RAPID COMMUNICATION Jpn. J. Cancer Res. **93**, 1183–1186, November 2002

Identification of the *CAB2/hCOS16* Gene Required for the Repair of DNA Double-strand Breaks on a Core Amplified Region of the 17q12 Locus in Breast and Gastric Cancers

Masahiko Nezu,[1,2] Michiko Nishigaki,[1] Tomoki Ishizuka,[1] Yoshitaka Kuwahara,[1] Chikako Tanabe,[1] Kazuhiko Aoyagi,[1] Hiromi Sakamoto,[1] Yasushi Saito,[2] Teruhiko Yoshida,[1] Hiroki Sasaki[1] and Masaaki Terada[1,3]

[1]Genetics Division, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045 and [2]The Second Department of Internal Medicine, Chiba University School of Medicine, 1-8-1 Inohana, Chuo-ku, Chiba 260-0856

We previously reported that *CAB1* and c-*ERBB-2* genes were found to be located in a core amplified region of the 17q12 locus, which is frequently amplified in various cancers. During identification of this core region, *CAB2*, a human homologue of the yeast *COS16* required for the repair of DNA double-strand breaks was cloned. Autofluorescence analysis of cells transfected with its GFP fusion protein demonstrated that CAB2 translocates into vesicles, suggesting that overexpression of CAB2 may decrease intercellular Mn^{2+} by accumulating it in the vesicles, in the same way as yeast COS16. This is the first report identifying all of the genes on the core amplified region of the 17q12 locus in breast and gastric cancers.

Key words: Gene amplification — CAB2/hCOS16 — 17q12

Amplification of various oncogenes has been examined in different types of human cancers, and in some cases a good correlation between oncogene amplification and the prognosis of particular types of cancer has been demonstrated.[1)] We previously reported that *HST1* and *INT2* on chromosome 11q13 were amplified in about 40% of esophageal cancers.[2–4)] By cosmid walking and exon trapping, we further demonstrated that this amplicon contained at least five genes, *HST1*, *INT2*, *MB38*, *EXP1*, and *cyclin D1*.[5,6)] It was reported by other laboratories that the *N-MYC* locus at 2p24, frequently amplified in neuroblastoma, contained the DEAD box gene, *HuDBP-RB/DDX1*,[7)] that the 12q13-14 locus amplified in human sarcoma or glioma contained six genes including *CDK4*, *SAS* and *MDM2*,[8,9)] and that the 20q12-13 locus amplified in breast and ovarian cancers contained the *BTAK* and *AIB1* genes.[10,11)] Amplification of c-*ERBB-2* was found in various cancers including well-differentiated types of gastric cancers.[12)] We previously cloned four genes, *CAB1*, *GRB7*, *A39* and *C51*, from 500 kb of yeast artificial clone (YAC) DNA containing the c-*ERBB-2* gene located on the 17q12 locus.[13,14)] These facts, together with the results from our laboratory and those of other laboratories, established that there are usually multiple expressed genes in a single amplicon, providing a growth advantage to cells. Therefore the identification of genes with increased expression in a core amplified region will be required.

[3] To whom correspondence should be addressed.
E-mail: mterada@ncc.go.jp

The predicted amino acid sequence of the *CAB1* gene had significant homology to the steroidogenic acute regulatory protein, StAR, which has an essential role in cholesterol transport to mitochondria. *GRB7* is an SH2 domain protein, and may be a signaling adaptor molecule of receptor-type tyrosine kinases.[15–17)] A putative gene, *C51* has been found to be derived from an alternative transcript of c-*ERBB-2*, transcribed from a novel promoter.[18)] These two genes, *CAB1* and c-*ERBB-2*, were reported to be always amplified in DNA samples of well-differentiated gastric cancer and breast cancer, demonstrating that they are present in a core amplified region on the 17q12 locus.[13,14)] The gene map is shown in Fig. 1A. During the identification of a major transcriptional start site of the novel promoter of the c-*ERBB-2* gene, approximately 3 kb and 4 kb mRNAs that hybridized with a DNA fragment containing an unknown exon were found to be overexpressed in 5 cancer cell lines, UACC-893, UACC-812, BT-474, SKBR-3 and SKOV-3, with amplification of both *CAB1* and c-*ERBB-2*, but not in 6 other cancer cell lines, BT-20, MDA-MB-436, ZR75-1, T47-D, MCF-7 and OVCAR-3 without amplification (Fig. 1B), demonstrating that a novel gene is present between *CAB1* and c-*ERBB-2*. This novel amplified gene, which must transcribe in the opposite direction to the c-*ERBB-2* gene, was designated as *CAB2* (a gene co-amplified with c-*ERBB-2*).

We cloned three cDNAs with the size of about 3 kb from a size-fractionated cDNA library of UACC-812 cells, and sequenced them. The full-length *CAB2* cDNA (AB088396) was found to consist of 2733 bp, encoding a

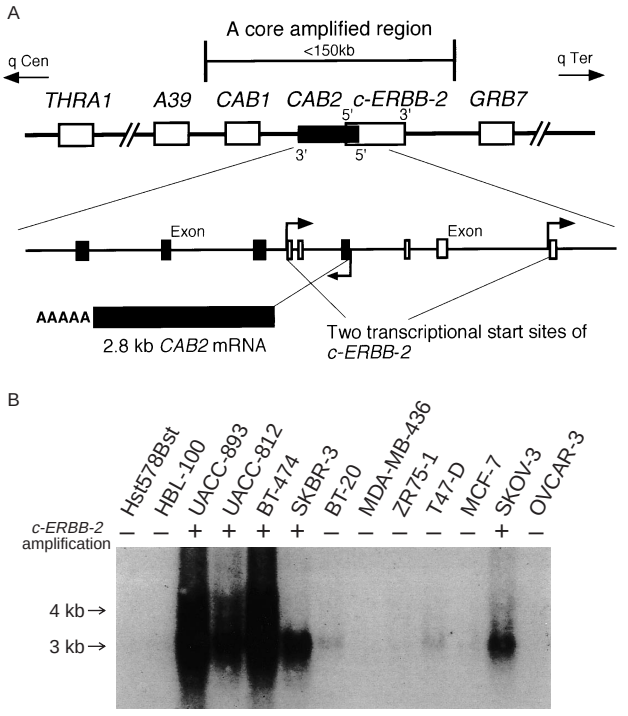

Fig. 1. Cloning of the *CAB2* gene from a core amplified region of the 17q12 locus. (A) Schematic presentation of gene map of the 17q12 region. The order and direction of *A39*, *CAB1*, *CAB2*, c-*ERBB-2* and *GRB7* genes were determined with chromatin fiber fluorescence *in situ* hybridization (FISH) of the EMBL3 clones containing each gene, which were derived from the YAC clone containing the c-*ERBB-2*.[14] (B) Northern blot analysis of the *CAB2* gene in the cancer cell lines. Ten micrograms of total RNA from 2 cell lines derived from mammary epithelial cells, Hst578Bst and HBL-100, 9 breast cancer cell lines, UACC-893, UACC-812, BT-474, SKBR-3, BT-20, MDA-MB-436, ZR75-1, T47-D and MCF7, and 2 ovarian cancer cell lines, SKOV-3 and OVCAR-3, was hybridized with a genomic DNA fragment containing the sequence corresponding to exons of the *CAB2* gene. +, a cell line with c-*ERBB-2* amplification; −, a cell line without c-*ERBB-2* amplification.

```
H. sapiens      1:M-AGLAARLVL-LAGAAALASGSQGD-REPVY-RDCVLQC---EEQ--NCSGGALNHFR- 50
C. elegans      1:MYIKWVVGCLL-ITCFVQL-EASPGD-RSIWY-QECTQVC---ISKY-NCS-TKFGTL-- 49
S. cerevisiae   1:MRLAVVVTLLVHCFLVTCSPGDNLDEFIDCTYACEYNRRCPNSQINYIDPETNMFHDIEF 60
                  *    ** ***          * ** *   *  *    *     * ***   *

H. sapiens     51:-SRQPIYMSLAGWTCRDDCKYECMWVTVGLYLQEGHKVPQFHGKWPFSRF-L-F-F--QE 104
C. elegans     50:-DWARG-DCF--W-CR----YDCMWDTIGHFDSNFGVVPQFHGKWPFLAIPLPFGFIIQE 100
S. cerevisiae  61:FDTPPLYSKLLFWDCISDCDYQCQHIITRWRIDEEEEIYQFHGKWPFLRVLGTQEFFSTI 120
                  * *  *  * ** ** * *** * *    *   ************      *  **

H. sapiens    105:-PASAVASFLNGLASLVMLCR---YRTFVPASSPMYHTCVAFAWVSLNAWFWSTVFHTRD 160
C. elegans    101:-PASMIFSLLNLFTVYKMLRR---FKKMQNLPNRTMW--LVYAHVGMFTWISSSLFHMFD 154
S. cerevisiae 121:FSIGNFIPHYKGFVKFSRIIREEGDRRRKNSRSILIWNYLYVTVAGMLAWTASSVFHCRD 180
                  ***   * ** *    ** *    *   *  *      *  * *** **  ***** **

H. sapiens    161:TDLQRKWTTSVPPVSYTQSTCAASGPWGCSTQ-L-WSSA-FRALLLLMLTVHVSYLSLIR 217
C. elegans    155:CDFTEK-MDYFGAYSFVLFALYVSVIFTKQLQFLGRGGPKYIQILFALVYLN-HFMKMMQ 212
S. cerevisiae 181:LIITEKLDYFFAGLTVLTGFHAIFARMTSMFLY-PKIAQAFTASVAAIFALHILRLYVDW 239
                  * ***    *  *          *   *   * *        * * *   **   *

H. sapiens    218:-FDYG--YN--LVANVAIGLVNVVWWLAWCLWNQRRLPHVRKC--VVVV--LLL-QGLS- 266
C. elegans    213:NFDYG--YN--MTCCIVFSLITTCLYVHHLYYRKRNLGSLQESD-IVLIR-LIIWANLST 266
S. cerevisiae 240:SYTYNMRFNIFFGVLQYILLIMLSCQNYHALQKQKLMGEFKKTAYSSFKRQIFKLCVIPI 299
                  ****  **        * **       * *   * *    *    *    *     **

H. sapiens    267:LL---ELLDF-PPLFWVLDAHAIWHISTIPVHVLFFSFLE-DDSLYLLKESEDKFKLD   319
C. elegans    267:AL---EILDF-TPVFWIFDSHSLFHLATIPIPIWWSDFLALYYDLDVASGKRSVLKMA   320
S. cerevisiae 300:LLVIVTTMAMSLELFDFFSYEWQIDAHALWHLCTIWPSWVLYDFFLEDYAYWGNRQLY   357
                  **   * ***  **** ** *      ***       ** *** *          *
```

Fig. 2. Comparison of the predicted amino acid sequences of CAB2 among different species. Residues from 1 to the end of CAB2 are shown aligned with corresponding regions of other sequences. The standard single-letter symbols for amino acid residues are used. Identical residues of at least two species among the CAB2 sequences of human, *C. elegans*, and *S. cerevisiae* are indicated by asterisks underneath.

319-amino-acid protein. Interestingly, *CAB2* was conserved among the amino acid sequences of the *Saccharomyces cerevisiae* gene *COS16* (AL355013) and *Caenorhabditis elegans* (NM072166) with 33% and 28% overall sequence identity, respectively (Fig. 2). We also found homologous sequences in a plant, *Arabidopsis thaliana* (AAAB01008960), with 38% identity, and *Drosophila melanogaster* (NM136379) with 37% identity (data not shown). By Southern blot analysis of the newly isolated cDNA, amplified signals were detected not only in cancer cell lines, but also in primary gastric cancer tissues with amplification of both *CAB1* and c-*ERBB-2* (data not shown), and 4 exons were identified in the 5′-upstream region of the novel promoter of c-*ERBB-2* (Fig. 1A), which has been sequenced previously.[18] These data clearly showed that *CAB2* is transcribed from the downstream region of the novel promoter of c-*ERBB-2* in the reverse direction (Fig. 1A). Expression of *CAB2* was examined in 21 adult human tissues: adrenal medulla, thyroid, adrenal cortex, stomach, spleen, thymus, prostate, testis, ovary, small intestine, colon, peripheral blood leukocytes, heart, brain, placenta, lung, liver, skeletal muscle, kidney, pancreas and mammary gland, and 4 fetal tissues: brain, lung, liver and kidney. Ubiquitous expression was observed, and the expression was high in the thyroid and placenta out of the 21 adult tissues and in all 4 fetal tissues (data not shown).

The yeast gene *COS16* has recently been isolated as a recessive suppressor gene of *cdc1* (Ts) temperature-sensitive growth, and it encodes a vacuolar membrane protein that appears to regulate intercellular Mn^{2+} homeostasis.[19] To compare the structures of COS16, CAB2 and a candidate homologue of *C. elegans*, transmembrane domains were predicted according to the Kyte and Doolittle hydropathy algorithm.[20] As shown in Fig. 3A, CAB2 and the candidate homologue for *C. elegans* are probably membrane proteins with about 10 transmembrane domains. Although overall sequence similarity among these three genes was not high, several key features such as size, membrane topology and some highly conserved amino acid sequences (for example, QFHGKWPF) were retained. Thus, we concluded that *CAB2* is a good candidate for a

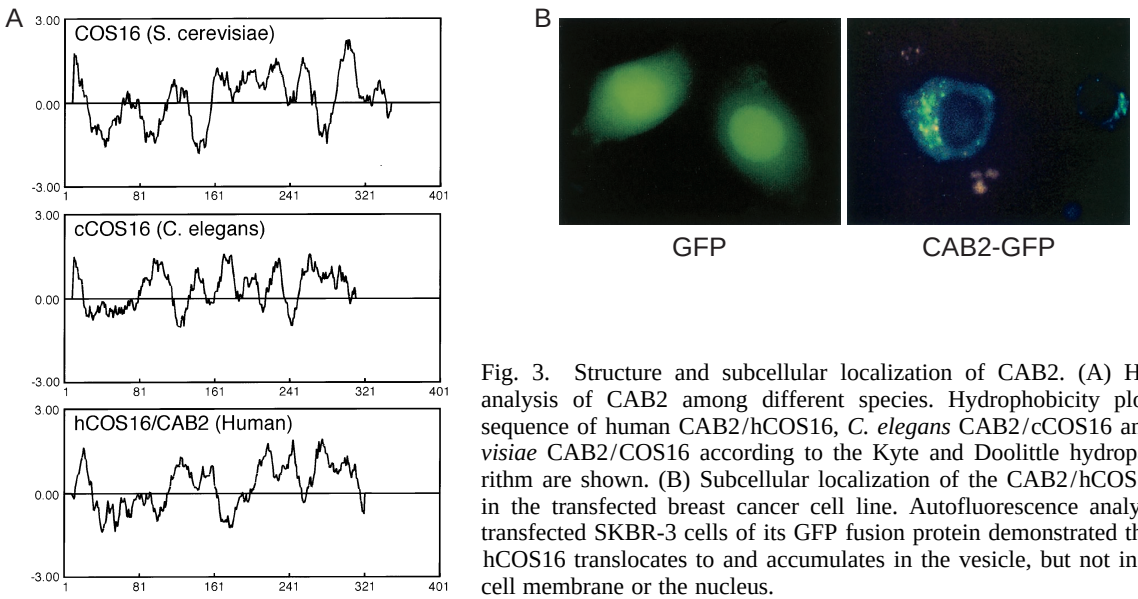

Fig. 3. Structure and subcellular localization of CAB2. (A) Hydropathy analysis of CAB2 among different species. Hydrophobicity plots of the sequence of human CAB2/hCOS16, *C. elegans* CAB2/cCOS16 and *S. cerevisiae* CAB2/COS16 according to the Kyte and Doolittle hydropathy algorithm are shown. (B) Subcellular localization of the CAB2/hCOS16 protein in the transfected breast cancer cell line. Autofluorescence analysis of the transfected SKBR-3 cells of its GFP fusion protein demonstrated that CAB2/hCOS16 translocates to and accumulates in the vesicle, but not in either the cell membrane or the nucleus.

human homologue of the yeast *COS16* gene. To examine the subcellular localization of the CAB2/hCOS16 protein, we expressed the full-size of *CAB2/hCOS16* fused in-frame to green fluorescent protein (GFP) and monitored the cellular localization of the fusion products. Autofluorescence analysis of SKBR-3 cells transfected with a vector encoding the GFP fusion protein demonstrated that CAB2/hCOS16 translocates to and accumulates in vesicles, but not in either the cell membrane or the nucleus (Fig. 3B). This is analogous to the finding that yeast COS16 accumulates in the vacuole,[19] an organelle known to sequester cytosolic Mn^{2+}, and suggests that overexpression of CAB2/hCOS16 may decrease intercellular Mn^{2+} by accumulating it in the vesicles. The yeast CDC1 protein has been reported to be a small subunit of DNA polymerase δ.[21, 22] Cdc1 mutants were sensitive to killing by the DNA-damaging agent methyl methanesulfonate (MMS) and γ-irradiation, and showed a low efficiency for the repair of DNA double-strand breaks.[23] Key molecules, including ataxia-telangiectasia mutated (ATM) kinase family members and DNA-activated protein kinase catalytic polypeptides, for the repair of DNA double-strand breaks have been reported to be Mn^{2+}-dependent enzymes.[24] Therefore, overexpression of *CAB2/hCOS16* may increase the sensitivity of the cells to DNA damage, and cause chromosome instability.

This work was supported in part by a Grant-in-Aid for the 2nd Comprehensive 10-Year Strategy for Cancer Control from the Ministry of Health, Labour and Welfare of Japan, by Grants-in-Aid from the Ministry of Health, Labour and Welfare of Japan, and by a grant for the Promotion of Fundamental Studies in Health Science from the Organization for Drug ADR Relief, R & D Promotion and Product Review of Japan. Nezu, M., Ishizuka, T., and Kuwahara, Y. were awardees of Research Resident Fellowships from the Foundation for Promotion of Cancer Research.

(Received July 24, 2002/Revised September 30, 2002/Accepted October 7, 2002)

REFERENCES

1) Brison, O. Gene amplification and tumor progression. *Biochim. Biophys. Acta*, **1155**, 25–41 (1993).
2) Sakamoto, H., Mori, M., Taira, M., Yoshida, T., Matsukawa, S., Shimizu, K., Sekiguchi, M., Terada, M. and Sugimura, T. Transforming gene from human stomach cancers and a noncancerous portion of stomach mucosa. *Proc. Natl. Acad. Sci. USA*, **83**, 3997–4001 (1986).
3) Yoshida, M., Wada, M., Satoh, H., Yoshida, T., Sakamoto, H., Miyagawa, K., Yokota, J., Koda, T., Kakinuma, M., Sugimura, T. and Terada, M. Human *HST1*(*HSTF1*) gene maps to chromosome band 11q13 and coamplifies with the *INT2* gene in human cancer. *Proc. Natl. Acad. Sci. USA*, **85**, 4861–4864 (1988).
4) Tsuda, T., Tahara, E., Kajiyama, G., Sakamoto, H., Terada, M. and Sugimura, T. High incidence of coamplification of *hst-1* and *int-2* genes in human esophageal carcinomas. *Cancer Res.*, **49**, 5505–5508 (1989).
5) Hamaguchi, M., Sakamoto, H., Tsuruta, H., Sasaki, H.,

Muto, T., Sugimura, T. and Terada, M. Establishment of a highly sensitive and specific exon-trapping system. *Proc. Natl. Acad. Sci. USA*, **89**, 9779–9783 (1992).
6) Akiyama, N., Tsuruta, H., Sasaki, H., Sakamoto, H., Hamaguchi, M., Ohmura, Y., Seto, M., Ueda, R., Hirai, H., Yazaki, Y., Sugimura, T. and Terada, M. Messenger RNA levels of five genes located at chromosome 11q13 in B-cell tumors with chromosome translocation t(11;14)(q13;q32). *Cancer Res.*, **54**, 377–379 (1994).
7) Suire, J. A., Thorner, P. S., Weitzman, S., Maggi, J. D., Dirk, P., Doyle, J., Hale, M. and Godbout, R. Co-amplification of *MYCN* and a DEAD box gene (*DDX1*) in primary neuroblastoma. *Oncogene*, **10**, 1417–1422 (1995).
8) Forus, A., Florenes, V. A., Maelandsmo, G. M., Meltzer, P. S., Fodstad, O. and Myklebost, O. Mapping of amplification units in the q13-14 region of chromosome 12 in human sarcomas: some amplica do not include *MDM2*. *Cell Growth Differ.*, **4**, 1065–1070 (1993).
9) Reifenberger, G., Reifenberger, J., Ichimura, K., Meltzer, P. S. and Collins, V. P. Amplification of multiple genes from chromosomal region 12q13-14 in human malignant gliomas: preliminary mapping of the amplicons shows preferential involvement of *CDK4*, *SAS*, and *MDM2*. *Cancer Res.*, **54**, 4299–4303 (1994).
10) Sen, S., Zhou, H. and White, R. A. A putative serine/threonine kinase encoding gene *BTAK* on chromosome 20q13 is amplified and overexpressed in human breast cancer cell lines. *Oncogene*, **14**, 2195–2200 (1997).
11) Anzick, S. L., Kononen, J., Walker, R. L., Azorsa, D. O., Tanner, M. M., Guan, X.-Y., Sauter, G., Kallionniemi, O.-A., Trent, J. M. and Meltzer, P. S. AIB1, a steroid receptor coactivator amplified in breast and ovarian cancer. *Science*, **277**, 965–968 (1997).
12) Yokota, J., Yamamoto, T., Miyajima, N., Toyoshima, K., Nomura, N., Sakamoto, H., Yoshida, T., Terada, M. and Sugimura, T. Genetic alterations of the c-*erb*B-2 oncogene occur frequently in tubular adenocarcinoma of the stomach and are often accompanied by amplification of the v-*erb*A homologue. *Oncogene*, **2**, 283–287 (1988).
13) Kishi, T., Sasaki, H., Akiyama, N., Ishizuka, T., Sakamoto, H., Aizawa, S., Sugimura, T. and Terada, M. Molecular cloning of human *GRB-7* co-amplified with *CAB1* and c-*ERBB*-2 in primary gastric cancer. *Biochem. Biophys. Res. Commun.*, **232**, 5–9 (1997).
14) Akiyama, N., Sasaki, H., Ishizuka, T., Kishi, T., Sakamoto, H., Onda, M., Hirai, H., Yazaki, Y., Sugimura, T. and Terada, M. Isolation of a candidate gene, *CAB1*, for cholesterol transport to mitochondria from the c-*ERBB*-2 amplicon by a modified cDNA selection method. *Cancer Res.*, **57**, 3548–3553 (1997).
15) Stein, D., Wu, J., Fuqua, S., Roonprapunt, C., Yajnik, V., D'Eustachio, P., Moskow, J. J., Buchberg, A. M., Osborne, C. K. and Margolis, B. 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 (1994).
16) Yokote, Y., Margolis, B., Heldin, C. H. and Claesson-Welsh, L. Grb7 is a downstream signaling component of platelet-derived growth factor α- and β-receptors. *J. Biol. Chem.*, **271**, 30942–30949 (1996).
17) Fiddes, R. J., Champbell, D. H., Janes, P. W., Sivertsen, S. P., Sasaki, H., Wallasch, C. and Daly, R. J. Analysis of Grb7 recruitment by heregulin-activated erbB receptors reveals a novel target selectivity for erbB3. *J. Biol. Chem.*, **273**, 7717–7724 (1998).
18) Nezu, M., Sasaki, H., Kuwahara, Y., Ochiya, T., Yamada, Y., Sakamoto, H., Tashiro, H., Yamazaki, M., Ikeuchi, T., Saito, Y. and Terada, M. Identification of a novel promoter and exons of the c-*ERBB2* gene. *Biochem. Biophys. Res. Commun.*, **258**, 499–505 (1999).
19) Paidhungat, M. and Garrett, S. Cdc1 and the vacuole coordinately regulate Mn^{2+} homeostasis in the yeast *Saccharomyces cerevisiae*. *Genetics*, **148**, 1787 (1998).
20) Kyte, J. and Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.*, **157**, 105 (1982).
21) MacNeill, S. A., Moreno, S., Reynolds, N., Nurse, P. and Fantes, P. A. The fission yeast Cdc1 protein, a homologue of the small subunit of DNA polymerase delta, binds to Pol3 and Cdc27. *EMBO J.*, **15**, 4613–4628 (1996).
22) Zuo, S., Gibbs, E., Kelman, Z., Wang, T. S., O'Donnell, M., MacNeill, S. A. and Hurwitz, J. DNA polymerase delta isolated from *Schizosaccharomyces pombe* contains five subunits. *Proc. Natl. Acad. Sci. USA*, **84**, 11244–11249 (1997).
23) Halbrook, J. and Hoekstra, M. F. Mutations in the *Saccharomyces cerevisiae* CDC1 gene affect double-strand-break-induced intrachromosomal recombination. *Mol. Cell. Biol.*, **14**, 8038 (1994).
24) Kim, S. T., Lim, D. S., Canman, C. E. and Kastan, M. B. Substrate specificities and identification of putative substrates of ATM kinase family members. *J. Biol. Chem.*, **274**, 37538–37543 (1999).