

Title: A CpG-SNP located within the *ARPC3* gene promoter is associated with hypertriglyceridemia in severely obese patients.

Authors: Juan de Toro-Martín^{1,2}, Frédéric Guénard^{1,2}, André Tchernof^{2,3}, Yves Deshaies^{3,4}, Louis Pérusse^{1,5}, Simon Biron⁶, Odette Lescelleur⁶, Laurent Biertho⁶, Simon Marceau⁶ and Marie-Claude Vohl^{1,2}

¹ Institute of Nutrition and Functional Foods (INAF), Laval University, Quebec, QC, Canada

² School of Nutrition, Laval University, Quebec, QC, Canada

³ Quebec Heart and Lung Institute Research Center, Quebec, QC, Canada

⁴ Department of Medicine, Laval University, Quebec, QC, Canada

⁵ Department of Kinesiology, Laval University, Quebec, QC, Canada

⁶ Department of Surgery, Laval University, Quebec, QC, Canada

Keywords: ARPC3, obesity, visceral adipose tissue, CpG-SNP, metabolic syndrome, triglycerides

Running title: A triglyceride-associated CpG-SNP in *ARPC3* gene

Contact info:

Marie-Claude Vohl, Ph.D.
Institute of Nutrition and Functional Foods (INAF)
Université Laval
2440 Hochelaga Blvd
Quebec, Quebec
Canada, G1V 0A6
Email: marie-claude.vohl@fsaa.ulaval.ca
Phone number: (418) 656-2131 ext. 4676

Supplementary Table 1. Ranked affinity transcription factor matrices for rs3759384 (C>T) containing sequences.

Rank	C	T	Difference $\log(P)$	Matrix ID	Matrix name
1	0.50	1.5×10^{-6}	-5.53	MA0095.1	YY1
2	0.76	0.08	-1.00	MA0036.1	GATA2
3	0.76	0.14	-0.75	MA0037.1	GATA3
4	0.15	0.69	0.67	MA0018.2	CREB1
5	0.13	0.51	0.59	MA0032.1	FOXC1
6	0.86	0.40	-0.34	MA0077.1	SOX9
7	0.45	0.97	0.33	MA0076.1	ELK4
8	0.45	0.93	0.32	MA0099.2	AP1
9	0.66	0.33	-0.30	MA0087.1	Sox5
10	0.75	0.40	-0.28	MA0151.1	ARID3A

Predicted transcription factor (TF) binding affinities were analyzed using sTRAP [27], which calculates the difference in binding affinity between common and rare allele-containing sequences. Sequences overlapping a differentially methylated CpG site (5 bp before and after) were used as input sequence. The JASPAR vertebrates database [28] was used as TF matrix file and human promoter sequences were introduced as background model to compare affinities. The table is sorted from the highest to the lowest absolute difference $\log(P)$ for two sequences. C and T stand for Benjamini-Hochberg-corrected P values for C- and T-carrying sequences, respectively.

Supplementary Figure 1.

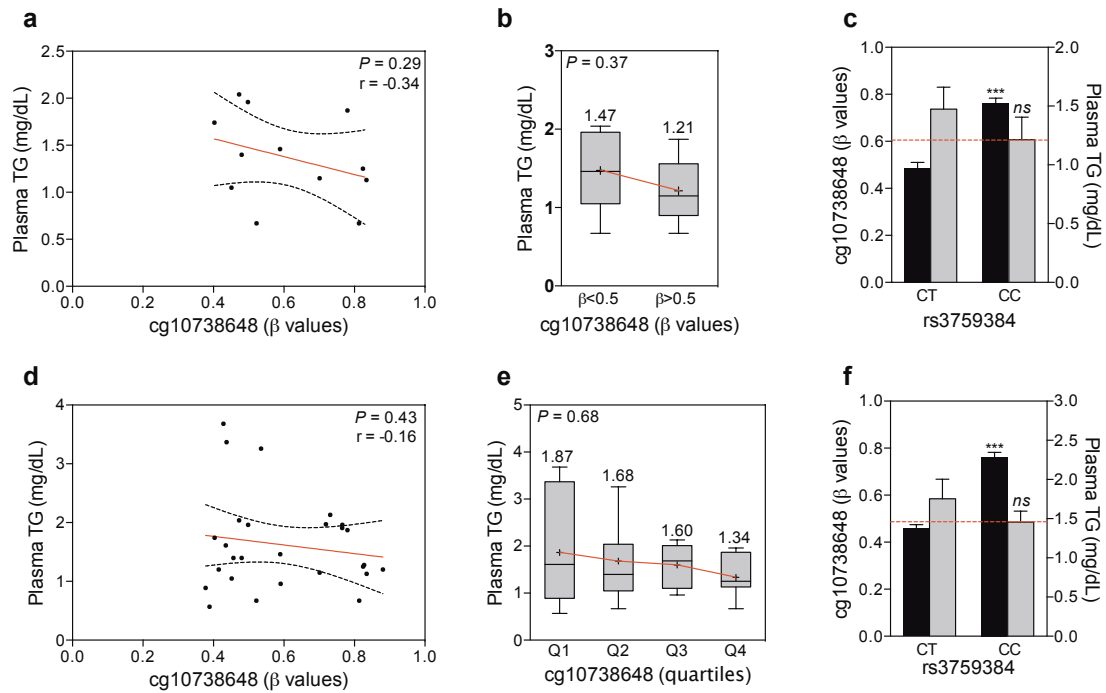


Figure S1. Association between methylation levels of cg10738648 and plasma triglyceride levels. a. Pearson's correlation between cg10738648 (β values) and plasma triglyceride (TG) levels (mg/dL). P value and Pearson's r coefficient are shown ($n=16$). **b.** Plasma TG levels by methylation levels of cg10738648 in visceral adipose tissue (VAT) shown as $\beta < 0.5$ (undermethylated) and $\beta > 0.5$ (overmethylated). P for 2-tailed Student t -test ($n=16$). **c.** Methylation levels of cg10738648 in VAT (left axis; black bars) shown as average β methylation ranging from 0 (unmethylated) to 1 (completely methylated) and plasma TG levels (right axis; grey bars). Values presented are adjusted least square means (LS-means \pm SEM) for each genotype and P values are adjusted for the effect of age and body mass index ($n=16$). *** $P < 1 \times 10^{-4}$, ns = not significant. **d.** Pearson's correlation between cg10738648 (β values) and plasma TG levels (mg/dL). P value and Pearson's r coefficient are shown ($n=31$). **e.** Plasma TG levels by methylation levels of cg10738648 in VAT shown as quartiles (Q). P for one-way ANOVA ($n=31$). **f.** Methylation levels of cg10738648 in VAT (left axis; black bars) shown as average β methylation ranging from 0 (unmethylated) to 1 (completely methylated) and plasma TG levels (right axis; grey bars). Values presented are adjusted least square means (LS-means \pm SEM) for each genotype and P values are adjusted for the effect of age and body mass index ($n=31$). *** $P < 1 \times 10^{-4}$, ns = not significant.