EFTUD2 missense variants disrupt protein function and splicing in mandibulofacial dysostosis Guion-Almeida type

Huw B. Thomas¹ | Katherine A. Wood¹ | Weronika A. Buczek¹ | Christopher T. Gordon²,³ | Véronique Pingault²,³,⁴ | Tania Attié-Bitach³,⁴,⁵ | Kathryn E. Hentges¹ | Vinod C. Varghese⁶ | Jeanne Amiel²,³,⁴ | William G. Newman¹,⁷ | Raymond T. O'Keefe¹

¹Division of Evolution and Genomic Sciences, School of Biological Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester Academic Health Science Centre, Manchester, UK
²Laboratory of Embryology and Genetics of Human Malformation, Institut National de la Santé et de la Recherche Médicale (INSERM) UMR 1163, Institut Imagine, Paris, France
³Paris Descartes-Sorbonne Paris Cité University, Institut Imagine, Paris, France
⁴Département de Génétique, Hôpital Necker-Enfants Malades, AP-HP, Paris, France
⁵INSERM UMR 1163, Institut Imagine, Paris, France
⁶All Wales Medical Genomics Service, Cardiff, UK
⁷Division of Evolution and Genomic Sciences, School of Biological Sciences, Faculty of Biology, Medicine and Health, Center for Genomic Medicine, St. Mary's Hospital, The University of Manchester, Manchester Academic Health Science Centre, Manchester, UK

Correspondence
Raymond T. O'Keefe, Division of Evolution and Genomic Sciences, School of Biological Sciences, Faculty of Biology, Medicine and Health, The University of Manchester, Manchester Academic Health Science Centre, Manchester M13 9PT, UK.
Email: rokeefe@manchester.ac.uk

Abstract
Pathogenic variants in the core spliceosome U5 small nuclear ribonucleoprotein gene EFTUD2/SNU114 cause the craniofacial disorder mandibulofacial dysostosis Guion-Almeida type (MFDGA). MFDGA-associated variants in EFTUD2 comprise large deletions encompassing EFTUD2, intragenic deletions and single nucleotide truncating or missense variants. These variants are predicted to result in haploinsufficiency by loss-of-function of the variant allele. While the contribution of deletions within EFTUD2 to allele loss-of-function are self-evident, the mechanisms by which missense variants are disease-causing have not been characterized functionally. Combining bioinformatics software prediction, yeast functional growth assays, and a minigene (MG) splicing assay, we have characterized how MFDGA missense variants result in EFTUD2 loss-of-function. Only four of 19 assessed missense variants cause EFTUD2 loss-of-function through altered protein function when modeled in yeast. Of the remaining 15 missense variants, five altered the normal splicing pattern of EFTUD2 pre-messenger RNA predominantly through exon skipping or cryptic splice site activation, leading to the introduction of a premature termination codon. Comparison of bioinformatic predictors for each missense variant revealed a disparity amongst different software packages and, in many cases, an inability to correctly predict changes in splicing subsequently determined by MG interrogation. This study highlights the need for laboratory-based validation of bioinformatic predictions for EFTUD2 missense variants.

Keywords
EFTUD2, mandibulofacial dysostosis Guion-Almeida type, minigene, missense variants, premRNA splicing, Snu114, splicing variants, yeast
1 | INTRODUCTION

The spliceosome is a large RNA/protein complex that is required for removal of intron regions from pre-messenger RNA (pre-mRNA; Wahl, Will, & Lührmann, 2009). The spliceosome is composed of five small nuclear ribonucleoproteins (snRNPs) that associate with the pre-mRNA and are dynamically remodeled to allow the two trans-estereification reactions required for intron removal from pre-mRNA (Will & Lührmann, 2011). Genetic variants in spliceosome-associated genes cause a number of human craniofacial disorders. Variants in the human spliceosome-associated genes TXNLI4A, EFTUD2, SF3B4, SNRPB, and EIF4A3 cause craniofacial disorders: Burn-McKeown syndrome, mandibulofacial dysostosis Guion-Almeida type (MFDGA), Nager syndrome/Rodriguez syndrome, cerebrocostomandibular syndrome, and Richieri-Costa-Pereira syndrome, respectively (Lehalle et al., 2015). In most of these craniofacial disorders, the disease variants inactivate one allele and are proposed to cause disease through haploinsufficiency. Patients with MFDGA possess a wide variety of variants within EFTUD2 (EFTUD2/Snu114) that are associated with disease through haploinsufficiency. Patients with MFDGA possess a wide variety of variants within EFTUD2 that potentially inactivate one EFTUD2 allele (Gordon et al., 2012; Huang et al., 2016; Lacour et al., 2019; Lehalle et al., 2014; Lines et al., 2012; Luquetti et al., 2013; Matsuo et al., 2017; Sarkar et al., 2015; Smigiel et al., 2015; Vincent et al., 2016; Voigt et al., 2013). EFTUD2/Snu114 is a GTPase, and a core USN snRNP protein that is present throughout the splicing cycle and regulates spliceosome remodeling (Frazer, Nancollis, & O’Keefe, 2008). The MFDGA disease-associated variants comprise small and large deletions, splice site variants, and nonsense and missense variants. These MFDGA disease-associated variants are present in a single allele in trans with a wild type, functional allele, consistent with haploinsufficiency. It is not entirely clear why a reduction in the amount of a core pre-mRNA splicing protein, required for the splicing of all pre-mRNAs, results in such a specific disease phenotype. However, recent cell and animal models of MFDGA have begun to analyze the consequences of reduced EFTUD2 expression (Beauchamp et al., 2019; Deml, Reis, Muheisen, Bick, & Semina, 2015; Lei et al., 2017; Wood et al., 2019).

Of all the variants in EFTUD2, the missense variants are of particular interest as several of these variants have been suggested to disrupt EFTUD2 protein function (Huang et al., 2016), but have not been tested for their function experimentally. We took advantage of the high conservation between EFTUD2 and its orthologue in yeast, SNU114, to test the function in vivo of 19 EFTUD2 missense variants associated with MFDGA. Functional assays in yeast revealed that only four missense variants in SNU114, modeling EFTUD2 missense variants in MFDGA, disrupted protein function. The viability of many MFDGA related SNU114 missense variants in the yeast functional assay suggested that EFTUD2 missense variants influenced EFTUD2 function in a different way. In fact, by subsequently using a minigene (MG) splicing assay, we determined that five EFTUD2 missense variants influenced the splicing of the EFTUD2 pre-mRNA to inactivate one allele in MFDGA. Thus, we have defined how missense variants in EFTUD2 can influence both EFTUD2 protein function and pre-mRNA splicing to cause MFDGA and provide support for the growing evidence that missense variants can influence splicing and should be routinely tested for splicing defects.

2 | MATERIAL AND METHODS

2.1 | Yeast SNU114 mutagenesis and functional analysis

Mutagenic primers were designed to introduce missense variants that corresponded to orthologous MFDGA-associated missense mutations in EFTUD2 following pairwise alignment of protein sequences (EFTUD2 NP_001245282, Snu114 AJ556599.1) by EMBOSS Needle (Figure S1). Where necessary, optimal yeast codon usage was used. Mutagenesis of the SNU114 gene in the plasmid pRS413-SNU114 was carried out using the Kunkel method (Kunkel, 1985) and mutagenic primers (Table S1), followed by sequencing to confirm the mutation. The mutant or wild type plasmids were transformed into the haploid Snu114 deletion strain (YSNU114KO1 MATa; his3α1; leu2Δ0; lys2Δ0; ura3Δ0; YKL173w::kanMX4; pRS416-Snu114) then grown on 5-fluoroorotic acid (5-FOA) to assay for function as previously described (Frazer, Lovell, & O’Keefe, 2009).

2.2 | Splicing minigene construction

A 3.8 kb fragment ("Fragment 3") of the pSpliceExpress MG splicing reporter vector (a gift from Stefan Stamm, Addgene 32485) (Kishore, Khanna, & Stamm, 2008) was amplified by polymerase chain reaction (PCR) using Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific), or isolated by restriction enzyme digestion with Nhel and BamHI. Similarly, EFTUD2 exons and at least 100 bp of flanking 5’ and 3’ intronic sequence were PCR-amplified from control genomic DNA also using Phusion High-Fidelity DNA Polymerase employing two pairs of primers to produce two overlapping (10–20 bp) fragments ("Fragment 1" and "Fragment 2") each with a single portion of overlap (6–10 bp) with the vector fragment ("Fragment 3") on one end. Where necessary, overlapping primer sequences for Fragments 1 and 2 were altered to introduce single-nucleotide variations into the exons corresponding to MFDGA-associated EFTUD2 missense variants (see Table 1 for full list of variants). Fragments 1, 2, and 3 were assembled using the Gibson method (Gibson, 2011) and transformed into competent bacteria. Successfully assembled vectors were isolated from candidate colonies and their sequence-verified by direct Sanger sequencing performed by Eurofins Genomics. Sequences of primers used for Gibson assembly of MG splicing constructs can be found in Table S1.

For the MG splicing assay, HEK293 cells were grown overnight to 40–60% confluency in 3 ml of Dulbecco’s modified Eagle’s medium high-glucose, DMEM (Sigma-Aldrich), supplemented
with 10% fetal bovine serum (Sigma-Aldrich) in tissue-culture treated six-well plates at 37°C and with 5% CO2. Cells were transiently transfected with at least 0.2 μg of MG vector (either wild type or mutant) using Lipofectamine (Thermo Fisher Scientific) and the manufacturer's standard protocol. Following 48 hr incubation at 37°C with 5% CO2, RNA was extracted using TRI Reagent® according to the manufacturer's instructions (Sigma-Aldrich). Extracted RNA was purified further using the RNeasy column clean-up kit (Qiagen), which included a DNase digestion step. cDNA was synthesized from up to 4 μg RNA (using an equal amount of RNA for each sample set) using Superscript Reverse Transcriptase (Thermo Fisher Scientific). Resulting cDNA was amplified by Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific) using "Minigene RT PCR-for" and "Minigene RT PCR-rev" primers (Table S1). Finally, PCR products were run on an agarose gel (1-3%) supplemented with SafeView nucleic acid stain (NBS Biologicals). Gels were visualized under a blue-light transilluminator and, where appropriate, bands of interest were extracted and purified using QIAquick gel extraction kit (Qiagen) followed by direct Sanger sequencing performed by Eurofins Genomics, to confirm splicing products.

### 2.3 Ethical compliance

Institutional ethical review and approval was granted, and informed consent was provided, for all data obtained from patients.
3 | RESULTS

3.1 | Functional analysis of disease-associated EFTUD2 missense variants in yeast

There are currently over 100 different characterized genetic variants in EFTUD2 (https://databases.lovd.nl/shared/variants/EFTUD2(unique), the majority of which are linked to craniofacial anomalies consistent with a diagnosis of MFDGA (Gordon et al., 2012; Huang et al., 2016; Lehalle et al., 2014; Lines et al., 2012; Luqueti et al., 2013; Sarkar et al., 2015; Smigel et al., 2015; Vincent et al., 2016). Here, we present five new missense variants in EFTUD2 (see Supporting Information Notes). It is now well established that haploinsufficiency of EFTUD2 is the primary cause of craniofacial phenotypes in patients with MFDGA (Lines et al., 2012). However, it is less clear how each individual pathogenic variant leads to a loss-of-function in that allele and, as such, haploinsufficiency. While many of the known variants in EFTUD2 are large enough to be clearly deleterious (e.g., multixen deletion or whole gene deletion), approximately 15–20% of MFDGA patients present with only a single allele missense variant. The effect of these missense variants on EFTUD2 function has not been characterized. Several of the EFTUD2 missense variants are located close to, or within the EFTUD2 protein G-domain or one of five G-box motifs thought to be important for GTP-binding (Bartels, Urlaub, Lührmann, & Fabrizio, 2003) (Figure S1). However, several other missense variants are distal to any clearly functional EFTUD2 protein domain and may not influence EFTUD2 function. Alignment of the yeast Snu114 and human EFTUD2 amino acid sequences reveals that 13 of 20 amino acids where missense variants occur in EFTUD2 are identical in yeast Snu114 (Figure S1).

To investigate the effect of EFTUD2 missense variants on EFTUD2 protein function, comparable variants were made in the orthologous yeast gene SNU114 by in vitro mutagenesis of the HIS3 plasmid pRS413-SNU114. SNU114 missense containing plasmids were transformed into a SNU114 haploid deletion strain where SNU114 function was complemented with the URA3 plasmid pRS416-SNU114. Transformed strains were grown on plates containing 5-FOA to remove the pRS416-SNU114 plasmid, leaving only the missense variant form of yeast Snu114 as a single allele. Missense variants that caused a loss-of-function change in the yeast Snu114 protein would be lethal after growth in the presence of 5-FOA. Nineteen different EFTUD2 missense variants, plus one nonsense variant used as a lethal control, were tested for function in yeast SNU114. In addition to the nonsense variant control, three variants (Gly217Glu, Ala470Arg, Thr147Ile) were lethal and one variant (His218Arg) displayed slow growth (Table 1). None of the missense variants exhibited any temperature sensitivity difference compared to control samples and all variants exhibited a similar growth pattern when grown in liquid culture or on solid media (data not shown). The lack of a functional defect in yeast Snu114, containing orthologous EFTUD2-associated missense variants, suggests that some variants do not influence EFTUD2 protein function and might be disrupting another process, potentially the splicing of the EFTUD2 pre-mRNA.

3.2 | EFTUD2 missense variants disrupt splicing of the EFTUD2 pre-mRNA

Since a number of EFTUD2 missense variants were found to be viable in yeast SNU114, these missense variants were assayed to determine if they resulted in loss of function by influencing the splicing of EFTUD2 pre-mRNA. An important difference between EFTUD2 and SNU114 is that EFTUD2 is a multixen gene, whereas yeast SNU114 is a single exon gene. Therefore, any missense variants which alter a splice site or splicing enhancer/silencer motif in EFTUD2 would not have been deleterious in the processing of the yeast SNU114 mRNA.

Initial characterization of EFTUD2 missense mutations involved the use of multiple bioinformatics tools to predict the influence on splicing of each variant (Desmet et al., 2009; Fairbrother, Yeh, Sharp, & Burge, 2002; Lim & Fairbrother, 2012; Mort et al., 2014; Reese, Eckman, Kulp, & Haussler, 1997; Schwarz, Cooper, Schuele, & Seelow, 2014; Shapiro & Senapathy, 1987; Xiong et al., 2015; Yeo & Burge, 2004). Bioinformatic tools with disparate predictive algorithms were employed to provide a broad range of possible outcomes. Outputs from the in silico tools varied from quantitative scores in the case of "MutPred Splice" and "Spliceman" to more qualitative splice predictions of donor/acceptor site gain or loss or the introduction of exonic splicing enhancers (ESEs). All EFTUD2 variants that we tested were predicted to have some altered splicing characteristics from at least one of the nine software tools used (Tables 2 and S2). However, there was considerable discordance between the in silico predictors. For example, the variants Arg262Trp and Pro778Ser both had very low splicing score from MutPred Splice suggesting both were likely to be splice neutral variants and there were also no alterations to potential splice donor or acceptor sites predicted for these variants from five other tools. However, both Arg262Trp and Pro778Ser achieved a high ranking score from the Spliceman software suggesting these particular missense variants may alter or disrupt the normal splicing pattern. Conversely, there was good agreement from all software packages on the variant Asn286Ser and the likelihood of this variant altering the splicing of EFTUD2 exon 10.

In total, 19 EFTUD2 missense variants and one nonsense variant (negative control) covering 13 exons were tested using a MG assay in HEK293 cells. In addition, for two MFDGA-associated variants (Gly224Arg and Lys620Asn), a further pair of “synthetic” variant constructs were designed, whereby the specific nucleotide was changed to an alternative (non-wild type) nucleotide while maintaining the same amino acid substitution. These synthetic variants provided an additional test of the specific nucleotide in each variant in changing the splice pattern for their parent transcripts. Neither of the synthetic variants designed have previously been seen in population databases and were therefore predicted to be as deleterious as their equivalent patient-associated variants. For each EFTUD2 missense variant investigated, the appropriate exon with 100 nucleotides of upstream and downstream intron sequence was amplified from human genomic DNA, and the appropriate missense variant was incorporated into each MG construct by specific overlapping primer
TABLE 2  Predictive analysis of MFDGA-associated EFTUD2 missense mutations using nine different analysis programs and summarized results of MG splicing assay

| EFTUD2 variant | Genomic change | MutPred splice score/prediction | SPANR | SR (%) | R-ESE | MT | SSF | MES | NNSplice | HSF | MG assay results |
|----------------|----------------|---------------------------------|-------|--------|-------|----|-----|-----|---------|-----|------------------|
| NP_004238.3:-p.(T143I) | NC_000017.10:-g.42960525G>A | 0.86, SAV | 43 | Loss of ESE, gain of ESE | Acceptor loss | Acceptor increase | Acceptor increase | No effect |
| NP_004238.3:-p.(G207E) | NC_000017.10:-g.42957006C>T | 0.8, SAV | 57 | Gain of 3 ESEs | Acceptor loss | Acceptor decrease | Acceptor decrease | No effect |
| NP_004238.3:-p.(H208R) | NC_000017.10:-g.42957003T>C | 0.4, SNV | 62 | Loss of ESE | Donor gain | Donor loss | Acceptor increase | Donor gain | No effect |
| NP_004238.3:-p.(G224R) | NC_000017.10:-g.42956956C>T | 0.57, SNV | 57 | Loss of ESE | Donor gain | Acceptor decrease | Acceptor increase | Acceptor gain | No effect |
| NP_004238.3:-p.(R262W) | NC_000017.10:-g.42953387G>A | 0.17, SNV | 72 | Gain of ESE | No effect |
| NP_004238.3:-p.(Q383H) | NC_000017.10:-g.42945175C>G | 0.82, SAV | 66 | Loss of two ESEs | Donor gain | Donor increase | Donor increase | Donor increase, acceptor gain | Cryptic splice site |
| NP_004238.3:-p.(Q436E) | NC_000017.10:-g.42941130G>C | 0.22, SNV | 53 | Gain of ESE | Donor increase, acceptor gain | Acceptor loss | Acceptor loss | Acceptor decrease | No effect |
| NP_004238.3:-p.(C476R) | NC_000017.10:-g.42940262A>G | 0.12, SNV | 68 | Donor gain | Acceptor decrease | Acceptor increase | No effect |
| NP_004238.3:-p.(G499D) | NC_000017.10:-g.42940192C>T | 0.21, SNV | 41 | Gain of ESE | Donor gain | Acceptor increase | No effect |
| NP_004238.3:-p.(R578*) | NC_000017.10:-g.42937401G>A | 0.31, SNV | 87 | Decreased exon 18 inclusion | Acceptor increase | Acceptor increase | Acceptor increase | Acceptor increase | No effect (control) |
| NP_004238.3:-p.(K620N) | NC_000017.10:-g.42937273C>G | 0.8, SAV | 73 | Decreased exon 18 inclusion | Gain of ESE | Donor loss, acceptor gain | Donor decrease | Donor increase, acceptor loss | No effect, acceptor loss | Spliced out (majority) |
| EFTUD2 variant | Genomic change | MutPred splice score/prediction | SPANR (%) | R-ESE | MT | SSF | MES | NNSplice | HSF | MG assay results |
|---------------|----------------|-------------------------------|------------|-------|----|-----|-----|----------|-----|-----------------|
| NP_004238.3:-p.(L637R) | NC_000017.10:-g.42936500A>C | 0.24, SNV | 69 | Donor gain |
| NP_004238.3:-p.(T678K) | NC_000017.10:-g.42934455G>T | 0.26, SNV | 70 | Loss of ESE, gain of ESE | Donor increase | Acceptor gain | Acceptor gain |
| NP_004238.3:-p.(G769R) | NC_000017.10:-g.42931679C>G | 0.27, SNV | 64 | Gain of ESE | Donor increase, acceptor increase | Acceptor gain | Acceptor gain | Donor loss, acceptor loss |
| NP_004238.3:-p.(P778S) | NC_000017.10:-g.42931652G>A | 0.15, SNV | 68 | No effect |
| NP_004238.3:-p.(A823T) | NC_000017.10:-g.42930758C>T | 0.8, SAV | 44 | Donor gain, acceptor loss | Donor loss, acceptor decrease | Acceptor increase | Acceptor decrease |
| NP_004238.3:-p.(E829K) | NC_000017.10:-g.42930740C>T | 0.32, SNV | 55 | Gain of 3 ESEs | Donor increase | Acceptor gain | Acceptor increase |
| NP_004238.3:-p.(H856Y) | NC_000017.10:-g.42929926G>A | 0.92, SAV | 65 | Donor increase | Donor gain | Donor gain, acceptor increase | Donor gain, acceptor increase |
| NP_004238.3:-p.(R938H) | NC_000017.10:-g.42929088C>T | 0.2, SNV | 64 | Donor gain | Donor increase, acceptor increase | Donor increase, acceptor increase | Donor increase |
| NP_004238.3:-p.(G224R) (synth) | NC_000017.10:-g.42956956C>T | 0.35, SNV | 67 | Increased exon 9 inclusion | Loss of ESE, gain of ESE | Donor increase | Acceptor increase | Acceptor increase |
| NP_004238.3:-p.(K620N) (synth) | NC_000017.10:-g.42937237C>G | 0.81, SAV | 60 | Decreased exon 18 inclusion | Donor loss, acceptor gain | Donor decrease | Donor decrease, acceptor loss | Donor decrease, acceptor loss |

Note: Programs include SPANR (Xiong et al., 2015), MutPred Splice (Mort et al., 2014), SR (Lim & Fairbrother, 2012), R-ESE (Fairbrother et al., 2002), MT (Schwarz et al., 2014), SSF (Shapiro & Senapathy, 1987), MES (Yeo & Burge, 2004), NNSplice (Reese et al., 1997), and HSF (Desmet et al., 2009). Empty cells represent no result for that particular program. Refer to individual programs for details of predictive algorithms employed.

Abbreviations: HSF, Human Splicing Finder; MES, MaxEntScan; MFDGA, mandibulofacial dysostosis Guion-Almeida type; MG, minigene; MT, Mutation Taster; R-ESE, Rescue-ESE; SAV, splice-altering variant; SNV, splice-neutral variant; SR, Spliceman ranking; SSF, SpliceSiteFinder.
design during PCR amplification. MG fragments were then cloned into the pSpliceExpress MG plasmid (Kishore et al., 2008). All constructs were validated by Sanger sequencing before transfection into HEK293 cells and RNA extracted for analysis of MG splicing patterns.

Exon 9 of EFTUD2 contains the highest number (n = 3) of MFDGA-associated missense variants of any exon within the EFTUD2 transcript (Table 1). In addition, the mutation characteristics of the Gly224Arg variant provided an opportunity to design a synthetic variant that differed from both the wild type and patient-associated variant but would maintain a consistent amino acid translation with the variant form. Neither the MFDGA-associated variant Gly224Arg nor its synthetic replicate (Gly224Arg-synth) showed any influence on MG splicing (Figure 1a). The variants of Gly207Glu and His208Arg in Snu114 (Gly217Glu and His218Arg, respectively) had already shown a deleterious effect in the yeast growth assay, therefore, it was unsurprising that Gly207Glu and His208Arg showed no impact on the correct splicing of Exon 9 in the MG assay (Figure 1a) as the deleterious effect seen in EFTUD2 protein is sufficient to account for the loss of function needed to cause MFDGA.

Two variants in EFTUD2 exon 10 were tested by MG assay, Arg262Trp and a previously unpublished MFDGA-associated variant Asn286Ser. No splicing defect was seen with the Arg262Trp variant but the Asn286Ser variant resulted in a shorter spliced product (Figure 1b). Sequencing of this Asn286Ser variant shorter product revealed that it was produced by the activation of a cryptic splice site within exon 10, leading to a shortening of exon 10 by 17 nucleotides (Figure 1c). Interestingly, the Asn286Ser variant nucleotide (c.857A>G) forms the +5 position of the new splice site, a splice site position that has previously been shown to be frequently mutated in human diseases (Buratti et al., 2007). The predicted consequence of this alternate cryptic splice site in vivo is a frameshift that introduces two consecutive stop codons downstream of the new donor splice site, which is predicted to result in nonsense-mediated decay (NMD) of this aberrant transcript. The deleterious nature of the Asn286Ser variant is well predicted by the bioinformatics analysis with all five predictors of splice site alterations suggesting either an increase in donor strength (SpliceSiteFinder, MaxEntScan, Human Splicing Finder) or a gain of a donor site (Mutation Taster, NNSplice) at the variant position (Table 2). Mutpred Splice predicted Asn286Ser to be a splice-altering variant and Rescue-ESE predicted that Asn286Ser variant leads to the loss of two ESE motifs (Table 2). This ESE loss is consistent with the activation of the cryptic splice site seen in our MG assay for the Asn286Ser variant.

A single variant (Gln383His) in EFTUD2 exon 13 was analyzed for its influence on splicing compared to the wild type sequence using the MG assay. As expected, the wild type MG construct resulted in the correct splicing of exon 13, producing a RT-PCR product of 408 bp in length (Figure 1d). However, the Gln383His variant sequence altered splicing of the MG transcript. In total, three clear bands were visible (Figure 1d). The smallest product was a skipped-exon product. A middle band, which is larger than the wild type spliced product, could not be sequenced accurately by Sanger sequencing. However, most interestingly, the largest product was approximately 800 bp in length (Figure 1d). Direct sequencing of this approximately 800 bp product revealed that the Gln383His variant sequence leads to a decrease in the strength of the exon 13 donor splice site and retention of the downstream intron. This retained intron transcript is also present in the wild type sample, albeit at a much lower proportion than that seen in the Gln383His variant.

Next, exon 18 variants were investigated including Lys620Asn, its synthetic counterpart, and a nonsense variant associated with MFDGA (Arg578Ter) (Figure 1e). The reason behind inclusion of the nonsense variant was to act as a negative control for the splicing assay and to assess what consequence a premature termination codon (PTC) had on the stability of the transcript produced from the MG assay. As expected, the presence of the nonsense variant had no influence on the normal splicing of exon 18, resulting in similar levels of transcript to that seen for the wild type MG construct. Thus, the inclusion of a PTC, in this instance, does not lead to detectable levels of degradation of the transcript via the NMD pathway. Both Lys620Asn and Lys620Asn-Synth variants resulted in a different splice pattern compared to wild type (Figure 1e). In each case a lower band where exon 18 is entirely skipped, and a larger band, which could not be accurately sequenced, were observed. Since exon 18 is 141 bp in length, its exclusion from any mature mRNA transcripts in vivo would not lead to a frameshift. However, the essential EFG-domain 2 of EFTUD2 is located within exons 18 and 19 (Figure S1) and would be removed by the absence of exon 18. In fact, removal of the equivalent amino acids from the yeast Snu114 that are encoded by the human EFTUD2 exon 18 results in a lethal phenotype in yeast (Table 1).

Analysis of EFTUD2 exon 25 variants Ala823Thr and Glu829Lys by MG assay revealed that the wild type construct produced two splice products in roughly equal proportions (Figure 1f), one being a product where exon 25 was entirely included, (Figure 1f, top band) and one where exon 25 was entirely excluded (Figure 1f, bottom band). Both missense variants Ala823Thr and Glu829Lys produced a similar splice pattern to that of the wild type construct. However, the ratio of included/excluded (top/bottom) band intensity was altered by both variants. Variant Ala823Thr produced a higher proportion of the skipped exon product (bottom band), whereas Glu829Lys variant produced a higher proportion of transcripts where the exon was included (Figure 1f, top band). While there is no difference in the sequence of the splice products produced between the wild type and variant MGs, the alteration of relative transcript proportions suggests the locations of both missense mutations are sufficient to influence the splicing patterns of the parent exon. All remaining variants tested (across exons 5, 15, 16, 19, 20, 23, 26, and 27) showed no difference in splicing patterns from their respective wild type constructs (Table 1).
FIGURE 1  Splicing outcomes of EFTUD2 exons using minigene (MG) assay. Schematic representation of the MG construct of missense variants located in exon 9 (a), exon 10 (b,c), exon 13 (d), exon 18 (e), and exon 25 (f). (c) Schematic representation of splicing outcomes of variant Asn286Ser compared to wild type. Relative location of missense variant is designated by horizontal red line. Flanking endogenous MG exons are shown in gray. Number above each MG exon designates the length of exon (bp). Representative RT-PCR results for each construct and their predicted sizes are shown with the make-up of splice products depicted by small schematics on the right-hand side of RT-PCR gel images. Molecular weight markers and their sizes are provided at the left of each gel image. RT-PCR, reverse transcription polymerase chain reaction
4 | DISCUSSION

Of the EFTUD2 variants associated with the human craniofacial disorder MFDGA to date, over 15% are missense. Many of these missense variants in EFTUD2 are located at highly conserved residues between eukaryotes and have, therefore, previously been assumed to influence EFTUD2 protein function or stability, acting as null alleles. These missense null alleles would cause haploinsufficiency consistent with the majority of disease-associated EFTUD2 variants that are predicted to result in loss of function (Lines et al., 2012). However, none of the missense variants had been functionally characterized beyond bioinformatic predictive software or homology studies. Here, we assayed the function of 19 MFDGA-associated missense variants in yeast SNU114, as the EFTUD2 and Snu114 proteins are highly conserved. Contrary to the assumption that all EFTUD2 missense variants influence EFTUD2 protein function, only three of the 19 homologous missense variants we tested in Snu114 were lethal and a further one resulted in a slow growing phenotype. The positioning of the lethal variants Gly217Glu and Thr147Ile within well-conserved EFTUD2/Snu114 domains can explain their lethality. For instance, Thr147Ile is located within the G1-box domain, an area associated with GTP and magnesium binding. The importance of Snu114 Threonine 147 is well established as Thr147Ile impacts associated with GTP and magnesium binding. The importance of EFTUD2 Switch II domain and the “switch II” domain, two important sub-domains of the larger GTP-binding domain. However, the adjacent His218Arg variant, which exhibited slow growth in yeast, falls outside of the G3-box domain but still lies within the “Switch II” domain. Recent cryo-EM structural analysis of the yeast spliceosome tri-snRNP has revealed that Histidine 218 plays an important role in inhibiting GTPase activity of Snu114 through hydrogen bonding to Tyrosine 403 of the core U5 snRNP protein Prp8 (Nguyen et al., 2016). Surprisingly, the remaining variant that was found to be lethal in our yeast functional assays (Ala470Arg) is not located at a conserved residue or within a conserved structural domain associated with Snu114 GTP/GDP binding function. However, Ala470Arg is located within a domain homologous to domain II of the eukaryotic translational elongation factor EF-2. Additionally, the residue change from Alanine (hydrophobic) to Arginine (hydrophilic) may be sufficient to disrupt the local protein structure and influence the function of Snu114. Conversely, variant Gly491Asp is located at a highly conserved residue within a G-domain, yet this variant was viable in our yeast growth assay. Absence of functional complementation in these SNU114 variants in yeast may result from decreased activity of the Snu114 protein or decreased levels or stability of the variant protein. Additionally, variants that reduce Snu114/EFTUD2 protein activity could also influence splicing, in general, causing a larger molecular defect.

One possible explanation as to why some EFTUD2 missense variants in yeast Snu114 protein do not impact function could be that the missense variants in EFTUD2 may be influencing another process to cause haploinsufficiency, for example, by altering the splicing of the EFTUD2 pre-mRNA. There is growing evidence that some missense variants do not change protein function but instead impact the splicing of pre-mRNA (Cartegni, Chew, & Krainer, 2002; Soukarieh et al., 2016; Sterne-Weiler & Sanford, 2014). To address whether any of the EFTUD2 missense variants influenced splicing, we introduced individual exons with EFTUD2 wild type or missense variants into MG splicing constructs and analyzed splicing after transfection into human cells. We found that five missense variants did influence the splicing of the EFTUD2 pre-mRNA by either inducing exon skipping, intron retention, or activation of a cryptic splice site. In four of the five cases of altered splicing, the native reading frame is changed introducing a premature stop codon and most likely inducing NMD of the mRNA transcript. Although the remaining case (Lys620Asn) resulted in exon 18 skipping, it is not predicted to change the translational reading frame when the exon is skipped. However, any truncated protein that may be translated from EFTUD2 mRNA with a skipped exon 18 would be lacking a large proportion of its EFG-domain-2, which would significantly disrupt EFTUD2 protein function. Therefore, as Lysine 620 would no longer be present, a previous assumption that changes to Lysine 620 could disrupt protein function alone, as it is a surface facing basic residue potentially involved in protein-protein interactions (Huang et al., 2016), is now unlikely.

While we have found that some missense variants in EFTUD2 influence EFTUD2 protein function and others influence the splicing of the EFTUD2 pre-mRNA, there are still some EFTUD2 missense variants where it is still unclear how they may be causing inactivation of an allele and haploinsufficiency, including Gly224Arg, Arg262Trp, Gln436Glu, Gly499Asp, Leu637Arg, Thr678Lys, Gly769Arg, Pro778Ser, His856Tyr, and Arg938His. With the exception of Gly224Arg, Leu637Arg, Thr678Lys, and His856Tyr, the remaining six variants are all located at highly conserved residues but were not found to be deleterious in either our yeast growth assay or MG assay. Of particular interest is the variant Arg262Trp, which is currently the most commonly found MFDGA-associated missense variant and has been identified in three unrelated families (Huang et al., 2016; Lines et al., 2012; Smigiel et al., 2015). With these missense variants where no functional defect has been determined, either the yeast system is not revealing any functional defect for an amino acid change or the cell type used for the MG assay does not express the relevant splicing factors to reveal mis-splicing. Alternatively, the variant sequence may lead to a splicing defect affecting a distal part of the EFTUD2 transcript not included in the MG construct. Finally, missense variants could create a novel binding site for miRNA that could inhibit translation of the mRNA produced from that allele (Brummer & Hausser, 2014; Ni & Leng, 2015).

In conclusion, our results reveal that not only do missense mutations lead to splicing errors and erroneous mRNA transcripts, but also highlights how delicately balanced and dynamically regulated the mechanisms of splicing are. Our MG assay has determined that, even in a wild type construct of EFTUD2 exon splicing, more than one splicing outcome may be produced constitutively, as seen for the EFTUD2 exon 25 wild type MG construct (Figure 1). Additionally, alteration of a single nucleotide, often distal to the wild type
splice site, can lead to dramatic changes in the splicing patterns of that transcript. This influence of single nucleotide changes was particularly true for the Lys620Asn synthetic variant produced, which led to a change in the proportions of spliced products compared to its paired MFDGA-associated variant.

Our comparison of bioinformatic results from splicing predictors with our in vitro splicing assay results also demonstrates how varied and inconsistent many software packages are in predicting the splicing outcomes based on sequence information alone. While in silico predictions are useful tools for uncovering potentially pathogenic point mutations or implicating a possible mechanism of action, the importance of in vitro or in vivo lab-based assays cannot be underestimated, in particular in a medical setting where a functional diagnosis is often important to furthering our understanding of the underlying disease/disorder. As with a solely bioinformatic approach, both assays utilized here can accommodate a lack of patient material by replicating the genetic variant using a wild type reference, and this is of particular importance for medical conditions (such as MFDGA), which are both vanishingly rare and genetically heterogeneous.

ACKNOWLEDGMENTS
We would like to thank Clair Byrne and Katie Evans who were involved in some of the initial functional characterization of EFTUD2 missense variants. This study was supported by a grant from the BBSRC (BB/N000358/1). WGN is supported by the Manchester NIHR Biomedical Research Centre (IS-BRC-1215-20007). CG and JA are supported by Université Sorbonne Paris-Cité Pôle de recherche et d’enseignement supérieur (project number SPC/JFG/2013-031), the Agence Nationale de la Recherche (project CranioRespiro), E-Rare (project CRANIRARE) and MSDAvenir (Devo-Decode project).

CONFLICT OF INTERESTS
The authors declare that there are no conflict of interests.

DATA AVAILABILITY STATEMENT
The data that support the findings of this study are available from the corresponding author upon reasonable request. Variant sequence data for novel variants presented in Table 1 and Supporting Notes have been submitted to LOVD3 (https://www.lovd.nl/).

REFERENCES
Barfod, C., Urlaub, H., Lührmann, R., & Fabrizio, P. (2003). Mutagenesis suggests several roles of Snu114p in pre-mRNA splicing. Journal of Biological Chemistry, 278(30), 28324–28334.
Beauchamp, M. C., Djedid, A., Daupin, K., Clokie, K., Kumar, S., Majewski, J., & Jerome-Majewska, L. A. (2019). Loss of function mutation of Eftud2, the gene responsible for mandibulofacial dysostosis with microcephaly (MFDGM), leads to pre-implantation arrest in mouse. PLOS One, 14(7):e0219280. https://doi.org/10.1371/journal.pone.0219280
Brenner, T. J., & Guthrie, C. (2005). Genetic analysis reveals a role for the C terminus of the Saccharomyces cerevisiae GTPase Snu114 during spliceosome activation. Genetics, 170(3), 1063–1080.
Brummer, A. & Haussler, J. (2014). MicroRNA binding sites in the coding region of mRNAs: Extending the repertoire of post-transcriptional gene regulation. BioEssays, 36(6), 617–626. https://doi.org/10.1002/bies.201300104
Buratti, E., Chivers, M., Kralovicova, J., Romano, M., Baralle, M., Krainer, A. R., & Vorechovsky, I. (2007). Aberrant splicing of human disease genes: Mutation pattern, nucleotide structure and comparison of computational tools that predict their utilization. Nucleic Acids Research, 35(13), 4250–4263. https://doi.org/10.1093/nar/gkm402
Cartegni, L., Chew, S. L., & Krainer, A. R. (2002). Listening to silence and understanding nonsense: Exonic mutations that affect splicing. Nature Reviews Genetics, 3(4), 285–298.
Deml, B., Reis, L. M., Muheisen, S., Bick, D., & Semina, E. V. (2015). EFTUD2 deficiency in vertebrates: Identification of a novel human mutation and generation of a zebrafish model. Birth defects research. Part A, Clinical and molecular teratology, 103(7), 630–640. https://doi.org/10.1002/bdra.23397
Desmet, F. O., Hamroun, D., Lalande, M., Collod-Beroud, G., Claustres, M., & Beroud, C. (2009). Human Splicing Finder: An online bioinformatics tool to predict splicing signals. Nucleic Acids Research, 37(9), e67. https://doi.org/10.1093/nar/gkp215
Fairbrother, W. G., Yeh, R. F., Sharp, P. A., & Burge, C. B. (2002). Predictive identification of exonic splicing enhancers in human genes. Science, 297(5583), 1007–1013. https://doi.org/10.1126/science.1073774
Frazer, L. N., Lovell, S. C., & O’Keefe, R. T. (2009). Analysis of synthetic lethality reveals genetic interactions between the GTPase Snu114p and snRNAs in the catalytic core of the Saccharomyces cerevisiae spliceosome. Genetics, 183, 497–515.
Frazer, L. N., Nancollis, V., & O’Keefe, R. T. (2008). The role of Snu114p during pre-mRNA splicing. Biochemical Society Transactions, 36(Pt 3), 551–553.
Gibson, D. G. (2011). Enzymatic assembly of overlapping DNA fragments. Method Enzymol., 498, 349–361. https://doi.org/10.1016/B978-0-12-385120-8.00015-2
Gordon, C. T., Petit, E., Oufadem, M., Decaestecker, C., Jourdain, A. S., Andrieux, J., ... Amiel, J. (2012). EFTUD2 haploinsufficiency leads to syndromic oesophageal atresia. Journal of Medical Genetics, 49(12), 737–746. https://doi.org/10.1136/jmedgenet-2012-101173
Huang, L. J., Vanstone, M. R., Hartley, T., Osmond, M., Barrowman, N., Allanson, J., ... Consortium, C. R. C. (2016). Mandibulofacial dysostosis with microcephaly: Mutation and database update. Human Mutation, 37(2), 148–154. https://doi.org/10.1002/humu.22924
Kishore, S., Khanna, A., & Stamm, S. (2008). Rapid generation of splicing reporters with pSpliceExpress. Gene, 427(1-2), 104–110. https://doi.org/10.1016/j.gene.2008.09.021
Kunkel, T. A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. Proceedings of the National Academy of Sciences of the United States of America, 82, 488–492.
Lacour, J. C., McBride, L., St Hilaire, H., Mundinger, G. S., Moses, M., Koon, J., ... Lacassie, Y. (2019). Novel de novo EFTUD2 mutations in 2 cases with MFDGM, initially suspected to have alternative craniofacial
Sarkar, A., Emrick, L. T., Smith, E. M., Austin, E. G., Yang, Y. P., Hunter, J. V., Reese, M. G., Eeckman, F. H., Kulp, D., & Haussler, D. (1997). Improved sequence-based prediction of exonic variants that disrupt splicing. Bioinformatics, 13(8), R19.

Lim, K. H., & Fairbrother, W. G. (2012). Splecem—A computational web server that predicts sequence variations in pre-mRNA splicing. Bioinformatics, 28(7), 1031–1032. https://doi.org/10.1093/bioinformatics/bts074

Lines, M. A., Huang, L., Schwartzenbruber, J., Douglas, S. L., Lynch, D. C., Beaulieu, C., ... Boycott, K. M. (2012). Haploinsufficiency of a spliceosomal GTPase encoded by EFTUD2 causes mandibulofacial dysostosis with microcephaly. American Journal of Human Genetics, 90(2), 369–377. https://doi.org/10.1016/j.ajhg.2011.12.023

Luqueti, D. V., Hing, A. V., Rieder, M. J., Nickerson, D. A., Turner, E. H., Smith, J., ... Cunningham, M. L. (2013). “Mandibulofacial dysostosis with microcephaly” caused by EFTUD2 mutations: Expanding the phenotype. American Journal of Medical Genetics, 161A(1), 108–113. https://doi.org/10.1002/ajmg.a.35696

Matsuo, M., Yamauchi, A., Ito, Y., Sakauchi, M., Yamamoto, T., Okamoto, N., ... Saito, K. (2017). Mandibulofacial dysostosis with microcephaly: A case presenting with seizures. Brain and Development, 39(2), 177–181. https://doi.org/10.1016/j.braindev.2016.08.008

Mort, M., Sterne-Weiler, T., Li, B., Ball, E. V., Cooper, D. N., Radiojoval, P., ... Mooney, S. D. (2014). MutPred splice: Machine learning-based prediction of exonic variants that disrupt splicing. Genome Biology, 15(1), R19. https://doi.org/10.1186/gb-2014-15-1-r19

Nguyen, T. H. D., Galej, W. P., Bai, X. C., Oubridge, C., Newman, A. J., Scheres, S. H. W., & Nagai, K. (2016). Cryo-EM structure of the yeast U4/U6/U5 tri-snRNP at 3.7 angstrom resolution. Nature, 530(7590), 298–304. https://doi.org/10.1038/nature16940

Ni, W. J., & Leng, X. M. (2015). Dynamic miRNA-mRNA paradigms: New faces of miRNAs. Biochemistry and Biophysics Reports, 4, 337–341. https://doi.org/10.1016/j.bbrep.2015.10.011

Reese, M. G., Eeckman, F. H., Kulp, D., & Haussler, D. (1997). Improved splice site detection in Genie. Journal of Computational Biology, 4(3), 311–323. https://doi.org/10.1089/cmb.1997.4.311

Sarkar, A., Emrick, L. T., Smith, E. M., Austin, E. G., Yang, Y. P., Hunter, J. V., ... Lalani, S. R. (2015). Novel de novo mutations in EFTUD2 detected by exome sequencing in mandibulofacial dysostosis with microcephaly syndrome. American Journal of Medical Genetics, 167(4), 914–918. https://doi.org/10.1002/ajmg.a.36948

Schwarz, J. M., Cooper, D. N., Schuelke, M., & Seelow, D. (2014). MutationTaster2: Mutation prediction for the deep-sequencing age. Nature Methods, 11(4), 361–362. https://doi.org/10.1038/nmeth.2890

Shapiro, M. B., & Senapathy, P. (1987). RNA splice junctions of different classes of eukaryotes: Sequence statistics and functional implications in gene expression. Nucleic Acids Research, 15(17), 7155–7174. https://doi.org/10.1093/nar/15.17.7155

Smigielski, R., Beznikow, N., Jakubiak, A., Bloch, M., Patkowski, D., Obersztyn, E., & Sasiadek, M. M. (2015). Phenotype analysis of Polish patients with mandibulofacial dysostosis type Guion-Almeida associated with esophageal atresia and choanal atresia caused by EFTUD2 gene mutations. Journal of Applied Genetics, 56(2), 199–204. https://doi.org/10.1007/s13353-014-0255-4

Soukarieh, O., Gaidrat, P., Hamieh, M., Drouet, A., Baert, S., Martins, A. (2016). Exoninc splicing mutations are more prevalent than currently estimated and can be predicted by using in silico tools. PLOS Genetics, 12(1):e1005756. https://doi.org/10.1371/journal.pgen.1005756

Sterne-Weiler, T., & Sanford, J. R. (2014). Exon identity crisis: Disease-causing mutations that disrupt the splicing code. Genome Biology, 15(1), 1–8. https://doi.org/10.1186/gb1450, Art 201.

Vincent, M., Genevieve, D., Ostertag, A., Marlin, S., Lacombe, D., Martin-Colognard, D., ... Collet, C. (2016). Treacher Collins syndrome: A clinical and molecular study based on a large series of patients. Genetics in Medicine, 18(1), 49–56. https://doi.org/10.1038/gim.2015.29

Voigt, C., Megarbane, A., Neveling, K., Czeschik, J. C., Albrecht, B., Callewaert, B., ...Wieczorek, D. (2013). Oto-facial syndrome and esophageal atresia, intellectual disability and zygotomic anomalies—Expanding the phenotypes associated with EFTUD2 mutations. Orphanet Journal of Rare Diseases, 8, 110. https://doi.org/10.1186/1750-1172-8-110

Wahl, M. C., Will, C. L., & Lührmann, R. (2009). The spliceosome: Design principles of a dynamic RNP machine. Cell, 136(4), 701–718.

Will, C. L., & Lührmann, R. (2011). Spliceosome structure and function. Cold Spring Harbor Perspectives in Biology, 3(7), a003707. https://doi.org/10.1101/cshperspect.a003707

Wood, K. A., Rowlands, C. F., Qureshi, W. M. S., Thomas, H. B., Bucek, W. A., Briggs, T. A., ... O’Keefe, R. T. (2019). Disease modelling of core pre-mRNA splicing factor haploinsufficiency. Human Molecular Genetics, 28, 3704–3723. https://doi.org/10.1093/hmg/ddz169

Xiong, J. W. (2017). Expanding the phenotypes associated with EFTUD2 mutations. Human Mutation, 38(7), a1103756. https://doi.org/10.1002/humu.23517

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Thomas HB, Wood KA, Bucek WA, et al. EFTUD2 missense variants disrupt protein function and splicing in mandibulofacial dysostosis Guion-Almeida type. Human Mutation. 2020;41:1372–1382. https://doi.org/10.1002/humu.24027