

1 High-throughput splicing assays identify missense and silent splice-disruptive *POU1F1* variants  
2 underlying pituitary hormone deficiency

3  
4 Peter Gergics<sup>1a</sup>, Cathy Smith<sup>1a</sup>, Hironori Bando<sup>1,12a</sup>, Alexander A. L. Jorge<sup>2a</sup>, Denise Rockstroh-  
5 Lippold<sup>3a, 13</sup>, Sebastian Vishnopska<sup>4a</sup>, Frederic Castinetti<sup>5a</sup>, Mariam Maksutova<sup>1</sup>, Luciani  
6 Renata Silveira Carvalho<sup>6</sup>, Julia Hoppmann<sup>3, 14</sup>, Julian Martinez Mayer<sup>4</sup>, Frédérique Albarel<sup>5</sup>,  
7 Debora Braslavsky<sup>7</sup>, Ana Keselman<sup>7</sup>, Ignacio Bergadá<sup>7</sup>, Marcelo Martí<sup>8</sup>, Alexandru Saveanu<sup>5</sup>,  
8 Anne Barlier<sup>5</sup>, Rami Abou Jamra<sup>11</sup>, Michael H. Guo<sup>9, 15</sup>, Andrew Dauber<sup>10,16</sup>, Marilena  
9 Nakaguma<sup>6</sup>, Berenice B Mendonça<sup>6</sup>, A Bilge Ozel<sup>1</sup>, Qing Fang<sup>1</sup>, Qianyi Ma<sup>1</sup>, Jun Z. Li<sup>1</sup>, Thierry  
10 Brue<sup>5\*</sup>, María Ines Pérez Millán<sup>4\*</sup>, Jacob O. Kitzman<sup>1\*</sup>, Ivo JP Arnhold<sup>6\*</sup>, Roland Pfaeffle<sup>3,11\*</sup>,  
11 Sally A. Camper<sup>1\*</sup>

12  
13 co-first author <sup>(a)</sup>

14 \* co-corresponding authors

15

16 <sup>1</sup> Department of Human Genetics, University of Michigan, Ann Arbor, MI, USA

17 <sup>2</sup> Genetic Endocrinology Unit (LIM25), Division of Endocrinology, Hospital das Clinicas da  
18 Faculdade de Medicina da Universidade de São Paulo (FMUSP), São Paulo, Brazil

19 <sup>3</sup> Department of Women's and Child Health, Division of Pediatric Endocrinology, University  
20 Hospital Leipzig, Leipzig, Germany

21 <sup>4</sup> Instituto de Biociencias, Biotecnología y Biología Traslacional (IB3), Departamento de  
22 Fisiología, Biología Molecular y Celular, Facultad de Ciencias Exactas y Naturales, Universidad  
23 de Buenos Aires, Ciudad de Buenos Aires, Argentina

24 <sup>5</sup> Aix Marseille University II, Hôpital Conception, Marseille, France

25 <sup>6</sup> Developmental Endocrinology Unit, Laboratory of Hormones and Molecular Genetics LIM/42,  
26 Division of Endocrinology, Hospital das Clinicas da Faculdade de Medicina da Universidade de  
27 São Paulo (FMUSP), São Paulo, Brazil

28 <sup>7</sup> Centro de Investigaciones Endocrinológicas “Dr. César Bergadá”, (CEDIE), FEI – CONICET –  
29 División de Endocrinología, Hospital de Niños Ricardo Gutiérrez, Ciudad de Buenos Aires,  
30 Argentina

31 <sup>8</sup> Instituto de Química Biología en Exactas y Naturales (IQUIBICEN-CONICET), Facultad de  
32 Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad de Buenos Aires,  
33 Argentina.

34 <sup>9</sup> Division of Endocrinology, Boston Children's Hospital and Department of Genetics, Harvard  
35 Medical School, Boston, Massachusetts, USA

36 <sup>10</sup> Cincinnati Center for Growth Disorders, Division of Endocrinology, Cincinnati Children's  
37 Hospital Medical Center, Cincinnati, Ohio, USA

38 <sup>11</sup> Institute of Human Genetics, University of Leipzig Hospitals and Clinics, Leipzig, Germany

39 <sup>12</sup> Current Address: Division of Diabetes and Endocrinology, Kobe University Graduate School  
40 of Medicine, Kobe, Japan

41 <sup>13</sup> Current Address: Beckman Coulter, Krefeld, Germany

42 <sup>14</sup> Current Address: Division of Paediatric Endocrinology and Diabetes, Department of  
43 Paediatric and Adolescent Medicine, University of Luebeck, Germany

44 <sup>15</sup> Current Address: Department of Neurology, University of Pennsylvania, Philadelphia, PA,  
45 19104, USA

46 <sup>16</sup> Current Address: Division of Endocrinology, Children's National Hospital, Washington, D.C.  
47 20010, USA

48

49 Email, ORCID ID

50 Peter Gergics [gergicsp@gmail.com](mailto:gergicsp@gmail.com), <http://orcid.org/0000-0002-3480-896X>

51 Cathy Smith: [smithcat@umich.edu](mailto:smithcat@umich.edu), <https://orcid.org/0000-0003-1518-9320>  
52 Hironori Bando: [hbando@med.kobe-u.ac.jp](mailto:hbando@med.kobe-u.ac.jp), <https://orcid.org/0000-0002-7421-2714>  
53 Alexander A L Jorge: [alexj@usp.br](mailto:alexj@usp.br), <https://orcid.org/0000-0003-2567-7360>  
54 Denise Rockstoh-Lippold: [d.rockstroh@gmx.de](mailto:d.rockstroh@gmx.de), <https://orcid.org/0000-0003-1874-2363>.  
55 Sebastian Vishnopolska: [sebasvishno@gmail.com](mailto:sebasvishno@gmail.com) <https://orcid.org/0000-0002-1729-9948>  
56 Frederic Castinetti: [Frederic.CASTINETTI@ap-hm.fr](mailto:Frederic.CASTINETTI@ap-hm.fr), <https://orcid.org/0000-0002-1808-8800>  
57 Mariam Maksutova: [mmaksut@umich.edu](mailto:mmaksut@umich.edu)  
58 Luciani Carvalho: [lucianic@gmail.com](mailto:lucianic@gmail.com), <https://orcid.org/0000-0002-7313-2130>  
59 Julia Hoppmann: [Julia.Hoppmann@uksh.de](mailto:Julia.Hoppmann@uksh.de), <https://orcid.org/0000-0003-1436-1707>  
60 Julian Martinez Mayer: [julianmm.7@gmail.com](mailto:julianmm.7@gmail.com), <https://orcid.org/0000-0003-3696-7291>  
61 Frederique Albarel: [Frederique.ALBAREL@ap-hm.fr](mailto:Frederique.ALBAREL@ap-hm.fr), <https://orcid.org/0000-0002-6995-2590>  
62 Debora Braslavsky: [dbraslavsky@cedie.org.ar](mailto:dbraslavsky@cedie.org.ar), <http://orcid.org/0000-0002-4501-7610>  
63 Ana Keselman: [keselman@cedie.org.ar](mailto:keselman@cedie.org.ar), <http://orcid.org/0000-0002-5612-0445>  
64 Ignacio Bergadá: [ibergada@cedie.org.ar](mailto:ibergada@cedie.org.ar), <https://orcid.org/0000-0001-6546-1949>  
65 Marcelo Martí: [marti.marcelo@gmail.com](mailto:marti.marcelo@gmail.com), <https://orcid.org/0000-0002-7911-9340>  
66 Alexandru Saveanu: [Alexandru.SAVEANU@ap-hm.fr](mailto:Alexandru.SAVEANU@ap-hm.fr), <https://orcid.org/0000-0001-6365-1538>  
67 Anne Barlier: [anne.barlier@ap-hm.fr](mailto:anne.barlier@ap-hm.fr), <https://orcid.org/0000-0002-3740-6173>  
68 Rami Abou Jamra: [rami.aboujamra@medizin.uni-leipzig.de](mailto:rami.aboujamra@medizin.uni-leipzig.de), <https://orcid.org/0000-0002-1542-1399>  
69  
70 Michael H. Guo: [mhguo1@gmail.com](mailto:mhguo1@gmail.com), <https://orcid.org/0000-0002-1357-6389>  
71 Andrew Dauber: [adauber@childrensnational.org](mailto:adauber@childrensnational.org), <https://orcid.org/0000-0003-4890-0262>  
72 Marilena Nakaguma: [mari.nakaguma@gmail.com](mailto:mari.nakaguma@gmail.com), <https://orcid.org/0000-0001-7962-8156>  
73 Berenice Mendonça: [beremen@usp.br](mailto:beremen@usp.br), <https://orcid.org/0000-0003-1762-1084>  
74 A Bilge Ozel: [aozel@umich.edu](mailto:aozel@umich.edu), <http://orcid.org/0000-0002-1112-4258>  
75 Qing Fang: [qing.fang@regeneron.com](mailto:qing.fang@regeneron.com), <https://orcid.org/0000-0002-8808-8523>  
76 Qianyi Ma: [qzm@umich.edu](mailto:qzm@umich.edu), <http://orcid.org/0000-0002-6555-7446>  
77 Jun Z. Li: [junzli@med.umich.edu](mailto:junzli@med.umich.edu), <http://orcid.org/0000-0001-6727-0812>.  
78 Thierry Brue: [Thierry.BRUE@ap-hm.fr](mailto:Thierry.BRUE@ap-hm.fr), <https://orcid.org/0000-0001-8482-6691>  
79 Maria Inés Pérez Millán: [mipmillan@gmail.com](mailto:mipmillan@gmail.com), <https://orcid.org/0000-0002-0486-109X>  
80 Jacob O. Kitzman: [kitzmanj@umich.edu](mailto:kitzmanj@umich.edu), <https://orcid.org/0000-0002-6145-882X>  
81 Ivo J P Arnhold: [iarnhold@usp.br](mailto:iarnhold@usp.br), <https://orcid.org/0000-0003-1739-1354>  
82 Roland Pfäffle: [Roland.Pfaeffle@medizin.uni-leipzig.de](mailto:Roland.Pfaeffle@medizin.uni-leipzig.de), <https://orcid.org/0000-0001-6754-3681>  
83 Sally A Camper: [scamper@umich.edu](mailto:scamper@umich.edu), <https://orcid.org/0000-0001-8556-3379>

84 The authors have no conflicts of interest to disclose

85

86 Key words: alternative splicing, growth hormone deficiency, multiplexed assays of variant  
87 effects, transcriptional regulation

88 **Abstract**

89 Pituitary hormone deficiency occurs in ~1:4,000 live births. Approximately 3% of the cases are  
90 due to mutations in the alpha isoform of POU1F1, a pituitary-specific transcriptional activator.  
91 We found four separate heterozygous missense variants in unrelated hypopituitarism patients  
92 that were predicted to affect a minor isoform, POU1F1 beta, which can act as a transcriptional  
93 repressor. These variants retain repressor activity, but they shift splicing to favor the expression  
94 of the beta isoform, resulting in dominant negative loss of function. Using a high throughput  
95 splicing reporter assay, we tested 1,080 single nucleotide variants in *POU1F1*. We identified  
96 113 splice disruptive variants, including 23 synonymous variants. We evaluated separate  
97 cohorts of hypopituitarism patients and found two different synonymous splice disruptive  
98 variants that co-segregate with hypopituitarism. This study underlines the importance of  
99 evaluating the impact of variants on splicing and provides a catalog for interpretation of variants  
100 of unknown significance in the *POU1F1* gene.

101 **Introduction**

102 *POU1F1* (formerly PIT-1) is a signature pituitary transcription factor that directly  
103 regulates the transcription of growth hormone (*GH*), prolactin (*PRL*), and both the alpha (*CGA*)  
104 and beta (*TSHB*) subunits of thyroid stimulating hormone<sup>1,2</sup>. In mice, *Pou1f1* is expressed after  
105 the peak expression of *Prop1* at E14.5 and remains expressed into adulthood<sup>3,4</sup>. A well-  
106 characterized mutant of *Pou1f1* (*Pit1<sup>dw/dw</sup>*) carries a spontaneous missense mutation (p.W251C)  
107 in the homeodomain that disrupts DNA binding<sup>4,5</sup>. The homozygous mutant mice have no  
108 somatotrophs, lactotrophs or thyrotrophs except for the *Pou1f1*-independent rostral tip  
109 thyrotrophs<sup>4,6-8</sup>. In humans, loss of *POU1F1* function typically results in GH, TSH and PRL  
110 deficiency<sup>9</sup>.

111 *POU1F1* undergoes an evolutionarily conserved program of alternative splicing<sup>10,11</sup>,  
112 resulting in a predominant isoform, alpha, that acts as a transcriptional activator and a minor  
113 isoform, beta, that acts as a transcriptional repressor<sup>12-14</sup>. In the human pituitary gland, the beta  
114 isoform comprises approximately 1-3% of *POU1F1* transcripts<sup>10,15</sup>. The *POU1F1* beta isoform  
115 transcript is created by utilization of an alternative splice acceptor sequence for exon 2, located  
116 78 bp upstream of the alpha acceptor, resulting in a 26 amino acid insertion that encodes an  
117 ETS1 binding domain. This insertion, which is absent in the alpha isoform, disrupts the  
118 transactivation domain at amino acid 48. The *POU1F1* alpha and beta isoforms have different  
119 activities depending on the context of the target gene<sup>12</sup>. For example, the *POU1F1* alpha  
120 isoform activates its own expression, but the beta isoform does not, and the beta isoform  
121 interferes with alpha isoform mediated activation<sup>14</sup>.

122 The first case of a recessive *POU1F1* loss of function was described in a patient with  
123 CPHD born to consanguineous parents<sup>16</sup>, and since then thirty-seven unique variants in  
124 *POU1F1* have been reported in patients with CPHD or IGHD<sup>17-23</sup>. A few dominant negative  
125 mutations have been reported: p.P76L alters the transactivation domain and causes completely  
126 penetrant IGHD<sup>24</sup>, p.K216E interferes with the ability of *POU1F1* to interact with retinoic acid

127 receptors and p300<sup>25</sup>, and p.R271W interferes with the ability of POU1F1 to be tethered to the  
128 nuclear matrix through MATR3, SATB1 and CTNNB1<sup>26</sup>. All of the reported mutations are  
129 located in domains shared by the alpha and beta isoforms of POU1F1 and were functionally  
130 tested using the alpha isoform only.

131         We found four missense variants, in four independent families, that shift splicing to favor  
132 the POU1F1 beta isoform almost exclusively, while retaining its transcriptional repressor activity  
133 on the *POU1F1* enhancer. We used a high throughput assay to identify 113 variants in and  
134 around exon 2 that cause exon skipping, cryptic splicing and isoform switching. We used this  
135 catalog to evaluate additional families with hypopituitarism and identified two unrelated patients  
136 carrying synonymous *POU1F1* variants that affect its splicing without changing the amino acid  
137 sequence. This study underscores the importance of evaluating splicing defects as a disease  
138 mechanism.

## 139 **Methods**

### 140 **Patients**

141 The studies were approved by ethical committees: the local Comit e de  tica e Pesquisa da  
142 Faculdade de Medicina da Universidade de S o Paulo (CEP-FMUSP) and the national Comit e  
143 nacional de  tica em pesquisa (CONEP) CAAE, 06425812.4.0000.0068; the Ethics Committee  
144 of the Faculty of Medicine, University of Leipzig (UL), Karl-Sudhoff-Institute for Medical History  
145 and Natural Sciences, K the-Kollwitz-Stra e 82, 04109 Leipzig, Germany; and the Comit e de  
146  tica en Investigaci n (Research Ethics Committee) of the Hospital de Ni os Ricardo Gutierrez  
147 (HNG), Gallo 1330, Ciudad aut noma de Buenos Aires, Argentina (CEI N  16.06). The  
148 GENHYPOPIT network collected anonymized information in a database declared to health  
149 authorities in accordance with local regulations at Aix-Marseille Universit  (AMU) - Conception  
150 Hospital (Assistance Publique - H pitaux de Marseille, AP-HM), and a declaration was made to  
151 the National Commission for Data Protection and Liberties (CNIL-France): 1991429 v 0.  
152 Patients or their parents signed a written informed consent to participate. Families 1, 3, and 6  
153 are historical cases that were referred to the GENHYPOPIT network for genetic testing. Limited  
154 information is available for Families 1 and 3, and they were lost for follow up. The University of  
155 Michigan Institutional Review Board (UM) found the study exempt because patient DNA  
156 samples were anonymized before exome sequencing at UM.

### 157 **DNA sequencing of patient samples**

158 Individuals from Families 1, 2, 4, and 5 underwent whole exome sequencing (WES).  
159 Representative *POU1F1* variants in Family 3 and 6 were discovered in a traditional CPHD  
160 candidate gene screening using Sanger sequencing (*PROP1*, *POU1F1*, *LHX3* and *LHX4*).  
161 WES of Families 1 and 5 was carried out at University of Michigan as previously described<sup>17</sup>.  
162 WES of Family 2 was performed at the Broad Institute as previously described<sup>27</sup>. WES of  
163 Family 4 was performed at the Institute of Human Genetics at University of Leipzig.

### 164 **Expression vectors and cell culture**

165 The open reading frame of either *POU1F1* isoform alpha (NM\_000306.3) or beta  
166 (NM\_001122757.2) was cloned into pcDNA3.1+/C-(K)-DYK. Site directed mutagenesis was  
167 used to obtain each of the variant *POU1F1* beta isoforms: p.S50A, p.I51S, p.L52W, and  
168 p.S53A (Genscript). A firefly luciferase reporter gene was constructed in pNBm81-luc with 14  
169 kb of the mouse *Pou1f1* 5' flanking sequences that includes early and late enhancers and the  
170 promoter, and 13 bp of the 5'UTR. Cloning was performed with Infusion HD (Clontech) or  
171 NEBuilder HiFi DNA Assembly (New England Biolabs). Plasmid sequences were confirmed by  
172 Sanger sequencing. The pRL-TK renilla (Promega) was used as a normalization control and  
173 pcDNA3.1(-) (Thermo-Fisher) to keep the total DNA constant. COS-7 cells were purchased  
174 from the American Type Culture Collection. Cells were maintained in Dulbecco's modified eagle  
175 medium (DMEM, Gibco, Grand Island, NY, USA) containing 10% fetal bovine serum and pen-  
176 strep (Gibco). Plasmids were transiently transfected into COS-7 cells using ViaFect  
177 Transfection Reagent (Promega, Madison, WI, USA). Luciferase activities were measured as  
178 suggested by the manufacturer (Dual-luciferase assay system; Promega).

#### 179 Exon trapping assay

180 Human *POU1F1* exon 2, flanked by partial intron 1 (85 bp upstream) and intron 2 (178 bp  
181 downstream), was cloned into the *BamHI* cloning site of the pSPL3 plasmid (Invitrogen) to  
182 create an exon trapping plasmid with a total insert size of 413 bp. Similarly, a minigene exon  
183 trapping plasmid was constructed that included the last 85 bp of intron 1 and the first 85 bp of  
184 intron 5, for a total insert size of 3,442 bp including exons 2, 3, 4, and 5. Site directed  
185 mutagenesis was used to create the desired variants. Plasmids were transiently transfected  
186 into COS-7 cells. Total RNA was purified with RNeasy mini (Qiagen). After reverse  
187 transcription, we analyzed exon trapping using RT-PCR with following primers; Primers SD6  
188 Forward (5'-TCT GAG TCA CCT GGA CAA CC- 3') and SA2 reverse (5'- ATC TCA GTG GTA  
189 TTT GTG AGC -3')<sup>28</sup>.

#### 190 *POU1F1* Saturation Mutagenesis



191 The cloned *POU1F1* fragment in pSPL3 was divided into four overlapping tiles of 150 bp each,  
192 spanning exon 2 plus flanking introns (79 bp upstream to 131 bp downstream). Mutant tile  
193 libraries containing every possible single nucleotide variant were synthesized as a single  
194 150mer oligonucleotide pool by Twist Bio. HiFi Assembly was used to replace each wild type  
195 tile with the respective mutant tile library amplified from the oligo pool. The resulting mutant  
196 minigene library pools were transformed in 10b *E. coli* (New England Biolabs), with a minimum  
197 coverage of 90 clones per mutation.

#### 198 Mutant library barcoding and sequencing

199 To tag each mutant minigene clone with a unique barcode, a random barcode sequence ( $N_{20}$ )  
200 was inserted by HiFi Assembly into the MscI site within the common 3' UTR. Subassembly  
201 sequencing<sup>29</sup> was used to pair each 3' UTR barcode with its linked variant(s) in cis. Briefly, a  
202 fragment starting with the *POU1F1* insert and ending at the  $N_{20}$  barcode (2.2 kb downstream)  
203 were amplified from the plasmid library DNA by PCR using 5'-phosphorylated primers. The  
204 resulting linear fragment was re-circularized by intramolecular ligation using T4 DNA ligase  
205 (NEB), to bring each barcode in close proximity to the mutagenized region. From this re-  
206 circularized product, paired-end amplicon sequencing libraries were generated, such that each  
207 reverse read contained a plasmid barcode and the paired forward read contained a sequence  
208 from the associated *POU1F1* insert. Barcode reads were clustered with starcode<sup>30</sup> (arguments  
209 "-d 1 -r 3") to generate a catalog of known barcodes. Variants were called within each barcode  
210 group using freebayes<sup>31</sup> and filtered to require majority support, and read depth  $\geq 4$  along the  
211 entire region targeted for mutagenesis.

#### 212 Pooled exon-trap transfection and RNA-seq

213 COS-7 cells were plated at  $5 \times 10^6$  cells/60 mm plate. Each was transfected with 4  $\mu$ g of the  
214 barcoded mutant exon-trap library using ViaFect reagent (3:1 ratio to DNA). After 24 hrs, RNA  
215 was purified as above, and 5  $\mu$ g of total RNA was used to prepare first-strand cDNA using the  
216 SuperScript III First-Strand Synthesis kit (Invitrogen) with oligo dT primers. Spliced transcript



243 HAL delta\_psi scores<sup>33</sup>, SPANR zdelta\_psi scores<sup>34</sup>, SpliceAI ds\_max scores<sup>35</sup>, and MMSplice  
244 delta\_logit\_psi scores<sup>36</sup> were obtained from their original publications without modification. To  
245 compute per-variant  $\Delta$ ESRseq scores<sup>37</sup>, we took the difference between the mean ESRseq z-  
246 scores of hexamers overlapping a variant position from that of hexamers overlapping the  
247 corresponding wildtype position. Precision-recall curves were obtained to summarize each  
248 algorithm's ability to predict the experimental determination of splice disruptiveness. For  
249 algorithms which output signed scores, area under the curve (prAUC) was separately computed  
250 using signed and absolute scores as input and the higher prAUC was taken.

#### 251 Selection of candidate RNA binding proteins (RBP)

252 RNACompete z-scores<sup>38</sup> were obtained from the cisBP-RNA database ([http://cisbp-](http://cisbp-rna.cabr.utoronto.ca)  
253 [rna.cabr.utoronto.ca](http://cisbp-rna.cabr.utoronto.ca)). At each position, wild-type and variant-containing z scores were taken as  
254 the maximum among the overlapping kmers, and the difference taken between the wild-type  
255 and variant scores. Motifs with high scoring matches (wildtype  $z \geq 3$ ) to the wild-type sequence  
256 in the beta variant cluster (c.143-1 to c.167) were then pursued further.

#### 257 Data availability

258 Raw sequence reads and processed counts are available at GEO (accession tbd). Jupyter  
259 notebooks to reproduce the processed dataset from raw counts are posted at GitHub (URL tbd).

## 260 **Results**

### 261 **Mutations in the POU1F1 beta coding region cause hypopituitarism**

262 We initially focused on four cases of hypopituitarism from different cohorts in Europe and South  
263 America (**Fig. 1A**). Affected individuals' presentation was variable, ranging from multiple  
264 hormone deficiency with pituitary stalk interruption (Family 1) to isolated GH deficiency (Family  
265 2) (**Table 1, Suppl. Fig. 1**). The affected individuals had severe short stature and responded  
266 well to GH therapy (**Fig. 1B**). To identify causal variants, we performed whole exome  
267 sequencing (WES) for individuals in three families. Combined with conventional Sanger  
268 sequencing in another family, this revealed four missense variants in exon 2 of the *POU1F1*  
269 beta isoform, each in an unrelated family (**Figure 1A, 2A**). The four patient *POU1F1* missense  
270 variants are absent from gnomAD and in-house population-matched exome databases<sup>39,40</sup>, and  
271 they are predicted to be damaging by several bioinformatic algorithms (**Table 1**). Remarkably,  
272 these variants clustered in four consecutive codons: c.148T>G (p.S50A), c.152T>G (p.I51S),  
273 c.155T>G (p.L52W), and c.157T>G (p.S53A) in NM 001122757.3 (**Table 1**). Only one of these  
274 (c.155T>G, Family 3) appears to be *de novo*; the others were dominantly inherited and co-  
275 segregate with hypopituitarism phenotypes, except for c.148T>G which was inherited from the  
276 apparently unaffected parent in Family 1, indicating that if causal, this variant is incompletely  
277 penetrant. The other parent in Family 1, the two affected children, and one unaffected relative  
278 also carried a variant of uncertain significance, *SIX3* p.P74R. No other variants in known  
279 hypopituitarism genes were detected.

### 280 **Sequence variants retain POU1F1 beta isoform repressor function**

281 We first asked whether these variants disrupt the ability of POU1F1 to bind and transactivate its  
282 own promoter<sup>41</sup> using a transient transfection assay (**Figure 2B**). As expected, a *Pou1f1*  
283 promoter reporter was strongly activated when co-transfected with cDNA of POU1F1 alpha  
284 isoform, which does not include the variant sites. Neither WT POU1F1 beta isoform, nor any of  
285 the four patient missense variants, showed significant activation of the *Pou1f1*-luc reporter.

286 Consistent with a repressive role for POU1F1 beta, co-transfection with alpha at a 1:1 ratio  
287 significantly suppressed activation compared to the equivalent amount of alpha isoform alone.  
288 The four POU1F1 beta variants and WT beta repressed POU1F1 alpha activity to a similar  
289 degree.

#### 290 **Patient missense variants disrupt normal *POU1F1* splicing to favor the beta isoform**

291 Alpha is normally the predominant *POU1F1* isoform, but its splice acceptor is predicted to be  
292 much weaker than the beta isoform acceptor 78 bp upstream (MaxEntScan<sup>42</sup>; scores, alpha: -  
293 3.63, beta: 6.96) (**Fig. 2C**). The beta isoform splice acceptor sequence and coding region are  
294 evolutionarily conserved in mammals and birds (**Fig. 2D**). We reasoned that splice repressor  
295 and/or enhancer sequences in *POU1F1* may dictate the normal balance of alpha over beta  
296 isoforms, and these may be disrupted by the four patient T>G transversions. To test the effect  
297 of these variants directly, we cloned *POU1F1* exon 2 and portions of the flanking introns into the  
298 exon trap splice reporter pSPL3 and introduced each variant by site directed mutagenesis (**Fig.**  
299 **2E**). These small minigenes were transfected into COS-7 cells, and RNA was analyzed by RT-  
300 PCR. The wild type minigene produced almost exclusively alpha isoform, while each of the  
301 patient variants predominantly produced the beta isoform (**Fig. 2F**). We also tested splicing with  
302 larger minigenes, which contain portions of intron 1 and intron 5 with intact exons 2, 3, 4 and 5  
303 as well as introns 2, 3 and 4, and obtained similar results, indicating the additional sequence  
304 context does not strongly influence the observed splicing pattern (**Fig. 2G**). We also tested two  
305 previously reported *POU1F1* variants in the longer minigene context. The c.214+1 G>T caused  
306 skipping of exon 2, as expected, resulting in an in-frame POU1F1 protein that lacks 80% of the  
307 transactivation domain<sup>43</sup>. This variant is associated with mild hypopituitarism. The p.P76L  
308 variant is located in the transactivation domain, enhances POU1F1 interaction with other  
309 proteins, and is associated with severe, dominant IGHD<sup>24</sup>. The effect of this variant on splicing  
310 had not been assessed previously, and we found that it produced predominantly alpha isoform  
311 expression, indistinguishable from wild type.

## 312 **Saturation mutagenesis screen for splice disruptive effects**

313 We set out to systematically identify splice disruptive variants in *POU1F1* exon 2 using a  
314 massively parallel splice reporter assay. We designed oligonucleotide pools containing every  
315 possible single nucleotide variant across exon 2 (150 bp) and 210 bp of the flanking introns  
316 (N=1080), and generated libraries in which these allelic series replaced the wild type *POU1F1*  
317 fragment in the pSPL3 reporter. To track the splicing outcomes associated with each mutation,  
318 we placed a degenerate 20mer barcode in the downstream 3' UTR. The mutant plasmid library  
319 was subjected to subassembly sequencing<sup>29</sup> to establish the pairing between each unique  
320 barcode and its associated *POU1F1* mutation. In total, the mutant library contained 255,023  
321 distinct barcoded clones, among which 188,772 (74.0%) had exactly one programmed mutation.  
322 Nearly every targeted mutation appeared in this library (1070/1080, 99.1%), with a high degree  
323 of redundancy (median 75.0 distinct barcodes/mutation, **Suppl. Fig. 2**).

324 The splice reporter library was transfected as a pool into COS-7 cells and processed  
325 similarly to the single mutation constructs. Spliced reporter transcripts were read out *en masse*  
326 using paired-end RNA-seq (**Fig. 3A**), with each forward read measuring an individual splicing  
327 outcome and the paired reverse read containing the 3' UTR barcode which identifies the  
328 mutation(s) present in the primary transcript. We performed 14 biological replicates, across  
329 which 94.2% (81.8-93.4%, mean 87.4%) of clone library single nucleotide variant associated  
330 barcodes were detected. As expected, alpha was the predominant *POU1F1* isoform (69.2% of  
331 reads overall), followed by exon 2 skipping (25.6%), and beta (1.6%). We created a catch-all  
332 category ('Other') for the remaining reads (3.6%) derived from the 262 other isoforms detected.  
333 Most of those noncanonical isoforms were only scarcely used; among them, the top 20  
334 accounted for >80% of the reads from that category. For each *POU1F1* variant, a percent  
335 spliced in (PSI) value was computed for each isoform (alpha, skip, beta, other), averaged over  
336 the associated barcodes. PSI values were highly reproducible across replicates (median  
337 pairwise Pearson's *r*: 0.92; **Suppl. Fig. 3**).

338 **Splice disruptive variants (SDVs) across *POU1F1* exon 2**

339 We assessed the splicing effect of 1,070 single nucleotide variants (**Figure 3A** and  
340 **Suppl. Fig. 4**). Of these, 113 (10.6%) were splice disruptive variants (SDVs), which we defined  
341 as those which increased usage of beta, skip, or other isoforms by  $\geq 1.5$  standard deviations.  
342 SDVs leading to increased skipping were the most frequent ( $n = 69/113$ ; 61.1%), followed by  
343 those creating other isoforms ( $n = 34/113$ ; 30.1%), and those which increased beta usage ( $n =$   
344  $29/113$ ; 25.7%), with some variants ( $n = 19/113$ ; 16.8%) impacting usage of multiple isoforms  
345 (**Suppl. Fig. 5**). Variants leading to each outcome tended to cluster in distinct regions; notably,  
346 the beta-increasing SDVs were located near the 5' end of the beta isoform. Intronic SDVs  
347 tended to lead to skipping. A few variants that increased skipping were scattered across exon 2,  
348 and there was some enrichment in the 5' end of the beta isoform coding region, but most were  
349 enriched near splice donor and acceptor sites: 27 of 28 intronic SDVs were within  $\pm 20$  bp of  
350 exon 2.

351 SDVs which created novel ("other") isoforms were nearly all located within the coding  
352 region unique to the beta isoform and at the alpha isoform acceptor site ( $n = 33/34$ ; 97.1%). Of  
353 these, most (27/34) create a novel acceptor AG dinucleotide that outcompetes the more distal,  
354 native alpha acceptor (**Suppl. Fig. 6**). Most of these (19/27) result in a frame-shifted transcript,  
355 and all are predicted to undergo non-sense mediated decay due to premature termination  
356 codons  $>50$  bp upstream of the final exon junction. In contrast, each of the six mutations in the  
357 alpha isoform acceptor dinucleotide activates another cryptic acceptor six bp downstream,  
358 leading to in-frame deletion of two codons (**Suppl. Fig. 7**). Out of 99 mutations creating a GT  
359 dinucleotide, only one was used as a novel splice donor, c.290:C>T located 4 bp upstream of  
360 the native donor.

361 We compared these SDVs to the variants found in Families 1-4. All four patient  
362 missense variants showed strongly increased beta isoform usage (beta isoform z-scores range:  
363 5.12-6.18), mirroring the results of individual minigene assays (**Fig. 3B**). Our results also

364 recapitulate previously described effects of two variants found in CPHD patients: first, an  
365 upstream intronic variant c.143-5:A>G<sup>44</sup> which led to increased beta usage and skipping (beta  
366 isoform  $z=3.48$ , skip  $z=3.13$ ; **Figure 3C**), and an essential splice donor variant c.292+1 G>T  
367 which led to near-complete skipping (skip  $z=5.05$ ; **Suppl. Fig. 7**)<sup>43</sup>.

368 We also examined the incidence of splice disruptive *POU1F1* variants in the general  
369 population. The gnomAD database contains 93 of the variants measured here; among those,  
370 eight (8.6%) are splice disruptive, all which are individually rare (minor allele frequency  $\leq$   
371  $1.6 \times 10^{-5}$ ; **Suppl. Fig. 8**). Overall, variants found in gnomAD were not significantly depleted for  
372 splice disruptive effects relative to randomly selected subsets of the tested single nucleotide  
373 variants ( $p=0.60$ , Fisher's Exact Test). Thus, *POU1F1* SDVs are tolerated to a similar extent as  
374 other predicted loss of function variants (stop gain, frameshift, splice site), which are observed  
375 throughout *POU1F1* at low frequencies in gnomAD.

#### 376 **Identification of a silent variant in *POU1F1* in a hypopituitarism patient**

377 We next examined the splicing impacts of synonymous variants, which would typically  
378 be given low priority during genetic screening due to their expected lack of coding impact. Of  
379 the 111 synonymous variants tested, 23 were splice disruptive (20.7%; **Suppl. Fig. 5**). We  
380 identified unrelated patients with IGHD carrying two of these synonymous SDV in the beta  
381 isoform coding region near the 5' end of exon 2 (**Fig. 3C**), both of which were absent in  
382 gnomAD and population-matched control databases. The first, c.150T>G (p.Ser50=), was  
383 found among an Argentinian cohort ( $n=171$ ) in a family with two individuals with severe short  
384 stature and IGHD (**Table 1**), for whom WES did not reveal any likely pathogenic variants in  
385 known CPHD or IGHD genes. The index case had pituitary hypoplasia, and the patient  
386 responded well to recombinant GH treatment. The second, c.153T>A (p.Ile51=), was found in a  
387 French family in relatives with severe IGHD. The parent's DNA was not available for testing,  
388 and the parent could be an unaffected carrier or an example of gonadal mosaicism. Each of



389 these two silent variants increased beta isoform usage (beta  $z=5.33$  for c.150T>G; beta  $z=1.79$   
390 for c.153T>A) to a degree similar to that of the four patient missense variants.

### 391 **Comparison to bioinformatic splicing effect predictions**

392 We examined how scores from splicing effect prediction algorithms compared with these  
393 experimental measurements. We scored each single nucleotide variant in the targeted region of  
394 *POU1F1* using SpliceAI<sup>35</sup>, MMSplice<sup>36</sup>, SPANR<sup>34</sup>, HAL<sup>33</sup> and ESRseq scores<sup>37</sup>. Among these,  
395 only SpliceAI predicted a high density of SDVs specific to the exon 2 beta region surrounding  
396 the patient variants (**Suppl. Fig. 9**). To benchmark each bioinformatic predictions, we took our  
397 SDV calls as a truth set and computed for each algorithm the area under the precision recall  
398 curve (**Suppl. Fig. 10**). SpliceAI was the most highly concordant with our results for both exonic  
399 variants (prAUC=0.790 vs other tools' range: 0.329-0.430) and intronic variants (prAUC=0.661  
400 versus other tools' range: 0.526-0.604). Nevertheless, SpliceAI disagreed with our  
401 measurements for numerous variants: at the minimum threshold needed to capture all six  
402 patient variants as disruptive (SpliceAI deltaMax score $\geq 0.19$ ), it achieved 67.3% sensitivity  
403 (n=37 SDVs not predicted by SpliceAI) and 97.2% specificity (n=27 variants predicted by  
404 SpliceAI but not identified by our assay) for predicting the SDVs we identified. The degree of  
405 concordance with SpliceAI was largely insensitive to the specific measurement threshold used  
406 to call variants as splice disruptive in our screen (**Suppl. Fig. 10**). Additional studies will be  
407 required to resolve the discordant predictions for variants observed during clinical screening.  
408

409 **Discussion**

410 We found six unrelated cases with CPHD or IGHD that can be explained by variants that  
411 shift splicing to favor the repressive beta isoform *POU1F1*. The missense variants, p.S50A,  
412 p.I51S, p.L52W, and p.S53A, retain repressive function. They act in a dominant negative  
413 manner by suppressing the ability of the *POU1F1* alpha isoform, expressed from the wild-type  
414 allele, to transactivate expression of *POU1F1* and other downstream target genes. Using  
415 saturation mutagenesis coupled to a high-throughput RNA-seq splicing readout, we  
416 systematically tested nearly every possible single nucleotide variant in or near *POU1F1* exon 2  
417 for splice disruptive potential (**Supplemental Table 1**). We identified 113 SDVs which similarly  
418 activate usage of the beta isoform or cause other aberrant splicing outcomes such as exon  
419 skipping.

420 This screen accurately identified the four missense variants we identified in patients. It  
421 also identified 23 synonymous splice disruptive variants, or over one-fifth of the possible  
422 synonymous variants in *POU1F1* exon 2. We identified two of these synonymous SDVs in  
423 unrelated families with IGHD, c.150T>G (p.Ser50=) and c.153T>A (p.Ile51=), each of which  
424 increased beta isoform usage similarly to the four patient missense variants that initially drew  
425 our attention. These findings underscore the need to closely examine variants for splice  
426 disruptive effects, particularly synonymous variants that could be overlooked by traditional  
427 exome sequencing filtering pipelines.

428 The clinical features varied amongst the six families, although they were consistent  
429 within a family. Families 1, 3, and 4 presented with CPHD, while Families 2, 5, and 6 had IGHD.  
430 Moreover, Family 4 developed hypocortisolism. The reason for this variability in presentation is  
431 unknown. However, there are precedents for variable clinical features and incomplete  
432 penetrance with other cases of hypopituitarism<sup>18</sup>. Approximately 50% of IGHD progresses to  
433 CPHD, and this can even occur when the mutated gene is only expressed in GH-producing cells,  
434 i.e. *GH1*<sup>45</sup>. Even individuals with the same *POU1F1* mutation (i.e. p.E230K) can present with

435 either IGHD or CPHD<sup>46</sup>, indicating a contributing role for genetic background, epigenetic, and/or  
436 environmental factors. Both affected relatives in Family 1 had stalk disruption, a phenotype not  
437 currently associated with any other *POU1F1* variants. This feature may be due to the presence  
438 of an additional variant in *SIX3*, p.P74R, that was carried by two unaffected relatives.  
439 Heterozygous loss of function of *SIX3* is associated with incompletely penetrant and highly  
440 variable craniofacial abnormalities, including CPHD and holoprosencephaly, and there is  
441 precedent in mice for *Six3* loss of function to exacerbate the phenotype caused by mutations in  
442 other CPHD genes such as *Hesx1*<sup>47-49</sup>.

443 Autosomal dominant inheritance is clear in Family 2, in which there were four affected  
444 individuals over three generations, as well as Families 4, 5, and 6. *POU1F1* acts as a  
445 heterodimer<sup>50</sup>. Some other dominant mutations in *POU1F1* act as negative effectors due to the  
446 ability of the mutant protein to interfere with the action of the wild type protein produced from the  
447 other allele<sup>25,51,52</sup>. The negative effect of *POU1F1* beta on the transactivation properties of  
448 *POU1F1* alpha are context dependent, with differential effects on *Gh*, *Prl* and *Pou1f1* reporter  
449 genes<sup>53</sup>. The strongest effect was reported for autoregulation of *POU1F1* expression via the  
450 distal, late enhancer; dampening the auto-activation of *POU1F1* expression, and adversely  
451 affecting differentiation of the entire *POU1F1* lineage and result in anterior lobe hypoplasia.

452 The lack of significant depletion for *POU1F1* SDVs among ostensibly healthy adult  
453 populations underscores the possibility of variable expressivity and/or penetrance for *POU1F1*  
454 splice-disruptive variants. This is consistent with the apparently unaffected parents in Families  
455 1 and 6. A subset of these variants, like c.222T>C which disrupts the alpha isoform acceptor  
456 and causes a frame-preserving two-codon deletion, may retain partial or complete function. Still  
457 others, may cause loss-of-function without dominant negative effects, and would not be  
458 expected to be strongly depleted.

459 In human genes, canonical splice site motifs contain less than half of the information  
460 content needed for proper splicing<sup>54</sup>. Additional specificity is provided by short (6-10 nt) motifs

461 termed exonic or intronic silencers and enhancers, which are bound by RNA binding proteins  
462 that promote or antagonize splicing<sup>55</sup>. Although transcriptome-wide atlases have been  
463 developed to map these sites<sup>37,56</sup>, and derive motif models<sup>57</sup>, it often remains unclear how  
464 genetic variants impact their binding and in turn the eventual splicing output. Our splicing effect  
465 map identifies a cluster of SDVs at the 5' end of the *POU1F1* exon 2, each of which increases  
466 the usage of the normally repressed beta isoform. These results suggest the presence of an  
467 exonic splice silencer (ESS) which may normally suppress utilization of the beta isoform  
468 acceptor. We mined the cisBP-RNA database<sup>38</sup> and identified nine candidate motifs with  
469 strong matches to the U-rich wild-type sequence in this region (c.143-1 to c.167) corresponding  
470 to known splicing factors including ELAVL1 (HuR), RALY, TIA1, and U2AF2 (**Suppl. Fig. 11**).  
471 All six patient variants replaced a U with another base (G in 5 of 6 bases), which may disrupt  
472 these motifs at high information content positions (**Suppl. Fig. 12**). Other variants predicted to  
473 disrupt these motifs tended to be beta-promoting more often than neutral in our map ( $p < 0.01$ ,  
474 Fisher's Exact test). These trends suggest that U-rich ESS serves to inhibit production of  
475 *POU1F1* beta and this inhibition is disrupted by CPHD-associated variants, although  
476 conclusively identifying the specific cognate binding factor will require further study.

477         These results extend the breadth of endocrine disorders caused by disrupted splicing.  
478 For example, in a large cohort with IGHD from Itabaianinha, Brazil, affected individuals are  
479 homozygous for a mutation in the splice donor dinucleotide (c.57 + 1G > A) in the growth  
480 hormone releasing hormone receptor gene (*GHRHR*)<sup>58</sup>. In addition, most mutations that cause  
481 dominant IGHD type II affect splicing of the growth hormone (*GH1*) gene<sup>59</sup>. Mutations in splice  
482 sites or splice enhancer sequences result in skipping exon 3 and production of a dominant-  
483 negative 17.5 kD isoform of growth hormone that lacks amino acids 32-71<sup>60</sup>. The severity of the  
484 disease is variable and correlates inversely with the ratio of 17.5 to 20 kD GH. Finally, severe  
485 short stature associated with Laron syndrome, or GH resistance, can be caused by generation  
486 of a cryptic splice site in the GH receptor gene. Individuals from El Oro and Loja in southern

487 Ecuador are homozygous for a p.180E codon variant (GAA to GAG) that do not change the  
488 amino acid encoded but create a splice acceptor site 24 nt upstream of the normally utilized  
489 site<sup>61</sup>. It is notable that antisense oligonucleotide therapies hold promise for treating diseases  
490 caused by abnormal splicing, including IGHD<sup>62,63</sup>.

491 Splicing disruption accounts for a significant minority of the genetic burden in endocrine  
492 disorders, as in human genetic disease more generally<sup>64,65</sup>. Some estimates from large-scale  
493 screens indicate that 10% of SNV within exons alter splicing, and a third of all disease  
494 associated SNVs impact splicing efficiency<sup>66</sup>. Variants at or near canonical splice sites are  
495 readily recognized as pathogenic<sup>67</sup>, and these can be identified predicted with high accuracy by  
496 algorithms such as SpliceAI. However, for exonic variants, particularly those farther from exon  
497 junctions, splicing defects may be more challenging to identify bioinformatically<sup>68-70</sup>. Efforts to  
498 interpret these variants will need to account for the functional impacts of changing the encoded  
499 protein sequence as well as its splicing. Finally, as our results illustrate, different variants in a  
500 single gene may lead to distinct splicing outcomes with diverse consequences ranging from the  
501 straightforward loss-of-function to dominant negative effects.

502

503

504 **Table 1**

Feature	Family 1		Family 2			Family 3	
	II.1	II.2	I.1	II.2 (Index)	II.4	III.1	II.1
<b>Cases</b>	Male	Female	Male	Female	Male	Female	Female
<b>Age at diagnosis 1st hormone deficiency / Hormone</b>	<5 yr / GH, TSH	<5 yr / GH, TSH	40s / GH	<5 yr / GH	< 5 yr / GH	<5 yr / GH	<5 yr / GH, TSH, PRL
<b>Height at diagnosis of GHD (SDS)</b>	na	na	-3.7	-5	-5.3	na	-4
<b>rhGH treatment (Yes/No)</b>	na	na	No	Yes	Yes	Yes	na
<b>Final height (cm / SDS)</b>	na	na	150 / -3.7	156.5/-0.9	165/-1.5	na	na
<b>Pituitary hormone deficiencies</b>	GH, TSH	GH, TSH	GH	GH	GH	GH	GH, TSH, PRL
<b>Biochemical assessment</b>							
GHSTs	na	na	clonidine, ITT	clonidine, ITT	clonidine	na	na
Maximum GH peak (ng/ml)	na	na	7.6	0.9	0.5	na	na
TSH (U/L)	na	na	0.6	0.7	1.7	na	na
Total T4 (ug/dL)	na	na	6.4	6	5.6-7.3	na	na
Free T4 (ng/dL)	na	na	0.7	0.6-1.1	0.6-0.9	na	na
Prolactin (ng/mL)	na	na	na	3.8 (pTRH 12)	3.2-6.6	na	na
Cortisol (ug/dL)	na	na	na	peak ITT 42	normal	na	na
LH/FSH	na	na	na	early puberty	normal puberty	na	na
<b>Pituitary MRI</b>	Disrupted stalk	Disrupted stalk	Normal	Normal	Normal	na	Normal
<b>Extrapituitary brain MRI</b>	na	na	na	na	na	na	na
<b>Dysmorphic features</b>	none noted	none noted	none noted	large forehead	none noted	na	na
<b>Molecular findings (all in heterozygous state)</b>	c.148T>G p.S50A	c.148T>G p.S50A	c.152T>G p.I51S	c.152T>G p.I51S	c.152T>G p.I51S	c.152T>G p.I51S	c.155T>G p.L52W
<b>In silico predictions</b>							
CADD	22.00			23.4			25.8
SIFT	damaging			damaging			damaging
PP2	benign			benign			probably damaging
Mutation taster	disease causing			disease causing			disease causing

505

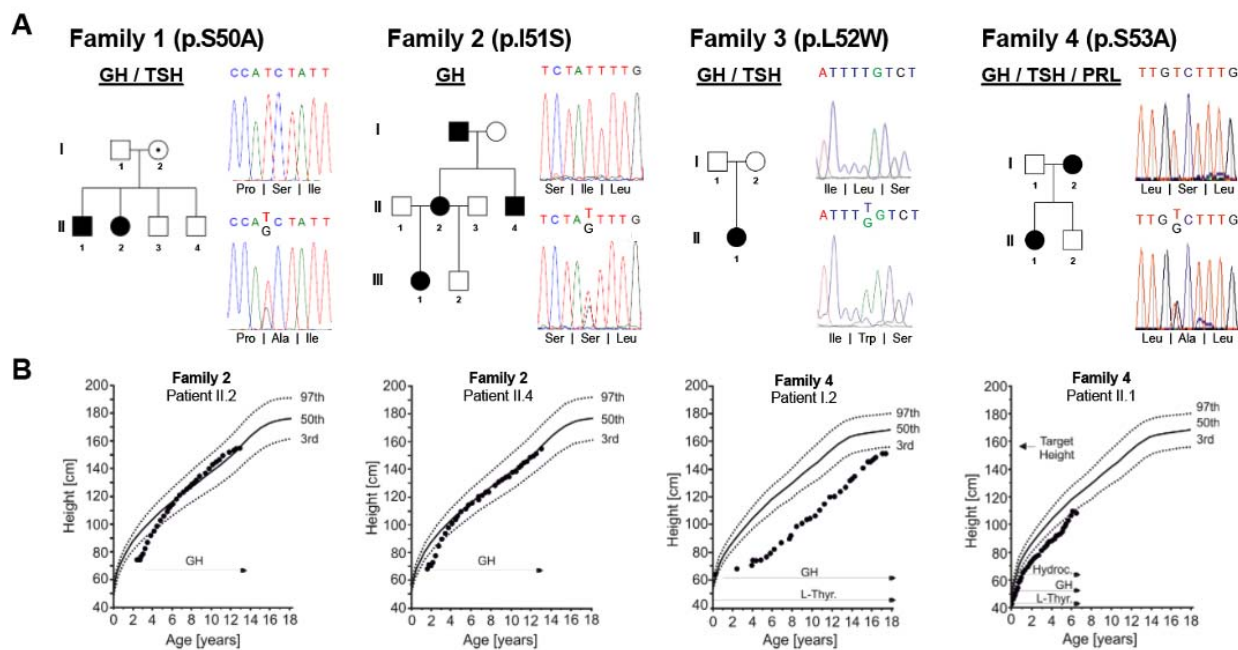
Feature	Family 4		Family 5		Family 6	
	I.2	II.1	I.1	II.1 (index)	II.1 (index)	II.2
<b>Cases</b>	Female	Female	Male	Male	Male	Male
<b>Age at diagnosis 1st hormone deficiency / Hormone</b>	<5 yr / TSH	<5 yr / TSH, GH	preteen / GH	<5 yr / GH	preteen / GH	preteen / GH
<b>Height at diagnosis of GHD (SDS)</b>	-5.42	-3.45	-4.2	-4.15	na	na
<b>rhGH treatment (Yes/No)</b>	Yes	Yes	Yes	Yes	Yes	na
<b>Final height (cm / SDS)</b>	150.9 / -2.67	still growing	147.9 / -3.66	still growing	na	na
<b>Pituitary hormone deficiencies</b>	GH, TSH, PRL, (ACTH)	GH, TSH, PRL, (ACTH)	GH	GH	GH	GH
<b>Biochemical assessment</b>						
GHSTs	glucagon, insulin, clonidine	Basal neonatal	Insulin- Ldopa	Arginine-Clonidine	ITT	ITT
Maximum GH peak (ng/ml)	ND	ND	2.9*	2.7**	3.7 mU/l	3.6 mU/l
TSH (U/L)	na	0.28	na	3.4	normal	normal
Total T4 (ug/dL)	na	na	na	na	na	na
Free T4 (ng/dL)	na	0.4	0.9	1.2	na	na
Prolactin (ng/mL)	2.0 mU/l	9 mU/l	8.3	4.4	na	na
Cortisol (ug/dL)	treated in 20's	36.2 nmol/l	13	13.2	na	na
LH/FSH	Delayed puberty, spontaneous pregnancy	Normal	Spontaneous puberty	0.1/0.75	spontaneous puberty	na
<b>Pituitary MRI</b>	Normal	Normal	na	APH	Normal	na
<b>Extrapituitary brain MRI</b>	1 cm left frontal and parietal lobe abnormality	Normal	Normal	Normal	na	na
<b>Dysmorphic features</b>	Intellectual disability, delayed puberty, strabismus, astigmatism, nystagmus, dysplastic thyroid gland	Macroglossia, bilateral hearing impairment, developmental delay, dysplastic thyroid gland	none noted	Short stature, frontal bossing, high pitched voice	na	na
<b>Molecular findings (all in heterozygous state)</b>	c.157T>G p.S53A	c.157T>G p.S53A	c.150T>G p.S50=	c.150T>G p.S50=	c.153T>A p.I51=	c.153T>A p.I51=
<b>In silico predictions</b>						
CADD	18.65					
SIFT	damaging					
PP2	benign					
Mutation taster	disease causing					

506

na: not available  
 GHSTs: growth hormone stimulation tests, ND: non-detectable, MRI: magnetic resonance imaging, rhGH: recombinant human growth hormone, APH: anterior pituitary hypoplasia  
 ITT: insulin-tolerance test  
 \*Cut-off 4.8 ng/ml; \*\*cut-off 10 ng/ml  
 †diagnosed in late 20s for I.2 and newborn for II.1

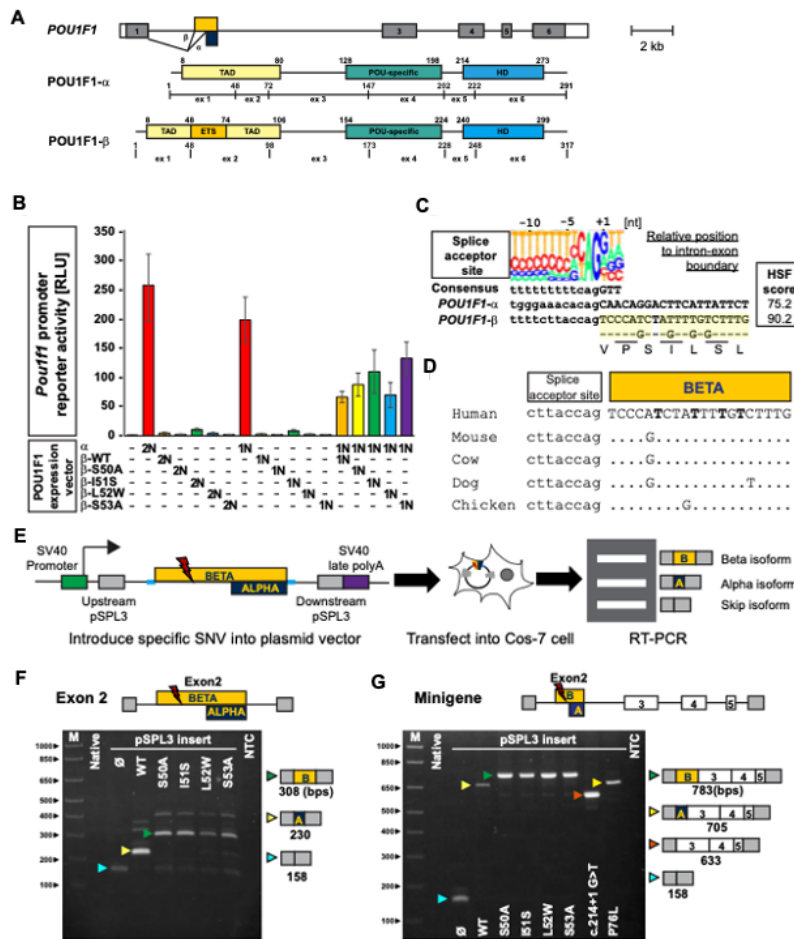
507

508 **Figures**



509  
 510 **Figure 1. Clinical characteristics of the variants of POU1F1 beta coding region**

511 **A.** Pedigrees and the sequence of the patients. Family 1-4 have variants in the POU1F1 beta  
 512 coding region: c.148T>G (p.S50A), c.152T>G (p.I51S), c.155T>G (p.L52W), and c.157T>G  
 513 (p.S53A). **B.** Growth curve of the patients from Family 2 and 4. GH replacement therapy was  
 514 effective in reaching ideal height.

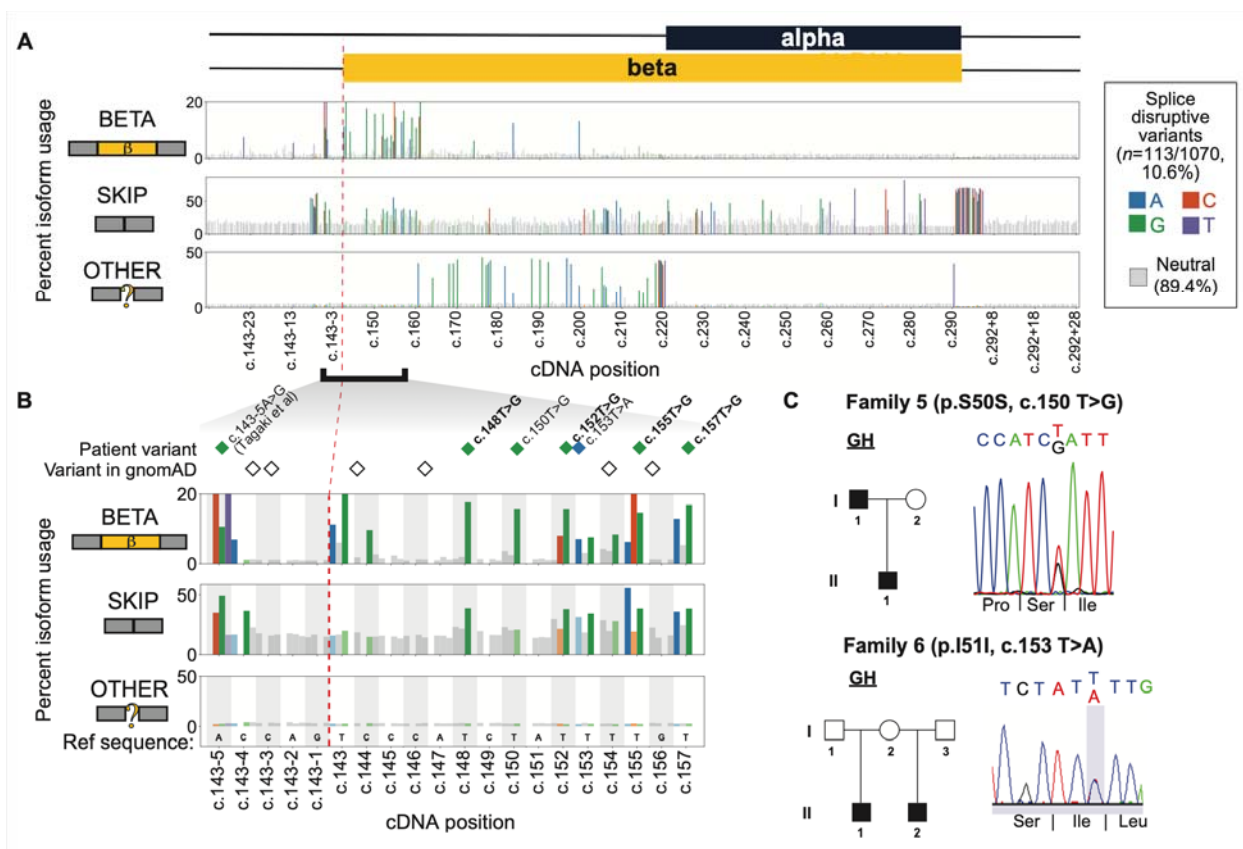


515

516 **Figure 2. Variants in the *POU1F1* beta coding region suppress the function of alpha**  
 517 **isoform and lead to splicing abnormality.**

518 **A.** Schematic of the human *POU1F1* gene and protein isoforms produced by use of alternate  
 519 splice acceptors at exon 2. The *Pou1f1* beta isoform has an insertion of 26 amino acids located  
 520 at amino acid 48 in the transactivation domain. **B.** COS-7 cells were transfected with a *Pou1f1*-  
 521 luciferase reporter gene and expression vectors for *POU1F1* alpha or beta isoforms either singly  
 522 or together in the ratios indicated (2N and 1N). WT *POU1F1* alpha has strong activation at 2N  
 523 and 1N dosages. WT and variants of *POU1F1* beta isoform have no significant activation over  
 524 background. A 50:50 mix of alpha and WT beta isoforms exhibited reduced activation. The  
 525 variant beta isoforms suppress alpha isoform mediated activation to a degree similar to WT. **C.**  
 526 Diagram of the splice acceptor site consensus and the genomic DNA sequence at the boundary  
 527 between intron 1 and splice sites utilized in exon 2 of the *POU1F1* gene<sup>71</sup> **D.** Evolutionary  
 528 conservation of the genomic sequence encoding *POU1F1* beta isoform in mammals and  
 529 chicken. **E.** Exon trapping assay with pSPL3 exon trap vector containing exon 2 of *POU1F1*  
 530 and portions of the flanking introns. **F.** Ethidium bromide-stained gel of exon trap products from  
 531 cells transfected with the indicated plasmid. Arrowheads indicate the expected products for  
 532 exon skipping (Blue), alpha isoform (Yellow), and beta isoform (green). **G.** *POU1F1* minigenes  
 533 spanning from intron 1 to 5, with all of the intervening exons, were engineered with the indicated  
 534 variants and assayed for splicing. WT and p.76L *POU1F1* splice to produce the alpha isoform,  
 535 the G>T change in the splice acceptor causes exon skipping (red arrow) and the other patient  
 536 variants all splice to produce *POU1F1* beta isoform.





537

538 **Figure 3. Splicing effect map in *POU1F1* exon 2 and flanking introns, and identification of**

539 **IGHD families with synonymous changes.**

540 **A.** Percent usage of *POU1F1* exon 2 beta (top panel), skip (middle), and other isoforms

541 (bottom) by variant position as measured by massively parallel minigene assay. Gray bars

542 denote splicing-neutral variants, while colored bars indicate the base pair change of each SDV.

543 Cropped intronic regions are shown in **Suppl. Fig. 4.**

544 **B.** A cluster of SDVs near the beta isoform splice acceptor leads to increased usage of the beta

545 isoform, and in some cases, increased exon skipping. Diamonds colored by the alternate allele

546 indicate patient variants, and empty diamonds indicate variants reported in gnomAD. Missense

547 variants' labels are in bold text.

548 **C.** Families 5 and 6 each had two individuals affected with IGHD and synonymous variants that

549 were splice disruptive. Pedigrees and Sanger sequence confirmation of variants are shown.

550

551 **Acknowledgements**

552 This work was supported by the National Institutes of Health (R01HD097096 to SAC), the Japan  
553 Society for Promotion of Science (HB), Grant 2013/03236-5 from the São Paulo Research  
554 Foundation (FAPESP) (IJPA), a grant from Pfizer (RP), and the Argentinean National Agency of  
555 Scientific and Technical Promotion, PICT 2016-2913 and PICT 2017-0002 (MIPM).

556

557 **Author Contributions**

558 Transfection and exon trapping: Peter Gergics

559 Families 1, 3 and 6 patient collection and candidate gene screening: Frederic Castinetti,

560 Frédérique Albarel, Alexandru Saveanu, Anne Barlier, Thierry Brue

561 Family 2 patient collection and candidate gene screening: Alexander Jorge, Luciani Renata

562 Silveira Carvalho, Marilena Nakaguma, Berenice B Mendonça, Ivo JP Arnhold

563 Family 4 patient collection and WES: Denise Rockstroh-Lippold, Julia Hoppmann, Rami Abou

564 Jamra, Roland Pfaeffle

565 Family 5 patient collection: Debora Braslavsky, Ana Keselman, Ignacio Bergadá

566 Family 1 WES and analysis: Qing Fang, A Bilge Ozel, Qianyi Ma, Jun Z. Li

567 Family 2 WES and analysis: Michael H. Guo, Andrew Dauber

568 Family 5 WES and analysis: Sebastian Vishnopolaska, Julian Martinez Mayer, Marcelo Martí,

569 María Ines Pérez Millán

570 High throughput mutagenesis and analysis: Cathy Smith, Mariam Maksutova, Jacob O. Kitzman

571 Wrote manuscript: Hironori Bando, Cathy Smith, Jacob O. Kitzman, Sally A. Camper

572

573 **Competing Interests statement**

574 The authors declare no competing interest in connection with this manuscript.

575 **Materials & Correspondence**

576 Sally A. Camper

577 Telephone: 734-763-0682

578 Email: [scamper@med.umich.edu](mailto:scamper@med.umich.edu)

579

580 Jacob O. Kitzman

581 Telephone: 734-764-9587

582 email: [kitzmanj@umich.edu](mailto:kitzmanj@umich.edu)

583

584 Mailing address:

585 Camper or Kitzman

586 Dept. Human Genetics

587 University of Michigan Medical School

588 4909 Buhl Building

589 1241 Catherine St.

590 Ann Arbor, MI. 48198-5618

591

592 Fax: 734-763-3784

593

594

595

## 596 References

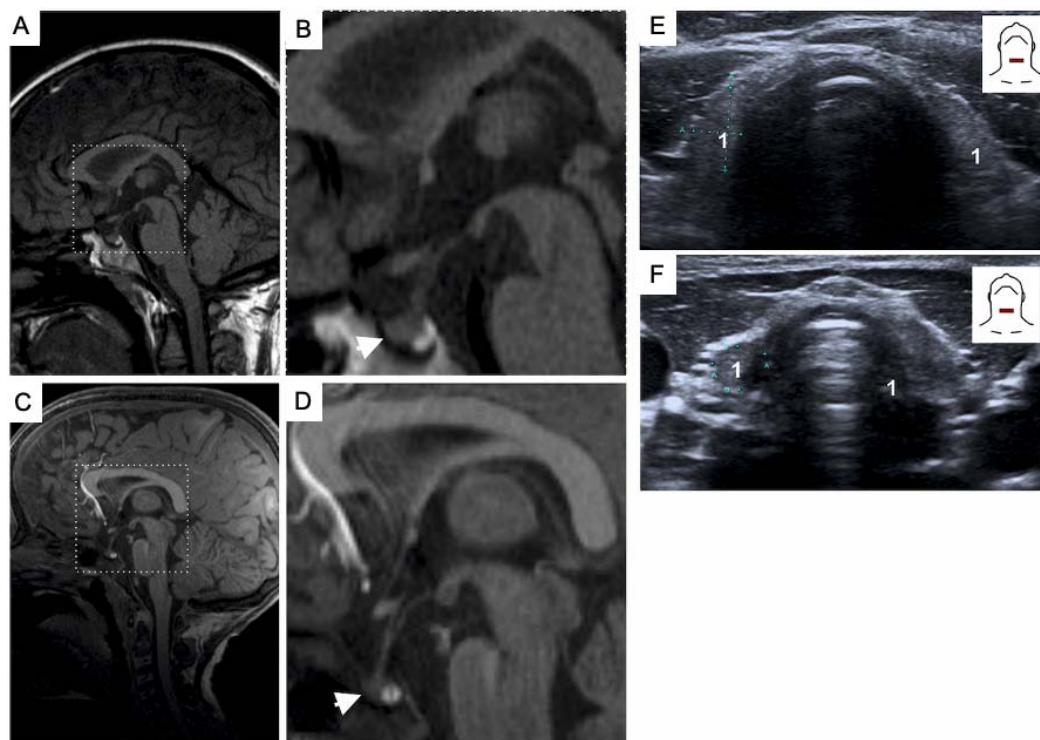
- 597 1 Ingraham, H. A. *et al.* The POU-specific domain of Pit-1 is essential for sequence-specific,  
598 high affinity DNA binding and DNA-dependent Pit-1-Pit-1 interactions. *Cell* **61**, 1021-  
599 1033, doi:10.1016/0092-8674(90)90067-o (1990).
- 600 2 Gordon, D. F. *et al.* Analysis of Pit-1 in regulating mouse TSH beta promoter activity in  
601 thyrotropes. *Mol Cell Endocrinol* **96**, 75-84, doi:10.1016/0303-7207(93)90097-4 (1993).
- 602 3 Davis, S. W., Keisler, J. L., Perez-Millan, M. I., Schade, V. & Camper, S. A. All Hormone-  
603 Producing Cell Types of the Pituitary Intermediate and Anterior Lobes Derive From  
604 Prop1-Expressing Progenitors. *Endocrinology* **157**, 1385-1396, doi:10.1210/en.2015-  
605 1862 (2016).
- 606 4 Li, S. *et al.* Dwarf locus mutants lacking three pituitary cell types result from mutations in the  
607 POU-domain gene pit-1. *Nature* **347**, 528-533, doi:10.1038/347528a0 (1990).
- 608 5 Camper, S. A., Saunders, T. L., Katz, R. W. & Reeves, R. H. The Pit-1 transcription factor  
609 gene is a candidate for the murine Snell dwarf mutation. *Genomics* **8**, 586-590 (1990).
- 610 6 Simmons, D. M. *et al.* Pituitary cell phenotypes involve cell-specific Pit-1 mRNA translation  
611 and synergistic interactions with other classes of transcription factors. *Genes Dev* **4**,  
612 695-711, doi:10.1101/gad.4.5.695 (1990).
- 613 7 Slabaugh, M. B., Lieberman, M. E., Rutledge, J. J. & Gorski, J. Growth hormone and  
614 prolactin synthesis in normal and homozygous Snell and Ames dwarf mice.  
615 *Endocrinology* **109**, 1040-1046, doi:10.1210/endo-109-4-1040 (1981).
- 616 8 Lin, S. C., Li, S., Drolet, D. W. & Rosenfeld, M. G. Pituitary ontogeny of the Snell dwarf  
617 mouse reveals Pit-1-independent and Pit-1-dependent origins of the thyrotrope.  
618 *Development* **120**, 515-522 (1994).
- 619 9 Pfaffle, R. & Klammt, J. Pituitary transcription factors in the aetiology of combined pituitary  
620 hormone deficiency. *Best Pract Res Clin Endocrinol Metab* **25**, 43-60,  
621 doi:10.1016/j.beem.2010.10.014 (2011).
- 622 10 Wallis, M. Evolution of the POU1F1 transcription factor in mammals: Rapid change of the  
623 alternatively-spliced beta-domain. *Gen Comp Endocrinol* **260**, 100-106,  
624 doi:10.1016/j.ygcen.2018.01.005 (2018).
- 625 11 Schanke, J. T., Conwell, C. M., Durning, M., Fisher, J. M. & Golos, T. G. Pit-1/growth  
626 hormone factor 1 splice variant expression in the rhesus monkey pituitary gland and the  
627 rhesus and human placenta. *J Clin Endocrinol Metab* **82**, 800-807,  
628 doi:10.1210/jcem.82.3.3791 (1997).
- 629 12 Konzak, K. E. & Moore, D. D. Functional isoforms of Pit-1 generated by alternative  
630 messenger RNA splicing. *Mol Endocrinol* **6**, 241-247, doi:10.1210/mend.6.2.1569967  
631 (1992).
- 632 13 Haugen, B. R., Wood, W. M., Gordon, D. F. & Ridgway, E. C. A thyrotrope-specific variant of  
633 Pit-1 transactivates the thyrotropin beta promoter. *J Biol Chem* **268**, 20818-20824  
634 (1993).
- 635 14 Jonsen, M. D., Duval, D. L. & Gutierrez-Hartmann, A. The 26-amino acid beta-motif of the  
636 Pit-1beta transcription factor is a dominant and independent repressor domain. *Mol*  
637 *Endocrinol* **23**, 1371-1384, doi:10.1210/me.2008-0137 (2009).
- 638 15 Consortium, G. T. The GTEx Consortium atlas of genetic regulatory effects across human  
639 tissues. *Science* **369**, 1318-1330, doi:10.1126/science.aaz1776 (2020).
- 640 16 Tatsumi, K. *et al.* Cretinism with combined hormone deficiency caused by a mutation in the  
641 PIT1 gene. *Nat Genet* **1**, 56-58, doi:10.1038/ng0492-56 (1992).
- 642 17 Fang, Q. *et al.* Genetics of Combined Pituitary Hormone Deficiency: Roadmap into the  
643 Genome Era. *Endocr Rev* **37**, 636-675, doi:10.1210/er.2016-1101 (2016).
- 644 18 Gergics, P. Pituitary Transcription Factor Mutations Leading to Hypopituitarism. *Exp Suppl*  
645 **111**, 263-298, doi:10.1007/978-3-030-25905-1\_13 (2019).

- 646 19 Birla, S. *et al.* Characterization of a Novel POU1F1 Mutation Identified on Screening 160  
647 Growth Hormone Deficiency Patients. *Horm Metab Res* **51**, 248-255, doi:10.1055/a-  
648 0867-1026 (2019).
- 649 20 Bas, F. *et al.* Precocious or early puberty in patients with combined pituitary hormone  
650 deficiency due to POU1F1 gene mutation: case report and review of possible  
651 mechanisms. *Hormones (Athens)* **17**, 581-588, doi:10.1007/s42000-018-0079-4 (2018).
- 652 21 Blum, W. F. *et al.* Screening a large pediatric cohort with GH deficiency for mutations in  
653 genes regulating pituitary development and GH secretion: Frequencies, phenotypes and  
654 growth outcomes. *EBioMedicine* **36**, 390-400, doi:10.1016/j.ebiom.2018.09.026 (2018).
- 655 22 Bertko, E. *et al.* Combined pituitary hormone deficiency due to gross deletions in the  
656 POU1F1 (PIT-1) and PROP1 genes. *J Hum Genet* **62**, 755-762, doi:10.1038/jhg.2017.34  
657 (2017).
- 658 23 Birla, S. *et al.* Identification of Novel PROP1 and POU1F1 Mutations in Patients with  
659 Combined Pituitary Hormone Deficiency. *Horm Metab Res* **48**, 822-827, doi:10.1055/s-  
660 0042-117112 (2016).
- 661 24 Sobrier, M. L. *et al.* Functional characterization of a human POU1F1 mutation associated  
662 with isolated growth hormone deficiency: a novel etiology for IGHD. *Hum Mol Genet* **25**,  
663 472-483, doi:10.1093/hmg/ddv486 (2016).
- 664 25 Cohen, L. E., Zanger, K., Brue, T., Wondisford, F. E. & Radovick, S. Defective retinoic acid  
665 regulation of the Pit-1 gene enhancer: a novel mechanism of combined pituitary  
666 hormone deficiency. *Mol Endocrinol* **13**, 476-484, doi:10.1210/mend.13.3.0251 (1999).
- 667 26 Skowronska-Krawczyk, D. *et al.* Required enhancer-matrin-3 network interactions for a  
668 homeodomain transcription program. *Nature* **514**, 257-261, doi:10.1038/nature13573  
669 (2014).
- 670 27 Guo, M. H. *et al.* Whole exome sequencing to identify genetic causes of short stature. *Horm*  
671 *Res Paediatr* **82**, 44-52, doi:10.1159/000360857 (2014).
- 672 28 Nisson, P. E., Ally, A. & Watkins, P. C. Protocols for trapping internal and 3'-terminal exons.  
673 *PCR Methods Appl* **4**, S24-39, doi:10.1101/gr.4.1.s24 (1994).
- 674 29 Hiatt, J. B., Patwardhan, R. P., Turner, E. H., Lee, C. & Shendure, J. Parallel, tag-directed  
675 assembly of locally derived short sequence reads. *Nat Methods* **7**, 119-122,  
676 doi:10.1038/nmeth.1416 (2010).
- 677 30 Zorita, E., Cusco, P. & Filion, G. J. Starcode: sequence clustering based on all-pairs search.  
678 *Bioinformatics* **31**, 1913-1919, doi:10.1093/bioinformatics/btv053 (2015).
- 679 31 Garrison, E. & Marth, G. Haplotype-based variant detection from short-read sequencing.  
680 *arXiv.org* **1207.3907v2** (2012).
- 681 32 Wu, T. D. & Watanabe, C. K. GMAP: a genomic mapping and alignment program for mRNA  
682 and EST sequences. *Bioinformatics* **21**, 1859-1875, doi:10.1093/bioinformatics/bti310  
683 (2005).
- 684 33 Rosenberg, A. B., Patwardhan, R. P., Shendure, J. & Seelig, G. Learning the sequence  
685 determinants of alternative splicing from millions of random sequences. *Cell* **163**, 698-  
686 711, doi:10.1016/j.cell.2015.09.054 (2015).
- 687 34 Xiong, H. Y. *et al.* RNA splicing. The human splicing code reveals new insights into the  
688 genetic determinants of disease. *Science* **347**, 1254806, doi:10.1126/science.1254806  
689 (2015).
- 690 35 Jaganathan, K. *et al.* Predicting Splicing from Primary Sequence with Deep Learning. *Cell*  
691 **176**, 535-548 e524, doi:10.1016/j.cell.2018.12.015 (2019).
- 692 36 Cheng, J. *et al.* MMSplice: modular modeling improves the predictions of genetic variant  
693 effects on splicing. *Genome Biol* **20**, 48, doi:10.1186/s13059-019-1653-z (2019).
- 694 37 Ke, S. *et al.* Quantitative evaluation of all hexamers as exonic splicing elements. *Genome*  
695 *Res* **21**, 1360-1374, doi:10.1101/gr.119628.110 (2011).

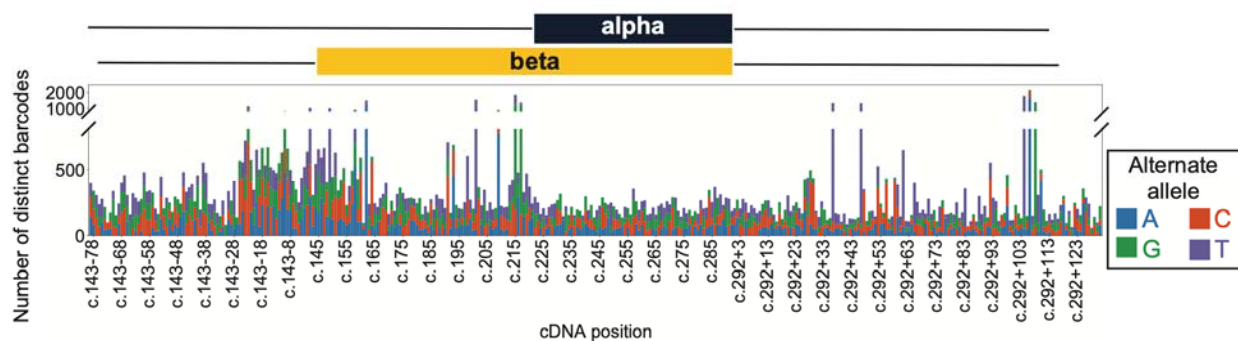
- 696 38 Ray, D. *et al.* A compendium of RNA-binding motifs for decoding gene regulation. *Nature*  
697 **499**, 172-177, doi:10.1038/nature12311 (2013).
- 698 39 Lerario, A. M. *et al.* SELAdb: A database of exonic variants in a Brazilian population referred  
699 to a quaternary medical center in Sao Paulo. *Clinics (Sao Paulo)* **75**, e1913,  
700 doi:10.6061/clinics/2020/e1913 (2020).
- 701 40 Vishnopolska, S. A. *et al.* Genetics and genomic medicine in Argentina. *Mol Genet Genomic*  
702 *Med*, doi:10.1002/mgg3.455 (2018).
- 703 41 Ho, Y., Cooke, N. E. & Liebhaber, S. A. An autoregulatory pathway establishes the definitive  
704 chromatin conformation at the pit-1 locus. *Mol Cell Biol* **35**, 1523-1532,  
705 doi:10.1128/MCB.01283-14 (2015).
- 706 42 Yeo, G. & Burge, C. B. Maximum entropy modeling of short sequence motifs with  
707 applications to RNA splicing signals. *J Comput Biol* **11**, 377-394,  
708 doi:10.1089/1066527041410418 (2004).
- 709 43 Inoue, H. *et al.* Identification of a novel mutation in the exon 2 splice donor site of the  
710 POU1F1/PIT-1 gene in Japanese identical twins with mild combined pituitary hormone  
711 deficiency. *Clin Endocrinol (Oxf)* **76**, 78-87, doi:10.1111/j.1365-2265.2011.04165.x  
712 (2012).
- 713 44 Takagi, M. *et al.* A novel heterozygous intronic mutation in POU1F1 is associated with  
714 combined pituitary hormone deficiency. *Endocr J* **64**, 229-234,  
715 doi:10.1507/endocrj.EJ16-0361 (2017).
- 716 45 Cerbone, M. & Dattani, M. T. Progression from isolated growth hormone deficiency to  
717 combined pituitary hormone deficiency. *Growth Horm IGF Res* **37**, 19-25,  
718 doi:10.1016/j.ghir.2017.10.005 (2017).
- 719 46 Turton, J. P. *et al.* Novel mutations within the POU1F1 gene associated with variable  
720 combined pituitary hormone deficiency. *J Clin Endocrinol Metab* **90**, 4762-4770,  
721 doi:10.1210/jc.2005-0570 (2005).
- 722 47 Solomon, B. D. *et al.* Analysis of genotype-phenotype correlations in human  
723 holoprosencephaly. *Am J Med Genet C Semin Med Genet* **154C**, 133-141,  
724 doi:10.1002/ajmg.c.30240 (2010).
- 725 48 Domene, S. *et al.* Mutations in the human SIX3 gene in holoprosencephaly are loss of  
726 function. *Hum Mol Genet* **17**, 3919-3928, doi:10.1093/hmg/ddn294 (2008).
- 727 49 Gaston-Massuet, C. *et al.* Genetic interaction between the homeobox transcription factors  
728 HESX1 and SIX3 is required for normal pituitary development. *Dev Biol* **324**, 322-333,  
729 doi:10.1016/j.ydbio.2008.08.008 (2008).
- 730 50 Holloway, J. M., Szeto, D. P., Scully, K. M., Glass, C. K. & Rosenfeld, M. G. Pit-1 binding to  
731 specific DNA sites as a monomer or dimer determines gene-specific use of a tyrosine-  
732 dependent synergy domain. *Genes Dev* **9**, 1992-2006, doi:10.1101/gad.9.16.1992  
733 (1995).
- 734 51 Rhodes, S. J. *et al.* A tissue-specific enhancer confers Pit-1-dependent morphogen  
735 inducibility and autoregulation on the pit-1 gene. *Genes Dev* **7**, 913-932 (1993).
- 736 52 Cohen, R. N. *et al.* The role of CBP/p300 interactions and Pit-1 dimerization in the  
737 pathophysiological mechanism of combined pituitary hormone deficiency. *J Clin*  
738 *Endocrinol Metab* **91**, 239-247, doi:10.1210/jc.2005-1211 (2006).
- 739 53 Theill, L. E., Hattori, K., Lazzaro, D., Castrillo, J. L. & Karin, M. Differential splicing of the  
740 GHF1 primary transcript gives rise to two functionally distinct homeodomain proteins.  
741 *EMBO J* **11**, 2261-2269 (1992).
- 742 54 Lim, L. P. & Burge, C. B. A computational analysis of sequence features involved in  
743 recognition of short introns. *Proc Natl Acad Sci U S A* **98**, 11193-11198,  
744 doi:10.1073/pnas.201407298 (2001).

- 745 55 Cartegni, L., Chew, S. L. & Krainer, A. R. Listening to silence and understanding nonsense:  
746 exonic mutations that affect splicing. *Nat Rev Genet* **3**, 285-298, doi:10.1038/nrg775  
747 (2002).
- 748 56 Cheung, R. *et al.* A Multiplexed Assay for Exon Recognition Reveals that an Unappreciated  
749 Fraction of Rare Genetic Variants Cause Large-Effect Splicing Disruptions. *Mol Cell* **73**,  
750 183-194 e188, doi:10.1016/j.molcel.2018.10.037 (2019).
- 751 57 König, J., Zarnack, K., Luscombe, N. M. & Ule, J. Protein-RNA interactions: new genomic  
752 technologies and perspectives. *Nat Rev Genet* **13**, 77-83, doi:10.1038/nrg3141 (2012).
- 753 58 Salvatori, R. *et al.* Familial dwarfism due to a novel mutation of the growth hormone-  
754 releasing hormone receptor gene. *J Clin Endocrinol Metab* **84**, 917-923,  
755 doi:10.1210/jcem.84.3.5599 (1999).
- 756 59 Alatzoglou, K. S. & Dattani, M. T. Phenotype-genotype correlations in congenital isolated  
757 growth hormone deficiency (IGHD). *Indian J Pediatr* **79**, 99-106, doi:10.1007/s12098-  
758 011-0614-7 (2012).
- 759 60 Shariat, N., Holladay, C. D., Cleary, R. K., Phillips, J. A., 3rd & Patton, J. G. Isolated growth  
760 hormone deficiency type II caused by a point mutation that alters both splice site  
761 strength and splicing enhancer function. *Clin Genet* **74**, 539-545, doi:10.1111/j.1399-  
762 0004.2008.01042.x (2008).
- 763 61 Berg, M. A., Guevara-Aguirre, J., Rosenbloom, A. L., Rosenfeld, R. G. & Francke, U.  
764 Mutation creating a new splice site in the growth hormone receptor genes of 37  
765 Ecuadorean patients with Laron syndrome. *Hum Mutat* **1**, 24-32,  
766 doi:10.1002/humu.1380010105 (1992).
- 767 62 Miletta, M. C., Lochmatter, D., Pektovic, V. & Mullis, P. E. Isolated growth hormone  
768 deficiency type 2: from gene to therapy. *Endocr Dev* **23**, 109-120,  
769 doi:10.1159/000341766 (2012).
- 770 63 Kuijper, E. C., Bergsma, A. J., Pijnappel, W. & Aartsma-Rus, A. Opportunities and  
771 challenges for antisense oligonucleotide therapies. *J Inherit Metab Dis*,  
772 doi:10.1002/jimd.12251 (2020).
- 773 64 Cummings, B. B. *et al.* Improving genetic diagnosis in Mendelian disease with transcriptome  
774 sequencing. *Sci Transl Med* **9**, doi:10.1126/scitranslmed.aal5209 (2017).
- 775 65 Scotti, M. M. & Swanson, M. S. RNA mis-splicing in disease. *Nat Rev Genet* **17**, 19-32,  
776 doi:10.1038/nrg.2015.3 (2016).
- 777 66 Soemedi, R. *et al.* Pathogenic variants that alter protein code often disrupt splicing. *Nat*  
778 *Genet* **49**, 848-855, doi:10.1038/ng.3837 (2017).
- 779 67 Lord, J. *et al.* Pathogenicity and selective constraint on variation near splice sites. *Genome*  
780 *Res* **29**, 159-170, doi:10.1101/gr.238444.118 (2019).
- 781 68 Khan, M. *et al.* Resolving the dark matter of ABCA4 for 1054 Stargardt disease probands  
782 through integrated genomics and transcriptomics. *Genet Med* **22**, 1235-1246,  
783 doi:10.1038/s41436-020-0787-4 (2020).
- 784 69 Chen, J. M. *et al.* The Experimentally Obtained Functional Impact Assessments of 5' Splice  
785 Site GT'GC Variants Differ Markedly from Those Predicted. *Curr Genomics* **21**, 56-66,  
786 doi:10.2174/1389202921666200210141701 (2020).
- 787 70 Dionnet, E. *et al.* Splicing impact of deep exonic missense variants in CAPN3 explored  
788 systematically by minigene functional assay. *Hum Mutat*, doi:10.1002/humu.24083  
789 (2020).
- 790 71 Ma, S. L. *et al.* Whole Exome Sequencing Reveals Novel PHEX Splice Site Mutations in  
791 Patients with Hypophosphatemic Rickets. *PLoS One* **10**, e0130729,  
792 doi:10.1371/journal.pone.0130729 (2015).
- 793

795 **Supplementary Figures and Legends**

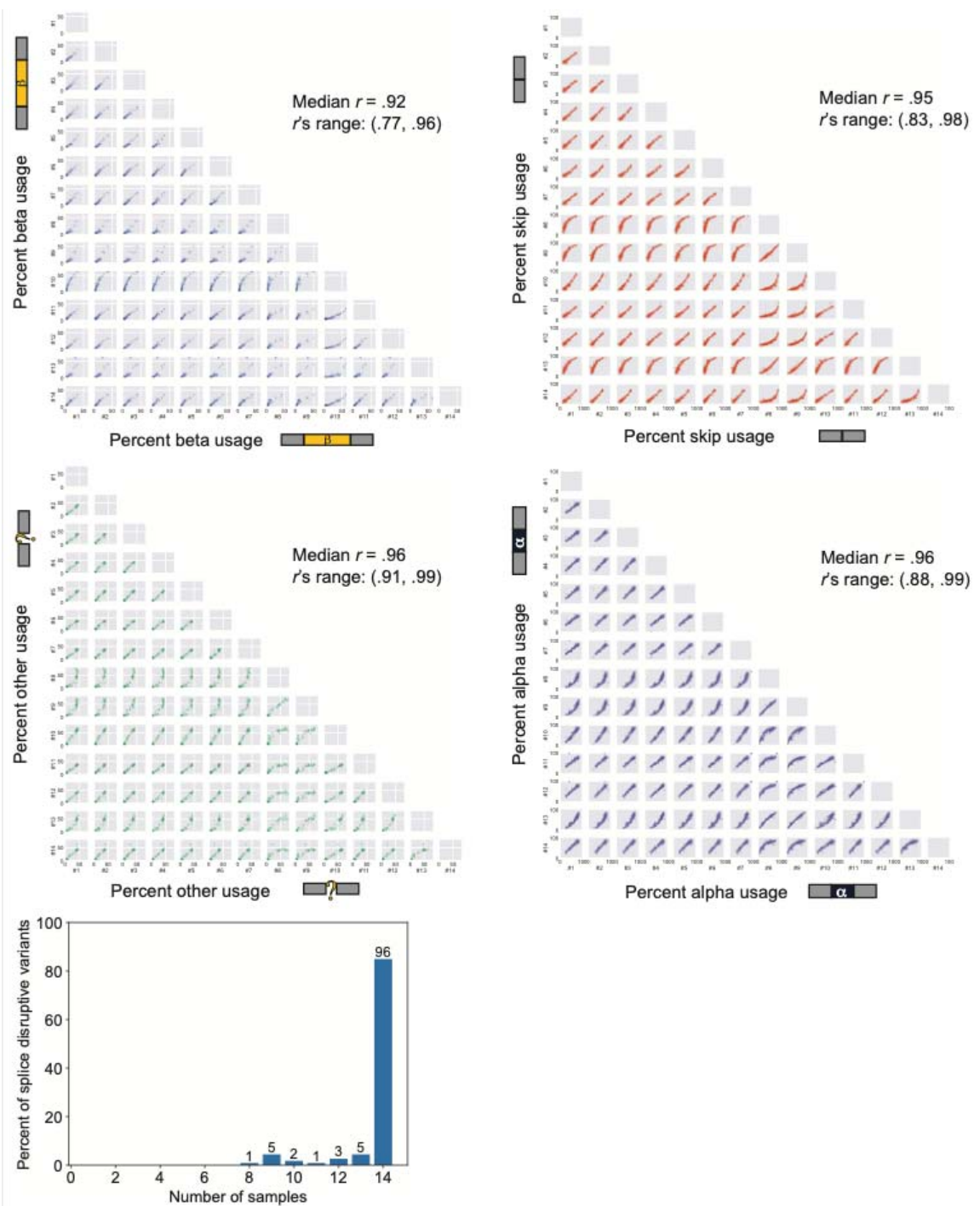


796  
797 **Supplementary Figure 1. Clinical information for Family 4.**  
798 Brain MRI of patient I.1 (A, B; as a teenager) and II.1 (C-D; as a pre-teen). Thyroid ultrasound  
799 of patient I.2 (E) and II.1 (F). <sup>1</sup>Arteria carotis communis.  
800  
801



802  
803 **Supplementary Figure 2. Completeness and uniformity of saturation mutagenesis.**  
804 Stacked barplot showing, for each *POU1F1* variant by position (x-axis) and allele (color), the  
805 number of distinct barcodes detected in RNA-seq data (median across replicates).  
806



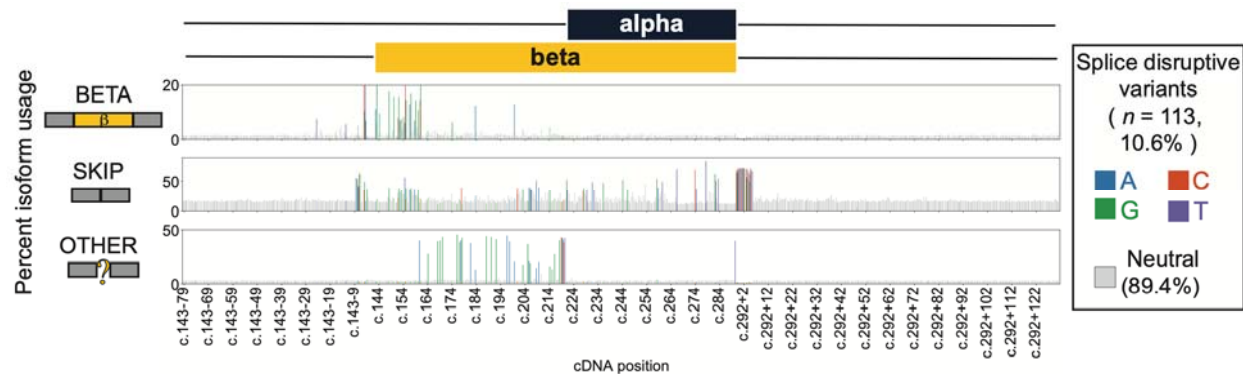


807  
808 **Supplementary Figure 3. Inter-replicate correlation.**

809 A. Pairwise scatterplots of percent isoform use for beta, skip, other, and alpha isoforms among  
810 the fourteen biological replicates. The median and range of Pearson's correlation values across  
811 samples are shown for each isoform.

812 B. Histogram plotting the number of replicate samples in which variants met the splice disruptive  
813 variant (SDV) criteria; all SDVs met threshold in  $\geq 8$  replicates, with 96/113 found in all 14  
814 replicates.

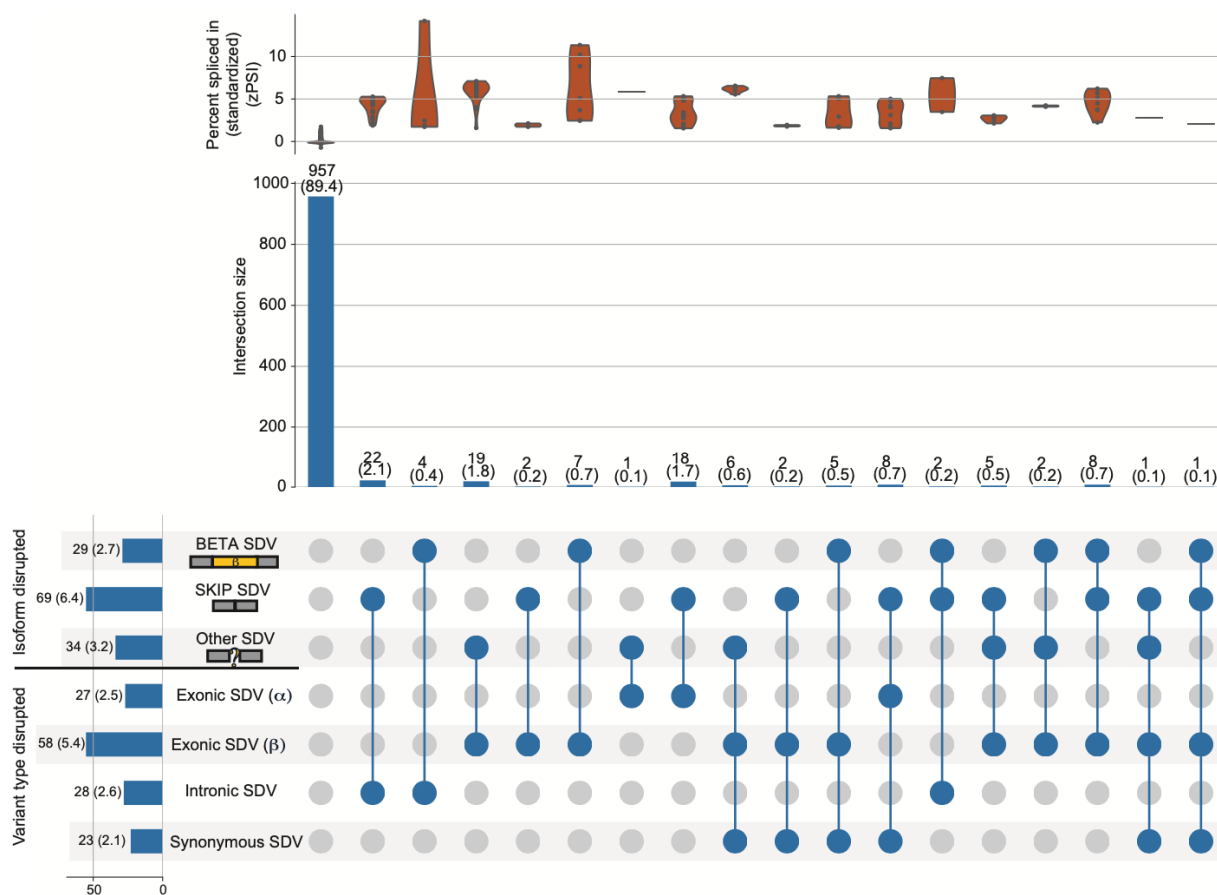
815



816  
817  
818  
819  
820  
821

### Supplementary Figure 4. Uncropped *POU1F1* splicing effect map

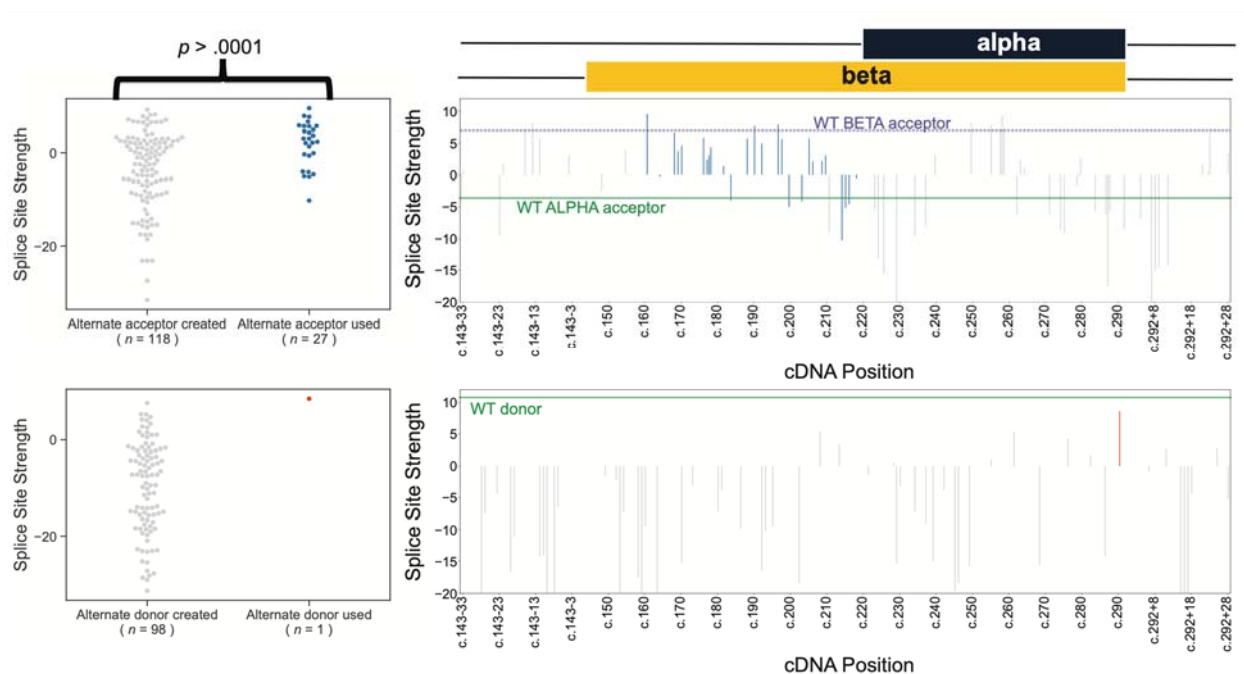
Uncropped version of **Figure 3A**, including cropped intronic regions lacking any SDVs. Isoform usage and variants are plotted as in **Figure 3A**.



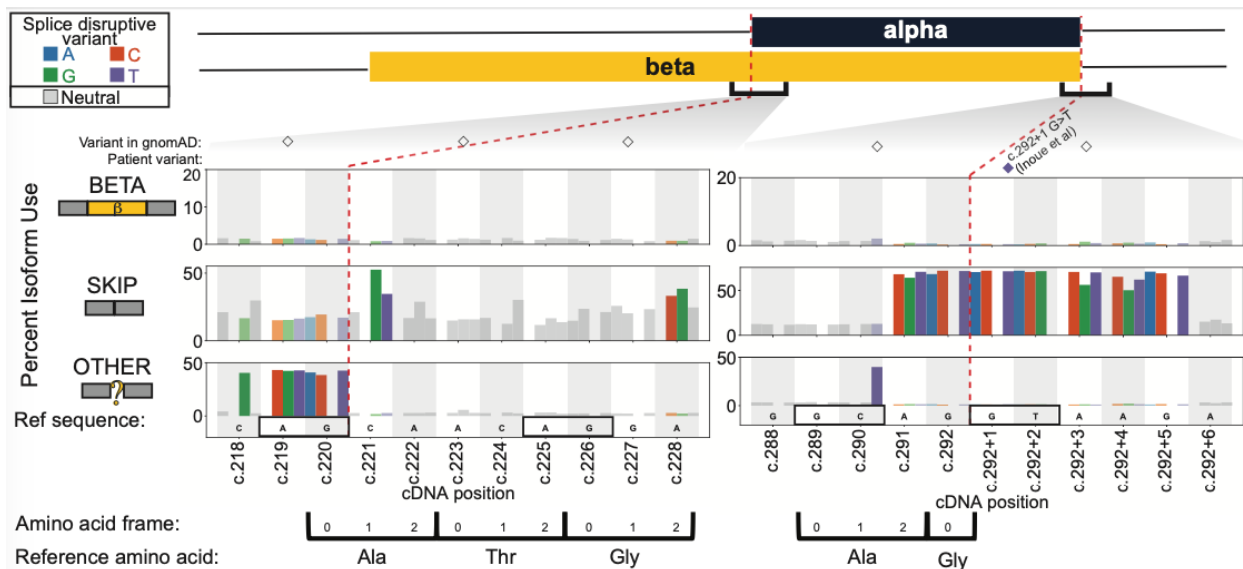
822  
823  
824  
825  
826  
827  
828  
829  
830

### Supplementary Figure 5. Splice disruptive variants by isoform and variant type.

Distributions of isoform usage z-scores for each subset of subsets are shown as violin plots. Count within each intersection (and % of total) are shown above vertical bars. Count within each subset prior to intersection (and % of total) are shown along horizontal bars. UpSet plot showing splice disruptive variants (SDVs), categorized by isoform (beta, skip, and other) and variant type (exonic, intronic, and synonymous). Filled circles denote membership in multiple categories (e.g., second column from the left indicates there are 22 intronic SDVs causing increased skipping).



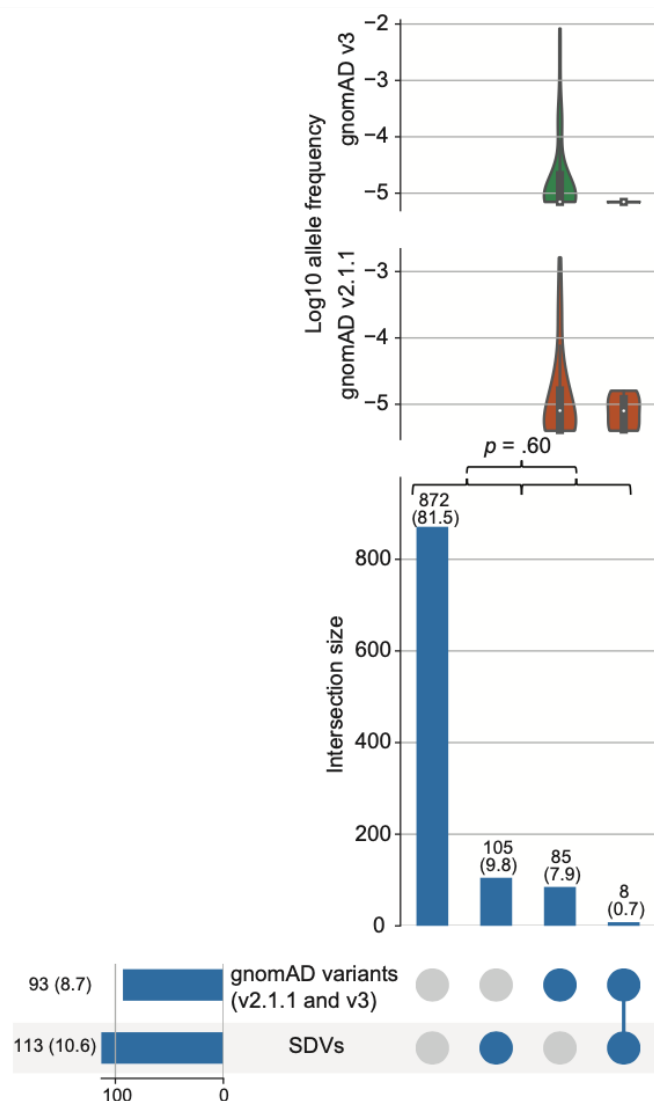
831  
832 **Supplementary Figure 6. Splice site strength for novel alternate donors and acceptors.**  
833 Splice site strength as predicted by MaxEntScan<sup>42</sup> for novel alternate splice acceptor and donor  
834 sites. *P*-value corresponds to a *t*-test comparing the splice site strength for motifs at seldom  
835 used novel acceptor sites (*n* = 118) vs. motifs that promote use of another isoform at novel  
836 acceptor sites (*n* = 27). Dashed line (purple) represents the splice site strength of the native  
837 beta acceptor site. Solid lines (green) indicate the splice site strength of the native alpha  
838 acceptor site and native donor site respectively. Splice site strength is truncated at -20 in the  
839 positional plots, but minimum is as low as -31.6 for novel acceptors and donors within this exon.  
840



841  
 842  
 843  
 844  
 845  
 846  
 847  
 848  
 849  
 850

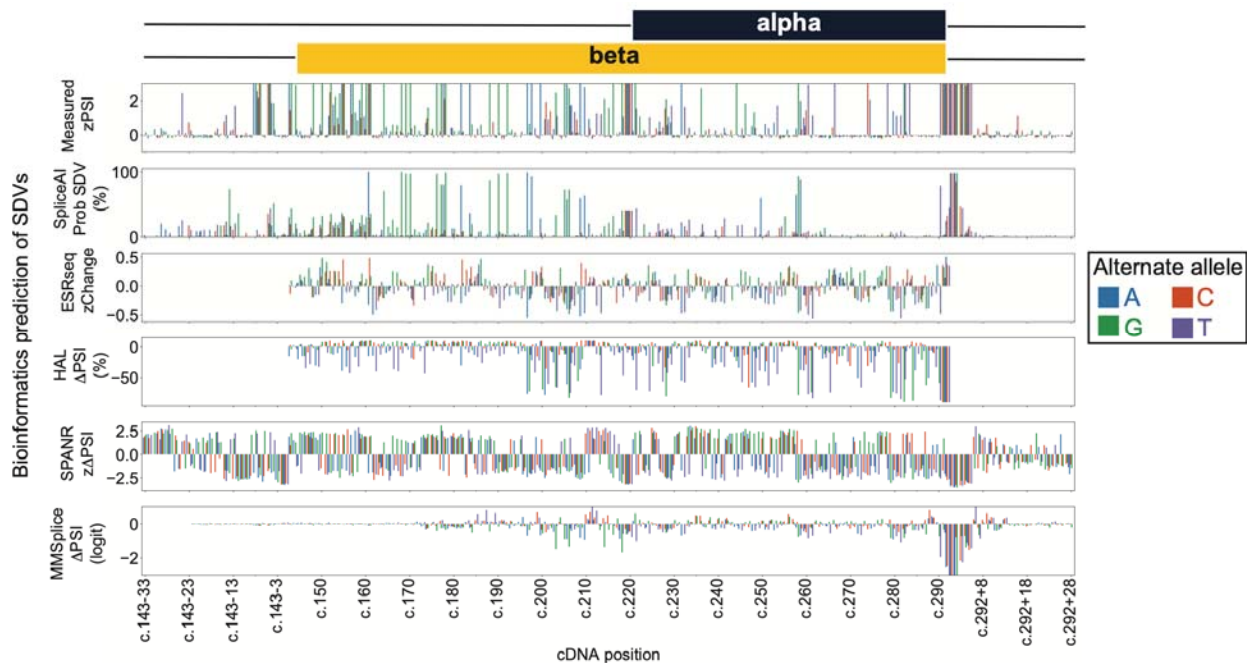
**Supplementary Figure 7. Alternate splice sites and frameshift mutations.**

Detailed view of splicing effect measurements, plotted as in **Figure 3B**, focusing on native alpha acceptor site (left) and native donor site (right). Colored and unfilled diamonds indicate patient variants (colored by alternate allele) and gnomAD variants, respectively. Canonical and cryptic splice sites are boxed, red dashed lines demarcate canonical exon boundaries, and coding frame and corresponding amino acids are indicated below.

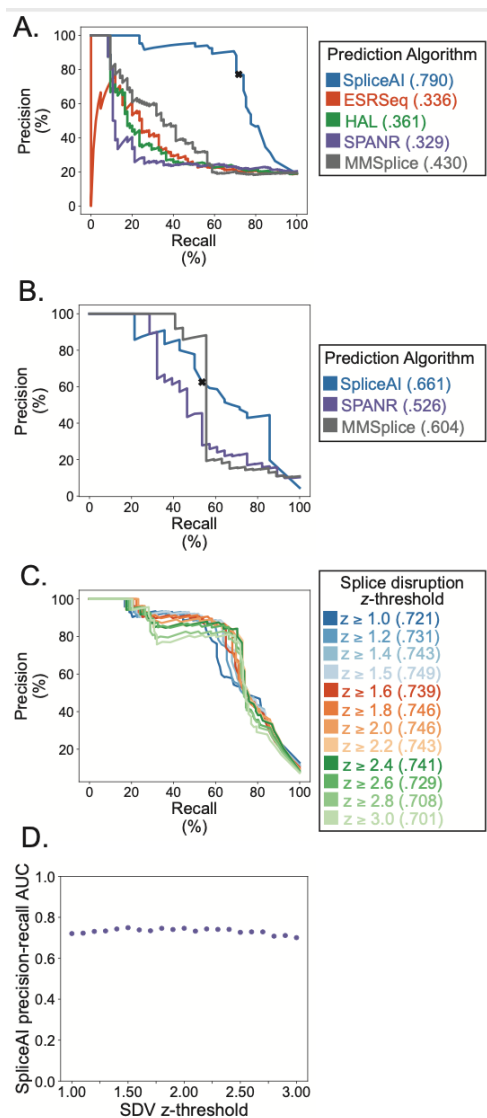


851  
 852 **Supplementary Figure 8. Splice disruptive variants (SDVs) in gnomAD.**  
 853 Violin plots of the log10 allele frequency for each variant found in gnomAD v2.1.1 (orange) and  
 854 v3 (green) within each subset are shown. Count within each intersection (and % of total) are  
 855 shown above vertical bars. Count within each subset prior to intersection (and % of total) are  
 856 shown along horizontal bars.  $P$ -value corresponds to a Fisher's exact test comparing the  
 857 proportion of splice disruptive variants between gnomAD variants and variants absent from  
 858 gnomAD. UpSet plots showing intersection of tested variants with or without splice disruptive  
 859 effects, and gnomAD variants.

860  
 861  
 862



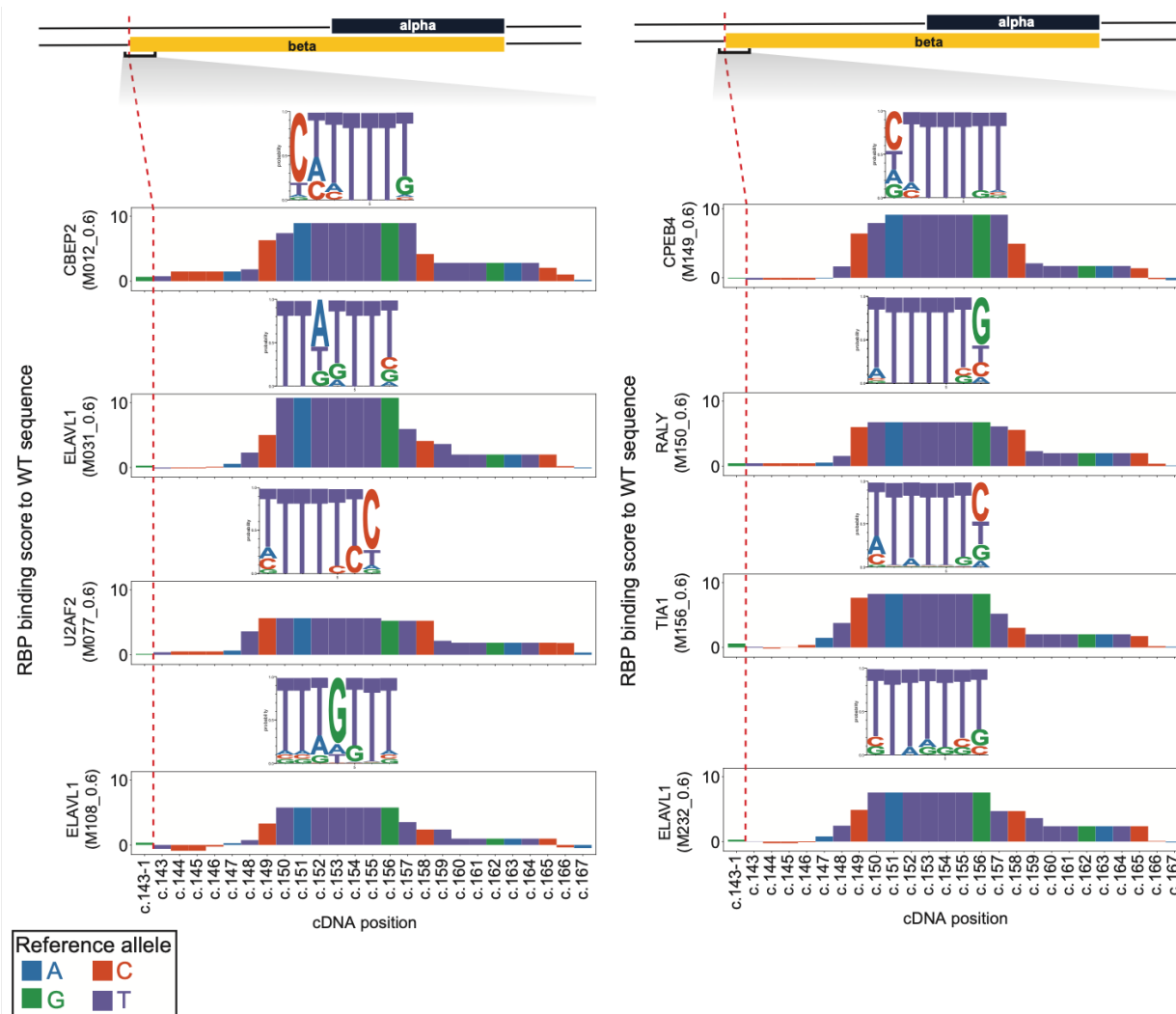
863  
864 **Supplementary Figure 9. *In silico* predictions of splice disrupting variants (SDV).**  
865 Barplots showing for each variant (color) at every position (x-axis) the splicing effect  
866 measurements (top y-axis) and splice disruption as predicted by SpliceAI, ESRSeq, HAL,  
867 SPANR, and MMSplice (from second from the top to bottom y-axes).  
868



869

870 **Supplementary Figure 10. Evaluation of *in silico* splicing effect predictions.**

871 **A.** Precision-recall curve showing the precision (y-axis) and recall (x-axis) of SpliceAI (blue)<sup>35</sup>,  
 872 ESRseq (orange)<sup>37</sup>, HAL (green)<sup>33</sup>, SPANR (purple)<sup>34</sup>, and MMSplice (gray)<sup>36</sup> to predict splice  
 873 disruptive variants (SDVs) in exonic regions. The 'x' is the at the minimum threshold where  
 874 SpliceAI predicts all of the patient variants as disruptive (SpliceAI deltaMax score ≥ 0.19). Area  
 875 under the curve (prAUC) is shown within the legend **B**. Same as in A but for intronic variants.  
 876 Since HAL and ESRseq values do not apply in noncoding regions so they are omitted from this  
 877 plot. **C.** Precision-recall curve of the precision (y-axis) and recall (x-axis) for SpliceAI prediction  
 878 of measured splice disruption across varying zPSI thresholds (range: 1 - 3) to call variants as  
 879 disruptive. prAUC for each threshold is shown within the legend. **D.** Scatterplot of SpliceAI  
 880 prAUC (y-axis) at varying splice disruption z-thresholds (x-axis).  
 881



882

883 **Supplementary Figure 11. RNA binding protein consensus binding motifs relative to**  
 884 **splice acceptor for *POU1F1* beta.**

885 Barplots displaying match scores (y-axis) for selected motifs defined by RNACompete<sup>38</sup> scored  
 886 against the wild-type *POU1F1* sequence beta region (positions c.143-1 to to c.167).





887

888 **Supplementary Figure 12. Changes in RNA binding protein motifs scores due to the**  
 889 **SNVs in *POU1F1* beta.**

890 Barplots show the change in maximal RNAcompete kmer score (y-axis) by variant and  
 891 position (x-axis), relative to the same motif scored against the wild-type *POU1F1*  
 892 sequence. Black stars indicate SDVs that promote beta isoform splicing.  
 893

894

895 **Supplemental Table 1. Summary functional effects of all variants tested near *POU1F1***  
896 **exon 2.**