bioRxiv preprint doi: https://doi.org/10.1101/064196; this version posted July 15, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Functional genetic characterization by CRISPR-Cas9 of two enhancers of *FOXP2* in a child with speech and language impairment

Torres-Ruiz, R.[1], Benítez-Burraco, A.[2,3], Martínez-Lage, M.[1], Rodríguez-Perales, S. [1,*], and García-Bellido, P.[3,4,*]

1. Molecular Cytogenetics Group. Centro Nacional Investigaciones Oncológicas (CNIO), Madrid, Spain.

2. Department of Philology, University of Huelva, Huelva, Spain.

3. Faculty of Modern Languages, University of Oxford, Oxford, United Kingdom.

4. Faculty of Linguistics, Philology and Phonetics, University of Oxford, Oxford, United Kingdom.

* Corresponding authors: Sandra Rodriguez-Perales: srodriguezp@cnio.es; Phone: (+34) 912 246 900; Fax: (+34) 912 246 911. Paloma Garcia-Bellido: paloma.garcia-bellido@mod-langs.ox.ac.uk Phone (+44) 1865 27049;Fax: (+44) 1865 270757.

ABSTRACT

Mutations in the coding region of the *FOXP2* transcriptor factor gene are known to cause speech and language impairment. Chromosomal rearrangements with breakpoints downstream the gene have been hypothesised to impair speech and cognitive abilities via physical separation of distant regulatory DNA elements. In this study, we used highly efficient targeted chromosomal deletions induced by the CRISPR/Cas9 genome editing tool to characterise two functional enhancers: FOXP2-$E^{proximal}$ and FOXP2-E^{distal}, located in the intergenic region between *FOXP2* and its adjacent *MDFIC* gene. FOXP2-E^{distal}, separated from *FOXP2* by a chromosomal rearrangement in a case of speech and language impairment, was demonstrated to be functional in a luciferase assay. Deletion of any of these two functional enhancers in a neuroectodermal tumor cell-line downregulates *FOXP2* and decreases FOXP2 protein levels, conversely it upregulates *MDFIC* and increases *MDFIC* protein levels. We expect these findings contribute to a deeper understanding of *FOXP2* and *MDFIC* may pace the development of speech and language in the brain.

KEYWORDS

FOXP2, MDFIC, speech and language impairment, Spanish, CRISPR-Cas9, enhancer

bioRxiv preprint doi: https://doi.org/10.1101/064196; this version posted July 15, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

ACKNOWLEDGMENTS

This project was supported by funds from The University of Oxford John Fell OUP Research Grant [121/435] awarded to its Principal Investigator Paloma Garcia-Bellido. This study was supported in part by funds from the Spanish National Research and Development Plan, Instituto de Salud Carlos III, and FEDER (FIS project no. PI14/01884 to Sandra Rodriguez-Perales), and in part by funds from the Spanish Ministry of Economy and Competitiveness (grant numbers FFI2014-61888-EXP and FFI-2013-43823-P to Antonio Benítez-Burraco). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

INTRODUCTION

Mutations in *FOXP2* gene are known to cause speech and language impairment (Vargha-Khadem et al. 2005; Zhao et al. 2010). Polymorphisms of the gene have been also associated to schizophrenia (Tolosa et al. 2010) and frontotemporal lobar degeneration (Padovani et al. 2010). *FOXP2* has been hypothesised to regulate the development and function in the adult state of brain areas involved in human language processing (Lai et al., 2003; Fisher and Scharff 2009), because of its known role in neurogenesis, neuron differentiation and migration patterns in the developing telencephalon in mice (Tsui et al. 2013; Chiu et al. 2014; Garcia-Calero et al. 2016). Pathogenic mutations in humans have proven to impair auditory-motor association learning in mice (Kurt et al. 2012). Nonetheless, the exact role of *FOXP2* in normal development is unknown. Common variants of the gene do not contribute appreciably to individual differences in language development (Mueller et al. 2016), nor in brain structure (Hoogman et al. 2014), although a *FOXP2* polymorphism has been associated with enhanced procedural learning of non-native speech sound categories (Chandrasekaran et al. 2015). Less is known about how the expression of the gene is modulated. The promoter of *FOXP2* contains four transcription start sites (Schroeder and Myers 2008), with multiple alternative splicing sites (Bruce and Margolis 2002). *FOXP2* also contains six ultraconserved regions in its introns (Schroeder and Myers 2008), as well as enhancers for LEF1, a transcription factor that drives expression of the gene in the central nervous system during embryogenesis (Bonkowsky et al. 2008). Interestingly, several microRNAS bind the 3'UTR of the gene and regulate the expression of *FOXP2* (Clovis et al. 2012; Shi et al. 2013; Fu et al. 2014b; Cuiffo et al. 2014).

bioRxiv preprint doi: https://doi.org/10.1101/064196; this version posted July 15, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Apart from gene mutations, chromosomal translocations involving the FOXP2- MDFIC intergenic region have been also associated to speech problems and cognitive impairment (Kosho et al. 2008). Microdeletions involving *FOXP2* and/or *MDFIC* and the region between these two genes have been found in subjects with speech delay and cognitive impairment (DECIPHER patients 262086, 292652, and 301696). We have recently reported on a young female harbouring a genomic complex rearrangement involving chromosomes 7 and 11, who presents with severe expressive and receptive speech and language impairment in both Castilian Spanish and Valencian (Moralli et al. 2015). Although the *FOXP2* coding region is intact, the breakpoint in 7q31.1 is located 205.5 kb downstream the 3' end of *FOXP2* and 22.8 kb upstream the 5' region of *MDFIC*. Becker et al. (2015) found and characterized a functional enhancer located 2.5 kb downstream the breakpoint. In our proband this element was maintained in chromosome 7, whereas *FOXP2* was rearranged to chromosome 11. A more robust approach, aimed at looking for changes in the expression level of the gene seems desirable in order to know if this enhancer regulates *FOXP2* expression.

The development of nuclease mediated genome editing tools, specially, of those based on clustering regularly interspaced short palindromic repeats (CRISPR) (Sakuma and Woltjen 2014; Torres-Ruiz and Rodriguez-Perales 2016), has emerged as a highly efficient way of inducing targeted chromosomal deletions and an accurate method to validate the functionality of enhancers (Cong et al. 2013; Mali et al. 2013). Here we report a detailed study of the intergenic region between the *FOXP2* and *MDFIC* genes. We have found that this region contains, apart from the enhancer reported in Becker et al. (2015), a second functional enhancer, FOXP2-E ^{proximal}. We performed targeted deletions of each regulatory element by CRISPR-Cas9 and found that both affect the expression levels of *FOXP2* and *MDFIC* in an opposite manner. We hypothesise therefore that the breakpoint in this case would cause *FOXP2* to be anomalously downregulated by the separation of FOXP2-^{distal} from *FOXP2*, while *MDFIC* to be anomalously upregulated by the separation of FOXP2-^{proximal} from *MDFIC*. These changes in the expression levels of these two genes may account for the observed language deficits in this case. We expect these findings contribute to a better understanding of how *FOXP2* is regulated.

bioRxiv preprint doi: https://doi.org/10.1101/064196; this version posted July 15, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

MATERIALS AND METHODS

Cell culture and electroporation

Cells of the non neuronal cell-line HEK293A (CRL-1573, ATCC, USA) and the neuronal cell-line SK-N-MC (HTB-10, ATCC, USA) were maintained under standard conditions in Dulbecco's modified Eagle's medium (DMEM) (Lonza), supplemented with 10% foetal bovine serum (FBS) (Life Technologies), 1% GlutaMAX (Life Technologies), and 10mg/ml penicillin/streptomycin (Life Technologies). The neuronal cell-line SK-SY-5Y (CRL-2266, ATCC, USA) was cultured in a 1:1 mixture of Dulbecco's modified Eagle's medium (DMEM) (Lonza) and F12 medium (Lonza) supplemented with 10% FBS (Life Technologies), 1% GlutaMAX (Life Technologies), and 10mg/ml penicillin/streptomycin (Life Technologies). Cells were cultured at 37ºC in a humidified incubator in a 5% CO_2 + 20% O_2 atmosphere.

For electroporation, we used the Neon Transfection System (Life Technologies). The manufacturer's protocols for HEK293A, SK-N-MC and SH-SY5Y cells were modified as follows. The three cell types were electroporated at 80% confluence. Cells were trypsinized and resuspended in R solution (Life Technologies). For SK-N-MC and SH-SY5Y, 10-µl tips were used to electroporate 2.5×10^6 cells with a single 50-ms pulse of 900 V. For HEK293A cells, 4×10^5 cells were electroporated with 10-µl tips using three 10-ms pulses of 1245 V. After electroporation, cells were seeded in a 24-well plate containing pre-warmed medium. When required, cells were sorted 72 h post-transfection.

Construction of Double-Guide Cas9-Encoding Plasmids

The parental pLV-U6$^{\#1}$H1$^{\#2}$-C9G vector has been described elsewhere (Torres et al. 2014b). Two gBlocks gene fragments were synthesized to clone sgRNA#1 and sgRNA#2 flanking the FOXP2-E^{proximal} and FOXP2-E^{distal} enhancer regions in the backbone vector using BsrGI and SpeI target sites.

Flow Cytometry and Cell Sorting

72 hours after electroporation, cells were trypsinized and washed with DPBS twice, counted, and resuspended in an appropriate volume of sorting buffer (PBS containing 1% FBS and antibiotics) for flow cytometry analysis. Immediately before cell sorting, samples were filtered through a 70-µm filter to remove any clumps or aggregates. Cell sorting was carried out in a Synergy 2L instrument (Sony Biotechnology

bioRxiv preprint doi: https://doi.org/10.1101/064196; this version posted July 15, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Inc.); flow cytometry was performed in a BD LSR Fortessa analyzer (BD Biosciences). Cells were sorted and seeded individually per well in a 96 well-plate.

Genomic DNA Extraction and PCR Analysis

Genomic DNA was extracted using standard procedures (Torres et al. 2014a). Briefly, 5-10×10^6 cells were either trypsinized or scraped, washed in PBS, pelleted, and lysed in 100mM NaCl, Tris (pH 8.0) 50mM, EDTA 100mM, and 1% SDS. After overnight digestion at 56ºC with 0.5 mg/ml of proteinase K (Roche Diagnostics), the DNA was cleaned by precipitation with saturated NaCl, and the clear supernatant was precipitated with isopropanol, washed with 70% ethanol, air-dried, and resuspended overnight at room temperature in 1xTE buffer. Serial DNA dilutions were quantified with a NanoDrop ND 1000 Spectrophotometer (NanoDrop Technologies).

Standard PCR was performed in a Veriti 96-well Thermal Cycler (Applied Biosystems) under the following conditions: template denaturation at 95°C for 1 min followed by 30 cycles of denaturation at 94°C for 30 s, annealing at 62.5°C for 30 s, extension at 72°C for 60 s, and a final extension of 5 min at 72°C. Primers used are listed in Supplementary Table 1.

RNA extraction and PCR

Total RNA was extracted from tissues and cell cultures using Trizol (Sigma-Aldrich), followed by treatment with RNase-free DNAse (Roche Applied Science). cDNA was synthesized from 500 ng of total RNA using the Superscript III First Strand cDNA Synthesis Kit (Life Technologies). Specific mRNAs were quantified by qRT-PCR using an ABI Prism 7900 HT Detection System (Applied Biosystems) and TaqMan detection. PCR was performed in 96-well plate microtest plates with TaqMan master mix (Thermo Fisher) for 40 cycles. In all experiments, mRNA amounts were normalized to the total amount of cDNA by using amplification signals for hGUSB. Each sample was determined in triplicate, and at least three independent samples were analysed for each experimental condition.

Western Blot

Proteins were extracted by standard procedures as previously described (Rodriguez-Perales et al. 2015) in the presence of Complete Protease Inhibitor Cocktail Tables (Roche Applied Science). Proteins were

bioRxiv preprint doi: https://doi.org/10.1101/064196; this version posted July 15, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

transferred to PVDF using TransFi (Invitrogen; Life Technologies), and membranes were probed for FOXP2 or MDFIC with monoclonal mouse anti-human FOXP2 or MDFIC antibodies (1/1000 or 1/500; BD Pharmigen) or for GAPDH (AbCam), with antibodies diluted 1/2500 in PBS/0.1% Tween-20 (PBS-T). Secondary antibodies were HRP-conjugated with goat anti-mouse IgG (1/1000) and goat anti-Rabbit (1/500; Dako, Barcelona, Spain), and blots were developed with ECL (GE Healthcare).

RESULTS

In silico search of enhancer regions

We first hypothesised that the breakpoint in 7q31.1 (114,888,284 hg38) affected the expression of *FOXP2* by physically separating some cis-acting distant element with an enhancer role. Accordingly, we searched in silico for putative enhancers in the intergenic region between *FOXP2* and *MDIFC* looking for the following hallmarks: DNAsa clusters, presence of histones with specific post-translational modifications (specifically histone H3, lysine 4 monomethylation (H3K4me1) and H3 lysine 27 acetylation (H3K27ac)), and ChIP-seq data provided by ENCODE of regions recruiting co-activators and co-repressors as revealed by chromatin immunoprecipitation followed by deep sequencing. We found two putative enhancers located at 120kb and 203.5kb downstream the end of *FOXP2*, respectively (Figure 1A and 1B). These putative enhancers (referred hereafter as FOXP2-$E^{proximal}$ and FOXP2-E^{distal}) span 6264bp (chr7:114,817,431-114,823,694 hg38) and 2300bp (chr7:114,900,989-114,903,302 hg38 equivalent to 114,541,370-114,542,201 hg19), respectively. FOXP2-E^{distal} is the one previously validated by luciferase assay by Becker et al. (2015); FOXP2-$E^{proximal}$ is a new putative regulatory element.

CRISPR deletion of FOXP2-$E^{proximal}$ and FOXP2-E^{distal}

We then tested in vitro the functionality of FOXP2-$E^{proximal}$ and FOXP2-E^{distal}. Since both putative enhancers are located in an intergenic region, we aimed at characterizing that both of them are functional with respect to *FOXP2* or/and *MDFIC*. We relied on a CRISPR genome editing approach to delete the entire predicted sequence of each enhancer. Accordingly, we designed two couples of sgRNAs targeting the flanking regions of either FOXP2-$E^{proximal}$ or FOXP2-E^{distal} (Figure 1C). Each sgRNA pair was cloned in the pLV-U6$^{\#1}$H1$^{\#2}$-C9G (Torres et al. 2014b) in order to couple the expression of the sgRNAs to the expression of Cas9 and a GFP reporter. Afterwards, we tested if the sgRNAs were able to induce the

bioRxiv preprint doi: https://doi.org/10.1101/064196; this version posted July 15, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

expected deletions. HEK293A cells were nucleofected with 2ug of pLV-U6$^{\#1}$H1$^{\#2}$-C9G plasmid targeting either FOXP2-E^{proximal} or FOXP2-E^{distal}. After 72 h, the DNA was isolated and analyzed. After designing PCR oligos that span the deleted regions (Supplementary Table 1), PCR assays were performed. They revealed efficient targeted deletions of the 6.2kb or the 2.3kb regions containing the entire sequence of FOXP2-E^{proximal} and FOXP2-E^{distal}, respectively (data not shown).

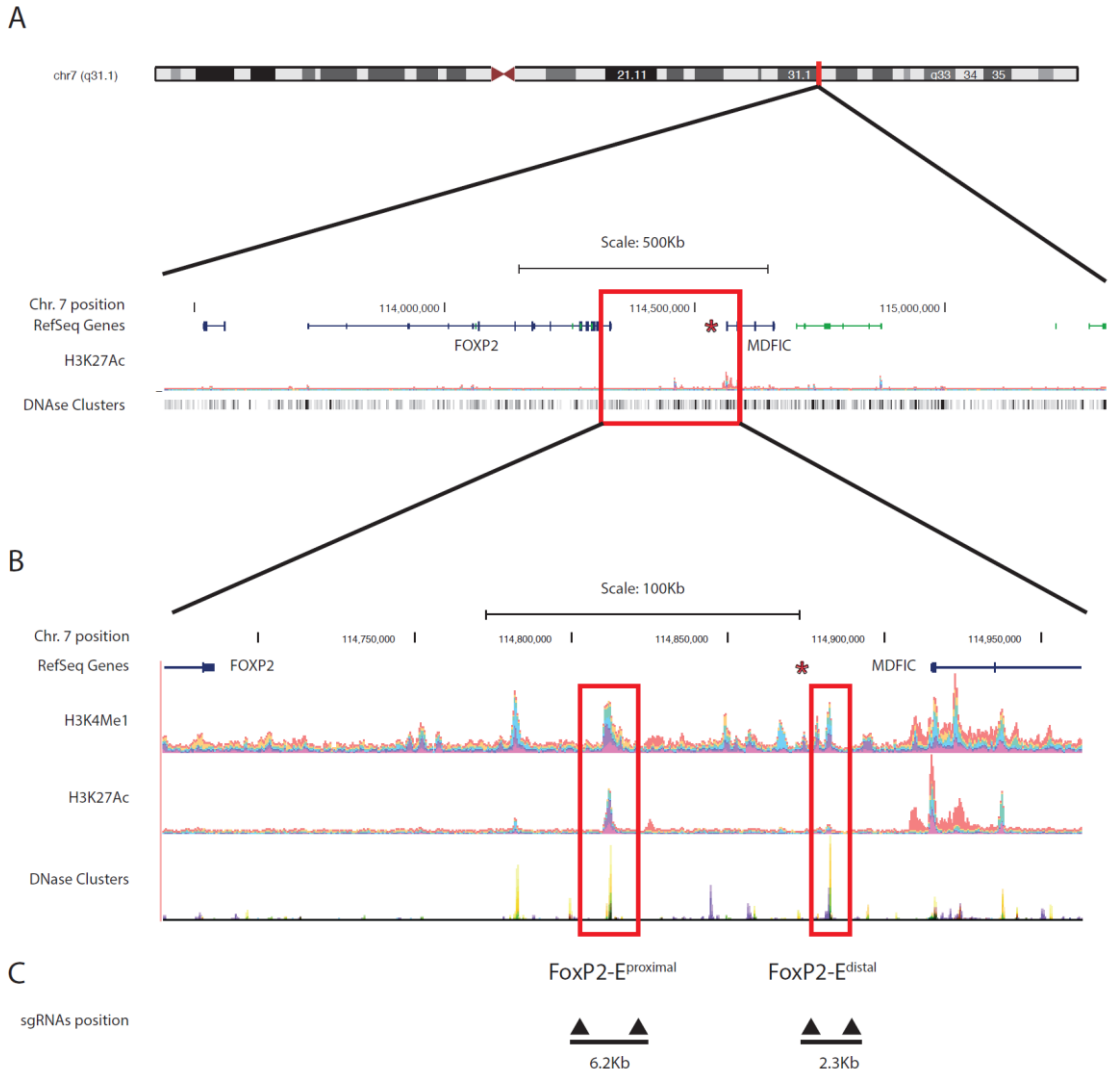

Figure 1. Identification of enhancer regions downstream *FOXP2* and upstream *MDFIC*. A. Genomic location of human *FOXP2* and *MDFIC* genes. Genes are depicted in blue. **B.** Detailed view of an Encode UCSC genome-browser snapshot showing bar graphs with a detailed representation of the locations of H3K4Me1 and H3K27Ac histone marks in human neurological cell lines. The squared regions in red show the locations of FOXP2-E^{proximal} and FOXP2-E^{distal}. The red asterisk shows the position of the breakpoint. **C.** Schematic representation of the location of he four sgRNAs flanking the 6.2kb region including FOXP2-E^{proximal} and the 2.3kb region including FOXP2-E^{distal}.

bioRxiv preprint doi: https://doi.org/10.1101/064196; this version posted July 15, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Neuronal cell lines defective in FOXP2-E^{proximal} or FOXP2-E^{distal}

We next used CRISPR to delete FOXP2-E^{proximal} or FOXP2-E^{distal} in two cell lines: SK-N-MC, a metastatic cell line derived from the supra-orbital area, which expresses *FOXP2* constitutively (although it does not express *MDFIC* at the same level), and SH-SY5Y, a neuroblastomic cell line which expresses neither *FOXP2* nor *MDFIC*. Cells were electroporated with 2ug of either pLV-U6$^{\#1}$H1$^{\#2}$-C9G- E^{proximal}, pLV-U6$^{\#1}$H1$^{\#2}$-C9G- E^{distal}, or with empty plasmids. After 72h the DNA was extracted and analysed. PCR and Sanger sequencing analyses confirmed the deletion of the 6.2kb or the 2.3kb fragment (Figures 2A and 2B). We then generated two clonal cell lines (one for each putative enhancer) by sorting GFP positive cells into 96-well plates for single cell colony expansion (data not shown). We confirmed by PCR that the cellular clones had expanded. These two cell lines were used for further expression analyses.

FOXP2 and *MDFIC* expression analyses

We next aimed to characterize in more detail the functionality of FOXP2-E^{proximal} and FOXP2-E^{distal}. We used RT-qPCR to determine the amount of *FOXP2* mRNA in the SK-N-MC cells transduced with either pLV-U6$^{\#1}$H1$^{\#2}$-C9G-E^{proximal}, or pLV-U6$^{\#1}$H1$^{\#2}$-C9G-E^{distal}, or an empty plasmid as a control. The expression of *FOXP2* was significantly reduced (2.9 fold change) compared to that of the control when FOXP2-E^{proximal} was deleted (Figure 2C, left). Likewise, *FOXP2* expression was decreased (2 fold change) when FOXP2-E^{distal} was deleted (Figure 2C left). We then measured the levels of expression of *MDFIC* after deletion of each enhancer. As shown in Figure 2C right, the expression of *MDFIC* was significantly increased when either FOXP2-E^{proximal} or FOXP2-E^{distal} were deleted (8.6 and 7.5 fold change, respectively). The experiment was replicated in SH-SY5Y cells, but no change of expression of *FOXP2* or *MDFIC* was detected (data not shown).

We next analysed by Western blot the amount of FOXP2 and MDFIC proteins in the SK-N-MC cells transduced with either pLV-U6$^{\#1}$H1$^{\#2}$-C9G-E^{proximal}, or pLV-U6$^{\#1}$H1$^{\#2}$-C9G-E^{distal}, or an empty plasmid, used as a control (Figure 2D). In line with our mRNA data, the deletion of FOXP2-E^{proximal} or FOXP2-E^{distal} was found to reduce the level of FOXP2 (Figure 2D top) and to increase the level of MDFIC (Figure 2D bottom).

bioRxiv preprint doi: https://doi.org/10.1101/064196; this version posted July 15, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

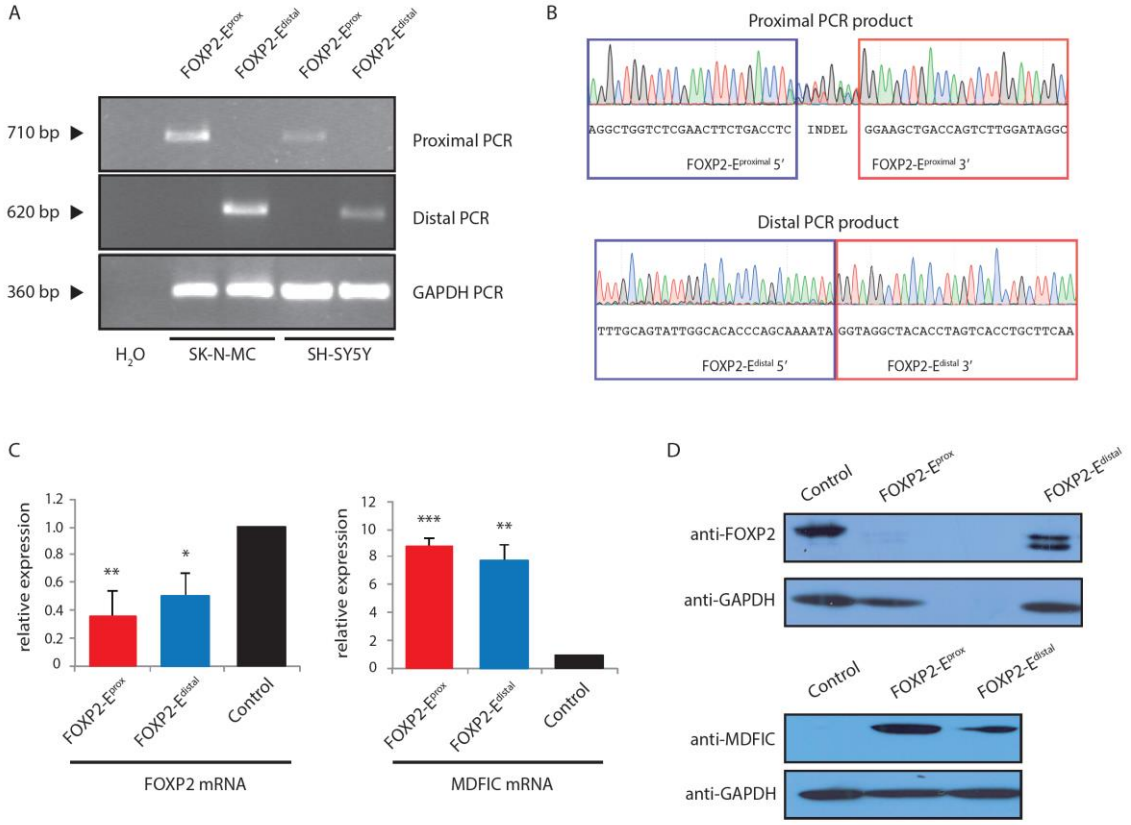

Figure 2. Molecular characterization of FOXP2-E^{proximal} and FOXP2-E^{distal}. A. PCR analysis. Two oligos flanking the deleted regions were used to amplify the genomic DNA from mutant SK-NMC and SH-SY-5Y clones. GAPDH genomic oligos were used to amplify a positive control region. Black triangles show the size of the PCR products. **B.** Representative Sanger sequencing chromatogram showing the boundaries of the FOXP2-E^{proximal} and FOXP2-E^{distal} deleted regions in SK-NMC cells. **C**. RT-qPCR analysis of SK-NMC cells with FOXP2-E^{proximal} or FOXP2-E^{distal} deletions. Samples are normalized to the average *FOXP2* (left) or *MDFIC* (right) signal between three SK-NMC wild type replicates. **D.** Western blot analysis of cell lysates of SK-NMC with FOXP2-E^{proximal} deleted, with FOXP2-E^{distal} deleted, and of control cells immunoblotted for FOXP2 (top) or MDFIC (bottom).

DISCUSSION

In this paper we have characterised in detail the role of two functional regulatory elements located downstream *FOXP2*, a gene important for speech and language (Fisher and Scharff 2009, Graham and Fisher 2013). Both elements affect the expression of *FOXP2*¸ but also that of the adjacent gene, *MDFIC*, a gene associated to developmental language and cognitive impairment (DECIPHER patients 262086, 292652, and 301696). MDFIC is highly expressed in the cerebellum during human embryonic development

bioRxiv preprint doi: https://doi.org/10.1101/064196; this version posted July 15, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

and in the thalamus after birth (Human Brain Transcriptome http://hbatlas.org/). These two brain regions seem to play an important role in language processing interacting in a dopaminergic cortico-striato-thalamic loop (Vargha-Khadem et al. 2005). Interestingly the cerebellum and thalamus of those bearing the R553H mutation in *FOXP2* associated to speech and language impairment exhibit changes in their grey matter suggesting that the modulation of brain volume may impact in sensorimotor performance (Watkins et al. 2002b). One of these enhancers, FOXP2-E^{distal}, had been previously found to be functional in a luciferase assay (Becker et al. 2015). We have been able to prove further that if deleted *FOXP2* becomes downregulated and the levels of FOXP2 protein are reduced. Conversely, its deletion upregulates *MDFIC* and increases the levels of MDFIC protein in the same SK-N-MC neuroectodermal tumor cell-line. The second enhancer, FOXP2-E^{proximal} was previously unknown. We have now found that it also upregulates *FOXP2* and downregulates *MDFIC*. These findings are coherent with previous reports of two genes being regulated by the same enhancer (Gould et al., 1997; Tsujimura et al. 2010), which in some cases has proven to facilitate recruitment of RNA polymerase II to promoters of both genes (Collins et al. 2012).

Our findings in a neuronal cell line give support to the view that the breakpoint in our proband which separated these two functional enhancers may have altered the expression levels of both *FOXP2* and *MDFIC* contributing to the observed speech and language deficits (Moralli et al 2015). Accordingly, we expect the expression of *FOXP2* to be downregulated and the expression of *MDFIC* to be upregulated because of the displacement of both enhancers with respect to both genes (whereas FOXP2-E ^{distal} and *MDFIC* remained in chromosome 7, FOXP2-E^{proximal} and *FOXP2* were rearranged to chromosome 11). The knockdown of *FoxP2* in zebra finch results in a shorter window for song learning and in less accurate song imitation and performance (Haesler et al., 2007). This resembles the inability that those carrying the R553H mutation of *FOXP2* show in repeating words and pseudowords (Watkins et al., 2002a). We expect that our findings also contribute to a better understanding of the role that this region may have played in the evolution of language. Differences in the expression levels of both *FOXP2* and *MDFIC* are expected between extinct hominins and modern humans, plausibly accounting for some of the presumed differences in their language abilities. Neanderthals bear the ancestral allele of a binding site for POU3F2 within intron 8 of *FOXP2*, which is more efficient in activating transcription (Maricic et al. 2013). Accordingly, higher levels of FOXP2 are expected in this hominin species. Likewise, the *MDFIC* locus is among the top five percent S score regions in modern humans (Green et al. 2010, table S37). Finally, both genes are

bioRxiv preprint doi: https://doi.org/10.1101/064196; this version posted July 15, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

functionally related to *RUNX2*, which encodes an osteogenic factor that controls the closure of cranial sutures and several aspects of brain growth, and that has been related to the changes that brought about our more globular brain (case) and our species-specific mode of cognition, including language (Boeckx and Benítez-Burraco 2014; Benítez-Burraco and Boeckx 2015).

We expect that our study, together with new available data about seed sequences of miRs in the 3'UTR region of *FOXP2* (Clovis et al. 2012; Shi et al. 2013; Fu et al. 2014a; Cuiffo et al. 2014), contributes to a deeper understanding of how *FOXP2*, but also *MDFIC*, are regulated, and how they each contribute to the development of brain function underlying language.

REFERENCES

Becker M, Devanna P, Fisher SE, Vernes SC (2015) A chromosomal rearrangement in a child with severe speech and language disorder separates FOXP2 from a functional enhancer. Mol Cytogenet 8:69. doi: 10.1186/s13039-015-0173-0

Benítez-Burraco A, Boeckx C (2015) Approaching motor and language deficits in autism from below: a biolinguistic perspective. Front Integr Neurosci 9:25. doi: 10.3389/fnint.2015.00025

Boeckx C, Benítez-Burraco A (2014) The shape of the human language-ready brain. Front Psychol 5:282. doi: 10.3389/fpsyg.2014.00282

Bonkowsky JL, Wang X, Fujimoto E, et al. (2008) Domain-specific regulation of foxP2 CNS expression by lef1. BMC Dev Biol 8:103. doi: 10.1186/1471-213X-8-103

Bruce HA, Margolis RL (2002) FOXP2: novel exons, splice variants, and CAG repeat length stability. Hum Genet 111:136–144. doi: 10.1007/s00439-002-0768-5

Chandrasekaran B, Yi H-G, Blanco NJ, et al. (2015) Enhanced procedural learning of speech sound categories in a genetic variant of FOXP2. J Neurosci 35:7808–7812. doi: 10.1523/JNEUROSCI.4706-14.2015

Chiu Y-C, Li M-Y, Liu Y-H, et al. (2014) Foxp2 regulates neuronal differentiation and neuronal subtype specification. Dev Neurobiol 74:723–738. doi: 10.1002/dneu.22166

Clovis YM, Enard W, Marinaro F, et al. (2012) Convergent repression of Foxp2 3'UTR by miR-9 and miR-132 in embryonic mouse neocortex: implications for radial migration of neurons. Development 139:3332–3342. doi: 10.1242/dev.078063

Collins PL, Henderson MA, Aune TM. (2012) Lineage-specific adjacent IFNG and IL26 genes share a common distal enhancer element. Genes Immun. 13:481-488. doi: 10.1038/gene.2012.22.

Cong L, Ran FA, Cox D, et al. (2013) Multiplex genome engineering using CRISPR/Cas systems. Science 339:819–823. doi: 10.1126/science.1231143

Coutinho P, Pavlou S, Bhatia S, et al. (2011) Discovery and assessment of conserved Pax6 target genes and enhancers. Genome Res 21:1349–1359. doi: 10.1101/gr.124115.111

Cuiffo BG, Campagne A, Bell GW, et al. (2014) MSC-regulated microRNAs converge on the transcription factor FOXP2 and promote breast cancer metastasis. Cell Stem Cell 15:762–774. doi:

bioRxiv preprint doi: https://doi.org/10.1101/064196; this version posted July 15, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

10.1016/j.stem.2014.10.001

Devanna P, Middelbeek J, Vernes SC (2014) FOXP2 drives neuronal differentiation by interacting with retinoic acid signaling pathways. Front Cell Neurosci 8:305. doi: 10.3389/fncel.2014.00305

Enard W (2011) FOXP2 and the role of cortico-basal ganglia circuits in speech and language evolution. Curr Opin Neurobiol 21:415–424. doi: 10.1016/j.conb.2011.04.008

Fisher SE, Scharff C (2009) FOXP2 as a molecular window into speech and language. Trends Genet 25:166–177. doi: 10.1016/j.tig.2009.03.002

French CA, Fisher SE (2014) What can mice tell us about Foxp2 function? Curr Opin Neurobiol 28:72–79. doi: 10.1016/j.conb.2014.07.003

Fu L, Shi Z, Luo G, et al. (2014a) Multiple microRNAs regulate human FOXP2 gene expression by targeting sequences in its 3' untranslated region. Mol Brain 7:71. doi: 10.1186/s13041-014-0071-0

Fu Y, Sander JD, Reyon D, et al. (2014b) Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat Biotechnol 32:279–284. doi: 10.1038/nbt.2808

Garcia-Calero E, Botella-Lopez A, Bahamonde O, et al. (2016) FoxP2 protein levels regulate cell morphology changes and migration patterns in the vertebrate developing telencephalon. Brain Struct Funct 221:2905–2917. doi: 10.1007/s00429-015-1079-7

Gould A, Morrison A, Sproat G, et al. (1997) Positive cross-regulation and enhancer sharing: two mechanisms for specifying overlapping Hox expression patterns. Genes Dev. 11:900-913. doi 10.1101/gad.11.7.900

Graham SA, Fisher SE (2013) Decoding the genetics of speech and language. Curr Opin Neurobiol 23:43–51. doi: 10.1016/j.conb.2012.11.006

Green RE, Krause J, Briggs AW, et al. (2010) A draft sequence of the Neandertal genome. Science 328:710–722. doi: 10.1126/science.1188021

Haesler S, Rochefort C, Georgi B, et al. (2007) Incomplete and inaccurate vocal imitation after knockdown of FoxP2 in songbird basal ganglia nucleus area X. PLoS Biol. 5:e321. doi: 10.1371/journal.pbio.0050321

Hallikas O, Palin K, Sinjushina N, et al. (2006) Genome-wide prediction of mammalian enhancers based on analysis of transcription-factor binding affinity. Cell 124:47–59. doi: 10.1016/j.cell.2005.10.042

Hammerschmidt K, Schreiweis C, Minge C, et al. (2015) A humanized version of Foxp2 does not affect ultrasonic vocalization in adult mice. Genes Brain Behav 14:583–590. doi: 10.1111/gbb.12237

Hoogman M, Guadalupe T, Zwiers MP, et al. (2014) Assessing the effects of common variation in the FOXP2 gene on human brain structure. Front Hum Neurosci 8:473. doi: 10.3389/fnhum.2014.00473

Konopka G, Bomar JM, Winden K, et al. (2009) Human-specific transcriptional regulation of CNS development genes by FOXP2. Nature 462:213–217. doi: 10.1038/nature08549

Kosho T, Sakazume S, Kawame H, et al. (2008) De-novo balanced translocation between 7q31 and 10p14 in a girl with central precocious puberty, moderate mental retardation, and severe speech impairment. Clin Dysmorphol 17:31–34. doi: 10.1097/MCD.0b013e3282f17688

Krause J, Lalueza-Fox C, Orlando L, et al. (2007) The derived FOXP2 variant of modern humans was shared with Neandertals. Curr Biol 17:1908–1912. doi: 10.1016/j.cub.2007.10.008

Kurt S, Fisher SE, Ehret G (2012) Foxp2 mutations impair auditory-motor association learning. PLoS ONE 7:e33130. doi: 10.1371/journal.pone.0033130

bioRxiv preprint doi: https://doi.org/10.1101/064196; this version posted July 15, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Mali P, Yang L, Esvelt KM, et al. (2013) RNA-guided human genome engineering via Cas9. Science 339:823–826. doi: 10.1126/science.1232033

Maricic T, Günther V, Georgiev O, et al. (2013) A recent evolutionary change affects a regulatory element in the human FOXP2 gene. Mol Biol Evol 30:844–852. doi: 10.1093/molbev/mss271

Moralli D, Nudel R, Chan MTM, et al. (2015) Language impairment in a case of a complex chromosomal rearrangement with a breakpoint downstream of FOXP2. Mol Cytogenet 8:36. doi: 10.1186/s13039-015-0148-1

Mueller KL, Murray JC, Michaelson JJ, et al. (2016) Common Genetic Variants in FOXP2 Are Not Associated with Individual Differences in Language Development. PLoS ONE 11:e0152576. doi: 10.1371/journal.pone.0152576

Mukamel Z, Konopka G, Wexler E, et al. (2011) Regulation of MET by FOXP2, genes implicated in higher cognitive dysfunction and autism risk. J Neurosci 31:11437–11442. doi: 10.1523/JNEUROSCI.0181-11.2011

Ocklenburg S, Arning L, Gerding WM, et al. (2013) FOXP2 variation modulates functional hemispheric asymmetries for speech perception. Brain Lang 126:279–284. doi: 10.1016/j.bandl.2013.07.001

Padovani A, Cosseddu M, Premi E, et al. (2010) The speech and language FOXP2 gene modulates the phenotype of frontotemporal lobar degeneration. J Alzheimers Dis 22:923–931. doi: 10.3233/JAD-2010-101206

Reimers-Kipping S, Hevers W, Pääbo S, Enard W (2011) Humanized Foxp2 specifically affects cortico-basal ganglia circuits. Neuroscience 175:75–84. doi: 10.1016/j.neuroscience.2010.11.042

Rodenas-Cuadrado P, Pietrafusa N, Francavilla T, et al. (2016) Characterisation of CASPR2 deficiency disorder--a syndrome involving autism, epilepsy and language impairment. BMC Med Genet 17:8. doi: 10.1186/s12881-016-0272-8

Rodriguez-Perales S, Torres-Ruiz R, Suela J, et al. (2015) Truncated RUNX1 protein generated by a novel t(1;21)(p32;q22) chromosomal translocation impairs the proliferation and differentiation of human hematopoietic progenitors. Oncogene. doi: 10.1038/onc.2015.70

Roll P, Vernes SC, Bruneau N, et al. (2010) Molecular networks implicated in speech-related disorders: FOXP2 regulates the SRPX2/uPAR complex. Hum Mol Genet 19:4848–4860. doi: 10.1093/hmg/ddq415

Sakuma T, Woltjen K (2014) Nuclease-mediated genome editing: At the front-line of functional genomics technology. Dev Growth Differ 56:2–13. doi: 10.1111/dgd.12111

Schreiweis C, Bornschein U, Burguière E, et al. (2014) Humanized Foxp2 accelerates learning by enhancing transitions from declarative to procedural performance. Proc Natl Acad Sci USA 111:14253–14258. doi: 10.1073/pnas.1414542111

Schroeder DI, Myers RM (2008) Multiple transcription start sites for FOXP2 with varying cellular specificities. Gene 413:42–48. doi: 10.1016/j.gene.2008.01.015

Shi Z, Luo G, Fu L, et al. (2013) miR-9 and miR-140-5p target FoxP2 and are regulated as a function of the social context of singing behavior in zebra finches. J Neurosci 33:16510–16521. doi: 10.1523/JNEUROSCI.0838-13.2013

Shlyueva D, Stampfel G, Stark A (2014) Transcriptional enhancers: from properties to genome-wide predictions. Nat Rev Genet 15:272–286. doi: 10.1038/nrg3682

Shriberg LD, Ballard KJ, Tomblin JB, et al. (2006) Speech, prosody, and voice characteristics of a mother and daughter with a 7;13 translocation affecting FOXP2. J Speech Lang Hear Res 49:500–525. doi: 10.1044/1092-4388(2006/038)

bioRxiv preprint doi: https://doi.org/10.1101/064196; this version posted July 15, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Spiteri E, Konopka G, Coppola G, et al. (2007) Identification of the transcriptional targets of FOXP2, a gene linked to speech and language, in developing human brain. Am J Hum Genet 81:1144–1157. doi: 10.1086/522237

Tolosa A, Sanjuán J, Dagnall AM, et al. (2010) FOXP2 gene and language impairment in schizophrenia: association and epigenetic studies. BMC Med Genet 11:114. doi: 10.1186/1471-2350-11-114

Torres R, Garcia A, Jimenez M, et al. (2014a) An integration-defective lentivirus-based resource for site-specific targeting of an edited safe-harbour locus in the human genome. Gene Ther 21:343–352. doi: 10.1038/gt.2014.1

Torres R, Martin MC, Garcia A, et al. (2014b) Engineering human tumour-associated chromosomal translocations with the RNA-guided CRISPR–Cas9 system. Nature Communications. doi: 10.1038/ncomms4964

Torres-Ruiz R, Rodriguez-Perales S (2016) CRISPR-Cas9 technology: applications and human disease modelling. Brief Funct Genomics elw025. doi: 10.1093/bfgp/elw025

Tsui D, Vessey JP, Tomita H, et al. (2013) FoxP2 regulates neurogenesis during embryonic cortical development. J Neurosci 33:244–258. doi: 10.1523/JNEUROSCI.1665-12.2013

Tsujimura Tl, HosoyaT, Kawamara S (2010) A single enhancer regulating the differential expression of duplicated red-sensitive opsine genes in zebra-fish. PLoS Genet. 2010 Dec 16;6 (12):e1001245. doi:10.1371/journal.pgen.1001245

Vargha-Khadem F, Gadian DG, Copp A, Mishkin M (2005) FOXP2 and the neuroanatomy of speech and language. Nat Rev Neurosci 6:131–138. doi: 10.1038/nrn1605

Vargha-Khadem F, Watkins K, Alcock K, et al. (1995) Praxic and nonverbal cognitive deficits in a large family with a genetically transmitted speech and language disorder. PNAS 92:930–933.

Vernes SC, Oliver PL, Spiteri E, et al. (2011) Foxp2 regulates gene networks implicated in neurite outgrowth in the developing brain. PLoS Genet 7:e1002145. doi: 10.1371/journal.pgen.1002145

Vernes SC, Spiteri E, Nicod J, et al. (2007) High-throughput analysis of promoter occupancy reveals direct neural targets of FOXP2, a gene mutated in speech and language disorders. Am J Hum Genet 81:1232–1250. doi: 10.1086/522238

Walker RM, Hill AE, Newman AC, et al. (2012) The DISC1 promoter: characterization and regulation by FOXP2. Hum Mol Genet 21:2862–2872. doi: 10.1093/hmg/dds111

Watkins KE, Dronkers NF, Vargha-Khadem F (2002a) Behavioural analysis of an inherited speech and language disorder: comparison with acquired aphasia. Brain 125:452–464.

Watkins KE, Vargha-Khadem F, Ashburner J, et al. (2002b) MRI analysis of an inherited speech and language disorder: structural brain abnormalities. Brain 125:465–478.

Wohlgemuth S, Adam I, Scharff C (2014) FoxP2 in songbirds. Curr Opin Neurobiol 28:86–93. doi: 10.1016/j.conb.2014.06.009

Zhao Y, Ma H, Wang Y, et al. (2010) Association between FOXP2 gene and speech sound disorder in Chinese population. Psychiatry Clin Neurosci 64:565–573. doi: 10.1111/j.1440-1819.2010.02123.x

bioRxiv preprint doi: https://doi.org/10.1101/064196; this version posted July 15, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplementary Table 1. Oligonucleotide and sgRNA sequences used in this study.

sgRNA sequences (IDT gblocks)

sgRNA_FOXP2_Edistal_1_5:

ccataACGCGTTGTACACGAACGCTGACGTCATCAACCCGCTCCAAGGAATCGCGGGCC
CAGTGTCACTAGGCGGGAACACCCAGCGCGCGTGCGCCCTGGCAGGAAGATGGCTGTGA
GGGACAGGGGAGTGGCGCCCTGCAATATTTGCATGTCGCTATGTGTTCTGGGAAATCAC
CATAAACGTGAAATGTCTTTGGATTTGGGAATCTTATAAGTTCTGTATGAGACCACTCT
TTCCC**G**cacacccagcaaaatacatGTTTTAGAGCTATGCTGGAAACAGCATAGCAAGT
TAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTA
CTAGTcgcta

sgRNA_FOXP2_Edistal_2_5:

ccataACGCGTTGTACACGAACGCTGACGTCATCAACCCGCTCCAAGGAATCGCGGGCC
CAGTGTCACTAGGCGGGAACACCCAGCGCGCGTGCGCCCTGGCAGGAAGATGGCTGTGA
GGGACAGGGGAGTGGCGCCCTGCAATATTTGCATGTCGCTATGTGTTCTGGGAAATCAC
CATAAACGTGAAATGTCTTTGGATTTGGGAATCTTATAAGTTCTGTATGAGACCACTCT
TTCCC**G**gcaaggtatattctctgagGTTTTAGAGCTATGCTGGAAACAGCATAGCAAGT
TAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTA
CTAGTcgcta

sgRNA_FOXP2_Edistal_1_3:

ccataCAATTGGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATA
CGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATT
AGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAAT
TATGTTTTAAAATGGACTATCATGTACACTTACCGTAACTTGAAAGTATTTCGATTTCT
TGGCTTTATATATCTTGTGGAAAGGACGAGGTACC**G**atctactcttctttagggtGTTT
TAGAGCTATGCTGGAAACAGCATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTG
AAAAAGTGGCACCGAGTCGGTGCTTTTTTACGCGTACTAGTcgcta

sgRNA_FOXP2_Edistal_2_3:

ccataCAATTGGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATA
CGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATT
AGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAAT
TATGTTTTAAAATGGACTATCATGTACACTTACCGTAACTTGAAAGTATTTCGATTTCT
TGGCTTTATATATCTTGTGGAAAGGACGAGGTACC**G**gaagagtagatcgcatgagGTTT
TAGAGCTATGCTGGAAACAGCATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTG
AAAAAGTGGCACCGAGTCGGTGCTTTTTTACGCGTACTAGTcgcta

bioRxiv preprint doi: https://doi.org/10.1101/064196; this version posted July 15, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

sgRNA_FOXP2_Eproximal_1_5:

ccataACGCGTTGTACACGAACGCTGACGTCATCAACCCGCTCCAAGGAATCGCGGGCC
CAGTGTCACTAGGCGGGAACACCCAGCGCGCGTGCGCCCTGGCAGGAAGATGGCTGTGA
GGGACAGGGGAGTGGCGCCCTGCAATATTTGCATGTCGCTATGTGTTCTGGGAAATCAC
CATAAACGTGAAATGTCTTTGGATTTGGGAATCTTATAAGTTCTGTATGAGACCACTCT
TTCCC**G**gtgatctcagctactcgggGTTTTAGAGCTATGCTGGAAACAGCATAGCAAG
TTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT
ACTAGTcgcta

sgRNA_FOXP2_Eproximal_2_5:

ccataACGCGTTGTACACGAACGCTGACGTCATCAACCCGCTCCAAGGAATCGCGGGCC
CAGTGTCACTAGGCGGGAACACCCAGCGCGCGTGCGCCCTGGCAGGAAGATGGCTGTGA
GGGACAGGGGAGTGGCGCCCTGCAATATTTGCATGTCGCTATGTGTTCTGGGAAATCAC
CATAAACGTGAAATGTCTTTGGATTTGGGAATCTTATAAGTTCTGTATGAGACCACTCT
TTCCC**G**ctcgaacttctgacctcagGTTTTAGAGCTATGCTGGAAACAGCATAGCAAGT
TAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTTA
CTAGTcgcta

sgRNA_FOXP2_Eproximal_1_3:

ccataCAATTGGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATA
CGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATT
AGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAAT
TATGTTTTAAAATGGACTATCATGTACACTTACCGTAACTTGAAAGTATTTCGATTTCT
TGGCTTTATATATCTTGTGGAAAGGACGAGGTACC**G**ctgtaataagatagcagggGTTT
TAGAGCTATGCTGGAAACAGCATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTG
AAAAAGTGGCACCGAGTCGGTGCTTTTTTACGCGTACTAGTcgcta

sgRNA_FOXP2_Eproximal_2_3:

ccataCAATTGGGGCAGGAAGAGGGCCTATTTCCCATGATTCCTTCATATTTGCATATA
CGATACAAGGCTGTTAGAGAGATAATTAGAATTAATTTGACTGTAAACACAAAGATATT
AGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTTAAAAT
TATGTTTTAAAATGGACTATCATGTACACTTACCGTAACTTGAAAGTATTTCGATTTCT
TGGCTTTATATATCTTGTGGAAAGGACGAGGTACC**G**tatggctgccacattccgtGTTT
TAGAGCTATGCTGGAAACAGCATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTG
AAAAAGTGGCACCGAGTCGGTGCTTTTTTACGCGTACTAGTcgcta

bioRxiv preprint doi: https://doi.org/10.1101/064196; this version posted July 15, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Primers qPCR

> qFoxP2_Fw
GCAGCAGAGATGGAAGATCA

> qFoxP2_Rv
AGTTGTCTTGCTGCCTGGAG

cDNA amplicon size:103 Estimated genomic amplicon size:108040

> qMDFIC_Fw
GTCCATTTGGGGAAATCCTT

> qMDFIC_Rv
CATTGCTCAGACCTGTGTGG

cDNA amplicon size:140 Estimated genomic amplicon size:37248

Primers Surveyor & deletion detection

sMDFIC Fw2
TGATCTCAGTGCAGGCAAA
sMDFIC Rv2
GTTGGACTAAGGTGCCAGTT

2314pb (deletion FOXP2 Distal)

sMDFIC Fw
TACTGTTTCATGGATGCTGACT
sMDFIC Rv
CCTTTGGCCACAGACTGAA

sFOXP2 Fw
GGGATAGCACTGGGAGAAATAC
sFOXP2 Rv
GCGGTGGCTCATTTCTGTA

6264 pb (deletion FOXP2 Proximal)

sFOXP2 Fw2
TTCTGCACCTTGGGTTAGG
sFOXP2 Rv2
AGGGTTGATTGATTGCCAGAG