TALEN-mediated single-base-pair editing identification of an intergenic mutation upstream of *BUB1B* as causative of PCS (MVA) syndrome

Hiroshi Ochiai[a,b,1], Tatsuo Miyamoto[a,1], Akinori Kanai[c], Kosuke Hosoba[a], Tetsushi Sakuma[b], Yoshiki Kudo[d], Keiko Asami[e], Atsushi Ogawa[e], Akihiro Watanabe[e], Tadashi Kajii[f], Takashi Yamamoto[b], and Shinya Matsuura[a,2]

[a]Department of Genetics and Cell Biology, Research Institute for Radiation Biology and Medicine, Hiroshima University, Hiroshima 734-8553, Japan; [b]Department of Mathematical and Life Sciences, Graduate School of Science, Hiroshima University, Higashi-Hiroshima 739-8526, Japan; [c]Department of Molecular Oncology, Research Institute for Radiation Biology and Medicine, Hiroshima University, Hiroshima 734-8553, Japan; [d]Department of Obstetrics and Gynecology, Graduate School of Biomedical Sciences, Hiroshima University, Hiroshima 734-8551, Japan; [e]Department of Pediatrics, Niigata Cancer Center Hospital, Niigata 951-8566, Japan; and [f]Hachioji, Tokyo 192-0023, Japan

Edited by Albert de la Chapelle, Ohio State University Comprehensive Cancer Center, Columbus, OH, and approved November 20, 2013 (received for review September 11, 2013)

Cancer-prone syndrome of premature chromatid separation with mosaic variegated aneuploidy [PCS (MVA) syndrome] is a rare autosomal recessive disorder characterized by constitutional aneuploidy and a high risk of childhood cancer. We previously reported monoallelic mutations in the *BUB1B* gene (encoding BUBR1) in seven Japanese families with the syndrome. No second mutation was found in the opposite allele of any of the families studied, although a conserved *BUB1B* haplotype and a decreased transcript were identified. To clarify the molecular pathology of the second allele, we extended our mutational search to a candidate region surrounding *BUB1B*. A unique single nucleotide substitution, G > A at ss802470619, was identified in an intergenic region 44 kb upstream of a *BUB1B* transcription start site, which cosegregated with the disorder. To examine whether this is the causal mutation, we designed a transcription activator-like effector nuclease–mediated two-step single-base pair editing strategy and biallelically introduced this substitution into cultured human cells. The cell clones showed reduced *BUB1B* transcripts, increased PCS frequency, and MVA, which are the hallmarks of the syndrome. We also encountered a case of a Japanese infant with PCS (MVA) syndrome carrying a homozygous single nucleotide substitution at ss802470619. These results suggested that the nucleotide substitution identified was the causal mutation of PCS (MVA) syndrome.

spindle assembly checkpoint | genome editing

Budding uninhibited by benzimidazole-related 1 (BUBR1) is an important protein of the spindle assembly checkpoint. Constitutional mutations in the *BUB1B* (budding uninhibited by benzimidazoles 1 homolog beta) gene encoding BUBR1 cause the rare human disorder premature chromatid separation (PCS) syndrome [Mendelian Inheritance in Man (MIM) 176430], also known as mosaic variegated aneuploidy (MVA) syndrome (MIM 257300). PCS (MVA) syndrome is characterized by PCS in >50% metaphase cells and a variety of mosaic aneuploidies (1, 2). Patient clinical findings include Dandy–Walker complex, postcerebellar cysts, hypoplasia of the cerebellar vermis, lissencephaly, cataracts, uncontrollable clonic seizures, polycystic kidneys, infantile obesity, and a high risk of malignancy including Wilms' tumor and rhabdomyosarcoma (3).

Both biallelic and monoallelic mutations of *BUB1B* have been identified in individuals with the syndrome (1, 2). Biallelic mutations were previously found in five of eight families (1), each of which had one null mutation in the first allele and another missense mutation in the second (opposite) allele. The null mutations result in a 50% reduction of BUBR1 function, whereas the missense mutations partially disrupt BUBR1 protein functions. It was therefore deduced that a >50% reduction of BUBR1 function is involved in the syndrome. We previously reported monoallelic *BUB1B* mutations in seven Japanese families (2), all of which had one null mutation in the first allele but no second mutation was found in the opposite allele despite the decrease in *BUB1B* transcripts and a conserved *BUB1B* haplotype. The molecular basis of the second alleles was therefore unknown. In this study, we searched for the mutation in the second allele and identified a unique SNP [ss802470619 (G/A)] in an intergenic region 44 kb upstream of *BUB1B* as a candidate mutation.

To prove that this is the disease-causing mutation, we used transcription activator-like effector nuclease (TALEN)-mediated single-base-pair editing to establish biallelically SNP-modified disease model cells for functional cytological assays. A TALEN consists of a customizable DNA binding domain and a DNA cleavage domain and offers the advantage of simple and convenient design and construction compared with other engineered endonucleases (EENs) such as zinc-finger nuclease (ZFN). TALENs can introduce DNA double-stranded breaks (DSBs) into a specific genomic site in pairs and induce the DNA damage response to mend such breaks (4–7). In mammalian cells, DSBs are mainly repaired by nonhomologous end-joining in which the two ends are processed and ligated together in a way that is frequently accompanied by nucleotide insertions and deletions (6–8). The

Significance

A single nucleotide substitution in an intergenic region upstream of *BUB1B* (encoding BUBR1) was identified as a candidate mutation for premature chromatid separation with mosaic variegated aneuploidy [PCS (MVA) syndrome], a cancer-prone genetic disorder. To prove that this is the causal mutation, we designed a unique genome editing strategy, transcription activator-like effector nuclease–mediated two-step single-base-pair editing, to biallelically introduce this substitution into cultured human cells. The cell clones showed chromosomal instability in the form of PCS and MVA, which are cellular hallmarks of the syndrome, suggesting that this is indeed the underlying mutation. This single-base-pair editing technique will be useful for investigations of noncoding variants of unknown functional relevance.

Author contributions: H.O., T.M., and S.M. designed research; H.O., T.M., A.K., K.H., T.S., and S.M. performed research; Y.K., K.A., A.O., A.W., and T.K. provided clinical information/analytic tools; and H.O., T.M., T.Y., and S.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

See Commentary on page 1233.

[1]H.O. and T.M. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: shinya@hiroshima-u.ac.jp.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1317008111/-/DCSupplemental.

cointroduction of EENs with a targeting vector and subsequent antibiotic selection previously enabled efficient homologous recombination-mediated genome editing to be performed in mammalian cells (9–11). Moreover, the introduction of a specific nucleotide substitution was demonstrated using ZFNs and either Cre/*loxP* or *piggyBac* technology (9, 11), as well as EENs and either plasmid donor vectors (11, 12) or single-stranded oligodeoxynucleotides (ssODNs) (11, 13–16) without antibiotic selection. In this study, we describe the successful use of a selection-based TALEN-mediated two-step single-base pair editing strategy in establishing model cells of PCS (MVA) syndrome.

Results

PCS (MVA) Syndrome Infant with No *BUB1B* Exonic Mutations. We studied a Japanese infant (KH01) with PCS (MVA) syndrome who developed rhabdomyosarcoma of the urinary bladder, but no microcephaly or Dandy–Walker anomalies, and had PCS in 67% of lymphocytes and mosaic aneuploidy in 17% (Fig. 1*A*). Western blot analysis revealed significant reduction of BUBR1 expression compared with normal individuals (Fig. 1*B*). However, no *BUB1B* mutations were found in coding regions, intron-exon boundaries, promoter regions, and untranslated regions.

In our previous study of monoallelic *BUB1B* mutations in seven Japanese families with PCS (MVA) syndrome, the ancestral haplotype, composed of microsatellite marker 26020GT, SNP 1046G/A (rs1801376) in exon 8 of *BUB1B*, and microsatellite marker D15S994, was 6G3 in five families and 6G8 in one family (Fig. 1 *C* and *D*) (2). Here, haplotype analysis of the infant (KH01) revealed a homozygous 6G3 haplotype, each allele being derived from the parents (Fig. 1*D*). The expression level of BUBR1 in the infant was higher than in a patient with PCS (MVA) syndrome (MY1) who is a compound heterozygote for a null mutation plus 6G3 haplotype and lower than in a heterozygous carrier with a null mutation (1833delT) (Fig. 1*B*). These results indicate that the infant is a homozygote for the second hypomorphic allele common to Japanese families.

Identification of a Single Nucleotide Substitution in an Intergenic Region Upstream of *BUB1B* by Deep Sequencing. We speculated that an unidentified mutation could be embedded around the *BUB1B* locus of the second allele. To search for this, we analyzed the haplotypes of a 1-Mb region in Japanese families and ascertained the common ancestral haplotype within the 200-kb interval from microsatellite marker 1790GT to 901722 (Fig. 1 *C* and *D*). We therefore determined the nucleotide sequences of the 200-kb region in a patient with PCS (MVA) syndrome (PCS1, family 1) who had a truncating *BUB1B* mutation plus 6G3 haplotype. A total of 186 known and 5 unknown SNPs were identified in this region. The frequencies of all known SNPs were >1% in the general population and were excluded as candidates for the mutation. Of the five unique SNPs, four did not cosegregate with the disorder so were excluded as candidates. Only SNP, ss802470619 (G/A), located 44 kb upstream of a *BUB1B* transcription start site, cosegregated with the disease (Fig. 1 *C* and *D*). The second hypomorphic alleles in Japanese families with PCS (MVA) syndrome all carried the nucleotide substitution G > A at ss802470619, and the infant (KH01) was homozygous for the nucleotide substitution (Fig. 1 *D* and *E*). The frequency of this minor allele in 398 Japanese individuals was 0.0025 (2/796). This substitution was therefore a strong candidate mutation for PCS (MVA) syndrome.

Introduction of Biallelic Single Nucleotide Substitutions into Cultured Human Cells Using TALEN-Mediated Two-Step Single-Base-Pair Editing. If the nucleotide substitution at ss802470619 was indeed the disease-causing mutation, its introduction into normal cultured human cells would result in a decrease of *BUB1B* transcripts and an increase of PCS frequency and MVA. To examine this possibility, we designed a TALEN-mediated two-step single-base-pair editing strategy (Fig. 2*A* and Fig. S1). The first step included TALEN-mediated targeted integration of a selection cassette into the SNP region. The targeting vector contained a puromycin-resistant gene and a herpes simplex virus

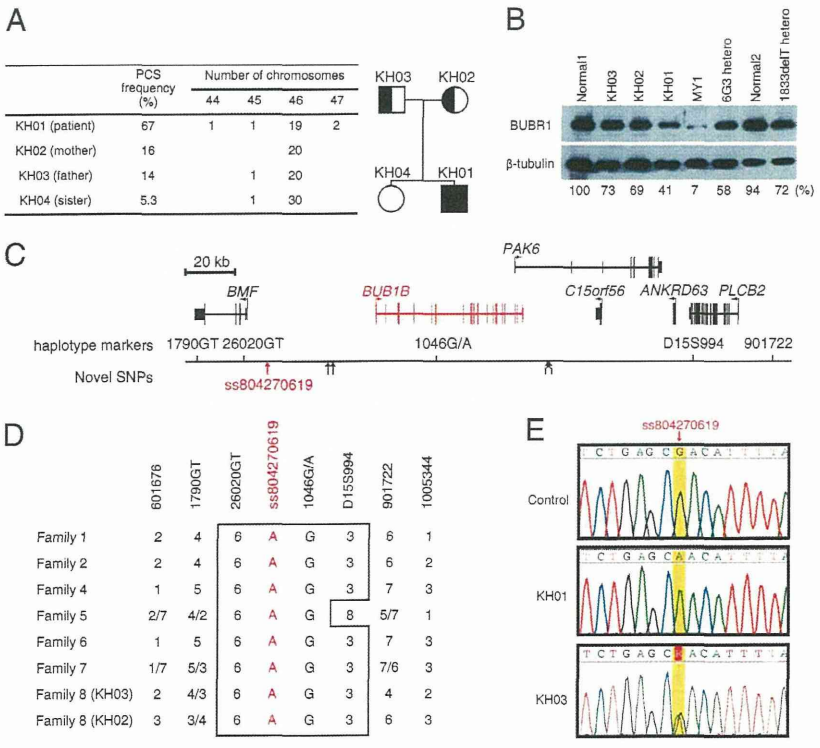

Fig. 1. An infant with PCS (MVA) syndrome homozygous for an intergenic single nucleotide substitution associated with the disease. (*A*) The frequencies of PCS and distribution of chromosome numbers in the infant (KH01) and family members (KH02–04). (*Right*) Family pedigree. (*B*) Western blot analysis of Epstein–Barr virus-transformed lymphoblastoid cell lines showing reduction of BUBR1 expression in the infant (KH01). MY1 is a patient with PCS (MVA) syndrome who is a compound heterozygote for a truncating *BUB1B* mutation (IVS10-5A > G) plus 6G3 haplotype. 6G3 hetero and 1833delT hetero represent heterozygous carriers with the 6G3 haplotype and the truncating *BUB1B* mutation 1833delT, respectively. The β-tubulin antibody was used as a loading control. Densitometric levels of BUBR1 bands normalized to those of β-tubulin bands are shown beneath each lane. (*C*) Schematic of genomic structure of *BUB1B* locus and haplotype markers used in this study. Upward pointing arrows represent unique single nucleotide variations found in this study. One of these, ss804270619, associated with PCS (MVA) syndrome, is highlighted in red. (*D*) Haplotypes of the second allele in Japanese families with PCS (MVA) syndrome. Families 1–7 were described previously (2), and family 8 is described in this study. (*E*) Sequence analysis of ss804270619 in a normal individual, the infant (KH01), and the father (KH03).

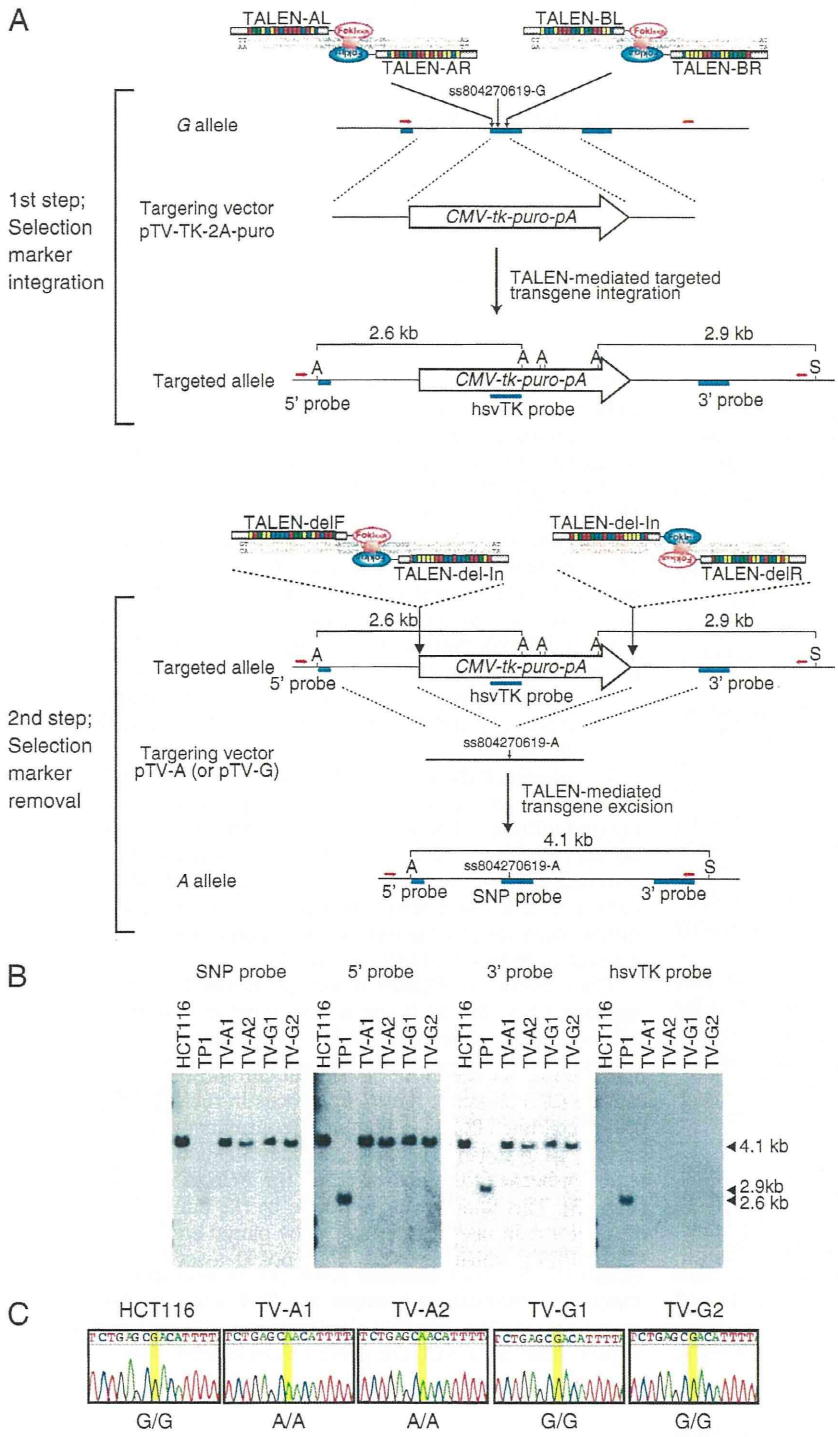

Fig. 2. Biallelic introduction of single nucleotide substitutions into an intergenic region located 44 kb upstream of *BUB1B* in cultured human cells. (*A*) The strategy for biallelic introduction of single nucleotide substitutions into cultured human cells. Red arrows indicate sites of primers used in PCR genotyping. Blue bars indicate Southern blot probes. A, ApaI; S, StuI. (*B*) Southern blot analyses showing TALEN-mediated biallelic insertion and biallelic excision of the selection cassettes in modified cell clones. (*C*) Sequence analyses revealed introduction of the single nucleotide substitutions at the SNP site.

thymidine kinase (hsvTK) gene separated by a 2A peptide sequence, allowing expression of the discrete protein products from a single ORF (17). TALEN expression vectors and the targeting vector were cotransfected into cultured human cells, and puromycin-resistant clones were positively selected. The second step involved targeted excision of the selection cassette and introduction of the single nucleotide substitution. The TALEN expression vectors encoding TALENs, which cut each end of the selection cassette in the genome, and the secondary targeting vector containing the single nucleotide substitution were transfected into the cells, and the single nucleotide-edited clones were negatively selected using ganciclovir treatment.

For gene targeting experiments, we used the human colon cancer cell line HCT116 because it has two copies of a WT *BUB1B* allele and shows a normal response at the colcemid-induced mitotic spindle checkpoint (2). We transfected both the TALEN expression vectors and the targeting vector (pTV-TK-2A-puro) into HCT116 cells. After puromycin selection, 96 colonies were picked up, and their genotypes were analyzed

Table 1. Summary of TALEN-mediated targeted cassette integration

Cell line	TALEN	Donor vector	Clones analyzed	Number of clones with monoallelically targeted locus	Number of clones with biallelically targeted locus		Cell line designation
					Without random integration	With random integration	
HCT116	TALEN-AL TALEN-AR TALEN-BL TALEN-BR	pTV-TK-2A-puro	96	75	3*	8	TP1 TP2 TP3

*These were designated as TP1, TP2, and TP3.

by PCR (Table 1). Of these, 75 clones were shown to be monoallelically and 11 to be biallelically targeted (Table 1). Southern blot analysis revealed that, of the 11 clones, three had no random integration of the vector in the genome. Therefore, we used the biallelically targeted clone TP1 for selection cassette removal.

Expression vectors encoding TALENs (TALEN-delF, -delR, and -delIn), which target both ends of the selection cassette (Fig. 2*A* and Fig. S1), and secondary targeting vectors pTV-G and pTV-A, each consisting of a 2-kb genomic region including the SNP (ss802470619) G and A, respectively, were generated (Fig. 2*A* and Fig. S1). We transfected the TALEN expression vectors and either pTV-G or pTV-A into TP1 cells. After negative selection using ganciclovir treatment, 32 colonies each from pTV-G– and pTV-A–transfected plates were picked up, and PCR genotyping and sequence analysis were carried out. In two pTV-G (TV-G1 and TV-G2) and two pTV-A (TV-A1 and TV-A2) clones, the selection cassettes were biallelically excised, as expected (Table 2). Southern blot analysis revealed no random integration of the vector in these clones (Fig. 2*B*). Sequence analysis demonstrated that the single nucleotide substitutions were correctly introduced in pTV-A cell clones (Fig. 2*C*).

EENs including ZFNs and TALENs can induce DSB at sites other than their intended targets (18–20), so we searched for potential TALEN off-target sites in the human genome using the TALE-NT web tool (21) (Table S1). We analyzed the top two most-likely off-target sites, but observed no sequence alterations in the pTV-A and -G cell clones (Table S1). We therefore used these four clones for functional analysis.

Functional Cytological Analysis of the Single Nucleotide–Modified Cell Clones. Expression levels of BUBR1 protein and *BUB1B* mRNA in pTV-G and pTV-A clones were analyzed. Western blot analysis showed reduced BUBR1 protein levels in both pTV-A clones compared with the parental cell line HCT116 and pTV-G clones (Fig. 3*A*). Consistent with this, reduced expression of *BUB1B* mRNA was detected by quantitative RT-PCR (qRT-PCR) analysis in pTV-A clones compared with HCT116 cells and pTV-G clones (Fig. 3*B*).

To compare the effects of BUBR1 reduction in the SNP-modified cells, we generated constitutive *BUB1B* knockdown cells using *BUB1B* shRNA from HCT116 cells (Fig. S2). Parental HCT116 cells showed relatively stable karyotypes (86% of cells have 45 chromosomes) as reported previously (Fig. 3*C*) (22), and 1% of metaphase was in PCS (Fig. 3*E*). By contrast, constitutive *BUB1B* knockdown cells, TV-A1, and TV-A2 cells showed 62%, 48%, and 44% aneuploidy, respectively (Fig. 3*C*), and 53.5%, 33%, and 31.5% metaphases, respectively, in PCS (Fig. 3 *D* and *E*), which are hallmarks of PCS (MVA) syndrome. TV-G1 and TV-G2 cells showed no apparent increase of aneuploidy (Fig. 3*C*) and only 2% and 2.5% metaphases in PCS, respectively (Fig. 3 *D* and *E*).

We examined the effect of colcemid treatment on the mitotic index of the cell clones. Control HCT116 cells showed an increased mitotic index, whereas *BUB1B* knockdown cells showed no appreciable increase. Both pTV-A cell lines showed an intermediate response between control and *BUB1B* knockdown cells (Fig. 3*F*). These results indicated that the pTV-A cells were mildly insensitive to the mitotic checkpoint, which is likely to be a result of reduced BUBR1 expression.

The reduction of *BUB1B* transcripts in the SNP-modified cells suggested that the SNP site was situated in a distal regulatory element, thus affecting *BUB1B* expression by a physical interaction with the promoter region and the formation of chromatin loops. We therefore carried out chromosome conformation capture (3C) assays at the *BUB1B* locus in the PCS (MVA) syndrome cell line (PCS1), a human osteosarcoma cell line (U2OS), and HCT116 cells (23). These cell lines all showed a physical interaction between the SNP site and the *BUB1B* promoter region (Fig. S3). This finding was supported by the public genomic data of a chromatin interaction analysis by paired-end tag sequencing (ChIA-PET), which revealed RNA pol II–associated long-range chromatin interactions between the SNP site and *BUB1B* promoter region (Fig. S4) (24). Furthermore, DNase hypersensitive regions from the ENCODE project were located close to the SNP

Table 2. Summary of TALEN-mediated targeted cassette excision

Cell line	TALEN	Donor vector	Number of clones analyzed	Number of clones with no excision	Number of clones with incorrectly excised	Number of clones with correctly excised		Cell line designation
						Without random integration	With random integration	
TP1	TALEN-delF TALEN-delR	pTV-G	32	19	10	2*	1	TV-G1 TV-G2
	TALEN-delIn	pTV-A	32	15	15	2*	0	TV-A1 TV-A2

*These clones were used for functional analysis.

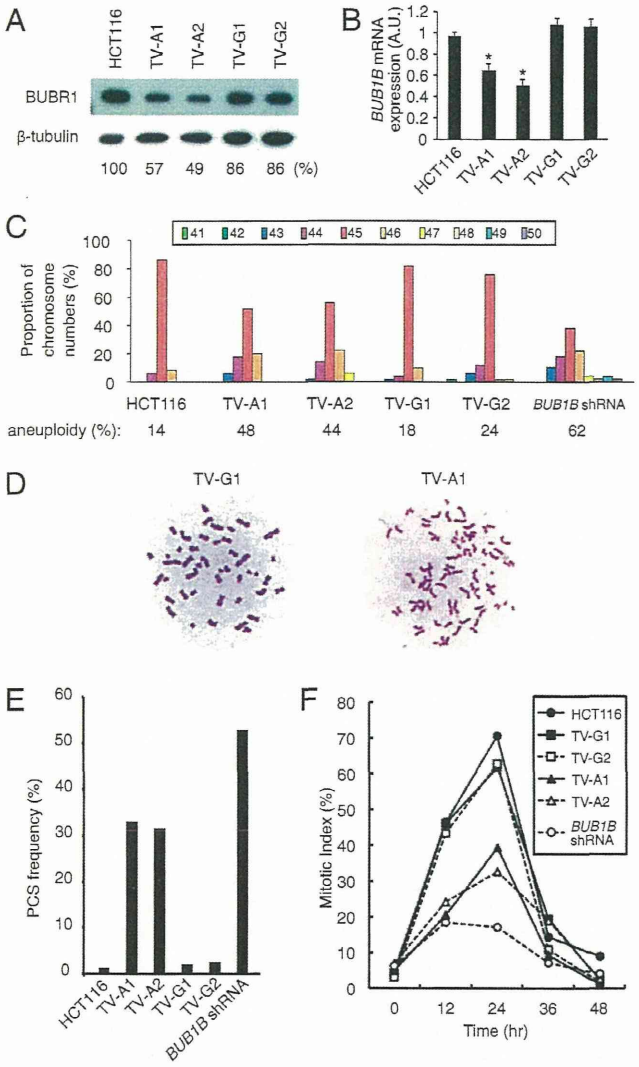

Fig. 3. The nucleotide substitution at ss804270619 affects *BUB1B* expression. (*A*) Western blot analysis showing reduction of BUBR1 in the cell clones with biallelically introduced substitutions. The β-tubulin antibody was used as a loading control. Densitometric levels of BUBR1 bands normalized to those of β-tubulin bands are shown beneath each lane. (*B*) Quantitative RT-PCR analysis showing reduction of *BUB1B* transcripts in the cell clones with biallelically introduced substitutions. Each gene expression analysis was normalized to *β-tubulin*. $n = 4$. Bars are ±SD. Asterisks indicate significant differences between pTV-A (TV-A1 and -A2) and other clones (HCT116, TV-G1 and -G2) ($P < 0.005$, two-tailed t test). (*C*) Metaphase chromosome analysis showing proportion of chromosome numbers in the cell clones. (*D*) Example of premature chromatid separation (PCS) in TV-A1 compared with intact chromatid cohesion in TV-G1. (*E*) PCS frequency of each cell clone showing that TV-A1 and TV-A2 clones exhibited high rates of PCS. (*F*) Mitotic index of the cell clones after treatment with colcemid indicating that those with biallelic substitutions at ss804270619 are mildly insensitive to the spindle assembly checkpoint.

site in several cell types (Fig. S4) (25). These results suggested that the region containing the SNP site is a distal regulatory element that affects *BUB1B* expression.

Discussion

We identified a single nucleotide substitution (G > A, ss802470619) in an intergenic region 44-kb upstream of *BUB1B* as a unique causative mutation of PCS (MVA) syndrome. Both biallelic and monoallelic *BUB1B* mutations have previously been found in individuals with the syndrome (1, 2). Biallelic mutations always consist of one truncating and one missense mutation. Until now, monoallelic mutations have been shown to be composed of one truncating mutation and a second mutation of unknown molecular basis. The present study has now identified this as an intergenic mutation. Both the missense mutations and the intergenic mutation partially disrupt the functions of the BUBR1 protein, whereas the allele with the protein truncating mutation is effectively null.

Bub1b KO mice were previously found to be embryonic lethal (26), suggesting that it is necessary to have one hypomorphic allele that allows for minimal production of BUBR1 protein in individuals with the syndrome. On the other hand, it was not known whether an individual with biallelic hypomorphic mutations would show the PCS (MVA) syndrome or a normal clinical phenotype. The infant studied here had biallelic hypomorphic mutations and indeed showed PCS (MVA) syndrome clinical phenotypes. Although reduced BUBR1 expression by the infant was demonstrated by Western blot analysis, expression levels were much higher than those of a patient with a truncating plus intergenic mutation but lower than those of a carrier with a truncating mutation of *BUB1B*. The infant had PCS in 67% of lymphocytes, which was lower than in the patient with a truncating plus intergenic mutation and higher than in the parents with heterozygous mutations (2). These results suggested that the biallelic hypomorphic mutations are causative of PCS (MVA) syndrome.

The single nucleotide substitution identified was situated in the intergenic region. This nucleotide was not strictly conserved among vertebrates and no obvious potential regulatory elements were predicted. Because it was difficult to evaluate this substitution for its effects on *BUB1B* expression by functional genomics studies alone, we showed experimentally that it is involved in the down-regulation of *BUB1B* expression using the TALEN-mediated two-step single-base-pair editing method. In cells with biallelic hypomorphic mutations, there was no change to the size of the *BUB1B* product, but expression levels were reduced compared with control cells. Moreover, the 3C assay showed that the region containing the SNP site physically interacted with the *BUB1B* promoter region in all cell lines studied. In addition, ENCODE data revealed that chromatin regions accessible by DNase hypersensitive analysis were located close to the SNP site in several cell types. These results suggest that the region containing the intergenic mutation acts as a *cis*-regulatory element for *BUB1B* expression, such that the nucleotide substitution impairs proper regulation of *BUB1B* expression. However, DNase hypersensitivity, a universal feature of active *cis*-regulatory sequences, was not apparent at the SNP site in HCT116 cells despite the reduced *BUB1B* expression in SNP-modified cells (Fig. S4). Further investigations will therefore be needed to understand the molecular mechanisms of how the intergenic mutation reduces *BUB1B* expression.

Successful introductions of single nucleotide substitutions have been reported using ZFNs and a targeting vector, followed by antibiotic selection and excision of the selection marker by either Cre/*loxP* or *piggyBac* technology (9, 11). However, in these cases, unwanted footprints were introduced besides the targeted sites. This makes it difficult to precisely evaluate the effects of single nucleotide substitutions on cellular functions. We therefore designed a TALEN-mediated two-step single-base pair editing strategy, which, although labor intensive because it required multiple sets of TALENs, two separate targeting constructs, and two rounds of gene targeting, was nevertheless able to introduce single nucleotide substitutions biallelically into the genome. Single-step gene correction has also been documented using EENs and either plasmid donor vectors (11, 12) or ssODNs (11, 13–16). Moreover, the clustered regularly interspaced short palindromic repeat (CRISPR)/Cas based RNA-guided DNA

endonucleases have emerged as a potentially facile and efficient alternative to ZFNs and TALENs (27, 28), with enhanced genome editing specificity achieved by a Cas9 nickase with paired guide RNAs that reduce off-target activity (29, 30). These editing schemes should facilitate the efficient introduction of single nucleotide substitutions into a specific genomic site.

A genomewide association study suggested that the greater part of disease- and trait-associated nucleotide variants is located within noncoding sequence, which complicates their functional evaluation (31). To overcome this problem, functional genomics studies could be used, including ChIP of transcription factor binding sites, chromatin regulators, and histone modification marks or DNase hypersensitivity analysis of accessible chromatin regions. When the nucleotide variants of interest have not been predicted, it can be difficult to define whether it is a causal mutation or merely a neutral polymorphism as reported here. Single-base-pair editing will therefore be useful to study the function of such nucleotide variants with unknown functional relevance.

Materials and Methods

Patient (KH01), a boy, was born at 39 wk of gestation to healthy, nonconsanguineous parents of Japanese origin: a 28-y-old mother (KH02) and a 27-y-old father (KH03) (family 8). The older sister (KH04) was normal. The birth weight of the patient was 2,775 g (−1.1 SD); length was 48.2 cm (−0.9 SD); and occipito-frontal head circumference (OFC) was 32.2 cm (−0.9 SD). He was referred to us at the age of 2 mo with embryonal rhabdomyosarcoma of the urinary bladder. He weighed 6,100 g (0.0 SD); length was 59.7 cm (−0.2 SD); and OFC was 36.7 cm (−1.8 SD). MRI of the brain showed no internal malformations. At the age of 4 mo, he developed uncontrollable clonic seizures, which were generalized after chemotherapy with vincristine. Chromosome analysis of peripheral blood lymphocytes revealed a 46, XY, t(14,18)(q11.1; p11.1) karyotype with PCS in 67% of cells and mosaic aneuploidy in 17%. Parental karyotypes were normal, with no increase in aneuploidy, but PCS in 14% of the father's and 16% of the mother's cells. At 8 mo of age, he underwent a urinary bladder tumorectomy. Now at the age of 8 y, he is in complete remission. He does not have cataracts or polycystic kidneys but still has uncontrollable clonic seizures and does not speak meaningful words.

The parents expected their third child, and amniocentesis was performed at the 16th week of pregnancy. Chromosome examination of the amniocyte cultures showed a normal karyotype and no increase of aneuploid cells. The haplotype of amniocyte DNA was heterozygous 6G3, which was derived from the mother. The fetus was thus determined to be a heterozygous carrier, and a normal male infant (KH05) was delivered after 40 wk of gestation.

Full descriptions of the haplotype analysis, cell culture, long-range PCR, and next-generation sequencing, plasmids, Western blot analysis, TALEN-mediated targeted gene integration, PCR genotyping, Southern blot analysis, TALEN-mediated targeted gene excision, qRT-PCR, chromosome analysis, mitotic checkpoint analysis, and 3C assay are detailed in *SI Materials and Methods*. The nucleotide sequences of pTV-TK-2A-puro are shown in Fig. S5. The sequences of oligonucleotides used in this study are listed in Table S2. The target nucleotide sequences and amino acid sequences of TALENs used in this study are shown in Table S3 and Fig. S6.

ACKNOWLEDGMENTS. We thank Drs. Tatsuro Ikeuchi and Yoshiyuki Matsumoto for analysis of patient samples and Drs. Takayuki Takachi, Jun Tohyama, and Masayuki Kubota for providing clinical information. We also thank Dr. Didier Trono for supplying the pLOX-TERT-iresTK vector (Addgene plasmid 12245); Dr. Feng Zhang for providing the pTALEN_v2 (NG) vector (Addgene plasmid 32190); Dr. Sumihare Noji for valuable comments; Dr. Ken-ichi Suzuki and Dr. Miyako Nakano for discussions; and Ms. Y. Tonouchi and T. Amimoto for technical assistance. This work was supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology (to S.M. and T.M.); a Grant-in-Aid for Scientific Research from the Ministry of Health, Labour and Welfare (to S.M.); a Grant-in-Aid for Japan Society for the Promotion of Science Fellows (to H.O.); research grants from the Naito Foundation (to S.M.) and from Tsuchiya Medical Foundation (to T.M.) and the Sasakawa Scientific Research Grant from The Japan Science Society (to H.O.).

1. Hanks S, et al. (2004) Constitutional aneuploidy and cancer predisposition caused by biallelic mutations in BUB1B. *Nat Genet* 36(11):1159–1161.
2. Matsuura S, et al. (2006) Monoallelic BUB1B mutations and defective mitotic-spindle checkpoint in seven families with premature chromatid separation (PCS) syndrome. *Am J Med Genet A* 140(4):358–367.
3. García-Castillo H, Vásquez-Velásquez AI, Rivera H, Barros-Núñez P (2008) Clinical and genetic heterogeneity in patients with mosaic variegated aneuploidy: Delineation of clinical subtypes. *Am J Med Genet A* 146A(13):1687–1695.
4. Wood AJ, et al. (2011) Targeted genome editing across species using ZFNs and TALENs. *Science* 333(6040):307.
5. Sander JD, et al. (2011) Targeted gene disruption in somatic zebrafish cells using engineered TALENs. *Nat Biotechnol* 29(8):697–698.
6. Tesson L, et al. (2011) Knockout rats generated by embryo microinjection of TALENs. *Nat Biotechnol* 29(8):695–696.
7. Miller JC, et al. (2011) A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol* 29(2):143–148.
8. Kim Y, et al. (2013) A library of TAL effector nucleases spanning the human genome. *Nat Biotechnol* 31(3):251–258.
9. Yusa K, et al. (2011) Targeted gene correction of α1-antitrypsin deficiency in induced pluripotent stem cells. *Nature* 478(7369):391–394.
10. Hockemeyer D, et al. (2011) Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol* 29(8):731–734.
11. Soldner F, et al. (2011) Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. *Cell* 146(2):318–331.
12. Urnov FD, et al. (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* 435(7042):646–651.
13. Ding Q, et al. (2013) A TALEN genome-editing system for generating human stem cell-based disease models. *Cell Stem Cell* 12(2):238–251.
14. Wefers B, et al. (2013) Direct production of mouse disease models by embryo microinjection of TALENs and oligodeoxynucleotides. *Proc Natl Acad Sci USA* 110(10):3782–3787.
15. Meyer M, Ortiz O, Hrabé de Angelis M, Wurst W, Kühn R (2012) Modeling disease mutations by gene targeting in one-cell mouse embryos. *Proc Natl Acad Sci USA* 109(24):9354–9359.
16. Chen F, et al. (2011) High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. *Nat Methods* 8(9):753–755.
17. Szymczak AL, et al. (2004) Correction of multi-gene deficiency in vivo using a single 'self-cleaving' 2A peptide-based retroviral vector. *Nat Biotechnol* 22(5):589–594.
18. Gupta A, Meng X, Zhu LJ, Lawson ND, Wolfe SA (2011) Zinc finger protein-dependent and -independent contributions to the in vivo off-target activity of zinc finger nucleases. *Nucleic Acids Res* 39(1):381–392.
19. Gabriel R, et al. (2011) An unbiased genome-wide analysis of zinc-finger nuclease specificity. *Nat Biotechnol* 29(9):816–823.
20. Pattanayak V, Ramirez CL, Joung JK, Liu DR (2011) Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nat Methods* 8(9):765–770.
21. Doyle EL, et al. (2012) TAL Effector-Nucleotide Targeter (TALE-NT) 2.0: Tools for TAL effector design and target prediction. *Nucleic Acids Res* 40(Web Server issue):W117-22.
22. Kienitz A, Vogel C, Morales I, Müller R, Bastians H (2005) Partial downregulation of MAD1 causes spindle checkpoint inactivation and aneuploidy, but does not confer resistance towards taxol. *Oncogene* 24(26):4301–4310.
23. Hagège H, et al. (2007) Quantitative analysis of chromosome conformation capture assays (3C-qPCR). *Nat Protoc* 2(7):1722–1733.
24. Li G, et al. (2012) Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* 148(1-2):84–98.
25. Birney E, et al.; ENCODE Project Consortium; NISC Comparative Sequencing Program; Baylor College of Medicine Human Genome Sequencing Center; Washington University Genome Sequencing Center; Broad Institute; Children's Hospital Oakland Research Institute (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 447(7146):799–816.
26. Wang Q, et al. (2004) BUBR1 deficiency results in abnormal megakaryopoiesis. *Blood* 103(4):1278–1285.
27. Cong L, et al. (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339(6121):819–823.
28. Mali P, et al. (2013) RNA-guided human genome engineering via Cas9. *Science* 339(6121):823–826.
29. Mali P, et al. (2013) CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol* 31(9):833–838.
30. Ran FA, et al. (2013) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154(6):1380–1389.
31. Maurano MT, et al. (2012) Systematic localization of common disease-associated variation in regulatory DNA. *Science* 337(6099):1190–1195.

Exp. Anim. 63(1), 79–84, 2014

—Original—

Screening Methods to Identify TALEN-Mediated Knockout Mice

Yoshiko NAKAGAWA[1)], Takashi YAMAMOTO[2)], Ken-Ichi SUZUKI[2)], Kimi ARAKI[1)], Naoki TAKEDA[1)], Masaki OHMURAYA[1)], and Tetsushi SAKUMA[2)]

[1)]*Center for Animal Resources and Development, Kumamoto University, 2–2–1 Honjo, Kumamoto 860-0811, Japan*
[2)]*Department of Mathematical and Life Sciences, Graduate School of Science, Hiroshima University, 1–3–1 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8526, Japan*

Abstract: Genome editing with site-specific nucleases, such as zinc-finger nucleases or transcription activator-like effector nucleases (TALENs), and RNA-guided nucleases, such as the CRISPR/Cas (clustered regularly interspaced short palindromic repeats/CRISPR-associated) system, is becoming the new standard for targeted genome modification in various organisms. Application of these techniques to the manufacture of knockout mice would be greatly aided by simple and easy methods for genotyping of mutant and wild-type pups among litters. However, there are no detailed or comparative reports concerning the identification of mutant mice generated using genome editing technologies. Here, we genotyped TALEN-derived enhanced green fluorescent protein (*eGFP*) knockout mice using a combination of approaches, including fluorescence observation, heteroduplex mobility assay, restriction fragment length polymorphism analysis and DNA sequencing. The detection sensitivities for TALEN-induced mutations differed among these methods, and we therefore concluded that combinatorial testing is necessary for the screening and determination of mutant genotypes. Since the analytical methods tested can be carried out without specialized equipment, costly reagents and/or sophisticated protocols, our report should be of interest to a broad range of researchers who are considering the application of genome editing technologies in various organisms.
Key words: genome editing, knockout mouse, TALEN, targeted mutagenesis

Introduction

Transcription activator-like effector (TALE) nuclease (TALEN)-mediated gene knockout technology is now applicable to a wide variety of cells and organisms [5]. Each TALEN comprises a TALE domain that binds to a specified DNA sequence and a nuclease domain derived from the *Fok*I restriction endonuclease. When a pair of TALENs designed for a specific genomic locus is introduced into embryos, a DNA double-strand break (DSB) occurs at the target site. DSBs are mainly repaired by error-prone non-homologous end-joining (NHEJ), resulting in randomly induced insertions and deletions that cause disruption of gene functions [7].

Conventionally, knockout mice have been created using an embryonic stem (ES) cell-mediated strategy based on spontaneous homologous recombination between genomic DNA and a targeting construct [2]. This method is time-consuming and requires several laborious processes, such as construction of a gene targeting vector, isolation of targeted ES cell clones, production of chimeras, test breeding for germline transmission and, in some cases, backcrossing to another inbred background. However, the use of TALENs for gene targeting

(Received 13 June 2013 / Accepted 12 August 2013)
Address corresponding: T. Sakuma, Department of Mathematical and Life Sciences, Graduate School of Science, Hiroshima University, 1–3–1 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8526, Japan

enables knockout mice to be produced in a short time because the TALEN mRNAs are simply injected into the embryos of any intended inbred strain. Recently, several groups reported that knockout mice and rats could be efficiently created by TALEN-mediated gene targeting [4, 6, 10, 14, 16]. The convenience of this technique means that TALEN-mediated gene knockout will become a major method for the production of genetically modified rodents in the near future. However, while ES cell-mediated chimeric mice can easily be determined by their coat color, with TALEN-mediated gene targeting it is difficult to distinguish genetically modified mice from wild-type pups unless a phenotype is apparent.

Here, we injected TALEN mRNAs targeting the enhanced green fluorescent protein (*eGFP*) gene into fertilized mouse eggs expressing eGFP ubiquitously under control of the *CAG* promoter (*pCAG*). Pups were analyzed by observation of green fluorescence, heteroduplex mobility assay (HMA), restriction fragment length polymorphism (RFLP) analysis and DNA sequencing to consolidate the method for detecting pups with TALEN-induced mutations.

Materials and Methods

TALEN plasmid construction and mRNA preparation

Synthesized TALE repeats were cloned into pBluescript SK and assembled using the Golden Gate cloning method [12]. The N- and C-terminal domains of TALE and the *FokI* nuclease domain were taken from pTALEN_v2 (Addgene, Cambridge, MA, USA) [13]. The *eGFP* TALEN target sequence was described previously [11] and is indicated in Fig. 1. TALEN mRNAs were synthesized from plasmids linearized by *SmaI* digestion using an mMessage mMachine T7 Ultra Kit (Life Technologies, Carlsbad, CA, USA) and purified with an RNeasy Mini Kit (Qiagen, Hilden, Germany) following the manufacturers' instructions and as previously described [11].

pCAG-eGFP mouse embryos

The parental *pCAG-eGFP* mouse strain has been deposited in the Center for Animal Resources and Development (CARD), Kumamoto University (B6;D2-*Tg (CAG-EGFP) 49SImeg*; CARD ID: 267; http://cardb.cc.kumamoto-u.ac.jp/transgenic/strainsDetailAction.do?strainId=267). The background strain is C57BL/6. The *pCAG-eGFP* gene, illustrated in Supplementary Fig. 1A, was confirmed to be one copy by Southern blot analysis (Supplementary Fig. 1B). Expression of the *eGFP* gene is detected throughout the whole body.

Fig. 1. Diagram depicting engineered TALENs binding to the *eGFP* gene. The TALENs comprise DNA-binding repeats (gray boxes), the N- and C-terminal domains of TALE (white boxes) and a *FokI* nuclease domain (black box). The left and right TALEN target sequences are underlined.

To obtain mouse embryos for TALEN injections, C57BL/6N female mice were induced to super-ovulate using pregnant mare serum gonadotropin (PMSG; Serotropin; ASKA Pharmaceutical, Tokyo, Japan) and human chorionic gonadotropin (hCG; Veterinary Puberogen; Novartis Animal Health, Tokyo, Japan) at 5 weeks of age, and then mated with male *pCAG-eGFP* mice described above. Fertilized eggs were collected from females displaying vaginal plugs.

Microinjection of TALEN mRNAs

TALEN mRNAs were diluted in RNase-free PBS at 100 or 150 ng/μl for injection of each TALEN into the pronuclei or cytoplasm of zygotes. Approximately 2–3 pL of capped mRNAs was injected into the zygotes. The injected embryos were cultured in potassium simplex optimized medium with amino acids (KSOM-AA) at 37°C in 5% CO_2 and 95% humidified air for 1 h. Surviving embryos were transferred to the oviducts of pseudo-pregnant ICR female mice.

Genomic PCR for HMA and DNA sequencing

Genomic DNA was extracted from the tail of each pup using a DNeasy Blood & Tissue Kit (Qiagen). Genomic PCR was performed using LA *Taq* DNA polymerase (TAKARA Biotechnology, Shiga, Japan) under the following conditions: 94°C for 2 min; followed by 94°C for 30 s, 64°C for 30 s, and 72°C for 20 s for 38 cycles. The PCR primers were as follows: 5′-CCTCGTGACCACCCTGACCTAC-3′ and 5′-CTGTTGTAGTTGTACTCCAGCTTGTGC-3′. The PCR products were subjected to agarose gel electrophoresis and ethidium bromide staining for the HMA.

Table 1. TALEN-mediated eGFP gene disruption in mice

Route	Dose (ng/μl)	Injected	Transferred	Newborns	eGFP disappeared pups*	Analyzed pups	HMA	RFLP	Mutants
pronucleus	100	141	130	51 (39.2%)	3 (5.9%)	NT	NT	NT	NT
pronucleus	150	128	116	32 (27.6%)	4 (12.5%)	NT	NT	NT	NT
cytoplasm	150	80	69	33 (47.8%)	4 (12.1%)	33	12 (36.4%)	11 (33.3%)	17 (51.5%)
no injection	–	–	21	8 (38.1%)	0 (0%)	NT	NT	NT	NT

*Not including mosaic pups. HMA, heteroduplex mobility assay. RFLP, restriction fragment length polymorphism. NT, not tested.

For DNA sequence analysis, the PCR products were subcloned into pGEM-T Easy (Promega, Madison, WI, USA). The plasmids were extracted and sequenced using a T7 (5'-TAATACGACTCACTATAGGG-3') or SP6 (5'-CATACGATTTAGGTGACACTATAG-3') primer with a BigDye Terminator Cycle Sequencing Kit (Life Technologies), and then analyzed using an ABI PRISM 3130 Genetic Analyzer (Life Technologies).

RFLP analysis

The PCR products were purified using a Wizard SV Gel and PCR Clean-Up System (Promega). The purified products were digested with 3 units of *Acc*II (TAKARA Biotechnology), and then subjected to agarose gel electrophoresis and ethidium bromide staining.

Results and Discussion

As a model for TALEN-mediated knockout, we used fertilized eggs from mice ubiquitously expressing *pCAG-eGFP*. All of the fertilized eggs were heterozygous for the *pCAG-eGFP* gene. We selected the same *eGFP* TALEN target sequence as described previously [11] and newly constructed TALEN expression vectors as described in the Materials and Methods.

First, to judge the efficiency and toxicity between pronuclear and cytoplasmic injections, we microinjected *eGFP* TALEN mRNAs at 100 or 150 ng/μl into the pronuclei of zygotes. After the microinjection, 92.2% (130/141) and 90.6% (116/128) of the TALEN-injected embryos survived, respectively (Table 1). Following transfer of the surviving embryos to pseudopregnant females, 51 pups were born from 130 transferred embryos (39.2%) at 100 ng/μl and 32 pups were born from 116 transferred embryos (27.6%) at 150 ng/μl. Observation of the pups for green fluorescence under ultraviolet light on the day of birth revealed that eGFP fluorescence was completely abolished in three pups (5.9%) at 100 ng/μl and in four pups (12.5%) at 150 ng/μl (Table 1). The fluorescence images and DNA sequences of two of the three eGFP-disrupted pups at 100 ng/μl are shown in Fig. 2A and 2B, respectively. In addition, we obtained some pups with mosaic disruption of eGFP fluorescence at 150 ng/μl (Supplementary Fig. 2).

Next, *eGFP* TALEN mRNAs were injected at 150 ng/μl into the cytoplasm of zygotes. After the microinjection, 86.3% (69/80) of TALEN-injected embryos survived (Table 1) and 33 pups were born from 69 transferred embryos (47.8%). On the other hand, after transfer of *pCAG-eGFP* mouse embryos that had not been injected with TALEN mRNAs, eight pups were born from 21 embryos (38.1%). Thus, toxicity was not observed after the microinjection of TALEN mRNAs into the cytoplasm. When we observed the pups for green fluorescence under ultraviolet light, we found four pups in which fluorescence was completely absent and four mosaic pups (12.1% each) (Tables 1 and 2). Although the percentage of the green fluorescence-disappeared pups differed little from pronuclear injection, the birth rate of the pups with cytoplasmic injection was much higher than that with pronuclear injection (Table 1). Therefore, we concluded that TALEN mRNAs should be injected into the cytoplasm rather than the pronuclei. We then examined several analytical methods using the 33 pups described above.

Genomic DNA was extracted from all pups, and subjected to genomic PCR. The amplified product, including the TALEN target site, was 264 bp if no mutation was present. The individual products were subjected to agarose gel electrophoresis for the HMA. A recent study demonstrated that PCR products including TALEN-induced mutations could be detected by the HMA [9]. The HMA is the easiest method for the detection of mutations, because it only requires the performance of agarose gel electrophoresis after the genomic PCR. If mutations are introduced in the target DNA fragments, the shifted bands appear in proportion to the mutation rate based on the formation of heteroduplexes between the

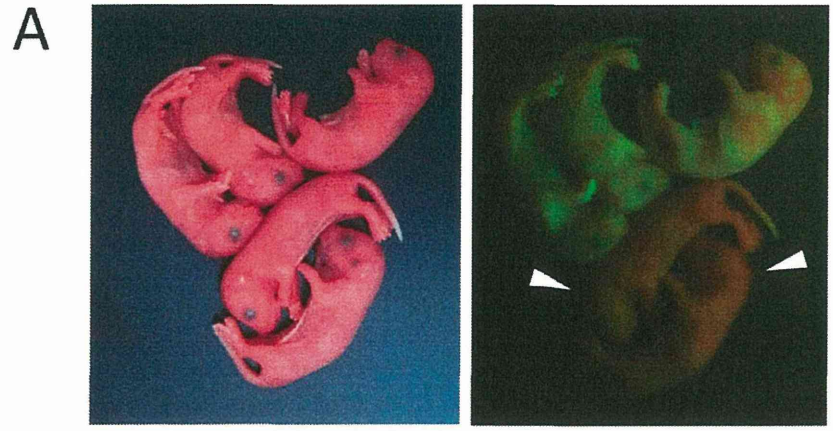

Fig. 2. TALEN-mediated disruption of the *eGFP* gene in mice. (A) Bright-field (left panel) and fluorescence microscopy (right panel) images of newborn mice. *eGFP* TALEN mRNAs were injected into fertilized eggs heterozygous for *eGFP*. Embryos were transferred to pseudopregnant females. The arrowheads indicate pups in which *eGFP* is disrupted. (B) eGFP sequences in pups displaying a disrupted *eGFP* gene (M1 and M2). The original sequence is shown at the bottom with the TALEN target sequences (underlined). Deletions are indicated by dashes.

Table 2. Summary of the analyses for mutant screening

Founders	1	2	3	4	5	6	7	8	9	10	11
eGFP disruption	mosaic	ND	ND	ND	ND	ND	mosaic	ND	ND	ND	ND
HMA	+	ND	ND	ND	+	+	ND	ND	ND	ND	+
RFLP	+	ND	ND	ND	+	ND	ND	ND	ND	ND	+
Sequence mutation	NT	NT	NT	NT	NT	+	+	NT	NT	NT	NT
Genotype	MT				MT	MT	MT				MT
Founders	12	13	14	15	16	17	18	19	20	21	22
eGFP disruption	mosaic	ND	ND	ND	ND	ND	ND	ND	+	+	mosaic
HMA	+	ND	ND	ND	ND	+	ND	ND	+	ND	ND
RFLP	+	ND	+/–	ND	ND	+	ND	ND	ND	+	ND
Sequence mutation	NT	NT	+	NT	NT	NT	NT	NT	NT	+	+
Genotype	MT		MT			MT			MT	MT	MT
Founders	23	24	25	26	27	28	29	30	31	32	33
eGFP disruption	ND	ND	+	+	ND	ND	ND	ND	ND	ND	ND
HMA	+	+	ND	+	ND	ND	+	ND	+	ND	ND
RFLP	+	+	ND	ND	ND	ND	+/–	ND	+	ND	ND
Sequence mutation	NT	NT	+	NT	ND	NT	NT	NT	NT	NT	NT
Genotype	MT	MT	MT	MT			MT		MT		

HMA, heteroduplex mobility assay. RFLP, restriction fragment length polymorphism. ND, not detected. NT, not tested. MT, mutant.

mutated alleles and non-mutated alleles. We confirmed shifted bands in 12 pups (#1, 5, 6, 11, 12, 17, 20, 23, 24, 26, 29 and 31) (Table 2 and Supplementary Fig. 3). In two pups (#21 and 29), we detected PCR products of 264 bp and about 200 bp. These smaller bands suggested that extensive deletions were induced in these pups.