Genetic and functional analysis of ZBTB24, which is mutated in type 2 ICF syndrome

Hirohisa Nitta

DOCTOR OF PHILOSOPHY

Department of Genetics

School of Life Science

The Graduate University for Advanced Studies

2013

Contents

Abstract

Abbreviations

Gene Symbols

Introduction

Materials & Methods

· Patients

· Cell lines

· Southern blotting

· Mutation analysis of the *DNMT3B* and *ZBTB24* genes in ICF patients

· Plasmids

· Immunofluorescent analysis

· Generation of *Zbtb24* knockout mice

Results

· Clinical characteristics of three ICF syndrome patients

· Diagnosis of the three patients as ICF2

· Identification of three novel *ZBTB24* mutations in the two ICF2 patients

· Requirement of the zinc finger domains of ZBTB24 for heterochromatin targeting

· DNA methylation- and repeat-sequence-independent heterochromatin targeting of ZBTB24

· *Zbtb24* knockout mice have hypomethylated satellite repeats

Discussion

Acknowledgements

References

Figures 1-9

Tables 1-3

Abstract

Centromeric and pericentromeric regions of the genome are important for stable chromosome structure, sister chromatid separation, and proper chromosome segregation. These regions are composed of highly condensed repetitive DNA sequences and form constitutive heterochromatin during interphase. Centromeric and pericentromeric regions contain different repetitive sequences: human centromeric and pericentromeric regions contain α-satellite repeat and satellite-2 and -3 repeats, respectively, and corresponding mouse regions contain minor satellite repeat and major satellite repeat, respectively. Transcription of these repeats is epigenetically repressed, and disruption of this repression causes transcription of the repeats and aberrant chromosome segregation. DNA methylation is important for the epigenetic repression and stabilization of the constitutive heterochromatin formed at these repeats. However, the mechanisms of establishment and maintenance of DNA methylation at the satellite repeats remain unclear.

In mammals, there are three DNA methyltransferases (DNMTs), DNMT1, DNMT3A, and DNMT3B. DNMT1 is a maintenance methyltransferase, which copies the pre-existing methylation patterns from the parental DNA strands to newly synthesized DNA strands during DNA replication. DNMT3A and DNMT3B are *de novo* methyltransferases, which methylate unmethylated DNA at CpG dinucleotides and create new methylation patterns.

Immunodeficiency, centromeric instability, and facial anomalies (ICF) syndrome is a rare autosomal recessive disorder that shows DNA hypomethylation at pericentromeric satellite-2 and

-3 repeats in chromosomes 1, 9, and 16. ICF syndrome is classified into two groups: type 1 (ICF1) patients have mutations in the *DNMT3B* gene, and about half of type 2 (ICF2) patients have mutations in the *ZBTB24* gene. Besides satellite-2 and -3 repeats, α-satellite repeats are also hypomethylated in ICF2. Importantly, all ICF1 patients have at least one hypomorphic *DNMT3B* allele perhaps because null mutants are lethal. In contrast, many ICF2 patients with *ZBTB24* mutations are null mutants. ZBTB24 protein belongs to the BTB-zinc finger domain family, which contains a BTB domain, an AT-hook domain, and eight C2H2-type zinc finger domains, but its biological function is totally unknown.

Although Dnmt3b null knockout mice show embryonic lethality at around E14.5-16.5 with multiple tissue defects, mice homozygous for ICF1 hypomorphic mutations develop to term and some of them survive to adulthood. These mice show phenotypes that are reminiscent of ICF1 patients, including hypomethylation of the satellite repeats, distinct cranial facial anomalies, and T-cell death. Therefore, *DNMT3B* has been experimentally proved to be the causative gene for ICF1. However, the role of ZBTB24 in satellite repeat methylation is unknown and there is no functional evidence that *ZBTB24* is a causative gene for ICF2.

In this study, I determined the sequences of *ZBTB24* in two Japanese and one Cape Verdean ICF2 patients, and found three novel *ZBTB24* mutations. A Japanese patient was homozygous for a novel missense mutation (C383Y), and a Cape Verdean patient was a compound heterozygote for a novel nonsense mutation (K263X) and a novel frameshift mutation (C327W fsX54). The second Japanese patient was homozygous for a previously reported nonsense mutation (R320X). The C383Y mutation abolished a C2H2 motif in one of the eight zinc finger

domains, and the other three mutations caused a complete or large loss of the zinc finger domains.

To examine the importance of the zinc finger domains, I performed immunofluorescent analysis with EGFP fusion proteins, and found that wildtype mouse Zbtb24 localizes at constitutive heterochromatin in NIH3T3 mouse fibroblast cells. In contrast, Zbtb24 mutant proteins possessing either of the mutations corresponding to C383Y and R320X are mislocalized from pericentrometic heterochromatin, suggesting that the zinc finger domains are required for proper intranuclear localization of this protein.

Next, I sought a factor(s) that is required for the proper intranuclear localization of this protein. Because Kaiso/ZBTB33, which belongs to the BTB-zinc finger domain family, recognizes methylated-CpG dinucleotides through its zinc finger domains, I first examined whether ZBTB24 protein recognizes methylated CpGs at constitutive heterochromatin. When wildtype mouse Zbtb24 was overexpressed in *Dnmt1, Dnmt3a, and Dnmt3b* triple knockout embryonic stem (TKO ES) cells, which lack DNA methylation genome-wide, the localization of the protein did not differ, compared with that in wildtype ES cells. This result indicates that DNA methylation is not required for the proper localization of Zbtb24 protein. Next, I examined whether Zbtb24 protein recognizes the DNA sequences of the satellite repeats, which largely differ between the mammalian species. When I expressed human ZBTB24 in mouse NIH3T3 cells, the human protein localized at constitutive heterochromatin, indicating that the DNA sequences of the satellite repeats are not required for the proper localization of Zbtb24 protein.

ZBTB24 might recognize some structual components of the chromatin or RNAs contained in the chromatin.

Finally, I generated *Zbtb24* knockout mice to examine whether a lack of Zbtb24 causes the distinctive phenotypes of ICF2 patients. Although ICF2 patients homozygous for *ZBTB24* null mutations are born alive, *Zbtb24* null mice were embryonic lethal at around E10.5. Importantly, yolk sac cells of *Zbtb24* null embryos clearly showed hypomethylation in major and minor satellite repeats although the phenotype was less clear in embryonic tissues. Therefore, I obtained the first experimental evidence that ZBTB24 is functionally involved in satellite repeat methylation, which further suggests that *ZBTB24* is a *bona fide* causative gene of ICF2. Since the yolk sac tissue contains blood islands, the source of hematopoietic stem cells, the *Zbtb24* null cells from this tissue may be useful in studying the developmental origin of immunodeficiency in ICF2 patients. Altogether, this study will provide a basis for understanding the pathogenesis of ICF2 and the mechanisms involved in formation of heterochromatin at the satellite repeats.

Abbreviations

BTB: BR-C, Ttk and Bab

DAPI: 4',6-DiAmidino-2-PhenylIndole

EDTA: Ethylenediamine Tetraacetic Acid

EGFP: Enhanced Green Fluorescent Protein

En2 SA: Engrailed2 Splice Acceptor

ES: Embryonic Stem

EuMMCR: European Mouse Mutant Cell Repository

FBS : Fetal Bovine Serum

FRT: FLP Recombinase Target

hBactP: human β-actin Promoter

ICF: Immunodeficiency, Centromeric instability, and Facial anomalies

ICF1: ICF syndrome type 1

ICF2: ICF syndrome type 2

IRES: Internal Ribosome Entry Sites

KO: KnockOut

kbp: kilobase pair

LIF : Leukemia Inhibiting Factor

LCL: Lymphoblastoid Cell Line

neo: neomycin resistant

pA: polyadenylation signal

PBS: Phosphate-Buffered Saline

PCR: Polymerase Chain Reaction

PFA: Paraformaldehyde

SDS: Sodium Dodecyl Sulfate

SSC: Saline Sodium Citrate

SSPE: Saline Sodium Phosphate EDTA

TKO: Triple KnockOut

WT: WildType

Gene Symbols

ATRX: Alpha thalassemia/mental retardation syndrome X-linked

BCL6: B cell lymphoma 6

CBX5: Chromobox homolog 5

DNMT1: DNA methyltransferase 1

DNMT3A: DNA methyltransferase 3A

DNMT3B: DNA methyltransferase 3B HELLS: Helicase, lymphoid-specific

HP1α: Heterochromatin protein 1 alpha

LSH: Lymphoid specific helicase

ZBTB24: zinc finger and BTB domain containing 24

ZBTB27: Zinc finger and BTB domain containing 27

ZBTB33: Zinc finger and BTB domain containing 33

Introduction

Immunodeficiency, centromeric instability, and facial anomalies (ICF) syndrome (MIM 242860) is an extremely rare autosomal recessive disorder accompanied by recurrent infectious diseases (Hulten, 1978; Tiepolo *et al.*, 1978). The immunodeficiency phenotype is characterized by reduction in various types of serum immunoglobulin. Peripheral blood cells of ICF patients contain only naive B cells, and lack memory B cells and gut plasma cells. Because naive ICF B-cells bear potentially autoreactive long heavy chain variable regions complementarity determining region 3, it is possible that negative selection breakdown and peripheral B-cell maturation blockage contribute to agammagloblinemia in the ICF syndrome (Blanco-Betancourt *et al.*, 2004). Facial anomalies distinctive for ICF syndrome patients include hypertelorism, low-set ears, epicanthic folds, and macroglossia.

The centromeric instability is manifested by undercondensation of the pericentromeric heterochromatin of chromosomes 1, 9, and 16, and multiradial chromosome configurations involving these regions (Maraschio *et al.*, 1988). These regions contain satellite-2 and -3 repeats, and the cytological defects are associated with DNA hypomethylation of the repeat sequences at the pericentromeric heterochromatin regions (Smeets *et al.*, 1994; Jeanpierre *et al.*, 1993; Miniou *et al.*, 1994). In addition, some of female ICF patients have a hypomethylated inactive X chromosome, although gene expression and replication timing of the inactive X chromosome are not affected by the hypomethylation (Bourc'his *et al.*, 1999).

About half of ICF syndrome patients are classified as type 1 (ICF1), which shows DNA hypomethylation of only satellite-2 and -3 repeats, and the other half are classified as type 2 (ICF2), which shows hypomethylation of α-satellite repeats, in addition to satellite-2 and -3 repeats, at the centromeric regions of all chromosomes (Jiang *et al*., 2005). ICF1 is caused by mutations in the *DNMT3B* gene, which encodes one of the two *de novo* DNA methyltransferases (Okano *et al*., 1999; Xu *et al*., 1999; Hansen *et al*., 1999). On the other hand, while about half of ICF2 patients possess mutations in the *ZBTB24* gene (de Greef *et al*., 2011), the gene(s) responsible for the other half of ICF2 patients is still unknown.

Methylation of cytosine residues in CpG dinucleotides plays a major role in gene repression, transposon silencing, and genome stabilization. The enzymatic activities catalyzing DNA methylation can be classified into two types. One is maintenance methylation, which is an activity to methylate unmethylated cytosine residues of hemimethylated CpGs during DNA replication, and DNA methyltransferase 1 (DNMT1) is responsible for this function (Gruenbaum *et al*., 1982; Bestor and Ingram, 1983). The other is *de novo* methylation, which is an activity to add methyl groups to cytosine at unmethylated CpGs, and DNMT3A and DNMT3B are known to catalyze this reaction (Okano *et al*., 1999). Genetic studies have indicated that DNA methylation is required for normal mouse development, because functional deficiency of Dnmt1 results in embryonic lethality at around E8.5-9.5 with a severe reduction of genomic DNA methylation (Li *et al*., 1992). Similarly, mice deficient for Dnmt3a show postnatal lethality, and mice deficient for Dnmt3b show embryonic lethality at around E14.5-16.5 with multiple tissue defects including liver hypotrophy, ventricular septal defect, and haemorrhage (Okano *et al*.,

1999). On the other hand, mice with homozygous *Dnmt3b* hypomorphic mutations similar to ICF1 patients can survive to adulthood. The mice show phenotypes similar to those of human ICF1 patients including hypomethylation of satellite repeats in all tissues examined, low body weight, distinct cranio-facial anomalies, and T cell apoptosis (Ueda *et al.*, 2006). *De novo* Dnmts methylate CpGs at particular genomic regions including repetitive sequences. The mechanism of target recognition by the enzymes is still largely unknown, although it is plausible that various proteins and/or other factors such as small RNAs help the enzymes to target particular regions. For example, deficiency of LSH/HELLS or ATRX, both of which are SWI/SNF-like chromatin remodeling proteins, results in changes in DNA methylation at repetitive sequences (Dennis *et al.*, 2001; Gibbons *et al.*, 2000). In fission yeast, small RNAs generated from repeat RNAs are involved in heterochromatin formation at pericentrometic regions (Lippman *et al*., 2004).

ZBTB24 protein, which is mutated in approximately half of ICF2 patients, has a BTB domain and an AT-hook domain in its N-terminal region, and eight C2H2-type zinc finger domains in its C-terminal region, but the molecular function of this protein is totally unknown. The other members of the BTB-zinc finger domain family include Kaiso/ZBTB33, which recognizes methylated DNA by its zinc finger domains (Prokhortchouk *et al*., 2001), and BCL6/ZBTB27, which is required for differentiation of naive B-cells into memory B-cells in germinal centers (Dent *et al*., 1997). Thus, ZBTB24 might be directly involved in regulation of DNA methylation of the repetitive sequences, and/or regulate transcription of genes or small RNAs, which are necessary for the DNA methylation.

In this study, I identified three novel *ZBTB24* mutations, all of which affected the zinc finger domains of the protein. I then showed the importance of the zinc finger domains for heterochromatin targeting, and revealed that DNA methylation and specific repeat sequences are not required for the localization. Although *ZBTB24* is mutated in many ICF2 patients, experimental validation of the importance of this protein as a causative gene for ICF2 syndrome has not been achieved. Therefore, I generated *Zbtb24* knockout mice and obtained first experimental evidence that ZBTB24 is functionally involved in satellite repeat methylation. The results strongly suggested that *ZBTB24* is a *bona fide* causative gene of ICF2.

Matetials & Methods

Patients

Three ICF2 patients were studied. Two of them were Japanese males (P6 and P7) and the last one was a Cape Verdean female (P8). Research protocols and consent forms were reviewed and approved by the ethical committee of each University/Medical Center/Hospital. Genomic DNA was isolated from peripheral blood leukocytes of the patients by a standard method.

Cell lines

NIH3T3 mouse fibroblast cells were maintained in Dulbecco's modified Eagle's medium (DMEM) (Sigma-Aldrich) supplemented with 10% FBS and penicillin/streptomycin at 37 ^{o}C under 5% CO_2. An ES cell line, which has a wildtype allele and a *Zbtb24* knockout allele (*Zbtb24*$^{tm1a(EUCOMM)Hmgu}$), with C57BL/6N (substrain JM8A3.N1) background carrying an agouti coat color gene was obtained from European Mouse Mutant Cell Repository (EuMMCR). Wildtype mouse ES cells (J1), Dnmt1/Dnmt3a/Dnmt3b triple knockout (TKO) ES cells (Tsumura *et al.*, 2006), and the ES cells carrying a *Zbtb24* knockout allele were maintained in DMEM supplemented with 15% FBS, 1x MEM Non-essential amino acids solution (Life Technologies), 0.1 mM β-mercaptoethanol, 500 U/ml LIF and penicillin/streptomycin on irradiated mouse embryonic fibroblasts at 37^{o}C under 5% CO_2.

Southern blot analysis

For human methylation analysis, genomic DNA of 500 ng was digested with 5 units of methylation sensitive *Bst*BI (New England Biolabs) for 12 h at 65°C to study satellite-2, or with 5 units of methylation sensitive *Hha*I (Nippon Gene) for 12 h at 37°C to study α-satellite. For mouse methylation analysis, genomic DNA of 500 ng was digested with 5 units of methylation sensitive HpyCH4IV (New England Biolabs) for 12 h at 37°C to study major satellite. To study minor satellite, genomic DNA of 500 ng was digested with 5 units of methylation sensitive *Hap*II (TAKARA) or 5 units of methylation insensitive *Msp*I (Nippon Gene), which is an isoschizomer of *Hap*II, for 12 h at 37°C. For detection of the *Zbtb24* knockout allele in the *Zbtb24* knockout ES cells, genomic DNA of 10 μg was digested with *Eco*RI (Nippon Gene) for 12 h at 37°C. The digested DNA fragments were separated by electrophoresis using 1% agarose gel, and transferred to Hybond XL nylon membranes (GE Healthcare) in 0.4 N NaOH. The membranes were hybridized with a ^{32}P-labeled satellite-2 probe (5′-TCGAGTCCATTCGATGAT-3′) for the *Bst*BI digests (Figure 2a), a ^{32}P-labeled α-satellite probe (5′-ATGTGTGCATTCAACTCACAGAGTTGAAC-3′) for the *Hha*I digests (Tagarro *et al.*, 1994) (Figure 2b), a ^{32}P-labeled major satellite probe (5′-CATGGAAAATGATAAACATCCACTT -3′) for the HpyCH4IV digests (Figure 9a, b), a ^{32}P-labeled minor satellite probe (5′-CCACACTGTAGAACATATTAGAT-3′) for the *Hap*II and *Msp*I digests (Figure 9a, b), or a ^{32}P-labeled *Zbtbr24* probe (Figure 7a), which is a PCR product labeled using Amersham Megaprime DNA labeling systems (GE Healthcare), for the *EcoR*I digests. The PCR primer sequences for generating the *Zbtb24* probe were: forward, 5′-

CTTACAGGTTAGGTTGCATCTCATAG-3′; reverse, 5′-GGAACCTGGTATCCTTCTAATGTAAG-3′. PCR was performed using TaKaRa Ex Taq (Takara Bio) under the following conditions: 35 cycles of 95°C for 30 sec, 60°C for 40 sec, and 72°C for 30 sec. Hybridization was carried out for 24 h at 42°C in 6 × SSC, 5 × Denhardt's solution, and 0.1% SDS for methylation analysis, or in 5 × SSPE, 5 × Denhardt's solution, 0.5% SDS, 50% formamide, and 50ug/ml Herring sperm DNA for detection of the *Zbtb24* knockout allele. After hybridization, the membranes were washed in 6 × SSC and 0.1% SDS for methylation analysis, or in 2 × SSC and 0.5% SDS for detection of *Zbtb24* knockout allele, three times at room temperature, followed by autoradiography.

Mutation analysis of the *DNMT3B* and *ZBTB24* genes in ICF patients

Mutations were identified in ICF2 patients by sequencing the PCR products together covering all the exons and exon/intron boundaries of the *DNMT3B* and *ZBTB24* genes. The PCR primer sequences are listed in Table 1. PCR was performed using KOD–plus (TOYOBO) under the following conditions: 35 cycles of 98°C for 10 sec, 60°C for 30 sec and 68°C for 30 sec. The PCR products were purified using QIAquick Gel Extraction Kit (QIAGEN). The sequencing reaction was performed using BigDye Terminator v3.1 Cycle Sequencing kit (Applied Biosystems) with either of the two PCR primers and analyzed on 3130xl Genetic Analyzer (Applied Biosystems).

Plasmids

Mouse *Zbtb24*, mouse *Dnmt3b*, and human *ZBTB24* cDNAs containing the entire coding regions were synthesized by RT-PCR from total RNA of mouse ES cells and human HEK293T cells, respectively, using PrimeScript II 1st strand cDNA Synthesis Kit (Takara Bio) with random hexamers as primers. These genes were cloned into pEGFP-C1 plasmid vector (Clontech). Primers used for the PCR were: mouse *Zbtb24* (forward, 5′-AAAGATCTATGGCAGACACAACCCCGGAGCCTT-3′; reverse, 5′-AAGAATTCCCACGATAAAGAATGGCTGGCTCACATCTT-3′); mouse *Dnmt3b* (forward, 5′-AAGAATTCAATGAAGGGAGACAGCAGACATC-3′; reverse, 5′-AAGGATCCCTATTCACAGGCAAAGTAGTCC-3′); and human *ZBTB24* (forward, 5′-AAAGATCTATGGCAGAAACATCGCCAGAGCCTTCT-3′; reverse, 5′-AAGAATTCAAACGTGTTTACAGGTCAGCTCTGCTCC-3′). PCR was performed using KOD–plus under the following conditions: 35 cycles of 98°C for 10 sec, 60°C for 30 sec and 68°C for 2 min. Mutations corresponding to those found in two Japanese ICF2 patients P6 and P7 were introduced into the mouse *Zbtb24* using PCR primers carrying the mutations. The primers used for PCR were: P6 (forward, 5′-ATGGAAAATATTTCAGCCAGAAAAG-3′; and reverse, 5′-ACTGATCACAGGTAAAAGACTTCTG-3′); P7 (forward, 5′-TGACCCTTCAAATGTAATGAGTGTG-3′; and reverse, 5′-TTCCCCCGTGTGGCGTCTCTGGTGGA-3′). PCR was performed using KOD–plus under the following conditions: 35 cycles of 98°C for 10 sec, 60°C for 30 sec and 68°C for 6 min.

Immunofluorescent analysis

Cells were transfected with EGFP fused wildtype or mutant *Zbtb24* expression plasmids using FuGENE HD Transfection Reagent (Promega) according to the manufactur's protocol. After 48 h of transfection, the cells grown on glass slides were fixed with 4% PFA at room temperature for 30 min, permeabilized by 0.1% Triton X at 4 ^{o}C for 3 min, and blocked by 1 x Block Ace (DS Pharma Biomedical) at room temperature for 1 h. Fixed cells were incubated with an anti-HP1α mouse monoclonal antibody (MAB3584; Merck Millipore) (1:1000 dilution) for 1 h at room temperature. Then the cells were incubated with Alexa Fluor 594 goat anti-mouse IgG (H+L) antibody (A11029; Life Technologies) (1:1000 dilution) for 1 h at room temperature. Nuclei were visualized by SlowFadeR Gold Antifade Reagent with DAPI (Life Technologies). Fluorescent images were taken using LSM510 confocul laser scanning microscopy (Carl Zeiss).

Generation of *Zbtb24* knockout mice

Chimeric mice were generated by injection of the ES cells carrying the *Zbtb24*$^{tm1a(EUCOMM)Hmgu}$ allele into host C57BL/6 blastocysts (Laboratory of Embryonic and Genetic Engineering, Medical Institute of Bioregulation, Kyushu University). Heterozygous mice were obtained from the chimeric mice, and maintained on a C57BL/6 background. Genotyping was performed by PCR using following primers: genome forward, 5′-GGAAATACCCTCTACTCTCCTAGC-3′; genome reverse, 5′-AAACATTCCAAGTCCAGCAGATAG-3′; and insert reverse, 5′-

AGGAGATGGAGAAGACTGAAGAAA-3′. genome forward and genome reverse primers were used for detecting wildtype allele, and genome forward and insert reverse primers were used for detecting the knockout allele. PCR was performed using TaKaRa Ex Taq (Takara Bio) under the following conditions: 35 cycles of 95°C for 30 sec, 58°C for 40 sec and 72°C for 30 sec.

Results

Clinical characteristics of three ICF syndrome patients

ICF syndrome is an extremely rare autosomal recessive disorder (Hulten, 1978; Tiepolo *et al.*, 1978). Through my collabolators in Japan and France, I obtained DNA of three ICF syndrome patients. Two patients were Japanese (ID: P6, P7), and one patient was Cape Verdean (ID: P8). Here I describe the clinical characteristics of these patients.

■**P6:** A 5-year-old Japanese boy born in 2003 from non-consanguineous parents. His birth weight was 2,880 g, height was 51.0 cm, and occipito-frontal circumference was 32.7 cm. He had craniofacial features including macrocephaly, hypertelorism, epicanthal folds, midface flatness, low nasal root, long and flat philtrum, and thick lips (upper < lower). He had a thin habitus with slender arms, knock-knee, and slightly irregular toes. He had hypoplastic primary teeth and bilateral hydronephrosis. His developmental quotient was 80, with delayed motor development and apparently normal mental development. He showed short stature because of hypopituitarism resulting from compression by Rathke's cleft cyst, which was treated with replacement of human growth hormone and thyroid hormone. He had skin abnormalities including eruption of blisters at the onset of fever, and freckle formation by exposure to sunlight with no butterfly erythema. He suffered from refractory diarrhea and recurrent respiratory tract infections. He finally died of fatal viral infection at the age of 7. His serum IgG2, IgA and IgM levels were decreased and the IgG3 level was increased (Table 2). The numbers of T cells and B

cells were in the normal range, but $CD27^+$ memory B cells were hardly detectable (Table 2). The activity of natural killer (NK) cell was low. The lymphocyte blastoid transformation test using phytohaemagglutinin (PHA) was in the normal range. He did not produce antibodies in response to vaccines or pathogens. A cytogenetic analysis revealed that 7% (2/30 cells) of peripheral blood leukocytes had a large deletion of the long arm of chromosome 1, designated as del(1)(q21) (Figure 1a).

■**P7:** A 39-year-old Japanese man born in 1971 from consanguineous parents. His birth weight was 2,600 g. He showed dysmorphic facial features including micrognathia and hypertelorism. He showed a delay in language development and had learning disability. He had recurrent pneumonia and sinusitis and was on intravenous immunoglobulin (IVIG) therapy. At the age of 39, he had ingravescence of tremor, articulation disorder, and left hemiplegia. He also showed involuntary movements of the right lower limb and left toes, choreoathetotic movement of the right upper limb, ambulation difficulty, increased tendon reflex, Babinski's reflex, and vesicorectal disorders accompanied by incontinence without sensory impairment. The brain magnetic resonance imaging showed cortical atrophy. His intelligence quotient (IQ) was 47. He had progressive multifocal leukoencephalopathy due to John Cunningham virus infection and died at the age of 41. His serum levels of IgG, IgG1, IgG2, IgA, and IgM were decreased (Table 2). The number of T cells was in the normal range, but that of B cells was decreased (Table 2). The NK cell activity and lymphocyte blastoid transformation count were also decreased. A

cytogenetic analysis revealed multiradial chromosome configurations with multiple p and q arms from chromosomes 1 and 16 (Figure 1b).

■**P8:** A 16-year-old Cape Verdean girl born in 1996 from nonconsanguineous parents (Dupont *et al.,* 2012). She had characteristic facial features including epicanthic folds. She had recurrent thrush caused by candida infection in the first year of life. Her initial speech was delayed. She had recurrent ear, nose, and throat infections and had a secondary hearing impairment. She suffered from recurrent bacterial pulmonary infections that caused bronchiectasis and atelectasis. She had been on IVIG therapy since she was 10 years old. She had a lobotomy (left lower lobe) at the age of 12 because of the severe bronchiectasis. She had mental retardation (IQ < 70) and did not acquire reading at the age of 16. Her serum IgG and IgA levels before IVIG treatment were in the normal range, but the IgM level was decreased (Table 2). The number of T cells was in the normal range, but that of $CD27^+$ memory B cells was decreased (Table 2). She had low antibody titers against tetanus (0.32 IU/ml), no antibody against pneumococcus, and no IgM blood group isohemagglutinin. T cells proliferated normally in response to PHA and tetanic toxin. A cytogenetic analysis of peripheral blood leukocytes showed that 8% of the cells had whole-arm deletions and/or breaks in chromosomes 1 and 16. Multibranched chromosomes were also observed (Dupont *et al.,* 2012).

Diagnosis of the three patients as ICF2

To determine whether the three patients belong to ICF1 or ICF2, I first examined the DNA methylation status of the satellite-2 and α-satellite repeats in the patients by Southern blot analysis using methylation-sensitive restriction enzymes. When satellite-2 was examined, the genomic DNAs from the three patients, as well as an ICF1 patient (P3) (Shirohzu *et al.*, 2002), gave multiple low-molecular-weight bands, indicating that this repeat was hypomethylated (Figure 2a). In contrast, the DNAs from the mother of P6 (P6M) and a normal control individual (N) were resistant to the digestion. When α-satellite was examined, the DNAs from all patients gave low-molecular-weight bands, whereas those from an ICF1 patient (P3) and a normal control (N) were resistant to the digestion (Figure 2b). Therefore, the three patients P6, P7, and P8 showed a DNA methylation pattern diagnostic of ICF2. Sequencing of the PCR products from the three patients revealed no mutation in all coding exons of the *DNMT3B* gene, further supporting that these patients are ICF2.

Identification of three novel *ZBTB24* mutations in the two ICF2 patients

Because it was reported that about a half of ICF2 patients have mutations in *ZBTB24* gene (de Greef *et al*., 2011), I looked for mutations in the gene, which is located on the long arm of chromosome 6 and is composed of seven exons. I found a total of four mutations, three of which were novel (Table 3 and Figure 3a,b). P6 was homozygous for a G-to-A substitution (c.1148G>A) causing an amino acid substitution at Cys-383 in a zinc finger domain (C383Y). His parents were heterozygous for this mutation although they were not consanguineous. P7 was homozygous for a C-to-T substitution (c.958C>T) causing a nonsense mutation at codon 320

(R320X). This mutation was previously reported in ICF2 (de Greef *et al.,* 2011). His parents were consanguineous and the mother was found to be heterozygous for the mutation. P8 was a compound heterozygote with a nonsense mutation (c.787A>T) at codon 263 (K263X) and a frame-shift mutation (c.980_981delGT) at codon 327, generating a termination codon at 54-codon downstream (C327W fsX54). Besides the missense mutation identified in P6 (C383Y), only one other missense mutation was previously reported (c.1222T>G, C408G) (Table 3 and Figure 3a) (de Greef *et al.,* 2011; Cerbone *et al.,* 2012). Interestingly, the affected cysteines are located in the fourth and fifth C2H2 motifs, respectively, and are highly conserved among vertebrates (Figure 3c). These results strongly suggest the importance of these motifs for the protein function and/or structure.

Requirement of the zinc finger domains of ZBTB24 for heterochromatin targeting

To gain insights into the biological function of the zinc finger domains of ZBTB24 and also the impact of ICF2 mutations on this function, I examined the subnuclear localization of the wildtype and mutant proteins. I used mouse NIH3T3 cells for the analysis because they have distinct pericentromeric heterochromatin foci seen as DAPI-dense regions. Because amino acid sequence of human and mouse ZBTB24 proteins, especially its zinc finger domains, is highly conserved (Figure 4a, homology: 93%), it is thought that these orthologues probably show similar subnuclear localization. Therefore, I transiently overexpressed wildtype mouse Zbtb24 protein fused with EGFP (EGFP-mZbtb24) in the cells. The transiently expressed EGFP-mZbtb24 (WT) showed colocalization with HP1α/CBX5, a heterochromatin marker protein, in

multiple DAPI-dense regions (Figure 4b). The EGFP-mDnmt3b (WT) fusion protein also showed heterochromatin localization (Figure 4b). In contrast, EGFP-mZbtb24 proteins with either of the mutations C382Y and R319X, equivalent to those identified in P6 (C383Y) and P7 (R320X), showed a rather dispersed distribution in the nuclei (Figure 4b), suggesting the importance of the zinc finger domains, especially the C2H2 motif of the domains, in heterochromatin targeting. (Please note that human ZBTB24 is one amino acid longer than mouse Zbit24.)

DNA methylation- and repeat-sequence-independent heterochromatin targeting of ZBTB24

To explore what factors are involved in heterochromatin targeting of ZBTB24, I examined two possible factors, DNA methylation and DNA sequences at the regions. First, I used DNA methylation deficient *Dnmt* TKO mouse ES cells, and observed that the complete loss of DNA methylation does not affect heterochromatin localization of the EGFP-mZbtb24 (WT) protein (Figure 5). This suggests that Zbtb24 does not require DNA methylation for proper localization.

Next, I examined whether human ZBTB24 protein recognizes the DNA sequences of mouse satellite repeat to localize at heterochromatin, because DNA sequences, length, and unit of human and mouse satellite repeats are largely different (Figure 6a); α-satellite and minor satellite constitute centromeric repeats in human and mouse, respectively, and satellite-2 (-3) and major satellite constitute pericentromeric repeats in human and mouse, respectively. The result showed that human ZBTB24 protein (EGFP-hZBTB24) co-localized with HP1α (a heterochromatin

marker protein) at DAPI dense regions in mouse NIH3T3 cells, indicating that the satellite DNA sequences are not required for proper localization of ZBTB24 (Figure 6b).

Zbtb24 knockout mice have hypomethylated satellite repeats

Although mutations in the *ZBTB24* gene are found in about half of ICF2 patients, *ZBTB24* has not been experimentally proved to be the causative gene. Therefore, I generated *Zbtb24* knockout mice to examine whether *ZBTB24* deficiency causes the ICF2 phenotype. An ES cell-line that carries a *Zbtb24* wildtype allele and a knockout (*Zbtb24*$^{tm1a(EUCOMM)Hmgu}$) allele was obtained from EuMMCR (Figure 7a). The knockout allele has an insertion of two cassettes in intron 2 of the *Zbtb24* gene. The two cassettes were flanked by two FRT sites (Figure 7a). The first cassette contains a splicing acceptor site (En2 SA), IRES, *β-galactosidase* gene (lacZ), and a polyadenylation signal (PA) to terminate transcription of the *Zbtb24* gene at exon 2, resulting in the generation of a fusion protein composed of truncated Zbtb24 (1-316 aa) and LacZ. The truncated Zbtb24 lacks seven of the eight zinc finger domains resembling R320X and C327W fsX54 mutations (Figure 3a). The second cassette contains a neomycin resistant (*neo*) gene, which was used as a marker for screening. The knockout allele also contained three loxP sites: first site was located between the two cassettes; second site was located after the second cassette in intron 2; the last site was located in intron 5 (Figure 7a). By Southern blot analysis, I detected wildtype allele as an 8-kbp band and the knockout allele as a 3-kbp band (Figure 7b), confirming the proper insertion of the cassettes in the *Zbtb24* gene in the ES cells.

The heterozygous *Zbtb24* (+/-) ES cells were used to generate chimeric mice. Two of the four chimeras showed successful germline transmission, and thus gave *Zbtb24*$^{+/-}$ mice. The *Zbtb24*$^{+/-}$ mice were healthy, and had normal fertility. I crossed female and male *Zbtb24*$^{+/-}$ mice to generate *Zbtb24* knockout (-/-) mice, but no *Zbtb24*$^{-/-}$ neonates were obtained (Figure 8a). Because *Zbtb24*$^{+/+}$ and *Zbtb24*$^{+/-}$ neonates were born normally in accordance with the expected Mendelian inheritance ratios (Figure 8a), *Zbtb24*$^{-/-}$ mice were thought to be embryonic lethal. To determine at which stage *Zbtb24*$^{-/-}$ embryos were lost, I examined embryos from embryonic day 9.5 (E9.5) onward. At E9.5 and E10.5, *Zbtb24*$^{-/-}$ embryos were still alive, but the E10.5 embryos exhibited severe growth retardation compared with *Zbtb24*$^{+/+}$ and *Zbtb24*$^{+/-}$ embryos. The phenotype of the *Zbtb24*$^{-/-}$ embryos varied among the individuals: they showed different sizes and different morphological defects (Figure 8a, 8c). From E11.5 onward, no *Zbtb24*$^{-/-}$ embryos were obtained. I also observed uterine scars at E9.5 and E10.5, possibly resulting from early resorption of *Zbtb24*$^{-/-}$ embryos.

To know whether *Zbtb24* knockout mice have hypomethylated satellite repeats similar to ICF2 patients, I examined the DNA methylation status of major and minor satellite repeats in whole embryos and yolk sacs (Samokhvalov *et al*., 2007) by Southern blot analysis using methylation-sensitive restriction enzymes. When I studied whole embryos at E10.5, I did not find clear methylation differences between the control (*Zbtb24*$^{+/+}$ and *Zbtb24*$^{+/-}$) and *Zbtb24*$^{-/-}$ genotypes at either major or minor satellite repeat (Figure 9a). In contrast, when I examined E10.5 yolk sac cells, major and minor satellite repeats were more methylated in the control

Zbtb24$^{+/+}$ and *Zbtb24*$^{+/-}$ genotypes than the *Zbtb24*$^{-/-}$ genotype (Figure 9b). This is the first experimental evidence that Zbtb24 is involved in DNA methylation at satellite repeats.

Discussion

In this study, I identified three novel and one previously reported *ZBTB24* mutations in three ICF2 patients (Table 3 and Figure 3a,b). Three of the four mutations were nonsense or frame-shift mutations (R320X, K263X and C327W fsX54). These mutations left the BTB domain and the AT-hook motif intact, but made the zinc finger domains partially or completely lost. Consistent with my finding, the majority of the *ZBTB24* mutations previously reported in ICF2 patients were also nonsense or frame-shift mutations (Table 3 and Figure 3a,b) (de Greef *et al.*, 2011; Cerbone *et al.*, 2012; Chouery *et al.*, 2012; Weemaes *et al.*, 2013). The last mutation that I identified altered a highly conserved cysteine in the C2H2 motif of the fourth zinc finger domain (C383Y). Interestingly, the only one missense mutation previously reported (C408G) affected a highly conserved cysteine in the fifth zinc finger domain (Figure 3c) (de Greef *et al.*, 2011; Cerbone *et al.*, 2012). These data highlight the importance of the cysteine residues of the C2H2 motifs of the ZBTB24 zinc finger domains in the pathogenesis of ICF2. However, the degree of defect in the ZBTB24 protein structure does not appear to be correlated with the severity of immunodeficiency and other clinical manifestations.

Because ICF2 patients have hypomethylated centromeric and pericentromeric repeats, ZBTB24 appears to play a role in the establishment and/or maintenance of DNA methylation at these repeats. Consistent with this function, I observed heterochromatic localization of the wildtype mouse Zbtb24 fusion protein, although how Zbtb24 targets heterochromatin is totally

unknown. In contrast, disruption of the zinc finger domains by either a missense mutation (C382Y) or a truncation mutation (R319X) abolished the heterochromatin targeting.

It is possible that ZBTB24 might recognize either the DNA methylation status or satellite repeat sequences to localize at heterochromatin. Therefore, I examined whether depletion of DNA methylation causes mislocalization of ZBTB24. My result showed that DNA methylation is not required for proper localization of ZBTB24, as wildtype mouse Zbtb24 protein localized at heterochromatin even in the absence of DNA methylation in Dnmt TKO mouse ES cells. Satellite DNA sequences were also not the determination factor of the subnuclear localization of ZBTB24, because human ZBTB24 protein localized at heterochromatin of mouse cells, which have totally different satellite repeat sequences. These results indicate that the zinc finger domains of ZBTB24 are important for the proper localization of this protein, and this localization does not involve methylated DNA or satellite repeat sequences. It is possible that ZBTB24 recognizes its target region through interactions with the other structural components of the chromatin, for example, heterochromatin proteins and/or associated RNAs.

Finally, in order to establish that *ZBTB24* is a causative gene for ICF2, I generated *Zbtb24* knockout mice. Although ICF2 patients with homozygous null-mutation in *ZBTB24* gene are born alive, *Zbtb24* deficient mice were embryonic lethal. It is possible that human and mouse ZBTB24 proteins have different functions and/or that human has a functionally redundant, compensatory protein(s). It is also possible that the genetic background of the *Zbtb24* knockout mice affected the phenotype: it might be interesting to bring these mice into strain backgrounds other than C57BL/6. I then showed that both major and minor satellite repeats were significantly

hypomethylated in *Zbtb24* knockout yolk sac cells compared with control cells at E10.5. This is the first experimental evidence that Zbtb24 is functionally involved in methylation of satellite repeats. In contrast, DNA methylation differences were not clear at either major or minor satellite repeats in E10.5 whole embryos between the different genotypes. Although it is possible that DNA methylation levels are specifically decreased in *Zbtb24* deficient yolk sac cells, some embryonic lineage cells such as immune linage cells might also be hypomethylated in *Zbtb24* knockout mice. Detailed analysis of DNA methylation levels at satellite repeats in different embryonic cells will be a next important work. Although there were differences between human and mouse, the hypomethylation of satellite repeats in yolk sac cells strongly supports that *ZBTB24* is a *bona fide* causative gene of ICF2.

It is known that the first hematopoiesis occurs in the blood islands of the yolk sac, and that hematopoietic stem cells in the yolk sac migrates to fetal liver to start secondary hematopoiesis (Samokhvalov *et al*., 2007). Thus the *Zbtb24* null cells from this tissue may be useful in studying the developmental origin of immunodeficiency in ICF2 patients. Also, because the *Zbtb24* knockout allele used in this study can be used to create conditional knockout mice, it might be helpful to generate B-cell specific *Zbtb24* knockout mice for studying the immunodeficiency phenotype of ICF2. Altogether, this study will provide a basis for understanding the pathogenesis of ICF2 and the mechanisms involved in formation of heterochromatin at the satellite repeats.

Acknowledgements

I would like to express my special thanks to Prof. Hiroyuki Sasaki and Dr. Motoko Unoki for their help and support throughout my PhD work. I would like to thank all patients and their families for participating to this study. I thank Drs Takeo Kubota, Tomoki Kosho, Tomonari Shigemura, Hiroshi Takahashi, Guillaume Velasco, Claire Francastel, and Capucine Picard for providing the DNA samples of ICF2 patients, and Dr. Masaki Okano for providing the Dnmt TKO mouse ES cells. I am grateful to the members of Laboratory of Embryonic and Genetic Engineering, Medical Institute of Bioregulation, Kyushu University for their help in generating *Zbtb24* knockout mice. I am also grateful to Drs Kenji Ichiyanagi, Takashi Sado, Yufeng Li, Tomoko Ichiyanagi, and Kei Fukuda for useful advice, and to Mr. Hiroyasu Furuumi for the technical assistance. I also thank all members of the Prof. Sasaki's laboratory for discussion and supports.

References

Bestor, T. H., and Ingram, V. M. 1983. Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with DNA. *Proc. Natl. Acad. Sci. U S A.* **80**: 5559-5563.

Blanco-Betancourt, C. E., Moncla, A., Milili, M., Jiang, Y. L., Viegas-Pequignot, E. M., Roquelaure, B., Thuret, I., and Schiff, C. 2004. Defective B-cell-negative selection and terminal differentiation in the ICF syndrome. *Blood* **103**: 2683-2690.

Bourc'his, D., Miniou, P., Jeanpierre, M., Molina Gomes, D., Dupont, J., De Saint-Basile, G., Maraschio, P., Tiepolo, L., and Viegas-Péquignot, E. 1999. Abnormal methylation does not prevent X inactivation in ICF patients. *Cytogenet. Cell Genet*. **84**: 245-252.

Cerbone, M., Wang, J., Van der Maarel, S. M., D'Amico, A., D'Agostino, A., Romano, A., and Brunetti-Pierri, N. 2012. Immunodeficiency, centromeric instability, facial anomalies (ICF) syndrome, due to ZBTB24 mutations, presenting with large cerebral cyst. *Am. J. Med. Genet. A.* **158A**: 2043-2046.

Chouery, E., Abou-Ghoch, J., Corbani, S., El Ali, N., Korban, R., Salem, N., Castro, C., Klayme, S., Azoury-Abou Rjeily, M., Khoury-Matar, R., Debo, G., Germanos-Haddad, M., Delague, V.,

Lefranc, G., and Mégarbané, A. 2012. A novel deletion in ZBTB24 in a Lebanese family with immunodeficiency, centromeric instability, and facial anomalies syndrome type 2. *Clin. Genet.* **82**: 489-493.

de Greef, J. C., Wang, J., Balog, J., den Dunnen, J. T., Frants, R. R., Straasheijm, K. R., Aytekin, C., van der Burg, M., Duprez, L., Ferster, A., Gennery, A. R., Gimelli, G., Reisli, I., Schuetz, C., Schulz, A., Smeets, D. F., Sznajer, Y., Wijmenga, C., van Eggermond, M. C., van Ostaijen-Ten Dam, M. M., Lankester, A. C., van Tol, M. J., van den Elsen, P. J., Weemaes, C. M., and van der Maarel, S. M. 2011. Mutations in ZBTB24 are associated with immunodeficiency, centromeric instability, and facial anomalies syndrome type 2. *Am. J. Hum. Genet.* **88**: 796-804.

Dennis, K., Fan, T., Geiman, T., Yan, Q., and Muegge, K. 2001. Lsh, a member of the SNF2 family, is required for genome-wide methylation. *Genes Dev.* **15:** 2940-2944.

Dent, A. L., Shaffer, A. L., Yu, X., Allman, D., and Staudt, L. M. 1997. Control of inflammation, cytokine expression, and germinal center formation by BCL-6. *Science* **276**: 589-592.

Dupont, C., Guimiot, F., Perrin, L., Marey, I., Smiljkovski, D., Le Tessier, D., Lebugle, C., Baumann, C., Bourdoncle, P., Tabet, A. C., Aboura, A., Benzacken, B., and Dupont, J. M. 2012. 3D position of pericentromeric heterochromatin within the nucleus of a patient with ICF syndrome. *Clin. Genet.* **82**: 187-192.

Gibbons, R. J., McDowell, T. L., Raman, S., O'Rourke, D. M., Garrick, D., Ayyub, H., and Higgs, D. R. 2000. Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the pattern of DNA methylation. *Nat. Genet.* **24:** 368-371.

Gruenbaum, Y., Cedar, H., and Razin, A. 1982. Substrate and sequence specificity of a eukaryotic DNA methylase. *Nature* **295**: 620-622.

Hassan, K. M., Norwood, T., Gimelli, G., Gartler, S. M., and Hansen, R. S. 2001. Satellite 2 methylation patterns in normal and ICF syndrome cells and association of hypomethylation with advanced replication. *Hum. Genet.* **109**: 452-462.

Hulten, M. 1978. Selective somatic pairing and fragility at 1q12 in a boy with common variable immunodeficiency. *Clin. Genet.* **14**: 294.

Jeanpierre, M., Turleau, C., Aurias, A., Prieur, M., Ledeist, F., Fischer, A., and Viegas-Pequignot, E. 1993. An embryonic-like methylation pattern of classical satellite DNA is observed in ICF syndrome. *Hum. Mol. Genet.* **2**: 731-735.

Jiang, Y. L., Rigolet, M., Bourc'his, D., Nigon, F., Bokesoy, I., Fryns, J. P, Hultén, M., Jonveaux, P., Maraschio, P., Mégarbané, A., Moncla, A., and Viegas-Péquignot, E. 2005. DNMT3B

mutations and DNA methylation defect define two types of ICF syndrome. *Hum. Mutat.* **25**: 56-63.

Kubota, T., Furuumi, H., Kamoda, T., Iwasaki, N., Tobita, N., Fujiwara, N., Goto, Y., Matsui, A., Sasaki, H., and Kajii, T. 2004. ICF syndrome in a girl with DNA hypomethylation but without detectable DNMT3B mutation. *Am. J. Med. Genet.* **129A**: 290-293.

Li, E., Bestor, T. H., and Jaenisch, R. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**: 915-926.

Lippman, Z., and Martienssen, R. 2004. The role of RNA interference in heterochromatic silencing. *Nature* **431**: 364-370.

Maraschio, P., Zuffardi, O., Dalla Fior, T., and Tiepolo, L. 1988. Immunodeficiency, centromeric heterochromatin instability of chromosomes 1, 9, and 16, and facial anomalies: the ICF syndrome. *J. Med. Genet.* **25**: 173-180.

Miniou, P., Jeanpierre, M., Blanquet, V., Sibella, V., Bonneau, D., Herbelin, C., Fischer, A., Niveleau, A., and Viegas-Péquignot, E. 1994. Abnormal methylation pattern in constitutive and facultative (X inactive chromosome) heterochromatin of ICF patients. *Hum. Mol. Genet.* **3**: 2093-2102.

Okano, M., Bell, D. W., Haber, D. A., and Li, E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**: 247-257.

Prokhortchouk, A., Hendrich, B., Jorgensen, H., Ruzov, A., Wilm, M., Georgiev, G., Bird, A., and Prokhortchouk, E. 2001. The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. *Genes Dev.* **15**: 1613-1618.

Samokhvalov, I. M., Samokhvalova, N. I., and Nishikawa, S. 2007. Cell tracing shows the contribution of the yolk sac to adult haematopoiesis. *Nature* **446**:1056-1061.

Shirohzu, H., Kubota, T., Kumazawa, A., Sado, T., Chijiwa, T., Inagaki, K., Suetake, I., Tajima, S., Wakui, K., Miki, Y., Hayashi, M., Fukushima, Y., and Sasaki, H. 2002. Three novel DNMT3B mutations in Japanese patients with ICF syndrome. *Am. J. Med. Genet.* **112**: 31-37.

Smeets, D. F., Moog, U., Weemaes, C. M., Vaes-Peeters, G., Merkx, G. F., Niehof, J. P., and Hamers, G. 1994. ICF syndrome: a new case and review of the literature. *Hum. Genet.* **94**: 240-246.

Tagarro, I., Ferna'ndez-Peralta, A. M., and Gonza'lez-Aguilera, J. J. 1994. Chromosomal localization of human satellites 2 and 3 by a FISH method using oligonucleotides as probes. *Hum. Genet.* **93**: 383-388.

Tiepolo, L., Maraschio, P., Gimelli, G., Cuoco, C., Gargani, G. F., and Romano, C. 1979. Multibranched chromosomes 1, 9 and 16 in a patient with combined IgA and IgE deficiency. *Hum. Genet.* **51**: 127-137.

Tuck-Muller, C. M., Narayan, A., Tsien, F., Smeets, D. F., Sawyer, J., Fiala, E. S., Sohn, O.S., and Ehrlich, M. 2000. DNA hypomethylation and unusual chromosome instability in cell lines from ICF syndrome patients. *Cytogenet. Cell Genet.* **89**: 121-128.

Tsumura, A., Hayakawa, T., Kumaki, Y., Takebayashi, S., Sakaue, M., Matsuoka, C., Shimotohno, K., Ishikawa, F., Li, E., Ueda, H. R., Nakayama, J., and Okano, M. 2006. Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes Cells* **11**: 805-814.

Ueda, Y., Okano, M., Williams, C., Chen, T., Georgopoulos, K., and Li, E. 2006. Roles for Dnmt3b in mammalian development: a mouse model for the ICF syndrome. *Development* **133**:1183-1192.

Weemaes, C. M., van Tol, M. J., Wang, J., van Ostaijen-Ten Dam, M. M., van Eggermond, M. C., Thijssen, P. E., Aytekin, C., Brunetti-Pierri, N., van der Burg, M., Graham Davies, E., Ferster, A., Furthner, D., Gimelli, G., Gennery, A., Kloeckener-Gruissem, B., Meyn, S., Powell, C., Reisli, I., Schuetz, C., Schulz, A., Shugar, A., van den Elsen, P. J., and van der Maarel, S. M. 2013. Heterogeneous clinical presentation in ICF syndrome: correlation with underlying gene defects. *Eur. J. Hum. Genet.* **21**: 1219-1225.

Xu, G. L., Bestor, T. H., Bourc'his, D., Hsieh, C. L., Tommerup, N., Bugge, M. , Hulten, M., Qu, X., Russo, J. J., and Viegas-Péquignot, E. 1999. Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* **402**: 187-191.

Figure 1

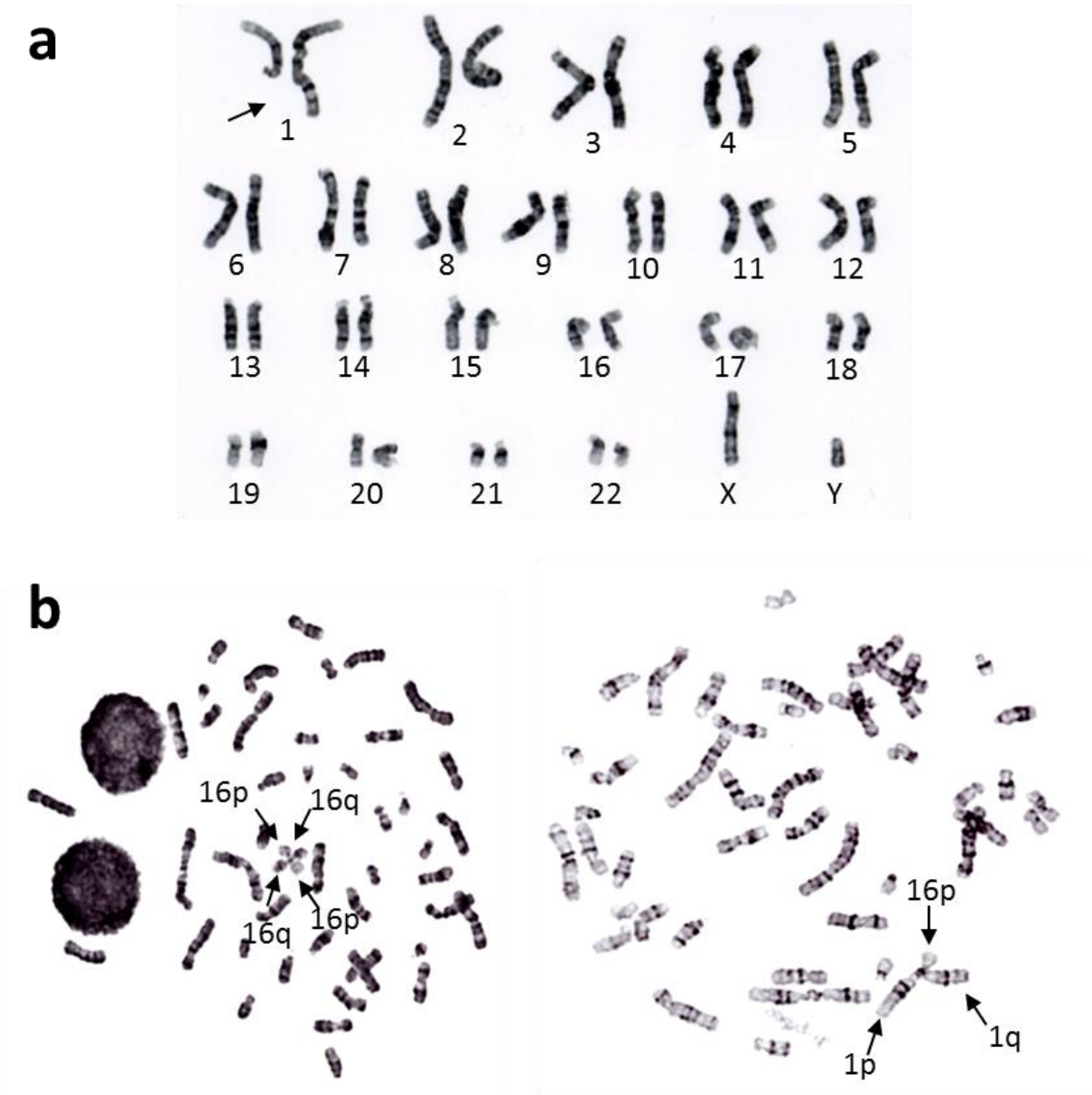

a
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
X
Y
b
16p
16q
16q
16p
16p
1p
1q

Figure 1 Chromosome abnormalities observed in the two Japanese ICF2 patients, P6 and P7. (**a**) Giemza-stained mitotic chromosomes of peripheral blood leukocytes from P6. About 7% of the cells lacked the long arm of chromosome 1 (arrow). (**b**) Representative mitotic chromosome spreads of P7. Multiradial configurations with multiple p and q arms derived from chromosomes 1 and 16 were observed (arrows).

Figure 2

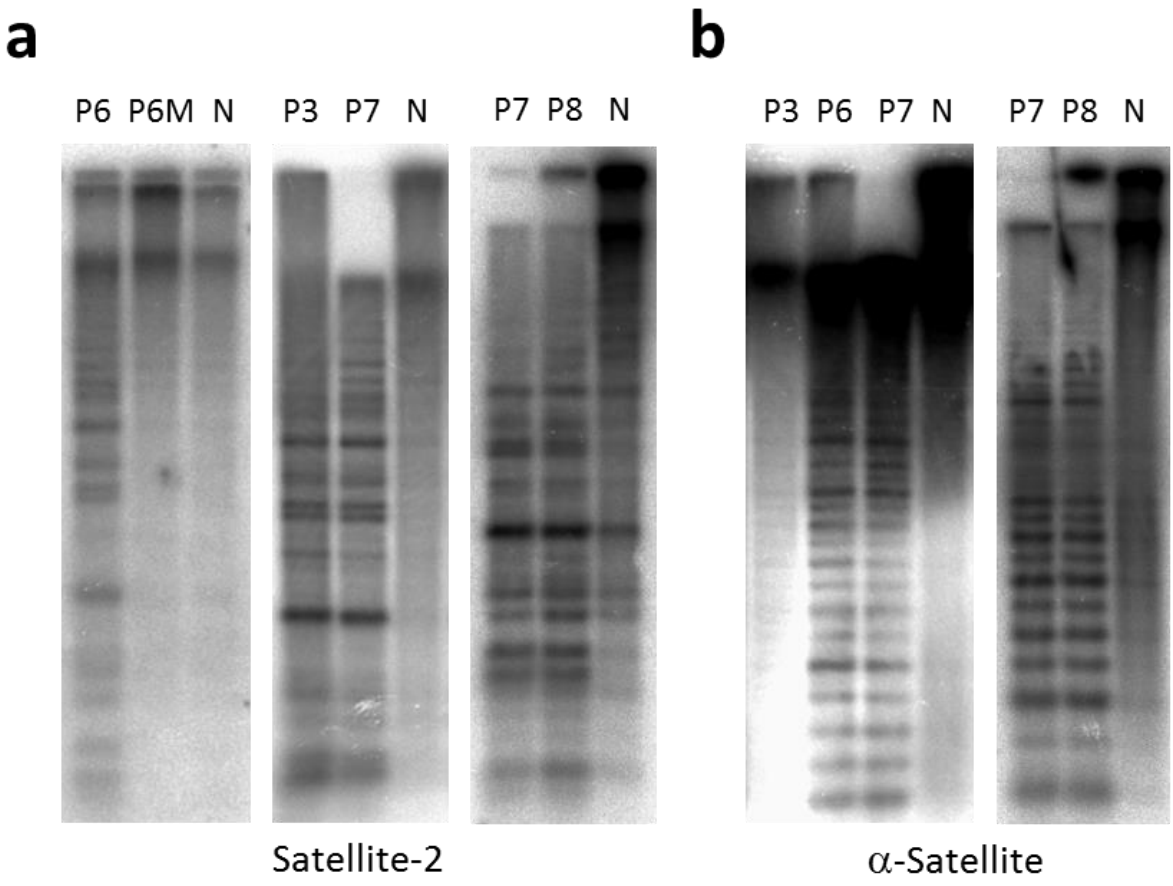

a
b
P6 P6M N
P3 P7 N
P7 P8 N
P3 P6 P7 N
P7 P8 N
Satellite-2
α-Satellite

Figure 2 DNA methylation statuses of the satellite-2 and α-satellite repeats in peripheral blood leukocytes from the three ICF2 patients. (**a**) Southern blot analyses with the satellite-2 probe. Genomic DNAs from P3 (previously diagnosed as ICF1), P6, P7, P8, the mother of P6 (P6M), and a normal control (N) were digested with a methylation-sensitive restriction enzyme *Bst*BI and separated on a 1% agarose gel before blotting. (**b**) Southern blot analyses with the α-satellite probe. Genomic DNAs from P3, P6, P7, P8, and a normal control (N) was digested with a methylation-sensitive restriction enzyme *Hha*I and separated by 1% agarose gel electrophoresis.

Figure 3

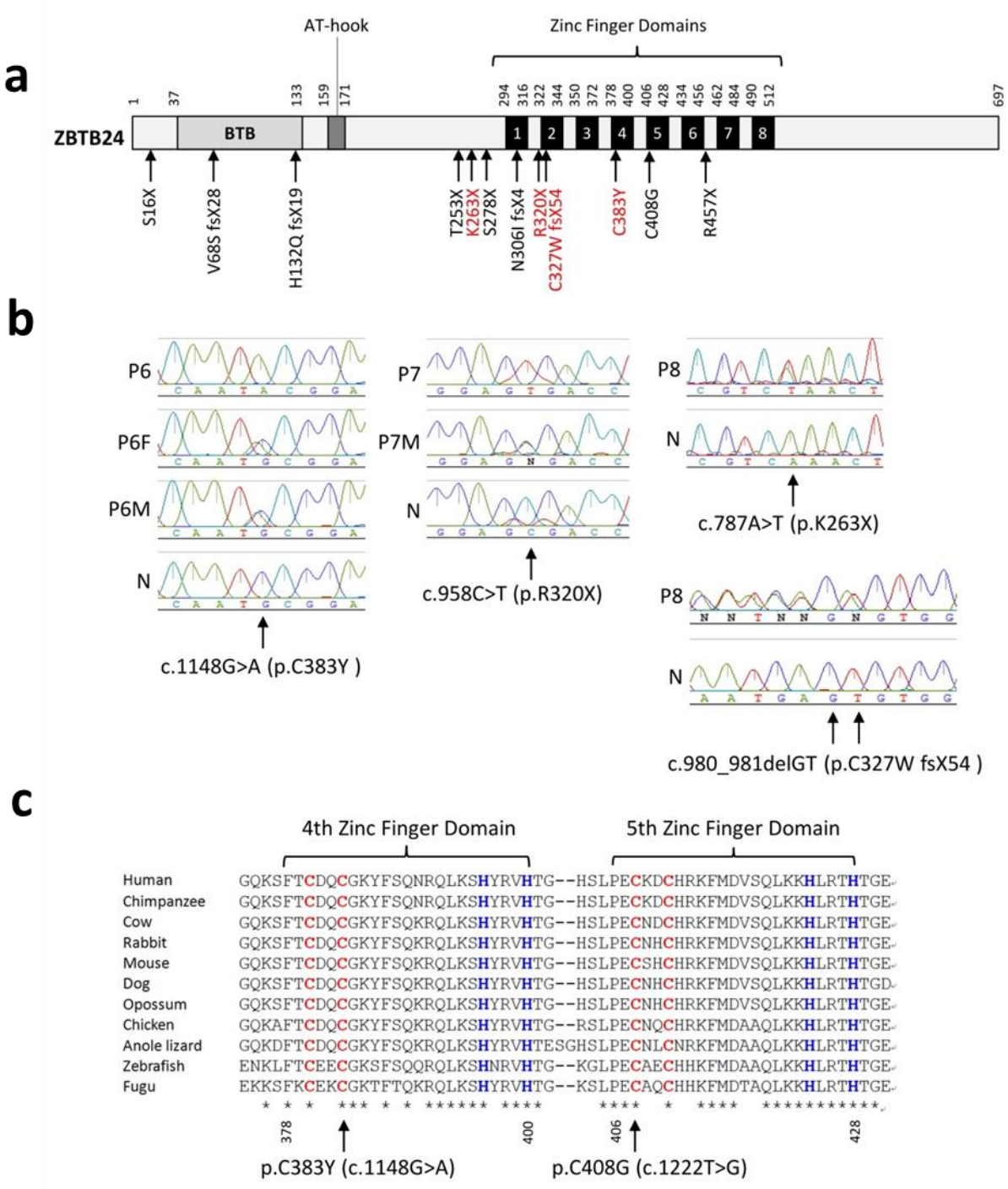

a
AT-hook
Zinc Finger Domains
ZBTB24
BTB
S16X
V68S fsX28
H132Q fsX19
T253X
K263X
S278X
N306I fsX4
R320X
C327W fsX54
C383Y
C408G
R457X
b
P6
P6F
P6M
N
c.1148G>A (p.C383Y)
P7
P7M
N
c.958C>T (p.R320X)
P8
N
c.787A>T (p.K263X)
P8
N
c.980_981delGT (p.C327W fsX54)
c
4th Zinc Finger Domain
5th Zinc Finger Domain
Human
Chimpanzee
Cow
Rabbit
Mouse
Dog
Opossum
Chicken
Anole lizard
Zebrafish
Fugu
p.C383Y (c.1148G>A)
p.C408G (c.1222T>G)

Figure 3 Mutation analysis of the *ZBTB24* gene in ICF2 patients. (**a**) A schematic representation of the ZBTB24 protein structure and locations of all reported mutations including those identified in this study (red) (Table 3). (**b**) Sequencing electrophoregrams of the regions containing the mutations. P6F is the father of P6 and P7M is the mother of P7. Other symbols are as in Figure 2. (**c**) Amino acid sequence alignments of the 4th and 5th zinc finger domains of the vertebrate ZBTB24 proteins (from 374G to 431E of the human ZBTB24 protein). The two amino acid substitutions, C383Y and C408G, altered the highly conserved cysteine in the C2H2 motifs. Asterisks indicate the positions of amino acids conserved across vertebrates. Two cysteins (red) and two histidines (blue) are highlighted.

Figure 4

a

```
Human   RICGRRKRPGGPEARCKDCGKVFKYNHFLAIHQRSHTGERPFKCNECGKGFAQKHSLQVH
Mouse   RLCGRKKRSSGPEARCKDCDRVFKYSHFLAIHQRRHTGERPFKCNECGKGFAQKHSLQVH
        * *** **  *********  **** ******** *************************

Human   TRMHTGERPYTCTVCSKALTTKHSLLEHMSLHSGQKSFTCDQCGKYFSQNRQLKSHYRVH
Mouse   TRMHTGERPYTCTVCGKALTTKHSLLEHMSLHSGQKSFTCDQCGKYFSQKRQLKSHYRVH
        *************** ********************************** **********

Human   TGHSLPECKDCHRKFMDVSQLKKHLRTHTGEKPFTCEICGKSFTAKSSLQTHIRIHRGEK
Mouse   TGHSLPECSHCHRKFMDVSQLKKHLRTHTGEKPFTCEICGKSFTAKSSLQTHIRIHRGEK
        ********  **************************************************

Human   PYSCGICGKSFSDSSAKRRHCILHTGKKPFSCPECNLQFARLDNLKAHLKIHSKEKHASD
Mouse   PYSCSICGKCFSDSSAKRRHCILHTGKKPFSCPECGLQFARLDNLKAHLKIHSKEKHTAD
        **** **** ************************* ********************  *
```

b

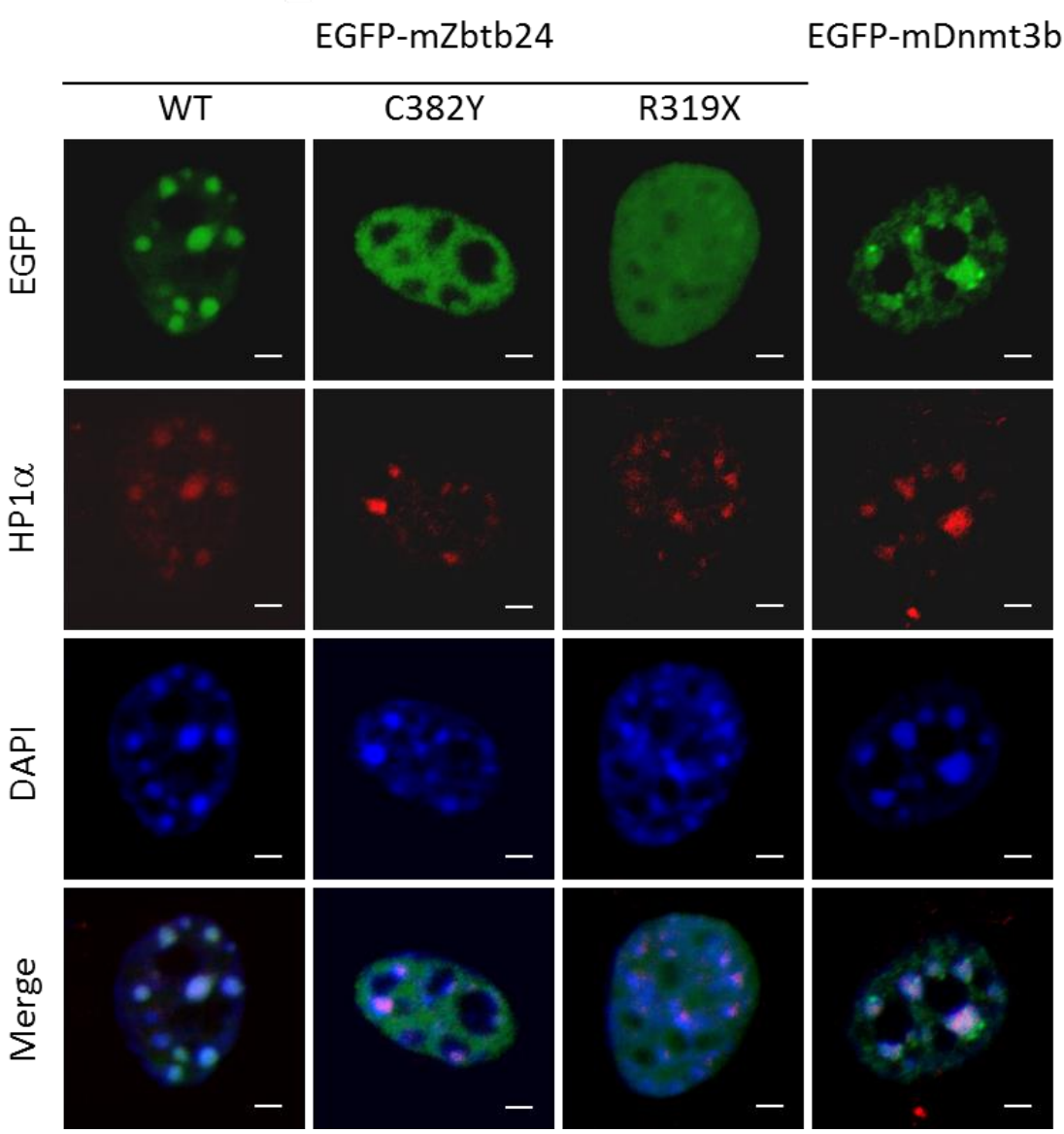


Figure 4 Subnuclear localization of mouse Zbtb24 fusion proteins. (**a**) Alignment of human and mouse zinc finger domains of ZBTB24 proteins (from 281R to 520D of the human ZBTB24 protein, and 280R to 519D of the mouse Zbtb24 protein). Asterisks indicate the positions of amino acids conserved between human and mouse. Cysteins (red) and histidines (blue) in eight C2H2 motifs are highlighted. An arrow indicates the position altered in human C383Y and mouse C382Y mutations, and an arrowhead indicates the arginine residue altered in human R320X and mouse R319X mutations. (**b**) Confocal microscopic images of NIH3T3 nuclei overexpressing EGFP-fused mouse Zbtb24 without (WT) or with ICF2 mutations (C382Y and R319X) (green). EGFP-Dnmt3b was also analyzed (green). Mouse C382Y is equivalent to human C383Y, and mouse R319X is equivalent to human C320X. HP1α (red) was used as a heterochromatin marker, and DAPI (blue) was used for DNA staining. Scale bar: 2 μm.

Figure 5

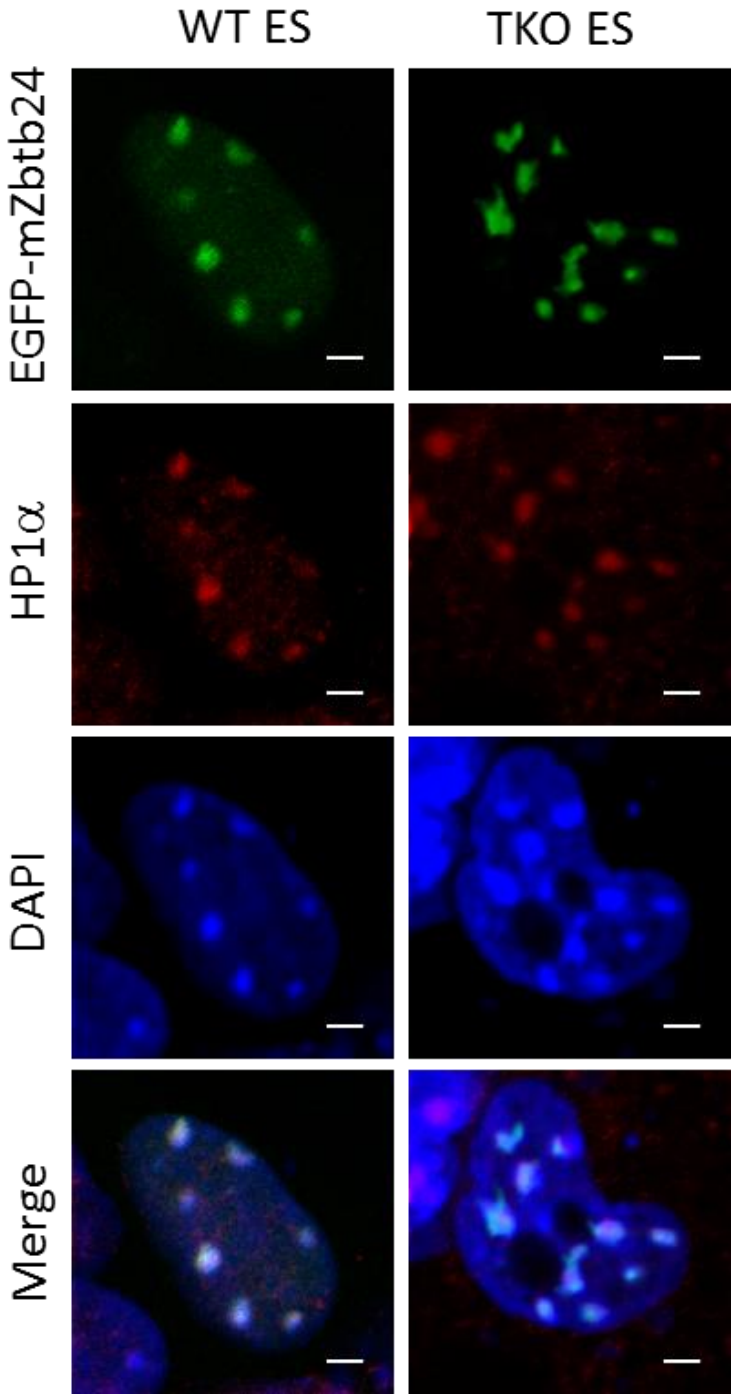


Figure 5 Subnuclear localization of mouse Zbtb24 fusion protein in DNA methylation deficient cells. Confocal microscopic images of WT ES (J1) and Dnmt TKO ES nuclei overexpressing EGFP-fused mouse wildtype Zbtb24 (EGFP-mZbtb24). HP1α (red) was used as a heterochromatin marker, and DAPI (blue) was used for DNA staining. Scale bar: 2 μm.

Figure 6

a

α-satellite (human, centromeric)

```
aattcccagtaacctccttgtgttgtgtgttcaactcacagagttgaacttttattta
gacagagcagattggaaacactctttttgtggaatttgcaagtggagatttcaggcgctt
tgaggttaaaggtagaaaagggaatatcttcgtataaaaactagacagaat
```

Satellite 2 (human, pericentromeric)

```
attccattagatgatgacccctttcatttccattcaatgagg
```

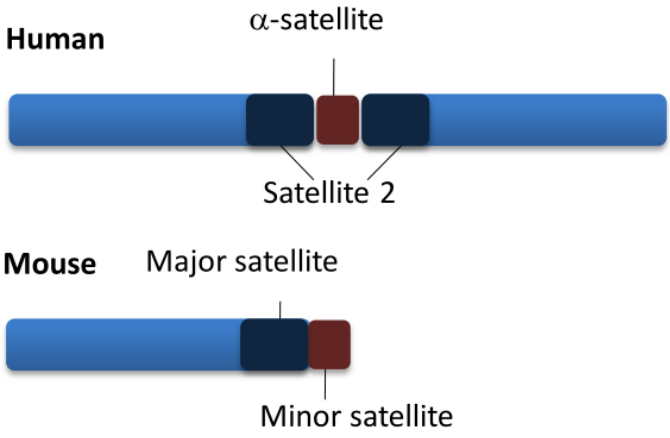
Minor satellite (mouse, centromeric)

```
ggaaaatgataaaaacctacactgtagaacatattagatgagtgagttacactgaaaaac
acattcgttggaaaccggcattgtagaacagtgtatatcaatgagttacaattagaaaca
t
```

Major satellite (mouse, pericentromeric)

```
tatggcgaggaaaaatcaaaaggtggaatatttagaaatgtccactgtaggacttgaaac
atggcaagaaaactgaaaatcatggaaaatgagaaacatgcccttgacgacttgaaaatg
acaaaatcactaaaaaacgtgaaaaatgagaaatgcacactgaaggacctggaatatggc
gagaaaactgaaaatcacggaaaatgagaaatacacactttaggacatgaa
```

b



c

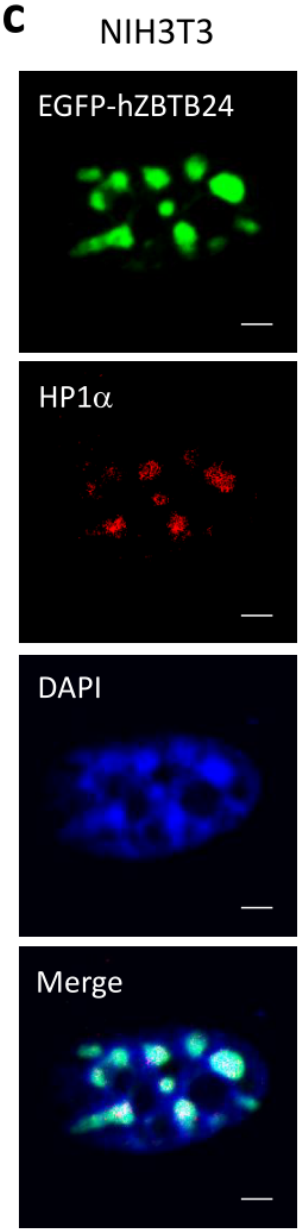
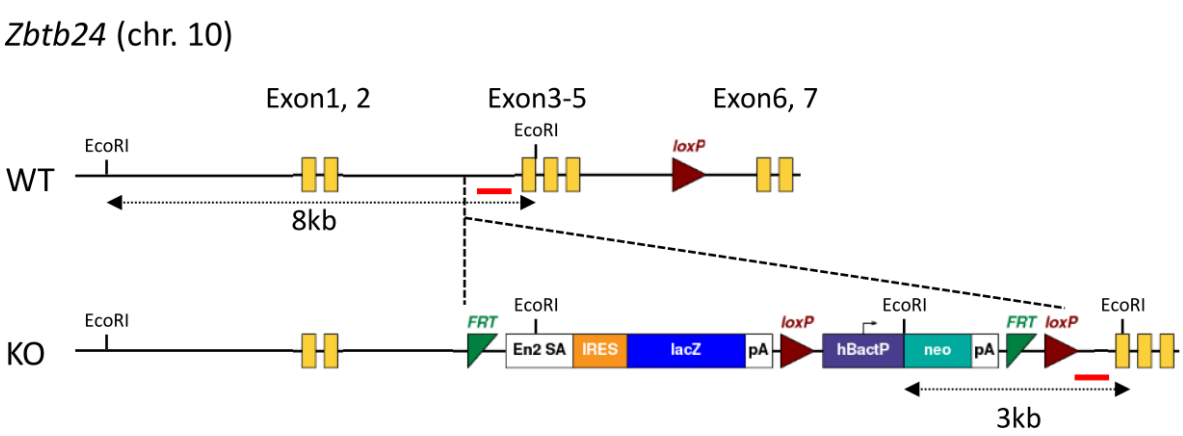


Figure 6 Subnuclear localization of human ZBTB24 fusion protein in mouse cells. (**a**) Satellite repeat sequences of human and mouse. A representative sequence of each satellite repeat unit is shown; α-satellite (Accession No. M27771), satellite 2 (Accession No. X03458), minor satellite (Accession No. X14462), and major satellite (Accession No. M32541). (**b**) Schematic representation of human and mouse chromosomes. (**c**) Confocal microscopic images of NIH3T3 nuclei overexpressing EGFP-fused human wildtype ZBTB24 (EGFP-hZBTB24). HP1α (red) was used as a heterochromatin marker, and DAPI (blue) was used for DNA staining. Scale bar: 2 μm.

Figure 7

a


Zbtb24 (chr. 10)
Exon1, 2
Exon3-5
Exon6, 7
EcoRI
EcoRI
loxP
WT
8kb
EcoRI
EcoRI
FRT
EcoRI
loxP
EcoRI
FRT
loxP
EcoRI
KO
En2 SA
IRES
lacZ
pA
hBactP
neo
pA
3kb

b

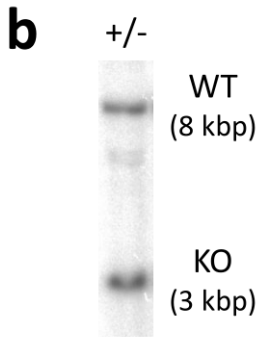

+/-
WT
(8 kbp)
KO
(3 kbp)

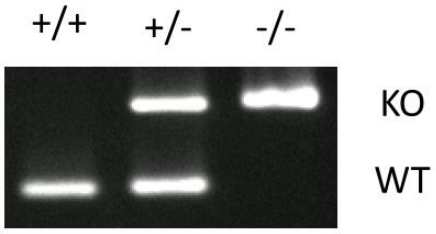
Figure 7 Structure and detection of the *Zbtb24* knockout allele. (**a**) Schematic representation of the genomic organization of the wildtype (WT) allele and knockout (KO) allele (*Zbtb24*$^{tm1a(EUCOMM)Hmgu}$). Yellow boxes indicate exons and red bars indicate the location of the probe for Southern blot analysis. FRT, FLP recombinase target; En2 SA, Engrailed2 splice acceptor; IRES, internal ribosome entry site; pA, polyadenylation signal; hBactP, human β-actin promoter; neo, neomycin resistance cassette. (**b**) Southern blot analysis of heterozygous ES cells carrying a wildtype allele (8 kbp) and a knockout allele (3 kbp). Genomic DNA was digested with *Eco*RI and hybridized with the probe shown in (**a**).

Figure 8

a

Stage	+/+	+/−	−/−	Total
E9.5	5 (33%)	8 (53%)	2 (13%)	15
E10.5	13 (31%)	21 (50%)	8 (19%)	42
E11.5	2 (33%)	4 (66%)	0 (0%)	6
E12.5	1 (9%)	10 (91%)	0 (0%)	11
E13.5	0 (0%)	6 (100%)	0 (0%)	6
E16.5	3 (20%)	12 (80%)	0 (0%)	15
Neonate	14 (29%)	34 (71%)	0 (0%)	48

b



c

Figure 8 Phenotype of *Zbtb24* knockout embryos. (**a**) Numbers and percentages of progenies obtained from *Zbtb24* heterozygous crosses. (**b**) Representative PCR genotyping results for wildtype, heterozygous, and homozygous E10.5 embryos. (**c**) Representative external appearance of E10.5 *Zbtb24*$^{+/-}$ and *Zbtb24*$^{-/-}$ embryos. Scale bar: 1.0 mm.

Figure 9

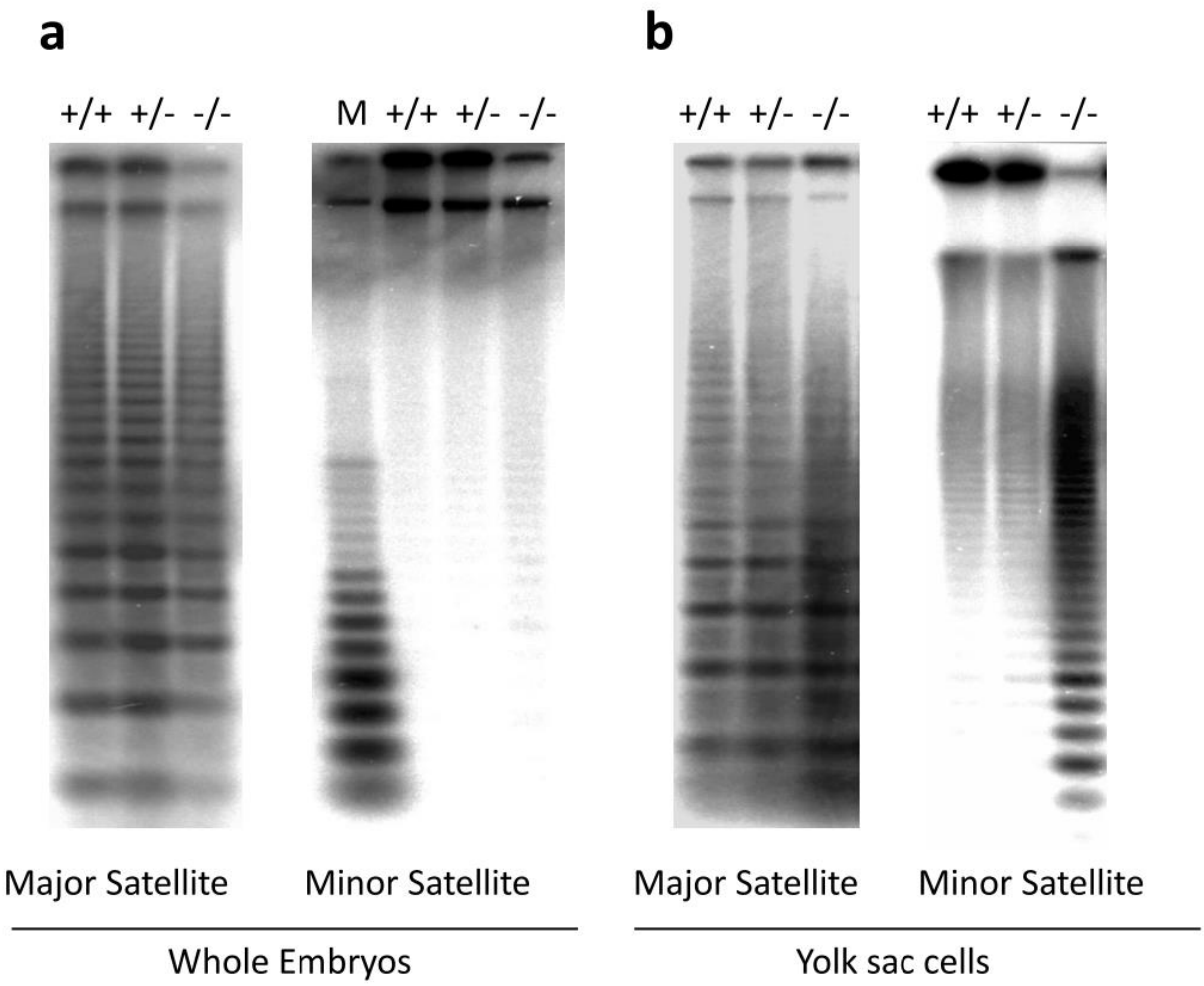

a
b
+/+ +/- -/-
M +/+ +/- -/-
+/+ +/- -/-
+/+ +/- -/-
Major Satellite
Minor Satellite
Major Satellite
Minor Satellite
Whole Embryos
Yolk sac cells

Figure 9 DNA methylation statuses of the major and minor satellite repeats in *Zbtb24* knockout cells. (**a**) DNA methylation analysis of major and minor satellite repeats in E10.5 *Zbtb24*$^{+/+}$, *Zbtb24*$^{+/-}$, and *Zbtb24*$^{-/-}$ whole-embryos. (**b**) DNA methylation analysis of major and minor satellite repeats in E10.5 *Zbtb24*$^{+/+}$, *Zbtb24*$^{+/-}$, and *Zbtb24*$^{-/-}$ yolk sac cells. M indicates methylation-insensitive *Msp*I digest.

Table1 Primer sequences for genomic PCR of *DNMT3B* and *ZBTB24* exons

Primer name	*Primer sequence*
ZBTB24-exon2-1F	5'-TTATCATCACTACTCCCAATTCTGC-3'
ZBTB24-exon2-1R	5'-AACTTTGTGGTTAAAGGAGACAGTG-3'
ZBTB24-exon2-2F	5'-GTCTTGGGATCATAGTTCTCATCTT-3'
ZBTB24-exon2-2R	5'-CTGCTGGAATTTATCTACACAGGTT-3'
ZBTB24-exon2-3F	5'-AGCCTTTACCAGGTCATAGACTTTT-3'
ZBTB24-exon2-3R	5'-ACCTTTTAGACAGTGCTTTCTCCTT-3'
ZBTB24-exon3-F	5'-TATACCATGCTTCTGGGGCTCT-3'
ZBTB24-exon3-R	5'-GTTGGAAATGATACTAGGCTGAGAA-3'
ZBTB24-exon4-F	5'-TAGATTGAAGGCTTCTGGGTATGTA-3'
ZBTB24-exon4-R	5'-ATACCCTATGTGAACGATATGTGCT-3'
ZBTB24-exon5-F	5'-TTATCTTCTTACACAACTCCCATCC-3'
ZBTB24-exon5-R	5'-GTGCTTTAACACAAATCACTGAAGA-3'
ZBTB24-exon6-F	5'-AAAGCCAAAGTAATAAACTGCACAC-3'
ZBTB24-exon6-R	5'-CTGTTTATATTGTGAAACCGAAGTG-3'
ZBTB24-exon7-1F	5'-GAGGGCATTATAAACATCAGTTAGC-3'
ZBTB24-exon7-1R	5'-CCGATTCTGTACATAACATCAATTTC-3'
ZBTB24-exon7-2F	5'-CTCCTGTTGAGCTGAAAGAATTAAA-3'
ZBTB24-exon7-2R	5'-TTCCCCCTAAATAATCAGGAAGTAG-3'
DNMT3B-exon2F	5'-CTAAATGGCATTGTTTGAAGGGG-3'
DNMT3B-exon2R	5'-GAATTGTGGTGGAGGTTGTCAG-3'
DNMT3B-exon3F	5'-AGCAAATCCCTGTGGCTGAC-3'
DNMT3B-exon3R	5'-TAGCTCGGCAACCCTCCATA-3'
DNMT3B-exon4F	5'-CATCTGTCGTGGTTGGTCTG-3'
DNMT3B-exon4R	5'-CCTCCAACAATAGTCCAAGAGC-3'
DNMT3B-exon5F	5'-CCAGTTGTCCTGAAGCTGGCTAC-3'
DNMT3B-exon5R	5'-CTTACGGGAAGCTTACCCAGGAG-3'
DNMT3B-exon6F	5'-GGAAAAGTTTCTTCAGCGGTCTC-3'
DNMT3B-exon6R	5'-AAAGAAAGGTCCGTGGTGCAT-3'
DNMT3B-exon7F	5'-TTATGGCAAGCTTTTTGTTGC-3'
DNMT3B-exon7R	5'-AGTTTTGTCCTTCCTCCCAGA-3'

Primer name	*Primer sequence*
DNMT3B-exon8F	5'-TCCCCAGGAATGGTCTCTTG-3'
DNMT3B-exon8R	5'-TCCGGTGGCCATCATGTTAT-3'
DNMT3B-exon9F	5'-CCTTCCTCTGTCACATACCACA-3'
DNMT3B-exon9R	5'-CCTTCGTATGCCTGCTCCTA-3'
DNMT3B-exon10F	5'-ATGCCACCTGCTAACATTTTG-3'
DNMT3B-exon10R	5'-ACCTGGAGGGTATTTGTGAGG-3'
DNMT3B-exon11-12F	5'-CAGTGTGGACCCCCTGTCAT-3'
DNMT3B-exon11-12R	5'-GGGCACAATTGGCAACTCAT-3'
DNMT3B-exon13F	5'-TGGCCTGAGGTGCAAAGAGT-3'
DNMT3B-exon13R	5'-CAAGGCAGCTGGGATGAGTG-3'
DNMT3B-exon14F	5'-GCAGGTCACAGCCAGTTGTC-3'
DNMT3B-exon14R	5'-CAACCACAGCCACCACAAAATAT-3'
DNMT3B-exon15F	5'-TCAGGGAAGCCCGTACTGCACAG-3'
DNMT3B-exon15R	5'-CGCCAGGTCTCTCAGCCATTGAG-3'
DNMT3B-exon16F	5'-CCAAATGCAGGCTCGTGTGGTACA-3'
DNMT3B-exon16R	5'-GTGACAAAGGGAGACTCCATCTCG-3'
DNMT3B-exon17F	5'-GCTGGATTGAACGCATGTGA-3'
DNMT3B-exon17R	5'-TGGTCACCAACTCCACTTCTGA-3'
DNMT3B-exon18F	5'-AAGTGGGTCCAGCTCTCTTTCC-3'
DNMT3B-exon18R	5'-TTACTATGGAGCGTAGGCAGCAC-3'
DNMT3B-exon19F	5'-GCCTGTGTCCCTGCTGTCAT-3'
DNMT3B-exon19R	5'-CCAGACCAAGAACGGGAAAGT-3'
DNMT3B-exon20F	5'-GGATCCGTGCCTCATCCATA-3'
DNMT3B-exon20R	5'-ACAGAGATCAACCCGCACGA-3'
DNMT3B-exon21F	5'-TCATGCCAGGATCATTTTCA-3'
DNMT3B-exon21R	5'-TTTGAGGATGGGACAGAACAA-3'
DNMT3B-exon22F	5'-CGTGATCTCGTTCATGGTCA-3'
DNMT3B-exon22R	5'-TACTCCTCGCCCCAGTAACA-3'
DNMT3B-exon23F	5'-CCTGGCTGGTTGAGGCTGTCAAC-3'
DNMT3B-exon23R	5'-GAGCACCGTGTTAGGCTGCTCCA-3'

Table 2 Immunological data of the three ICF2 patients.

	Japanese Patients			*Cape Verdean Patient*	
Immunological Data	*P6*	*P7*	*Normal Range*	*P8*	*Normal Range*
IgG (mg/dl)	1088	108	570-1700	815	570-1700
IgG1 (mg/dl)	457	50.8	320-748	ND[a]	320-748
IgG2 (mg/dl)	41	34	208-754	ND	208-754
IgG3 (mg/dl)	420	38.1	6.6-88.3	ND	6.6-88.3
IgG4 (mg/dl)	6	16.7	4.8-108	ND	4.8-108
IgA (mg/dl)	29	69	110-410	128	110-410
IgM (mg/dl)	16	16	33-190	18	33-190
IgE (IU/ml)	<25	<5.0	<170	ND	<170
T cell	ND	99%	66-89%	ND	1000-3900 cells/mm^3
T $CD3^+$	90%	97%	58-84%	2624 cells/mm^3	1000-3900 cells/mm^3
T $CD4^+$	65%	25%	25-54%	1120 cells/mm^3	560-2700 cells/mm^3
T $CD8^+$	22%	76%	23-56%	1536 cells/mm^3	330-1400 cells/mm^3
$CD4^+/CD8^+$	3	0.32	0.4-2.3	ND	0.4-2.3
B cell	ND	1%	4-13%	352 cells/mm^3	270-860 cells/mm^3
B $CD20^+$	13%	ND	7-30%	ND	70-660 cells/mm^3
B $CD27^+$	0.34%	ND	>15%	2%	>15%
PHA^+ (cpm)	23200	3150	20500-56800	ND	20500-56800
Control (cpm)	120	103	127-456	ND	127-456
NK cell activity (%)	5	8	18-40	ND	18-40

[a]ND, not determined

Table 3 Genetic data of ICF2 patients

Patient ID	*Origin*	*Consanguineous marriage*	*Gender*	*Birth year*	*Status*	*Nucleotide change*	*Amino-acid change*	*Reference*
P6	Japanese	No	Male	2003	Died at age 7	c.1148G>A	p.C383Y	Present study
P7	Japanese	Yes	Male	1971	Died at age 41	c.958C>T	p.R320X	Present study
P8	Cape Verdean	No	Female	1996	16 years old (2013)	c.787A>T, c.980_981delGT	p.K263X, p.C327W fsX54	Present study
P1[a]	Turkish	Yes	Male	2006	4 years old (2011)	c.917delA	p.N306I fsX4	de Greef *et al.*
P2[a]	Scottish	Yes	Female	1987	Died at age 13	c.47C>G	p.S16X	de Greef *et al.*
P3[a]	Turkish	Yes	Female	1983	Died at age 11	c.958C>T	p.R320X	de Greef *et al.*
P5[a]	Turkish	Yes	Male	1997	13 years old (2011)	c.501dup	p.V168S fsX28	de Greef *et al.*
P6[ad]	German	No	Male	1998	Died at age 4	c.833C>G, c.1222T>G	p.S278X, p.C408G	de Greef *et al.*
P7[ad]	German	No	Female	2001	9 years old (2011)	c.833C>G, c.1222T>G	p.S278X, p.C408G	de Greef *et al.*
P10[a]	Italian	No	Male	1981	Adult (2011)	c.1369C>T	p.R457X	de Greef *et al.*
P1[bd]	Lebanese	Yes	Male	1997	NA[e]	c.396_397delTA	p.H132Q fsX19	Chouery *et al.*
P2[bd]	Lebanese	Yes	Male	1998	NA[e]	c.396_397delTA	p.H132Q fsX19	Chouery *et al.*
P3[bd]	Lebanese	Yes	Male	2003	NA[e]	c.396_397delTA	p.H132Q fsX19	Chouery *et al.*
No ID	Moroccan	No	Male	2003	8 years old (2012)	c.1222T>G	p.C408G	Cerbone *et al.*
P40[c]	Turkish	Yes	Male	2000	NA[e]	c.759C>G	p.T253X	Weemaes *et al.*
P55[c]	Turkish	No	Male	2010	NA[e]	c.958C>T	p.R320X	Weemaes *et al.*

[a]IDs in de Greef *et al.* [b]IDs in Chouery *et al.* [c]IDs in Weemaes *et al.* [d]These patients are siblings. [e]NA, not available.