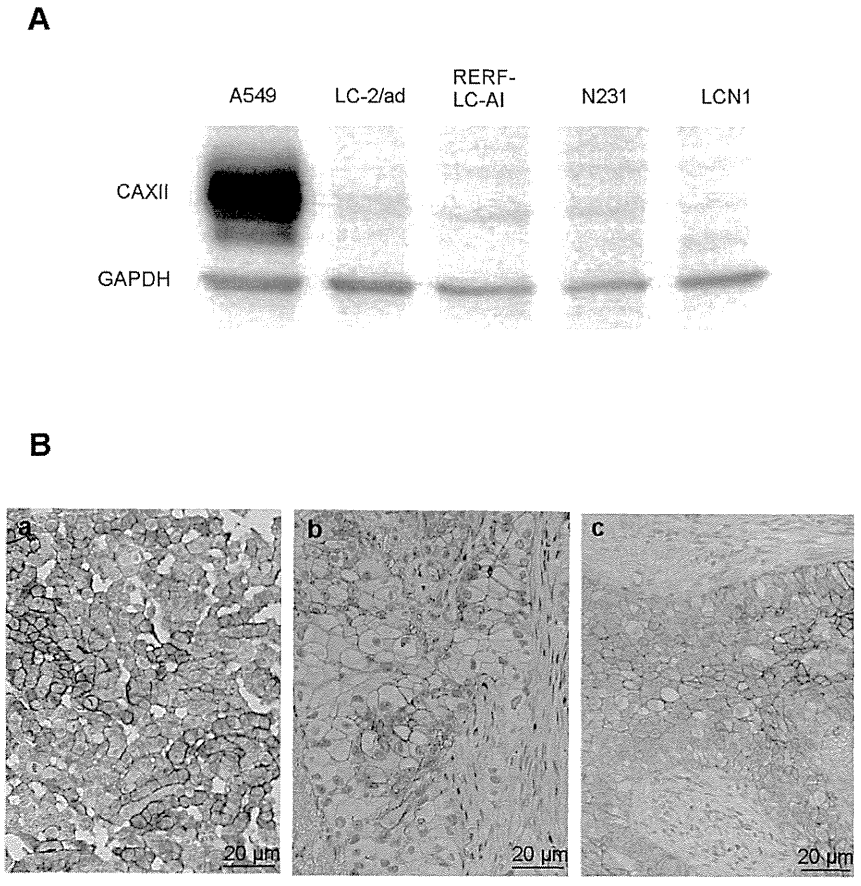

Figure 2. Expression of CAXII antibody in lung cancer cell lines and tissues. (A) Immunoblot analysis of CAXII in lung cancer cell lines. CAXII was detected as an approximately 40-kDa protein with A549 cells. (B) Immunostaining of CAXII in A549 cells (a), adenocarcinoma (b), and squamous cell carcinoma (c) of the lung, and each showed membranous staining of CAXII.
doi:10.1371/journal.pone.0033952.g002

5. Serum CAXII in patients with lung cancer

The serum CAXII levels were significantly higher in lung cancer patients than in healthy controls in the training set ($P<0.0001$). Relative values of serum CAXII levels ranged from 0.101 to 4.01 (median: 1.520) in lung cancer patients, but 0.006 to 1.679 (median: 0.290) in healthy controls (Fig. 3 A). In lung cancer, CAXII serum levels of SCC patients were significantly higher than those of AD patients ($P = 0.03$) (Fig. 3 A). The area under the ROC curve (AUC) between lung cancers and healthy controls was 0.794 (Fig. 3 B). When an optimal cut-off value of 0.387 for CAXII was applied, the diagnostic sensitivity and specificity for lung cancer were 82.9 and 70.0, respectively, and the negative and positive predictive values were 0.617 and 0.863, respectively. Furthermore, within SCCs, serum CAXII levels were significantly higher in patients with well- and moderately differentiated tumors than those with poorly differentiated ones ($P = 0.027$) (Fig. 4 A), and tended to be higher in patients with a tumor size of less than 3 cm rather than more than 3 cm ($P = 0.0538$). However, there was no difference in the smoking history of patients (Fig. 4 B). CAXII levels in stage I, II, and III ADs were 1.501, 0.704, and 1.001, respectively, and CAXII levels in stage I, II, and III SCCs were 1.764, 2.093, and 1.854, respectively. These data were summarized in Table 2. No relations between the CAXII serum levels and tumor stage or presence of metastasis were identified for either ADs or SCCs. To further confirm the utility of serum CAXII levels as a sero-diagnostic marker, 56 additional samples of sera were analyzed by dot blot analysis as a validation study. The serum CAXII levels were also significantly higher in lung cancer patients than in healthy controls in the validation set ($P = 0.030$). Relative values of serum CAXII levels ranged from 0.000 to 8.023 (median: 3.921) in lung cancer patients, but 0.000 to 8.331 (median: 2.806) in healthy controls (Fig. 5). When an optimal cut-off value of 3.086 for applied, the diagnostic sensitivity and specificity for lung cancer were 65.4 and 70.0, respectively.

Discussion

In this study, aiming to discover useful sero-diagnostic markers for lung cancer, we generated monoclonal antibodies using lung AD-derived A549 cells as antigens. From the obtained 188 antibodies, we focused on an antibody recognizing CAXII, and explored its clinical utility as a sero-diagnostic marker for lung cancer. This random immunization method is expected to yield antibodies against tumor-specific proteins with post-translational modifications, which are difficult to obtain by conventional immunization methods. Actually, several authors have reported that monoclonal antibodies generated by this method are useful as diagnostic and prognostic markers for cancers [5,17,19]. Battke *et al.* [20] established a 6A10 antibody recognizing CAXII using a similar immunization methodology. However, the obtained antibodies were limited to those only reacting with cell surface antigens because of using flow cytometry for the screening of hybridomas. In the present study, the hybridomas were immunohistochemically screened which facilitated the obtaining of

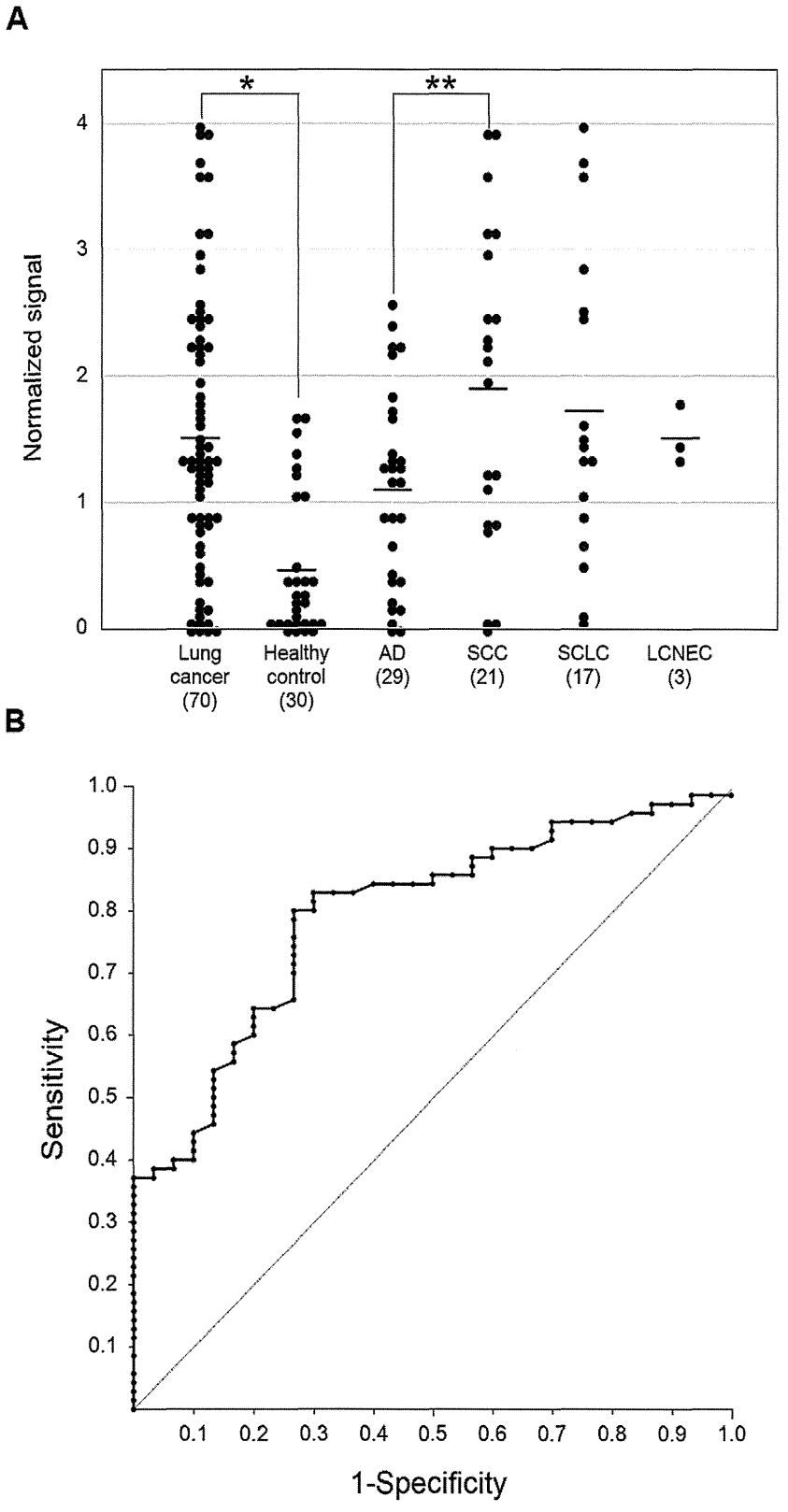

Figure 3. Serum CAXII levels in patients with lung cancer and healthy controls in the training set. Serum CAXII levels in patients with lung cancer and healthy controls. (A) The median CAXII level in the sera from healthy controls was 0.29, and that in sera from lung cancer patients was 1.52. Serum CAXII levels were significantly higher in lung cancer patients (*P<0.001). Furthermore, serum CAXII levels were higher in SCCs than ADs (**P = 0.0381). (B) Receiver-operating characteristic curve analysis of CAXII as a serum marker for lung cancer. The corresponding areas under the curves were 0.794 for CAXII. With a 70.0% specificity, the sensitivity of CAXII for lung cancer was 82.9%, at a cut-off value corresponding to 0.387.
doi:10.1371/journal.pone.0033952.g003