Title: Contributions of function-altering variants in genes implicated in pubertal timing and body mass for self-limited delayed puberty

Authors – Sasha R Howard,¹ Leonardo Guasti,¹ Ariel Poliandri,¹ Alessia David,⁴ Claudia P Cabrera,²,³ Michael R Barnes,²,³ Karoliina Wehkalampi,⁵ Stephen O’Rahilly,⁶ Catherine E Aiken,⁷ Anthony P Coll,⁶ Marcella Ma,⁶ Debra Rimmington,⁶ Giles SH Yeo,⁶ Leo Dunkel¹*

Affiliations –
¹Centre for Endocrinology, William Harvey Research Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, EC1M 6BQ, UK
²Centre for Translational Bioinformatics, William Harvey Research Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, EC1M 6BQ, UK
³NIHR Barts Cardiovascular Biomedical Research Unit, Queen Mary University of London, London, EC1M 6BQ, UK
⁴Centre for Integrative Systems Biology and Bioinformatics, Department of Life Sciences, Imperial College London, London, SW7 2AZ, UK
⁵Children’s Hospital, Helsinki University Hospital and University of Helsinki, Helsinki, FIN-00029 HUS, Finland
⁶University of Cambridge Metabolic Research Labs and MRC Metabolic Diseases Unit, Institute of Metabolic Science, Addenbrooke’s Hospital, Cambridge, CB2 0QQ, UK
⁷Department of Obstetrics and Gynaecology, University of Cambridge; NIHR Cambridge Comprehensive Biomedical Research Centre, Cambridge, UK
*Correspondence: l.dunkel@qmul.ac.uk

Word count: 3684

Keywords: Puberty; delayed puberty; FTO; paediatric endocrinology, constitutional delay

Short title: Role of GWAS genes in self-limited delayed puberty

Declaration of interest:
The authors declare no conflict of interest

Funding:
SRH is funded by The Wellcome Trust (102745), Rosetrees Trust (M222) and the Barts and the London Charity (417/1551). AP is funded by Bart's & The London Charity Strategic Research Grant (MEAG16R). LG is funded by the Biotechnology and Biological Sciences Research Council (BB/L002671/1). LD is partly supported by the Academy of Finland (14135).

MRB and CPC are funded by the National Institutes for Health Research (NIHR), and this work forms part of the portfolio of translational research of the NIHR Biomedical Research Unit at Barts. AD is funded by the MRC (MR/K021613/1). DR, GSHY, SOR and APC are funded by the Medical Research Council (MRC) Metabolic Disease Unit (MRC_MC_UU_12012/1) and animal work was carried out with the assistance of MRC Disease Model Core of the Wellcome Trust MRC Institute of Metabolic Sciences (MRC_MC_UU_12012/5).
1. I, the designated corresponding author, on behalf of myself and my co-authors, hereby transfer and assign all right, title, and interest, including copyright and any moral rights, in and to the manuscript named in this submission (called the Work hereafter) to the Endocrine Society (ES). If ES ultimately declines to publish the Work in an ES journal, all rights in and to the Work will revert to the author(s).

2. I, and all co-authors, warrant that the Work intended for publication is original and has not been published other than as an abstract or preprint in any language or format and has not been submitted elsewhere for print or electronic publication consideration. We further warrant that the Work does not contain any material that is defamatory or the publication of which would violate any copyright or other personal, intellectual, property, contract, or proprietary right of any person or entity.

3. I warrant that each person listed as an author participated in the Work in a substantive way and is prepared to take public responsibility for it. All authors consent to the investigation of any improprieties that may be alleged regarding the Work. Each author further releases and holds harmless the Endocrine Society from any claim or liability that may arise therefrom.

4. I warrant that I am authorized to accept the terms of this agreement on behalf of myself and all co-authors.

   ___ I accept the terms of the agreement on behalf of myself and all co-authors.

5. I certify that neither I nor my co-authors have a conflict of interest as described above that is relevant to the subject matter or materials included in this Work.
Contributions of function-altering variants in genes implicated in pubertal timing and body mass for self-limited delayed puberty

Abstract
Context: Self-limited delayed puberty (DP) is often associated with delay in physical maturation, but whilst highly heritable the causal genetic factors remain elusive. Genome-wide association studies of the timing of puberty have identified multiple loci for age of menarche (AAM) in females and voice break in males, particularly in pathways controlling energy balance.

Objective/Main outcome measures: We aimed to assess the contribution of rare variants in such genes to the phenotype of familial DP.

Design/Patients: We performed whole exome sequencing (WES) in 67 pedigrees (125 individuals with DP and 35 unaffected controls) from our unique cohort of familial self-limited DP. Using a WES filtering pipeline one candidate gene (FTO) was identified. In silico, in vitro and mouse model studies were performed to investigate the pathogenicity of FTO variants and timing of puberty in FTO+/− mice.

Results: We identified potentially pathogenic, rare variants in genes in linkage disequilibrium with GWAS of AAM loci in 283 genes. Of these, 5 genes were implicated in the control of body mass. After filtering for segregation with trait one candidate, FTO, was retained. Two FTO variants, found in 14 affected individuals from 3 families, were also associated with leanness in these DP patients. One variant (p.Leu44Val) demonstrated altered demethylation activity of the mutant protein in vitro. Fto+/− mice displayed a significantly delayed timing of pubertal onset (p <0.05).
Conclusions: Mutations in genes implicated in body mass and timing of puberty in the general population may contribute to the pathogenesis of self-limited DP.

Introduction

Puberty is the maturational process of the reproductive endocrine system that results in adult height and body proportion, in addition to the capacity to reproduce. A minimum level of energy availability is required for the onset of puberty, whilst increased fat mass has been shown to be associated with precocious onset of puberty\(^1,2\). However, a role for genes connected with regulation of body mass have not been clearly demonstrated in pubertal timing.

The existence of genetic heterogeneity in pubertal timing is supported by several large genome wide association studies (GWAS) of the age of menarche (AAM)\(^3-5\). Evidence (P< 5 x 10\(^{-8}\)) for 123 signals at 106 genomic loci has been identified. Many of these loci were associated with Tanner staging in both sexes, suggesting this data is applicable to both men and women\(^6,7\).

The first of many GWAS loci associated with AAM was the developmental gene \(LIN28B\)\(^3,8\). Additional signals in genes involved in energy homeostasis and growth have been found near \(LEPR-LEPROT\), which encodes the leptin receptor. Leptin (a key regulator of body mass) is an important permissive signal for the onset of puberty\(^9\). In addition to leptin signaling, overlap with several genes implicated in body mass index was found, including \(FTO\), \(SEC16B\), \(TMEM18\), and \(NEGR1\) (Supplementary Table 1)\(^5\). Whether such
genes may regulate pubertal timing exclusively via impact on fat mass or via other BMI-independent mechanisms is unknown\(^{(10)}\).

Disordered pubertal timing affects up to 5% of adolescents and is associated with adverse health and psychosocial outcomes\(^{(11-14)}\). Self-limited delayed puberty (DP) represents the extreme end of normal pubertal timing, and is defined as the absence of testicular enlargement in boys or breast development in girls at an age that is 2 to 2.5 standard deviations (SD) later than the population mean\(^3\). DP may be an isolated feature of the condition or be associated with constitutional delay in growth that can manifest from early childhood.

DP segregates within families, usually with an autosomal dominant pattern of inheritance\(^{(15,16)}\). Despite strong heritability in most cases the genetic basis of DP remains elusive\(^{(17)}\). Moreover, the relevance of genetic factors influencing timing of puberty in the general population to patients with extreme pubertal delay has not been explored. Given the importance of energy balance for reproductive health, genes identified by AAM GWAS that relate to energy homeostasis are of particular interest. Our multi-generational DP families provide a highly valuable resource to investigate these candidate genes in familial DP.

**Materials and Methods**

**Patients**

The patients selected for this study are taken from a previously described, accurately phenotyped and characterized, Finnish DP patient cohort\(^{(19)}\).

Diagnosis is based on objective evidence of a delayed pubertal growth spurt
rather than self-recall. Patients referred with DP to specialist paediatric care in central and southern Finland (1982-2004) were identified. All patients (n=492) met the diagnostic criteria for self-limited DP, defined as the onset of Tanner genital stage II (testicular volume >3 ml) >13.5yr in boys or Tanner breast stage II >13.0yr in girls (i.e. two SD later than average pubertal development). Pubertal growth spurt in probands was more than 2 SD later than average: age at acceleration of pubertal growth (take-off) beyond 13.8 and 12.2 yr and age at peak height velocity (PHV) later than 15.6 and 13.7 yr in males and females, respectively (21).

Chronic illness and undernutrition was excluded by medical history, clinical examination, and routine laboratory tests. HH, if suspected, was excluded by spontaneous pubertal development at follow-up. In the 50% of patients who choose to have pubertal induction via the use of exogenous sex steroids, all patients were followed up until the point of full pubertal development (Tanner stage G4+ or B4+) to ensure development did not arrest off treatment.

Families of the DP patients were invited to participate, with information about medical history and pubertal timing obtained by structured interviews and from archived height records. The criteria for DP in probands' family members were one or more of: 1) age at takeoff or 2) PHV occurring 1.5 SD beyond the mean, i.e. age at takeoff exceeding 12.9 and 11.3 yr, or age at PHV exceeding 14.8 and 12.8 yr in males and females, or 3) age at attaining adult height more than 18 or 16 yr, in males and females, respectively (19). Previous linkage analysis from this cohort did not find evidence for linked families sharing chromosomal segments identical by
descent, suggesting a founder effect is unlikely to be responsible for this phenotype (19).

Written informed consent was obtained from all participants. The study protocol was approved by the Ethics Committee for Pediatrics, Adolescent Medicine and Psychiatry, Hospital District of Helsinki and Uusimaa (extended to encompass Kuopio, Tampere and Turku University Hospitals) (570/E7/2003). UK ethical approval was granted by the London-Chelsea NRES committee (13/LO/0257). The study was conducted in accordance with the guidelines of The Declaration of Helsinki.

Genetic Analysis

Genetic analysis was performed in 160 individuals from the 67 most extensive families from our cohort with DP. These included 67 probands (male n=57, female n=10), 58 affected family members (male n=36, female n=22) and 35 unaffected family members (male, n=13, female n=22). Whole exome sequencing (WES) was performed on DNA extracted from peripheral blood leukocytes. Variants were analyzed and filtered for potential causal variants in Ingenuity Variant Analysis (Qiagen) using filters for quality control, predicted functional annotation, minor allele frequency (MAF), and GWAS relevance (Figure 1). GWAS relevance filtering allowed identification of those remaining variants that lay within genes in linkage disequilibrium with 106 GWAS loci associated with AAM (n=760) (5). Filters for genes implicated in body mass regulation were applied using a biological context filter with pathway analysis. Variants were filtered for segregation with trait in family members using conventional Sanger sequencing.
Targeted exome sequencing using a Fluidigm array of the remaining candidate gene identified post-filtering was then performed in a further 42 cohort families (288 individuals, 178 with DP; male=106, female=69 and 110 controls; male=55, female=58, Figure 1). Whole gene rare variant burden testing was performed post sequencing.

Growth Pattern Analysis

The pattern of prepubertal growth in the individuals carrying FTO variants was analyzed by using five screening parameters: 1) height for age standard deviation score (HSDS); 2) body mass index (BMI; calculated as weight in kilograms divided by height in meters squared) for age SDS (BMI SDS); 3) HSDS distance from target height (TH) (TH formula = 0.791 × mean parental height SDS − 0.147 for girls and 0.886 × mean parental height SDS − 0.071 for boys; 4) change in height SDS (ΔHSDS); 5) change in BMI SDS (ΔBMI SDS) across time with free age intervals. The calculations of the age-specific and sex-specific normal values for ΔHSDS and ΔBMI SDS were based on longitudinal reference measurements (22). Normality of linear growth was tested by using auxological screening rules based on data from >70,000 healthy Finnish children (23).

In silico Analysis

The FTO experimentally solved structure (PDB identifier: 4cxx) was used to study the structural effect of FTO variants. The following interactions involved in protein stability were considered: i) salt bridges; ii) hydrogen bonds (H-bond); and iii) disulphide bridge (S-S bridge). N-glycosylation sites were determined based on the consensus sequence Asn-X-Thr/Ser (X= any amino
acid, except proline). The DSSP program was used to calculate surface accessibility and Disopred3 \(^{(24)}\) to predict disordered protein regions.

**Functional Annotation of FTO mutant proteins**

Cloning of wild-type human FTO cDNA into pET302/NT-His has been described previously \(^{(25)}\). The p.Leu44Val and p.Ala163Thr point mutations were introduced using PCR-mediated mutagenesis (Quickchange II, Agilent Technologies) using primers FTO_L44V FOR: 5'-GAATTCTATCAGCAGTGCAAGGAATATCCTAAACTAATTCT-3', REV: 5'-AGAATTAGTTTAGATATTTACCTGCCACTGCGATAGAATTC-3' and FTO_A163T FOR: 5'-CACAGCATCCTCATTAGTCTTCTTTGGCAGCAA-3', REV: 5'-TTGCTGCCAAAGAGAAGACTAATGAGGATGCTGTG-3' and verified by sequencing. An RNase-cleavage assay \(^{(26)}\) was used to measure the demethylation activity of FTO on 3-methyl-uridine (3-meU). Recombinant wild-type and mutant FTO expression plasmids were transformed into *Escherichia coli* BL21-Gold (DE3) (Stratagene) and cultured in LB broth and 50 μg/ml carbenicillin. Expression of the cloned gene was induced by the addition of IPTG (isopropyl-β-D-1-thiogalactopyranoside) at 1 mM final concentration at 15°C for 4 h. The cells were harvested and pellets resuspended in lysis buffer [50 mM HEPES-KOH (pH 8.0), 2 mM 2-mercaptoethanol, 5% glycerol and 300 mM NaCl] before digestion with lysozyme (1 mg/ml). The cleared lysate was supplemented with imidazole (final concentration 10 mM) before mixing with 1 ml of pre-washed Ni-NTA (Ni2+-nitrilotriacetate) beads (Qiagen). After binding for 1 h in the cold, the mixture was washed with lysis buffer supplemented with increasing concentrations of imidazole. FTO was eluted with 2 ml of lysis buffer.
containing 250 mM imidazole. The eluate was concentrated with a 30 kDa molecular-mass cut-off concentrator (Sartorius Stedim) with buffer changing to 20 mM HEPES-KOH (pH 8), 5 % glycerol and 50 mM NaCl. Purified proteins were snap-frozen and stored at −80°C. Protein purity was estimated by Comassie Blue staining after resolving by SDS/PAGE (4–12 % gradient gels; Invitrogen).

Dose response of FTO on 3-meU demethylation: Recombinant FTO proteins were assayed as previously described (26). Each protein, at different protein concentrations from 0 -1000 nM, was assayed in a reaction containing 100 nM substrate, 75 μM Fe(NH4)2(SO4)2, 300 μM 2-OG, 2 mM ascorbate, 50 μg/ml BSA and 62.5 μg/μl of RNase A in 50 mM Tris/HCl buffer at pH 7.0. Samples were prepared in duplicate in a dark flat-bottomed 96-well plate and the FAM (6-carboxyfluorescein) emission was measured for 30 min at a wavelength of 520 nm with excitation at 485 nm. The measurement was performed at room temperature (25°C) using a microplate reader [Infinite M1000, Tecan]. Wild type (WT) FTO protein and catalytically inactive mutant p.Arg316Gln (R316Q) served as positive and negative controls respectively.

Mouse experiments

Fto deficient mice were a generous gift from Prof. Roger Cox (MRC Harwell, Oxford) and were genotyped as previously described (27). This research is regulated under the Animals (Scientific Procedures) Act 1986 Amendment Regulations 2012 following ethical review by the University of Cambridge Animal Welfare and Ethical Review Body (AWERB). Animals were kept under controlled temperature (22°C) and a 12-h light, 12-h dark schedule (lights on
07:00–19:00). Standard chow (Special Diet Services) and water were available *ad libitum*.

For the vaginal opening study female Fto heterozygous mice (*Fto*+/-) (n=45) and their WT littermates (n=24) were taken from either a male Fto WT x female Fto+/- cross or a male Fto+/- x female Fto WT cross. From P21 (day of weaning) all female mice were weighed and visual examination of the vagina was carried out by placing the mouse on top of a cage lid and lifting the tail vertically away from the body. No excessive force was involved. First day of vaginal opening was recorded when a complete opening was observed.

For all experiments, data are expressed as the mean ± SEM. To determine statistical significance, we used the unpaired t test (2-tailed) using SPSS Software (version 24). A p value of <0.05 was considered statistically significant.

**Results**

Variants in GWAS genes implicated in body mass were identified following exome sequencing in families with self-limited delayed puberty

WES performed in the 67 largest and best phenotyped families from our cohort (160 individuals: a total of 125 individuals with DP, male=93, female=32; and 35 controls, male=13, female=22), identified 6,952,773 variants after quality control (Figure 1). Filtering to identify high quality, rare, predicted deleterious variants not present in control subjects selected 12,371 variants in 7,470 genes. Of these 7,470 genes, 238 were found to be in linkage disequilibrium with a GWAS locus for timing of puberty, and 5 of these
238 were genes implicated in body mass regulation or growth by pathway
analysis. Of these 5 genes, 4 (GPD2, GHR, ESR1 and VDR) were found to
have only variants that did not segregate with the DP trait in family members.
The remaining candidate gene, FTO (Fat mass and obesity-associated
protein, ENSG00000140718, gene identification number 79068), has been
previously described in the literature as involved in pathways of energy
homeostasis and growth\(^5\), and is known to act as an Fe(II) 2-OG (2-
oxoglutarate) -dependent dioxygenase to repair alkylated DNA and RNA by
demethylation\(^26\). FTO contributes to the regulation of energy balance, and
thus to the regulation of body size and fat accumulation.

Two variants in FTO (NM_001080432.2: c.130C>G p.Leu44Val and
NM_001080432.2: c.487G>A (rs145884431) p.Ala163Thr) were identified in
three families from our cohort and found in one or fewer control subjects (rare
variant burden testing adjusted p = 0.058). Both variants are rare (MAF <
0.2%) heterozygous missense variants and predicted benign or tolerated by
>2/5 prediction software tools.

Families with potentially pathogenic FTO variants display autosomal
dominant inheritance of DP phenotype and low body mass.

The family identified with the p.Ala163Thr variant (family 1) and both of the
families with the p.Leu44Val variant (families 2 and 3) displayed the typical
autosomal inheritance pattern of the DP trait, with perfect segregation (Figure
2, panel A). Affected individuals from family 1 with the p.Ala163Thr variant
and from family 3 with the p.Leu44Val variant were particularly underweight in
childhood, with the two probands from these families (individuals 1.III.2 and
3.III.2) falling into the thinness grade 2 category\(^28\) before puberty (Figure 2,
family members carrying FTO variants had ISO-BMI values in the lower range (<23) (Figure 2 and Supplementary Fig. 1-3, Table 1). In addition, both of the probands from families 2 and 3 who carry the p.Leu44Val displayed faltering growth in early childhood. Both displayed significant deflection from previous height measurements in the 2 years following birth, as well as height significantly below target height in later adolescence associated with delayed pubertal growth (Figure 2, panels C and D) (22).

**In silico analysis of potential mutations**

We carried out *in silico* analysis using the solved structure of FTO (PDB identifier: 3lfm) to determine the possible pathogenicity of the identified variants. The hydrophobic residue Leucine 44 is part of a solvent-exposed alpha helix on the surface. Substitution with Valine is not predicted to alter the structure of FTO or interaction with iron molecules or DNA. However, L44 and other residues in the same solvent-exposed alpha helix form a motif (Supplementary Fig. 4 and 5), which is highly conserved across placental mammals but not reptiles, birds or fish (Supplementary Fig. 6). This motif (residues 36-48) forms a patch on the FTO protein surface (Supplementary Fig. 7). This may act as a mammal-specific interaction site (between FTO and another protein), required for FTO function for example in reproductive development. In this scenario, a small change in side chain volume, such as Leucine-to-Valine, may have a subtle effect in protein-protein interaction and lead to a change in FTO activity *in vivo.*

Alanine 163 is a hydrophobic, not highly conserved residue (Supplementary Fig. 8, panel A). Alanine 163 is at the end of the H4 alpha helix and the
beginning of a long, disordered region (Supplementary Fig. 8, panel B), which connects helices H4 and H5 (Supplementary Fig. 8, panel C).

**FTO p.Leu44Val mutant protein displays reduced demethylase activity in vitro**

We carried out functional characterization of the identified mutant FTO proteins (p.Leu44Val and p.Ala163Thr) as compared to WT protein. A previously verified RNase-cleavage assay was used to measure the demethylation activity of FTO on 3-meU (26). Although kinetic activity of the mutant protein p.Ala163Thr did not vary from WT using this assay, mutant protein p.Leu44Val showed an approximately 20% lower kinetic activity than WT activity (Figure 3).

**FTO deficiency in vivo results in delayed vaginal opening in mice**

In order to examine the influence of FTO activity on pubertal timing in an *in vivo* model, we examined timing of puberty in mice deficient for FTO in the heterozygous state (*Fto*^+/-^), in keeping with the human genotype identified. *Fto*^−/−^ mice were not selected for these experiments because of their poor postnatal health (29). *Fto*^−/−^ mice had significantly delayed timing of vaginal opening (VO) (mean postnatal day +/- SEM: 27.20 +/- 0.44 in wild-type (n=24) vs 28.56 +/- 0.48 in *Fto*^+/−^ mice (n=45), p =0.047), an event which reflects the pubertal rise in estradiol (30) (Figure 4). Mean body weight of the *Fto*^+/−^ group was not significantly different to the WT mice (mean body weight (in g) +/- SEM: 11.64 +/- 0.21 in wild-type vs 11.45 +/- 0.14 in *Fto*^+/−^ mice, p=0.467) (Figure 5).

Using simple linear modelling, *Fto* genotype of the pup (Het vs WT) explained approximately 3% of the total variation in timing of VO. Consideration of an
additional factor, maternal genotype, improved the model by increasing the significance of the association between pup genotype and timing of VO slightly (p=0.04), and accounted for 6% of the total variation in timing of VO. In contrast, paternal genotype decreased the significance and total variation accounted for by the model.

Discussion

Genome wide association studies of AAM in the general population have attempted to unravel the complex conundrum of which genetic factors influence the timing of puberty. Despite many loci being identified, clear evidence for the role of particular genes and pathways is for the most part lacking. Those genes lying within pathways of energy metabolism and growth appear promising, with the discovery of the role of Lin28B in C.elegans development\(^3\) and the importance of leptin as a permissive signal in triggering the onset of puberty\((9,31)\).

The inheritance of DP is known to be under strong genetic influence with commonly an autosomal dominant inheritance pattern, and thus represents a useful basis for the investigation of puberty genetics. Notably, self-limited or constitutional DP is often associated with slow maturation throughout childhood, implicating growth and energy metabolism pathways in its pathogenesis. Previously, genes in such pathways identified through GWAS have not been screened in patients with DP.

Our results have identified variants in \(FTO\) as a potential contributory factor in the development of self-limited DP in three pedigrees from our large cohort of patients with familial DP. \(FTO\) (fat mass and obesity associated gene) was the
first obesity-susceptibility gene identified through GWAS and continues to be the locus with the largest effect on body mass index (BMI) and obesity risk. Those DP patients identified with FTO variants from our study showed reductions in body mass. The FTO variants carried by our DP patients may result in reduced fat mass, which would in turn contribute to a delay in the timing of pubertal onset. This delay may be mediated directly through reduced leptin levels. Although we do not routinely measure leptin levels in DP patients, leptin levels have been shown to be significantly lower in pubertal-age patients with self-limited DP.

Notably, in an in vivo model Fto+/− mice had a significantly delayed onset of puberty as compared to WT mice. In the 7 days preceding puberty onset, however, body weight was not significantly different between the two pup genotype groups. Previous studies have demonstrated that Fto−/− mice show a 30-40% reduction in body weight by 6 weeks of age and that transgenic mice with additional copies of Fto show a dose-dependent increase in body and fat mass. However, the relationship between FTO genotype, fat mass and leptin levels remains somewhat unclear. Fto deficient mice do become obese when subjected to a high fat diet, although they remain sensitive to the anorexigenic effects of leptin.

Moreover, it is possible that FTO gene dosage may have an effect on energy homeostasis independent of effects on fat mass, including on the balance between catabolic and anabolic pathways. FTO has been identified as an amino acid sensor acting, via mTOR, to influence appropriate levels of development and translation. FTO is expressed within the hypothalamus in several sites critical for energy balance, including in the arcuate nucleus.
within proopiomelanocortin (POMC) neurons (37, 38). In one study Fto levels in
the arcuate nuclei of fasted mice fell by up to 60%, and this was not rescued
by leptin administration. Other studies have shown conflicting results in the
effects on Fto mRNA levels of fasting, depending on whether whole
hypothalamus or arcuate nucleus were studied and on the length of fast (38).
However, Fto−/− mice display blunted starvation-induced Npy mRNA
induction (29). More recent studies have suggested that Fto may influence the
metabolic outcomes of a high fat diet via hypothalamic signaling pathways
acting independently of body weight (34). Mutations in FTO, including those
with greatly reduced demethylase activity (e.g. pR316Q, Figure 3), have been
identified in human subjects associated with both lean and obese phenotypes
(25). We were not able in our study to identify the mechanism by which the
p.Alα163Thr variant might affect protein function; although no reduction in
demethylation activity was demonstrated it is possible that this variant may
produce a deleterious effect by another route, for example defects in post-
translational modification or protein degradation.
Thus, FTO may be important for signaling energy sufficiency
and the ‘healthy energy balance’ required for pubertal onset. Our in
silico analysis suggests that the p.Leu44Val mutation we have identified may
represent a mammal-specific interaction site between FTO and another
protein (or DNA), important for FTO function in terms of reproductive
development. Moreover, maternal genotype may contribute to pubertal timing,
as demonstrated from our Fto+/− mice data. A reproductive phenotype present
in Fto heterozygote mothers could expose pups to a suboptimal environment
that could influence their puberty timing.
Finally, our finding of maturational delay in growth in early childhood in the two probands with p.Leu44Val mutation is of interest. Constitutional delay in growth is seen in a subset of patients with DP, and our findings implicate mutations in energy pathway genes in the pathogenesis of patients with such a phenotype.

Overall, our discovery of two rare variants in FTO associated with self-limited DP in our large familial cohort, and of delayed vaginal opening in FTO-deficient mice, provides evidence that perturbations in pathways of energy homeostasis and growth may potentially produce a phenotype of DP. We note that despite this extensive analysis, only three of 67 probands were identified with potentially pathogenic variants in such pathways, highlighting the high degree of heterogeneity in the genetic basis of self-limited DP. These findings merit further exploration in our own cohort and in other populations, including sub-group analysis of DP patients with low BMI from early childhood.

Acknowledgements

We are very grateful to the patients and families who contributed their time, medical information, and DNA samples to this study.
References


2. He Q, Karlberg J. Bmi in childhood and its association with height gain, timing of puberty, and final height. Pediatric research 2001; 49:244-251


Barash IA, Cheung CC, Weigle DS, Ren H, Kabigting EB, Kuijper JL, Clifton DK, Steiner RA. Leptin is a metabolic signal to the reproductive system. Endocrinology 1996; 137:3144-3147

Yeo GS. The role of the FTO (Fat Mass and Obesity Related) locus in regulating body size and composition. Molecular and cellular endocrinology 2014; 397:34-41


Sedlmeyer IL. Pedigree Analysis of Constitutional Delay of Growth and Maturation: Determination of Familial Aggregation and Inheritance Patterns. Journal of Clinical Endocrinology & Metabolism 2002; 87:5581-5586


26. Ma M, Harding HP, O'Rahilly S, Ron D, Yeo GS. Kinetic analysis of FTO (fat mass and obesity-associated) reveals that it is unlikely to function as a sensor for 2-oxoglutarate. The Biochemical journal 2012; 444:183-187


33. Church C, Moir L, McMurray F, Girard C, Banks GT, Teboul L, Wells S, Bruning JC, Nolan PM, Ashcroft FM, Cox RD. Overexpression of Fto leads
to increased food intake and results in obesity. Nature genetics 2010; 42:1086-1092.


Whole exome sequencing was initially performed on DNA extracted from peripheral blood leukocytes of 160 individuals from the 67 most extensive families from our cohort (125 with DP and 35 controls), with exome capture on a Nimblegen V2 or Agilent V5 platform and sequencing on the Illumina Hiseq 2000. The exome sequences were aligned to the UCSC hg19 reference genome. Picard tools and the genome analysis toolkit were used to mark PCR duplicates, realign around indels, recalibrate quality scores and call variants. Variants were then analyzed further and filtered for potential causal variants using filters for quality control, predicted functional annotation, minor allele frequency (MAF), segregation with trait and GWAS relevance (See methods for further information on filtering criteria). Targeted exome sequencing using a Fluidigm array of a candidate gene identified post-filtering was then performed in a further 42 families from the same cohort (288 individuals, 178 with DP and 110 controls). Variants post targeted re-sequencing were filtered using the same criteria as the whole exome sequencing data. Functional annotation of the variants as described elsewhere in methods. DP – delayed puberty.

Functional annotation of the variants as described elsewhere in methods. DP – delayed puberty.
Panel A: Squares indicate male family members, circles female family members. Black symbols represent clinically affected, grey represent unknown phenotype, clear symbols represent unaffected individuals. The arrow with ‘P’ indicates the proband in each family and ‘us’ indicates unsequenced due to lack of DNA from that individual. The mutation in each family is given next to the family number; a horizontal black line above an individual’s symbol indicates they are heterozygous for the variant as confirmed by either whole exome sequencing or Fluidigm array, and verified by Sanger sequencing. A red dot indicates the individual was underweight (thinness grade 2 or more significant) and ‘?’ indicates that BMI information for that individual is not available. Panels B-D: BMI and height standard deviation score (SDS) charts for the probands of each of the three pedigrees (family 1.III.2, family 2.III.5 and family 3.III.2). Underweight values are shown in red, green dots indicate a significant deflection from previous height measurements and orange dots indicate significant deflection from target height. Normal values, based on data from >70,000 healthy Finnish children, have been previously published [22].

**Figure 3 – Demethylation assay assessing kinetic activity of mutant versus wild type FTO proteins.**

FTO activity is proportional to the concentration present in the reaction. Demethylase activity is likely to be related to the ability of FTO to function as a sensor for cellular metabolism (36). The R316Q mutant is enzymatically dead across all concentrations tested. The A163T and L44V mutants showed
demethylase activity towards methylated-uridine in a dose-dependent manner
but with different affinities. WT – wild-type

Figure 4 – Timing of vaginal opening in wild-type (WT) and FTO<sup>−/−</sup> heterozygous (Het) mice.
Cumulative percentages of mice displaying vaginal opening by postnatal day are shown for WT and FTO<sup>−/−</sup> mice. WT mice n=24, FTO<sup>−/−</sup> n=45; p <0.05 by un-paired t test.

Figure 5 – Mean body weight (g) for wild type (WT) and Fto<sup>−/−</sup> (Het) mice in 7 days prior to vaginal opening
Mean body weight (g) +/- SEM: 11.64 +/- 0.21 in wild-type (n=24) vs 11.45 +/- 0.14 in Fto<sup>−/−</sup> mice (n=45), p=0.467 by un-paired t test. Error bars show SEM for each group each day.
<table>
<thead>
<tr>
<th>Case</th>
<th>Sex</th>
<th>Amino acid alteration</th>
<th>Height SDS at age 4 yrs</th>
<th>Height SDS at age 8/9 yrs</th>
<th>Height SDS at age 18 yrs</th>
<th>ISO-BMI at 18 yrs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.II.1</td>
<td>M</td>
<td>p.Ala163Thr</td>
<td>-</td>
<td>1.1</td>
<td>1.7</td>
<td>16.9</td>
</tr>
<tr>
<td>1.III.2</td>
<td>M</td>
<td>p.Ala163Thr</td>
<td>1.1</td>
<td>0.5</td>
<td>1.1</td>
<td>17.1</td>
</tr>
<tr>
<td>(P)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.III.1</td>
<td>F</td>
<td>p.Ala163Thr</td>
<td>0.9</td>
<td>1.0</td>
<td>1.1</td>
<td>17.3</td>
</tr>
<tr>
<td>1.II.5</td>
<td>M</td>
<td>p.Ala163Thr</td>
<td>-1.0</td>
<td>-1.0</td>
<td>-0.4</td>
<td>-</td>
</tr>
<tr>
<td>(P)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.III.5</td>
<td>M</td>
<td>p.Leu44Val</td>
<td>-0.9</td>
<td>-1.4</td>
<td>-1.5</td>
<td>18.8</td>
</tr>
<tr>
<td>2.III.6</td>
<td>M</td>
<td>p.Leu44Val</td>
<td>-1.1</td>
<td>-1.3</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2.II.2</td>
<td>M</td>
<td>p.Leu44Val</td>
<td>-</td>
<td>-0.8</td>
<td>-0.8</td>
<td>20.5</td>
</tr>
<tr>
<td>2.III.1</td>
<td>M</td>
<td>p.Leu44Val</td>
<td>0</td>
<td>-1.4</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3.II.2</td>
<td>M</td>
<td>p.Leu44Val</td>
<td>-</td>
<td>-1.0</td>
<td>-0.9</td>
<td>18.6</td>
</tr>
<tr>
<td>3.III.2</td>
<td>M</td>
<td>p.Leu44Val</td>
<td>-0.9</td>
<td>-1.1</td>
<td>-1.3</td>
<td>18.7</td>
</tr>
<tr>
<td>(P)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.II.3</td>
<td>M</td>
<td>p.Leu44Val</td>
<td>-0.4</td>
<td>-0.1</td>
<td>22.7</td>
<td></td>
</tr>
<tr>
<td>3.III.3</td>
<td>M</td>
<td>p.Leu44Val</td>
<td>-0.1</td>
<td>0.2</td>
<td>0.5</td>
<td>17.8</td>
</tr>
</tbody>
</table>

Table 1 – Clinical data of probands with FTO variants

Height is expressed in s.d. score (SDS) for national reference data for Finland at 4 years of age and at either 8 years for girls or 9 years for boys. Normal limits: delta HSDS <1.21, distance to target height at 4 yrs <1.76, distance to target height at 8/9 yrs <1.72(22). P – proband.
Figure 1 – Flowchart of WES (whole exome sequencing) filtering strategy to identify candidate genes.
Figure 2 – Pedigrees and auxological data of the families with potentially pathogenic FTO variants
Figure 3 – Demethylation assay assessing kinetic activity of mutant versus wild type FTO proteins.
Figure 4 – Timing of vaginal opening in wild-type (WT) and $FTO^{het}$ heterozygous (Het) mice.
Figure 5 – Mean body weight (g) for wild type (WT) and Fto<sup>+</sup>/− (Het) mice in 7 days prior to vaginal opening.
Supplementary Data

Contributions of function-altering variants in genes implicated in pubertal timing and body mass for self-limited delayed puberty

Sasha R Howard, Leonardo Guasti, Ariel Poliandri, Alessia David, Claudia P Cabrera, Michael R Barnes, Karoliina Wehkalampi, Stephen O’Rahilly, Catherine E Aiken, Anthony P Coll, Marcella Ma, Debra Rimmington, Giles SH Yeo, Leo Dunkel

Table of Contents

Supplementary Methods: Patient details and Genetic Analysis

Supplementary Fig. 1: Auxological data of the family members from Family 1 with potentially pathogenic FTO variant.

Supplementary Fig. 2: Auxological data of the family members from Family 2 with potentially pathogenic FTO variant.

Supplementary Fig. 3: Auxological data of the family members from Family 3 with potentially pathogenic FTO variant.

Supplementary Fig. 4: Surface representation of FTO bound to 3'-methylthymidine.

Supplementary Fig. 5: 3D structure of FTO bound to 3'-methylthymidine and iron (PDB 3flm).

Supplementary Fig. 6: Multiple sequence alignment between human FTO and its orthologues.

Supplementary Fig. 7: Tertiary structure of FTO local to L44 residue

Supplementary Fig. 8: p.A163T sequence and structural analysis.

Supplementary Table 1: Genes involved in energy metabolism and growth pathways implicated in the timing of puberty in the general population from genome wide association studies
Supplementary Methods:

Patient Details
For family members diagnosis of DP was based on PHV occurring 1.5 SD beyond the mean, i.e. age at takeoff exceeding 12.9 and 11.3 yr, or age at PHV exceeding 14.8 and 12.8 yr in males and females. This 1.5SD cutoff has sensitivity of 98% in identifying boys with Tanner stages G2 later than 14.0 years and sensitivity of 97% in identifying girls with Tanner stage B2 later than 13.0 years.

Genetic Analysis
Genetic analysis was performed in 160 individuals from the 67 most extensive families from our cohort with DP. These included 67 probands (male n=57, female n=10), 58 affected family members (male n=36, female n=22) and 35 unaffected family members (male, n=13, female n=22). Whole exome sequencing (WES) was performed on DNA extracted from peripheral blood leukocytes, using a Nimblegen V2 or Agilent V5 platform and Illumina HiSeq 2000 sequencing. The exome sequences were aligned to the UCSC hg19 reference genome. Picard tools and the genome analysis toolkit were used to mark PCR duplicates, realign around indels, recalibrate quality scores and call variants.

Variants were analyzed and filtered for potential causal variants in Ingenuity Variant Analysis (Qiagen) using filters for quality control, predicted functional annotation, minor allele frequency (MAF), and GWAS relevance (Figure 1). Quality control included thresholds for call quality, read depth and upstream pipeline filtering. Predicted functional annotation involved prioritizing nonsense, exonic missense, splice site variants, structural or promoter
changes, or variants deleterious to a microRNA. Filtering by MAF entailed including those variants with minor allele frequency (MAF) <1% in the 1000 Genomes database, the NHLBI exome variant server, and ExAC and gnoMAD databases. GWAS relevance filtering allowed identification of those remaining variants that lay within genes in linkage disequilibrium with 106 GWAS loci associated with AAM\textsuperscript{1}. All genes in linkage disequilibrium with these GWAS AAM loci (using inclusive limits: D’ > 0.8; r\textsuperscript{2}: no limit) were selected using the Broad institute SNAP tool (SNP annotation and proxy search). Linkage disequilibrium data was calculated using Haploview 4.0, based on phased genotype data from the International HapMap Project and the 1000 Genomes Project. A total of 760 genes were selected using this SNAP tool, and ‘GWAS relevance filtering’ allowed identification of those remaining variants that lay within these 760 genes\textsuperscript{2}. Filters for genes implicated in body mass regulation were applied using a biological context filter with pathway analysis. Variants were then filtered for segregation with trait: variants present in $\geq n$-1 affected individuals (where n = number of affected individuals in a given pedigree) and not present in more than one unaffected individual being retained. Family members were screened using conventional Sanger sequencing. Targeted exome sequencing using a Fluidigm array of the remaining candidate gene identified post-filtering was then performed in a further 42 families from our cohort (288 individuals, 178 with DP; male=106, female=69 and 110 controls; male=55, female=58, Figure 1). Variants post targeted re-sequencing were filtered using the same criteria as the WES data: quality
control, predicted functional annotation, minor allele frequency and segregation with trait.

Whole gene rare variant burden testing was performed post sequencing. Fisher’s exact test was used to compare the prevalence of deleterious variants in our cohort with the Finnish population, using the ExAC Browser (Exome Aggregation Consortium (ExAC), Cambridge, MA: http://exac.broadinstitute.org, accessed September 2015). All variants from the ExAC database with minor allele frequency <1%, predicted to be deleterious by Polyphen-2 \(^3\) or SIFT \(^4\), were included in the analysis. A multiple comparison adjustment was applied post hoc using the Benjamini & Hochberg method \(^5\), as detailed in \(^6\). Variants were confirmed via Sanger sequencing.
Supplementary Figures

Family 1 (p.A163T)

Family 1 (II.5)

Supplementary Fig. 1 - Auxological data of the family members from Family 1 with potentially pathogenic FTO variant.
BMI and height standard deviation score (SDS) charts for the family members with FTO variant. Underweight values are shown in red, green dots indicate a significant deflection from previous height measurements and orange dots indicate significant deflection from target height. Normal values, based on data from >70,000 healthy Finnish children, have been previously published

7.
Supplementary Fig. 2 - Auxological data of the family members from Family 2 with potentially pathogenic FTO variant. BMI and height standard deviation score (SDS) charts for the family members with FTO variant. Underweight values are shown in red, green dots indicate a significant deflection from previous height measurements and orange dots indicate significant deflection from target height. Normal values, based on data from >70,000 healthy Finnish children, have been previously published. 
Supplementary Fig. 3 - Auxological data of the family members from Family 3 with potentially pathogenic FTO variant.
BMI and height standard deviation score (SDS) charts for the family members with FTO variant. Underweight values are shown in red, green dots indicate a significant deflection from previous height measurements and orange dots indicate significant deflection from target height. Normal values, based on data from >70,000 healthy Finnish children, have been previously published. 

Family 3 (p.L44V)
Supplementary Fig. 4 - Surface representation of FTO (here presented in grey) bound to 3-methylthymidine (in yellow). Leucine 44 is presented in blue.

Supplementary Fig. 5 - 3D structure of FTO bound to 3-methylthymidine and iron (PDB 3fim)
Supplementary Fig. 6 – Multiple sequence alignment between human FTO and its orthologues.

Orthologous sequences from other species were retrieved from Ensembl. Species are classified in “Placental mammals”, “Birds & Reptiles” and “Fish” according to Ensembl classification. For each species, FTO Uniprot accession number (in blue) and entry name (in black) are presented at the beginning of the row. The position of deleterious mutations on the human FTO sequence is indicated by a red arrow. L44 and other surrounding residues part of the same alpha helix form a motif, which is highly conserved across placental mammals but not in reptiles, birds and fish.

Supplementary Fig. 7 – Tertiary structure of FTO local to L44 residue. Residue L44 is presented in blue and other residues located on the same alpha helix and conserved across placental mammals are presented in green. FTO structure is shown as cartoon on the left and as surface on the right. 3′-methylthymidine is presented as yellow spheres in the cartoon representation.
Supplementary Fig. 8 - p.A163T sequence and structural analysis.
Panel A: a multiple sequence alignment shows that alanine at position 163 is not conserved. Panel B: the position of Ala163 (indicated by a red arrow) is presented in relation to FTO secondary structure (H, alpha helix; C, residues that are not part of an alpha helix or a beta strand). Residues predicted to be disordered are indicated with an asterisk. Panel C: FTO tertiary structure. Ala163 (presented in red) is at the end of FTO alpha helix H4 and is not predicted to disrupt FTO function. FTO structure is visualized using visualisation program (http://www.pymol.org/).

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>FTO</th>
<th>GNPDA2</th>
<th>LEKR1</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEC16B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TMEM18</td>
<td></td>
<td>LIN28B</td>
<td>LEPR-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LEPROT</td>
</tr>
<tr>
<td>NEGR1</td>
<td>SIX6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNN13K</td>
<td></td>
<td>CENPW-</td>
<td>NCO7</td>
</tr>
</tbody>
</table>

Supplementary Table 1 – Genes involved in energy metabolism and growth pathways implicated in the timing of puberty in the general population from genome wide association studies (adapted from Perry et al1).
References


