benzidine (DAB) substrate chromogen system (Dako Cytomation). The sections were further processed for enhancement of the DAB reaction products by use of a Dako Envision Kit (Dako Cytomation).

Full-Length Cloning and Direct Sequencing of the TA14_391 Isoform

To determine the full-length TAF1 isoform, including the 5′ end of the transcript, a CapSite cDNA library derived from human brain (317-04041 [Nippon Gene]) was used. The CapSite cDNA libraries consist of cDNAs in which the 5′ cap structure (m7Gppp) of eukaryotic mRNA is replaced with a synthetic oligoribonucleotide to label the 5′ end of the cDNA, enabling identification of the 5′ end sequence by PCR.[10] For amplification of the 3′ end, a whole Marathon-Ready cDNA library derived from human brain (BD Biosciences Clontech) was used. The reaction mixture contained 1 μl of the library, 200 mM of each primer, 0.16 mM of each deoxyribonucleotide diphosphate (dNTP), 1× BD Advantages2 PCR buffer, and 1× BD Advantages2 Polymerase Mix (639300 [BD Biosciences Clontech]) in a total volume of 50 μl. The PCR consisted of denaturation for 30 s at 94°C, followed by 5 cycles for 5 s at 94°C and 10 min at 70°C and then 20 cycles for 5 s at 94°C and 10 min at 78°C. An aliquot of the first PCR product was used for the second PCR reaction under the same conditions as the first-round PCR but with different primers. We used a primer set for the first PCR—first RDT primer (Nippon Gene) and TA3_r_5070 (5′-GGTATCATACAAATCAGGAGGCTT-3′)—and then used a set for the second heminested PCR—second primer and TA3_r_5070. These primers were purified by PAGE extraction. Alternative exon 34′–specific primers were designed to prevent erroneous amplification due to PCR slippage. For amplification of the 3′ end, a whole Marathon-Ready cDNA library derived from human brain (BD Biosciences Clontech) was used along with the primer set TA6_f_5032 (5′-CCCTACACGCCTCAGGCTA-3′) and TA2_r_7606 (5′-CCAGCATACATAACAAACACAGAAG-3′) under the same conditions as those used for the CapSite cDNA but as a single PCR reaction. Direct sequencing was done on these PCR products. After purification by a PCR Product Pre-Sequencing kit, cycle sequencing was performed using a BigDye Terminator v3.1 Cycle Sequencing kit, with 20 internal primers for a second PCR product from CapSite cDNA and 8 internal primers for the PCR product from Marathon-Ready cDNA. Direct sequencing was done on two long PCR products by use of the following primers: TAF_f_377, TAF_f_755, TAF_f_1152, TAF_f_1528, TAF_f_1889, TAF_f_2580, TAF_f_2945, TAF_f_3311, TAF_f_3689, TAF_f_4069, TAF_f_4413, TAF_f_4780, TAF_r_2273, TAF_r_2598, TAF_r_2979, TAF_r_3331, TAF_r_5164, TAF_r_5505, TAF_r_5862, TAF_r_6231, TAF_r_6469, TAF_r_6855, TAF_r_7164, TAF_r_7553, and TAF_r_7606 (table 3).

Results

Go to:

Entire Genomic Sequence of the *DYT3* Region

Two series of BAC libraries were constructed using DNA from a patient with XDP who had a disease-specific haplotype[8,9] in the *DYT3* region. From these libraries, a continuous BAC contig consisting of eight BAC clones was then generated to cover the *DYT3* region between the *GJB1* and *CXCR3* genes (fig. 1*a*). By applying a shotgun sequencing strategy, we accurately determined the complete DNA sequence of the BAC contig, with 5.7-fold redundancy. The total sequence length of the BAC contig was 463,567 bp. A comparison between our sequence from the patient with XDP and a reference sequence from National Center for Biotechnology Information (NCBI) build 30 showed a total of 159 sequence variants: 89 single-nucleotide substitutions, 68 small insertions/deletions (indels), 1 retrotransposal insertion, and 1 large (1,666-bp) deletion (all variants are listed in table 5). Of these variants, 53 were known SNPs, comprising 50 substitutions and 3 indels. Of the 68 indels, 62 were repetitive units of STRs, and the other 6 indels were also located in certain types of STRs. The large deletion was a direct repeat sequence spanning 1,666 bp. The retrotransposal insertion in intron 32 of the *TAF1* gene (fig. 1*a*) was 2,627 bp in length, which is categorized as an SVA (short interspersed nuclear element, VNTR, and Alu composite) retrotransposon.[12,13] The SVA retrotransposon insertion in the *DYT3* region had been never reported. However, none of these variants was located in any exon or promoter of the annotated genes that have experimentally verified coding sequences, including their alternative splicing exons. In the region between *DXS10017* and *DXS10018*, there were 25 variants, which included the SVA insertion and the 1,666-bp deletion. We confirmed these variants in an ethnic panel (see the "Material and Methods" section) that included patients with XDP