The American Journal of Human Genetics, Volume *105*

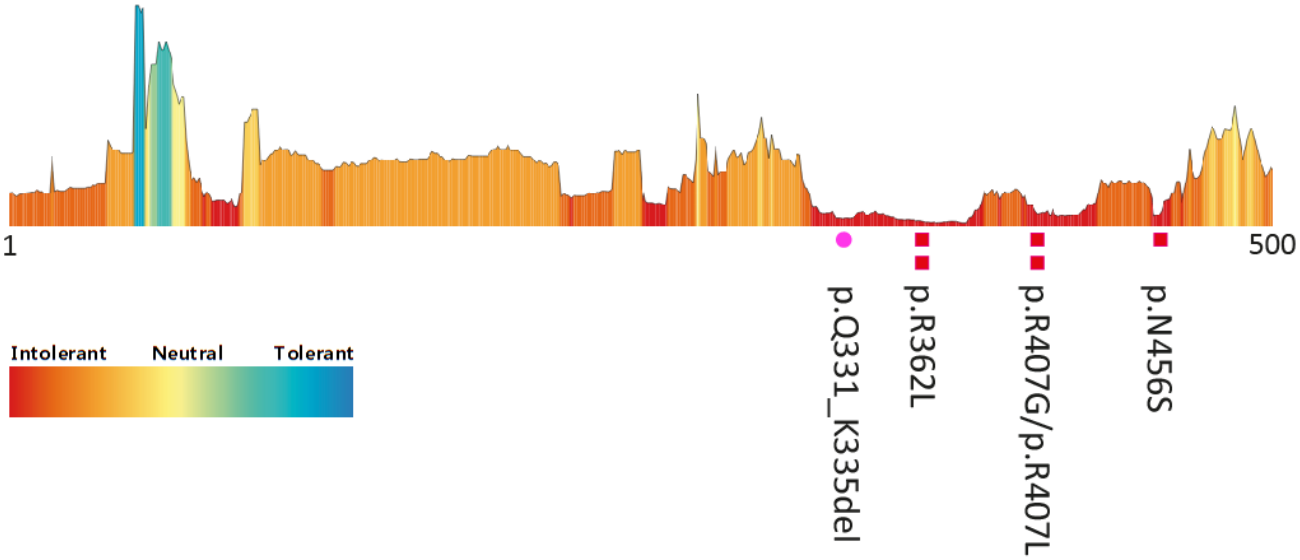
Supplemental Data

De Novo Variants Disturbing the Transactivation Capacity of POU3F3 Cause a Characteristic Neurodevelopmental Disorder

Lot Snijders Blok, Tjitske Kleefstra, Hanka Venselaar, Saskia Maas, Hester Y. Kroes, Augusta M.A. Lachmeijer, Koen L.I. van Gassen, Helen V. Firth, Susan Tomkins, Simon Bodek, The DDD Study, Katrin Õunap, Monica H. Wojcik, Christopher Cunniff, Katherine Bergstrom, Zoë Powis, Sha Tang, Deepali N. Shinde, Catherine Au, Alejandro D. Iglesias, Kosuke Izumi, Jacqueline Leonard, Ahmad Abou Tayoun, Samuel W. Baker, Marco Tartaglia, Marcello Niceta, Maria Lisa Dentici, Nobuhiko Okamoto, Noriko Miyake, Naomichi Matsumoto, Antonio Vitobello, Laurence Faivre, Christophe Philippe, Christian Gilissen, Laurens Wiel, Rolph Pfundt, Pelagia Deriziotis, Han G. Brunner, and Simon E. Fisher

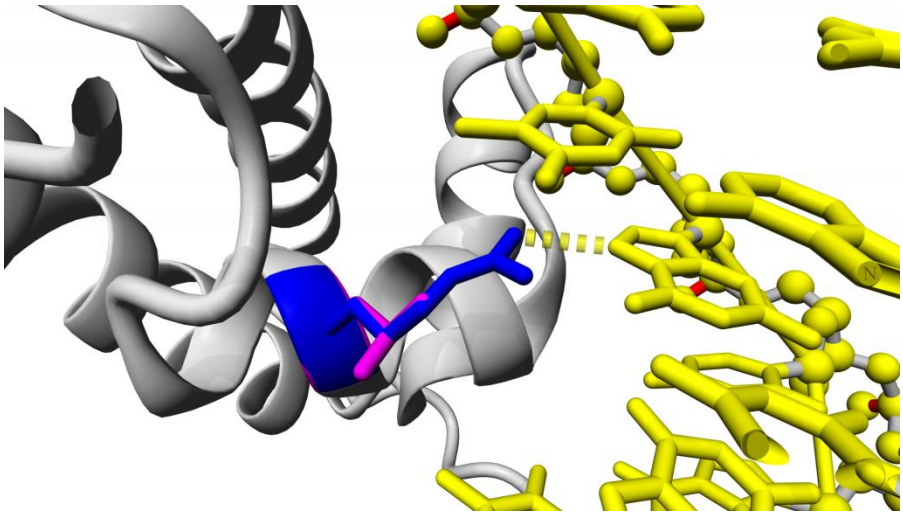
Supplemental Figures

Figure S1: Non-truncating variants visualised in tolerance landscape of POU3F3

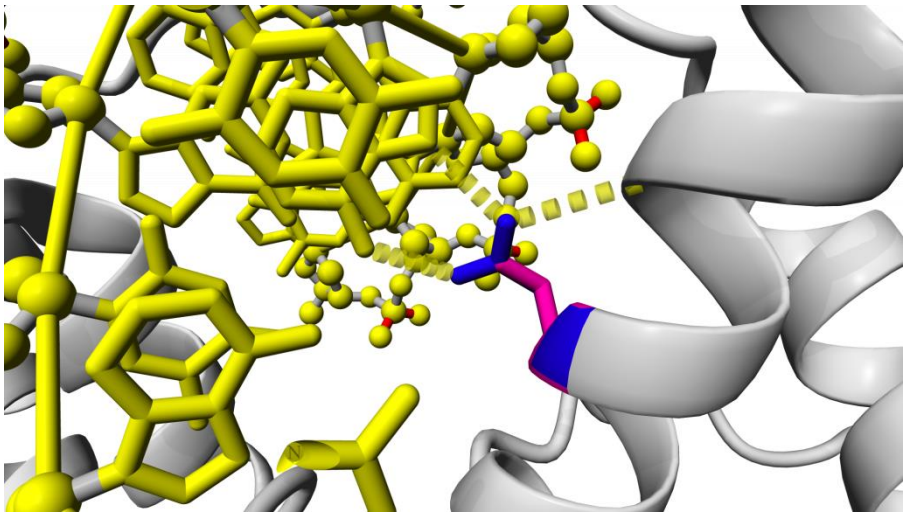


Tolerance landscape of the POU3F3 protein based on transcript NM_006236.2 (ENST00000361360.2) visualized via the MetaDome web server[1]. The tolerance landscape is computed based on single nucleotide variants present in the gnomAD database. It is calculated as a missense over synonymous ratio in a sliding window of 21 residues over the entire POU3F3 protein. The green and blue peaks correspond to regions more tolerant to missense variation, and the red valleys indicate intolerant regions. The locations of the non-truncating variants in our cohort are displayed within the tolerance landscape of POU3F3. All these variants are located in regions that are highly intolerant to missense variation.

Figure S2: Three-dimensional visualisation for the p.(Arg362Leu) and p.(Asn456Ser) variants



p.(Arg362Leu)

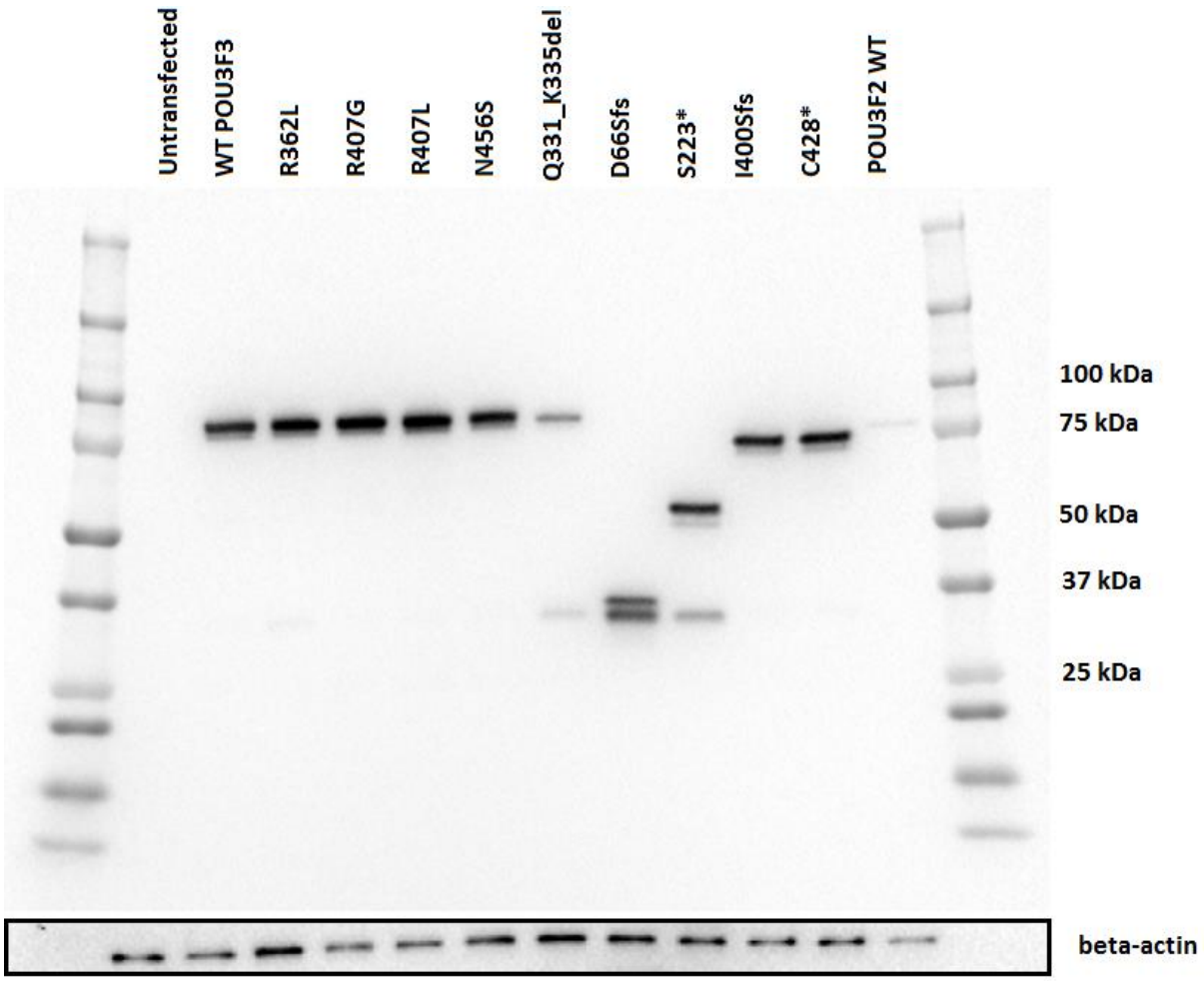


p.(Asn456Ser)

A detailed visualization of the three-dimensional modeling analysis for the two missense variants affecting amino acids that directly bind to the major groove of DNA . Wild-type residues are shown in blue, while substitutions (caused by variants) are shown in magenta. DNA is depicted in yellow.

A) Arg362 (blue) is able to form hydrogen bonds with guanine in the DNA binding site. This binding is disturbed by substitution into leucine (magenta).
B) Asn456 (blue) is able to form hydrogen bonds with adenine in the DNA binding site. This binding is disturbed by substitution into serine (magenta).

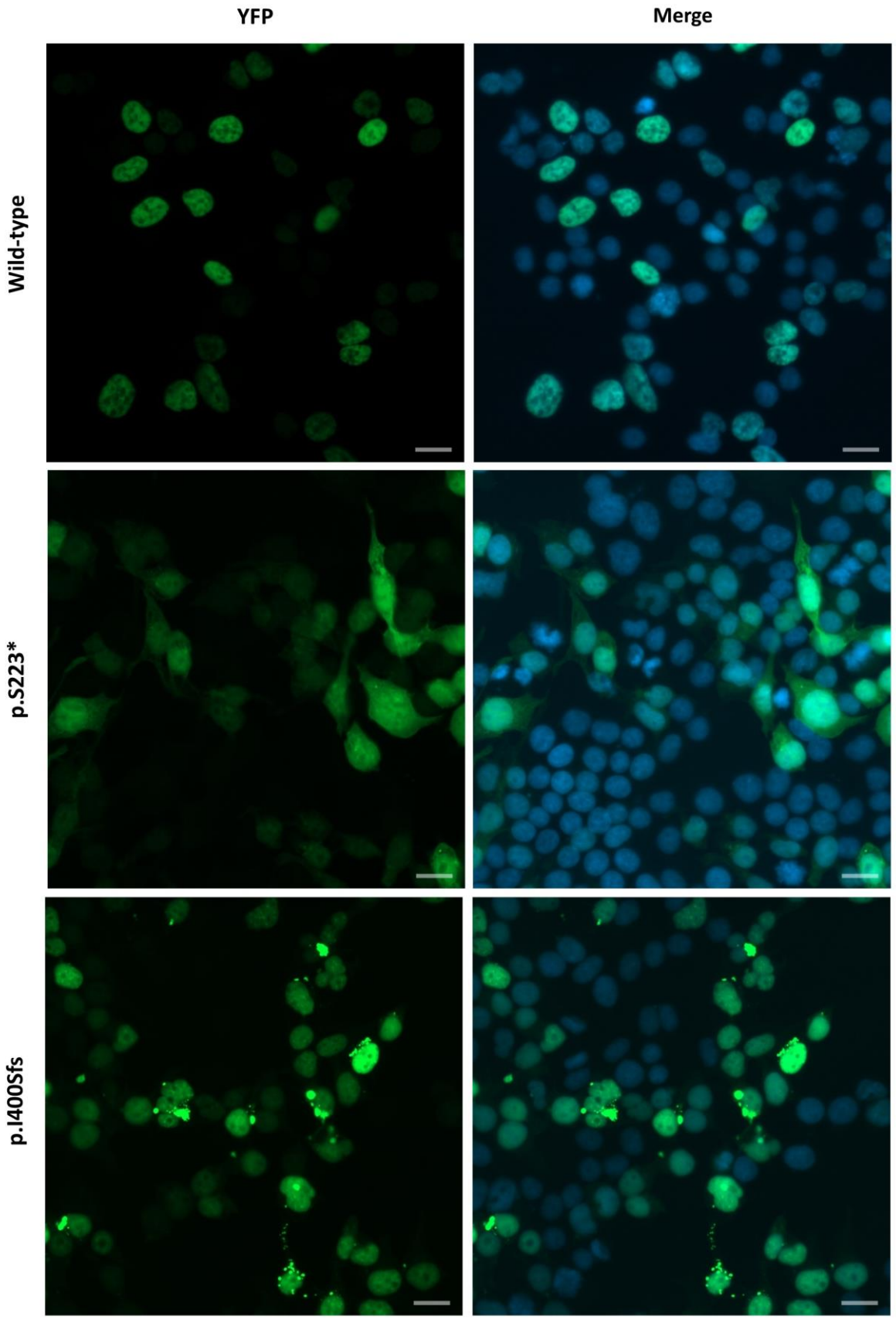
Figure S3: Immunoblot analysis



Western blot analysis of whole-cell lysates expressing eleven different YFP-tagged constructs, probed with an anti-EGFP antibody: wild-type POU3F3, nine different POU3F3 variants and wild-type POU3F2. The immunoblot was stripped and then re-probed with beta-actin as a loading control. All different expressed YFP-fusion proteins are visible at the expected molecular weights.

Figure S4: Aberrant subcellular localization patterns in a subset of cells for four variants

A) Direct fluorescence imaging of cells expressing YFP-tagged variants of the POU3F3 protein: wild-type POU3F3, and POU3F3 with the p.(Ser223*) variant and the p.(Ile400Serfs*16) variant. In addition to the cytoplasmic localization of the p.(Ser223*) variant, both variant conditions show perinuclear aggregates in a subset of cells. Nuclei are stained with DAPI (blue). Scale bar = 10µm.



B) Direct fluorescence imaging of cells expressing YFP-tagged variants of the POU3F3 protein: wild-type POU3F3, and POU3F3 with the p.(Gln331_Lys335del) variant and the p.(Cys428*) variant. Both variant conditions show an aberrant localization pattern within the nucleus in a subset of cells. Nuclei are stained with DAPI (blue). Scale bar = 10µm.

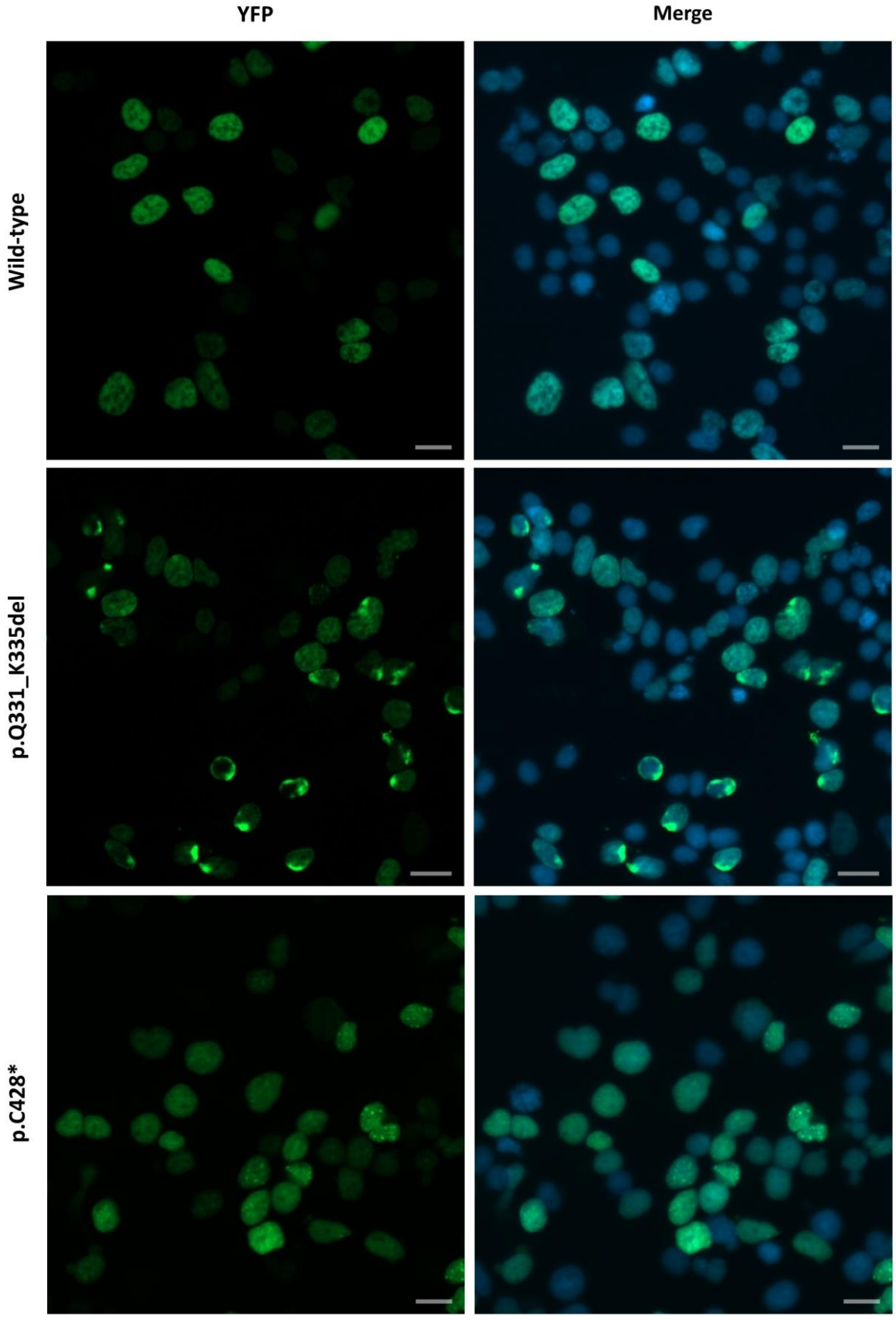
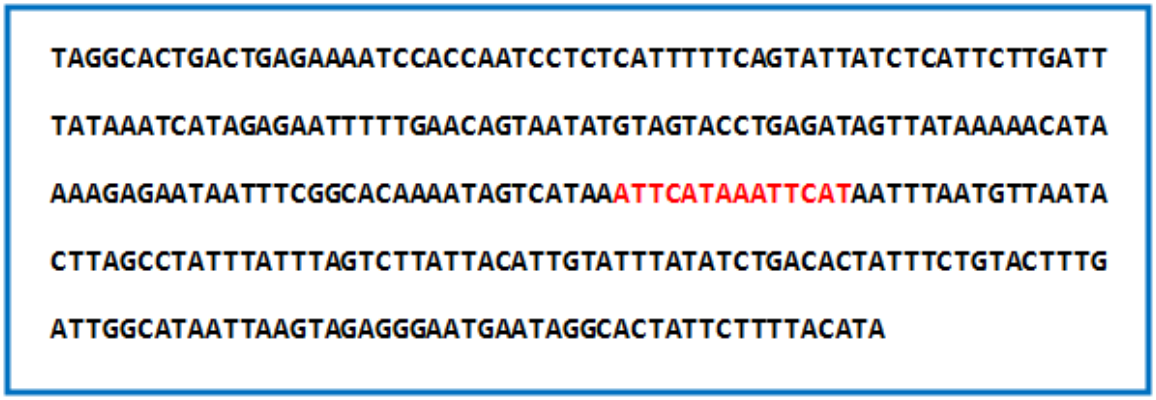


Figure S5: POU3F3 binding site in intronic region of *FOXP2*



This figure shows the ~300bp region of intron 8 of FOXP2 (chr7:114,289,482-114,289,778 (hg19/GRCh37)), that was cloned into a luciferase reporter vector to investigate transcriptional activation. The previously described POU3F2 consensus binding site[2] is shown in red.

Supplemental Material and Methods

Research subjects

Informed consent was obtained from all participating families. For all pictures of probands in this study, specific consent to publish clinical photographs was obtained. All procedures in this study matched the local ethical guidelines of the participating centres, and are in accordance with the Declaration of Helsinki. Probands with possible pathogenic POU3F3 variants were found using the GeneMatcher website[3], the Decipher Database[4] and matchbox[5], and table S1 contains details on which specific matchmaking platform was used to identify each individual.

Exome sequencing, variant filtering and annotation

Exome sequencing and variant filtering were performed as previously described[6-17]. In individuals 1-17, whole exome sequencing and variant filtering was performed using a trio approach, in which sequencing was performed in the proband and both parents. In individual 18 and 19, whole exome sequencing was performed with a duo approach, in which the proband (individual 18) and the mother (individual 19) were sequenced. In all individuals, the POU3F3 variant was considered to be the most likely variant contributing to the phenotype, and there were no additional pathogenic or likely pathogenic variants reported. Additional variants (SNVs and CNVs) considered to be possibly pathogenic and/or to possibly contribute to the phenotype, are listed in Table S1. All variants in this study are annotated with respect to the NM_006236.2 transcript.

Cell culture and transfection

Human embryonic kidney cells 293 (HEK293) cells were grown in Dulbecco's Modified Eagle's Medium (Gibco) that was supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin mix (both Gibco). The cells were cultured at 37°C in an incubator with 5% CO_2. Transient transfection was performed with GeneJuice Transfection Reagent (Merck MilliPore) according to the manufacturer's instructions.

Cloning of DNA constructs and site-directed mutagenesis

A synthetic clone of wildtype POU3F3 cDNA with flanking restriction sites EcoR1 and Xba1 in a pUC57-vector was synthesized by GenScript. The POU3F3 insert was subcloned using the EcoR1/Xba1 restriction sites into a modified pLuc and pYFP vector as previously described [18].

Variants in POU3F3 constructs were created using site-directed mutagenesis (SDM) with the QuikChange II Site-Directed Mutagenesis Kit (Agilent) according to the manufacterer's instructions. The following mutated constructs were created (corresponding SDM primers in parentheses; F = Forward primer, R = Reverse primer): p.Arg362Leu (F: 5'-ACCACCATCTGCCTCTTCGAGGCCCTG-3'; R: 5'-CAGGGCCTCGAAGAGGCAGATGGTGGT-3'), p.Arg407Gly (F: 5'-CAGGGCCGCAAGGGCAAGAAGCGGA-3'; R: 5'-TCCGCTTCTTGCCCTTGCGGCCCTG-3'), p.Arg407Leu (F: 5'-CAGGGCCGCAAGCTCAAGAAGCGGACC-3'; R: 5'-GGTCCGCTTCTTGAGCTTGCGGCCCTG-3'), p.Asn456Ser (F: 5'-GCGGGTCTGGTTCTGCAGTCGGCGCCA-3'; R: 5'-TGGCGCCGACTGCAGAACCAGACCCGC-3'), p.Gln331_Lys335del (F: 5'-CCTGCGTGAAGCCCAGCTTGAACTGCTTGG-3'; R: 5'-CCAAGCAGTTCAAGCTGGGCTTCACGCAGG-3'), p.Asp66Serfs (F: 5'-GCCTACCGGGGGTCCCGTCCTCTGT-3'; R: 5'-ACAGAGGACGGGACCCCCGGTAGGC-3'), p.Ser223* (F: 5'-CCGGGCTGCTAGTAGAGCAGACTCTGC-3'; R: 5'-GCAGAGTCTGCTCTACTAGCAGCCCCG-3'), p. Ile400Serfs (F: 5'-GCGCCGCGATTTGTCGATGCTTGTGGGG-3'; R: 5'-CCCCACAAGCATCGACAAATCGCGGCGC-3') and p.Cys428* (F: 5'-CACTTCCTCAAGTGACCCAAGCCCTCCGC-3'; R: 5'-GCGGAGGGCTTGGGTCACTTGAGGAAGTG-3'). After SDM, variants were validated using Sanger sequencing, and POU3F3 inserts were subcloned into new pLuc and pYFP vectors. An 'empty YFP-vector' (modified pYFP expression vector without POU3F3 insert) was used as a control construct for the luciferase assays.

POU3F2 cDNA was cloned into TOPO vector using the following primers: 5'-GAGGATCCTGGCGACCGCAGCGTCTAACCAC-3' (Forward primer) and 5'-GATCTAGATTACTGGACGGGCGTCTGCACCCCG-3' (Reverse primer). The POU3F2 insert was then

subcloned into modified pLuc and pYFP vectors (as previously described; [18]) using BamHI and XbaI restriction sites.

To create the firefly reporter construct for luciferase assays, the previously described POU3F2 binding site in FOXP2 (Figure S5) was cloned into TOPO from gDNA using the following primers: 5'-CTCGAGTAGGCACTGACTGAGAAAATC-3' (Forward primer) and 5'-AGATCTATATGTAAAAGAATAGTGCCT-3' (Reverse primer). The binding site was then subcloned into a p.GL4.23 vector (Promega) using the restriction sites BglII and XhoI.

Control constructs used in the BRET assays (pYFP-vector with NLS-insert, and pRLuc-vector with NLS-insert) were made as previously described[18]. All constructs were validated by Sanger sequencing.

Subcellular localization

HEK293 cells were seeded on poly-D-lysine (Sigma) coated coverslips, and transfected after 24 hours. At 36 hours post-transfection, the cells were fixed by incubation in 4% Paraformaldehyde (Electron Microscopy Sciences) for ten minutes at room temperature. Coverslips were mounted onto slides using Vectashield mounting medium for fluorescence with DAPI (Vector). The proteins of interest were expressed as fusion proteins to YFP. Fluorescence images were obtained with an Axiovert A-1 fluorescence microscope and ZEN imaging software (Zeiss).

BRET assay

HEK293 cells were transfected 24 hours after plating in 96 well plates, with pairs of Renilla luciferase and YFP fusion proteins, as previously described[18]. Luciferase substrate (EnduRen; Promega) was added at 60μM 36 hours post-transfection. After four hours of incubation, the emission was measured using a TECAN F200 PRO or M200 PRO microplate reader using the Blue1 and Green1 filter sets. To determine the YFP-fusion protein expression level, fluorescence measurements were taken using a filter and a dichroic mirror suitable for green fluorescent protein (GFP) fluorescence (excitation 480nm, emission 535nm). The corrected BRET ratio was obtained using the following

formula [$Green1_{(experimental\ condition)}/Blue1_{(experimental\ condition)}$] - [$Green1_{(control\ condition)}/Blue1_{(control\ condition)}$]. The BRET assay set-up that was used for this study is discussed in more detail in Deriziotis et al. [18].

Luciferase assay

HEK293 cells were transfected 24 hours after seeding in 96 well plates with the firefly luciferase reporter construct (2µl of 36nM), a pGL4.74 (*hRluc*/TK) Renilla Reniformis luciferase construct (2µl of 36nM; Promega) and a YFP-expression construct or empty YFP-expression vector (6µl of 36nM). Firefly luciferase and *Renilla* luciferase activities were measured 36 hours post-transfection using the Dual-Luciferase Reporter Assay System (Promega) and a TECAN F200 PRO microplate reader.

Western blotting

Whole cell lysates were collected 40 hours post-transfection by treatment with RIPA buffer (Cell Signaling Technology) supplemented with 0.1mM PMSF (Sigma), Protease Inhibitor Cocktail (Roche) and 1mM DTT (Sigma). Cells were lysed for 30 minutes at 4°C followed by centrifugation for 10 minutes at 13,000rpm at 4°C. Laemmli buffer (Bio-Rad) was added to the supernatants, and the proteins were loaded on 4-15% Mini Protean-TGX Precast Gels (Bio-Rad) and transferred onto polyvinylidene fluoride membranes (Bio-Rad). Membranes were blocked in 5% milk in PBS-T (Phospate Buffered Saline supplemented with 0.1% Tween) for 1.5h at room temperature and then probed with 1:8000 mouse anti-EGFP (Clontech) in 1% milk in PBS-T at 4°C overnight. This was followed by incubation with 1:3000 HRP-conjugated goat anti-mouse (Bio-Rad) in 1% milk at room temperature for 1h. Proteins were visualized with Novex ECL Chemluminescent Substrate Reagent Kit (Invitrogen), using the ChemiDoc XRS+ System (Bio-Rad). To check for equal loading of proteins, the blot was subsequently stripped for 25 minutes in Re-blot Plus Strong stripping solution (Millipore) at room temperature, and blocked in 5% milk in PBS-T for 30 minutes at room temperature, followed by incubation with 1:1000 mouse anti-beta-actin (Sigma) in 1% milk overnight at 4°C, and incubation with 1:3000 HRP-conjugated goat anti-mouse (Bio-Rad) in 1% milk at room temperature.

Three-dimensional modeling

The exact three-dimensional structure of human POU3F3 is not known. Therefore, we created a homology model based on the crystal structure of the mouse POU3F1 structure (PDB file 2XSD)[19]. The human POU3F3 and mouse POU3F1 sequences show 94% identity over 147 residues in the C-terminal domain, containing both the POU-specific and the POU-homeo domain. We used the YASARA & WHAT IF Twinset modeling algorithm with standard parameters for modeling and subsequent analysis[20; 21].

Supplemental Acknowledgements

Individuals 1 and 2 were part of the DDD study cohort. The DDD study presents independent research commissioned by the Health Innovation Challenge Fund [grant number HICF-1009-003], a parallel funding partnership between Wellcome and the Department of Health, and the Wellcome Sanger Institute [grant number WT098051]. The views expressed in this publication are those of the author(s) and not necessarily those of Wellcome or the Department of Health. The study has UK Research Ethics Committee approval (10/H0305/83, granted by the Cambridge South REC, and GEN/284/12 granted by the Republic of Ireland REC). The research team acknowledges the support of the National Institute for Health Research, through the Comprehensive Clinical Research Network. This study makes use of DECIPHER (http://decipher.sanger.ac.uk), which is funded by the Wellcome.

Supplemental References

1. Wiel, L., Baakman, C., Gilissen, D., Veltman, J.A., Vriend, G., and Gilissen, C. (2019). MetaDome: Pathogenicity analysis of genetic variants through aggregation of homologous human protein domains. Hum Mutat doi: 10.1002/humu.23798. [Epub ahead of print]
2. Maricic, T., Gunther, V., Georgiev, O., Gehre, S., Curlin, M., Schreiweis, C., Naumann, R., Burbano, H.A., Meyer, M., Lalueza-Fox, C., et al. (2013). A recent evolutionary change affects a regulatory element in the human FOXP2 gene. Mol Biol Evol 30, 844-852.
3. Sobreira, N., Schiettecatte, F., Valle, D., and Hamosh, A. (2015). GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Human mutation 36, 928-930.
4. Firth, H.V., Richards, S.M., Bevan, A.P., Clayton, S., Corpas, M., Rajan, D., Van Vooren, S., Moreau, Y., Pettett, R.M., and Carter, N.P. (2009). DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. Am J Hum Genet 84, 524-533.
5. Arachchi, H., Wojcik, M.H., Weisburd, B., Jacobsen, J.O.B., Valkanas, E., Baxter, S., Byrne, A.B., O'Donnell-Luria, A.H., Haendel, M., Smedley, D., et al. (2018). matchbox: An open-source tool for patient matching via the Matchmaker Exchange. Hum Mutat 39, 1827-1834.
6. Deciphering Developmental Disorders, S. (2015). Large-scale discovery of novel genetic causes of developmental disorders. Nature 519, 223-228.
7. Farwell, K.D., Shahmirzadi, L., El-Khechen, D., Powis, Z., Chao, E.C., Tippin Davis, B., Baxter, R.M., Zeng, W., Mroske, C., Parra, M.C., et al. (2015). Enhanced utility of family-centered diagnostic exome sequencing with inheritance model-based analysis: results from 500 unselected families with undiagnosed genetic conditions. Genet Med 17, 578-586.
8. Hempel, M., Cremer, K., Ockeloen, C.W., Lichtenbelt, K.D., Herkert, J.C., Denecke, J., Haack, T.B., Zink, A.M., Becker, J., Wohlleber, E., et al. (2015). De Novo Mutations in CHAMP1 Cause Intellectual Disability with Severe Speech Impairment. Am J Hum Genet 97, 493-500.
9. Flex, E., Niceta, M., Cecchetti, S., Thiffault, I., Au, M.G., Capuano, A., Piermarini, E., Ivanova, A.A., Francis, J.W., Chillemi, G., et al. (2016). Biallelic Mutations in TBCD, Encoding the Tubulin Folding Cofactor D, Perturb Microtubule Dynamics and Cause Early-Onset Encephalopathy. Am J Hum Genet 99, 962-973.
10. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature 536, 285-291.
11. Lelieveld, S.H., Reijnders, M.R., Pfundt, R., Yntema, H.G., Kamsteeg, E.J., de Vries, P., de Vries, B.B., Willemsen, M.H., Kleefstra, T., Lohner, K., et al. (2016). Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. Nature Neuroscience 19, 1194-1196.
12. Thevenon, J., Duffourd, Y., Masurel-Paulet, A., Lefebvre, M., Feillet, F., El Chehadeh-Djebbar, S., St-Onge, J., Steinmetz, A., Huet, F., Chouchane, M., et al. (2016). Diagnostic odyssey in severe neurodevelopmental disorders: toward clinical whole-exome sequencing as a first-line diagnostic test. Clin Genet 89, 700-707.
13. Smith, E.D., Radtke, K., Rossi, M., Shinde, D.N., Darabi, S., El-Khechen, D., Powis, Z., Helbig, K., Waller, K., Grange, D.K., et al. (2017). Classification of Genes: Standardized Clinical Validity Assessment of Gene-Disease Associations Aids Diagnostic Exome Analysis and Reclassifications. Hum Mutat 38, 600-608.
14. Bauer, C.K., Calligari, P., Radio, F.C., Caputo, V., Dentici, M.L., Falah, N., High, F., Pantaleoni, F., Barresi, S., Ciolfi, A., et al. (2018). Mutations in KCNK4 that Affect Gating Cause a Recognizable Neurodevelopmental Syndrome. Am J Hum Genet 103, 621-630.
15. Bouman, A., Waisfisz, Q., Admiraal, J., van de Loo, M., van Rijn, R.R., Micha, D., Oostra, R.J., and Mathijssen, I.B. (2018). Homozygous DMRT2 variant associates with severe rib malformations in a newborn. Am J Med Genet A 176, 1216-1221.

16. Gibson, K.M., Nesbitt, A., Cao, K., Yu, Z., Denenberg, E., DeChene, E., Guan, Q., Bhoj, E., Zhou, X., Zhang, B., et al. (2018). Novel findings with reassessment of exome data: implications for validation testing and interpretation of genomic data. Genet Med 20, 329-336.
17. Suzuki, T., Behnam, M., Ronasian, F., Salehi, M., Shiina, M., Koshimizu, E., Fujita, A., Sekiguchi, F., Miyatake, S., Mizuguchi, T., et al. (2018). A homozygous NOP14 variant is likely to cause recurrent pregnancy loss. J Hum Genet 63, 425-430.
18. Deriziotis, P., Graham, S.A., Estruch, S.B., and Fisher, S.E. (2014). Investigating protein-protein interactions in live cells using bioluminescence resonance energy transfer. J Vis Exp. 87:e51438.
19. Jauch, R., Choo, S.H., Ng, C.K., and Kolatkar, P.R. (2011). Crystal structure of the dimeric Oct6 (POU3f1) POU domain bound to palindromic MORE DNA. Proteins 79, 674-677.
20. Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. J Mol Graph 8, 52-56, 29.
21. Krieger, E., Koraimann, G., and Vriend, G. (2002). Increasing the precision of comparative models with YASARA NOVA--a self-parameterizing force field. Proteins 47, 393-402.