ONLINE MUTATION REPORT

Expression analysis of an *FGFR2* IIIc 5′ splice site mutation (1084+3A→G)

R Kan, S R F Twigg, J Berg, L Wang, F Jin, A O M Wilkie

J Med Genet 2004;**41**:e108 (http://www.jmedgenet.com/cgi/content/full/41/8/e108). doi: 10.1136/jmg.2004.018507

Sequence variations within splice sites may pose problems in the interpretation of their pathogenic effect, especially when these variations occur outside the highly conserved /gt (donor or 5′ site) and ag/ (acceptor or 3′ site) consensus dinucleotides that immediately flank most exons. A commonly used method to evaluate the probable effect of a sequence variation on splicing is to calculate the Shapiro-Senapathy (SS) score, which is based on the extended splice site consensus sequence.[1–3] Here we present a sequence variation in a 5′ splice site of the gene encoding fibroblast growth factor receptor type 2 (*FGFR2*) that maintains a consensus nucleotide at the variant position, but nevertheless causes a switch to the use of a cryptic 5′ splice site within the upstream (IIIc) exon. This variant is present in three generations of a family and manifests with mild features of Crouzon syndrome.

Key points

- Interpretation of the pathogenic effect of sequence changes outside the highly conserved donor and acceptor sites at splice site junctions can be problematic.
- We analysed a family in which members with mild features of Crouzon syndrome are heterozygous for a sequence variation with a 5′ splice site of *FGFR2* (1084+3A→G).
- Analysis of RNA revealed that the A→G sequence change results in activation of an upstream cryptic donor splice site. The consequence is synthesis of *FGFR2* mRNA lacking 51 nucleotides in the IIIc exon.
- Although both A and G are consensus nucleotides at the +3 position of the 5′ splice site, the A→G substitution reduces the strength of the splice site.
- Our data elucidate the mechanism of the 1084+3A→G mutation and confirm its pathogenicity.

PATIENTS AND METHODS

Following informed consent blood samples were collected from members of the family pedigree shown in fig 1. DNA was isolated by proteinase K treatment and phenol chloroform extraction. RNA was also extracted from the blood of affected individual II-2 and cDNA synthesised by standard techniques.

Mutation screening of exons IIIa and IIIc of *FGFR2*[4] was undertaken in DNA from the proband using WAVE (3500HT; Transgenomic) denaturing high performance liquid chromatography.[5] Sequencing was performed with Big Dye (version 3) on an ABI 3100 DNA sequencer.

Oligonucleotide hybridisation was carried out on cDNA derived from individual II-2. Reverse transcriptase PCR (RT-PCR) using primers 4F and 9R located respectively in *FGFR2* exons 6 and 12 (exon numbering according to Kan et al[5]; see fig 2A), was performed as previously described.[7] The RT-PCR product was separated by electrophoresis on a 2% agarose gel, blotted onto a nylon membrane (Zeta-Probe, Bio-Rad), and the presence of various potential splicing products was analysed by hybridisation to specific ^{32}P radiolabelled oligonucleotide probes[7] SPL7/8 (IIIa–IIIb), SPL7/9 (IIIa–IIIc) and SPL7/11 (IIIa-exon 12). An additional probe 9C/TM (5′-TGCTTGGCGGCGCCTGGAA-3′) was used to detect splicing of the cryptic donor splice site in exon IIIc to the downstream transmembrane exon 11.

RESULTS AND DISCUSSION

The family pedigree is shown in fig 1A and the facial phenotype is illustrated in fig 1B. The proband III-3 presented aged 3½ years with scaphocephaly and exorbitism. He was found to have sagittal and partial metopic synostosis, which was corrected surgically. His mother II-2, uncle II-3 (fig 1B) and grandfather I-1 (not illustrated) had a similar facial appearance but had not undergone any reconstructive surgery.

In view of the features of mild Crouzon syndrome we screened exons IIIa and IIIc of *FGFR2*. This revealed heterozygosity for a nucleotide transition, 1084+3A→G, located at the third (intronic) base of the 5′ splice site lying downstream of the alternatively spliced exon IIIc. The mutation was independently confirmed by demonstrating a new cleavage site for the restriction enzyme *Acc*I (not shown). The same substitution was identified by DNA sequencing in three other clinically affected family members, I-1, II-2, and II-3, but not in two clinically unaffected individuals at 50% prior risk (II-4 and II-5). Although this nucleotide change has been described twice before, in a child with Pfeiffer syndrome whose parents were normal on DNA testing,[10] and in a clinically unclassified patient with sagittal and lambdoid synostosis,[11] these reports did not include verification of sample relationships, nor experimental data documenting the pathogenic mechanism of the mutation. We therefore sought to obtain direct evidence of an effect of this putative mutation on splicing.

The results of the oligonucleotide hybridisations are shown in figs 2B–E. Whereas the normal control sample only showed the expected exon IIIc splice product (fig 2B, lane N), two additional shorter products were identifiable in the patient sample. DNA sequencing and hybridisation to specific oligonucleotides showed that one of these arose from direct splicing of exon IIIa to exon 12 (fig 2E), whereas the other utilised a cryptic splice site within exon IIIc (5′-CGGgtaatt-3′; new intronic nucleotides in lower case) (figs 2B, C). This cryptic splice site was previously known, both because it is

Abbreviations: *FGFR2*, fibroblast growth factor receptor type 2; RT-PCR, reverse transcriptase PCR; SS, Shapiro-Senapathy

activated by the pathogenic synonymous substitution 1032G→A in exon IIIc (Ala344Ala mutation),[4] creating the donor splice site 5′-CAGgtaatt-3′[12 13] and because the wild-type cryptic sequence is utilised in a mutant allele with a 6 nucleotide insertion that destroys the normal 5′ splice site.[14] There was no evidence for ectopic expression of the alternatively spliced IIIb exon (fig 2D).[7]

The activation of the cryptic splice site by the 1084+3A→G initially seemed surprising, because the consensus at +3 is either A or G, with A being slightly more prevalent than G in a large survey of 5′ splice sites.[1 2] However the SS score of splice site strength is somewhat lower for the mutated than the authentic splice site (75.39 wild-type against 72.25 mutated), and the difference is accentuated by several other methods used to calculate splice site strength (table 1). Interestingly, by two of these methods (SS and ΔG), the cryptic IIIc splice site is actually stronger than the authentic one. This suggests that only a mild reduction in strength of the wild-type site, as occurs in the case of the 1084+3A→G mutation, is sufficient to tip the balance towards use of the cryptic splice site. This analysis also shows that the cryptic site is further strengthened by the synonymous 1032G→A mutation (table 1).

In the complex process of splice site recognition, the RNA sequence in U1 small nuclear ribonucleoprotein plays an important role by binding to the 5′ splice site at the −2 to +6 positions.[18] The effect of the 1084+3A→G mutation is to reduce complementarity of the authentic 5′ site to the U1 sequence, whereas the 1032G→A mutation enhances complementarity of the cryptic site (see ΔG values in table 1). The consequence of both mutations is similar: the preferential synthesis of *FGFR2* mRNA lacking 51 nucleotides in the IIIc exon, thus encoding a protein that lacks 17 amino acids. It has previously been noted that the 1032G→A mutation is associated with a milder phenotype than most *FGFR2* mutations[19] and, based on the clinical phenotype of the present family, this also appears to be the case for 1084+3A→G.

Ohno et al[20] proposed that +3A→G mutations at 5′ splice sites are pathogenic when both the +4, and one or other of the +5/+6 positions, are not complementary to U1 in the wild-type sequence; the authentic *FGFR2* IIIc 5′ splice site indeed obeys this rule (+4 and +5 are non-complementary; table 1). In summary, our data elucidate the mechanism of the 1084+3A→G mutation and confirm its pathogenicity. Although a G at +3 is frequently recognised as part of the

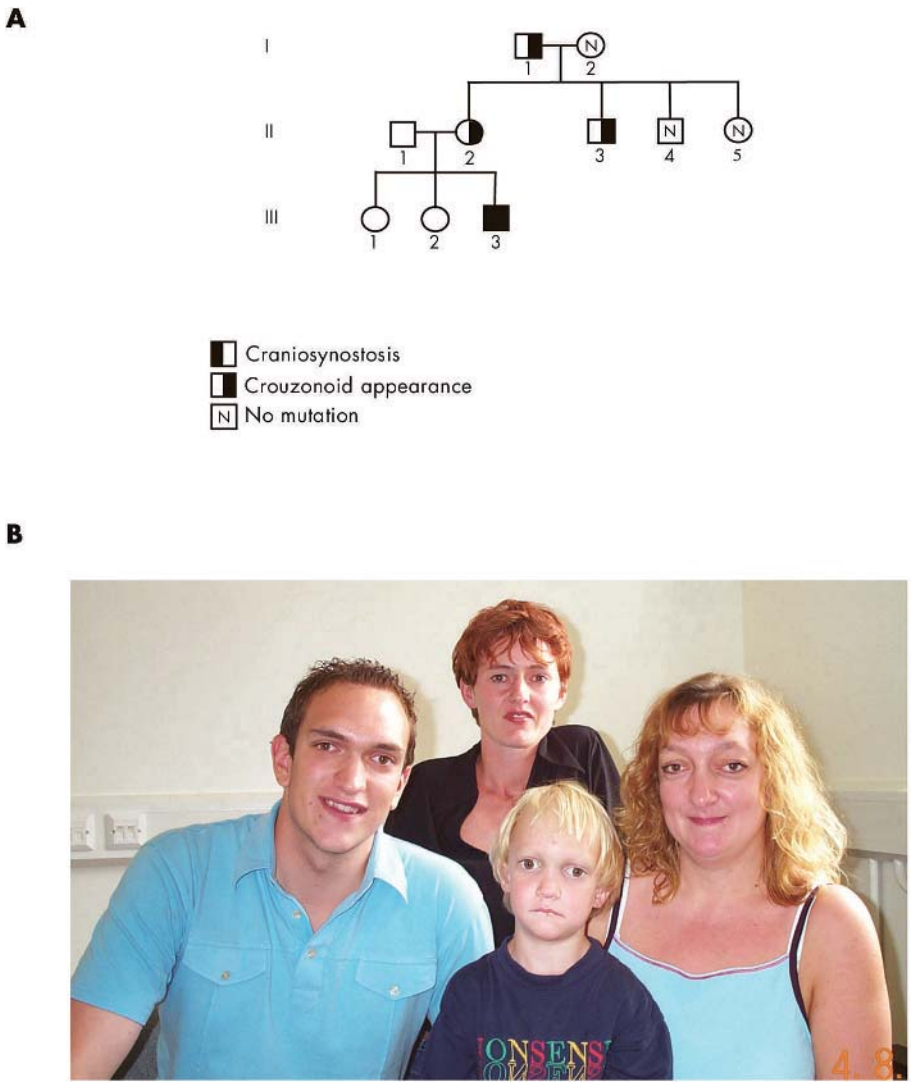

Figure 1 Clinical features of the family. (A) Pedigree showing transmission of the phenotype through three generations. (B) Facial appearance of family members (from left to right) II-3, II-5 (unaffected), III-3, and II-2 *(Photograph with family's permission).*

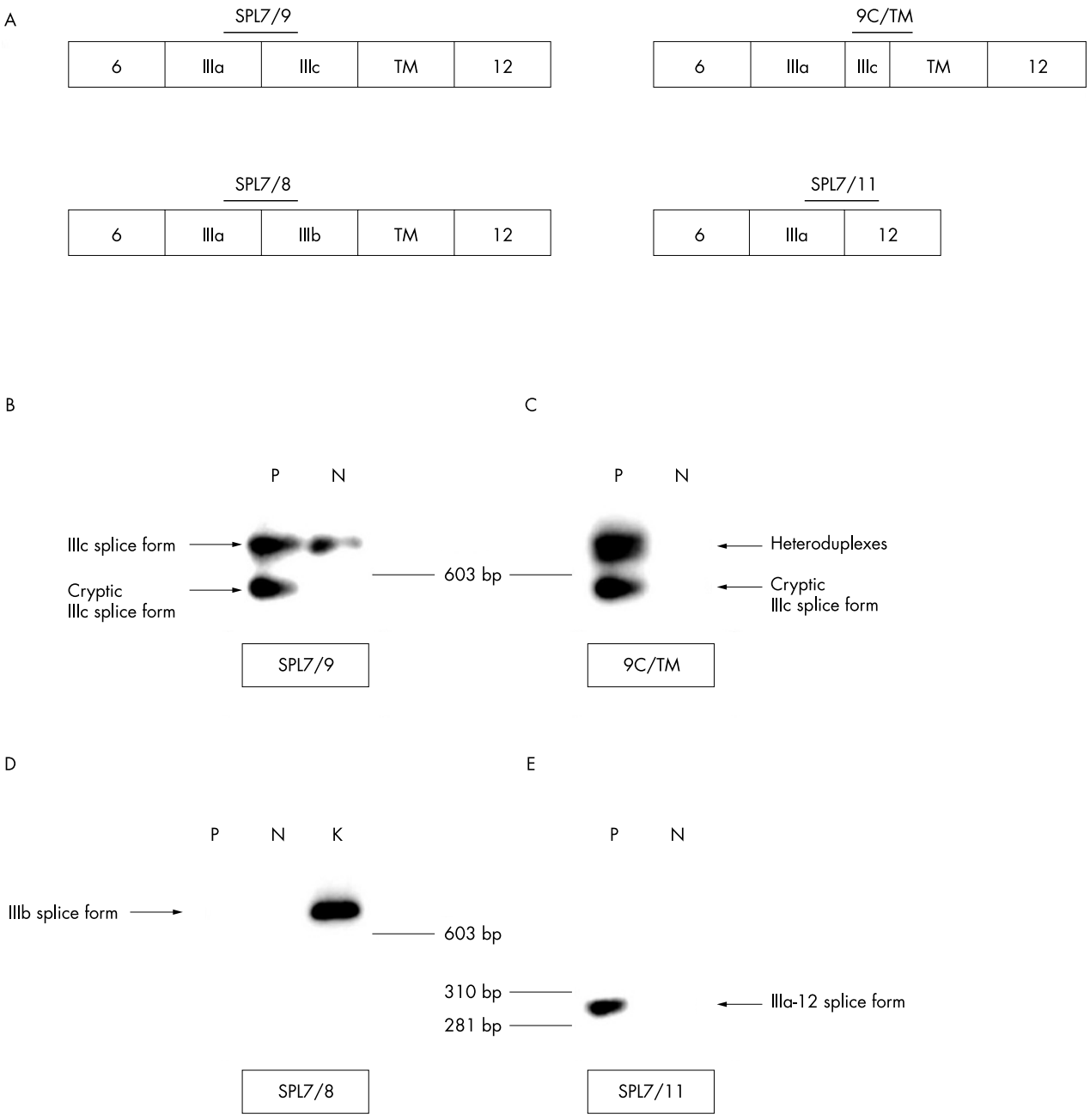

Figure 2 Analysis of *FGFR2* splicing. (A) Part of *FGFR2* showing various splice forms and their detection by different oligonucleotides; note the obligatory physiological alternative splice between either the IIIb or IIIc exons.[6] (B–E) Blot hybridisation analysis of RT-PCR products using primers 4F (exon 6) and 9R (exon 12). The originating RNA was extracted from whole blood from II-2 (heterozygous for the 1084+3A→G mutation) (P), whole blood from a normal control (N), and control keratinocytes (K). (B) Oligonucleotide SPL7/9 matches the IIIa-IIIc boundary and normally detects a 640 bp product. In the patient, an additional smaller product is detected. (C) This product (589 bp) specifically hybridises to the 9C/TM oligonucleotide, showing that it is generated by use of the cryptic 5′ splice site within exon IIIc. The larger fragments are heteroduplexes. (D) An oligonucleotide (SPL7/8) diagnostic for the 643 bp IIIb splice form does not hybridise to the patient lane, but is positive in a keratinocyte sample that characteristically expresses this alternative splice form. (E) The patient sample also expresses lower amounts of a smaller 292 bp product that is identified using SPL7/10 as a IIIa-exon 12 splice form. This product was not present in control cDNA, but has been observed previously[7 8] and might have a physiological function.[9] Size markers are HaeIII-digested φX174 DNA.

5′ splice consensus, in a small number of documented cases +3A is nevertheless essential to accomplish normal splicing.

ACKNOWLEDGEMENTS

We thank the family for their help with this study and N Elanko and A O'Rourke for undertaking the mutation analysis.

Authors' affiliations

R Kan, S R F Twigg, A O M Wilkie, Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Oxford, UK
R Kan, L Wang, F Jin, Institute of Genetics and Developmental Biology and Laboratory for Human and Animal Genetic Studies, Chinese Academy of Sciences, Beijing, P. R. China

Table 1 Comparison of 5′ splice sites

		−3	−2	−1	+1	+2	+3	+4	+5	+6		SS	ΔG	MAX	MDD	MM	NN
Authentic IIIc		C	A	G	g	t	a	t	a	t		75.39	−8.8	7.88	11.58	7.24	0.88
1084+3A→G		C	A	G	g	t	g	t	a	t		72.25	−8.5	3.09	8.48	3.93	0.19
Cryptic IIIc		C	G	G	g	t	a	a	t	t		78.71	−9.3	6.79	9.78	5.97	0.84
Mutant cryptic IIIc		C	A	G	g	t	a	a	t	t		86.21	−9.7	8.55	13.08	7.72	0.87
Authentic IIIb		A	A	G	g	t	a	a	c	a		80.10	−7.9	8.92	12.68	6.99	0.94
U1 snRNA	3′-	G	U	C	C	A	U	U	C	A	-5′						

SS, Shapiro and Senapathy matrix;
ΔG, RNA duplex stability with U1 snRNA. (Turner energy rules: www.bioinfo.rpi.edu/~zukerm/rna/energy)[15];
MAX, Maximum entropy model;
MDD, Maximum dependence decomposition model;
MM, First order Markov model. (http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html)[16];
NN, Neural network splice site prediction. (www.fruitfly.org/seq_tools/splice).[17]

R Kan, Faculty of Life Science, Inner Mongolia University, Huhhot, P. R. China
J Berg, Clinical Genetics, Guy's and St Thomas' Hospital NHS Trust, London, UK (present address: Molecular and Cellular Pathology, Division of Pathology and Neuroscience, Ninewells Hospital and Medical School, Dundee, UK)

This work was supported by the China Scholarship Council (to Rui Kan) and the Wellcome Trust.

Conflicts of interest: none declared.

Correspondence to: Dr S R F Twigg, Weatherall Institute of Molecular Medicine, The John Radcliffe Hospital, Oxford OX3 9DS, UK; stwigg@hammer.imm.ox.ac.uk

Revised version received 11 March 2004
Accepted for publication 11 March 2004

REFERENCES

1 **Shapiro MB**, Senapathy P. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res* 1987;**15**:7155–74.
2 **Senapathy P**, Shapiro MB, Harris NL. Splice junctions, branch point sites and exons: sequence statistics, identification and applications to the genome project. *Methods Enzymol* 1990;**183**:252–78.
3 **Roca X**, Sachidanandam R, Krainer AR. Intrinsic differences between authentic and cryptic 5′ splice sites. *Nucleic Acids Res* 2003;**31**:6321–33.
4 **Reardon W**, Winter RM, Rutland P, Pulleyn LJ, Jones BM, Malcolm S. Mutations in the fibroblast growth factor receptor 2 gene cause Crouzon syndrome. *Nat Genet* 1994;**8**:98–103.
5 **Kan S-H**, Elanko N, Johnson D, Cornejo-Roldan LR, Cook J, Reich EW, Tomkins S, Verloes A, Twigg SRF, Rannan-Eliya S, McDonald-McGinn DM, Zachai EH, Wall SA, Muenke M, Wilkie AOM. Genomic screening of fibroblast growth factor receptor 2 reveals a wide spectrum of mutations in patients with syndromic craniosynostosis. *Am J Hum Genet* 2002;**70**:472–86.
6 **Miki T**, Fleming TP, Bottaro DP, Rubin JS, Ron D, Aaronson SA. Expression cDNA cloning of the KGF receptor by creation of a transforming autocrine loop. *Science* 1991;**251**:72–5.
7 **Oldridge M**, Zackai EH, McDonald-McGinn DM, Iseki S, Morriss-Kay GM, Twigg SRF, Johnson D, Wall SA, Jiang W, Theda C, Jabs EW, Wilkie AOM. De novo *Alu*-element insertions in *FGFR2* identify a distinct pathological basis for Apert syndrome. *Am J Hum Genet* 1999;**64**:446–61.
8 **Katoh M**, Hattori Y, Sasaki H, Tanaka M, Sugano K, Yazaki Y, Sugimura T, Terada M. K-sam gene encodes secreted as well as transmembrane receptor kinase. *Proc Natl Acad Sci U S A* 1992;**89**:2960–4.
9 **Wells A**, Marti U. Signalling shortcuts: cell-surface receptors in the nucleus? *Nat Rev Mol Cell Biol* 2002;**3**:697–702.
10 **Cornejo-Roldan LR**, Roessler E, Muenke M. Analysis of the mutational spectrum of the FGFR2 gene in Pfeiffer syndrome. *Hum Genet* 1999;**104**:425–31.
11 **Mulliken JB**, Steinberger DS, Kunze S, Muller U. Molecular diagnosis of bilateral coronal synostosis. *Plast Reconstr Surg* 1999;**104**:1603–15.
12 **Li X**, Park W-J, Pyeritz RE, Jabs EW. Effect on splicing of a silent FGFR2 mutation on Crouzon syndrome. *Nat Genet* 1995;**9**:232–3.
13 **Del Gatto F**, Breathnach R. A Crouzon syndrome synonymous mutation activates a 5′ splice site with the IIIc exon of the FGFR2 gene. *Genomics* 1995;**27**:558–9.
14 **Meyers GA**, Day D, Goldberg R, Daentl DL, Przylepa KA, Abrams LJ, Graham JM Jr, Feingold M, Moeschler JB, Rawnsley E, Scott AF, Jabs EW. FGFR2 exon IIIa and IIIc mutations in Crouzon, Jackson-Weiss, and Pfeiffer syndromes: evidence for missense changes, insertions, and a deletion due to alternative splicing. *Am J Hum Genet* 1996;**58**:491–8.
15 **Serra MJ**, Turner DH. Predicting thermodynamic properties of RNA. *Methods Enzymol* 1995;**259**:241–61.
16 **Yeo G**, Burge CB. Maximum entropy modelling of short sequence motifs with applications to RNA splicing signals. In: Miller W, Vingron M, Istrail S, Pevzner P, Waterman M, eds. *Proceedings of the 7th annual international conference on computational molecular biology (RECOMB03).* New York: ACM Press, 2003:322–31.
17 **Brunak S**, Engelbrecht J, Knudsen S. Prediction of human mRNA donor and acceptor sites from the DNA sequence. *J Mol Biol* 1991;**220**:49–65.
18 **Madhani HD**, Guthrie C. Dynamic RNA-RNA interactions in the spliceosome. *Annu Rev Genet* 1994;**28**:1–26.
19 **Steinberger D**, Reinhartz T, Unsold R, Müller U. FGFR2 mutation in clinically nonclassifiable autosomal dominant craniosynostosis with pronounced phenotypic variation. *Am J Med Genet* 1996;**66**:81–6.
20 **Ohno K**, Brengman JM, Felice KJ, Cornblath DR, Engel AG. Congenital end-plate acetylcholinesterase deficiency caused a nonsense mutation and an A→G splice-donor-site mutation at position +3 of the collagenlike-tail-subunit gene (COLQ): how does G at position +3 result in aberrant splicing? *Am J Hum Genet* 1999;**65**:635–44.