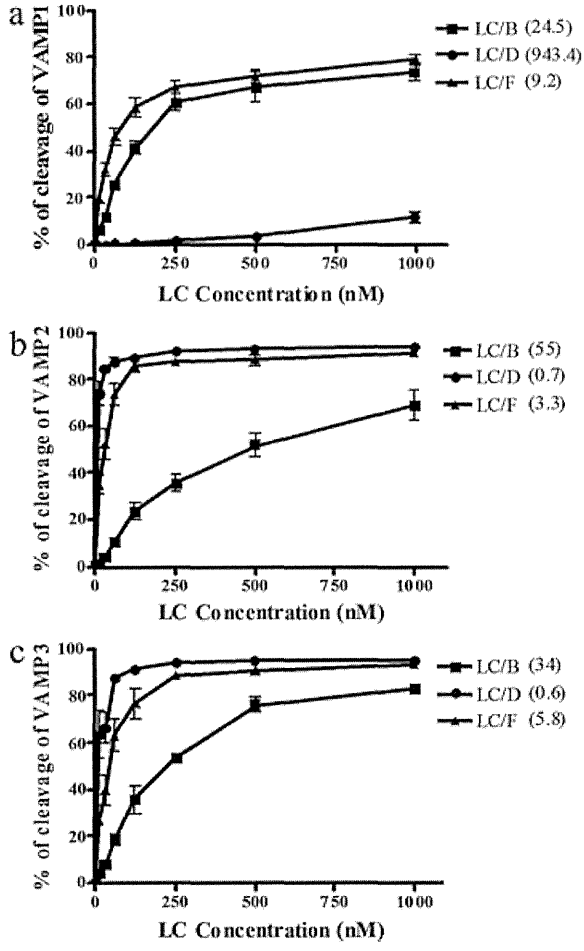

Fig. 2. Cleavage of VAMP1, 2, and 3 by LC/B, /D, and /F. Various concentrations of LC/B (1–430), D (1–436), and F (1–439) were incubated with 5 μM (a) VAMP1, (b) VAMP2, and (c) VAMP3 for 10 min at 37°C, and the reaction mixtures examined using SDS-PAGE and Coomassie Blue staining. Cleavage products were quantified by densitometry. Values in parentheses indicate the concentration of LC (nM) for 10% cleavage of a substrate. Standard error bars represent the average of three different experiments.

4-fold higher *Kcat* value when compared to that for wild-type VAMP1 (Table 3). This indicates that residue 48 is important for the cleavage of VAMP1 by LC/D.

Identification of BoNT-LC VAMP3 cleavage sites

To date, the precise sites of human VAMP3 cleavage by BoNT have not been clearly defined. To identify the scissile bonds, human VAMP3 was incubated with LC/B, /D, or /F for 1 hr at 37°C, and the cleaved C-terminal fragment of human VAMP3 subjected to amino acid sequencing of the N-terminal residues (data not shown). The products of LC/B, /D, and /F reactions yielded the sequences FETSA, LSELD, and KLSEL, respectively. These results indicate that the VAMP3 cleavage site of LC/B, /D, and /F corresponds to that of VAMP1 or 2 (Fig. 3).

DISCUSSION

Large numbers of gene families have been identified since the decoding of the sequence of the human genome. One of these, the VAMP family, comprises nine SNARE proteins that are crucial for membrane fusion. In the present paper, we carried out a comprehensive study of the substrate specificity and kinetic characteristics of LC/B, /D, and /F for the cleavage of human VAMP family proteins. While it has been reported that human VAMP1, 2, and 3 are digested by BoNT/B, D, F, and G (34), no reports have investigated other human VAMP family proteins as potential target molecules of BoNT. Recently, the resources and expression technology necessary for human proteomics on a whole-proteome scale have been developed, including the versatile Gateway vector system, which provides a foundation for the simple and efficient production of human proteins. In the present study, we utilized this system to determine the substrate specificity of BoNT for multiple recombinant human VAMP family proteins. While our data demonstrated that LC/B, /D, and /F are able to cleave VAMP1, 2, and 3, other VAMP family proteins were resistant to cleavage.

Table 3. Kinetic constants of LC/B, /D, and /F

	Km (μM)			*Kcat* (S^{-1})			*Kcat*/*Km* (S^{-1} μM^{-1})		
Substrate	LC/B	LC/D	LC/F	LC/B	LC/D	LC/F	LC/B	LC/D	LC/F
Wild type									
VAMP1	19.2 ± 1.7	28.9 ± 3.8	19.0 ± 3.0	3.96	0.59	16.12	0.21	0.02	0.85
VAMP2	29.9 ± 3.5	28.0 ± 2.4	24.5 ± 1.2	4.68	146.60	34.37	0.16	5.23	1.40
VAMP3	11.6 ± 1.2	11.1 ± 0.8	15.0 ± 1.3	3.50	113.41	28.57	0.30	10.22	1.90
VAMP1 mutant									
E42D	14.8 ± 1.2	16.6 ± 2.2	24.8 ± 1.9	3.71	0.49	17.22	0.25	0.03	0.69
I48M	24.7 ± 2.3	15.1 ± 0.8	9.9 ± 0.9	4.33	119.98	13.66	0.18	7.92	1.38
S81T	28.2 ± 3.0	25.2 ± 6.0	27.3 ± 2.6	4.66	0.55	17.77	0.17	0.02	0.65

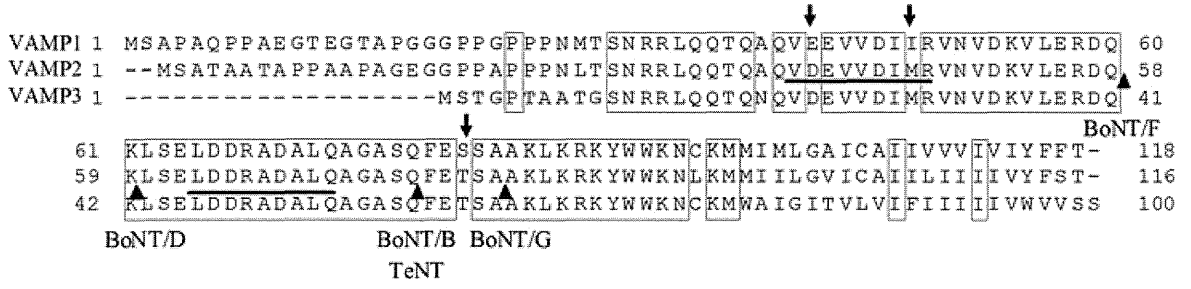

Fig. 3. Comparison of the amino acid sequences of human VAMP1, 2, and 3. Boxes indicate identical residues. The VAMP1 or 2 cleavage sites of BoNT/B, /D, /F, /G, and TeNT are denoted by arrowheads, whereas the VAMP1 residues that were mutated in the present study are shown by arrows. The V1 SNARE motif (residues 39–47) and the V2 SNARE motif (residues 63–71) in VAMP2, which are involved in substrate recognition (36, 40), are underlined.

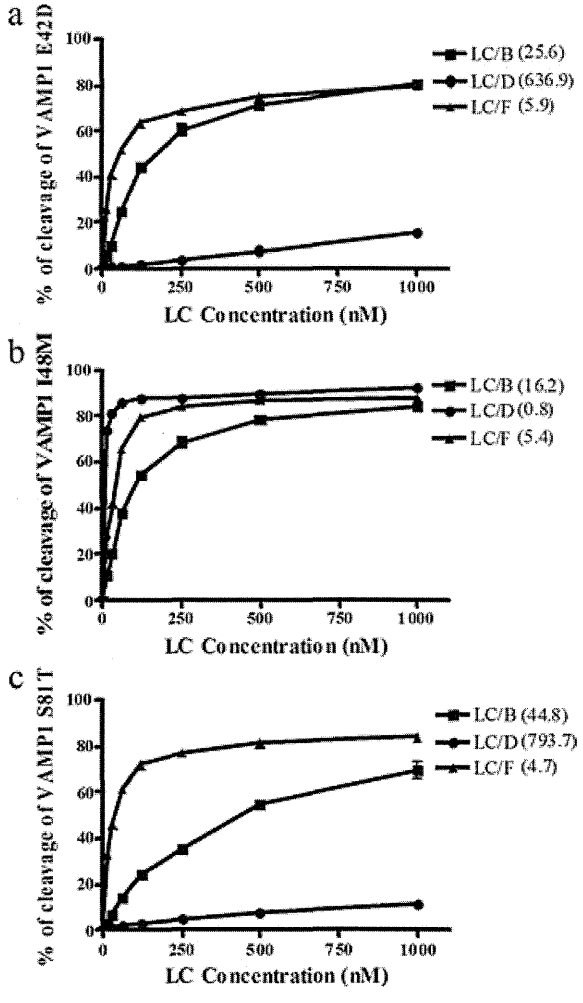

Fig. 4. Cleavage of VAMP1 mutants by LC/B, /D, and /F. Various concentrations of LC/B (1–430), /D (1–436), and /F (1–439) were incubated with 5 μM (a) VAMP1 E42D, (b) I48M, and (c) S81T mutants for 10 min at 37°C. Reaction mixtures were resolved by SDS-PAGE and proteins were detected by Coomassie Blue staining. The amount of cleaved VAMP1 mutants was quantified by densitometry. Values in parentheses indicate the concentration of LC (nM) for 10% cleavage of a substrate. Standard error bars represent the average of three different experiments.

Interestingly, we found that LC/D exhibited extremely low catalytic activity towards VAMP1 when compared to VAMP2 and 3. Our point mutation analysis elucidated that the low activity of LC/D for VAMP1 is due to the substitution of the Met found in VAMP2 and 3 with an Ile at residue 48 in VAMP1. This residue is located in a V1 SNARE motif. Previously, Yamasaki *et al.* reported that rat VAMP1 required an approximately 3700-fold higher concentration of BoNT/D than VAMP2 for *in vitro* cleavage to occur (35). Furthermore, it has been demonstrated that Met 46 in rats (equivalent to Ile 48 of human VAMP1), which is located within the V1 SNARE motif, is a key residue in the interaction of BoNT/D with rat VAMP2 (35, 36). This has recently been confirmed by the findings of a study that examined the substrate residues involved in the interaction of rat VAMP2 with BoNT/D using systematic mutagenesis, as it was demonstrated that Met 46 of rat VAMP2 had a significant effect on the low cleavage rate (37). These observations are consistent with our results, which found that the substitution of Met 48 with Ile in human VAMP1 is responsible for the low catalytic activity of LC/D towards this protein. Previously, Dolman *et al.* showed that the susceptibility of monkeys to BoNT/D given orally was much lower relative to other BoNT (38). Because residue 48 of VAMP1 of human and monkey was Ile, the lower susceptibility of monkeys to BoNT/D given orally may be due to the substitution at residue 48.

Despite the high sequence identity of human VAMP1, 2, and 3, the characteristics of human VAMP3 that make it a suitable substrate for BoNT are poorly understood. Unlike neuronal VAMP1 and 2, VAMP3 is widely expressed and is found in the sorting and recycling endosomes (10). In this study, we demonstrated for the first time that human VAMP3 is cleaved by LC/B, /D, and /F, and that these cleavage sites are located at the Gln 59-Phe 60, Lys 42-Leu 43, and Gln 41-Lys 42 peptide bonds, which correspond to those of VAMP1 or 2. Furthermore, our data clearly illustrate that all LC have higher substrate specificity to the non-neuronal SNARE isoform VAMP3 than to the neuronal SNARE isoforms VAMP1 or VAMP2. In a previous

study, it has been revealed that VAMP3 is involved in not only constitutive exocytosis, but also regulates exocytosis events such as TNF-α trafficking in macrophages (39, 40). As BoNT are well-established tools for investigating vesicle fusion and studying neurotransmitter release mechanisms in neuronal cells, our findings therefore suggest that BoNT/B, /D, and /F may be useful in the elucidation of intracellular vesicular trafficking routes and the mechanisms of membrane fusion processes in the VAMP3-related exocytotic pathway.

In conclusion, our study has elucidated the substrate specificity and kinetic characteristics of BoNT-LC/B, /D, and /F for the cleavage of nine human VAMP family proteins prepared from a 'human protein factory' resource. We demonstrated that although all LC cleave VAMP1, 2, and 3, other VAMP family proteins are not cleaved by LC. Furthermore, we found that LC/D exhibited extremely low catalytic activity for VAMP1 compared with VAMP2 and 3, and elucidated that the substitution of Met with Ile at residue 48 in human VAMP1 is responsible for this low activity. Further studies analyzing the substrate specificities of BoNT for other human SNARE family proteins are currently in progress, as understanding the substrate specificity and kinetic characteristics of BoNT for human SNARE proteins will provide insight valuable in the development of novel clinical uses for BoNT.

ACKNOWLEDGMENT

This work was supported by a grant-in-aid for Scientific Research (B, 21380188) from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of JAPAN.

DISCLOSURE

We have no disclosure.

REFERENCES

1. Montecucco C., Schiavo G. (1994) Mechanism of action of tetanus and botulinum neurotoxins. *Mol Microbiol* **13:** 1–8.
2. Schiavo G., Matteoli M., Montecucco C. (2000) Neurotoxins affecting neuroexocytosis. *Physiol Rev* **80:** 717–66.
3. Arnon S.S., Schechter R., Inglesby T.V., Henderson D.A., Bartlett J.G., Ascher M.S., Eitzen E., Fine A.D., Hauer J., Layton M., Lillibridge S., Osterholm M.T., O'toole T., Parker G., Perl T.M., Russell P.K., Swerdlow D.L., Tonat K. (2001) Botulinum toxin as a biological weapon: medical and public health management. *JAMA* **285:** 1059–70.
4. Glogau R.G. (2002) Review of the use of botulinum toxin for hyperhidrosis and cosmetic purposes. *Clin J Pain* **18:** S191–7.
5. Comella C.L., Pullman S.L. (2004) Botulinum toxins in neurological disease. *Muscle Nerve* **29:** 628–44.
6. Bock J.B., Matern H.T., Peden A.A., Scheller R.H. (2001) A genomic perspective on membrane compartment organization. *Nature* **409:** 839–41.
7. Hong W. (2005) SNAREs and traffic. *Biochim Biophys Acta* **1744:** 493–517.
8. Jahn R., Scheller R.H. (2006) SNAREs–engines for membrane fusion. *Nat Rev Mol Cell Biol* **7:** 631–43.
9. Jahn R., Sudhof T.C. (1999) Membrane fusion and exocytosis. *Annu Rev Biochem* **68:** 863–911.
10. Lin R.C., Scheller R.H. (2000) Mechanisms of synaptic vesicle exocytosis. *Annu Rev Cell Dev Biol* **16:** 19–49.
11. Feng D., Crane K., Rozenvayn N., Dvorak A.M., Flaumenhaft R. (2002) Subcellular distribution of 3 functional platelet SNARE proteins: human cellubrevin, SNAP-23, and syntaxin 2. *Blood* **99:** 4006–14.
12. Polgar J., Chung S.H., Reed G.L. (2002) Vesicle-associated membrane protein 3 (VAMP-3) and VAMP-8 are present in human platelets and are required for granule secretion. *Blood* **100:** 1081–3.
13. Galli T., Chilcote T., Mundigl O., Binz T., Niemann H., De Camilli P. (1994) Tetanus toxin-mediated cleavage of cellubrevin impairs exocytosis of transferrin receptor-containing vesicles in CHO cells. *J Cell Biol* **125:** 1015–24.
14. Proux-Gillardeaux V., Gavard J., Irinopoulou T., Mege R.M., Galli T. (2005) Tetanus neurotoxin-mediated cleavage of cellubrevin impairs epithelial cell migration and integrin-dependent cell adhesion. *Proc Natl Acad Sci USA* **102:** 6362–7.
15. Luftman K., Hasan N., Day P., Hardee D., Hu C. (2009) Silencing of VAMP3 inhibits cell migration and integrin-mediated adhesion. *Biochem Biophys Res Commun* **380:** 65–70.
16. Steegmaier M., Klumperman J., Foletti D.L., Yoo J.S., Scheller R.H. (1999) Vesicle-associated membrane protein 4 is implicated in trans-Golgi network vesicle trafficking. *Mol Biol Cell* **10:** 1957–72.
17. Mallard F., Tang B.L., Galli T., Tenza D., Saint-Pol A., Yue X., Antony C., Hong W., Goud B., Johannes L. (2002) Early/recycling endosomes-to-TGN transport involves two SNARE complexes and a Rab6 isoform. *J Cell Biol* **156:** 653–64.
18. Zeng Q., Subramaniam V.N., Wong S.H., Tang B.L., Parton R.G., Rea S., James D.E., Hong W. (1998) A novel synaptobrevin/VAMP homologous protein (VAMP5) is increased during in vitro myogenesis and present in the plasma membrane. *Mol Biol Cell* **9:** 2423–37.
19. Advani R.J., Yang B., Prekeris R., Lee K.C., Klumperman J., Scheller R.H. (1999) VAMP-7 mediates vesicular transport from endosomes to lysosomes. *J Cell Biol* **146:** 765–76.
20. Galli T., Zahraoui A., Vaidyanathan V.V., Raposo G., Tian J.M., Karin M., Niemann H., Louvard D. (1998) A novel tetanus neurotoxin-insensitive vesicle-associated membrane protein in SNARE complexes of the apical plasma membrane of epithelial cells. *Mol Biol Cell* **9:** 1437–48.
21. Pocard T., Le Bivic A., Galli T., Zurzolo C. (2007) Distinct v-SNAREs regulate direct and indirect apical delivery in polarized epithelial cells. *J Cell Sci* **120:** 3309–20.
22. Wong S.H., Zhang T., Xu Y., Subramaniam V.N., Griffiths G., Hong W. (1998) Endobrevin, a novel synaptobrevin/VAMP-like protein preferentially associated with the early endosome. *Mol Biol Cell* **9:** 1549–63.
23. Antonin W., Holroyd C., Tikkanen R., Honing S., Jahn R. (2000) The R-SNARE endobrevin/VAMP-8 mediates homotypic fusion of early endosomes and late endosomes. *Mol Biol Cell* **11:** 3289–98.

24. Wang C.C., Ng C.P., Lu L., Atlashkin V., Zhang W., Seet L.F., Hong W. (2004) A role of VAMP8/endobrevin in regulated exocytosis of pancreatic acinar cells. *Dev Cell* **7:** 359–71.
25. Zhang T., Wong S.H., Tang B.L., Xu Y., Hong W. (1999) Morphological and functional association of Sec22b/ERS-24 with the pre-Golgi intermediate compartment. *Mol Biol Cell* **10:** 435–53.
26. Hatsuzawa K., Hirose H., Tani K., Yamamoto A., Scheller R.H., Tagaya M. (2000) Syntaxin 18, a SNAP receptor that functions in the endoplasmic reticulum, intermediate compartment, and cis-Golgi vesicle trafficking. *J Biol Chem* **275:** 13713–20.
27. Zhang T., Hong W. (2001) Ykt6 forms a SNARE complex with syntaxin 5, GS28, and Bet1 and participates in a late stage in endoplasmic reticulum-Golgi transport. *J Biol Chem* **276:** 27480–7.
28. Fukasawa M., Varlamov O., Eng W.S., Sollner T.H., Rothman J.E. (2004) Localization and activity of the SNARE Ykt6 determined by its regulatory domain and palmitoylation. *Proc Natl Acad Sci USA* **101:** 4815–20.
29. Tai G., Lu L., Wang T.L., Tang B.L., Goud B., Johannes L., Hong W. (2004) Participation of the syntaxin 5/Ykt6/GS28/GS15 SNARE complex in transport from the early/recycling endosome to the trans-Golgi network. *Mol Biol Cell* **15:** 4011–22.
30. Goshima N., Kawamura Y., Fukumoto A., Miura A., Honma R., Satoh R., Wakamatsu A., Yamamoto J., Kimura K., Nishikawa T., Andoh T., Iida Y., Ishikawa K., Ito E., Kagawa N., Kaminaga C., Kanehori K., Kawakami B., Kenmochi K., Kimura R., Kobayashi M., Kuroita T., Kuwayama H., Maruyama Y., Matsuo K., Minami K., Mitsubori M., Mori M., Morishita R., Murase A., Nishikawa A., Nishikawa S., Okamoto T., Sakagami N., Sakamoto Y., Sasaki Y., Seki T., Sono S., Sugiyama A., Sumiya T., Takayama T., Takayama Y., Takeda H., Togashi T., Yahata K., Yamada H., Yanagisawa Y., Endo Y., Imamoto F., Kisu Y., Tanaka S., Isogai T., Imai J., Watanabe S., Nomura N. (2008) Human protein factory for converting the transcriptome into an in vitro-expressed proteome. *Nat Methods* **5:** 1011–7.
31. Chen S., Hall C., Barbieri J.T. (2008) Substrate recognition of VAMP-2 by botulinum neurotoxin B and tetanus neurotoxin. *J Biol Chem* **283:** 21153–9.
32. Arndt J.W., Chai Q., Christian T., Stevens R.C. (2006) Structure of botulinum neurotoxin type D light chain at 1.65 A resolution: repercussions for VAMP-2 substrate specificity. *Biochemistry* **45:** 3255–62.
33. Agarwal R., Binz T., Swaminathan S. (2005) Structural analysis of botulinum neurotoxin serotype F light chain: implications on substrate binding and inhibitor design. *Biochemistry* **44:** 11758–65.
34. Humeau Y., Doussau F., Grant N.J., Poulain B. (2000) How botulinum and tetanus neurotoxins block neurotransmitter release. *Biochimie* **82:** 427–46.
35. Yamasaki S., Baumeister A., Binz T., Blasi J., Link E., Cornille F., Roques B., Fykse E.M., Sudhof T.C., Jahn R., Niemann H. (1994) Cleavage of members of the synaptobrevin/VAMP family by types D and F botulinal neurotoxins and tetanus toxin. *J Biol Chem* **269:** 12764–72.
36. Pellizzari R., Mason S., Shone C.C., Montecucco C. (1997) The interaction of synaptic vesicle-associated membrane protein/synaptobrevin with botulinum neurotoxins D and F. *FEBS Lett* **409:** 339–42.
37. Sikorra S., Henke T., Galli T., Binz T. (2008) Substrate recognition mechanism of VAMP/synaptobrevin-cleaving clostridial neurotoxins. *J Biol Chem* **283:** 21145–52.
38. Dolman C.E., L. Murakami. (1961) Clostridium botulinum type F with recent observations on other types. *J Infect Dis* **109:** 107–28.
39. Murray R.Z., Kay J.G., Sangermani D.G., Stow J.L. (2005) A role for the phagosome in cytokine secretion. *Science* **310:** 1492–5.
40. Pellizzari R., Rossetto O., Lozzi L., Giovedi S., Johnson E., Shone C.C., Montecucco C. (1996) Structural determinants of the specificity for synaptic vesicle-associated membrane protein/synaptobrevin of tetanus and botulinum type B and G neurotoxins. *J Biol Chem* **271:** 20353–8.

Cancer Science
The official journal of the Japanese Cancer Association JCA

Cancer-testis antigen BORIS is a novel prognostic marker for patients with esophageal cancer

Koji Okabayashi,[1,2] Tomonobu Fujita,[1] Junichiro Miyazaki,[1] Tsutomu Okada,[1] Takashi Iwata,[1] Nobumaru Hirao,[1] Shinobu Noji,[1] Nobuo Tsukamoto,[1] Naoki Goshima,[3] Hirotoshi Hasegawa,[2] Hiroya Takeuchi,[2] Masakazu Ueda,[2] Yuko Kitagawa[2] and Yutaka Kawakami[1,4]

[1]Division of Cellular Signaling, Institute for Advanced Medical Research, and [2]Department of Surgery, Keio University School of Medicine, Tokyo; [3]National Institute of Advanced Industrial Science and Technology, Tokyo, Japan

(Received February 14, 2012/Revised May 29, 2012/Accepted May 30, 2012/Accepted manuscript online June 7, 2012/Article first published online July 16, 2012)

Esophageal squamous cell cancer (ESCC) is one of the most common lethal tumors in the world, and development of new diagnostic and therapeutic methods is needed. In this study, cancer-testis antigen, BORIS, was isolated by functional cDNA expression cloning using screening technique with serum IgG Abs from ESCC patients. BORIS was previously reported to show cancer-testis antigen like expression, but its immunogenicity has remained unclear in cancer patients. BORIS was considered to be an immunogenic antigen capable of inducing IgG Abs in patients with various cancers, including four of 11 ESCC patients. Immunohistochemical study showed that the BORIS protein was expressed in 28 of 50 (56%) ESCC tissues. The BORIS expression was significantly associated with lymph node metastasis in ESCC patients with pT1 disease ($P = 0.036$). Furthermore, the patients with BORIS-positive tumors had a poor overall survival (5-year survival rate: BORIS-negative 70.0% *vs* BORIS-positive 29.9%, *log-rank* $P = 0.028$) in Kaplan–Meier survival analysis and log-rank test. Multivariate Cox proportional hazard model demonstrated that BORIS expression was an independent poor prognostic factor (hazard ratio = 4.158 [95% confidence interval 1.494–11.57], $P = 0.006$). Downregulation of BORIS with specific siRNAs resulted in decreased cell proliferation and invasion ability of ESCC cell lines. BORIS may be a useful biomarker for prognostic diagnosis of ESCC patients and a potential target for treatment including by BORIS-specific immunotherapy and molecular target therapy. (*Cancer Sci* 2012; 103: 1617–1624)

Esophageal squamous cell cancer (ESCC) is one of the most common lethal tumors in the world, and the 5-year survival rate has been reported to be only about 20% due to advance disease, local relapse, and distant metastasis.[1] Despite recent progress in chemotherapy with or without radiotherapy,[2] new diagnostic and treatment methods need to be developed for patients with esophageal cancer.

Immune responses, as evidenced by the intratumoral presence of $CD4^+$ and $CD8^+$ T-cells, have been reported in esophageal cancer patients.[3] In fact, NY-ESO-I, which was originally isolated by cDNA cloning using serum IgG Abs (SEREX: serological identification of antigens by recombinant expression cloning) obtained from esophageal cancer patients, has recently been considered a promising antigen to use as a target for various cancer immunotherapies.[4] NY-ESO-I is one of the cancer-testis (CT) antigens, which are expressed in various cancers but only in germline cells in normal tissues,[4] which is considered to be an immunologically privileged organ, because spermatogenic cells do not express MHC class I or class II molecules on their surface, and a blood-testis barrier consisting of Sertoli cells is present in the seminiferous tubules.[5] Given this theoretical background, a CT-antigen-specific immune response has been considered an ideal reaction, which might lead to tumor-specific destruction.

Many clinical trials of immunotherapies targeting NY-ESO-I have recently been conducted.[6,7] Although no antitumor effect was observed in some trials,[7] a recent trial of immunization with cholesterol-bearing hydrophobized pullulan formulated NY-ESO-I protein showed some antitumor effects in patients with esophageal cancer and induced specific immune responses.[6] Adoptive transfer of NY-ESO-I–specific $CD4^+$ T cells in a melanoma patient resulted in dramatic tumor reduction.[8] In addition to their usefulness in immunotherapy, some CT antigens have been reported to be potential biomarkers.[9] We therefore attempted to identify additional CT antigens that might be useful in developing new diagnostic and therapeutic methods for cancer patients, especially esophageal cancer patients.[9,10]

In this study, we identified a human CT antigen, BORIS, by screening a testis cDNA library with serum IgG from esophageal squamous cell carcinoma (ESCC) patients. We demonstrated that BORIS was involved in ESCC cell proliferation and invasion and was a potential biomarker for esophageal cancer patients with a poor prognosis.

Materials and Methods

Cell lines, tissue specimens, and sera. The cell lines used in the study were esophageal squamous cell carcinoma cell lines, TE2, TE3, TE4, TE5, TE6, TE7, TE8, TE9, TE10, TE11, TE12, TE13, TE14, and TE15 (Tohoku University, Sendai, Japan); melanoma cell lines, SKmel23, SKmel28, 888mel, A375mel, 1363mel, 928mel, 624mel, 501Amel, 586mel, 526mel, 501mel, 397mel, and 1362mel (Surgery Branch, NCI, NIH, Bethesda, MD, USA); colon cancer cell line, COLO205 (JCRB, Osaka, Japan); breast cancer cell line HS578 (American Type Culture Collection (ATCC), Manassas, VA, USA); stomach cancer cell lines, MKN1, MKN7, MKN28, MKN46, and MKN74 (Yamagata University, Yamagata, Japan); endometrial cancer cell line SNGII (Keio University, Tokyo, Japan); prostate cancer cell line LNCaP (ATCC); bladder cancer cell line, KU7 (Keio University); and brain tumor cell line U87MG (ATCC). All cell lines were maintained in 10% FBS RPMI 1640 medium. COS-7, African Green Monkey kidney fibroblast-like cell line, (ATCC) was grown in DMEM supplement with 10% FBS.

Cancerous tissue and adjacent non-cancerous tissue were excised from the surgical specimens of patients who underwent surgery for various cancers at Keio University Hospital without any other preoperative adjuvant treatment, and the tissues were immediately frozen in liquid nitrogen. Sera obtained from cancer patients and healthy volunteers were frozen in freezers

[4]To whom correspondence should be addressed.
E-mail: yutakawa@z5.keio.jp

doi: 10.1111/j.1349-7006.2012.02355.x

maintained at −80°C. These clinical specimens were retrieved between 1999 and 2001. All clinical data were collected from medical records. This study was performed with the approval of the ethics committee of Keio University School of Medicine.

Histopathological findings. Serial 4 μm-thick tissue sections were then fixed with 10% formalin, embedded in paraffin, and stained with H&E. Several paraffin-embedded tissue blocks from the lesion were selected in each case, and Elastica-van Gieson stain was used for the evaluation of vascular invasion. Sections from the selected blocks were stained with monoclonal antibody (Ab) D2-40 immunohistochemistry (IHC) (1:200; Signet Laboratory, Dedham, MA, USA) to evaluate lymphatic invasion as previously mentioned.[11,12]

Histopathological findings were assessed by pathological experts. Tumor stage, including depth of invasion and lymph node metastasis, was categorized according to Union for International Cancer Control (UICC) TNM stage version 6. Histological grade of ESCC was as follows: well differentiated, moderately differentiated and poorly differentiated. Both lymphatic invasion and vascular invasion were evaluated in the entire tumor tissue on several section lines and categorized to positive and negative.

SEREX cDNA cloning of human tumor antigens. SEREX cDNA cloning was performed as previously reported.[4,13,14] Briefly, a normal testis cDNA library containing 1.2×10^7 plaque-forming units was immunoscreened with a mixture of sera (1:100 dilution) from four ESCC patients with stage III. After DNA sequencing of the clones that were isolated, they were analyzed by comparison with genetic databases at the National Center for Biotechnology Information.

Reverse transcription-PCR and quantitative PCR. Total RNA was isolated from esophageal cancer cell lines by using an RNeasy mini kit (QIAGEN GmbH, Hilden, Germany), and it was used to synthesize cDNAs with an oligo (dT). BORIS expression was determined by PCR with the BORIS-specific primers 3′-CAGGCCCTACAAGTGTAACGACTGCAA-5′ and 3′-GCATTCGTAAGGCTTCTCACCTGAGTG-5′. Quantitative analysis of BORIS expression was performed by using Power SYBR Green PCR Master Mix (Applied Biosystems, Foster City, CA, USA). GAPDH was used as an internal control.

Evaluation of immunogenicity using phage plaque assay and ELISA. A phage plaque assay was performed to evaluate the immunogenicity of the antigens isolated in various cancer patients and healthy individuals as previously reported.[15,16] This immunoreactivity of BORIS-specific IgG Abs was evaluated by staining with 1:200 diluted *Escherichia coli*-adsorbed sera from cancer patients, including esophagus, lung, stomach, colon, pancreas, endometrium, ovary, kidney, bladder, prostate, melanoma and acute myelogenous leukemia, and 30 healthy individuals. ELISA to evaluate BORIS specific IgG Ab titers was prepared using the wheat germ cell-free protein system as previously reported.[17] Recombinant BORIS protein was coated to immune plates. After the incubation with sera diluted 1–100 and wash with PBS-Tween, IgG specifically bound to the recombinant BORIS protein was detected by anti-human IgG-HRPO and tetramethylbenzidene.

Immunohistochemical staining. Rabbit polyclonal BORIS-specific Ab, 18337 (Abcam, Cambridge, UK), was used as the primary Ab. To confirm the BORIS specificity of the Ab, specific recognition was evaluated using COS-7 cells transfected with BORIS cDNA. Briefly, the BORIS cDNA were subcloned in the mammalian expression vector, *pcDNA3.1*. COS-7 cells were transfected with the recombinant plasmid *pcDNA3.1-BORIS* using Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's protocol. After 48 h incubation, cells were fixed in 4% paraformaldehyde in PBS for 10 min at room temperature and were incubated at room temperature for 1 h with the BORIS Ab (1/200 dilution), and then incubated at 37°C for 30 min with the secondary goat anti-rabbit IgG. IHC of cancer tissue samples was performed as follows. Paraffin-embedded specimens were incubated at 37°C for 1 h with the BORIS Ab (1/200 dilution), and then incubated at 37°C for 30 min with the secondary goat anti-rabbit IgG. Staining was graded according to the number of positive tumor cells as follows: negative, focal staining or <5% of the cells stained; weak, >5–20% of the cells stained; moderate, >20–50% of the cells stained; strong, >50% of the cells stained. Two independent investigators blinded to the patients' clinical information evaluated all specimens.

siRNA studies. The target sequences of the siRNAs for BORIS were: #1: 5′-UUAAGGUGAUUCCUCAGGAGGGUGA a -3′ and #3: 5′-UUCAGUCUUCAUCUGAAGAAGGGUG 3′ (Invitrogen). We used Stealth RNAi Negative Control Kit with medium GC content (Invitrogen) as a negative control. Esophageal squamous cell cancer cell lines, including TE5 and TE10, were transfected with dsRNAs by using Lipofectamine 2000 (Invitrogen). After silencing for 48 h, cells were re-plated at a density of 3×10^3 cells in a 96-well plate, and cell proliferation was assayed by using WST-1 Cell Proliferation System (Takara, Kyoto, Japan). In the invasion assays, cells were plated in Biocoat Matrigel invasion chambers (BD Biosciences, San Jose, CA, USA) at a cell density of 2.5×10^4 per chamber in serum-free medium supplemented with or (outer chamber) or not supplemented with (inner chamber) 10% FBS. After incubation for 22 h, cells were fixed and counted after staining with Diff-Quik stain (Sysmex, Kobe, Japan).

Statistical analyses. Statistical testing for associations between expression of BORIS protein and various clinicopathological factors was performed by using the Fisher's exact test or Student's *t*-test. Overall survival curves were evaluated by the Kaplan–Meier survival analysis, and the statistical significance was evaluated by the log-rank test. Associations between various factors and survival were assessed by the Cox proportional hazard model for univariate and multivariate analysis, which estimated hazard ratio (HR) and 95% confidence interval (CI), respectively. A P-value < 0.05 was considered a statistically significant difference. All statistical analyses were performed using Stat View software (version 5.0) (SAS Institute., Cary, NC, USA).

Results

Isolation of new cancer-testis antigens by SEREX using sera from patients with esophageal cancer. To isolate novel CT antigens that are expressed in normal testis and various cancers, we screened a normal human testis cDNA library containing more than 1.0×10^6 recombinant clones with a 1:100 diluted mixture of sera from four patients with stage III esophageal squamous cancer, and 39 cDNA clones encoding 13 different antigens were isolated (Table 1). The most frequently isolated antigen was BORIS, a protein homologous to multifunctional transcription factor CCCTC-binding factor (CTCF) with zinc finger domain, which has been reported to be involved in epigenetic reprogramming in germ line cells.[18] BORIS was previously reported to exhibit CT-antigen-like expression,[18] but its immunogenicity had not been demonstrated. Therefore, BORIS was for the first time shown to be a CT antigen recognized by cancer patients' sera IgG Abs. The other CT antigens such as NY-ESO-1 and melanoma antigens (MAGE), which have been previously isolated by the SEREX method were not isolated in this experiment, possibly because these CT antigens are not frequently isolated dominant antigens. The second most frequently isolated antigen was paraneoplastic-protein-like 5 (PNMA5), a member of the paraneoplastic Ma antigen family, which includes PNMA1, PNMA2, PNMA3, and PNMA6. The

doi: 10.1111/j.1349-7006.2012.02355.x

Table 1. cDNAs isolated by SEREX with serum IgG from esophageal squamous cell cancer (ESCC) patients

Gene symbol	No. of clones	Location	Function
BORIS	18	20q13.2	Transcriptional factor
PNMA5	6	Xq28	Unknown
ANKHD1	4	5q31.3	Cytoskeleton
KIF20B	2	10q23.31	Cytokinesis
DPY19L	1	8q22.1	Unknown
PFKP	1	10q15.2	Phosphofructokinase
NSD1	1	5q23	Histone methyltransferase
NOD1	1	7p15-p14	Apoptosis
GCC2	1	2q12.3	Maintaining Golgi structure
GPR108	1	19p13.3	Unknown
CPE	1	4q32.3	Carboxypeptidase
NUDCL	1	7p13-p12	Mitosis
TBC1D4	1	13q21.33	GTPase-activating protein

SEREX, serological identification of antigens by recombinant expression cloning.

antibodies evoked by PNMA1, PNMA2, and PNMA3 were previously considered to be markers for paraneoplastic limbic and brain-stem dysfunction.[19] In this study, we performed a detailed analysis on the immunological and clinical characteristics of BORIS.

Cancer- testis-antigen-like expression of BORIS in various human cancers. Expression of BORIS was first evaluated by RT-PCR in various normal human tissues and cancer cell lines. BORIS was found to be expressed in testis among the normal tissues and in various cancer cell lines and tissues, including 7 of 15 (47%) ESCC cell lines (Fig. 1), 2 of 5 (40%) endometrial cancer cell lines, and 7 of 12 (58%) endometrial cancer tissues (Table 2). Expression of BORIS protein was then evaluated by IHC with rabbit anti-BORIS polyclonal Ab. Specific recognition of BORIS by the Ab was confirmed by specific staining of COS-7 cells transfected with BORIS (Fig. 2A). The BORIS protein staining in normal testis was mainly observed in nucleus of spermatogonia and spermatocytes (Fig. 2B), as previously reported,[18] and different levels of cytoplasmic BORIS protein expression were detected in 28 of 50 esophageal squamous cancers (56%) (negative: 22, weak: 1, moderate: 11, strong: 16) (Fig. 2D–I) without expression in adjacent non-cancerous regions (Fig. 2C). Therefore, BORIS may be a cancer testis antigen frequently expressed in various cancers, particularly in ESCC and endometrial cancer.

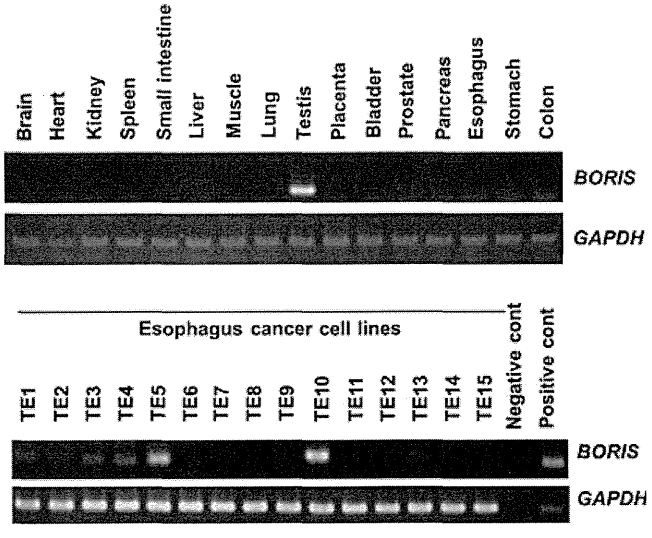

Fig. 1. BORIS is expressed in esophageal squamous cell cancer (ESCC) cell lines and testis among normal tissues. BORIS was expressed in 7 of 15 esophageal cancer cell lines (RT-PCR).

Table 2. Expression of BORIS mRNA in various cancer cell lines and cancer tissues and presence of BORIS-specific IgG in sera from patients with various cancers

Origin	Expression of BORIS (RT-PCR)		Anti-BORIS IgG in sera (positive/total)
	Cell line (positive/total)	Cancer tissue (positive/total)	
Esophagus	7/15		4/11
Lung	5/11		
Adenocarcinoma	3/5	5/9	
Squamous cancer	2/4	1/5	
Small cell carcinoma	0/2	0/1	
Stomach	2/5	3/6	
Colon	2/7	6/10	1/11
Pancreas	1/7	5/17	0/11
Endometrium	2/5	7/12	8/11
Ovary	0/4		
Kidney	1/8	1/4	1/11
Bladder	2/5		1/11
Prostate	1/4		
Melanoma	8/11	5/13	2/11
Acute myelogenous leukemia	0/2		
Healthy individuals			0/30

High immunogenicity of BORIS in patients with ESCC. The immunogenicity of BORIS in patients with various cancers was then evaluated by using the phage plaque assay, which detects BORIS-specific IgG Abs in the serum of patients with various cancers. No BORIS-specific IgG Abs was detected in the serum of 30 healthy donors, but it was detected in serum from patients with various cancers (Table 2). In particular, it was detected in four sera of the 11 (36%) patients with ESCC and eight of the 11 (73%) patients with endometrial cancer.

To evaluate titers of serum BORIS-specific IgG, ELISA was prepared using recombinant BORIS protein generated by wheat germ cell-free system.[17] When the cut-off level was set as above 3 standard deviations of the mean value obtained from 19 healthy individual sera, six of 25 (24%) esophageal cancer patients had positive anti-BORIS IgG Ab, which is comparable with the result of the phage plaque assay in repeated experiments (Fig. 3). However, only one of 25 patients with endometrial cancer was positive for serum BORIS specific Ab. The reason for this discrepancy is not clear. It could be the technical problem of the phage assay or different epitopes recognized by BORIS specific IgG of the endometrial cancer patients. Nevertheless, the presence of BORIS specific Ab in serum of the ESCC patients with esophageal cancer was confirmed by two different assays. Therefore, BORIS was an immunogenic antigen in patients with ESCC and other cancers.

All four ESCC patients positive for BORIS IgG had Stage III disease. Cancer tissue from six of these 11 ESCC patients was available for IHC, but only two of the six cancer tissue samples were BORIS-positive. One of these two patients had a very good outcome (alive and relapse-free at 88 months after surgical excision) and was positive for BORIS-specific serum IgG, but the other patient had a poor outcome (died with liver

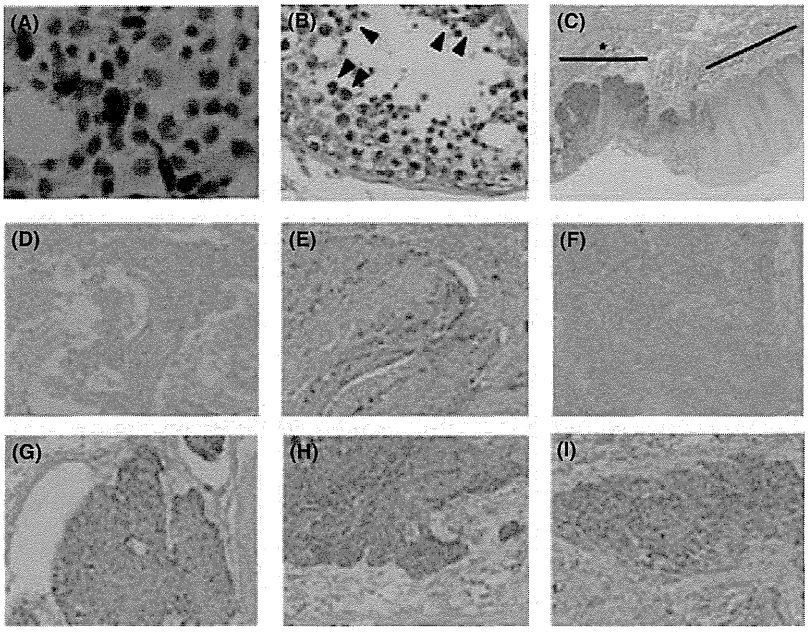

Fig. 2. Differential expression of the BORIS protein in esophageal squamous cell cancer (ESCC) tissues and normal testis. (A) Specific recognition of BORIS transfected COS-7 by the polyclonal Ab. Only BORIS transfected COS-7 cells were stained in immunohistochemistry (IHC). (B) Staining of spermatocytes in normal testis (×400). All the cells related with spermatogenesis (arrow head) were stained. (C) The expression of BORIS in cancer cells (*) is much higher than those normal cells. Heterogenous BORIS staining of ESCC tissues (D–I). (D) negative (×200): focal staining or <5% of the cells stained; (E, F) moderate staining (×200): 25–50% of the cells stained; (G, H, I) strong staining (×200): >50% of the cells stained. BORIS expression was mainly localized to the cytoplasm.

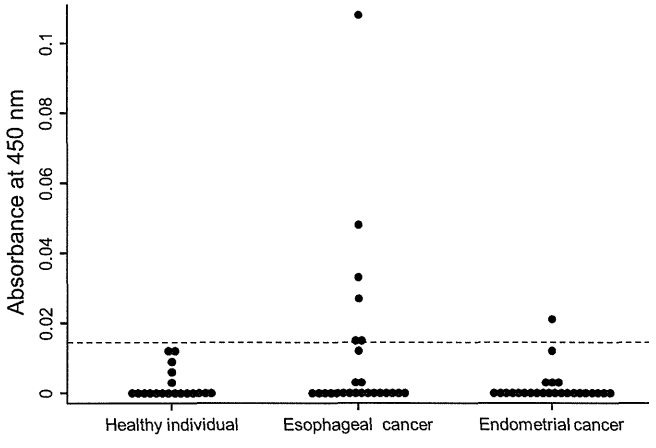

Fig. 3. Enzyme-linked immunosorbent assay (ELISA) for BORIS specific antibody. Scatter plots represent optical density measurements of serum reactivity of healthy individuals, esophageal cancer and endometrial cancer with the purified recombinant BORIS protein. The cut-off of the assay is represented by a dotted line.

and lung metastasis at 18 months after surgery) and was no serum IgG for BORIS. Further study should be conducted in regard to the possibly better prognosis in the group of patients with both BORIS-positive cancer and positive serum IgG.

Poor outcome of patients with BORIS-positive ESCC. Correlations between BORIS protein expression and various clinicopathological features were evaluated by IHC. Greater than "weak" staining was defined as BORIS-positive in the analysis. There was no significant correlation between age, gender, depth of invasion, lymph node metastasis, tumor grade, lymphatic invasion, vascular invasion or stage and BORIS expression (Table 3A). However, incidence of lymph node metastasis in pT1 ESCC expressing BORIS was 65.1% (8/13), which is higher than 29.3–34% reported in the previous reports,[20,21] suggesting that BORIS may play an role in formation of lymph node metastasis in pT1 ESCC. We therefore performed subclass analysis on the relationship between lymph node metastasis and BORIS expression in patients with pT1 ESCC. We found that BORIS expression was significantly correlated with lymph node metastasis ($P = 0.036$) in pT1 ESCC (Table 3B), suggesting that BORIS is associated with metastatic activity of ESCC cells in the early stage. There was no significant association between BORIS expression and lymph node metastasis in patients with pT2/3. The reason for no significant correlation between BORIS expression and lymph node metastasis in all the ESCC patients evaluated may be due to an already high incidence of lymph node metastasis in patients with pT2/3 ESCC. There was no significant correlation between BORIS expression and the number of metastatic lymph node (BORIS-positive 4.7 ± 5.6 *vs* BORIS-negative 3.8 ± 4.7; $P = 0.566$).

Overall survival was also significantly associated with expression of BORIS (5-year survival rate: BORIS-negative 70.0% *vs* BORIS-positive 29.9%, *log-rank* $P = 0.028$) (Fig. 4A). In the patients with pT2/3 who highly involved lymph node metastasis, patients with BORIS expression showed significantly poor prognosis than those without BORIS expression (5-year survival rate: BORIS-positive 0% *vs* BORIS-negative 60%, *log-rank* $P = 0.015$) (Fig. 4B), suggesting that the BORIS expression was associated with the cancer progression in both early and late stages. The univariate analysis showed that depth T2/3, lymph node metastasis, vascular invasion, and BORIS expression were significantly correlated with poor outcome (Table 4). The multivariate analysis using these four factors was the independent prognostic factor for a poor outcome among them (HR = 4.158 [95% CI: 1.494–11.57], $P = 0.006$) (Table 4). Therefore,

doi: 10.1111/j.1349-7006.2012.02355.x

Table 3. Correlation of BORIS expression and various clinicopathological features in patients with esophageal squamous cell cancer (ESCC)

(A) Correlation of BORIS expression and various clinicopathological features

	Total (n = 50)	BORIS positive (n = 28)	BORIS negative (n = 22)	P-value
Age median (range)	59.5 (39–80)	61.5 (39–80)	59 (50–69)	0.277
Gender				
M	49	27	22	>0.999
F	1	1	0	
Depth				
T1	18	13	5	0.202
T2	5	2	3	
T3	27	13	14	
N				
N0	15	6	9	0.214
N1	35	22	13	
Tumor grade				
Well	13	6	7	0.437
Moderately	30	19	11	
Poorly	7	3	4	
ly				
(–)	7	5	2	0.644
(+)	43	23	20	
v				
(–)	21	13	8	0.671
(+)	29	15	14	
Stage				
I	10	5	5	0.437
II	17	11	6	
III	22	12	10	
IV	1	0	1	

(B) Correlation of BORIS expression with lymph node metastasis in T1 stage disease

	BORIS positive (n = 13)	BORIS negative (n = 5)	P-value
T1N0	5	5	0.036*
T1N1	8	0	

*P < 0.05. CRT, chemoradiotherapy; ly, lymphatic invasion; N, lymph node metastasis; v, vascular invasion.

BORIS may be a novel prognostic factor for a poor outcome of patients with ESCC.

Involvement of BORIS in the cell proliferation and invasive ability of ESCC cells. To identify the mechanisms responsible for the increased lymph node metastasis and poor outcome of patients with BORIS-expressing ESCC cells, we evaluated the cell proliferation and invasive ability of BORIS-positive squamous ESCC cell lines TE5 and TE10 that had been treated with BORIS-specific siRNA #1, #3 or control siRNA. BORIS expression was inhibited at least at 96 h after the siRNA transfection by BORIS-specific siRNA #1 and #3 (Fig. 5A). Transfection of TE5 and TE10 with siRNA #1 or #3 inhibited cell proliferation (Fig. 5B), and invasion in a Matrigel invasion assay (Fig. 5C). The molecules involved in typical epithelial to mesenchymal transition (EMT) were not changed after the BORIS specific siRNA transfection, indicating that BORIS may enhance cancer cell invasion not through EMT. These findings suggested that BORIS expression might cause the increased lymph node metastasis and a poor outcome as a result of increased cell proliferation and invasion.

Discussion

BORIS has previously been reported to exhibit cancer-testis-antigen-like expression in various human cancers.[18] Immunization with DNA-based mouse BORIS vaccine results in the generation of anti-tumor $CD8^+$-cytotoxic lymphocytes in murine mammary 4T1 tumor models,[22] but immunogenicity of BORIS had never been reported in humans. In this study, we confirmed that BORIS is an immunogenic antigen in patients with various cancers, particularly ESCC and endometrial cancer. Analysis by RT-PCR showed that BORIS was frequently expressed in esophageal squamous cancer cell lines (7/15, 47%) and endometrial cancer cell lines (2/5, 40%).

The presence of IgG indicated that BORIS-specific $CD4^+$ helper T cells had been induced in those patients. Adoptive transfer of cultured NY-ESO-1–specific $CD4^+$ T cells was recently reported to induce tumor regression in a melanoma patient through induction of $CD8^+$ cytotoxic T cells (CTLs) specific for multiple endogenous tumor antigens.[8] Thus, BORIS-specific $CD4^+$ T cells may be generated and be useful for adoptive immunotherapy by inducing CTLs for multiple tumor antigens. It may also be possible to generate BORIS-specific $CD8^+$ CTLs and develop BORIS-specific active immunization and adoptive immunotherapy. We have attempted to generate BORIS specific CTL, which recognizes cancer cells, but failed to obtain such CTL using five possible HLA-A24 binding peptides predicted by computer programs. Further study is needed to confirm BORIS-specific T cells in patients. In the mouse model, immunization with protein- and DNA-based vaccines was reported to induce BORIS-specific CTL along with BORIS-specific Ab.[23,24]

BORIS is transcriptionally-silenced in normal tissues except germ line cells, but ectopically expressed in various cancer cells,[18,25–27] through DNA hypomethylation of the promoter region[26,28,29]. BORIS was previously reported to express in various normal cells and may be involved in the regulation of cellular functions. However, the expression of BORIS in cancer cells is much higher than those normal cells as shown in our study (Fig. 1). Immunohistochemistry showed that BORIS expression was strongly correlated with metastasis in pT1 ESCC and with poor overall survival in all ESCC patients, indicating that BORIS may be a novel diagnostic biomarker for patients with ESCC. BORIS/CTCF mRNA ratio was recently reported to be significantly associated with DNA hypomethylation and poor prognosis of patients with epithelial ovarian cancer.[30] BORIS has previously been reported to induce other CT antigens, including NY-ESO-1[31,32] and MAGE-A1,[33] by binding to their promoter regions. However, clear correlations between the expression of BORIS and these CT antigens including MAGE-A1 and NY-ESO-1 were not observed in our study (data not shown). These CT antigens have not been reported to have significant impact on the lymph node metastasis and poor prognosis in patients with ESCC.[34–36] Therefore, BORIS appears to be associated with a poor prognosis not through expression of other CT antigens.

Increased lymph node metastasis and poor prognosis in patients with BORIS overexpression may be explained by the increased proliferative and invasive ability in the BORIS expressing ESCC cell lines shown in this study. The decreased ESCC cell proliferation and invasive ability after knockdown of BORIS with siRNA suggests that BORIS may play on a crucial role of ESCC metastasis. In addition, there have been several reports on possible involvement of BORIS in cancer development.[25,37] BORIS and its paralog

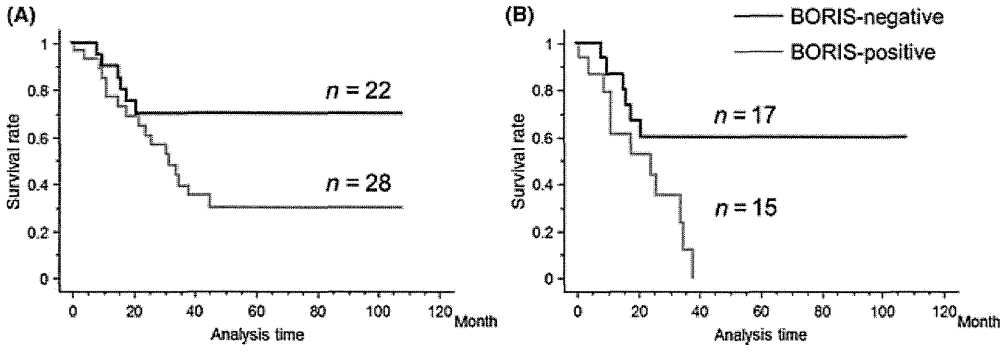

Fig. 4. Poor survival of patients with BORIS positive esophageal squamous cell cancer (ESCC). Overall survival rate was calculated by Kaplan–Meier method. (A) The 5-year overall survival rate was 70.0% in the BORIS-negative group (n = 22) and 29.9% in the BORIS-positive group (n = 28), and the difference between the two groups was significant (n = 0.028). (B) The 5-year survival rate in the patients with pT2/3 (n = 32) was 60% in BORIS-negative group (n = 17) and 0% in BORIS-positive group (n = 15), and the difference between the two groups was significant (P = 0.015).

Table 4. Univariate and multivariate analysis for an independent prognostic factor for esophageal squamous cell cancer (ESCC) patients

	Univariate analysis			Multivariate analysis		
	Hazard ratio	95% CI	*P*-value	Hazard ratio	95% CI	*P*-value
Age	0.987	0.931–1.046	0.650			
Depth						
T1	1	–	–	1	–	–
T2/3	2.711	1.062–6.922	0.037*	2.726	0.785–9.465	0.114
N						
N0	1.000	–	–	1	–	–
N1	3.508	1.186–10.371	0.023*	1.593	0.419–6.057	0.494
Tumor grade						
Well	1	–	–			
Moderately	1.920	0.637–5.794	0.247			
Poorly	1.655	0.416–6.670	0.471			
ly						
(–)	1.000	–	–			
(+)	2.634	0.616–11.263	0.192			
v						
(–)	1	–	–	1	–	–
(+)	2.625	1.075–6.409	0.034*	1.615	0.417–6.250	0.488
BORIS expression						
(–)	1	–	–	1	–	–
(+)	2.722	1.068–6.937	0.036*	4.158	1.494–11.57	0.006**

*P < 0.05; **P < 0.01. CI, confidence interval; HR, hazard ratio; ly, lymphatic invasion; N, lymph node metastasis; v, vascular invasion.

imprinting regulator CTCF,[18,29] are transcription factors containing the same zinc-finger domain that enables them to bind to differentially methylated regions (DMRs) of genomic DNA.[38] The most notable DMR where BORIS and CTCF competitively bind is in the region upstream of non-coding functional mRNA H19.[38–40] CTCF binds the methylated DMR in the methylated paternal allele, which represses H19 transcription.[38,41] However, overexpression of H19 has been reported in BORIS-positive human cancer cells, through binding of BORIS to this DMR region.[42] Knockdown of H19 mRNA in human bladder cancer cell lines caused significant retardation of tumor growth when implanted in nude mice, and inhibited expression of angiogenic factors in liver cancer cell lines.[43] This DMR is differentially methylated in most normal human tissues, with the paternal allele being methylated and the maternal allele unmethylated.[38] Expression of H19 in both paternal and maternal alleles has been identified in 50% ESCC patients.[44] These observations may suggest that BORIS enhances ESCC proliferation, and invasive ability through induction of H19.

It was previously reported that continued high titer of serum anti-p53 Ab after surgery was significantly correlated with poor prognosis in patient with ESCC.[37] In our study, serum anti-BORIS IgG as detected in 36% (4/11) of patients with ESCC and in 73% (8/11) of those with endometrial cancer, respectively. Thus, serum anti-BORIS IgG may also be a useful marker for prognosis of the patients with ESCC and endometrial cancer after surgery, although we were not able to evaluate follow-up of the anti-BORIS IgG titer in this study. Further investigation is required.

In conclusion, BORIS was found to be an immunogenic cancer-testis antigen that is capable of inducing serum IgG in patients with various cancers, particularly ESCC and endometrial cancer. BORIS is also an independent marker of a poor prognosis, possibly because of the increased proliferation and invasive ability of BORIS-positive ESCC cancer cells. BORIS

doi: 10.1111/j.1349-7006.2012.02355.x

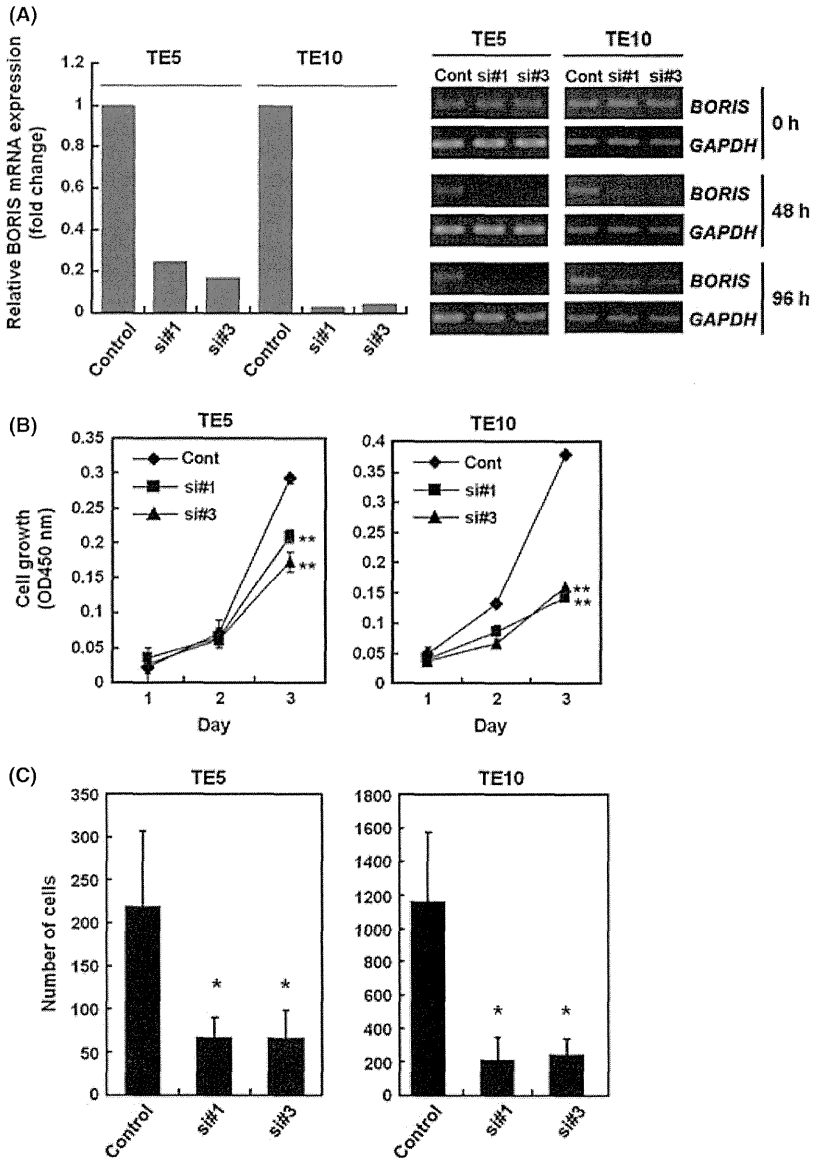

Fig. 5. BORIS is involved in the cell proliferation and invasive ability of esophageal squamous cell cancer (ESCC) cell lines. (A) BORIS expression was inhibited by BORIS-specific siRNA #1 and #3 in ESCC cell lines, TE5 and TE10, shown by qPCR at 48 h (left panel) and reverse transcription-polymerase chain reaction (RT-PCR) at each time point (right panel). (B) Cell proliferation of TE5 and TE10 was significantly inhibited by transfection with BORIS-specific siRNA #1 and #3 when measured by WST-1 assay. The number of cells was determined by absorbance of O.D. 450 nm. $^{**}P < 0.01$ (C) Matrigel invasion assay showed the invasive ability of TE5 and TE10 was significantly inhibited by siRNA #1 and #3. $^{*}P < 0.05$. We used Stealth RNAi Negative Control Kit with medium GC content (Invitrogen) as a negative control.

may therefore be useful in the development of new diagnostic and therapeutic methods for ESCC patients.

Acknowledgments

This work was supported by Grants-in-Aid for Scientific Research (14104013, 17016070, 23240128) and the Project for Development of Innovative Research on Cancer Therapeutics (P-Direct) from the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Japan Society for Promotion of Science, and a Grant-in-Aid for Cancer Research from the Ministry of Health, Labour, and Welfare (15-10, 15-17).

Disclosure Statement

The authors have no conflict of interest.

References

1 Cunningham D, Allum WH, Stenning SP *et al.* Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. *N Engl J Med* 2006; **355**: 11–20.
2 Tepper J, Krasna MJ, Niedzwiecki D *et al.* Phase III trial of trimodality therapy with cisplatin, fluorouracil, radiotherapy, and surgery compared with surgery alone for esophageal cancer: CALGB 9781. *J Clin Oncol* 2008; **26**: 1086–92.
3 Cho Y, Miyamoto M, Kato K *et al.* CD4+ and CD8+ T cells cooperate to improve prognosis of patients with esophageal squamous cell carcinoma. *Cancer Res* 2003; **63**: 1555–9.
4 Chen YT, Scanlan MJ, Sahin U *et al.* A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening. *Proc Nat Acad Sci USA* 1997; **94**: 1914–8.
5 Bart J, Groen HJ, van der Graaf WT *et al.* An oncological view on the blood-testis barrier. *Lancet Oncol* 2002; **3**: 357–63.
6 Uenaka A, Wada H, Isobe M *et al.* T cell immunomonitoring and tumor responses in patients immunized with a complex of cholesterol-bearing hydrophobized pullulan (CHP) and NY-ESO-1 protein. *Cancer Immun* 2007; **7**: 9.
7 Bender A, Karbach J, Neumann A *et al.* LUD 00-009: phase 1 study of intensive course immunization with NY-ESO-1 peptides in HLA-A2 positive patients with NY-ESO-1-expressing cancer. *Cancer Immun* 2007; **7**: 16.

8 Hunder NN, Wallen H, Cao J *et al.* Treatment of metastatic melanoma with autologous CD4+ T cells against NY-ESO-1. *N Engl J Med* 2008; **358**: 2698–703.
9 Hirata D, Yamabuki T, Miki D *et al.* Involvement of epithelial cell transforming sequence-2 oncoantigen in lung and esophageal cancer progression. *Clin Cancer Res* 2009; **15**: 256–66.
10 Ishikawa N, Takano A, Yasui W *et al.* Cancer-testis antigen lymphocyte antigen 6 complex locus K is a serologic biomarker and a therapeutic target for lung and esophageal carcinomas. *Cancer Res* 2007; **67**: 11601–11.
11 Fukunaga M. Expression of D2-40 in lymphatic endothelium of normal tissues and in vascular tumours. *Histopathology* 2005; **46**: 396–402.
12 Kahn HJ, Marks A. A new monoclonal antibody, D2-40, for detection of lymphatic invasion in primary tumors. *Lab Invest* 2002; **82**: 1255–7.
13 Okada T, Akada M, Fujita T *et al.* A novel cancer testis antigen that is frequently expressed in pancreatic, lung, and endometrial cancers. *Clin Cancer Res* 2006; **12**: 191–7.
14 Okada T, Noji S, Goto Y *et al.* Immune responses to DNA mismatch repair enzymes hMSH2 and hPMS1 in patients with pancreatic cancer, dermatomyositis and polymyositis. *Int J Cancer* 2005; **116**: 925–33.
15 Ishida T, Obata Y, Ohara N *et al.* Identification of the HERV-K gag antigen in prostate cancer by SEREX using autologous patient serum and its immunogenicity. *Cancer Immun* 2008; **8**: 15.
16 Naora H, Montz FJ, Chai CY, Roden RB. Aberrant expression of homeobox gene HOXA7 is associated with mullerian-like differentiation of epithelial ovarian tumors and the generation of a specific autologous antibody response. *Proc Nat Acad Sci USA* 2001; **98**: 15209–14.
17 Goshima N, Kawamura Y, Fukumoto A *et al.* Human protein factory for converting the transcriptome into an in vitro-expressed proteome. *Nat Methods* 2008; **5**: 1011–7.
18 Loukinov DI, Pugacheva E, Vatolin S *et al.* BORIS, a novel male germline-specific protein associated with epigenetic reprogramming events, shares the same 11-zinc-finger domain with CTCF, the insulator protein involved in reading imprinting marks in the soma. *Proc Nat Acad Sci USA* 2002; **99**: 6806–11.
19 Voltz R, Gultekin SH, Rosenfeld MR *et al.* A serologic marker of paraneoplastic limbic and brain-stem encephalitis in patients with testicular cancer. *N Engl J Med* 1999; **340**: 1788–95.
20 Kim DU, Lee JH, Min BH *et al.* Risk factors of lymph node metastasis in T1 esophageal squamous cell carcinoma. *J Gastroenterol Hepatol* 2008; **23**: 619–25.
21 Ando N, Ozawa S, Kitagawa Y, Shinozawa Y, Kitajima M. Improvement in the results of surgical treatment of advanced squamous esophageal carcinoma during 15 consecutive years. *Ann Surg* 2000; **232**: 225–32.
22 Loukinov D, Ghochikyan A, Mkrtichyan M *et al.* Antitumor efficacy of DNA vaccination to the epigenetically acting tumor promoting transcription factor BORIS and CD80 molecular adjuvant. *J Cell Biochem* 2006; **98**: 1037–43.
23 Ghochikyan A, Mkrtichyan M, Loukinov D *et al.* Elicitation of T cell responses to histologically unrelated tumors by immunization with the novel cancer-testis antigen, brother of the regulator of imprinted sites. *J Immunol* 2007; **178**: 566–73.
24 Mkrtichyan M, Ghochikyan A, Davtyan H *et al.* Cancer-testis antigen, BORIS based vaccine delivered by dendritic cells is extremely effective against a very aggressive and highly metastatic mouse mammary carcinoma. *Cell Immunol* 2011; **270**: 188–97.
25 D'Arcy V, Pore N, Docquier F *et al.* BORIS, a paralogue of the transcription factor, CTCF, is aberrantly expressed in breast tumours. *Br J Cancer* 2008; **98**: 571–9.
26 Woloszynska-Read A, James SR, Link PA, Yu J, Odunsi K, Karpf AR. DNA methylation-dependent regulation of BORIS/CTCFL expression in ovarian cancer. *Cancer Immun* 2007; **7**: 21.
27 Risinger JI, Chandramouli GV, Maxwell GL *et al.* Global expression analysis of cancer/testis genes in uterine cancers reveals a high incidence of BORIS expression. *Clin Cancer Res* 2007; **13**: 1713–9.
28 Renaud S, Pugacheva EM, Delgado MD *et al.* Expression of the CTCF-paralogous cancer-testis gene, brother of the regulator of imprinted sites (BORIS), is regulated by three alternative promoters modulated by CpG methylation and by CTCF and p53 transcription factors. *Nucleic Acids Res* 2007; **35**: 7372–88.
29 Klenova EM, Morse HC 3rd, Ohlsson R, Lobanenkov VV. The novel BORIS + CTCF gene family is uniquely involved in the epigenetics of normal biology and cancer. *Semin Cancer Biol* 2002; **12**: 399–414.
30 Woloszynska-Read A, Zhang W, Yu J *et al.* Coordinated cancer germline antigen promoter and global DNA hypomethylation in ovarian cancer: association with the BORIS/CTCF expression ratio and advanced stage. *Clin Cancer Res* 2011; **17**: 2170–80.
31 Kang Y, Hong JA, Chen GA, Nguyen DM, Schrump DS. Dynamic transcriptional regulatory complexes including BORIS, CTCF and Sp1 modulate NY-ESO-1 expression in lung cancer cells. *Oncogene* 2007; **26**: 4394–403.
32 Hong JA, Kang Y, Abdullaev Z *et al.* Reciprocal binding of CTCF and BORIS to the NY-ESO-1 promoter coincides with derepression of this cancer-testis gene in lung cancer cells. *Cancer Res* 2005; **65**: 7763–74.
33 Vatolin S, Abdullaev Z, Pack SD *et al.* Conditional expression of the CTCF-paralogous transcriptional factor BORIS in normal cells results in demethylation and derepression of MAGE-A1 and reactivation of other cancer-testis genes. *Cancer Res* 2005; **65**: 7751–62.
34 Akcakanat A, Kanda T, Tanabe T *et al.* Heterogeneous expression of GAGE, NY-ESO-1, MAGE-A and SSX proteins in esophageal cancer: Implications for immunotherapy. *Int J Cancer* 2006; **118**: 123–8.
35 Fujita S, Wada H, Jungbluth AA *et al.* NY-ESO-1 expression and immunogenicity in esophageal cancer. *Clin Cancer Res* 2004; **10**: 6551–8.
36 Akcakanat A, Kanda T, Koyama Y *et al.* NY-ESO-1 expression and its serum immunoreactivity in esophageal cancer. *Cancer Chemother Pharmacol* 2004; **54**: 95–100.
37 Smith IM, Glazer CA, Mithani SK *et al.* Coordinated activation of candidate proto-oncogenes and cancer testes antigens via promoter demethylation in head and neck cancer and lung cancer. *PLoS ONE* 2009; **4**: e4961.
38 Nguyen P, Cui H, Bisht KS *et al.* CTCFL/BORIS is a methylation-independent DNA-binding protein that preferentially binds to the paternal H19 differentially methylated region. *Cancer Res* 2008; **68**: 5546–51.
39 Lyko F, Brenton JD, Surani MA, Paro R. An imprinting element from the mouse H19 locus functions as a silencer in Drosophila. *Nat Genet* 1997; **16**: 171–3.
40 Fedoriw AM, Stein P, Svoboda P, Schultz RM, Bartolomei MS. Transgenic RNAi reveals essential function for CTCF in H19 gene imprinting. *Science* 2004; **303**: 238–40.
41 Kurukuti S, Tiwari VK, Tavoosidana G *et al.* CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. *Proc Nat Acad Sci USA* 2006; **103**: 10684–9.
42 Ulaner GA, Vu TH, Li T *et al.* Loss of imprinting of IGF2 and H19 in osteosarcoma is accompanied by reciprocal methylation changes of a CTCF-binding site. *Hum Mol Genet* 2003; **12**: 535–49.
43 Matouk IJ, DeGroot N, Mezan S *et al.* The H19 non-coding RNA is essential for human tumor growth. *PLoS ONE* 2007; **2**: e845.
44 Hibi K, Nakamura H, Hirai A *et al.* Loss of H19 imprinting in esophageal cancer. *Cancer Res* 1996; **56**: 480–2.

doi: 10.1111/j.1349-7006.2012.02355.x

The EMBO Journal (2012) 31, 4020–4034 | © 2012 European Molecular Biology Organization | All Rights Reserved 0261-4189/12
www.embojournal.org

THE EMBO JOURNAL

Alternative 3′-end processing of long noncoding RNA initiates construction of nuclear paraspeckles

Takao Naganuma[1], Shinichi Nakagawa[2], Akie Tanigawa[1], Yasnory F Sasaki[1], Naoki Goshima[3] and Tetsuro Hirose[1,*]

[1]Functional RNomics Team, Biomedicinal Information Research Center, National Institute of Advanced Industrial Science and Technology (AIST), Tokyo, Japan, [2]RNA Biology Laboratory, RIKEN Advanced Science Institute, Wako, Japan and [3]Biological Systems Control Team, Biomedicinal Information Research Center, National Institute of Advanced Industrial Science and Technology (AIST), Tokyo, Japan

Paraspeckles are unique subnuclear structures built around a specific long noncoding RNA, NEAT1, which is comprised of two isoforms produced by alternative 3′-end processing (NEAT1_1 and NEAT1_2). To address the precise molecular processes that lead to paraspeckle formation, we identified 35 paraspeckle proteins (PSPs), mainly by colocalization screening with a fluorescent protein-tagged full-length cDNA library. Most of the newly identified PSPs possessed various putative RNA-binding domains. Subsequent RNAi analyses identified seven essential PSPs for paraspeckle formation. One of the essential PSPs, HNRNPK, appeared to affect the production of the essential NEAT1_2 isoform by negatively regulating the 3′-end polyadenylation of the NEAT1_1 isoform. An *in vitro* 3′-end processing assay revealed that HNRNPK arrested binding of the CPSF6–NUDT21 (CFIm) complex in the vicinity of the alternative polyadenylation site of NEAT1_1. *In vitro* binding assays showed that HNRNPK competed with CPSF6 for binding to NUDT21, which was the underlying mechanism to arrest CFIm binding by HNRNPK. This HNRNPK function led to the preferential accumulation of NEAT1_2 and initiated paraspeckle construction with multiple PSPs.

The EMBO Journal (2012) **31**, 4020–4034. doi:10.1038/emboj.2012.251; Published online 7 September 2012
Subject Categories: RNA
Keywords: alternative RNA processing; noncoding RNA; nuclear body; RNA-binding proteins

Introduction

Recent postgenomic transcriptome analyses reveal that many nonprotein-coding transcripts, so-called noncoding RNAs (ncRNAs), are transcribed from large portions of mammalian genomes (Carninci *et al*, 2005; Kapranov *et al*, 2007). The limited numbers of long ncRNAs that have been characterized exhibit diverse functions, as well as cell type-specific expression and localization to subcellular compartments (Prasanth and Spector, 2007; Mercer *et al*, 2009; Wang and Chang, 2011). Most of the newly discovered ncRNAs are likely transcribed by RNA polymerase II. However, extensive analyses of the subcellular localization of human transcripts reveal that ncRNAs are enriched in the cell nucleus, suggesting that they play diverse roles in nuclear events (Kapranov *et al*, 2007; Prasanth and Spector, 2007).

The mammalian cell nucleus is highly organized. It is composed of distinct nuclear bodies that contain proteins or RNAs characteristic of particular nuclear processes. To date, ~10 different nuclear bodies have been characterized (Spector, 2006). Several long ncRNAs, such as Xist, Gomafu (Miat), Malat1 (NEAT2), NEAT1 (MENε/β), TUG1, and GRC-RNAs, localize to specific nuclear bodies (Clemson *et al*, 1996, 2009; Hutchinson *et al*, 2007; Sone *et al*, 2007; Sasaki *et al*, 2009; Sunwoo *et al*, 2009; Zheng *et al*, 2010; Yang *et al*, 2011). In particular, Malat1 localizes to nuclear speckles, where it regulates alternative splicing by modulating the phosphorylation status of Serine/Arginine (SR)-splicing factors (Tripathi *et al*, 2010). Malat1 controls growth signal-responsive gene expression through its association with unmethylated polycomb 2 protein (Yang *et al*, 2011).

Paraspeckles are recently discovered nuclear bodies that are usually detected in cultured cell lines as a variable number of foci found in close proximity to the nuclear speckles. Paraspeckles contain characteristic RNA-binding proteins, including paraspeckle protein 1 (PSPC1), RBM14, NONO, CPSF6, and SFPQ (Fox *et al*, 2002; Dettwiler *et al*, 2004; Prasanth *et al*, 2005). PSPC1, NONO, and SFPQ share common domain structures comprised of two RNA-recognition motifs (RRMs). Collectively, these three proteins comprise the *Drosophila melanogaster* behaviour and human splicing (DBHS) protein family (Bond and Fox, 2009).

The discovery of the specific paraspeckle localization of NEAT1 ncRNA opened a new window in paraspeckle research (Chen and Carmichael, 2009; Clemson *et al*, 2009; Sasaki *et al*, 2009; Sunwoo *et al*, 2009). NEAT1 ncRNA are transcribed from a genetic locus called familial tumour syndrome multiple endocrine neoplasia (MEN) type I on human chromosome 11 (Guru *et al*, 1997) and are comprised of two isoform transcripts, 3.7-kb NEAT1_1 (MENε) and 23-kb NEAT1_2 (MENβ). Both RNAs are produced from the same promoter. Alternatively, they can be processed at the 3′-end to produce a canonically polyadenylated NEAT1_1 and a noncanonically processed NEAT1_2. RNase P recognizes the tRNA-like structure and cleaves it to form the nonpolyadenylated 3′-end of NEAT1_2 (Sunwoo *et al*, 2009). The knockdown of NEAT1 ncRNA leads to the disintegration of paraspeckles, suggesting that these ncRNAs serve as a core structural component (Chen and Carmichael, 2009; Clemson *et al*, 2009; Sasaki *et al*, 2009; Sunwoo *et al*, 2009). However, the biological function of paraspeckles and the role(s) of NEAT1 ncRNA remain to be elucidated.

*Corresponding author. Functional RNomics Team, Biomedicinal Information Research Center, National Institute of Advanced Industrial Science and Technology (AIST), 2-4-7 Aomi, Koutou, Tokyo 135-0064, Japan. Tel.: +81 3 3599 8521; Fax: +81 3 3599 8579;
E-mail: tets-hirose@aist.go.jp

Received: 28 April 2012; accepted: 14 August 2012; published online: 7 September 2012

We recently found that paraspeckles were not essential for viability and development in a mouse model under normal conditions, suggesting that they play roles under certain stress conditions (Nakagawa *et al*, 2011). It has been noted that CTN-RNA, an isoform of mCat2 mRNA, is retained specifically in the paraspeckle. Intriguingly, the long 3′-untranslated region (UTR) of CTN-RNA is cleaved by an unidentified endoribonuclease upon exposure to certain stresses, which leads to the export of processed mCat2 mRNA for cytoplasmic translation (Prasanth *et al*, 2005). The CTN-RNA 3′-UTR contains a long inverted-repeat sequence that is capable of forming intramolecular double-stranded RNAs that are A-to-I edited. The hyperedited CTN-RNAs are enriched in the paraspeckles. Thus, paraspeckles are thought to suppress the expression of hyperedited transcripts through nuclear retention (Prasanth *et al*, 2005). Inverted Alu repeat sequences are commonly found in the 3′-UTRs of multiple mRNAs in human cells (Chen *et al*, 2008). This finding suggests that the expression of these transcripts is suppressed by a nuclear retention mechanism.

We previously reported that two paraspeckle-localized DBHS family proteins, SFPQ and NONO, are required for paraspeckle integrity and for the accumulation of NEAT1_2 but not NEAT1_1 (Sasaki *et al*, 2009). These results suggest that NEAT1_1 alone is unable to maintain paraspeckle integrity. By contrast, overexpressed NEAT1_1 is reportedly capable of increasing the number of paraspeckles, which suggests that it is the functional isoform for paraspeckle formation (Clemson *et al*, 2009). An electron microscopic study revealed the location of the NEAT1_2 and NEAT1_1 isoforms. The common NEAT1 region and NEAT1_2 3′-terminal region were located at the paraspeckle periphery, whereas the NEAT1_2 middle region was located in the paraspeckle interior. These findings suggest the importance of NEAT1_2 for the maintenance of paraspeckle integrity (Souquere *et al*, 2010).

In this study, the essential components for paraspeckle formation were determined. Plasmid rescue experiments revealed that NEAT1_2 but not NEAT1_1 is a necessary RNA for *de novo* paraspeckle formation. To analyse the detailed process of paraspeckle formation, we sought to identify unknown paraspeckle components. RNAi analyses identified additional factors, each with distinct roles, which were indispensible for paraspeckle formation. One of the essential PSPs was involved in the alternative 3′-end processing of NEAT1. This protein arrested the canonical NEAT1_1 3′-end processing, which led to preferential selection for the noncanonical processing of the NEAT1_2 3′-end. Our data provide important insights into the process of paraspeckle formation on the specific nuclear-retained long ncRNAs.

Results

The NEAT1_2 ncRNA isoform is essential for paraspeckle formation

We first attempted to clarify which NEAT1 isoform(s) were required for *de novo* paraspeckle formation. MEFs were prepared from NEAT1 knockout mice (MEF−/−) (Nakagawa *et al*, 2011), in which paraspeckles were absent, for rescue experiments with the expression plasmid of either the NEAT1_1 or NEAT1_2 isoform. The expression levels of NEAT1_1 and NEAT1_2 from the plasmids were comparable (Supplementary Figure S1C). Many of the paraspeckle-like foci that were detectable with both anti-SFPQ antibody immunostaining and NEAT1 RNA-FISH appeared when NEAT1_2 but not NEAT1_1 was transiently expressed from the plasmid (Figure 1A; Supplementary Figure S1A and B). This result indicates that NEAT1_2 is an authentic RNA component that is capable of *de novo* paraspeckle formation.

To prove that the rescued foci exhibited characteristics common to endogenous paraspeckles, transfected MEF(−/−) were treated with actinomycin D. Rescued foci did not appear with actinomycin D treatment. Instead, SFPQ–Flag (as a cotransfected marker) and endogenous PSPs relocated to perinucleolar caps, from which NEAT1_2 ncRNA was absent (Figure 1B). Paraspeckles reportedly display actinomycin D-induced disruption and the concomitant relocation of protein components (Fox *et al*, 2002; Shav-Tal *et al*, 2005). In the present study, the overexpression of either NEAT1_2 or NEAT1_1 in NIH3T3 cells led to elevated nuclear paraspeckle numbers; however, NEAT1_2 was more stimulatory than NEAT1_1 (Figure 1C; Supplementary Figure S1D). Taken together, these results indicate that the NEAT1_2 isoform is an essential RNA component for paraspeckle formation, and the NEAT1_1 isoform is not essential but can contribute to paraspeckle formation only when the NEAT1_2 isoform is present.

Identification of new paraspeckle components

We previously reported that two RNA-binding proteins that are essential for paraspeckle formation, NONO and SFPQ, preferentially bind to and stabilize the NEAT1_2 isoform (Sasaki *et al*, 2009). To obtain further insights into the paraspeckle structure, additional PSPs were searched for by employing the human full-length cDNA resource (FLJ Clones) available from the authors' affiliate (Maruyama *et al*, 2012). In this cDNA collection, the intact protein-coding regions of 18 467 human proteins are fused with Venus fluorescent protein. FLJ Clones provides information concerning the intracellular localization of >18 000 human proteins, through the transfection of each cDNA clone (Figure 2A).

Initially, 68 cDNA clones whose products exhibited the typical localization pattern of paraspeckle-like nuclear foci were selected. The identities of the foci were determined by immunostaining the endogenous SFPQ, to see if the Venus signals overlapped with the SFPQ signals (Figure 2B; Supplementary Figure S2A; Supplementary Table S1) but not with the signals of other nuclear bodies (Supplementary Figure S3B). This screening led to the eventual selection of 34 cDNA clones. Endogenous proteins corresponding to the identified cDNA clones were immunostained with their respective antibodies, when available (Figure 2C; Supplementary Figure S2B; Supplementary Table S1). The correct paraspeckle localization of all 27 examined proteins was confirmed, and no false positives were identified.

As a second screen, we confirmed the relocation of each Venus-fusion protein upon actinomycin D treatment. All 34 fusion proteins relocated to the perinucleolar caps (Supplementary Table S1). These caps corresponded to the destination of the endogenous PSPs (Figure 2D; Supplementary Figure S3A), but were distinct from those of the nucleolar or Cajal body proteins (Supplementary Figure S3B).

Therefore, 34 cDNA clones, designated PSP3 through PSP36, were confirmed as new PSPs. Additionally, TARDBP

(TDP43), which was recently reported to interact prominently with NEAT1 ncRNA in the brain from frontotemporal lobar degeneration (FTLD) patients (Tollervey *et al*, 2011), was confirmed to localize to the paraspeckle in HeLa cells by immunostaining of endogenous protein (Supplementary Figure S2B) and detection of TARDBP-Venus localization (Supplementary Figure S2A).

A comparison of all of the PSPs (Table I; Figure 4) indicated that most possessed canonical RNA-binding domains (Burd and Dreyfuss, 1994): 20 proteins with RRMs, two proteins with KH motifs, and five proteins with RGG boxes. Eight proteins possessed one or more zinc-finger motifs, which are involved in RNA binding (Brown, 2005). Three related RNA-binding proteins, FUS, EWSR1, and TAF15, which are known as significant disease-related proteins (Law *et al*, 2006), were all identified as PSPs. We also identified CPSF6, NUDT21, and CPSF7, which are components of the CFIm complex that regulates the 3′-end processing of mRNA, as PSPs.

Although SR-rich splicing factors are commonly enriched in nuclear speckles, SRSF10 was exceptionally enriched in paraspeckles rather than nuclear speckles. The set of PSPs included several abundant heterogenous nuclear ribonucleoproteins (hnRNPs) (A1, A1L, F, H1, H3, K, R, and UL1). Although the numbers were limited, some of the PSPs possessed putative DNA-binding domains, including the AT-hook, Zn finger, homeodomain, and SAP domain. Several of the PSPs possessing RNA-binding domains (e.g., NONO, SFPQ, EWSR1, FUS, TAF15, and HNRNPUL1) are known to be involved in transcription. This observation suggests that the paraspeckle may serve as a platform of transcription and subsequent RNA processing events (see Discussion).

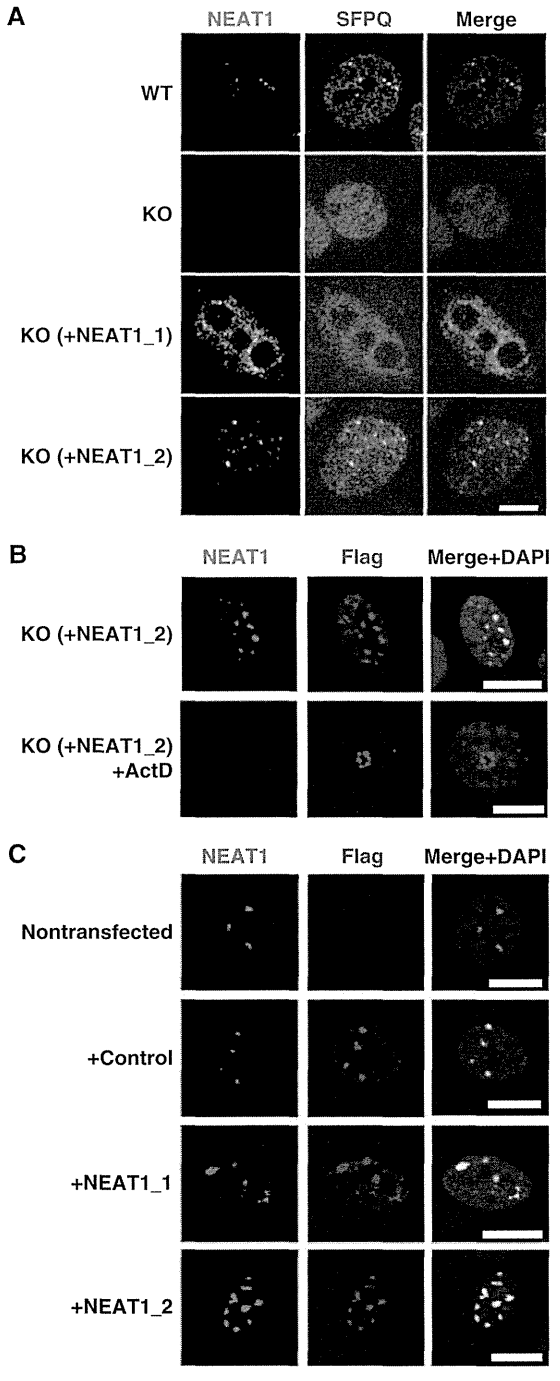

Figure 1 NEAT1_2 is a potent RNA component for paraspeckle formation. (**A**) NEAT1_2 but not NEAT1_1 rescues paraspeckle formation. Intact paraspeckles were detected by RNA-FISH with the antisense probe of mouse NEAT1 ncRNA and coimmunostaining of endogenous SFPQ. Paraspeckles, which were observed in WT MEF cells, were undetectable in MEF cells prepared from NEAT1-knockout mice (KO). Paraspeckles were detected in KO-MEF cells transfected with a plasmid expressing NEAT1_1 (KO + NEAT1_1) or NEAT1_2 (KO + NEAT1_2). (**B**) Effect of actinomycin D treatment on the reformed paraspeckle-like foci. KO-MEF cells were cotransfected with plasmids expressing NEAT1_2 ncRNA and SFPQ–Flag. Transfected cells were treated with 0.3 μg/ml actinomycin D for 4 h. Reformed paraspeckle-like foci were visualized with RNA-FISH of NEAT1 and coimmunostained with anti-Flag M2 antibody. (**C**) NEAT1_2 ncRNA is more competent than NEAT1_1 ncRNA at elevating the number of paraspeckles. NIH3T3 cells were cotransfected with expression plasmids of control (+control), NEAT1_1 (+NEAT1_1), or NEAT1_2 (+NEAT1_2), together with SFPQ–Flag. The counted paraspeckle numbers are shown in Supplementary Figure S1D. Scale bar, 10 μm.

Functional assignments for each PSP identified additional essential factors for paraspeckle formation

To investigate the respective roles of the newly identified and known PSPs in paraspeckle construction, each PSP was knocked down with at least two independent siRNAs (Supplementary Table S2). The resultant changes in paraspeckle appearance (i.e., proportion [%] of cells possessing intact paraspeckles) and in the levels of each NEAT1 isoform (summarized in Supplementary Table S3) were examined.

The PSPs were classified into three distinct categories according to the proportion of paraspeckle-possessing cells after RNAi. After treatment with control siRNA, 88% of the cells examined possessed paraspeckles; this group was defined as the control (Ctl) group (100%). Categories 1, 2, and 3 included PSPs whose RNAi led to a marked decrease (⩽30% of Ctl), substantial decrease (30–75% of Ctl), or no obvious change (⩽75% of Ctl) in the proportion of paraspeckle-possessing cells, respectively (Supplementary Table S3; Supplementary Figure S4). Investigation of the NEAT1 levels by RNase protection assays (RPAs) (Figure 3B) revealed that category 1 could be divided into two subcategories, in which the NEAT1_2 levels were markedly

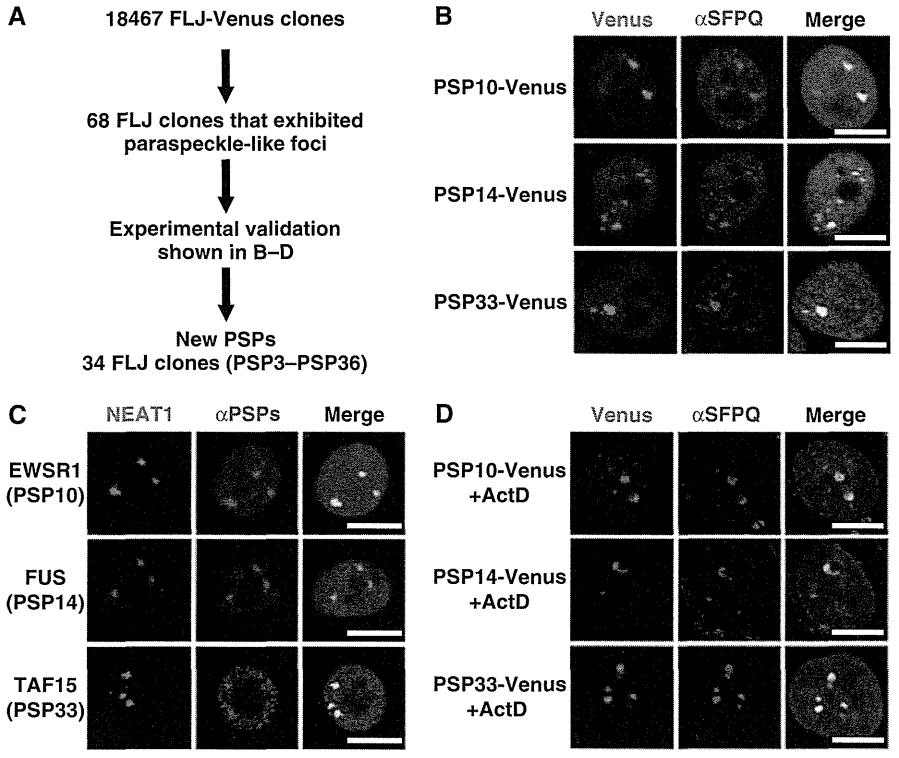

Figure 2 Identification of novel PSP components. (**A**) Experimental strategy to identify new PSPs. (**B**) Selection of FLJ-Venus clones that localize to paraspeckle-like nuclear foci. Paraspeckles were visualized by the immunostaining of SFPQ. Three representatives (PSP10, PSP14, and PSP33) of the new PSPs are shown. (**C**) Confirmation of the paraspeckle localization of endogenous PSP counterparts. Antibodies against each counterpart of PSP10, PSP14, and PSP33 (EWSR1, FUS and TAF15, respectively) were employed to monitor the localization (see Supplementary Table S4). NEAT1 ncRNA was used as a paraspeckle marker. (**D**) Effect of actinomycin D treatment on the localization patterns of selected PSPs. Localization of the Venus clones in **B** was monitored after actinomycin D treatment. SFPQ is an endogenous PSP control. Data regarding other PSPs are shown in Supplementary Figures S2 and S3. Scale bar, 10 μm.

diminished to ⩽30% (category 1A) or were unchanged (category 1B). Similarly, category 3 was divided into two subcategories, in which NEAT1_1 levels were either diminished to ⩽30% (category 3A) or were unchanged (category 3B). The PSP categorization is summarized in Figure 4 and Table I. Representative data from each of the five categories are shown in Figure 3C–G.

Several PSPs were unable to be categorized because their expression was either undetectable in HeLa cells (PSP5, 9, and 23) or they showed inconsistent paraspeckle phenotypes between treatments with the two siRNAs (PSP16, 17, 18, and 32). Several PSPs were tentatively categorized according to the consistent paraspeckle phenotypes between treatments with two siRNAs, although the RPA data were highly variable (PSP10, 15, 25, and 27).

Three PSPs are involved in NEAT1 isoform synthesis by modulating alternative 3′-end processing

Because the NEAT1 isoforms share an identical 5′-terminus, they are likely produced by alternative 3′-end processing. The 3′-ends of NEAT1_1 and NEAT1_2 are formed by two distinct mechanisms: canonical polyadenylation and RNase P cleavage, respectively. The above RNAi experiments identified factors that are involved in this alternative 3′-processing event.

Two PSPs, CPSF6 and PSP24/NUDT21, form a heterodimer (CFIm complex) to facilitate the 3′-end processing of alternatively processed mRNAs (Kim *et al*, 2010). These PSPs also appear to act in NEAT1_1 3′-end processing. We observed that the RNAi of NUDT21 or CPSF6 markedly diminished NEAT1_1 levels and simultaneously increased the NEAT1_2 level (Figure 5A; Supplementary Figure S5B).

The CFIm complex binds to UGUA sequences located upstream of the canonical polyadenylation signal (PAS) and recruits the general 3′-end processing machinery to polyadenylation sites (Venkataraman *et al*, 2005). Sequence searches revealed that five UGUA sequences are clustered 42–169 nt upstream of the PAS (AAUAAA) for NEAT1_1 3′-end processing (Figure 6A). This result strongly suggests that CFIm facilitates the 3′-end processing of NEAT1_1 through binding to the UGUA sequences.

Intact paraspeckles remained detectable after treatment with RNAi for either NUDT21 or CPSF6 (Figure 5B), even though NEAT1_1 was obliterated (Figure 5A; Supplementary Figure S5B). This result confirms that NEAT1_1 is dispensable for paraspeckle formation. Although RNAi elimination of PSP10/EWSR1 upregulated NEAT1_1, this condition did not affect NEAT1_2 (Supplementary Figure S4E). This observation suggests that EWSR1 may control the stability of NEAT1_1.

PSP20/HNRNPK is a new member of category 1A that is required for NEAT1_2 accumulation. Treatment with HNRNPK RNAi disrupted the paraspeckles and decreased the NEAT1_2 level, but simultaneously elevated the NEAT1_1 level (>2-fold) (Figure 5A and B; Supplementary Figure S5B). This finding was not observed with an RNAi of

Table I Paraspeckle proteins

Proteins			Accession	RNA-binding motifs	Other motifs	Category
PSP #	HUGO	Synonyms				
New paraspeckle proteins						
PSP3	AHDC1		Q5TGY3		AT hook	3B
PSP4	AKAP8L		Q9ULX6		2 ZnF C2H2s	3B
PSP5	CELF6	BRUNOL6	Q96J87	3 RRMs		ND
PSP6	CIRBP		Q14011	RRM		3B
PSP7	CPSF7	CFIm59	Q8N684	RRM		2
PSP8	DAZAP1		Q96EP5	2 RRMs		1B
PSP9	DLX3		O60479		homeodomain	ND
PSP10	EWSR1		Q01844	RRM	ZnF RanBP2	3B
PSP11	FAM98A		Q8NCA5			2
PSP12	FAM113A		Q9H1Q7			2
PSP13	FIGN		Q5HY92			2
PSP14	FUS	TLS	P35637	RRM	ZnF RanBP2	1B
PSP15	HNRNPA1		P09651	2 RRMs		2
PSP16	HNRNPA1L2		Q32P51	2 RRMs		ND
PSP17	HNRNPF		P52597	3 RRMs		ND
PSP18	HNRNPH1		P31943	3 RRMs		ND
PSP19	HNRNPH3		P31942	2 RRMs		1B
PSP20	HNRNPK		P61978	3 KHs		1A
PSP21	HNRNPR		O43390	3 RRMs		2
PSP22	HNRNPUL1		Q9BUJ2		SAP, SPRY	2
PSP23	MEX3C		Q5U5Q3	2 KHs	ZnF RING	ND
PSP24	NUDT21	CFIm25	O43809		NUDIX hydrolase	3A
PSP25	RBM3		P98179	RRM		3B
PSP26	RBM4B	RBM30	Q9BQ04	2 RRMs	ZnF CCHC	3B
PSP27	RBM7		Q9Y580	RRM		3B
PSP28	RBM12		Q9NTZ6	3 RRMs		2
PSP29	RBMX		P38159	RRM		3B
PSP30	RUNX3		Q13761			3B
PSP31	SRSF10	FUSIP1, SRp38	O75494	RRM	RS	2
PSP32	SS18L1		O75177			ND
PSP33	TAF15		Q92804	RRM	ZnF RanBP2	3B
PSP34	UBAP2L		Q14157			3A
PSP35	ZC3H6		P61129		3 ZnF C3H1s	3B
PSP36	ZNF335		Q9H4Z2		13ZnF C2H2s	3B
	TARDBP	TDP43, ALS10	Q13148	2 RRMs		2
Known paraspeckle proteins						
	CPSF6	CFIm68	Q16630	RRM		3A
	NONO	p54nrb	Q15233	2 RRMs		1A
	PSPC1	PSP1	Q8WXF1	2 RRMs		3B
	RBM14	PSP2, CoAA	Q96PK6	2 RRMs		1A
	SFPQ	PSF	P23246	2 RRMs		1A

any other category I protein (Supplementary Figure S4A), which suggests that HNRNPK facilitates NEAT1_2 synthesis, rather than stabilization, by modulating NEAT1_1 3′-end processing.

RT–qPCR measurement of NEAT1 ncRNA coimmunoprecipitated with anti-NUDT21 antibody (αNUDT21) revealed that the NEAT1_1/NEAT1_2-overlapped region but not the NEAT1_2-specific region was markedly increased with HNRNPK RNAi (Supplementary Figure S5C and D). This finding indicates that the association of NUDT21 with NEAT1_1 was accelerated by HNRNPK elimination *in vivo*.

The HNRNPK-eliminated cells were transfected with a plasmid for siRNA-resistant HNRNPK (K#2 and K#3 in Figure 5C), and the ratio of the NEAT1_1 and NEAT1_2 isoforms was measured (NEAT1_1/NEAT1_2, referred to as the NEAT1 ratio). As the amount of exogenous HNRNPK increased (lower panels of Figure 5C), the NEAT1 ratio proportionally decreased (Figure 5C). Moreover, exogenous HNRNPK rescued the defect of paraspeckle formation (Figure 5D and E). These results indicate that HNRNPK is responsible for NEAT1_2 synthesis, which determines the NEAT1 ratio and consequent paraspeckle formation.

Previous SELEX analyses have identified CU-rich stretches as preferred binding sequences for HNRNPK (Thisted *et al*, 2001). We identified a CU-rich stretch (UCCCCUU) that perfectly matched a SELEX-derived sequence, which was present in a region adjacent to the canonical PAS (blue box in Figure 6A) that is conserved in rodents. Therefore, HNRNPK likely binds to the CU-rich stretch and interferes with NEAT1_1 3′-end processing, resulting in the preferential synthesis of NEAT1_2. These data indicate that HNRNPK modulates the alternative 3′-end processing for NEAT1_2 synthesis that initiates paraspeckle formation.

HNRNPK binding arrests the CFIm-dependent 3′-end processing of NEAT1_1 in vitro

To investigate how HNRNPK controls alternative 3′-end processing, the NEAT1_1 3′-end processing reaction was recapitulated in HeLa cell nuclear extract (HNE). A ^{32}P-radiolabelled RNA substrate that contained 303 nt spanning the processing site

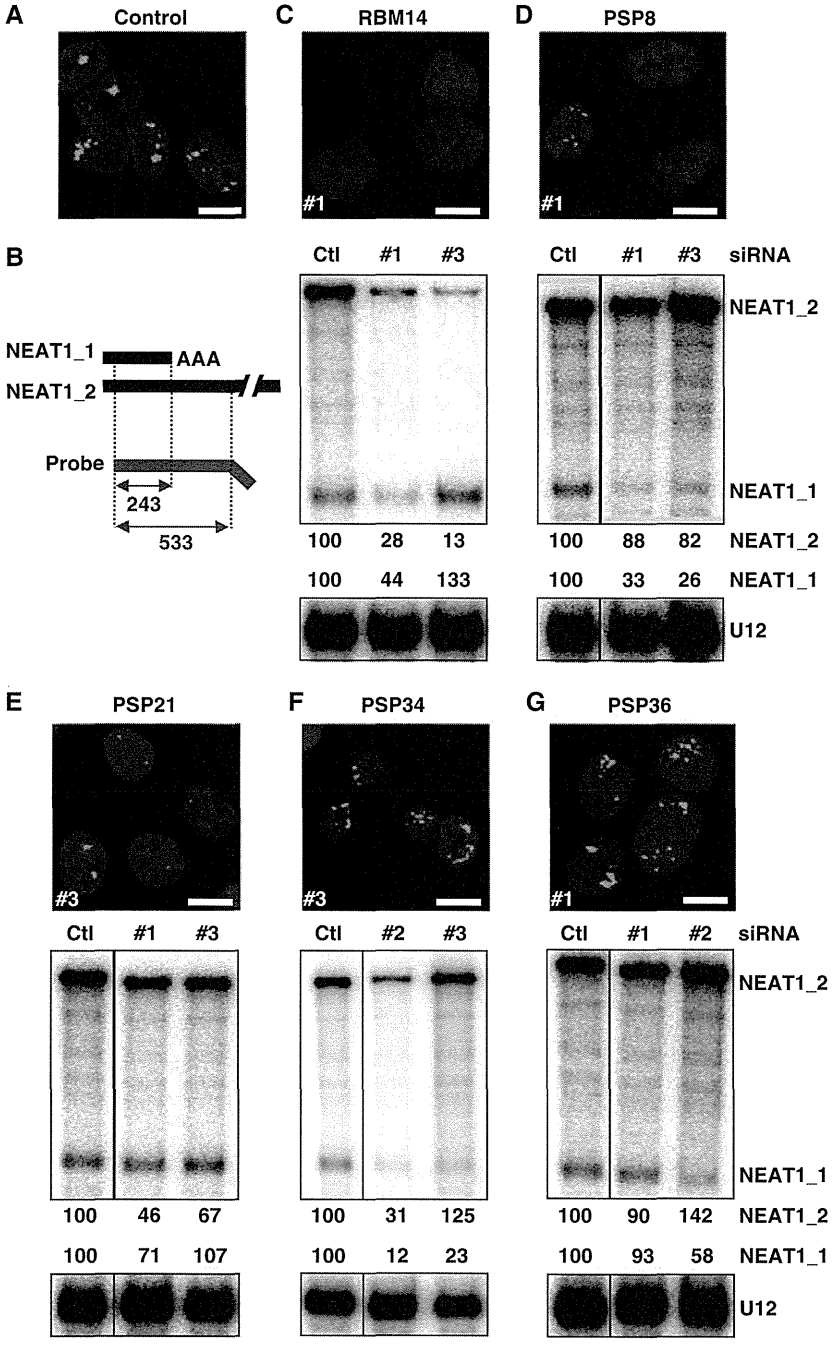

Figure 3 Functional assignment of new PSPs in paraspeckle formation by extensive RNAi treatment. Paraspeckle appearance and NEAT1 levels were monitored by RNA-FISH and RPA to detect NEAT1_1 and NEAT1_2. (**A**) Paraspeckles in cells treated with control siRNA. (**B**) Schema for the RPA probe used and the protected fragments (with size, nt) of NEAT1_1 and NEAT1_2. Data for a representative from each category (1A: RBM14, 1B: PSP8/DAZAP1, 2: PSP21/HNRNPR, 3A: PSP34/UBAP2L, 3B: PSP36/ZNF335) are shown in (**C–G**), respectively. The siRNA numbers (#) used in the RNA-FISH analysis are shown at the lower left of each photo. For RPA, the ratio of band intensities of the two isoforms, normalized by those of U12 snRNA, is shown below (Ctl was defined as 100%). RNAi data regarding all PSPs are compiled in Supplementary Figure S4. Their quantified data are shown in Supplementary Table S3. The siRNAs used are shown in Supplementary Table S6. Scale bar, 10 μm. Figure source data can be found with the Supplementary data.

of NEAT1_1 (Figure 6A) was incubated in HNE. Because no Mg^{2+} was added to the *in vitro* reaction, the endonucleolytic cleavage solely produced the 209-nt processed RNA without subsequent polyadenylation. The cleavage product was detectable after incubation for 30 min (Figure 6B; Supplementary Figure S6A). Processed RNA was not generated from an RNA substrate with a mutated PAS (PAS-mut in Supplementary Figure S6A), which indicated that accurate RNA processing was recapitulated *in vitro*.

To check the roles of CFIm and HNRNPK in the 3′-end processing of NEAT1_1 *in vitro*, RNA substrates in which the putative CFIm-binding sequences (CFBS) or HNRNPK-binding sequence (KBS) were mutated (CFIm-mut and K-mut, respectively, in Figure 6A) were applied to the *in vitro*

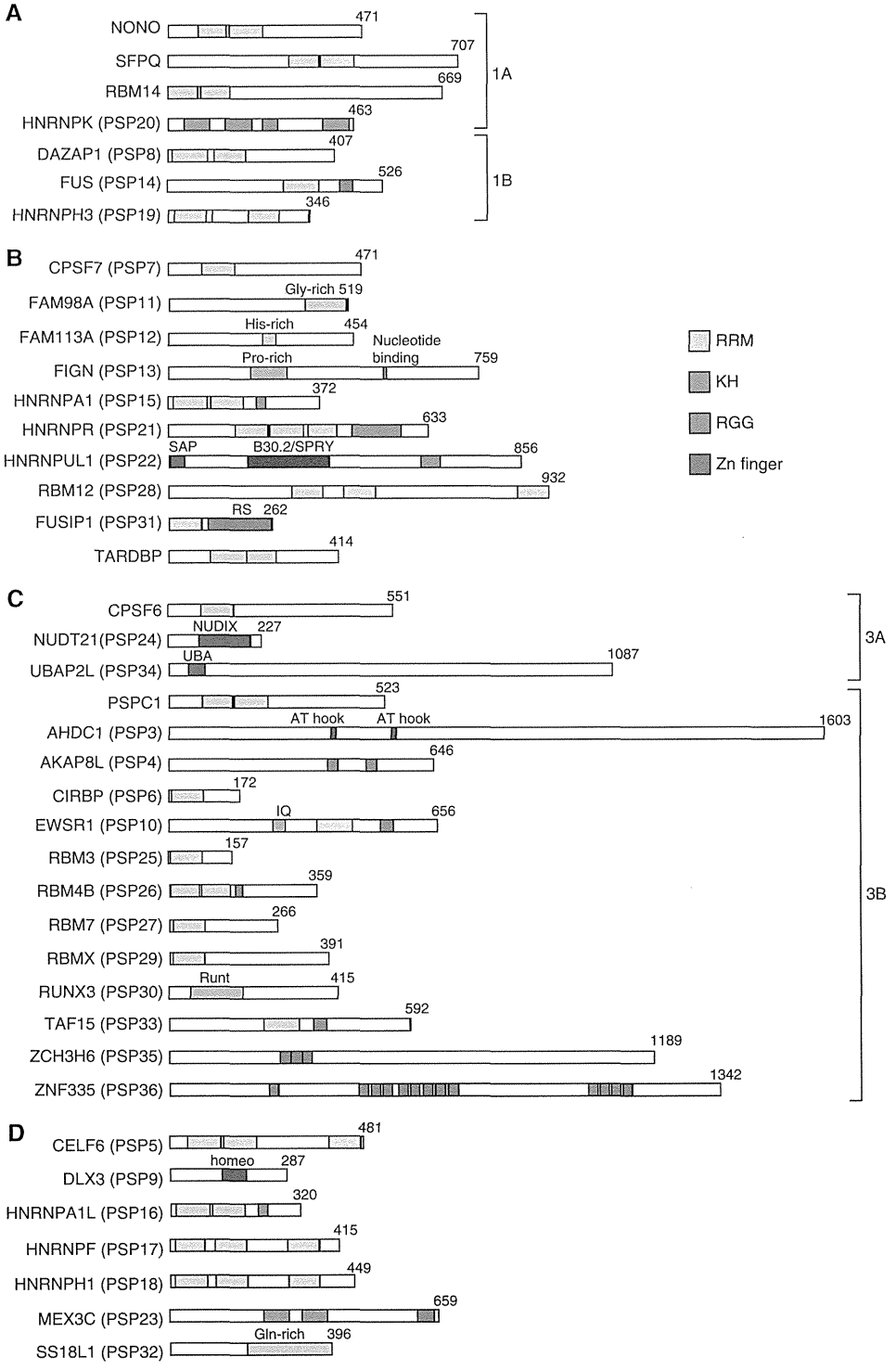

Figure 4 Compilation of PSPs. Schematics of the major domains of the PSPs that belong to category 1 (**A**), 2 (**B**), and 3 (**C**) are grouped and shown. Uncategorized PSPs are shown in (**D**). Subcategories in categories 1 and 3 are shown as 1A and 1B in (**A**) and 3A and 3B in (**C**). The amino-acid length of each PSP is shown in the right corner. The colour codes of four putative RNA-binding domains are shown on the right.

processing reaction. Time-course experiments revealed that CFIm-mut exhibited marked deceleration of the processing compared with the wild-type substrate (WT) (Figure 6B and C), confirming the reported evidence that CFIm facilitates 3′-end processing through its association with CFBS (Venkataraman *et al*, 2005). By contrast, K-mut accelerated *in vitro* processing (Figure 6B and C), which supported our *in vivo* result that HNRNPK acts to suppress the 3′-end processing of NEAT1_1. The results of a gel mobility shift assay with recombinant HNRNPK protein (r-K) confirmed

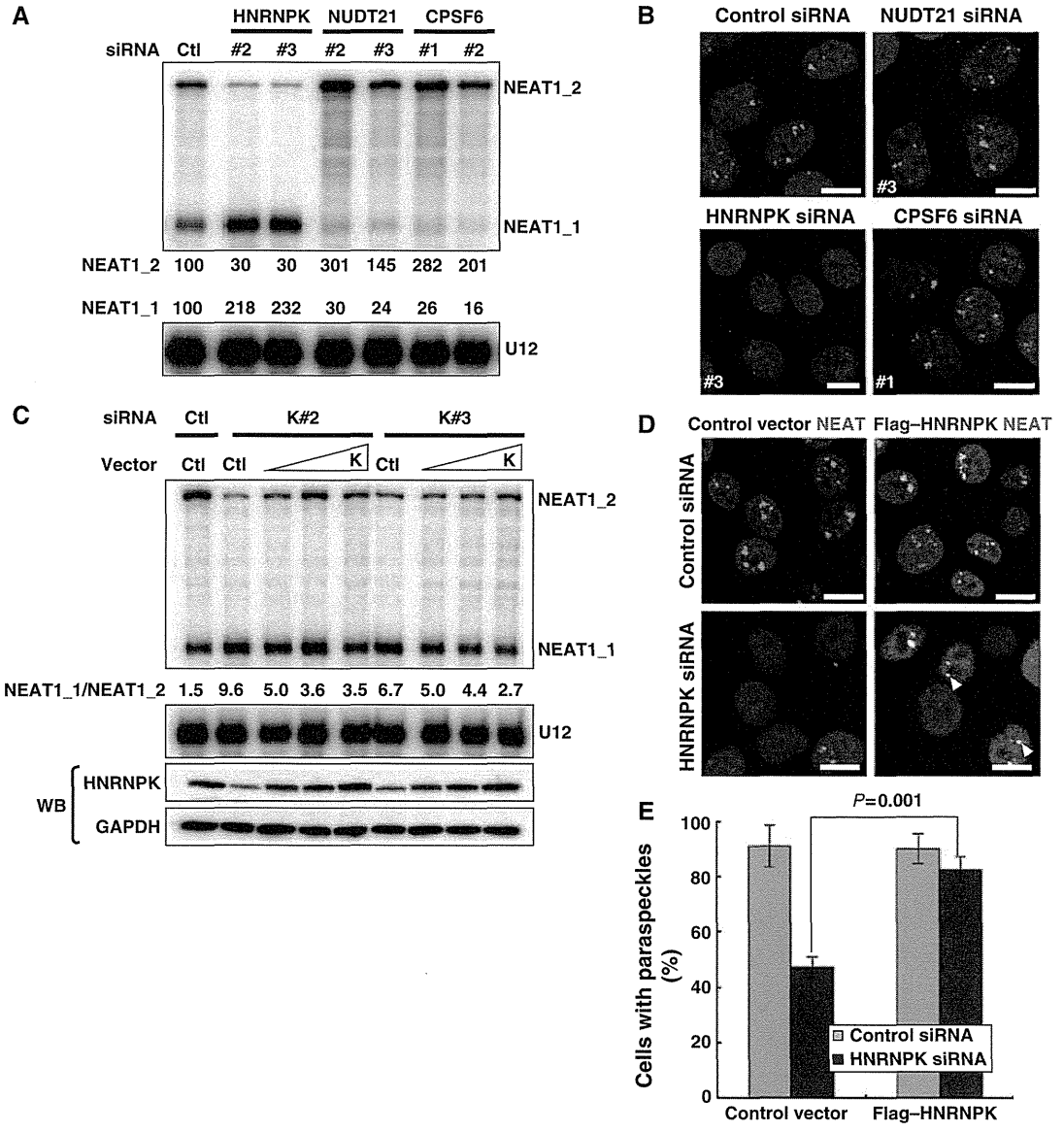

Figure 5 Alternative 3′-end processing of NEAT1 is the initial essential step underlying paraspeckle formation. (**A**) PSPs required for the alternative 3′-end processing of NEAT1. RPA was performed as in Figure 3. Data are shown for NUDT21 and CPSF6, which are required for NEAT1_1 3′-end processing, and HNRNPK, which is required for NEAT1_2 synthesis by interfering with NEAT1_1 3′-end processing. (**B**) Paraspeckle appearance in HeLa cells treated with siRNAs against NUDT21, CPSF6, or HNRNPK. (**C**) Plasmid rescue of a defect of NEAT1_2 synthesis in HNRNPK-eliminated cells. Two siRNAs against HNRNPK (K#2 and K#3) were used to eliminate endogenous HNRNPK, and HNRNPK rescue plasmid (K) was introduced at three concentrations (1–5 μg). NEAT1 ncRNA levels were measured by RPA as in Figure 3. The ratios of NEAT1_1 to NEAT1_2 (NEAT1_1/NEAT1_2) are shown below the upper panel. GADPH and HNRNPK were detected by western blotting (WB). (**D**) Paraspeckle formation is rescued by plasmid expression of HNRNPK. The siRNAs and rescue plasmids used are shown on the left and top, respectively. Paraspeckles were detected by RNA-FISH of NEAT1. Transfected cells were visualized by immunostaining with αFlag. Arrowheads indicate paraspeckles that formed in the rescued cells. Scale bar, 10 μm. (**E**). Quantification of the results in (**D**). Cells possessing more than one paraspeckle-like focus were counted. Total cell numbers counted are the siRNAs used are shown in Supplementary Table S6. *P*-value was calculated by Student's *t*-test. The cell numbers counted for control and HNRNPK-eliminated cells were 152 and 136, respectively.

that r-K binding was mostly abolished by the KBS mutation (Figure 6D), indicating that HNRNPK binds to KBS and arrests the CFIm-dependent 3′-end processing of NEAT1_1.

HNRNPK arrests the RNA binding of CFIm for NEAT1_1 3′-end processing

We examined the effect of HNRNPK on the RNA binding of CFIm during *in vitro* processing. Proteins that bound to WT during *in vitro* processing were visualized by UV-crosslinking (Figure 7A, lanes 1 and 4). Immunoprecipitation with αNUDT21 and αCPSF6 revealed that the UV-crosslinked ~68- and ~25-kDa proteins were efficiently precipitated by each antibody (Figure 7A), indicating that they corresponded to CPSF6 and NUDT21, respectively.

To assess whether the extra HNRNPK interfered with the RNA binding of CFIm, r-K was added to the *in vitro* processing. On the WT substrate, r-K markedly interfered with the UV-crosslinking of CPSF6 and NUDT21 in a