Genomic Analyses Identify Rare Variants in Genes Associated with Age at Menarche in Patients Affected with Anorexia Nervosa and Support a Role for Puberty Timing in Anorexia Nervosa Risk

Nicolas Lebrun¹, Philibert Duriez², Julia Clarke³, Philip Gorwood¹,², Nicolas Ramoz¹, Thierry Bienvenu¹,⁴*  

¹Université de Paris, Institute of Psychiatry and Neuroscience of Paris (IPNP), INSERM UMR1266, « Genetic vulnerability to addictive and psychiatric disorders » team, Paris, France  
²GHU Paris Psychiatrie et Neurosciences, CMME, Hôpital Sainte-Anne, F-75014 Paris, France  
³Service de Psychiatrie de l’enfant et de l’adolescent, Hôpital Universitaire Robert Debré, Paris, France  
⁴Laboratoire de Génétique et de Biologie Moléculaires, Hôpital Cochin, Assistance Publique- Hôpitaux de Paris-Centre Université de Paris, France  

*Corresponding Author: Dr. Thierry Bienvenu, Institut de Psychiatrie et de Neurosciences, 102 rue de la Santé, 75014 Paris, France, E-mail: thierry.bienvenu@inserm.fr  

Received: 08 August 2020; Accepted: 10 August 2020; Published: 03 September 2020  

Abstract  
Puberty is a critical risk period for eating disorders, especially in girls. Previous reports showed that the association between puberty and eating disorders is mainly due to genetic factors, but the nature of these factors is unclear. To identify these genetic factors, we carried out exome analysis in 10 girls with anorexia nervosa compared to 10 unaffected women controls and identified low-frequency variants in genes among a list of 185 puberty-associated genes. After filtering data, 9 controls and 9 affected girls had at least one potential pathogenic rare (PPR) variant. In controls, 35 PPR variants were identified located (in 31 genes) whereas 52 PPR variants (in 38 genes) were identified in patients with AN. Among the genes identified only in the AN patients, we found a significant enrichment as compared to the general population and the control sample, in variants previously identified as associated with age at menarche (TNRC6A, LAMB2 and FAAH2). Two AN patients presented at least one rare missense variant in these genes. Moreover, nine patients with AN carried out missense/frameshift variants in puberty-associated genes previously identified by GWAS studies. Our results suggest that rare variants in genes...
involved in the timing of puberty such as TNRC6A, LAMB2 and FAAH2 may predispose to anorexia nervosa susceptibility during puberty.

**Keywords:** Anorexia nervosa; Puberty; Exome; TNRC6A; LAMB2; FAAH2

1. 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. The time of human puberty is highly variable, and sexually dimorphic. Known influences on the timing of this event in mammals include the light/dark cycle, leptin levels, and the increased expression of neurokinin B, kisspeptin, and their receptors, NK3R and KISS1R [1-3]. Moreover, over 20 genes carrying rare pathogenic variants have been identified in pubertal disorders, many of them encode critical components of the hypothalamic-pituitary-gonadal (HPG) axis [4]. Recent genome-wide association studies (GWAS) have also identified more than 100 candidate genes at loci associated with age at menarche (AAM) or voice breaking in males [5, 6].

Disordered pubertal timing affects up to 5% of adolescents and is associated with adverse health and psychosocial outcomes. Interestingly, early pubertal timing increases risk for eating disorders. For example, adolescents reporting a younger age at menarche are more likely to report bulimic symptomatology, and women with bulimia nervosa (BN) report an earlier age of menarche compared to women with anorexia nervosa (AN) or healthy controls. A population based study of adolescents also indicated that the median ages of onset for AN, BN and subthreshold eating disorders ranged between 12.3 and 12.6 years-old, which are quite similar to the average age of menarche in the general population (12.4 years-old) [7]. However, the mechanism by which pubertal timing increases risk for eating disorders is unclear. One of the major hypothesis is that the association between pubertal timing and eating disorders could be linked to common biological phenomena. Twin studies showed that eating disorder symptomatology is heritable. Genetic effects for eating disorders are also almost not present in prepubertal girls and increase markedly to approximately 50%-60% in advancing-pubertal girls and young adults [8-12]. Taken together, these findings raise the possibility that the genes involved in age of menarche play a role in the genetic influence on eating disorders. Up to now, only few studies showed that genetic factors that predispose girls to early menarche also increase eating disorders [13].

To better identify these genetic factors predisposing to eating disorders, we carried out an exome analysis in 10 girls with anorexia nervosa and in 10 unaffected women as controls, and searched for low-frequency variants among a list of 185 puberty-associated genes.
2. Participants

Patients with AN were recruited at the Clinique des Maladies Mentales et de l’Encéphale (CMME) (Hôpital Sainte-Anne, Paris, France). Diagnosis was made according to Diagnostic and Statistical Manual of Mental Disorders 5 (DSM5) scores. Individuals that consented to participation in the study had blood drawn for the collection of DNA and received a diagnostic interview administered by a trained research team member. Consensus diagnosis was obtained by agreement of clinicians specializing in AN. Eight individuals were considered as having a restrictive-subtype anorexia nervosa and two with the bulimia subtype. Three of these eating disorders (ED) patients are prepubertal at the time of the examination. Exclusion criteria included other types of ED than AN, confounding psychiatric diagnoses, for example, psychotic disorders, medical or neurological conditions causing weight loss. This cohort of patients consisted of 10 female individuals (13 to 47 years-old). Control group consists of 10 unaffected women. These controls were used to determine the frequency of rare variant in the different candidate genes.

3. Methods

3.1 Whole exome analysis

All individuals provided a signed informed written consent. Peripheral blood samples were collected from patients and controls in EDTA tubes. Genomic DNA was extracted from blood samples by a standard procedure using the Maxwell system (Promega, Madison, WI, USA) and the LEV DNA BLOOD AS 1290 kit (Promega, Madison, WI, USA). Whole exome capture, next-generation sequencing and data analysis were carried out by Cochin Institute and Paris Descartes Platform, respectively. Briefly, libraries were prepared from 3 μg genomic DNA extracted from whole blood using an optimized SureSelect Human Exome kit (Agilent) following the manufacturer's instructions. Captured, purified and clonally amplified libraries targeting the exome were then sequenced on a Nextseq 500 instrument (Illumina) according to the manufacturer's recommendations. Obtained sequence reads were aligned to the human genome (hg19) using BWA software. Downstream processing was carried out with the Genome analysis toolkit (GATK), SAMtools and Picard Tools (http://picard.sourceforge.net). Single-nucleotide variants and indels were subsequently called by the SAMtools suite (mpileup, bcftools, vcfutil). All calls with a read coverage ≤20× and a Phred-scaled SNP quality of ≤20 were filtered out. We only searched for variants in a list of 185 puberty-associated genes (23 genes with known mutations in patients diagnosed with Kallmann syndrome, normosmic hypogonadotrophic hypogonadism or precocious puberty; 145 candidate genes near genetic variants identified in large-scale GWAS studies for age at menarche in women; 7 genes associated with age at menarche; and 10 literature-based genes) [14-17]. Variants were annotated with an in-house Paris Descartes bioinformatics platform pipeline based on the Ensembl database (release 67). Four consecutive filters were applied to narrow down the variants: 1- exonic or located in consensus splice junction (+/- 2 bp); 2- non-synonymous; 3- located in a gene identified by an NM_accession number (NM_); and 4- minor allele frequency (MAF)<1% (in gnomAD database https://gnomad.broadinstitute.org/). In all cases, sequences variant in a gene are numbered starting from the first
base of the ATG codon, numbering based on the reference sequence. Description of the sequence (Human Genome Variation Society, http://www.hgvs.org/mutnomen/recs.html) was done with the assistance of Alamut Visual software version 2.4.2 (Interactive Biosoftware, Rouen, France).

3.2 Sanger sequencing
DNA variants of interest were confirmed by PCR amplification and direct sequencing. DNA primer sequences and PCR conditions are available upon request to the corresponding author. Sequencing reactions were carried out with the BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems, Courtaboeuf, France) and loaded on the ABI 3130XL genetic analyser (Applied Biosystems, Courtaboeuf, France). Results were visualized using the Sequencher software (Gene codes, Ann Arbor, United States).

3.3 In silico analysis of missense variants
To find the influence of each missense variant on its protein structure that may have an important role in disease susceptibility, in silico analysis was performed. For each variant, bioinformatics tools (SIFT, MutationTaster, Polyphen-2) (https://www.interactive-biosoftware.com/alamut-visual/) were used to identify non-synonymous variant as possibly or probably damaging. These prediction tools are based on the degree of conservation of amino acid residues in sequence alignments derived from closely related sequences.

4. Results
4.1 Rare variants identified in AN patients
We performed exome sequencing for ten affected girls and 10 unaffected women as controls. In total 122,380 to 140,758 single nucleotide variants and 16,162 to 19,483 small deletions/insertions (1-10 bp) were identified across the exomes [18]. To identify genetic factors involved in anorexia and age of menarche, we search for low-frequency variants in genes among a list of 185 puberty-associated genes (Table 1). After filtering our data, it resulted in 9 affected individuals (9/10) at least one variant (no variant in patient NR279) (Table 2). In these patients, the number of variants varied from 2 (NR227) to 11(NR224) for each affected AN individual. In total, 52 variants were identified including 11 variants not previously described in all genomic databases (Table 2). One variant was a frameshift variant (VGLL3), two were in frame deletions (SCRI, GALNT10) and all other variants were missense variants (Table 2). In total, variants were identified in 38 of the tested genes.
<table>
<thead>
<tr>
<th>Gene stable ID</th>
<th>Chromosome</th>
<th>Gene start (bp)</th>
<th>Gene end (bp)</th>
<th>Karyotype</th>
<th>band</th>
<th>Gene name</th>
<th>RefSeq match transcript</th>
<th>RefSeq mRNA ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENSG00000176473</td>
<td>19</td>
<td>40191426</td>
<td>40215575</td>
<td>q13.2</td>
<td>MKL2/MAP3K10</td>
<td>NM_002446.3</td>
<td>Hou et al. 2017</td>
<td></td>
</tr>
</tbody>
</table>

**Table 1:** List of the puberty-associated genes selected in the analysis of exome.
Table 2: Rare variants (<1% in NM_; <0.01% in gnomAD (in grey)) identified in 185 genes involved in puberty-associated genes in 10 patients affected by anorexia nervosa (2 bulimia type NR86 and NR180; 8 restrictive type).
In contrast, after filtering data, 9 healthy women (9/10) had at least one variant (no variant was observed in one control) (Table 3). In these subjects, the number of variants varied from 2 to 7 for each control individual. In total, 35 variants were identified including 6 variants not previously described in all genomic databases (Table 3). One variant was a frameshift variant (BDNF), and all other variants were missense variants (Table 3). In total, variants were identified in 31 candidate genes.

Table 3: Rare variants (<1% NM_; <0.01% in gnomAD (in grey)) identified in genes involved in puberty-associated genes in 10 unaffected women (in 9/10 women).
4.2 Prediction of pathogenicity of variants
Among the 52 puberty-related (PR) missense variants, 12 were considered as deleterious and probably damaging by all bioinformatic softwares. Among these potential pathogenic variants, two (NPTXR, PTPRN2) were not described in all genomic databases, and 4 (FAAH2, LAMB2, GHRHR, SCRIB) were considered as extremely rare (<0.1%). Moreover, 18 variants were considered as deleterious or probably damaging by one bioinformatic software (GHR, DST (x3), TNRC6A (x2), PROKR2, LRP1B, CSMD1, NPHP3, WDR11, VDR, TENM2, GALNT10, WWP2, IGF2R, RBM6, UBA7).

In contrast, in controls, among the 35 PR variants, only 6 were considered as deleterious and probably damaging by all bioinformatic softwares (ALOX15B;EIF4G1; GALNT10; TCF7; WSCD1; CADPS2). Among these potential pathogenic variants, only one (TCF7) was not described in all genomic databases, and all were considered as extremely rare (<0.1%). Moreover, 8 variants were considered as deleterious or probably damaging by one bioinformatic software.

4.3 Variants in candidate genes with significant sex-biases expression
The fact that AN patients were predominantly female (with nine females affected for each male case observed) and that onset of its condition is classically around puberty highly suggested that genes with sex biases expression were potential candidates to explain AN vulnerability. Recently, O’Brien and colleagues (http://dx.doi.org/10.1101/483636) identified 2726 autosomal transcripts with significant sex biases expression. Among these transcripts, we found 8 of our 38 candidate genes (CSMD1, TNRC6A, LRP1B, NPHP3, WDR11, PEX2, SCRIB, RBM6). Moreover, using the SAGD database (http://bioinfo.life.hust.edu.cn/SAGD), we found that 18 of our candidate genes were found to be sex-associated genes which varied expression between females and males in Human (PTPRD, CSMD1, VDR, GHRHR, MKNR3, HNF4A, ALOX15B, TENM2, PTPRK, SIX6, NPTXR, WSCD1, GPRC5B, PROKR2, LAMB2, UBA7, BEGAIN, GALNT10). Interestingly, BEGAIN, GALNT10, GHRHR, GPRC5B, HNF4A, LAMB2, MAGEL2, PTPRK, SIX6, UBA7, and VDR showed a pituitary-specific sex-biased expression [19].

5. Discussion
Puberty is a critical risk period for psychiatric disorders and more specifically for eating disorders (i.e., anorexia nervosa, bulimia nervosa) and their symptoms [7]. ED incidence increases across the pubertal period and becomes female predominant, and genetic influences on eating disorders significantly increase [20]. Ten years ago, a twin study showed that genetic factors might be important for pubertal risk for eating disorders in girls [21]. This study showed that there was essentially no genetic influence on eating disorder symptoms in pre-adolescent female twins, but significant genetic effects in twins during late adolescence. This conclusion was confirmed by two other twin studies [13, 22]. Moreover, investigations of shared genetic risk factors between eating disorder symptoms (e.g.,
overall eating disorder symptoms, dieting) and age at menarche or early pubertal timing showed that genetic risk factors (rather than environmental influences) entirely account for the co-occurrence of early menarche and eating disorder behaviors in girls [13].

To identify genetic factors that appear to become “activated” during the pubertal period increasing the phenotypic risk for AN in girls, we searched for rare variants in genes associated with age at menarche in 10 women affected by anorexia nervosa.

We identified an excess of rare variants in genes associated with age at menarche in AN patients as compared to controls. When its number is corrected by the total number of variant identified in the exome, the difference was not significant (Table 4). However, interestingly, we identified a significant enrichment of extremely rare variants and variants in genes previously identified associated with age of menarche (i.e. TNC6CA, LAMB2 and FAAH2) (Table 4) [15]. Indeed, in 2015, Lunetta and colleagues found that five missense variants (in ALMS1, LAMB2, TNRC6A, TACR3, PRKAG1) and two X-chromosome loci (IGSF1, FAAH2) were associated with age of menarche [15]. In the present study, one AN patient carried three rare variants in these genes (p.G1225D in LAMB2, p.G1602S in TNRC6A, and p.R274C in FAAH2) and another AN patient two rare variants in TNRC6A (p.Y1681C, p.S1927C). TNRC6A encodes a trinucleotide repeat-containing gene 6A protein, a member of the trinucleotide repeat containing 6 protein family. It functions in post-transcriptional gene silencing through the RNA interference (RNAi) and microRNA pathways [23]. It associates with messenger RNAs and Argonaute proteins in cytoplasmic bodies known as GW-bodies or P-bodies [24]. Variants identified in our two AN patients (NR15, NR224) were located in the C-terminal part of the protein (amino acids 1603-1622, PABPC1-interacting motif-2), known to be important for translational repression, and may impair RNAi and microRNA-induced gene silencing affecting metabolic pathways [25]. LAMB2 encodes the Laminin subunit beta-2 protein, a member of the family of extracellular matrix glycoproteins, major non collagenous constituent of basement membranes implicated in a wide variety of biological processes including cell adhesion, differentiation, migration, signaling, and neurite outgrowth. Recently, compound heterozygous missense variants were identified in patients with congenital nephrotic syndrome, ocular abnormalities and neurodevelopmental delay [26]. Interestingly, Laminin b2 is highly expressed during pituitary development and may explain the abnormal parenchyma of the anterior pituitary gland in this patient. Taking into account the role of the hypothalamus-pituitary-adrenal axis in eating disorders, variant in LAMB2 may play a role in the biological vulnerability to ED. Finally, FAAH2 encodes the Fatty acid amide hydrolase 2 (EC 3.5.1.99, oleamide hydrolase 2, anandamide amidohydrolase 2), a member of the serine hydrolase family of enzymes. Fatty acid amide hydrolase 2 degrades endocannabinoids and defects in this enzyme have been associated with neurologic and psychiatric disorders [27]. Patient carrying the missense variant p.A458S in this gene presented hypotonia, autistic features, seizures, but also anxiety disorder, suicidal ideas and feeding difficulties. Interestingly, we previously identified a rare variant (p.H312N) at the hemizygous state in a male patient with AN [28].
Table 4: Comparison between the frequency of variants (taking into account their frequency in gnomAD database) in the cohort of AN patients and of control women.* genes associated with age of menarche (n=7) (15), ** genes involved in rare Mendelian disorders of puberty (n=23) [29], *** genes associated with age of menarche (n=145) [29], $ Literature-based genes [16-17, 30-32] (n=10).

In the present study we identified rare variants in genes initially associated with age of menarche by GWAS analysis and confirmed by other approaches. This is the case for MKRN3 (NR227) and VGLL3 (NR224). Up to now, one of the most relevant gene underlying central precocious puberty (CPP) is the makorin ring finger protein 3 gene (MKRN3) [33]. MKRN3 is a single-exon, maternally imprinted gene, which is expressed only from the paternal allele. Consequently, only paternally-inherited loss-of-function mutations in MKRN3 cause CPP. Interestingly, MKRN3 is expressed in the mouse and human hypothalamus, a structure containing integrative systems that support physiological processes such as food intake and energy expenditure [34]. VGLL3 is a transcriptional co-factor that binds TEAD family transcription factors to regulate events ranging from wing development in fly, to muscle fibre composition and immune function in mice. Interestingly, Vgll3 play a role in the sex-biased disorders such as autoimmune diseases and may participate to the sex-biased rate of anorexia nervosa. We hypothesize that these genes account for, at least part of, the underlying shared genetic factors between age at menarche and eating disorders [13].

We also identified 29 GWAS puberty-associated genes via coding variation, without significant enrichment as compared to general population and our control sample including 10 unaffected women. The majority of these genes...
are localized in candidate regions for age of menarche and puberty, but were not clearly involved in puberty. For example, this is the case for the rare variant in PTPRN2 (NR224). PTPRN2 is also known la-2beta and was identify as a candidate associated with pediatric obesity [35, 36].

6. Conclusions
To conclude, although pubertal timing increases risk for eating disorders, the nature of the association between pubertal timing and eating disorders is unclear [37-39]. Puberty involves many physical, psychological, and emotional changes all of which likely play an important and complex role in the association between pubertal timing and eating disorders. The genes involved in puberty timing likely interact with aspects of the environment to increase vulnerability to develop ED at puberty, and previous reports have shown that the association between puberty and eating disorders is due to genetic factors [13]. Our results suggest that rare variants in genes involved in puberty timing such as TNRC6A, LAMB2 and FAAH2 may contribute to the vulnerability to AN during puberty and support causal links with anorexia nervosa susceptibility.

Disclosure and Conflict of Interest
The authors report no conflict of interest.

Funding Source
This study was supported by Institut National de la Recherche Médicale INSERM and by Fondation Maladies Rares.

Acknowledgments
We thank the patients for their enthusiastic participation and all physicians from the different medical and psychiatric centres.

Credit Author Statement
Thierry Bienvenu and Nicolas Ramoz had full access to all data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Concept and Design: Thierry Bienvenu, Nicolas Ramoz; Acquisition, analysis and interpretation of the data: Thierry Bienvenu, Nicolas Lebrun, Philip Gorwood, Nicolas Ramoz; Drafting of the manuscript: Thierry Bienvenu, Philip Gorwood, Nicolas Ramoz; Patient’s inclusion: Julia Clarke, Philibert Duriez; Technical analysis: Nicolas Lebrun.

Compliance with Ethical Standards
Ethical approval All procedures performed in this study were in accordance with the ethical standards of our national research committee (Comité Protection Personnes Paris Pitié Salpêtrière N°33-03) and with the 1964
Helsinki declaration and its later amendments. Informed written consent was obtained from all patients included in the study.

References

Citation: Nicolas Lebrun, Philibert Duriez, Julia Clarke, Philip Gorwood, Nicolas Ramoz, Thierry Bienvenu, Genomic Analyses Identify Rare Variants in Genes Associated with Age at Menarche in Patients Affected with Anorexia Nervosa and Support a Role for Puberty Timing in Anorexia Nervosa Risk. Journal of Psychiatry and Psychiatric Disorders 4 (2020): 293-306.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license 4.0