



HAL
open science

Biallelic Mutations in P4HTM Cause Syndromic Obesity

Sadia Saeed, Lijiao Ning, Alaa Badreddine, Muhammad Usman Mirza,
Mathilde Boissel, Roohia Khanam, Jaida Manzoor, Qasim Janjua, Waqas
Khan, Bénédicte Toussaint, et al.

► **To cite this version:**

Sadia Saeed, Lijiao Ning, Alaa Badreddine, Muhammad Usman Mirza, Mathilde Boissel, et al..
Biallelic Mutations in P4HTM Cause Syndromic Obesity. *Diabetes*, 2023, 72 (9), pp.1228-1234.
10.2337/db22-1017 . inserm-04303603

HAL Id: inserm-04303603

<https://inserm.hal.science/inserm-04303603v1>

Submitted on 23 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 **Biallelic mutations in *P4HTM* cause syndromic obesity**

2

3 Sadia Saeed^{1,2,3§}, Lijiao Ning^{2,3}, Alaa Badreddine^{2,3}, Muhammad Usman Mirza⁴, Mathilde
4 Boissel^{1,2,3}, Roohia Khanam⁵, Jaida Manzoor⁶, Qasim M Janjua⁷, Waqas I. Khan⁸, Bénédicte
5 Toussaint^{2,3}, Emmanuel Vaillant^{2,3}, Souhila Amanzougarene^{2,3}, Mehdi Derhourhi^{2,3}, John F
6 Trant⁴, Anna-Maria Siegert⁹, Brian Y. H. Lam⁹, Giles S.H. Yeo⁹, Layachi Chabraoui¹⁰,
7 Asmae Touzani¹¹, Abhishek Kulkarni¹², I. Sadaf Farooqi⁹, Amélie Bonnefond^{1,2,3*},
8 Muhammad Arslan^{5*} and Philippe Froguel^{1,2,3*§}

9

10 ¹Department of Metabolism, Digestion and Reproduction, Imperial College London, London,
11 UK; ²Inserm UMR 1283, CNRS UMR 8199, EGID, Institut Pasteur de Lille, Lille, France,
12 ³University of Lille, Lille University Hospital, Lille, France; ⁴ Department of Chemistry and
13 Biochemistry, University of Windsor, Windsor ON, Canada, ⁵School of Life Sciences, Forman
14 Christian College, Lahore, Pakistan, ⁶Department of Paediatric Endocrinology, Children's
15 Hospital, Lahore, Pakistan, ⁷Department of Physiology and Biophysics, National University
16 of Science and Technology, Sohar, Oman, ⁸The Children Hospital and the Institute of Child
17 Health, Multan, Pakistan, ⁹Medical Research Council Metabolic Diseases Unit, Wellcome-
18 MRC Institute of Metabolic Science - Metabolic Research Laboratories, University of
19 Cambridge, Cambridge, United Kingdom, ¹⁰Laboratory of Biochemistry and Molecular
20 Biology -Faculty of Medicine and Pharmacy -University V Mohamed Rabat, Morocco,
21 ¹¹Children 's Hospital of Rabat and Laboratory of Biochemistry and Molecular Biology -
22 Faculty of Medicine and Pharmacy -University V Mohamed Rabat -Morocco, ¹²Department
23 of Paediatric Endocrinology, Sir HN Reliance Foundation & SRCC Children's Hospital,
24 Mumbai, India

25 *These authors contributed equally to the paper

26

27 §Corresponding authors: Sadia Saeed, s.saeed08@imperial.ac.uk; Philippe Froguel,
28 p.froguel@imperial.ac.uk

29

30

31

32

33

34

35

36 **ABSTRACT**

37 We previously demonstrated that 50% of children with obesity from consanguineous families
38 from Pakistan carried pathogenic variants in known monogenic obesity genes. Here, we have
39 discovered a novel monogenetic recessive form of severe childhood obesity, using an in-
40 house computational staged approach. This included analysis of whole-exome sequencing
41 data of 366 children with severe obesity, 1,000 individuals of the Pakistani PROMIS study,
42 and 200K participants of the UK Biobank, to prioritize genes harbouring rare homozygous
43 variants with putative effect on human obesity. We identified five rare or novel homozygous
44 missense mutations predicted deleterious in five consanguineous families in *P4HTM* encoding
45 Prolyl 4-Hydroxylase Transmembrane (P4H-TM). We further found two additional
46 homozygous missense mutations in children with severe obesity of Indian and Moroccan
47 origin. Molecular dynamics simulation suggested that these mutations destabilized the active
48 conformation of the substrate binding domain. Most carriers also presented with hypotonia,
49 cognitive impairment and/or developmental delay. Three of the five probands died of
50 pneumonia during the ~2 years of the follow up. *P4HTM* deficiency is a novel form of
51 syndromic obesity affecting 1.5% of our children with obesity associated with high mortality.
52 P4H-TM is a hypoxia inducible factor that is necessary for survival and adaptation under
53 oxygen deprivation but the role of this pathway in energy homeostasis and obesity
54 pathophysiology remains to be elucidated.

55 INTRODUCTION

56 Previously, our genetic studies on children with severe obesity from consanguineous families
57 of Pakistan (severely obese Pakistani population/SOPP cohort), enabled a successful genetic
58 diagnosis in ~50% of obese participants (1-5). This predominantly includes identification of
59 pathogenic mutations in genes that play a key role in the hypothalamic leptin-melanocortin
60 pathway (e.g. *LEP*, *LEPR* and *MC4R*), regulating appetite and energy homeostasis. Early-
61 onset, severe obesity accompanied with severe hyperphagia are the most common phenotypes
62 caused by these mutations.

63 There are also several well-known pleiotropic monogenic obesity syndromes where obesity is
64 one of a number of clinical anomalies often including intellectual disability, structural
65 abnormalities, dysmorphic features, vision and hearing impairments, maladaptive behaviour,
66 and organ- and cell-specific anomalies, as observed in Bardet-Biedl, Alstrom, Carpenter and
67 Cohen syndromes (6), amongst others. Our endeavour of unravelling genetic causality in
68 SOPP also led to the identification of probands with these obesity syndromes (7) and also
69 uncovered that bi-allelic mutations in *ADCY3* are a novel genetic cause of syndromic obesity
70 (8).

71 Aiming at a further investigation of the genetic basis of severe obesity and to identify
72 additional causes of obesity in undiagnosed cases from SOPP, we developed and employed an
73 in-house systematic analysis of the sequencing data using MiST method (9) in combination
74 with rigorous and step-wise data filtration. Additionally, we carried out molecular dynamics
75 (MD) simulations to determine consequences of the mutations on the encoded proteins'
76 stability and conformation. As a result, we have identified several patients with severe obesity
77 carrying homozygous, deleterious, missense mutations in *P4HTM* encoding the enzyme prolyl
78 4-hydroxylase transmembrane (P4H-TM), an atypical member of the hypoxia inducible
79 factors prolyl 4-hydroxylases (HIF-P4Hs), that are necessary for adaptation under reduced

80 oxygen supply. This is the only member of HIF-P4Hs that predominantly expresses in brain
81 but mechanism by which P4H-TM causes obesity is still unknown, opening new avenues of
82 research in the physiology of energy balance.

83

84 **MATERIALS AND METHODS**

85 *Genetic and Statistical analysis*

86 Initially, all probands with severe obesity from SOPP ($n=456$) were systematically screened
87 for mutations in *LEP* and *MC4R* through Sanger sequencing. The probands found negative for
88 pathogenic or likely pathogenic mutations in these two genes, were further analysed either
89 through conventional whole-exome sequencing or augmented whole exome sequencing
90 (CoDE-seq) described in detail elsewhere (10; 11).

91 Analysis of exome data was carried out stage-wise: **Stage I** All the genes known to be linked
92 with monogenic obesity including syndromic obesity, were searched for pathogenic or likely
93 pathogenic mutations by following the rules and guidelines of the American College of
94 Medical Genetics and Genomics (**Supplementary Table 1**) (12). **Stage II** Gene-centric
95 analysis using the MiST method was performed (9) on exome data from severely obese
96 subjects with yet undiagnosed genetic causality, in SOPP ($n=366$) and 1,003 normal-weight
97 subjects from PROMIS (Pakistan Risk of Myocardial Infarction Study) cohort. Participants in
98 both groups belong to the same geographical region allowing for a better comparison, to
99 identify genes harbouring a significant burden of rare (*i.e.* with a minor allele frequency
100 below 1%), homozygous and potentially deleterious (according to SIFT and PolyPhen)
101 variants among the cases with obesity. In **Stage III**, the genes identified with a significant p-
102 value at **Stage II** (*i.e.* $P_{\pi} < 0.05$) were then investigated against 200K exome data from UK
103 Biobank (Application #67575), where we assessed the association between null variants (*i.e.*

104 nonsense, frameshift, canonical ± 1 or 2 splice sites, initiation codon) per gene and body mass
105 index (BMI) using MiST. Our statistical analyses were adjusted for age, gender and the first 5
106 genetic principal components, to take into account any potential ethnicity confounds in the
107 cohort. The details of the MiST analysis have been provided elsewhere (13). In the **Stage IV**,
108 only the genes with mutations in 3 or more unrelated families were considered as to be likely
109 involved in human obesity. Finally, in **Stage V**, literature survey and further clinical
110 investigation of the affected probands was carried out (**Figure 1**).

111 **Biochemical analysis**

112 Metabolic markers including leptin, insulin and cortisol were determined using commercially
113 available ELISA kits (Monobind, Lake Forest).

114 **Molecular modelling**

115 The molecular modelling was carried out for four of the seven mutations (p.Glu155Lys,
116 p.Val297Met, p.Val316Ile and p.Gly433Ala), where crystal structural information was
117 available. Here the dynamic and structural consequences of the P4HTM mutations on
118 substrate binding were assessed through homology modelling followed by molecular
119 dynamics (MD) simulations. The detailed procedure is given in **Supplementary**
120 **Information**.

121 ***P4HTM* expression in murine and human hypothalamus**

122 Recently we generated an integrated single-cell atlas, namely HypoMap, of the murine
123 hypothalamus. It consists of 384,924 cells collected from 18 different single-cell studies,
124 covering various regions of the hypothalamus(14). The publicly available Seurat data object
125 was used to assess the expression of *P4htm* in cells captured within the atlas.

126 To examine the expression for *P4HTM* in human hypothalamus, we extracted the
127 hypothalamic cells from a single-nucleus sequencing dataset from adult human brains (15),
128 available via the Neuroscience Multi-omics Archive (NeMO, RRID:SCR_016152). Briefly,
129 the count data from 134,471 nuclei labelled ‘hypothalamus’ from ROIGroupCoarse was
130 extracted from the original loom file, and the count table was then imported into Seurat 4.1.1
131 for log-normalisation; variable feature selection (3000 genes); data reduction via principal
132 component analysis (PCA, 90 PCs); and uniform manifold approximation projection (UMAP
133 using PCs 1:50) was used for data visualisation.

134

135 **RESULTS**

136 The stepwise analysis and rigorous filtration criteria (**Figure 1**) resulted in the identification
137 of only one gene putatively involved in monogenic obesity: *P4HTM* (encoding Prolyl 4-
138 Hydroxylase, Transmembrane). Through gene-centric analysis the burden of rare,
139 homozygous, and potentially deleterious variants in *P4HTM* was significantly associated with
140 obesity risk ($P_{\pi} = 0.013$; $\pi_{\text{hat}} = 14$; 95% CI_{2.5} = -26; $P_{\tau} = 1.0$). In the next step (*i.e.* Stage
141 III) using the 200K exome data from UK Biobank, we assessed the association between null
142 variants (**Supplementary Table 2**) in *P4HTM* and adiposity. We found that these null
143 variants were significantly associated with higher BMI levels ($P_{\pi} = 0.017$; $\pi_{\text{hat}} = 1.4$; 95%
144 CI 0.25–2.5; $P_{\tau} = 0.89$). Notably, all the null variants that were detected in the UK biobank
145 were heterozygous.

146 A total of seven mutations were identified in *P4HTM* in seven unrelated probands (all males)
147 in consanguineous families. From the SOPP cohort, five different homozygous missense
148 mutations (p.Pro45Leu, p.Val297Met, p.Val316Ile, p.Gly433Ala and p.Arg509Pro) were
149 identified in 5 unrelated probands (**Table 1; Supplementary Figure 1**). A further screening
150 of *P4HTM* in other cohorts identified 2 additional subjects with severe obesity carrying
151 homozygous missense mutations (p.Gly116Asp and p.Glu155Lys), from consanguineous
152 families of Indian and Moroccan origin, respectively. All the mutations identified and
153 reported here are either novel or with a very rare minor allele frequency (≤ 0.0002). Of these,
154 p.Glu155Lys is present in EF-hand domain whereas p.Val316Ile, p.Gly433Ala and
155 p.Arg509Pro are in prolyl 4-hydroxylase domain (**Table 1; Supplementary Figure 1**).

156 The seven mutation carriers presented severe, early-onset obesity between 0.3 to 3 years with
157 a BMI SDS for age ≥ 4 . In addition to obesity, the carriers presented with hypotonia,
158 developmental delay and cognitive impairment. In addition, a variable set of other

159 abnormalities was also observed in individual cases. Three of 5 probands from SOPP died
160 from respiratory infection during the first ~2 years of the follow up (**Supplementary Figures**
161 **1,2, Table 2**).

162 The prediction of the stability of the protein was carried out for four out of seven mutations
163 where crystal structural information was available. The quantitative stability changes ($\Delta\Delta G$)
164 upon mutations indicated a destabilizing effect and revealed an increase in flexibility
165 (**Supplementary Table 3**). Through extensive molecular dynamics simulations, mutants
166 showed fluctuations in both catalytic domains, EF hand (residues 190 - 290) domain, and the
167 double-stranded β -helix (DSBH) fold (residues 310 - 460) (**Figure 2**). Overall, the structural
168 analysis revealed domain movements compared to its *wt-P4HTM*. Details of the structural
169 interpretation of mutations are provided in the **Supplementary Information**.

170 Finally, using ‘HypoMap’, an integrated reference atlas of the murine hypothalamus (14), we
171 found that *P4HTM* is expressed ubiquitously in all regions of the hypothalamus, and in all
172 cell-types, including neurons, oligodendrocytes, astrocytes and tanycytes (**Supplementary**
173 **Figure 3**). Similar expression of P4HTM was observed in the hypothalamic cells from a
174 single-nucleus sequencing dataset from adult human brains. This broad expression profile is
175 consistent with the observed complex phenotype, beyond severe obesity, of the patients
176 (**Supplementary Figure 5**). Non-fasting serum levels of metabolic hormones, leptin, insulin
177 and cortisol, determined in the five affected patients from the SOPP were within the normal
178 range except in one proband that presented hyperinsulinemia (Table 2).

179

180

181 **DISCUSSION**

182 We report seven paediatric patients with severe obesity carrying rare or novel bi-allelic
183 missense mutations in *P4HTM*, representing a form of monogenic syndromic obesity,
184 following a stepwise gene-centric analysis. The affected subjects were previously found
185 negative for all known obesity genes. The phenotypes of all the affected subjects, in addition
186 to obesity, included hypotonia, intellectual disability, and developmental delay. Importantly,
187 these neurologic characteristics are consistent with the recently proposed HIDEA syndrome
188 (acronym for **h**ypotonia, **i**ntellectual **d**isability, and **e**ye **a**bnormalities). This syndrome was
189 first described in 2014 in a single Finnish family carrying pathogenic mutations in three genes
190 - transketolase (*TKT*), prolyl 4-hydroxylase transmembrane (*P4HTM*), and ubiquitin specific
191 peptidase 4 (*USP4*) (16). HIDEA syndrome was later suggested to be only caused by biallelic
192 pathogenic mutations in *P4HTM*, and associates with hypotonia, hypoventilation, intellectual
193 disability, dysautonomia, developmental delay and eye abnormalities (17). To date, 10
194 families have been reported with homozygous or compound heterozygous mutations in
195 *P4HTM* (16-19). Obesity phenotype was not emphasized in any of the reports, but our
196 secondary assessment of the published phenotypes revealed that in 2 families, mutation
197 carriers presented obesity (17; 19).

198 It is noteworthy that all *P4HTM* mutations found in the SOPP cohort are missense, which is
199 also the case for the three previously published individuals with severe obesity, from the same
200 family(19). On the contrary, all the other families showing severe neurological phenotypes
201 that characterize HIDEA but lacking obesity (with the exception of one individual), have loss-
202 of-function mutations (i.e. stop-gain, frame-shift and splice site). It is possible that this
203 phenotypic heterogeneity reflects an allelic heterogeneity responsible for variable impairment
204 of brain function, that affects or not the central control of appetite and/or energy balance.

205 *P4HTM* encodes the enzyme P4H-TM, an atypical member of the hypoxia inducible factors
206 prolyl 4-hydroxylases (HIF-P4Hs) that are necessary for survival and adaptation under
207 reduced oxygen supply and oxygen deprivation (20). Four members of the HIF-P4H family
208 have been identified so far (21). In the scenario of normal oxygen supply - normoxia, the HIF-
209 P4Hs get activated resulting in post-translational hydroxylation of HIF- α by converting
210 prolines to 4-hydroxyprolines leading to its degradation. In the event when cells suffer from
211 hypoxia, the HIF- α related hydroxylation reactions mediated by the P4Hs are suppressed
212 resulting in HIF- α , accumulation and its dimerization with HIF- β and migration into the
213 nucleus for activation of several target genes responsible for maintaining oxygen homeostasis
214 (22). Unlike other P4Hs members, P4H-TM is a transmembrane enzyme with a high degree of
215 expression in the brain, predominantly in the hippocampus, amygdala and hypothalamic
216 regions (23). However, its physiological function in the brain is yet unknown (24; 25). Our
217 data suggest that genetic disruption of *P4HTM* also causes childhood obesity, possibly
218 through dysregulation of appetite.

219 The very high mortality at a very young age of the children with *P4HTM* mutations, of the
220 SOPP cohort (7 documented deaths including 3 probands and 4 family mutation carriers) is an
221 ongoing tragedy. Furthermore, the deaths are apparently not related to central defects but
222 possibly to lung hypoventilation following pneumonia and possibly associated with defective
223 immunity (as reported in leptin deficient children).

224 Gene-centric analysis using MiST (mixed effects score tests) method was employed to detect
225 this gene. This is an efficient approach specifically designed to prioritise rare variants by
226 using a set-based score testing framework(26). This method uses a set of two score statistics
227 derived by testing the group effect by variant characteristics and the heterogeneity effect
228 making it a powerful and efficient approach. Previously, similar analysis used by us led to the

229 identification of a significant burden of pathogenic variants in genes linked with monogenic
230 diabetes among patients clinically diagnosed with common type 2 diabetes (27).

231 In summary, pathogenic mutations in *P4HTM* cause, in the SOPP cohort, severe obesity
232 associated with neurologic features of the HIDEA syndrome in 1.5% of our probands. It is not
233 restricted to Pakistan as we identified two other patients from consanguineous families of
234 Indian and Moroccan origin. The severity of this obesity associated syndrome with a high risk
235 of mortality, highlights the advisability of screening *P4HTM* in young patients with severe
236 obesity that also present (or express) hypotonia and intellectual disability/developmental
237 delay, for a timely and appropriate management and care. Beyond linking *P4HTM* to obesity,
238 we have demonstrated the power of stepwise statistical analysis using MiST followed by
239 stringent filtration steps in bringing to surface new genes that are also associated with obesity.

240

241 **ACKNOWLEDGMENTS**

242 We are grateful to all individuals included in the different cohort studies. We thank
243 FrédéricAllegaert and Timothée Beke for technical assistance. We also thank Mickaël Canouil
244 for his help in statistical analyses. This work was supported by funding from the Medical
245 Research Council (MRC) MR/S026193/1 (P.F.) and the Pakistan Academy of Sciences
246 (M.A). This research has been conducted using the UK Biobank Application #67575. We
247 thank “France Génomique” consortium (ANR-10-INBS-009). This study was funded by the
248 French National Research Agency (ANR-10-LABX-46 [European Genomics Institute for
249 Diabetes] to PF and AB), the French National Research Agency (ANR-10-EQPX-07-01
250 [LIGAN-PM] to PF and AB), the European Research Council (ERC GEPIDIAB – 294785, to
251 PF; ERC Reg-Seq – 715575, to AB) and the National Center for Precision Diabetic Medicine
252 – PreciDIAB, which is jointly supported by the French National Agency for Research (ANR-

253 18-IBHU-0001), by the European Union (FEDER), by the Hauts-de-France Regional Council
254 and by the European Metropolis of Lille (MEL). Additional support was provided by the
255 Natural Sciences and Engineering Research Council of Canada (DG-2018-06338). G.S.H.Y.
256 is supported by the Medical Research Council (MRC Metabolic Diseases Unit
257 (MC_UU_00014/1)), S.S and A.M.S is supported by a project grant from the Medical
258 Research Council (MR/S026193/1). I.S.F is supported by the Wellcome (207462/Z/17/Z),
259 National Institute for Health and Care Research (NIHR) Cambridge Biomedical Research
260 Centre, Botnar Fondation, Bernard Wolfe Health Neuroscience Endowment and a NIHR
261 Senior Investigator Award. M.U.M. and J.F.T. wish to recognize that this work was made
262 possible by the facilities of the Compute Ontario (<https://www.computeontario.ca>) and the
263 Digital Research Alliance of Canada (www.alliancecan.ca).

264

265 **AUTHORS CONTRIBUTION**

266 SS, AB, MA and FP designed the study and wrote the first draft of the paper. SS, RH, JM,
267 WIK, LC, AT, AK, ISF, and MA recruited the samples. RH, QMJ and MA performed
268 biochemical analysis. SS, A Badreddine, BT, EV, SA, MD and AB performed whole exome
269 sequencing and analyzed the genetic data. LN and MB carried out statistical analysis. MUM
270 and JFT carried out molecular modelling. ANM, BYH and GSHY performed expression
271 analysis in murine hypothalamus. All authors contributed to the final version of the
272 manuscript.

273

274 **CONFLICT OF INTEREST**

275 The authors have declared that no conflict of interest exists.

277 **REFERENCES**

- 278
- 279
- 280 1. Saeed S, Arslan M, Manzoor J, Din SM, Janjua QM, Ayesha H, Ain Q-t, Inam L, Lobbens S, Vaillant E.
- 281 Genetic causes of severe childhood obesity: a remarkably high prevalence in an inbred population of
- 282 Pakistan. *Diabetes* 2020;69:1424-1438
- 283 2. Saeed S, Bech PR, Hafeez T, Alam R, Falchi M, Ghatei MA, Bloom SR, Arslan M, Froguel P. Changes
- 284 in levels of peripheral hormones controlling appetite are inconsistent with hyperphagia in leptin-
- 285 deficient subjects. *Endocrine* 2014;45:401-408
- 286 3. Saeed S, Bonnefond A, Manzoor J, Philippe J, Durand E, Arshad M, Sand O, Butt TA, Falchi M,
- 287 Arslan M. Novel LEPR mutations in obese Pakistani children identified by PCR-based enrichment and
- 288 next generation sequencing. *Obesity* 2014;22:1112-1117
- 289 4. Saeed S, Bonnefond A, Manzoor J, Shabir F, Ayesha H, Philippe J, Durand E, Crouch H, Sand O, Ali
- 290 M. Genetic variants in LEP, LEPR, and MC4R explain 30% of severe obesity in children from a
- 291 consanguineous population. *Obesity* 2015;23:1687-1695
- 292 5. Saeed S, Butt TA, Anwer M, Arslan M, Froguel P. High prevalence of leptin and melanocortin-4
- 293 receptor gene mutations in children with severe obesity from Pakistani consanguineous families.
- 294 *Molecular genetics and metabolism* 2012;106:121-126
- 295 6. Goldstone AP, Beales PL. Genetic obesity syndromes. 2008;
- 296 7. Saeed S, Arslan M, Manzoor J, Din SM, Janjua QM, Ayesha H, Ain QT, Inam L, Lobbens S, Vaillant E,
- 297 Durand E, Derhourhi M, Amanzougarene S, Badreddine A, Berberian L, Gaget S, Khan WI, Butt TA,
- 298 Bonnefond A, Froguel P. Genetic Causes of Severe Childhood Obesity: A Remarkably High Prevalence
- 299 in an Inbred Population of Pakistan. *Diabetes* 2020;69:1424-1438
- 300 8. Saeed S, Bonnefond A, Tamanini F, Mirza MU, Manzoor J, Janjua QM, Din SM, Gaitan J, Milochau A,
- 301 Durand E, Vaillant E, Haseeb A, De Graeve F, Rabearivelo I, Sand O, Queniat G, Boutry R, Schott DA,
- 302 Ayesha H, Ali M, Khan WI, Butt TA, Rinne T, Stumpel C, Abderrahmani A, Lang J, Arslan M, Froguel P.
- 303 Loss-of-function mutations in ADCY3 cause monogenic severe obesity. *Nat Genet* 2018;50:175-179
- 304 9. Sun J, Zheng Y, Hsu L. A unified mixed-effects model for rare-variant association in sequencing
- 305 studies. *Genetic epidemiology* 2013;37:334-344
- 306 10. Montagne L, Derhourhi M, Piton A, Toussaint B, Durand E, Vaillant E, Thuillier D, Gaget S, De
- 307 Graeve F, Rabearivelo I. CoDE-seq, an augmented whole-exome sequencing, enables the accurate
- 308 detection of CNVs and mutations in Mendelian obesity and intellectual disability. *Molecular*
- 309 *metabolism* 2018;13:1-9
- 310 11. Saeed S, Janjua QM, Haseeb A, Khanam R, Durand E, Vaillant E, Ning L, Badreddine A, Berberian L,
- 311 Boissel M. Rare Variant Analysis of Obesity-Associated Genes in Young Adults With Severe Obesity
- 312 From a Consanguineous Population of Pakistan. *Diabetes* 2022;71:694-705
- 313 12. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E.
- 314 Standards and guidelines for the interpretation of sequence variants: a joint consensus
- 315 recommendation of the American College of Medical Genetics and Genomics and the Association for
- 316 Molecular Pathology. *Genetics in medicine* 2015;17:405-423
- 317 13. Bonnefond A, Boissel M, Bolze A, Durand E, Toussaint B, Vaillant E, Gaget S, Graeve FD,
- 318 Dechaume A, Allegaert F. Pathogenic variants in actionable MODY genes are associated with type 2
- 319 diabetes. *Nature Metabolism* 2020;2:1126-1134
- 320 14. Steuernagel L, Lam BY, Klemm P, Dowsett GK, Bauder CA, Tadross JA, Hitschfeld TS, del Rio Martin
- 321 A, Chen W, de Solis AJ. HypoMap—a unified single-cell gene expression atlas of the murine
- 322 hypothalamus. *Nature metabolism* 2022:1-18
- 323 15. Siletti K, Hodge RD, Mossi Albiach A, Hu L, Lee KW, Lönnerberg P, Bakken TE, Ding S-L, Clark M,
- 324 Casper T. Transcriptomic diversity of cell types across the adult human brain. *bioRxiv*
- 325 2022:2022.2010. 2012.511898
- 326 16. Kaasinen E, Rahikkala E, Koivunen P, Miettinen S, Wamelink MM, Aavikko M, Palin K, Myllyharju J,
- 327 Moilanen JS, Pajunen L, Karhu A, Aaltonen LA. Clinical characterization, genetic mapping and whole-

328 genome sequence analysis of a novel autosomal recessive intellectual disability syndrome. *Eur J Med*
329 *Genet* 2014;57:543-551

330 17. Rahikkala E, Myllykoski M, Hinttala R, Vieira P, Nayebzadeh N, Weiss S, Plomp AS, Bittner RE,
331 Kurki MI, Kuismin O, Lewis AM, Vaisanen ML, Kokkonen H, Westermann J, Bernert G, Tuominen H,
332 Palotie A, Aaltonen L, Yang Y, Potocki L, Moilanen J, van Koningsbruggen S, Wang X, Schmidt WM,
333 Koivunen P, Uusimaa J. Biallelic loss-of-function P4HTM gene variants cause hypotonia,
334 hypoventilation, intellectual disability, dysautonomia, epilepsy, and eye abnormalities (HIDEA
335 syndrome). *Genet Med* 2019;21:2355-2363

336 18. Hay E, Wilson LC, Hoskins B, Samuels M, Munot P, Rahman S. Biallelic P4HTM variants associated
337 with HIDEA syndrome and mitochondrial respiratory chain complex I deficiency. *Eur J Hum Genet*
338 2021;29:1536-1541

339 19. Maddirevula S, Ben-Omran T, AlMureikhi M, Eyaid W, Arabi H, Alkuraya H, Alfaifi A, Alfalah AH,
340 Alsaif HS, Abdulwahab F, Alfadhel M, Alkuraya FS. Further delineation of HIDEA syndrome. *Am J Med*
341 *Genet A* 2020;182:2999-3006

342 20. Sanghani NS, Haase VH. Hypoxia-Inducible Factor Activators in Renal Anemia: Current Clinical
343 Experience. *Adv Chronic Kidney Dis* 2019;26:253-266

344 21. Leinonen H, Rossi M, Salo AM, Tiainen P, Hyvärinen J, Pitkänen M, Sormunen R, Miinalainen I,
345 Zhang C, Soininen R. Lack of P4H-TM in mice results in age-related retinal and renal alterations.
346 *Human Molecular Genetics* 2016;25:3810-3823

347 22. Byts N, Sharma S, Malm T, Kaakinen M, Korhonen P, Jaakkonen L, Keuters M, Huuskonen M,
348 Pietilä I, Koistinaho J. Inactivation of mouse transmembrane prolyl 4-hydroxylase increases blood
349 brain barrier permeability and ischemia-induced cerebral neuroinflammation. *Journal of Biological*
350 *Chemistry* 2022;298

351 23. Leinonen H, Rossi M, Salo AM, Tiainen P, Hyvärinen J, Pitkänen M, Sormunen R, Miinalainen I,
352 Zhang C, Soininen R, Kivirikko KI, Koskelainen A, Tanila H, Myllyharju J, Koivunen P. Lack of P4H-TM in
353 mice results in age-related retinal and renal alterations. 2016;

354 24. Oehme F, Ellinghaus P, Kolkhof P, Smith TJ, Ramakrishnan S, Hütter J, Schramm M, Flamme I.
355 Overexpression of PH-4, a novel putative proline 4-hydroxylase, modulates activity of hypoxia-
356 inducible transcription factors. *Biochemical and biophysical research communications* 2002;296:343-
357 349

358 25. Laitala A, Aro E, Walkinshaw G, Mäki JM, Rossi M, Heikkilä M, Savolainen E-R, Arend M, Kivirikko
359 KI, Koivunen P. Transmembrane prolyl 4-hydroxylase is a fourth prolyl 4-hydroxylase regulating EPO
360 production and erythropoiesis. *Blood, The Journal of the American Society of Hematology*
361 2012;120:3336-3344

362 26. Su YR, Di C, Bien S, Huang L, Dong X, Abecasis G, Berndt S, Bezieau S, Brenner H, Caan B, Casey G,
363 Chang-Claude J, Chanock S, Chen S, Connolly C, Curtis K, Figueiredo J, Gala M, Gallinger S, Harrison T,
364 Hoffmeister M, Hopper J, Huyghe JR, Jenkins M, Joshi A, Le Marchand L, Newcomb P, Nickerson D,
365 Potter J, Schoen R, Slattery M, White E, Zanke B, Peters U, Hsu L. A Mixed-Effects Model for Powerful
366 Association Tests in Integrative Functional Genomics. 2018;

367 27. Bonnefond AA-O, Boissel MA-O, Bolze A, Durand E, Toussaint B, Vaillant E, Gaget S, Graeve F,
368 Dechaume A, Allegaert F, Guilcher DL, Yengo L, Dhennin V, Borys JM, Lu JA-O, Cirulli ET, Elhanan G,
369 Roussel R, Balkau B, Marre M, Franc S, Charpentier G, Vaxillaire M, Canouil MA-O, Washington NL,
370 Grzymalski JJ, Froguel PA-O. Pathogenic variants in actionable MODY genes are associated with type 2
371 diabetes. 2020;

372

373

374

375

376
377

378
379
380

381 **Table 1** Homozygous mutations identified in *P4HTM* (NM_177938.2) among severely obese
 382 children from consanguineous populations.

| ID | Mutation | PolyPhen | SIFT | Origin (study) | MAF overall (in GnomAD) | MAF SA in GnomAD) |
|--------------------|---------------------------|-------------------|-------------|-----------------------|--------------------------------|--------------------------|
| Proband I | c.1298G>C; p.Gly433Ala | Probably Damaging | Damaging | Pakistan (SOPP) | NA | NA |
| Proband II | c.946G>A; p.Val316Ile | Probably Damaging | Damaging | Pakistan (SOPP) | 0.000003 | 0.000 |
| Proband III | c.1526G>C; p.Arg509Pro | Probably Damaging | Damaging | Pakistan (SOPP) | NA | NA |
| Proband IV | c.889G>A; p.Val297Met | Probably Damaging | Damaging | Pakistan (SOPP) | 0.00001 | 0.00013 |
| Proband V | c.134C>T; p.Pro45Leu | Possibly Damaging | Damaging | Pakistan (SOPP) | 0.0002 | 0.0021 |
| Proband VI | c.347G>A; p.Gly116Asp | Probably Damaging | Damaging | Indian (GOOS) | 0.000005 | 0.00003 |
| Proband VII | c.463G>A; p.Glu155Lys | Probably Damaging | Damaging | Morocco | NA | NA |

383 MAF SA: Minor Allele Frequency South Asian, NA: Not available.

Table 2 Clinical profile of severely obese children identified with balletic mutations in *P4HTM*.

| ID | Proband I | Proband II | Proband III | Proband IV | Proband V | Proband VI | Proband VII |
|----------------------------------|---|---|--|---|--------------------------------------|--|--------------------------------|
| Age at first recruitment (Years) | 4.5 | 0.9 | 5.0 | 3.9 | 9.6 | 3.0 | 4.0 |
| Sex M/F | M | M | M | M | M | M | M |
| Age of obesity onset (Years) | 0.9 | 0.3 | 1.0 | 0.6 | 3.0 | NA | NA |
| SDS (BMI for age) | 6.8 | 6.7 | 6.8 | 8.0 | 5.7 | NA | NA |
| Alive/deceased | Deceased (5.2 years of age due to severe pneumonia) | Deceased (1.5 years of age due to pneumonia and high-grade fever) | Deceased (6 years of age; cause of death unknown) | Alive | Alive | Alive | Alive |
| Leptin (ng/ml) | 3 | 18 | 10 | 16 | 12 | NA | 41.9 |
| Insulin (uIU/ml) | 26 | 19 | 122 | 5 | 12 | NA | NA |
| Cortisol (ug/dl) | 11 | 7 | 11 | 10 | 6 | NA | NA |
| Intellectual disability | Yes | Yes | Yes | Yes | No | Yes | Yes |
| Hypotonia | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Deaths in family | The sisters at the age of 2 and 3.5 years and one cousin at age of 1.5 died as the result of kidney dysfunction (proteinuria) | His elder brother who shared the similar clinical phenotype died at the age of 2 years due to pneumonia | NR | NR | NR | NR | NR |
| Birth by caesarean section | Yes | Yes | No | Yes | No | NA | NA |
| Additional features | Polydipsia and polyuria, eye abnormalities including astigmatism, reported stubborn and aggressive in temperament | Suffered hepatomegaly, asthma and abdominal pain | Delayed milestones, aggressive behaviour, self-mutilation/self-harm, neonatal hypotonia, Thalassemia | Scoliosis, abnormal gait (needs support for walking, pneumonia twice but recovered, aggressive and stubborn | Asthma, breathing issue, sleep apnea | Epilepsy, hypoventilation, developmental and cognitive impairment, autonomic dysfunction | Hearing issues, cryptorchidism |

F: Female; M: Male; NR: Not reported; SDS: Standard Deviation Score.

Figure Legends:

Figure 1. Schematic presentation of stepwise analyses and filtration strategy used in this study

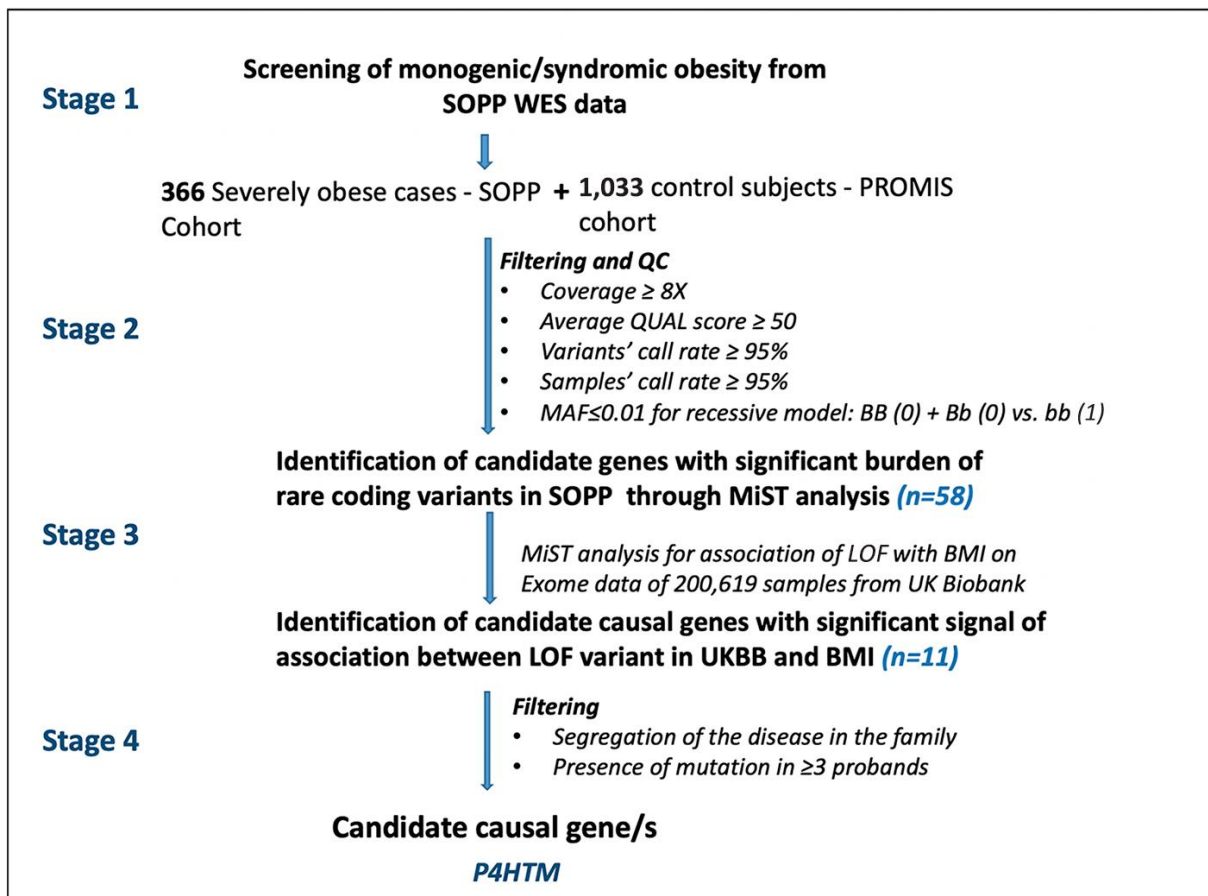


Figure 2. Structural interpretation of *P4HTM* missense variants. **A)** The *P4HTM* structure is presented as a secondary structure topology, where the EF domain is shown in orange, the DSBH core in green, and the rest of the protein in yellow. **B)** Root-mean-square-deviation of mutants and wild-type (black) obtained over a period of 100 ns. The individual domains are mentioned, while iron-binding residues are highlighted in the yellow section. **C)** The MD simulated conformations (10 snapshots after every 20 ns) of mutants and wild type. Different regions of *P4HTM* are colored to increase the visibility of backbone deviations. **D)** The movement of domains (EF and DSBH) are displayed during MD simulations. **E)** Superimposition of MD simulated conformations of FG-2216 (sticks representation in distinct colour) in the active site of mutants and wild-type. Fe^{2+} is positioned inside the active site of *P4HTM* (molecular surface representation). **F)** Molecular mechanics generalized born surface area (MM-GBSA) binding free energies (kcal/mol) for mutants and wild-type/FG-2216 complex.

