ARTICLE

Correct mRNA Processing at a Mutant TT Splice Donor in *FANCC* Ameliorates the Clinical Phenotype in Patients and Is Enhanced by Delivery of Suppressor U1 snRNAs

Linda Hartmann,[1] Kornelia Neveling,[2] Stephanie Borkens,[1] Hildegard Schneider,[3] Marcel Freund,[3] Elke Grassman,[4] Stephan Theiss,[5] Angela Wawer,[6] Stefan Burdach,[6] Arleen D. Auerbach,[7] Detlev Schindler,[2] Helmut Hanenberg,[3,8,9,*] and Heiner Schaal[1,9,*]

The U1 small nuclear RNA (U1 snRNA) as a component of the major U2-dependent spliceosome recognizes 5′ splice sites (5′ss) containing GT as the canonical dinucleotide in the intronic positions +1 and +2. The c.165+1G>T germline mutation in the 5′ss of exon 2 of the Fanconi anemia C (*FANCC*) gene commonly predicted to prevent correct splicing was identified in nine FA patients from three pedigrees. RT-PCR analysis of the endogenous *FANCC* mRNA splicing pattern of patient-derived fibroblasts revealed aberrant mRNA processing, but surprisingly also correct splicing at the TT dinucleotide, albeit with lower efficiency. This consequently resulted in low levels of correctly spliced transcript and minute levels of normal posttranslationally processed FANCD2 protein, indicating that this naturally occurring TT splicing might contribute to the milder clinical manifestations of the disease in these patients. Functional analysis of this *FANCC* 5′ss within splicing reporters revealed that both the noncanonical TT dinucleotide and the genomic context of *FANCC* were required for the residual correct splicing at this mutant 5′ss. Finally, use of lentiviral vectors as a delivery system to introduce expression cassettes for TT-adapted U1 snRNAs into primary *FANCC* patient fibroblasts allowed the correction of the DNA-damage-induced G2 cell-cycle arrest in these cells, thus representing an alternative transcript-targeting approach for genetic therapy of inherited splice-site mutations.

Introduction

Fanconi anemia (FA) is an autosomal or X-linked recessive inherited DNA instability disorder that is caused by germline mutations in at least thirteen genes (*FANCA* [MIM 607139], *FANCB* [MIM 300515], *FANCC* [MIM 227645], *FANCD1* [MIM 600185], *FANCD2* [MIM 227646], *FANCE* [MIM 600901], *FANCF* [MIM 603467], *FANCG* [MIM 602956], *FANCI* [MIM 609053], *FANCJ* [MIM 609054], *FANCL* [MIM 608111], *FANCM* [MIM 609644], and *FANCN* [MIM 610355]) involved in DNA-damage response.[1–3] The heterogenous clinical phenotype of this disease is characterized by variable congenital abnormalities, progressive bone marrow failure, and a predisposition to malignancies, primarily acute myelogenous leukemia (AML [MIM 601626]) and solid tumors.[4–8] Cells from FA patients exhibit a distinctive cellular phenotype of hypersensitivity to DNA crosslinking agents such as mitomycin C (MMC) and diepoxybutane (DEB), a phenotype that can be assessed as increased chromosomal breakage in metaphase analysis and as G2 cell-cycle arrest by flow cytometry.[9–11]

Defects in the FA pathway can be grouped into three distinct categories based on the biochemical characteristics of the cells in response to DNA damage.[1–3] The first group includes the classical FA genes *FANCA*, *FANCB*, *FANCC*, *FANCE*, *FANCF*, *FANCG*, and *FANCL*, whose protein products physically interact in a multiprotein core complex including FANCM. The main function of this FA core complex with E3 ubiquitin ligase activity appears to be the posttranslational activations of FANCD2 and FANCI by monoubiquitination of specific lysine residues.[2,12] Phosphorylation of FANCD2 and FANCI and several other core-complex members is also an essential regulatory step in the FA pathway.[13,14] The second group is formed by FANCD2 and FANCI, which form dimers and in their function depend on each other.[14–18] For FANCD2, we established that despite the fact that all live-born patients have at least one hypomorphic allele, the overall clinical characteristics of biallelic germline defects are more severe compared to the classical FA core genes.[19] The third category, in which FANCD2 and FANCI are normally ubiquitinated, is formed by the *late* FA genes *BRCA2*(*FANCD1*)[20] and the gene encoding its stabilizer *PALB2*(*FANCN*),[21] by the gene encoding the BRCA1-associated helicase BRIP1 (FANCJ),[22] and probably also by the atypical *FANCM*.[23] Defects in *FANCD1* and *FANCN* are associated with the most severe form of FA due to an extremely high incidence of malignancies in these children before 7 years of

[1]Institute of Virology, Heinrich-Heine-University, D-40225 Düsseldorf, Germany; [2]Institute of Human Genetics, Biocenter, University of Würzburg, D-97074 Würzburg, Germany; [3]Department of Pediatric Hematology, Oncology and Clinical Immunology, Children's Hospital, Heinrich-Heine-University, D-40225 Düsseldorf, Germany; [4]Translational Trials Development and Support Laboratory, Experimental Hematology and Cancer Biology, Cincinnati Children's Hospital, Cincinnati, OH 45229, USA; [5]Result GmbH, Friedenstr. 39, D-40219 Düsseldorf, Germany; [6]Children's Hospital Schwabing, Technische Universität München, D-80804 München, Germany; [7]International Fanconi Anemia Registry, Human Genetics and Hematology, The Rockefeller University, New York, NY 10021, USA; [8]Department of Pediatrics, Wells Center for Pediatric Research, Riley Hospital for Children, Indiana University School of Medicine, Indianapolis, IN 46202, USA
[9]These authors contributed equally to this work
*Correspondence: hanenberg@uni-duesseldorf.de (H.H.), schaal@uni-duesseldorf.de (H.S.)
DOI 10.1016/j.ajhg.2010.08.016.

age.[21,24,25] Little is known yet about the phenotypical consequences of biallelic defects in *FANCJ*,[22] *FANCM*,[23] and the recently identified FA-like *RAD51C* (MIM 602774).[26]

Additional factors have been identified that considerably influence the clinical phenotype in FA patients with germline mutations in the same gene. Well established is the impact of certain mutations in *FANCC* as the first FA gene identified on the clinical manifestations and the natural course of the disease.[4] In the North American International FA Registry (IFAR), the biallelic c.456+4A>T mutation (formerly IVS4+4A>T), almost exclusively found in the Ashkenazi-Jewish population, and the mutation c.1642C>T (p.R548X) in *FANCC* are strongly correlated with increased numbers of severe congenital malformation, earlier onsets of marrow failure, and overall poorer survivals compared to the patients with the Northern European exon 2 (c.67delC [p.Q13X]) *FANCC* and patients with *FANCA* or *FANCG* mutations.[4,27] Molecular analysis demonstrated that this milder phenotype of the exon 2 mutation was due to the presence of an N-terminally truncated FANCC protein with residual activity, TRP50, from an internal translation initiation site (methionine 55) in exon 2.[28] This truncated FANCC was not present in *FANCC* c.456+4A>T homozygous cells.[28] Surprisingly, Futaki et al. subsequently showed that the c.456+4A>T mutation in Japanese patients is not associated with a severe clinical phenotype, demonstrating that the ethnic background can be a major modifier of the clinical disease characteristics.[29] The clinical manifestations of FA are also influenced by mutations that result in minute levels of normal protein[19,21,24,25,27,30–33] or in mutant proteins with residual functions.[28,34,35] Polymorphisms in detoxifier genes may also play a role in the course of the disease.[36,37]

The mutation spectrum of germline mutations in FA includes single-base-pair substitutions within splice sites, either in the splice donor, also called 5′ splice site (5′ss), or in the splice acceptor (3′ss).[19,21,25,27,32,33] The canonical GT dinucleotide within the 5′ss of the major U2-type introns is found in more than 99% of all annotated human 5′ss.[38,39] In 0.7% of human 5′ss, GC is utilized as well.[38,39] Both the GT-AG and GC-AG splice sites at exon-intron boundaries are both processed by the U2-dependent spliceosome including the U1, U2, U4/6, and U5 small nuclear ribonucleoprotein particles (snRNPs; for review see [40]). In addition, a minor class of introns with AT-AC boundaries coexist (<0.1% of annotated 5′ss) and are excised by the U12-dependent spliceosome composed of the snRNPs U11, U12, U4atac/U6atac, and U5.[41–43] Pathogenic 5′ss mutations affecting the canonical GT dinucleotide at position +1 are by far the most frequently diagnosed pathogenic splice-site mutations, assumed to inevitably cause aberrant splicing.[44,45] Generally, for 5′ss mutations, the extent to which splicing is impaired depends on the weakening of the RNA duplex that forms between the 5′ss and all 11 nucleotides of the 5′-end of the endogenous U1 small nuclear RNA (snRNA).[46,47]

Since the seminal report in 1986,[48] compensatory mutations in U1 snRNA have been known to have the capability to suppress 5′ss mutations; however, to what extent which nucleotide of the 5′ss can be compensated is still unpredictable. Mutation-adapted U1 snRNAs designed to rescue specific pathogenic mutations have been shown to partially restore correct splicing of artificial minigene constructs in established cell lines.[49–51] So far, stable expression of engineered U1 snRNA molecules correcting aberrant splicing of endogenous mRNAs and thereby also the disease phenotype has not been reported for primary human cells.

Current gene therapy approaches for hematopoietic disorders typically use lentiviral vector systems to insert a wild-type cDNA copy of the mutant gene into hematopoietic stem cells as target cells.[52–54] For some genetic defects, pursuing this strategy might not be feasible because of the large size of the wild-type cDNA[20] or the profound toxicity of the overexpressed normal proteins.[19,22] An alternative approach for genetic therapies for mutations affecting splicing is a mutation-specific correction of mRNA processing that, incidentally, would allow the natural transcriptional control of the affected gene to be maintained.

Here, we report the clinical and molecular consequences of the c.165+1G>T splice-donor mutation in intron 2 of *FANCC* in patients from three FA families, a mutation that converts the canonical GT into a TT dinucleotide. Analysis of the endogenous *FANCC* splicing pattern of patient-derived fibroblasts revealed correct mRNA processing at this mutant splice donor resulting in low levels of correctly spliced transcript and residual E3 ligase activity of the partially reconstituted FA core complex. Functional characterization of this 5′ss mutation within minigene constructs carrying *FANCC* exon 2 as the middle exon revealed that both the overall complementarity of this pathogenic 5′ss to U1 snRNA and the genomic context were required for correct splicing at this mutant splice donor. In addition, we demonstrate that stable expression of TT-adapted artificial U1 snRNAs is capable of correcting the phenotype of primary FA patient fibroblasts. This suggests that lentiviral delivery of a splice-donor-mutation-adapted U1 snRNAs might be a viable alternative strategy for genetic therapy of certain FA complementation groups with founder mutations in 5′ss.

Material and Methods

Patients

Index patients from two Arabian families and one mixed Arabian/British couple were diagnosed as affected by FA on the basis of moderate clinical features (see Table 1) in combination with the hypersensitivity of their primary skin fibroblasts to mitomycin C (MMC)[10] or diepoxybutane (DEB),[9] respectively (data not shown). One family with five affected children was treated at the Children's Hospital Schwabing, Technical University of Munich, Germany. Two families were treated in North America or the United Kingdom and enrolled in the International Fanconi Anemia Registry (IFAR)[4] of the Rockefeller University, New York.

Table 1. Clinical Characteristics of the Nine FA-C Patients

Family/ Affected	Sex	Paternal Mutation	Maternal Mutation	Congenital Abnormalities	Age at BMF	Diagnosis to SCT	SCT at Age	Reason for SCT	Last Follow-Up
526/1	M	c.165+1G>T	c.165+1G>T	Café-au-lait spots	5.5	10.5	16	BMF	Alive at 20 years, 3.5 years after SCT
526/2	F	c.165+1G>T	c.165+1G>T	Café-au-lait spots	8	1	9	BMF	Alive at 29 years, 20 years after SCT
526/3	F	c.165+1G>T	c.165+1G>T	Café-au-lait spots	13	2	19	BMF	Alive at 19.5 years, 0.5 years after SCT
526/4	F	c.165+1G>T	c.165+1G>T	Café-au-lait spots	6	-	-		Alive at 11 years of age
526/5	F	c.165+1G>T	c.165+1G>T	Café-au-lait spots	7.5	-	-		Alive at 8 years of age
640/1	M	c.165+1G>T	c.1-250del	Café-au-lait spots, malrotated kidney, ureter duplication, microphallus	5.5	-	-		Died at 13.5 years of age
640/2	M	c.165+1G>T	c.1-250del	Café-au-lait spots, microphallus	7	8.5	15.5	BMF	Died at 16 years of age
1159/1	M	c.165+1G>T	c.165+1G>T	Café-au-lait spots	11.5	1	12.5	BMF	Died at 12.5 years of age
1159/2	F	c.165+1G>T	c.165+1G>T	none	none	-			Alive at 4.5 years of age

Shown are the clinical characteristics of the nine FA-C patients from three pedigrees numbered 526, 640, and 1159, respectively. The gender, the paternal and maternal *FANCC* mutations, the café-au-laits spots and the major congenital abnormalities, the age at the onset of bone marrow failure, the time interval from diagnosis until stem cell transplantation, the age at and the indication for transplantation, and the last follow-up are shown.

All families signed informed consent approved by the local ethics committees.

Cells and *FANCC* Sequencing

Fibroblasts and genomic DNA samples were derived by standard methods.[55,56] Complementation studies with oncoretroviral vectors containing normal *FANCA*, *FANCC*, *FANCE*, *FANCF*, and *FANCG* cDNAs were performed at the Department of Pediatric Hematology, Oncology, and Clinical Immunology, Children's Hospital, Heinrich-Heine-University, Düsseldorf, Germany, or at the Translational Trials Development and Support Laboratory, Experimental Hematology and Cancer Biology, Cincinnati Children's Hospital, similarly as described.[11,57,58] Primary fibroblasts were immortalized by transduction with a lentiviral vector expressing the SV40 large T antigen (H.H., unpublished data). *FANCC* sequencing was performed as described previously.[59] The maternally inherited genomic deletion leading to exon 2 and 3 skipping in the family 640 was detected on the transcriptional level by use of the *FANCC* exon 1 (#3244) and exon 4 (#3245) or exon 5 (#2911) primers, respectively, as described below under RT-PCR Assay. The updated nomenclature for *FANCC* mutations is based on the RefSeq number NM_000136.2.

Plasmids

FANCC exon 2 including 137 nucleotides of the upstream intron and 22 nucleotides of the downstream intron was amplified from genomic DNA by PCR in intron 1 and intron 2 listed in Table S1A (available online) and inserted via EcoRI-XhoI into the three-exon, two-intron splicing reporter construct described previously.[60,61] Mutant derivatives were generated by PCR mutagenesis with Pwo-Polymerase (Roche) and oligonucleotides carrying mutations to generate mutated constructs (Table S1C). For a second splicing reporter construct, the subgenomic region of *FANCC* including 676 bp of the 3′ part of intron 1, exon 2, intron 2, and exon 3 was amplified from control DNA by PCR with the primers #3715 and #3717 (Table S1C) using Expand High Fidelity DNA Polymerase (Roche). The 5′ splice-site mutation was introduced by PCR mutagenesis with the primers #3718 and #3719 with the QuickChange XL Site-Directed Mutagenesis Kit (Stratagene). All constructs were controlled by sequencing.

For the construction of the mutation-adapted U1 snRNA expression plasmid pUCBU1αTT, the BglII-XhoI fragment of pUCBΔU1[47] was substituted for a PCR fragment amplified with 5′-primer #2040 carrying the BglII site and the specific mutation and the 3′-primer #1131 and pUCBU1 (kindly provided by Alan Weiner, Seattle, WA) as the template. For construction of the pUCBU1-TTcom and pUCBU1mt, the BglII-XhoI fragment pUCBU1αTT was substituted for the PCR fragment amplified with 5′-primer #2809 and 3′-primer #1131 and pUCBU1αTT as the template. Oligonucleotides were synthesized as previously described[62] or were purchased from Metabion (Martinsried, Germany). Sequences of all constructs are available on request.

Cell Culture and Transfection

For the splicing reporter assay, 2.5×10^5 HeLa cells[63] in Dulbecco's modified Eagle's medium (DMEM; Invitrogen) supplemented with 10% fetal bovine serum (Pan Biotech), 2 mM L-glutamine, and 50 U/ml penicillin and streptomycin (both Invitrogen), were seeded per 6-well plate 24 hr before transfection. Cells were transfected with 1 μg of the splice reporter constructs or their mutated derivatives with FuGENE 6 according to the manufacturer's protocol (Roche Molecular Biochemicals). For cotransfection experiments, cells were transfected with 1 μg of pXGH5 encoding human growth hormone[64] and 2 μg of the respective plasmid.

Primary fibroblast strains were established by standard cell-culture procedures and maintained in complete DMEM in high-humidity incubators in an atmosphere of 5% (v/v) CO_2 and 5% (v/v) O_2.[65] For splicing analysis of the endogenous *FANCC* transcript, fibroblasts were seeded in T75 flasks and grown to approximately 80% confluency. For transfection of immortalized fibroblasts, cells were seeded 24 hr before transfection and transfected with 16 μg of the respective plasmid and 8 μg of

pXGH5 by using FuGENE 6 (Roche Molecular Biochemicals). For both assays, RNA was isolated 30 hr after transfection.

RT-PCR Assay

To isolate total RNA, we used Gene Elute Mammalian Total RNA Kit (Sigma-Aldrich). RT-PCR was performed with the SuperScript TM III RT-PCR System with Platinum Taq Polymerase (Invitrogen). For analysis of the splicing pattern, prior to reverse transcription, 4 μg (for the endogenous *FANCC* transcript) or 1 μg of total RNA (for the splicing reporter transcript) was subjected to DNase I digestion with 10 U of DNase I at 70°C for 5 min (Roche), and 2 μl of the DNase I-digested RNA samples were reverse transcribed with SuperScript TM III RT-PCR System with Platinum Taq Polymerase using the 5′ *FANCC* exon 1 primer #3244 and 3′ *FANCC* exon 4 primer #3245 or 3′ *FANCC* exon 5 primer #2911 (Table S1B). As negative control for remaining DNA contamination of each sample, a second assay was performed with Platinum Taq Polymerase (Invitrogen). As a control for transfection efficiency, human growth hormone (hGH) mRNA was detected with the primer pair #1273 and #1274 (Table S1B). Spliced mRNA expressed from the reporter construct was amplified with the primers #481 and #559 (Table S1B). To allow semiquantitative assessment of the spliced endogenous transcripts, PCR analysis was performed within the linear range. PCR products were separated on 6% nondenaturing polyacrylamide gels, stained with ethidium bromide, and visualized with the Lumi-Imager F1 (Roche Molecular Biochemicals).

FANCD2 Immunoblotting

FANCD2 immunoblotting was performed as described[19] with minor modifications: immortalized FA fibroblasts transduced with the different vectors were seeded in T75 flasks and grown to approximately 70% confluence and exposed to 2 mM hydroxyurea (HU, Sigma-Aldrich) for 24 hr.

Lentiviral-Vector Production

HEK 293T and HT1080 cells (CRL-12103; ATCC, Manassas, VA) grown in complete DMEM medium were used to produce and titrate the lentiviral vectors (LVs) carrying the U1 snRNA variants. For the production of LVs, 6×10^6 293 T cells were plated per 10 cm cell-culture dish 24 hr prior to transfection (using polyethylenimine [PEI, Aldrich]) with 6 μg of pCD/NL-BH,[66] 6 μg of an expression plasmid coding for vesicular stomatitis virus G protein,[67] and 6 μg of pCL1NPB-U1. Supernatants were harvested 48 hr after transfection and filtered through a 0.45 μm filter. Functional NEO titers of the LV vectors were determined in HT1080 cells, plated at 3.5×10^4 cells per well in 6-well plates the day before, and then infected with different dilutions of the LV preparations. The next day, cells were washed and incubated for 7 days with fresh medium containing 0.8 mg/ml G418 (Invitrogen). Colonies were fixed with methanol and stained with methylenblue. Titers were calculated as described previously,[55,56] usually obtaining 10^6–10^7 infectious virus particles/ml.

Phenotypic Correction of *FANCC* Fibroblasts

Primary patient-derived *FANCC* fibroblasts were transduced with either pCL1NPB-U1 wt or mutant U1 snRNA derivatives, or as control, with MFCPN and murine stem cell virus (MSCV) at equivalent multiplicities of infection.[11] G418-resistant cells were cultured for 72 hr with or without 33 nM of MMC (Sigma-Aldrich). Cells were harvested by trypsinization and washed with 1% (w/v) bovine serum albumin fraction V (BSA) in PBS (Invitrogen). The cell pellets were resuspended in PBS and fixed overnight in 98% ethanol at −20°C. After centrifugation (600 × g, 4°C), resuspended cell pellets were incubated with 100 μg/ml RNase (Invitrogen) in PBS for 15 min at 37°C. Cell pellets were resuspended in staining buffer containing 0.5% (w/v) BSA and 10 μg/ml propidiumiodid (PI; Sigma) in PBS. DNA histograms were recorded with the flow cytometer FACSCalibur (Becton Dickinson, Heidelberg, Germany). Quantitative assessment was performed with the ModFit software (Verity Software House).

Results

Mild Clinical Manifestations of FA in the Nine FA-C Patients

For complementation-group assignment, primary skin fibroblasts of three index FA patients from two consanguineous families of Arabian ancestry and one mixed Arabian/British couple were transduced with gammaretroviral vectors expressing one of the following cDNAs: *FANCA*, *FANCC*, *FANCE*, *FANCF*, or *FANCG*.[11,57,58] Transduced fibroblast cells were exposed to 33 nM of MMC for 3 days and then harvested for cell-cycle analysis by flow cytometry as described previously.[11,57,58] The cell-cycle distribution of the fibroblasts revealed that overexpression of the *FANCC* cDNA specifically corrected the characteristic DNA crosslinker hypersensitivity of the patients' cells (data not shown). Table 1 shows the clinical features of the nine FA-C patients. Among the four male and five female patients, only the two patients from the family 640 with mixed ethnic background had typical severe congenital malformations as described for FA patients.[27] Calculated for all nine affected individuals, the mean number of 0.45 malformations per patient was similar to that described for the "European" c.67delC mutation in exon 2[27] but different from the c.456+4A>T (IVS4+4A>T) and c.1642C>T (p.R548X) mutations described in Ashkenazi Jewish and North American patients.[4,27]

The c.165+1G>T Splice-Donor Mutation in Primary Cells of FA-C Patients Allowed Correct Splicing albeit at a Reduced Level

Genomic sequencing confirmed that all three index patients carried a base-pair substitution in the 5′ splice site (5′ss) of *FANCC* exon 2, c.165+1G>T, converting the highly conserved GT dinucleotide within the 5′ss of *FANCC* exon 2 to a TT dinucleotide (Figures 1A and 1B). Patients from the pedigrees 526 and 1159 were homozygous for this point mutation, whereas patients from the pedigree 640 carried a maternally inherited genomic deletion (Table 1), leading to the skipping of exons 2 and 3 in the *FANCC* mRNA (data not shown). For analysis of the phenotypic consequence at the RNA level, RT-PCR analysis on mRNA from primary patient fibroblasts from pedigree 526 was performed. In contrast to normal control, we found four distinct splice products contributing to 33%, 27%, 25%, and 15% of the transcripts, respectively (Figure 1C). Direct sequencing of the amplified products

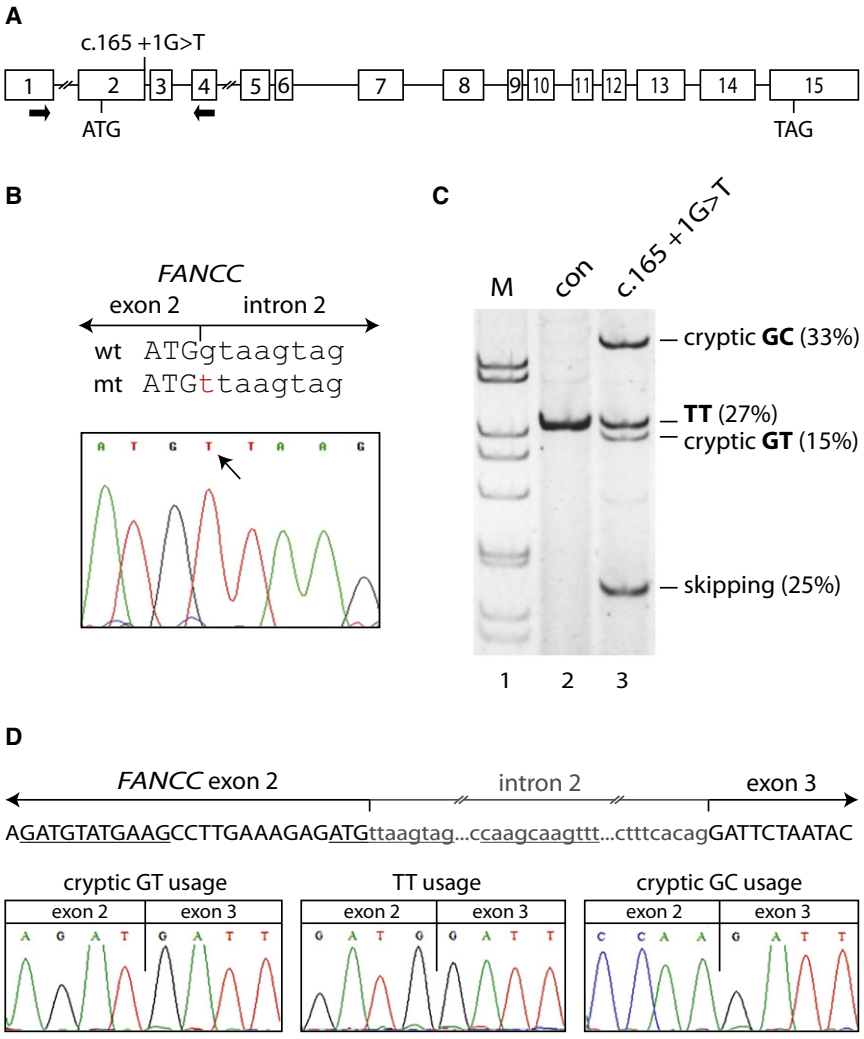

Figure 1. Homozygous c.165+1G>T Splice-Donor Mutation in Primary Cells of the Index FA-C Patient from Pedigree 526 Allowed Correct Splicing at Low Level
(A) Structure of the *FANCC* exons showing the position of the c.165+1G>T mutation and the translational start codon in exon 2 (1st coding exon). The position of RT-PCR primers is depicted.
(B) Sequence result of the genomic DNA carrying the biallelic c.165+1G>T (IVS2+1G>T) 5′ss mutation.
(C) Semiquantitative RT-PCR analysis of the splicing pattern of the *FANCC* mRNA in normal (con) and patient-derived primary fibroblasts (IVS2+1G>T).
(D) Schematic drawing highlighting the positions of the 5′ss used in cells with the biallelic c.165+1G>T mutation and cDNA sequencing results from primary patient-derived fibroblasts.

revealed that the three *FANCC* transcripts of aberrant size either lacked the translational start codon because of skipping of *FANCC* exon 2 (25%, skipping) or encoded mutant open reading frames with premature translation termination (33% and 15%, cryptic GC and GT). Remarkably, the fourth amplified product (27%, TT) was the normal wild-type *FANCC* transcript (Figure 1D). Therefore, the c.165+1G>T splice-donor mutation still enabled normal *FANCC* transcript processing, albeit at lower efficiency compared to the wild-type canonical 5′ss.

Increased Complementarity to U1 snRNA Specifically Reconstituted Splicing at the TT Dinucleotide in a Heterologous Splicing Reporter Construct

For systematic analysis of this unusual pathogenic *FANCC* splice donor, the *FANCC* exon 2 with flanking intronic nucleotides was inserted into a three-exon splicing reporter construct (Figure 2A) that we have established previously.[60] HeLa cells were transfected with plasmids carrying either the wild-type GT or mutant TT *FANCC* exon 2 splice donor and analyzed for their splicing pattern by RT-PCR analysis. Although the intrinsic strength of the wild-type *FANCC* 5′ss is relatively high, because of the high degree of complementarity to the U1 snRNA (Figure 2C), the analytical gel (Figure 2B) demonstrated that the recognition of the wild-type *FANCC* exon 2 was not as effective as expected (lane 2) and that the mutant 5′ss was not recognized at all (lane 3). These splicing patterns indicated a dependency of correct exon 2 recognition on the genomic context in *FANCC* that was missing in this minigene construct. To rank the intrinsic strength of the wild-type *FANCC* exon 2 splice donor among human 5′ss, we analyzed a representative group of 43,464 annotated 5′ss from constitutively spliced human exons with the HBond algorithm.[47,68] In this analysis, all annotated human splice-donor sites had an average HBond score of 15.001 ± 2.591 standard deviation (SD), compared to the HBond score of 18.7 of this *FANCC* 5′ss, which thereby is ranked at the 92.3 percentile of all splice-donor sites in this data set.

Since we and others have previously shown that increasing the complementarity between a 5′ss and the U1 snRNA can improve the recognition of 5′ss and can compensate for the lack of supportive context,[47,68] we additionally changed positions −3 and −2 of the mutant TT 5′ss to nucleotides complementary to the 5′-end of the endogenous U1 snRNA (Figure 2C). As shown in Figure 2B, lane 5, these two additional nucleotide adaptations facilitated inclusion of the *FANCC* exon 2 with the mutant TT splice donor. Direct sequencing of this splice product, however, revealed that splicing in this reporter transcript occurred not only at the TT dinucleotide at the authentic exon-intron border but was also shifted to the GT dinucleotide one position upstream of TT (Figure 2D, TT com −3/−2). The existence of this second transcript is

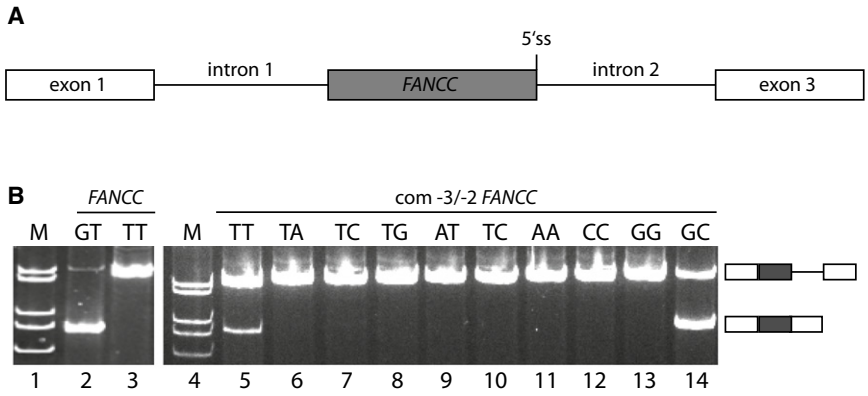

Figure 2. Improved Complementarity to U1 snRNA Reconstituted Splicing at the TT Dinucleotide in a Heterologous Splicing Reporter Construct

(A) Schematic drawing of the 2-intron splicing reporter harboring the *FANCC* exon 2 with its strong 5′ss (HBS 18.7).

(B) RT-PCR analysis of transfected HeLa cells. *FANCC* indicates the splicing reporter constructs that harbor either the wild-type GT 5′ss (lane 2) or the TT 5′ss (lane 3) found in the FA-C patients; com −3/−2 *FANCC* denotes the 5′ss with improved complementarity to U1 snRNA at positions −3 and −2. Dinucleotides that were tested at the +1 and +2 positions in the constructs with improved complementarity at positions −3 and −2 are indicated in the figure (lanes 5 to 14). The spliced products are schematically shown on the right.

(C) Pattern of the HBond formations between the *FANCC* TT 5′ss and U1 snRNA and the improved version of this splice site with enhanced complementarity to the U1 snRNA at positions −2 and −3.

(D) Direct sequencing results of the spliced PCR products. S denotes G or C, K denotes G or T, W denotes A or T, and N denotes any nucleotide.

remarkable, because all available algorithms for splice-donor sites unequivocally predict that the intrinsic strength of the GT dinucleotide at −1 is very weak (e.g., the HBond score is 2.3) as a result of the low complementary to the U1 snRNA in this base-pairing frame.

To further clarify whether splicing at this artificially improved TT 5′ss is simply determined by the overall complementarity to U1 snRNA or also by the position of the GT in the −1 register, we generated reporter constructs that carried different dinucleotides at positions +1 and +2 (Figure 2B). Noteworthy here is that the substitution of T at position +2 maintains the GT within the −1 register yet reduces the overall complementarity to the endogenous U1 snRNA in the original base-pairing frame. In contrast, substitution of the mismatching T at position +1 for A or C does not affect the overall complementarity in the original base-pairing frame but specifically destroys the GT in the −1 register. Thus, if the GT in the −1 register is important for recognition of the mutant TT 5′ss, the TA dinucleotide that specifically increases U1 complementarity in the −1 register should allow more efficient splice-site recognition. As shown in Figure 2B, lanes 5 to 14, and confirmed by sequencing (Figure 2D), splicing in this construct only occurred at the two physiological splice-donor sites GT and GC or if a TT dinucleotide was present at position +1 and +2. These data demonstrated that a mutant TT splice-donor site can be functional in a heterologous context if this site is highly complementary to the U1 snRNA. These results also suggested that the complementarity of the −1 GT register to the U1 snRNA is of less importance, because the TA dinucleotide despite higher complementarity did not allow splicing at this site (Figure 2B, lane 6).

Artificial TT-Adapted U1 snRNA Improved Correct mRNA Processing at the *FANCC* TT 5′ss within the Splicing Reporter

Because the mutant TT 5′ss of *FANCC* exon 2 was recognized only in the splicing reporter after increasing its complementarity to the U1 snRNA, we next asked whether compensatory mutations within the 5′ end of the U1 snRNA would also enable usage of the mutant *FANCC* TT 5′ss. To this end, two artificial U1 snRNAs were constructed (Figure 3A): the U1 snRNA αTT contains a single compensatory mutation complementary to the TT dinucleotide, and the fully complementary (TTcom) U1 snRNA TTcom matches every position of the mutant *FANCC* TT 5′ splice site. Whereas cotransfecting HeLa cells with the wild-type U1 snRNA and the minigene splicing reporter did not alter the splicing pattern of the construct (Figure 3B, lane 2), cotransfection of either U1 snRNA αTT or TTcom

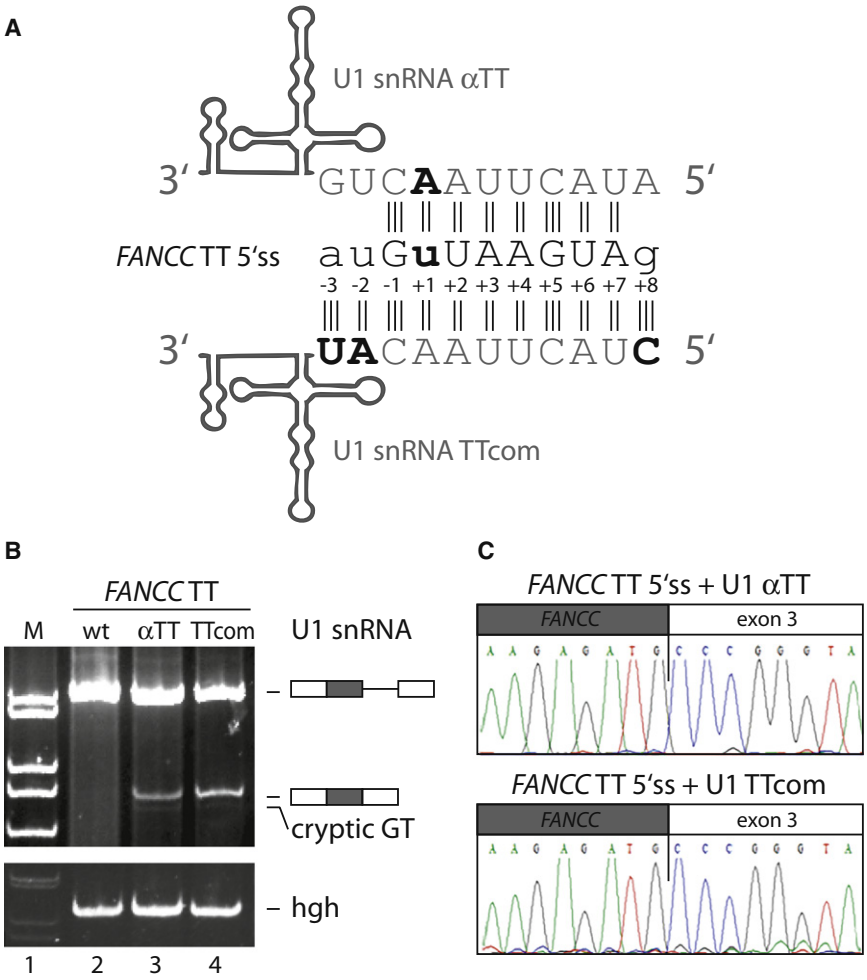

Figure 3. TT-Adapted U1 snRNA Restored Usage of the *FANCC* TT 5′ss within the Minigene Splicing Reporter

(A) Schematic illustration of two specifically TT-adapted U1 snRNAs and sketch of the HBond formation with the *FANCC* TT 5′ ss (positions are numbered). HBonds are indicated by vertical lines. U1 snRNA αTT contains a single compensatory mutation whereas U1 snRNA TTcom was engineered to be able to base-pair with each position of the *FANCC* TT 5′ splice site. The mutant nucleotide at the 5′ end of U1 in each case is shown in bold.

(B) RT-PCR analysis of HeLa cells transfected with the splicing reporter containing the TT 5′ss found in the FA-C patients. U1 wt, U1 αTT, and U1 TTcom indicate cotransfections with the wild-type or TT-adapted U1 snRNA expression plasmid pUCBU1 (lanes 2 to 4). RT-PCR analysis of hGH was performed to monitor transfection efficiency.

(C) Sequence results of the splice junctions.

partially restored recognition (8% and 12%, respectively) of the mutant *FANCC* TT 5′ss (Figure 3B, lanes 3 and 4). Here, sequence analysis of the splice products confirmed that splicing exclusively occurred at the correct exon-intron border (Figure 3C).

Combining the results from the last two experimental settings of fully adapting either the 5′ss to the endogenous U1 snRNA or the U1 snRNA to the mutant splice site was striking: the exclusive use of the noncanonical TT as splice site was not simply determined by the free energy of the RNA duplex formed between the splice donor and the matching U1 snRNA (which was identical in both cases) but was predominantly dependent on the 5′ss sequence itself. Importantly, these results implied that ectopic expression of a TT-adapted U1 snRNA will improve the recognition of the mutant *FANCC* TT 5′ss within the endogenous *FANCC* transcript in patient-derived cells without the risk of activating aberrant splicing at the GT in the −1 register or elsewhere in the genome.

Ectopic Expression of the TT-Adapted U1 snRNAs Specifically Enhanced the Amount of the Endogenous In-Frame Transcript in FA Patient-Derived Fibroblasts

Because the TT-adapted U1 snRNAs improved the usage of the pathogenic *FANCC* TT 5′ss in the reporter system, we immortalized the primary fibroblasts from pedigree 526 with a lentivirus expressing the SV40 large T cDNA and then analyzed whether transfection of the biallelic *FANCC* c.165+1G>T fibroblasts with the TT-adapted U1 snRNA constructs enhanced the levels of endogenous TT-spliced in-frame transcripts. By using RT-PCR primers in the 5′UTR and in exon 4 to distinguish the different endogenous *FANCC* transcripts in the biallelic *FANCC* c.165+1G>T patients' cells (Figure 1A), we showed that transfection of patient-derived fibroblasts with the TT-adapted U1 snRNAs U1 αTT or the U1 TTcom specifically increased the amount of the TT-spliced in-frame transcript from 30% to 56% and 58%, respectively (Figures 4A and 4B). Direct sequencing of both splice products confirmed that they were accurately spliced at the correct exon-intron border (Figure 4C). Concomitantly, the amount of all aberrantly spliced transcripts decreased, indicating that the two mutant TT-adapted U1 snRNAs were capable of improving exon recognition and thereby facilitating production of the correct in-frame transcript. Thus, ectopic expression of the TT-adapted artificial U1 snRNAs significantly increased the usage of the pathogenic TT 5′ss in the patients' fibroblasts.

Phenotypic Correction of *FANCC*-Mutant Fibroblasts by Integrating Lentivirus-Mediated Expression of TT-Adapted U1 snRNA

Permanent suppression of splice-donor mutations in cells still actively dividing requires that the mutation-adapted U1 snRNA integrate into the genome of the mutant cells. Because retroviruses are evolutionary optimized

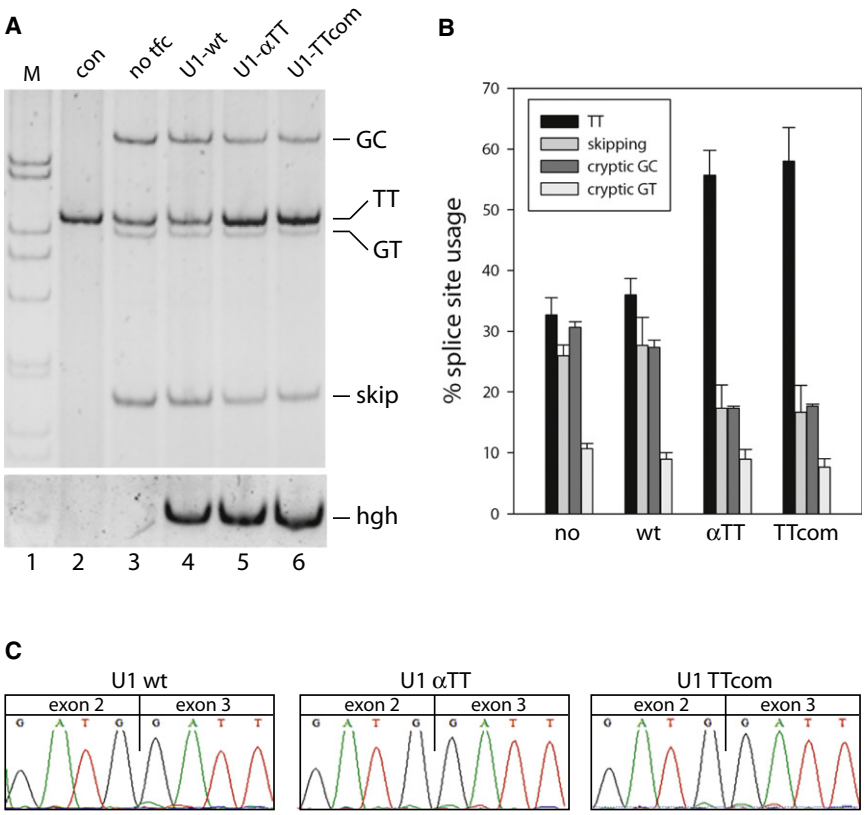

Figure 4. Ectopic Expression of the TT-Adapted U1 snRNAs Specifically Enhanced the Amount of the Endogenous In-Frame Transcript in Fibroblasts from the Index Patient in Pedigree 526

(A) RT-PCR analysis of endogenous transcripts in normal (con) and in patient-derived immortalized fibroblasts. Overexpression of the TT-adapted or wild-type U1 snRNAs in patient-derived fibroblasts is indicated at the top (lanes 4 to 6). RT-PCR analysis of hGH was performed to monitor transfection efficiency. GC corresponds to usage of a cryptic GC splice site downstream of *FANCC* exon 2, and GT denotes the usage of a cryptic GT splice site within *FANCC* exon 2. Skip denotes skipping of *FANCC* exon 2, and TT indicates the usage of the TT splice donor resulting in the correct in-frame transcript.

(B) Quantification of the relative splice-site usage in patient-derived immortalized fibroblasts from three independent transfection experiments (mean ± standard error of the mean).

(C) Sequence results of the splice junctions of the TT spliced transcripts. The transfection of the corresponding U1 snRNA expression plasmid is indicated above.

gene-delivery systems for stably introducing foreign cDNA into the cellular DNA, we transferred the two TT-adapted U1 snRNA expression cassettes into a LV that coexpressed the neomycin phosphotransferase (neoR) cDNA in the opposite orientation (Figure 5A). To determine whether expression for the TT-adapted U1 snRNAs achieved with the LV construct was sufficient for phenotypic correction of the DNA crosslinker hypersensitivity, we transduced primary fibroblasts of the index patient from family 526 with vectors expressing the wild-type or the two mutant U1 snRNAs. As controls, fibroblasts were transduced with a retroviral vector expressing the wild-type *FANCC* cDNA or the corresponding retroviral and lentiviral control vectors, expressing the *neoR* cDNA only. G418-resistant fibroblasts were exposed for 3 days to 33 nM MMC and then analyzed by flow cytometry for their cell-cycle distribution.

Transduction of the patient's fibroblasts carrying the pathogenic c.165 +1G>T mutation on both alleles with a retroviral vector containing the wild-type *FANCC* cDNA (MFCPN) corrected the MMC-induced G2 arrest, whereas cells transduced with the mock vector (MSCV) exhibited a prominent G2 phase arrest typical for FA (Figure 5B). Expression of both TT-adapted U1 snRNAs (αTT, TTcom) in the LV vector significantly improved the MMC-induced cell-cycle arrest whereas stable expression of the wild-type U1 snRNA did not influence the cell-cycle distribution of the primary cells (Figure 5B).

For further analyses, the G418-resistant primary fibroblasts from the index patient from pedigree 526 transduced with the different constructs (Figure 5B) were immortalized with a LV vector that expressed the SV40 large T antigen. RT-PCR analysis of the endogenous *FANCC* transcript in these fibroblasts confirmed the correct usage of the mutant TT splice donor c.165 +1G>T, because *FANCC* exon 2 recognition was clearly improved by lentivirus-mediated expression of both TT-adapted U1 snRNAs (Figure 5C). An indication of a functional FA core complex with normal FANCC as essential component is the monoubiquitination of FANCD2 in response to exposure to DNA interstrand crosslinking agents.[69] Because FANCD2 immunoblot analysis on primary fibroblasts is difficult because of the minute amounts of FANCD2 present, the immortalized fibroblasts from Figure 5C were exposed to 2 mM hydroxyurea for 24 hr and then protein harvested as published.[19,69] Results revealed that 526/1 fibroblasts (see Table 1) with the biallelic *FANCC* c.165+1G>T mutation already had minute levels of monoubiquitinated protein even in the absence of any further genetic modifications of the cells (Figure 5D). This indicated that the low levels of TT-spliced endogenous in-frame transcript encoded residual functional FANCC that was active in the FA core complex. Finally, results showed that transduction of the *FANCC* c.165+1G>T fibroblasts with each of the two TT-adapted synthetic U1 snRNA expression constructs increased the levels of monoubiquitinated FANCD2 corresponding to the mRNA splice pattern shown in Figure 5C. Whereas expression of U1 snRNA TTcom resulted in FANCD2 monoubiquitination levels comparable to the those induced by expression of the wild-type *FANCC*

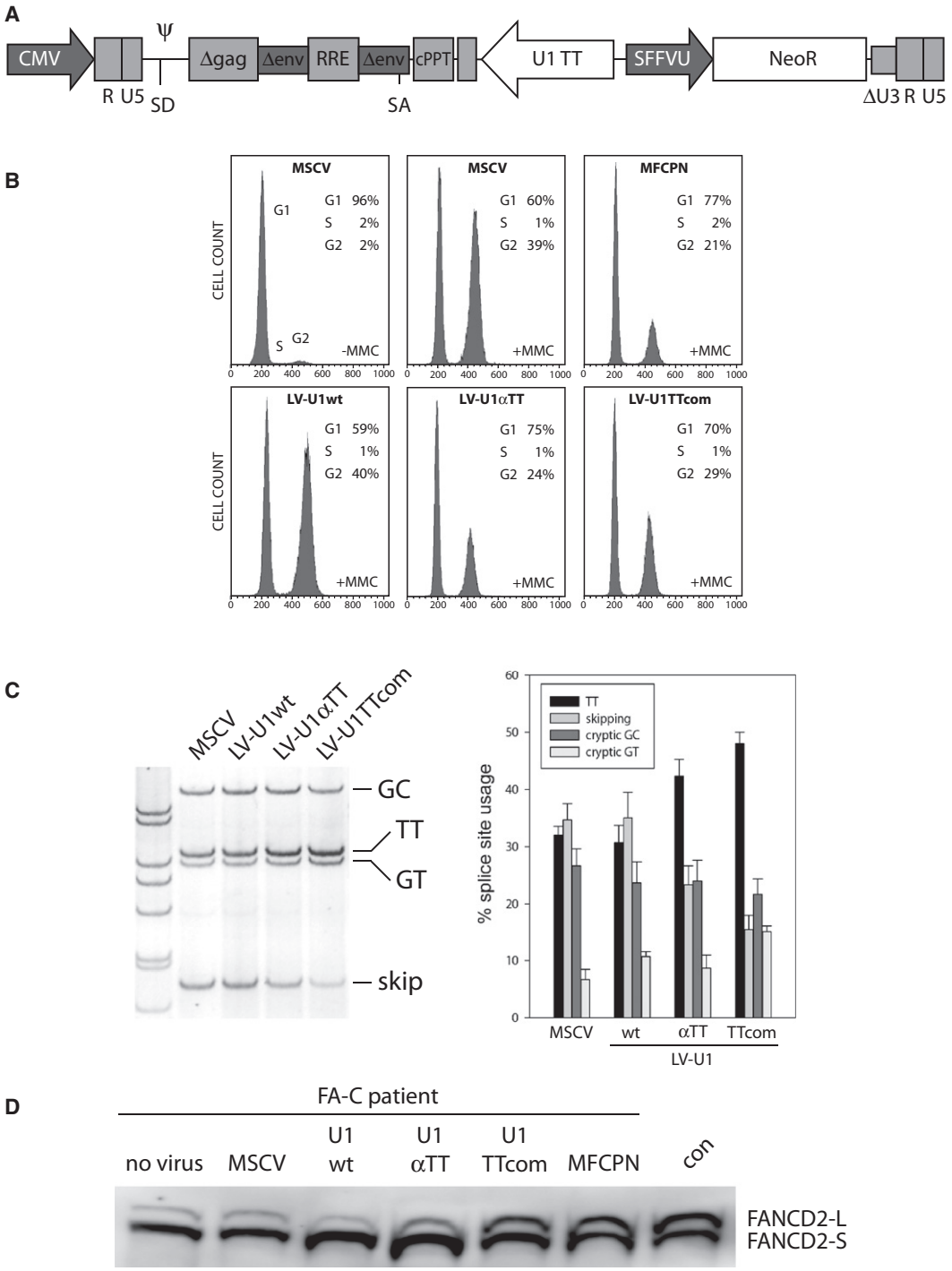

Figure 5. Phenotypic Correction of Primary Biallelic c.165+1G>T *FANCC* Fibroblasts of the Index Patient from Pedigree 526 by Lentivirus-Mediated Expression of TT-Adapted U1 snRNAs

(A) Scheme of the lentiviral vector introducing the U1 snRNA expression cassette into target cells.

(B) Cell-cycle distribution of primary $G418^R$ fibroblasts with the *FANCC* c.165+1G>T mutation transduced with the different retroviral vectors after exposure to 33 nM MMC for 72 hr by flow cytometry. Shown are the histograms of one representatitive analysis and the means of two independent polyclonal transduced primary cell populations. Transduction of the *FANCC*-mutant fibroblasts with a lentiviral vector harboring the TT-adapted U1 snRNAs significantly improved the MMC-induced cell-cycle arrest close to levels achieved by retroviral overexpression of the wild-type *FANCC* cDNA (MFCPN).

(C) RT-PCR analysis of the endogenous *FANCC* transcript in retro- or lentivirus transduced and subsequently immortalized 526 fibroblasts from (B) and quantitative assessment of splice-site usage from three independent RNA preparations (mean ± SEM).

(D) FANCD2 immunoblot from retro- or lentivirus transduced and subsequently immortalized fibroblasts from pedigree 526 after exposure to 2 mM hydroxyurea for 24 hr. The monoubiquitinated and the nonmodified forms of the FANCD2 protein are labeled as D2-L and D2-S, respectively.

cDNA (MFCPN) and present in the normal control (con), expression of the U1 αTT was less effective on both the mRNA and the protein level (Figures 5C and 5D).

These data indicated that lentivirus-mediated stable expression of TT-adapted U1 snRNA molecules can lead to continuous production of sufficient amounts of endogenous functional *FANCC* transcript for restoration of the FA pathway and correction of the cellular FA phenotype of DNA crosslinker hypersensitivity, thus demonstrating the potential of lentivirus-mediated transfer of splice-site-mutation-adapted U1 snRNA genes as curative therapeutic strategy for splice-donor site mutations in FA.

Discussion

The phenotypic correction of primary fibroblasts from three index FA patients by retroviral complementation led to the identification of a biallelic 5′ss mutation c.165+1G>T of the *FANCC* gene in two patients and a monoalleic mutation in the third, confirming that *FANCC* is the gene deficient in these three FA pedigrees. RT-PCR analysis on primary fibroblasts with biallelic mutations showed that the mutant TT 5′ss resulted in aberrant splicing, where the majority of transcripts (73%) were nonfunctional because of exon skipping or the usage of cryptic splice-donor sites. Unexpectedly, RT-PCR analysis also revealed that this mutant TT 5′ss facilitated correct recognition of the normal exon-intron boundaries, leading to reduced amounts of normal transcripts (27%).

In order to perform a systematic analysis of the mutant TT 5′ss in *FANCC*, we constructed a heterologous three-exon, two-intron splicing reporter with *FANCC* exon 2 as the middle exon. In HeLa cells, initial analyses of the splicing pattern revealed that inclusion of the wild-type exon was not exclusive and that the exon with the mutant 5′ss was not recognized at all. This suggested that the *FANCC* exon 2 recognition did not solely occur by exon definition but also depended on additional flanking sequences, both not present in the reporter construct. We therefore substituted exon 2 and downstream reporter sequences with authentic subgenomic *FANCC* sequences from the 3′-end of intron 1 up to exon 3. Transfection of this reporter construct into HeLa cells clearly confirmed that the genomic context of *FANCC* was also required for efficient recognition of the *FANCC* exon 2 5′ss (Figure S1B).

In our reporter system shown in Figure 2, exon recognition efficiency could be restored by additional nucleotide substitutions in positions −3 and −2, thereby improving the base-pairing with the U1 snRNA in the TT register. This improved complementarity allowed us to demonstrate that splicing specifically occurred only when GT, GC, or the mutant TT dinucleotides were present at positions +1 and +2. Here, GT was the most and TT the least efficient for correct mRNA processing. All other analyzed dinucleotides at positions +1 and +2 did not support splicing at this site. Sequencing of the RT-PCR-amplified TT-spliced transcripts from this reporter construct also revealed additional splicing at position −1, which has been recently characterized for the atypical 5′ss and is assumed to precede base-pairing in a shifted register.[70] This additional splicing at −1 cannot simply be explained by the increased complementarity to the U1 snRNA in the −1 register, because the TA dinucleotide in the splicing reporter that would otherwise have specifically increased base-pairing to U1 snRNA at −1 (gcaGTAAagta, HBond score 9.0 versus gcaGTtAagta, HBond score 1.9) did not allow splicing (Figure 2B, lane 6). This suggested that the wild-type U1 snRNA paired with the mutant *FANCC* TT splice donor in the canonical register. As confirmation, in the natural context, however, splicing of the mutant TT splice-donor site exclusively occurred immediately upstream of the TT dinucleotide at the correct exon-intron border, presumably because of additional sequences in the endogenous gene context.

Until now, only two dinucleotides have been thought to be recognized by the U2-dependent spliceosome as 5′ss in mammalian genomes, GT (>99%) and GC (0.5%).[38,39] Exon recognition with splicing at a TT dinucleotide under physiological conditions is not known to occur in mammals at all. There is only one report in the literature on TT splicing from 1987 that analyzed a +1G>T mutation in the 5′ss of the large human β-globin (HBB) intron 2 in in vitro assays.[71] This study in a HBB splicing reporter revealed a new cleavage site at GT in position −1, CAGG/GTGAGTCT>CAG/GTTGAGTCT. Splicing in this construct then could be shifted to TT splicing at +1 upon insertion of a second site mutation (c.494G>A; CAAG/TTGAGTCT) that rendered this artificial *HBB* splice donor with two mutations very similar to the mutant *FANCC* 5′ss described here. Common to both functional TT splice sites is their overall high complementarity to U1 snRNA in the TT register, which is considerably above the average HBond score of all human annotated 5′ss. Calculated with GT instead of TT, the *FANCC* and *HBB* second site mutations have HBond scores of 18.7 and 18.2, respectively. Because only 10% of our data set of annotated human 5′ss have an HBond score of ≥ 18.2 (S.T. and H.S., unpublished data), an overall high complementarity to U1 snRNA in the TT register might be a prerequisite for TT splicing. This high complementarity would also apply for the recently described atypical splice-donor sites,[70] where, according to our findings reported here, splicing should occur at both positions, GT at −1 and TT at +1. However, transcripts at TT in +1 were not described in this report,[70] possibly because of the fact that the RT-PCR bands of the splice products were not directly sequenced but cloned into plasmids prior to sequencing, or because of additional 5′ss sequence and context requirements, respectively, not identified yet. Given that high complementarity at least appears to be one prerequisite for TT splicing and this occurs only in a small percentage of human 5′ss, we caution here that a G>T mutation at +1 does not necessarily abrogate functional splicing in all cases, because these mutant noncanonical TT 5′ss might

be recognized by an endogenous, as-yet-unidentified spliceosomal U1 snRNA.

Recently, expression of a U1 snRNA variant, U1A7, that is complementary to the mutant *FANCC* TT 5′ss was identified in HeLa cells.[72] This U1A7 snRNA, however, did not enable TT splicing, neither in our own analyses (data not shown) nor in the work by Roca and Krainer,[70] most likely because of a nonfunctional snRNA body, thus suggesting that this U1A7 snRNA might be a transcript of a pseudogene.[70] Alternatively, the U1A7 snRNP might be delayed in its biogenesis, and thus its suppression capability might simply not have been detected in the transient transfection assays.

In the FA-C patients with the c.165+1G>T mutation described here, we predict that the low levels of normal mRNA transcript led to low minute levels of normal FANCC as indicated by the residual FANCD2 monoubiquitination seen upon hydroxyurea (HU) challenge. We hypothesize that this residual activity of the FA pathway is the reason for the milder clinical phenotype seen in these patients compared to other *FANCC* mutations. This is in line with findings that in certain FA complementation groups, such as FA-D2 and FA-D1, at least one hypomorphic allele with residual protein activity appears mandatory for the survival of patients with biallelic germline mutations.[6,19,73] Remarkably, the exclusive presence of severe congenital abnormalities in the two patients from the 640 family with monoallelic c.165+1G>T mutations (Table 1) clearly suggests that there might even be an allele-dose effect visible here.

So far, genetic therapies aimed at correcting the underlying deficiency in hematopoietic stem cells utilized integrating retroviral vector systems to introduce a normal cDNA copy of the affected gene into the target cell. In the present study, we showed that understanding the phenotypic consequence of splice-donor mutations at the mRNA level can be instrumental to develop therapeutic strategies to correct an aberrantly processed message. Since the initial report in 1986, compensatory mutations in U1 snRNA have been known to have the capability to suppress 5′ss mutations.[48] The suppressive efficiency of these altered U1 snRNAs, however, depends on the individual mutation and often can only be assessed by functional testing. Using two different cellular systems with the c.165+1G>T mutation, HeLa cells with the splicing reporter carrying the mutant *FANCC* 5′ss and the *FANCC*-mutated cells with the endogenous c.165+1G>T mutation demonstrated that transfection of the cells with the two TT-adapted U1 snRNA expression plasmids facilitated correct exon 2 recognition for both U1 snRNAs, thereby increasing the fraction of *FANCC* exon 2 in-frame spliced pre-mRNA in both approaches. Although correction of pre-mRNA processing in minigene constructs with pathogenic splice sites has been reported by a few groups,[49–51,74] correction of the endogenous transcript and correction of the disease phenotype of primary human cells that are deficient in a cellular transcript has not been reported so far. In a mouse model for spinal muscular atrophy, Meyer et al. elegantly showed in primary murine cells recently that the germline expression of an artificial U7 snRNA, which promoted inclusion of the mutant SMN2 exon 7, can efficiently complement the muscle tissue and significantly extend the limited lifespan of these animals.[74]

Here, we showed in primary cells from a patient with a monogenetic recessive disorder that stable expression of mutation-adapted U1 snRNAs can be utilized to rescue the pathological phenotype of these cells. Using lentivirus-based vectors as delivery systems for the U1 snRNA expression cassette allows stable integration of the U1 snRNA expression cassette into the target cell genome in dividing and nondividing cells,[52] e.g., hematopoietic stem cells and also retina cells as target cells for genetic correction.[75,76] Interestingly, the level of functional restoration of the FA/BRCA pathway in transduced cells differed between the two U1 snRNAs that were specifically adapted for the mutant *FANCC* exon 2 5′ss. The minimally adapted U1 snRNA αTT almost achieved a correction level of cells where the normal *FANCC* cDNA was overexpressed. Although it would appear likely that an increased complementarity of the TT-adapted U1 snRNA to the pathogenic 5′ss will more efficiently generate correct transcripts and also reduce the potential for deleterious off-target effects, surprisingly, the U1 snRNA TTcom with higher complementarity seemed to be less efficient in correcting the cell-cycle arrest in the primary FA cells. A more advanced strategy combining our U1 snRNA-based approach with efforts to support U1 snRNA binding by artificially recruited SR proteins[77] should be further developed to achieve the most efficient correction of a pathogenic 5′ss mutation on the RNA level. Correction of the endogenous transcript would also obviate the inability to deliver large genes and ensure that the natural fine-tuning of the endogenous protein remains intact. Therefore, correction of pathological mRNA processing at mutant splice sites might be an attractive gene-therapy approach for certain FA complementation groups with either very large genes or toxicity of the overexpressed genes such as *BRCA2*(*FANCD1*)[20] or *FANCD2*.[78] This mutation-specific approach might also be feasible in other genetic disorders with deficiencies in other genes such as *ATM* (MIM 607585)[79] and *NF1* (MIM 613113)[80] with a high percentage of 5′ss mutations.

Supplemental Data

Supplemental Data include one figure and one table and can be found with this article online at http://www.cell.com/AJHG/.

Acknowledgments

We would like to thank affected individuals and their families for providing samples for this study. We thank Alan M. Weiner (Seattle, University of Washington) for providing the plasmid pUCBU1 containing the U1 snRNA expression cassette. We are indebted to Imke Meyer (Düsseldorf), Birgit Gottwald (Würzburg), Richard Friedl (Würzburg), and Waltraud Kuss (Düsseldorf) for

expert technical assistance. Research in our laboratories was supported by the Jürgen-Manchot-Stiftung (L.H., H.S.), the Stiftung für AIDS-Forschung, Düsseldorf (H.S.), the Cancer Society Northrhine-Westphalia (H.S.), the Forschungskommission of the Heinrich-Heine-University, Düsseldorf (M.F., H.H., H.S.), the Deutsche Forschungsgemeinschaft (DFG SPP1230) (H.H.), the National Institutes of Health (NIH-NHLBI/P01 HL053586-14 and NIH-NCI/R01 CA138287-01) (H.H.), the Bundesministerium für Bildung und Forschung (BMBF) network for Inherited Bone Marrow Failure Syndromes (H.H., D.S., H.S.), the Fanconi Anemia Research Fund (H.H.), the Deutsche Fanconi-Anämie-Hilfe and the Aktionskreis Fanconi-Anämie funds (D.S., H.H.), the Schroeder-Kurth-Fund (D.S.), intramural funding from the Cincinnati Children's Hospital Research Foundation (E.G.), the BMBF Foamyviral Network for Genetic Therapy of FA (H.H.), and the NIH Clinical and Translational Science Award (CTSA) grant 5UL1RR024143-04 to The Rockefeller University Hospital (A.D.A.).

Received: June 18, 2010
Revised: August 26, 2010
Accepted: August 31, 2010
Published online: September 23, 2010

Web Resources

The URLs for data presented herein are as follows:

BIC database, http://research.nhgri.nih.gov/projects/bic/
Ensembl Genome Browser, http://www.ensembl.org/
FA mutation database, http://www.rockefeller.edu/fanconi/mutate/
FA Registry (IFAR), http://www.rockefeller.edu/labheads/auerbach/clinresearch.php
HBond Score, http://www.uni-duesseldorf.de/rna/
Human Genome Variation Society, http://hgvs.org/
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/

References

1. Wang, W. (2007). Emergence of a DNA-damage response network consisting of Fanconi anaemia and BRCA proteins. Nat. Rev. Genet. *8*, 735–748.
2. de Winter, J.P., and Joenje, H. (2009). The genetic and molecular basis of Fanconi anemia. Mutat. Res. *668*, 11–19.
3. Moldovan, G.L., and D'Andrea, A.D. (2009). How the fanconi anemia pathway guards the genome. Annu. Rev. Genet. *43*, 223–249.
4. Kutler, D.I., Singh, B., Satagopan, J., Batish, S.D., Berwick, M., Giampietro, P.F., Hanenberg, H., and Auerbach, A.D. (2003). A 20-year perspective on the International Fanconi Anemia Registry (IFAR). Blood *101*, 1249–1256.
5. Green, A.M., and Kupfer, G.M. (2009). Fanconi anemia. Hematol. Oncol. Clin. North Am. *23*, 193–214.
6. Neveling, K., Endt, D., Hoehn, H., and Schindler, D. (2009). Genotype-phenotype correlations in Fanconi anemia. Mutat. Res. *668*, 73–91.
7. Pinto, F.O., Leblanc, T., Chamousset, D., Le Roux, G., Brethon, B., Cassinat, B., Larghero, J., de Villartay, J.P., Stoppa-Lyonnet, D., Baruchel, A., et al. (2009). Diagnosis of Fanconi anemia in patients with bone marrow failure. Haematologica *94*, 487–495.
8. Auerbach, A.D. (2009). Fanconi anemia and its diagnosis. Mutat. Res. *668*, 4–10.
9. Auerbach, A.D. (1993). Fanconi anemia diagnosis and the diepoxybutane (DEB) test. Exp. Hematol. *21*, 731–733.
10. Seyschab, H., Friedl, R., Sun, Y., Schindler, D., Hoehn, H., Hentze, S., and Schroeder-Kurth, T. (1995). Comparative evaluation of diepoxybutane sensitivity and cell cycle blockage in the diagnosis of Fanconi anemia. Blood *85*, 2233–2237.
11. Schindler, D., Friedl, R., Gavvovidis, I., Kalb, R., Neveling, K., Linka, Y., Hanenberg, H., Kubbies, M., and Hoehn, H. (2007). Application of cell cycle testing in Fanconi anemia. In Fanconi Anemia: A Paradigmatic Disease for the Understanding of Cancer and Aging, D. Schindler and H. Hoehn, eds. (Basel: Karger), pp. 110–130.
12. Alpi, A.F., and Patel, K.J. (2009). Monoubiquitylation in the Fanconi anemia DNA damage response pathway. DNA Repair (Amst.) *8*, 430–435.
13. Ishiai, M., Kitao, H., Smogorzewska, A., Tomida, J., Kinomura, A., Uchida, E., Saberi, A., Kinoshita, E., Kinoshita-Kikuta, E., Koike, T., et al. (2008). FANCI phosphorylation functions as a molecular switch to turn on the Fanconi anemia pathway. Nat. Struct. Mol. Biol. *15*, 1138–1146.
14. Knipscheer, P., Räschle, M., Smogorzewska, A., Enoiu, M., Ho, T.V., Schärer, O.D., Elledge, S.J., and Walter, J.C. (2009). The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. Science *326*, 1698–1701.
15. Sims, A.E., Spiteri, E., Sims, R.J., 3rd, Arita, A.G., Lach, F.P., Landers, T., Wurm, M., Freund, M., Neveling, K., Hanenberg, H., et al. (2007). FANCI is a second monoubiquitinated member of the Fanconi anemia pathway. Nat. Struct. Mol. Biol. *14*, 564–567.
16. Dorsman, J.C., Levitus, M., Rockx, D., Rooimans, M.A., Oostra, A.B., Haitjema, A., Bakker, S.T., Steltenpool, J., Schuler, D., Mohan, S., et al. (2007). Identification of the Fanconi anemia complementation group I gene, FANCI. Cell. Oncol. *29*, 211–218.
17. Smogorzewska, A., Matsuoka, S., Vinciguerra, P., McDonald, E.R., 3rd, Hurov, K.E., Luo, J., Ballif, B.A., Gygi, S.P., Hofmann, K., D'Andrea, A.D., and Elledge, S.J. (2007). Identification of the FANCI protein, a monoubiquitinated FANCD2 paralog required for DNA repair. Cell *129*, 289–301.
18. Yuan, F., El Hokayem, J., Zhou, W., and Zhang, Y. (2009). FANCI protein binds to DNA and interacts with FANCD2 to recognize branched structures. J. Biol. Chem. *284*, 24443–24452.
19. Kalb, R., Neveling, K., Hoehn, H., Schneider, H., Linka, Y., Batish, S.D., Hunt, C., Berwick, M., Callen, E., Surralles, J., et al. (2007). Hypomorphic mutations in the gene encoding a key Fanconi anemia protein, FANCD2, sustain a significant group of FA-D2 patients with severe phenotype. Am. J. Hum. Genet. *80*, 895–910.
20. Howlett, N.G., Taniguchi, T., Olson, S., Cox, B., Waisfisz, Q., De Die-Smulders, C., Persky, N., Grompe, M., Joenje, H., Pals, G., et al. (2002). Biallelic inactivation of BRCA2 in Fanconi anemia. Science *297*, 606–609.
21. Reid, S., Schindler, D., Hanenberg, H., Barker, K., Hanks, S., Kalb, R., Neveling, K., Kelly, P., Seal, S., Freund, M., et al. (2007). Biallelic mutations in PALB2 cause Fanconi anemia subtype FA-N and predispose to childhood cancer. Nat. Genet. *39*, 162–164.
22. Levran, O., Attwooll, C., Henry, R.T., Milton, K.L., Neveling, K., Rio, P., Batish, S.D., Kalb, R., Velleuer, E., Barral, S., et al.

(2005). The BRCA1-interacting helicase BRIP1 is deficient in Fanconi anemia. Nat. Genet. *37*, 931–933.
23. Singh, T.R., Bakker, S.T., Agarwal, S., Jansen, M., Grassman, E., Godthelp, B.C., Ali, A.M., Du, C.H., Rooimans, M.A., Fan, Q., et al. (2009). Impaired FANCD2 monoubiquitination and hypersensitivity to camptothecin uniquely characterize Fanconi anemia complementation group M. Blood *114*, 174–180.
24. Offit, K., Levran, O., Mullaney, B., Mah, K., Nafa, K., Batish, S.D., Diotti, R., Schneider, H., Deffenbaugh, A., Scholl, T., et al. (2003). Shared genetic susceptibility to breast cancer, brain tumors, and Fanconi anemia. J. Natl. Cancer Inst. *95*, 1548–1551.
25. Wagner, J.E., Tolar, J., Levran, O., Scholl, T., Deffenbaugh, A., Satagopan, J., Ben-Porat, L., Mah, K., Batish, S.D., Kutler, D.I., et al. (2004). Germline mutations in BRCA2: Shared genetic susceptibility to breast cancer, early onset leukemia, and Fanconi anemia. Blood *103*, 3226–3229.
26. Vaz, F., Hanenberg, H., Schuster, B., Barker, K., Wiek, C., Erven, V., Neveling, K., Endt, D., Kesterton, I., Autore, F., et al. (2010). Mutation of the RAD51C gene in a Fanconi anemia-like disorder. Nat. Genet. *42*, 406–409.
27. Gillio, A.P., Verlander, P.C., Batish, S.D., Giampietro, P.F., and Auerbach, A.D. (1997). Phenotypic consequences of mutations in the Fanconi anemia FAC gene: An International Fanconi Anemia Registry study. Blood *90*, 105–110.
28. Yamashita, T., Wu, N., Kupfer, G., Corless, C., Joenje, H., Grompe, M., and D'Andrea, A.D. (1996). Clinical variability of Fanconi anemia (type C) results from expression of an amino terminal truncated Fanconi anemia complementation group C polypeptide with partial activity. Blood *87*, 4424–4432.
29. Futaki, M., Yamashita, T., Yagasaki, H., Toda, T., Yabe, M., Kato, S., Asano, S., and Nakahata, T. (2000). The IVS4 + 4 A to T mutation of the fanconi anemia gene FANCC is not associated with a severe phenotype in Japanese patients. Blood *95*, 1493–1498.
30. Tachibana, A., Kato, T., Ejima, Y., Yamada, T., Shimizu, T., Yang, L., Tsunematsu, Y., and Sasaki, M.S. (1999). The FANCA gene in Japanese Fanconi anemia: Reports of eight novel mutations and analysis of sequence variability. Hum. Mutat. *13*, 237–244.
31. Demuth, I., Wlodarski, M., Tipping, A.J., Morgan, N.V., de Winter, J.P., Thiel, M., Gräsl, S., Schindler, D., D'Andrea, A.D., Altay, C., et al. (2000). Spectrum of mutations in the Fanconi anaemia group G gene, FANCG/XRCC9. Eur. J. Hum. Genet. *8*, 861–868.
32. Yamada, T., Tachibana, A., Shimizu, T., Mugishima, H., Okubo, M., and Sasaki, M.S. (2000). Novel mutations of the FANCG gene causing alternative splicing in Japanese Fanconi anemia. J. Hum. Genet. *45*, 159–166.
33. Levran, O., Diotti, R., Pujara, K., Batish, S.D., Hanenberg, H., and Auerbach, A.D. (2005). Spectrum of sequence variations in the FANCA gene: An International Fanconi Anemia Registry (IFAR) study. Hum. Mutat. *25*, 142–149.
34. Faivre, L., Guardiola, P., Lewis, C., Dokal, I., Ebell, W., Zatterale, A., Altay, C., Poole, J., Stones, D., Kwee, M.L., et al; European Fanconi Anemia Research Group. (2000). Association of complementation group and mutation type with clinical outcome in fanconi anemia. Blood *96*, 4064–4070.
35. Huck, K., Hanenberg, H., Gudowius, S., Fenk, R., Kalb, R., Neveling, K., Betz, B., Niederacher, D., Haas, R., Göbel, U., et al. (2006). Delayed diagnosis and complications of Fanconi anaemia at advanced age—A paradigm. Br. J. Haematol. *133*, 188–197.
36. Cumming, R.C., Lightfoot, J., Beard, K., Youssoufian, H., O'Brien, P.J., and Buchwald, M. (2001). Fanconi anemia group C protein prevents apoptosis in hematopoietic cells through redox regulation of GSTP1. Nat. Med. *7*, 814–820.
37. Davies, S.M., Radloff, G.A., DeFor, T.E., Levran, O., Batish, S.D., Hanenberg, H., and Auerbach, A.D. (2005). GST genotype may modify clinical phenotype in patients with Fanconi anaemia. Br. J. Haematol. *131*, 118–122.
38. Burset, M., Seledtsov, I.A., and Solovyev, V.V. (2000). Analysis of canonical and non-canonical splice sites in mammalian genomes. Nucleic Acids Res. *28*, 4364–4375.
39. Burset, M., Seledtsov, I.A., and Solovyev, V.V. (2001). SpliceDB: Database of canonical and non-canonical mammalian splice sites. Nucleic Acids Res. *29*, 255–259.
40. Wahl, M.C., Will, C.L., and Lührmann, R. (2009). The spliceosome: Design principles of a dynamic RNP machine. Cell *136*, 701–718.
41. Hall, S.L., and Padgett, R.A. (1996). Requirement of U12 snRNA for in vivo splicing of a minor class of eukaryotic nuclear pre-mRNA introns. Science *271*, 1716–1718.
42. Tarn, W.Y., and Steitz, J.A. (1996). Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns. Science *273*, 1824–1832.
43. Tarn, W.Y., and Steitz, J.A. (1996). A novel spliceosome containing U11, U12, and U5 snRNPs excises a minor class (AT-AC) intron in vitro. Cell *84*, 801–811.
44. Krawczak, M., Thomas, N.S., Hundrieser, B., Mort, M., Wittig, M., Hampe, J., and Cooper, D.N. (2007). Single base-pair substitutions in exon-intron junctions of human genes: Nature, distribution, and consequences for mRNA splicing. Hum. Mutat. *28*, 150–158.
45. Roca, X., Olson, A.J., Rao, A.R., Enerly, E., Kristensen, V.N., Børresen-Dale, A.L., Andresen, B.S., Krainer, A.R., and Sachidanandam, R. (2008). Features of 5′-splice-site efficiency derived from disease-causing mutations and comparative genomics. Genome Res. *18*, 77–87.
46. Kammler, S., Leurs, C., Freund, M., Krummheuer, J., Seidel, K., Tange, T.O., Lund, M.K., Kjems, J., Scheid, A., and Schaal, H. (2001). The sequence complementarity between HIV-1 5′ splice site SD4 and U1 snRNA determines the steady-state level of an unstable env pre-mRNA. RNA *7*, 421–434.
47. Freund, M., Asang, C., Kammler, S., Konermann, C., Krummheuer, J., Hipp, M., Meyer, I., Gierling, W., Theiss, S., Preuss, T., et al. (2003). A novel approach to describe a U1 snRNA binding site. Nucleic Acids Res. *31*, 6963–6975.
48. Zhuang, Y., and Weiner, A.M. (1986). A compensatory base change in U1 snRNA suppresses a 5′ splice site mutation. Cell *46*, 827–835.
49. Pinotti, M., Rizzotto, L., Balestra, D., Lewandowska, M.A., Cavallari, N., Marchetti, G., Bernardi, F., and Pagani, F. (2008). U1-snRNA-mediated rescue of mRNA processing in severe factor VII deficiency. Blood *111*, 2681–2684.
50. Pinotti, M., Balestra, D., Rizzotto, L., Maestri, I., Pagani, F., and Bernardi, F. (2009). Rescue of coagulation factor VII function by the U1+5A snRNA. Blood *113*, 6461–6464.
51. Tanner, G., Glaus, E., Barthelmes, D., Ader, M., Fleischhauer, J., Pagani, F., Berger, W., and Neidhardt, J. (2009). Therapeutic strategy to rescue mutation-induced exon skipping in rhodopsin by adaptation of U1 snRNA. Hum. Mutat. *30*, 255–263.

52. Kohn, D.B., and Candotti, F. (2009). Gene therapy fulfilling its promise. N. Engl. J. Med. *360*, 518–521.
53. Vassilopoulos, G., and Rethwilm, A. (2008). The usefulness of a perfect parasite. Gene Ther. *15*, 1299–1301.
54. Williams, D.A. (2008). Foamy virus vectors come of age. Mol. Ther. *16*, 635–636.
55. Hanenberg, H., Xiao, X.L., Dilloo, D., Hashino, K., Kato, I., and Williams, D.A. (1996). Colocalization of retrovirus and target cells on specific fibronectin fragments increases genetic transduction of mammalian cells. Nat. Med. *2*, 876–882.
56. Hanenberg, H., Hashino, K., Konishi, H., Hock, R.A., Kato, I., and Williams, D.A. (1997). Optimization of fibronectin-assisted retroviral gene transfer into human CD34+ hematopoietic cells. Hum. Gene Ther. *8*, 2193–2206.
57. Hanenberg, H., Batish, S.D., Pollok, K.E., Vieten, L., Verlander, P.C., Leurs, C., Cooper, R.J., Göttsche, K., Haneline, L., Clapp, D.W., et al. (2002). Phenotypic correction of primary Fanconi anemia T cells with retroviral vectors as a diagnostic tool. Exp. Hematol. *30*, 410–420.
58. Chandra, S., Levran, O., Jurickova, I., Maas, C., Kapur, R., Schindler, D., Henry, R., Milton, K., Batish, S.D., Cancelas, J.A., et al. (2005). A rapid method for retrovirus-mediated identification of complementation groups in Fanconi anemia patients. Mol. Ther. *12*, 976–984.
59. Verlander, P.C., Lin, J.D., Udono, M.U., Zhang, Q., Gibson, R.A., Mathew, C.G., and Auerbach, A.D. (1994). Mutation analysis of the Fanconi anemia gene FACC. Am. J. Hum. Genet. *54*, 595–601.
60. Betz, B., Theiss, S., Aktas, M., Konermann, C., Goecke, T.O., Moslein, G., Schaal, H., and Royer-Pokora, B. (2010). Comparative in silico analyses and experimental validation of novel splice site and missense mutations in the genes MLH1 and MSH2. J. Cancer Res. Clin. Oncol. *136*, 123–134.
61. Kammler, S., Otte, M., Hauber, I., Kjems, J., Hauber, J., and Schaal, H. (2006). The strength of the HIV-1 3′ splice sites affects Rev function. Retrovirology *3*, 89.
62. Schaal, H., Klein, M., Gehrmann, P., Adams, O., and Scheid, A. (1995). Requirement of N-terminal amino acid residues of gp41 for human immunodeficiency virus type 1-mediated cell fusion. J. Virol. *69*, 3308–3314.
63. Maddon, P.J., Dalgleish, A.G., McDougal, J.S., Clapham, P.R., Weiss, R.A., and Axel, R. (1986). The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. Cell *47*, 333–348.
64. Selden, R.F., Howie, K.B., Rowe, M.E., Goodman, H.M., and Moore, D.D. (1986). Human growth hormone as a reporter gene in regulation studies employing transient gene expression. Mol. Cell. Biol. *6*, 3173–3179.
65. Schindler, D., and Hoehn, H. (1988). Fanconi anemia mutation causes cellular susceptibility to ambient oxygen. Am. J. Hum. Genet. *43*, 429–435.
66. Mochizuki, H., Schwartz, J.P., Tanaka, K., Brady, R.O., and Reiser, J. (1998). High-titer human immunodeficiency virus type 1-based vector systems for gene delivery into nondividing cells. J. Virol. *72*, 8873–8883.
67. Pietschmann, T., Heinkelein, M., Heldmann, M., Zentgraf, H., Rethwilm, A., and Lindemann, D. (1999). Foamy virus capsids require the cognate envelope protein for particle export. J. Virol. *73*, 2613–2621.
68. Hartmann, L., Theiss, S., Niederacher, D., and Schaal, H. (2008). Diagnostics of pathogenic splicing mutations: Does bioinformatics cover all bases? Front. Biosci. *13*, 3252–3272.
69. Garcia-Higuera, I., Taniguchi, T., Ganesan, S., Meyn, M.S., Timmers, C., Hejna, J., Grompe, M., and D'Andrea, A.D. (2001). Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway. Mol. Cell *7*, 249–262.
70. Roca, X., and Krainer, A.R. (2009). Recognition of atypical 5′ splice sites by shifted base-pairing to U1 snRNA. Nat. Struct. Mol. Biol. *16*, 176–182.
71. Aebi, M., Hornig, H., and Weissmann, C. (1987). 5′ cleavage site in eukaryotic pre-mRNA splicing is determined by the overall 5′ splice region, not by the conserved 5′ GU. Cell *50*, 237–246.
72. Kyriakopoulou, C., Larsson, P., Liu, L., Schuster, J., Söderbom, F., Kirsebom, L.A., and Virtanen, A. (2006). U1-like snRNAs lacking complementarity to canonical 5′ splice sites. RNA *12*, 1603–1611.
73. Popp, H., Kalb, R., Fischer, A., Lobitz, S., Kokemohr, I., Hanenberg, H., and Schindler, D. (2003). Screening Fanconi anemia lymphoid cell lines of non-A, C, D2, E, F, G subtypes for defects in BRCA2/FANCD1. Cytogenet. Genome Res. *103*, 54–57.
74. Meyer, K., Marquis, J., Trüb, J., Nlend Nlend, R., Verp, S., Ruepp, M.D., Imboden, H., Barde, I., Trono, D., and Schümperli, D. (2009). Rescue of a severe mouse model for spinal muscular atrophy by U7 snRNA-mediated splicing modulation. Hum. Mol. Genet. *18*, 546–555.
75. Bainbridge, J.W., Smith, A.J., Barker, S.S., Robbie, S., Henderson, R., Balaggan, K., Viswanathan, A., Holder, G.E., Stockman, A., Tyler, N., et al. (2008). Effect of gene therapy on visual function in Leber's congenital amaurosis. N. Engl. J. Med. *358*, 2231–2239.
76. Aiuti, A., and Roncarolo, M.G. (2009). Ten years of gene therapy for primary immune deficiencies. Hematology Am. Soc. Hematol. Educ. Program *2009*, 682–689.
77. Marquis, J., Meyer, K., Angehrn, L., Kämpfer, S.S., Rothen-Rutishauser, B., and Schümperli, D. (2007). Spinal muscular atrophy: SMN2 pre-mRNA splicing corrected by a U7 snRNA derivative carrying a splicing enhancer sequence. Mol. Ther. *15*, 1479–1486.
78. Timmers, C., Taniguchi, T., Hejna, J., Reifsteck, C., Lucas, L., Bruun, D., Thayer, M., Cox, B., Olson, S., D'Andrea, A.D., et al. (2001). Positional cloning of a novel Fanconi anemia gene, FANCD2. Mol. Cell *7*, 241–248.
79. Sandoval, N., Platzer, M., Rosenthal, A., Dörk, T., Bendix, R., Skawran, B., Stuhrmann, M., Wegner, R.D., Sperling, K., Banin, S., et al. (1999). Characterization of ATM gene mutations in 66 ataxia telangiectasia families. Hum. Mol. Genet. *8*, 69–79.
80. Wimmer, K., Roca, X., Beiglböck, H., Callens, T., Etzler, J., Rao, A.R., Krainer, A.R., Fonatsch, C., and Messiaen, L. (2007). Extensive in silico analysis of NF1 splicing defects uncovers determinants for splicing outcome upon 5′ splice-site disruption. Hum. Mutat. *28*, 599–612.