

## ORIGINAL ARTICLE

# A GWAS follow-up of obesity-related SNPs in SYPL2 reveals sex-specific association with hip circumference

J. de Toro-Martín<sup>1,2</sup>, F. Guénard<sup>1,2</sup>, A. Tchernof<sup>2,3</sup>, Y. Deshaies<sup>3,4</sup>, L. Pérusse<sup>1,5</sup>, S. Biron<sup>6</sup>, O. Lescelleur<sup>6</sup>, L. Biertho<sup>6</sup>, S. Marceau<sup>6</sup> and M.-C. Vohl<sup>1,2</sup>

<sup>1</sup>Institute of Nutrition and Functional Foods (INAF), Laval University, Quebec, Quebec, Canada; <sup>2</sup>School of Nutrition, Laval University, Quebec, Quebec, Canada; <sup>3</sup>Quebec Heart and Lung Institute Research Center, Quebec, Quebec, Canada; <sup>4</sup>Department of Medicine, Laval University, Quebec, Quebec, Canada; <sup>5</sup>Department of Kinesiology, Laval University, Quebec, Quebec, Canada; <sup>6</sup>Department of Surgery, Laval University, Quebec, Quebec, Canada;

Received 16 June 2016; revised 7 September 2016; accepted 15 September 2016

Address for correspondence: M-C Vohl, PhD, Institute of Nutrition and Functional Foods (INAF), Laval University, 2440 Hochelaga Blvd, Quebec, Quebec, Canada, G1V 0A6. E-mail: marie-claude.vohl@fsaa.ulaval.ca

## Summary

### Objective

A novel single-nucleotide polymorphism (SNP) associated with morbid obesity was recently identified by exome sequencing. The purpose of this study was to follow up this low-frequency coding SNP located within the *SYPL2* locus and associated with body mass index in order to reveal novel associations with obesity-related traits.

### Methods

The body mass index-associated SNP (rs62623713 A>G [chr1:109476817/hg19]) and two tagging SNPs within the *SYPL2* locus, rs9661614 T>C (chr1:109479215) and rs485660 G>A (chr1:109480810), were genotyped in the obesity ( $n=3,017$ ) and the infogene ( $n=676$ ) cohorts, which were further combined, leading to a larger cohort of 3,693 individuals. Association testing was performed by general linear models in the obesity cohort and validated by joint analysis in the combined cohort.

### Results

rs9661614 and rs485660 were significantly associated with hip circumference (HC) in the obesity cohort, with heterozygotes exhibiting a significantly lower HC. These results were validated by joint analysis for rs9661614 (false discovery rate [FDR]-corrected  $P=7.5 \times 10^{-4}$ ) and, to a lesser extent, for rs485660 (FDR corrected  $P=3.9 \times 10^{-2}$ ). The association with HC remained significant for rs9661614 when tested independently in women (FDR-corrected  $P=1.7 \times 10^{-2}$ ), but not for rs485660 (FDR-corrected  $P=0.2$ ). Both associations were absent in men.

### Conclusions

This study reveals strong evidence for a novel association between rs9661614 (T>C) and HC in women, which likely reflects a preferential association of *SYPL2* to a gynoid profile of fat distribution. The study findings support a clinical significance of *SYPL2* worth considering when assessing risk factors associated with obesity.

**Keywords:** Hip circumference, obesity, replication study, *SYPL2*.

## Introduction

Genome-wide association studies (GWAS) have identified susceptibility loci for obesity (1). These association studies usually focus on discovering novel genetic associations with body weight or body mass index (BMI) by using large and heterogeneous populations, as well as vast amounts of genetic markers. Findings from these works have been successfully replicated in a number of cases,

leading to the identification of consolidated BMI-associated genes, such as *FTO* or *MC4R* (2,3), currently being the object of active research to elucidate their actual functional relevance (4,5). Moreover, GWAS meta-analysis has spread as an effective method to test the consistency of genome-wide associations and to discover new obesity risk variants across different ancestry populations (6,7). However, for most of the loci identified as BMI associated, the causal genes and pathways

involved, as well as their physiological role, are yet poorly known. In this sense, whole exome sequencing has emerged as a useful tool to focus on genetic variability at coding regions, which allows the identification of rare variants with higher effect on BMI (8,9). Nevertheless, only a few studies have been focused on following up these obesity-associated single-nucleotide polymorphisms (SNPs) with additional analyses to test their impact on phenotype traits other than body weight or BMI (10,11).

Another critical point of association studies not always taken into consideration is the way these genetic variants impact phenotype traits depending on the sex of individuals. This issue becomes particularly relevant when genetic associations are related to alterations of the metabolic profile or to certain pathologies notably influenced by sex, e.g. type 2 diabetes or cardiovascular diseases (12). A recent meta-analysis revealed a large variability among genetic associations with obesity-related traits depending on sex, such as the waist-to-hip ratio (WHR) (13). Recently, a low-frequency coding variant identified by exome sequencing within the synaptophysin-like 2 gene (*SYPL2*), also known as mitsugumin 29 (*MG29*), a gene coding for a protein mainly involved in calcium homeostasis in skeletal muscle (14), was found to be significantly associated with morbid obesity (9). Importantly, this polymorphism showed a relatively high penetrance on BMI in a study cohort showing an overrepresentation of women with obesity (9).

In the present two-stage study, we followed up this BMI-associated SNP of the *SYPL2* locus, in order to reveal potential novel associations with other metabolic and anthropometric traits related to obesity, which were further tested independently in men and women to determine a potential sex-specific effect.

## Methods

### Study cohorts

The obesity cohort, used in the first stage of this study for discovery purposes, was composed strictly of individuals with obesity ( $\text{BMI} \geq 30 \text{ kg m}^{-2}$ ) and severe obesity ( $\text{BMI} \geq 35 \text{ kg m}^{-2}$ ). A total of 3,017 Caucasian patients (942 men and 2,075 women) undergoing biliopancreatic diversion with duodenal switch at the Quebec Heart and Lung Institute (Quebec City, Quebec, Canada) formed this cohort. The surgical protocol, blood sample collection and the standardized procedures to measure anthropometric and metabolic parameters are described elsewhere (15). The infogene cohort was composed of 676 Caucasian subjects with or without obesity (277 men and 399 women) recruited between 2004 and 2006

through radio and newspaper advertisements, as well as by a newsletter shared through the Laval University network for previous studies (16,17). The collection of anthropometric and metabolic measurements of the infogene subjects has been previously described (16). The obesity cohort was combined with the infogene cohort in the second stage of this study, resulting in a larger and heterogeneous cohort, composed of 3,693 subjects (1,219 men and 2,474 women) with obesity and severe obesity, as well as individuals without obesity, and used for validation purposes. This study was approved by the Laval University and Quebec Heart and Lung Institute Ethics Committees and was conducted in accordance with the 1964 Helsinki Declaration.

### Single-nucleotide polymorphism genotyping

The first SNP selected for genotyping was the rare variant previously associated with BMI and located at *SYPL2* exon 4 (rs62623713 A>G [chr1:109476817/hg19]) (9). Because exome sequencing is not able to identify common SNPs located within untranscribed regions, additional tagging SNPs were added to the association study in order to cover most of the genetic variability within the *SYPL2* locus. Selection of additional tagging SNPs within the *SYPL2* locus and surrounding regions (2.5 kb upstream and downstream) was carried using the tagger selection algorithm of the Haploview software (Massachusetts Institute of Technology, Cambridge, MA, USA) (18) and considering the CEU panel (Utah residents with Northern and Western European ancestry) of the latest release of HapMap (release 28, Phase II+III data). Using this tagging SNP selection, we identified rs9661614 T>C (chr1:109479215; intron variant) and rs485660 G>A (chr1:109480810; 3'-UTR variant) located in the vicinity of rs62623713 (2.4 and 4.0 kb, respectively). These two additional tagging SNPs covered 100% of *SYPL2* genetic variability considering common genetic variants with minor allele frequencies higher than 5% and high linkage disequilibrium (LD;  $r^2 > 0.8$ ). Selected SNPs were genotyped in both the obesity and the infogene cohorts using TaqMan probes (Applied Biosystems, Foster, CA, USA). Genomic DNA was extracted from the blood buffy coat using the GenElute Blood Genomic DNA kit (Sigma, St. Louis, MO, USA). Genotypes were determined using the 7500 Fast Real-Time PCR System (Applied Biosystems), and they were analysed using the high-throughput array technology QuantStudio 12K Flex System, coupled with Taqman OpenArray Technology (Life Technologies, Carlsbad, CA, USA). Haplotype reconstruction and individual diplotype assignments were inferred from genotype data using PLINK v1.07 (PLINK, Boston, MA, USA) (19) and PHASE v2.1.1 (20) software (University of

Washington, Seattle, WA, USA), with default parameters and using the CEU panel of the 1000 Genomes Project (Phase 3) as the reference population.

### Statistical analyses

A two-stage association study was carried out. Statistical analyses were first performed in the obesity cohort, which was randomly subdivided into two smaller sub-cohorts (discovery and replication) to test for associations separately, and formed by 1,513 (472 men and 1,041 women) and 1,511 (471 men and 1,040 women) subjects, respectively. In order not to be too restrictive at this exploratory stage, a nominal  $P \leq 0.05$  found in both discovery and replication sub-cohorts was used to determine an association to be further tested in the entire obesity cohort. Significant associations were further validated by means of a joint analysis in the combined cohort composed of the obesity and the infogene cohorts. False discovery rate-corrected  $P$  (FDR-corrected  $P \leq 0.05$ ) was applied for multiple-testing correction when testing for associations in the obesity and the combined cohorts.

Association tests were performed using the analysis of variance (general linear models, type III sum of squares) under an additive model of inheritance, and adjusted for the effects of age, sex and BMI. Genotype by sex ( $G \times S$ ) and genotype by BMI ( $G \times B$ ) interaction terms were added into separate models one at a time. Pairwise comparisons among genotype groups were performed using least square means, and statistically significant differences were determined with Bonferroni adjusted  $P$ -values ( $P_{\text{bon}} \leq 0.01$ ). Quantitative anthropometric and metabolic traits tested for associations were waist (WC) and hip circumference (HC), WHR, triglycerides (TG), HDL-cholesterol (HDL-C), LDL-cholesterol (LDL-C) and total cholesterol, total cholesterol to HDL-C ratio, fasting glucose and blood pressure (systolic and diastolic). Variables that were non-normally distributed were transformed to approximate a normal distribution (inverse transformed: TG and HDL-C; log<sub>10</sub> transformed: fasting glucose and total cholesterol to HDL-C ratio). Diplotype-based association tests were performed using diplotypes composed of SNPs showing statistically significant associations independently. The analysis of variance adjusted by the same variables was also applied for diplotype-based tests.

The proportion of phenotypic variance explained by the genotype was calculated as the ratio of the type III sum of squares because of the SNP effect to the sum of squares of the model. With the statistical significance set to  $\alpha = 0.05$  and  $\beta = 0.10$ , the statistical power to detect significant associations was higher than 99% in both the obesity and the combined cohorts. Statistical analyses were performed using SAS software version 9.3 (SAS

Institute, Cary, NC, USA). Statistical power analyses were performed using G\*Power (version 3.1.9.2) (Heinrich Heine-Universität, Düsseldorf, Germany) (21).

## Results

### Anthropometric, metabolic and genetic characteristics of the study cohorts

The obesity cohort was composed of 3,017 subjects with obesity and severe obesity (942 men and 2,075 women) with a mean BMI = 50.5 kg m<sup>-2</sup>, which almost doubled to that observed in the infogene cohort (BMI = 27.8 kg m<sup>-2</sup>), which ranged from 16 to 40 kg m<sup>-2</sup>. The infogene cohort, composed of 676 subjects with and without obesity (277 men and 399 women), was combined with the obesity cohort, resulting in a larger and heterogeneous cohort integrated by 3,693 subjects (1,219 men and 2,474 women), with a mean BMI = 46.3 kg m<sup>-2</sup>, and ranging from 16 to 100 kg m<sup>-2</sup>. Mean values of all the anthropometric and metabolic variables tested in association studies are depicted in Table 1. All the genotyped SNPs were in Hardy–Weinberg equilibrium (Table 2). The rare variant rs62623713 showed a MAF < 5%, while the other two SNPs showed a MAF > 20% (Table 2). rs9661614 showed moderate LD with rs62623713 ( $r^2 = 0.23$ ) and relatively high LD with rs485660 ( $r^2 = 0.62$ ).

### Two SYPL2 polymorphisms are significantly associated with hip circumference in the obesity cohort

The sex-based randomization of the obesity cohort resulted in two smaller sub-cohorts with the same proportion of men (31.2%) and women (68.8%) consisting of 1,513 and 1,511 subjects, respectively. Association tests performed independently in these two obesity sub-cohorts (adjusted for age, sex and BMI) revealed several nominal associations ( $P < 0.05$ ) between SYPL2 SNPs and quantitative phenotype traits (data not shown), but only two SNPs showed a significant association with any of the phenotype traits analysed in both the discovery (d) and the replication (r) obesity sub-cohorts, namely, rs9661614 ( $P_d = 7.3 \times 10^{-4}$ ;  $P_r = 5.3 \times 10^{-3}$ ) and rs485660 ( $P_d = 2.2 \times 10^{-2}$ ;  $P_r = 4.6 \times 10^{-2}$ ), which were significantly associated with HC. Although rs62623713 also showed a significant association with HC in the discovery sub-cohort ( $P_d = 4.6 \times 10^{-2}$ ), it was lost in the replication sub-cohort ( $P_r = 0.42$ ). The validation of these associations was performed in the entire obesity cohort. For multiple testing purposes, all the phenotype traits and SNPs were included in the model. Results showed that both SNPs that were significantly associated with

**Table 1** Metabolic and anthropometric characteristics of the study populations

	Obesity	Infogene	Obesity + Infogene
Number of subjects (men/women)	3,017 (942/2,075)	676 (277/399)	3,693 (1,219/2,474)
Age	43.7 ± 10.9	37.9 ± 11.3	42.6 ± 11.2
Anthropometric profile			
Weight (kg)	138.6 ± 27.8	78.2 ± 18.0	127.6 ± 35.2
Height (cm)	165.5 ± 9.5	167.8 ± 9.4	165.9 ± 9.5
Body mass index (kg m <sup>-2</sup> )	50.5 ± 8.3	27.8 ± 5.7	46.3 ± 11.8
Hip circumference (cm)	147.4 ± 16.3	106.2 ± 10.4	139.6 ± 22.3
Waist circumference (cm)	139.5 ± 17.0	90.4 ± 16.1	130.4 ± 25.4
Waist-to-hip ratio	0.95 ± 0.10	0.85 ± 0.10	0.93 ± 0.11
Metabolic profile			
Fasting glucose (mmol L <sup>-1</sup> )	6.62 ± 2.32	5.75 ± 1.10	6.46 ± 2.18
Triglycerides (mmol L <sup>-1</sup> )	1.77 ± 1.00	1.23 ± 0.80	1.67 ± 0.99
Total cholesterol (mmol L <sup>-1</sup> )	4.59 ± 0.95	4.60 ± 0.99	4.59 ± 0.96
HDL cholesterol (mmol L <sup>-1</sup> )	1.24 ± 0.35	1.39 ± 0.42	1.27 ± 0.37
LDL cholesterol (mmol L <sup>-1</sup> )	2.59 ± 0.84	2.88 ± 0.95	2.64 ± 0.87
Total to HDL cholesterol ratio	3.92 ± 1.22	3.62 ± 1.49	3.87 ± 1.28
Blood pressure			
Systolic blood pressure	136.3 ± 16.8	119.3 ± 10.8	133.1 ± 17.1
Diastolic blood pressure	82.0 ± 11.0	77.7 ± 8.4	81.2 ± 10.7

Data are expressed as mean ± standard deviation.

HC in the two obesity sub-cohorts exhibited highly significant associations (rs9661614  $P = 7.9 \times 10^{-6}$ ; rs485660  $P = 2.2 \times 10^{-6}$ ) that held after correction for multiple testing (rs9661614 FDR-corrected  $P = 2.6 \times 10^{-4}$ ; rs485660 FDR-corrected  $P = 1.4 \times 10^{-3}$ ). With BMI, age and sex included as co-variables in the model, 0.90% and 0.50% of HC variance were explained by rs9661614 and rs485660, respectively. Least square means post hoc analyses revealed that rs9661614 heterozygotes were characterized by narrower hips, as compared with common ( $P_{\text{bon}} = 0.0002$ ;  $\Delta\text{HC} = -1.5$  cm) and rare homozygotes ( $P_{\text{bon}} = 0.0004$ ;  $\Delta\text{HC} = -2.7$  cm). In contrast, rs485660 heterozygotes showed a significant HC reduction only when compared with rare homozygotes ( $P_{\text{bon}} = 0.003$ ;  $\Delta\text{HC} = -2.7$  cm), but not to common homozygotes ( $P_{\text{bon}} = 0.06$ ;  $\Delta\text{HC} = -0.9$  cm). Although rare homozygotes of both rs9661614 and rs485660 showed wider hips than common homozygotes, this increase in HC did not reach statistical significance (Table 3).

**Table 2** Genetic features of SYPL2 SNPs in the study populations

	Obesity		Infogene		Obesity + Infogene	
	MAF (%)	HWE <i>P</i> -value	MAF (%)	HWE <i>P</i> -value	MAF (%)	HWE <i>P</i> -value
rs9661614	0.27	0.30	0.26	0.69	0.27	0.25
rs485660	0.23	0.49	0.22	0.89	0.23	0.48
rs62623713	0.04	0.29	0.03	0.15	0.04	0.19

HWE, Hardy–Weinberg equilibrium; MAF, minor allele frequency; SNP, single-nucleotide polymorphism.

Association between rs9661614 and rs485660 with hip circumference was validated by a joint analysis

Further validation of the association found with rs9661614, rs485660 and HC in the obesity cohort was performed by means of a joint analysis. Again, both SNPs showed a significant association with HC in the combined cohort (rs9661614 FDR-corrected  $P = 7.5 \times 10^{-4}$ ; rs485660 FDR-corrected  $P = 3.9 \times 10^{-2}$ ). rs9661614

**Table 3** Hip circumference differences among genotypes of SYPL2 SNPs

		Obesity	Obesity + Infogene
		HC (cm)	HC (cm)
rs9661614	TT	147.0 ± 0.3 <sup>a</sup>	139.2 ± 0.2 <sup>a</sup>
	TC	145.5 ± 0.3 <sup>b</sup>	137.9 ± 0.2 <sup>b</sup>
	CC	148.1 ± 0.6 <sup>a</sup>	140.0 ± 0.5 <sup>a</sup>
rs485660	GG	146.7 ± 0.2 <sup>a,b</sup>	138.9 ± 0.2 <sup>a,b</sup>
	GA	145.8 ± 0.3 <sup>b</sup>	138.2 ± 0.3 <sup>b</sup>
	AA	148.5 ± 0.8 <sup>a</sup>	140.3 ± 0.6 <sup>a</sup>

Values are adjusted least square means (LS-means ± SEM) derived from analysis of variance (general linear models) adjusted for age, sex and BMI. For each SNP and cohort, superscript lowercase letters stand for significantly different genotype means with Bonferroni adjusted *P*-values ( $P_{\text{bon}} \leq 0.01$ ) following post hoc LS-means pairwise comparisons.

BMI, body mass index; HC, hip circumference; LS-means, least square means; SEM, standard error of the mean; SNP, single-nucleotide polymorphism.



showed similar results as those obtained in the obesity cohort, with heterozygotes exhibiting a significant reduction of HC, as compared with common ( $P_{\text{bon}}=0.0002$ ;  $\Delta\text{HC}=-1.3$  cm) and rare homozygotes ( $P_{\text{bon}}=0.001$ ;  $\Delta\text{HC}=-2.1$  cm; Table 3). Likewise, rs485660 showed significantly narrower hips compared with rare homozygotes ( $P_{\text{bon}}=0.005$ ;  $\Delta\text{HC}=-2.2$  cm), but did not differ from common homozygotes ( $P_{\text{bon}}=0.06$ ;  $\Delta\text{HC}=-0.8$  cm). Rare homozygotes of both SNPs continued to show the widest hips but were not statistically different from common homozygotes (Table 3). Although the percentage of HC variance explained by rs9661614 and rs485660 was reduced in the joint analysis, the impact of rs9661614 (0.24%) almost doubled to that of rs485660 (0.13%), similar to the ratio observed in the obesity cohort.

#### Diplotype-based analysis reveals the heterozygote disadvantage of rs9661614

Haplotype reconstruction with the two SNPs significantly associated with HC, rs9661614 (T>C) and rs485660 (G>A), in this order, led to the identification of three major haplotypes in both the obesity (TG: 72.5, CA: 23.0 and CG: 4.3%) and the combined (TG: 72.8, CA: 22.8, CG: 4.2%) cohorts. Haplotypes were scattered among four diplotypes with frequencies higher than 5% (Table 4). Diplotype-based analyses revealed a significant association with HC in the obesity cohort ( $P=6.3 \times 10^{-6}$ , FDR-corrected  $P=5.7 \times 10^{-5}$ ), which remained highly significant in the combined cohort ( $P=1.5 \times 10^{-5}$ , FDR-corrected  $P=1.6 \times 10^{-4}$ ). As shown in Table 4, the reduction of HC previously observed in heterozygotes for rs9661614 (TC) and rs485660 (GA) was mirrored herein,

**Table 4** Hip circumference differences among SYPL2 diplotypes

Diplotypes	Obesity		Obesity + Infogene	
	Freq (%)	HC (cm)	Freq (%)	HC (cm)
CA/CA	5.5	148.6 ± 0.8 <sup>a</sup>	5.6	140.3 ± 0.6 <sup>a</sup>
TG/TG	53.0	147.0 ± 0.3 <sup>a</sup>	54.8	139.1 ± 0.2 <sup>a</sup>
TG/CA	32.8	145.7 ± 0.3 <sup>b</sup>	33.7	138.0 ± 0.3 <sup>b</sup>
TG/CG	6.1	144.4 ± 0.7 <sup>b</sup>	6.0	137.0 ± 0.6 <sup>b</sup>

Values are adjusted least square means (LS-means ± SEM) derived from analysis of variance (general linear models) adjusted for age, sex and BMI. For each SNP and cohort, lowercase letters a and b stand for significantly different genotype means with Bonferroni adjusted  $P$ -values ( $P_{\text{bon}} \leq 0.01$ ) following post hoc LS-means pairwise comparisons.

BMI, body mass index; HC, hip circumference; Freq, diplotype frequency; LS-means, least square means; SEM, standard error of the mean.

with subjects carrying the TG/CA diplotype showing a significant reduction of HC, as compared with the common diplotype (TG/TG) in both the obesity ( $P_{\text{bon}}=0.002$ ;  $\Delta\text{HC}=-1.3$  cm) and the combined cohorts ( $P_{\text{bon}}=0.006$ ;  $\Delta\text{HC}=-1.1$  cm). This HC reduction was even more pronounced in subjects carrying the TG/CG diplotype, formed by heterozygotes of rs9661614 (TC) and common homozygotes of rs485660 (GG), in both the obesity ( $P_{\text{bon}}=0.0004$ ;  $\Delta\text{HC}=-2.6$  cm) and the combined cohorts ( $P_{\text{bon}}=0.006$ ;  $\Delta\text{HC}=-2.3$  cm; Table 4). Although not significantly different from the common diplotype (TG/TG), individuals carrying the rare diplotype (CA/CA) had the widest hips in both cohorts, with as much as 4.2 cm ( $P_{\text{bon}}=0.006$ ; obesity cohort) and 3.3 cm ( $P_{\text{bon}}=0.001$ ; combined cohorts) more than the diplotype with the narrowest hips (TG/CG; Table 4).

#### Associations with hip circumference occur in a sex-specific manner

No  $G \times S$  interaction with rs9661614 was identified in both the obesity ( $P=0.08$ ) and the combined cohort ( $P=0.1$ ). No further significant  $G \times S$  interactions were found with rs485660 in the obesity ( $P=0.35$ ) or the combined cohort ( $P=0.42$ ). Nevertheless, because body fat distribution is a sex-specific characteristic of obesity, the significant associations found with HC were tested independently in men and women in the joint analysis. As shown in Table 5, the association of rs9661614 with HC previously observed remained highly significant when tested independently in women in the combined cohorts ( $P=5.3 \times 10^{-4}$ , FDR-corrected  $P=1.7 \times 10^{-2}$ ). Indeed, rs9661614 heterozygotes exhibited a significantly lower HC as compared with common homozygotes ( $P_{\text{bon}}=0.0005$ ;  $\Delta\text{HC}=-1.4$  cm). On the other hand, the association of rs485660 with HC was lost in women ( $P=1.9 \times 10^{-2}$ , FDR-corrected  $P=0.2$ ), and both associations were no longer significant in men (Table 5). Diplotype-based association test performed in the

**Table 5** Genotype and diplotype sex-dependent associations with hip circumference in the combined cohort

	Men		Women	
	$P$ -value	FDR $P$	$P$ -value	FDR $P$
rs9661614	$6.6 \times 10^{-3}$	0.2	$5.3 \times 10^{-4}$	$1.7 \times 10^{-2}$
rs485660	$6.4 \times 10^{-2}$	0.7	$1.9 \times 10^{-2}$	0.2
Diplotype	$3.7 \times 10^{-2}$	0.4	$2.8 \times 10^{-4}$	$3.1 \times 10^{-3}$

$P$ -values are derived from analysis of variance (general linear models) adjusted for age and BMI.

BMI, body mass index; FDR  $P$ , false discovery rate-corrected  $P$ -value.

combined cohorts showed that the impact of diplotypes formed by rs9661614 and rs485660 on HC was restricted almost exclusively to women ( $P=2.8 \times 10^{-4}$ , FDR-corrected  $P=3.1 \times 10^{-3}$ ), with TG/CA ( $P_{\text{bon}}=0.002$ ;  $\Delta\text{HC}=-1.2$  cm) and TG/CG ( $P_{\text{bon}}=0.001$ ;  $\Delta\text{HC}=-2.4$  cm) showing a significantly lower HC.

## Discussion

Results of this study revealed a novel genetic association between the obesity-related gene *SYPL2* and HC. The most robust association was found with the intron variant rs9661614 (T>C), with heterozygotes exhibiting a significant reduction of HC, as compared with common and rare homozygotes. It is worth highlighting that rs9661614 was in partial LD with rs62623713, the low-frequency missense variant located at exon 4 of *SYPL2* and significantly associated with BMI (9). Because most genetic associations are usually indirect markers for the actual susceptibility variants and given the LD pattern observed herein, we were not able to determine the actual functional relevance of such SNPs. Indeed, the observed effect could be due to an unknown functional SNP in LD with rs9661614 in *SYPL2* or in neighbouring genes. As a noncoding polymorphism, we may however speculate on a potential role as an enhancer element of nearby genes, to be associated to a DNase hypersensitivity site, or even to be mapped to an open chromatin region, may affect the functionality of transcriptional activity of nearby genes (22). Indeed, *SYPL2* has been previously analysed in relation to the cholesterol-associated locus 1p13, a region with a complex architecture and encompassed by several genes where a noncoding polymorphism was revealed as the causal variant for altered LDL-C through altering *SORT1* expression (23). Thus, the consistent association found between rs9661614 and HC in subjects with obesity, further validated in a more heterogeneous population, supports the relevance of *SYPL2* as a potential susceptibility gene for obesity. Nevertheless, further functional genomics approaches focused on assessing the actual impact of *SYPL2* polymorphisms are still required.

The present two-stage study was performed first in a large obesity cohort, in order to discover novel associations with metabolic or anthropometric traits related to obesity in a population composed exclusively of subjects with obesity and severe obesity. The associations found in the obesity cohort were further validated by means of a joint analysis, where the infogene cohort composed of individuals with and without obesity (according to BMI) was merged with the obesity cohort, leading to a larger group and a broader range of adiposity values for further testing such associations. The rationale for performing

such a joint analysis was based on maximizing statistical power despite the use of a restrictive multiple-testing correction method to determine statistically significant associations, as previously reported (24).

The more prominent effect of *SYPL2* polymorphisms on HC was observed in women, suggesting a preferential association of this gene with a gynoid profile, metabolically less harmful than an android distribution of body fat (25). The genetic influence on body fat distribution has been broadly analysed (26), and two recent GWAS meta-analyses of WHR-associated loci have revealed that genes associated with anthropometric indices reflecting body fat distribution, such as WHR, WC or HC, usually are sexually dimorphic, with a higher impact in women (27,28). It is worth highlighting that the percentage of variance of WC (27) and WHR (28) attributed to locus-independent effects in these two studies was estimated to be less than 0.05%, whereas a variation of 0.24% was attributed to rs9661614 in the joint analysis of the present work. In terms of absolute effect size, this corresponds to less than 0.5 cm of WC change in the previous study (27), whereas it represents more than 2 cm of HC reduction in rs9661614 heterozygotes here. This significant difference in genetic penetrance has been previously attributed to increased power because of the use of extreme phenotypes, as compared with studies performed in more heterogeneous populations (29). This could also explain the lower % variance in HC attributed to rs9661614 in the joint analysis, as compared with the results obtained in the obesity cohort.

Consistent with the aforementioned findings, and given that large hips have been associated with a better prognosis of metabolic and cardiovascular complications (30–32), the results obtained herein also revealed what looks like a heterozygote disadvantage, more evident for rs9661614, whose impact on HC remained highly significant through the different cohorts analysed. Diplotype-based analyses reinforced this assumption by showing that diplotypes formed by rs9661614 heterozygotes had a major impact on HC. Particularly relevant was the reduction of HC observed in diplotypes formed by rs9661614 heterozygotes and rs485660 common homozygotes, which suggested that such reduction could be attributed exclusively to the presence of rs9661614 heterozygotes. Heterozygote disadvantage is a genetic feature worth taking into consideration in association studies, and several cases of heterozygote (dis)advantage have been recently reported regarding obesity. One of the most relevant has recently revealed a significant relationship between a polymorphism within the obesity-related gene *FTO* (33) and aortic valve stenosis (34), also reporting a sex-specific effect of this *FTO* polymorphism, showing a higher penetrance in women (34). As

mentioned earlier, results obtained in the present study also pointed to a sexual dimorphism of rs9661614. Because HC is a distinctive feature of body fat distribution between men and women, these results highlight a female-specific genetic association closely related to a gynoid profile of fat distribution.

It is worth highlighting that WHR, WC and HC are anthropometric features with an important sex-dependent effect, which makes risk thresholds different for men and women (35). Concretely, HC is a feature closely related to obesity in women, whose pattern of body fat distribution implies a greater accumulation of a relatively less deleterious subcutaneous fat in the lower half of the body, as compared with men, whose fat excess is more prone to accumulate in the form of visceral fat in the abdominal area (25,36). Based on these assumptions, the significant associations found herein between rs9661614 and HC have a physiological basis that highlights the sex-specific impact of genetic variants, as previously reported in a recent meta-analysis regarding WHR (13). Interestingly, WC and HC, taken into account independently, have been revealed as better predictors for the development of obesity-related complications than the WHR in certain situations, as extensively reviewed in (37). Recently, HC but not WHR was inversely associated to high TG and low HDL-C in a Chinese population of women with obesity, but not in men (38). Likewise, another study highlighted the relevance of introducing HC as an independent factor when assessing the mortality risk of central obesity (39).

Some strengths of this study worth considering are the large number of subjects making up the cohorts used for association studies, which provide an adequate statistical power to reveal associations with relatively weak effects. These strengths, together with the consistency of the association shown by rs9661614 with HC throughout this two-stage study, make the results a rather robust basis to consider *SYPL2* as a gene relevant to body fat distribution. In this sense, although other associations were found between *SYPL2* polymorphisms and phenotypes (data not shown), they either did not show sufficient consistency when tested in the different cohorts or did not hold after correction for multiple testing. Interestingly, because HC exhibited very significant correlations with the majority of the phenotype traits analysed (data not shown), we cannot rule out a side effect of this *SYPL2* polymorphism on these traits through an effect mediated by their impact on HC. Based on these results and those obtained previously showing that mice lacking *SYPL2* exhibited reduced body weight (40), the following step is to deepen our understanding of the functional relevance of this SNP on target tissues related to obesity, such as visceral or subcutaneous adipose tissues.

## Conclusion

Results reported in this study showed strong evidence for an association between an intron variant located within the obesity-related *SYPL2* gene and HC in women, which likely reflects a preferential association of this gene with a gynoid profile of fat distribution.

## Funding

JTM is funded by post-doctoral fellowships from the Institute of Nutrition and Functional Foods (INAF) and from *Fonds Germain-Brisson* (Laval University). MCV is the Canada Research Chair in Genomics Applied to Nutrition and Health. This study was supported by a grant-in-aid from the Heart and Stroke Foundation of Canada (G-14-0005824).

## Conflict of interest statement

No conflict of interest was declared.

## Author contributions

JTM performed statistical analysis, interpreted the data and drafted the manuscript; MCV and FG conceived and designed the research; AT, YD and LP participated in the elaboration of the study design. SB, PM, OL, LB and SM sampled blood from the study subjects. All authors read and approved the final manuscript.

## Acknowledgements

The authors acknowledge the invaluable collaboration of surgeons Picard Marceau, Frédéric-Simon Hould and Stéfane Lebel of the Québec Heart and Lung Institute (QHLI) bariatric surgery team, who have contributed to create the clinical setup necessary for the study and sampling of adipose tissues for this project. The authors also acknowledge the collaboration of the biobank staff of the QHLI. The authors would also like to acknowledge the contribution of Alain Houde and Catherine Raymond for technical assistance.

## References

1. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature* 2015; **518**: 197–206.
2. Graff M, Ngwa JS, Workalemahu T, et al. Genome-wide analysis of BMI in adolescents and young adults reveals additional insight into the effects of genetic loci over the life course. *Hum Mol Genet* 2013; **22**: 3597–3607.
3. Speliotes EK, Willer CJ, Berndt SI, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet* 2010; **42**: 937–948.

4. Claussnitzer M, Dankel SN, Kim K-H, et al. *FTO* obesity variant circuitry and adipocyte browning in humans. *N Engl J Med* 2015; **373**: 895–907.
5. Evans DS, Calton MA, Kim MJ, et al. (eds). Genetic association study of adiposity and melanocortin-4 receptor (*MC4R*) common variants: replication and functional characterization of non-coding regions Hribal ML (ed.). *PLoS One* 2014; **9**: e96805.
6. Wen W, Zheng W, Okada Y, et al. Meta-analysis of genome-wide association studies in East Asian-ancestry populations identifies four new loci for body mass index. *Hum Mol Genet* 2014; **23**: 5492–5504.
7. Monda KL, Chen GK, Taylor KC, et al. A meta-analysis identifies new loci associated with body mass index in individuals of African ancestry. *Nat Genet* 2013; **45**: 690–696.
8. Jiao H, Kulyté A, Näslund E, et al. Whole exome sequencing suggests *LAMB3* as a susceptibility gene for morbid obesity. *Diabetes* 2016; **65**: 2980–2989.
9. Jiao H, Arner P, Gerdhem P, et al. Exome sequencing followed by genotyping suggests *SYPL2* as a susceptibility gene for morbid obesity. *Eur J Hum Genet* 2015; **23**: 1216–1222.
10. Park SL, Cheng I, Pendergrass SA, et al. Association of the *FTO* obesity risk variant rs8050136 with percentage of energy intake from fat in multiple racial/ethnic populations: the PAGE study. *Am J Epidemiol* 2013; **178**: 780–790.
11. Poveda A, Ibáñez ME, Rebato E. Common variants in *BDNF*, *FAIM2*, *FTO*, *MC4R*, *NEGR1*, and *SH2B1* show association with obesity-related variables in Spanish Roma population. *Am J Hum Biol* 2014; **26**: 660–669.
12. van Vliet-Ostaptchouk JV, Nuotio M-L, Slagter SN, et al. The prevalence of metabolic syndrome and metabolically healthy obesity in Europe: a collaborative analysis of ten large cohort studies. *BMC Endocr Disord* 2014; **14**: 9.
13. Winkler TW, Justice AE, Graff M, et al. The influence of age and sex on genetic associations with adult body size and shape: a large-scale genome-wide interaction study. *PLoS Genet* 2015; **11**: e1005378.
14. Brotto MAP, Nagaraj RY, Brotto LS, Takeshima H, Ma JJ, Nosek TM. Defective maintenance of intracellular Ca<sup>2+</sup> homeostasis is linked to increased muscle fatigability in the MG29 null mice. *Cell Res* 2004; **14**: 373–378.
15. Robitaille J, Després J-P, Pérusse L, Vohl M-C. The PPAR- $\gamma$  P12A polymorphism modulates the relationship between dietary fat intake and components of the metabolic syndrome: results from the Québec Family Study. *Clin Genet* 2003; **63**: 109–116.
16. Paradis A-M, Godin G, Pérusse L, Vohl M-C. Interaction between familial history of obesity and fat intakes on obesity phenotypes. *J Nutrigenet Nutrigenomics* 2009; **2**: 37–42.
17. Cormier H, Thifault É, Gameau V, et al. Association between yogurt consumption, dietary patterns, and cardio-metabolic risk factors. *Eur J Nutr* 2015; **55**: 577–587.
18. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005; **21**: 263–265.
19. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.
20. Stephens M, Donnelly P. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *Am J Hum Genet* 2003; **73**: 1162–1169.
21. Faul F, Erdfelder E, Lang A-G, Buchner A. G\*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods* 2007; **39**: 175–191.
22. Zhang F, Lupski JR. Non-coding genetic variants in human disease. *Hum Mol Genet* 2015; **24**: R102–R110.
23. Musunuru K, Strong A, Frank-Kamenetsky M, et al. From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. *Nature* 2010; **466**: 714–719.
24. Skol AD, Scott LJ, Abecasis GR, Boehnke M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat Genet* 2006; **38**: 209–213.
25. Tchernof A, Després J-P. Pathophysiology of human visceral obesity: an update. *Physiol Rev* 2013; **93**: 359–404.
26. Zillikens MC, Yazdanpanah M, Pardo LM, et al. Sex-specific genetic effects influence variation in body composition. *Diabetologia* 2008; **51**: 2233–2241.
27. Lindgren CM, Heid IM, Randall JC, et al. Genome-wide association scan meta-analysis identifies three loci influencing adiposity and fat distribution. *PLoS Genet* 2009; **5**: e1000508.
28. Heid IM, Jackson AU, Randall JC, et al. Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. *Nat Genet* 2010; **42**: 949–960.
29. Berndt SI, Gustafsson S, Mägi R, et al. Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. *Nat Genet* 2013; **45**: 501–512.
30. Seidell JC, Perusse L, Despres J-P, Bouchard C. Waist and hip circumferences have independent and opposite effects on cardiovascular disease risk factors: the Québec Family Study. *Am J Clin Nutr* 2001; **74**: 315–321.
31. Snijder MB, Zimmet PZ, Visser M, Dekker JM, Seidell JC, Shaw JE. Independent and opposite associations of waist and hip circumferences with diabetes, hypertension and dyslipidemia: the AusDiab Study. *Int J Obes Relat Metab Disord* 2004; **28**: 402–409.
32. Pio-Magalhães JA, Ferreira-Sae MC, Souza FA, et al. Hip circumference is associated with high density lipoprotein cholesterol response following statin therapy in hypertensive subjects. *J Endocrinol Invest* 2011; **34**: 680–684.
33. Frayling TM, Timpson NJ, Weedon MN, et al. A common variant in the *FTO* gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 2007; **316**: 889–894.
34. Thron C, Akhyari P, Godehardt E, Lichtenberg A, Rütger U, Seehaus S. *FTO* is associated with aortic valve stenosis in a gender specific manner of heterozygote advantage: a population-based case-control study. *PLoS One* 2015; **10**: e0139419.
35. Grundy SM, Cleeman JI, Daniels SR, et al. Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute Scientific Statement. *Circulation* 2005; **112**: 2735–2752.
36. Karastergiou K, Smith SR, Greenberg AS, Fried SK. Sex differences in human adipose tissues – the biology of pear shape. *Biol Sex Differ* 2012; **3**: 13.
37. Cameron AJ, Magliano DJ, Söderberg S. A systematic review of the impact of including both waist and hip circumference in risk models for cardiovascular diseases, diabetes and mortality. *Obes Rev* 2013; **14**: 86–94.
38. Katz EG, Stevens J, Truesdale KP, Cai J, Adair LS, North KE. Hip circumference and incident metabolic risk factors in Chinese men and women: the People's Republic of China study. *Metab Syndr Relat Disord* 2011; **9**: 55–62.
39. Cameron AJ, Magliano DJ, Shaw JE, et al. The influence of hip circumference on the relationship between abdominal obesity and mortality. *Int J Epidemiol* 2012; **41**: 484–494.
40. Nishi M, Komazaki S, Kurebayashi N, et al. Abnormal features in skeletal muscle from mice lacking mitsugumin29. *J Cell Biol* 1999; **147**: 1473–1480.