J Community Genet (2012) 3:319–322
DOI 10.1007/s12687-012-0081-5

SHORT COMMUNICATION

Allelic clustering and ancestry-dependent frequencies of rs6232, rs6234, and rs6235 *PCSK1* SNPs in a Northern Ontario population sample

Francine Sirois · Nadine Kaefer · Krista A. Currie · Michel Chrétien · Kabwe K. Nkongolo · Majambu Mbikay

Received: 22 June 2011 / Accepted: 20 January 2012 / Published online: 4 February 2012

Abstract The *PCSK1* (proprotein convertase subtilisin/kexin type 1) locus encodes proprotein convertase 1/3, an endoprotease that converts prohormones and proneuropeptides to their active forms. Spontaneous loss-of-function mutations in the coding sequence of its gene have been linked to obesity in humans. Minor alleles of two common non-synonymous single-nucleotide polymorphisms (SNPs), rs6232 (T>C, N221D) and rs6235 (C>G, S690T), have been associated with increased risk of obesity in European populations. In this study, we compared the frequencies of the rs6232 and rs6234 (G>C, Q665E) SNPs in Aboriginal and Caucasian populations of Northern Ontario. The two SNPs were all relatively less frequent in Aboriginals: The minor allele frequency of the rs6232 SNP was 0.01 in Aboriginals and 0.08 in Caucasians ($P<4.10^{-6}$); for the rs6234 SNP, it was 0.20 and 0.32, respectively ($P<0.001$). Resequencing revealed that the rs6234 SNP variation was tightly linked to that of the rs6235 SNP, as previously reported. Most interestingly, all carriers of the rs6232 SNP variation also carried the rs6234/rs6235 SNP clustered variations, but not the reverse, suggesting the former occurred later on an allele already carrying the latter. These data indicate that, in Northern Ontario Aboriginals, the triple-variant *PCSK1* allele is relatively rare and might be of lesser significance for obesity risk in this population.

Keywords Proprotein convertase · Obesity · Genetic polymorphism · Population genetics · Aboriginals · Caucasians

Electronic supplementary material The online version of this article (doi:10.1007/s12687-012-0081-5) contains supplementary material, which is available to authorized users.

F. Sirois · N. Kaefer · M. Chrétien · M. Mbikay
Chronic Disease Program, Ottawa Hospital Research Institute,
725 Parkdale Avenue,
Ottawa, ON K1Y 4E9, Canada

M. Chrétien · M. Mbikay (✉)
Department of Biochemistry, Microbiology and Immunology,
Faculty of Medicine, University of Ottawa,
Ottawa, ON, Canada
e-mail: mmbikay@ohri.ca

K. A. Currie · K. K. Nkongolo
Department of Biology, Faculty of Sciences,
Laurentian University,
Sudbury, ON, Canada

Introduction

Proprotein convertase 1/3 (PC1/3) is a serine endoprotease involved in the proteolytic activation of prohormones and proneuropeptides in the secretory pathway of endocrine and neuroendocrine cells (Seidah and Chretien 1999). It is the product of the 14-exon *PCSK1* (proprotein convertase subtilisin/kexin-type 1) gene mapping on human chromosome 5. It is biosynthesized in the endoplasmic reticulum as a secretory zymogen which gets activated in downstream secretory compartments by successive autocatalytic removal of segments at the amino and at the carboxyl termini (Muller and Lindberg 1999).

PC1/3 is highly expressed in hypothalamic nuclei that controls appetite (Dong et al. 1997). Its substrates in the brain and the periphery include precursors of hormones and neuropeptides that regulate feeding and food processing. Induced PC1/3 genetic deficiency in mouse causes greater body weight gain (Zhu et al. 2002; Lloyd et al. 2006; Mbikay et al. 2007). Spontaneous non-synonymous mutations in its sequence have been linked to obesity in humans

(Jackson et al. 1997, 2003; Farooqi et al. 2007). In a case–control association study in populations of European ancestry, Benzinou et al. (2008) were the first to show that three common non-synonymous single-nucleotide polymorphisms (SNPs) in the *PCSK1* locus are strongly linked to risk of obesity. Because Aboriginal populations of Canada are disproportionably burdened by obesity (Liu et al. 2006), we sought to determine the frequencies of these SNPs in a sample population of Aboriginal Canadians from Northern Ontario.

Subjects

This study was conducted with archival blood samples. These samples were collected to determine the frequency and distribution of genetic variations in Northern Ontario Aboriginals and Caucasians (Currie et al. 2005). Ancestry was the only information asked of participants. Unrelated subjects were recruited at three Northern Ontario regional hospitals located in Espanola (100 Aboriginals, 100 Caucasians), Sioux Lookout (100 Aboriginals), and Wikwemikong (100 Aboriginals) (Currie et al. 2005). Aboriginal participants were predominantly of the Cree-Ojibwa groups. The recruitment was conducted under a protocol approved by the Human Research Ethics Board of each hospital. Participants were explained the nature of the study and were asked for their informed consent and a blood sample. All samples were anonymized.

Methods

Genotyping of the rs6232 (T>C) SNP was conducted on the MX 3005P thermocycler (Stratagene, La Jolla, CA) using a Taqman assay based on the presence or absence of fluorescence due to degradation of allele-specific fluorochrome-conjugated probes as previously described (Sirois et al. 2008). Genotyping of the rs6234 (G>C) SNP was based on polymerase chain reaction (PCR) and melting curve analysis of differential fluorescence loss during heat-induced dehybridization of the probes from the SNP-containing region of genomic amplicon (Didenko 2001). It was conducted on a LightCycler instrument (Roche, Laval, QC).

Results and discussion

The nonsynonymous rs6234 and rs6235 SNPs are located in exon 14; they consist of G>C and C>G transversions and cause Q665E and S690T amino acid substitutions, respectively (Benzinou et al. 2008). Genomic PCR amplicons of this exon were produced using lymphocyte DNA from 50 Espanola participants, 25 Caucasians, and 25 Aboriginals; the DNA amplicons was sequenced (see Electronic supplementary material and Methods online). Twelve of the 50 participants were double heterozygotes for both the rs6234 and the rs6235 SNPs, confirming the tight linkage between these two SNPs (Benzinou et al. 2008).

The rs6234 SNP (G>C) was genotyped in all 400 DNA samples. There was 100% concordance between the genotyping results and those of resequencing for the 50 Espanola samples. Overall, the minor allele frequency (MAF) for this SNP in Aboriginals and Caucasians was 0.20 and 0.32, respectively ($P<0.001$ by Fisher's exact test). The G/G, G/C, and C/C genotype frequencies were respectively 0.62, 0.32, and 0.04 in Aboriginals and 0.43, 0.50, and 0.07 in Caucasians (Table 1).

The rs6232 SNP (T>C) was also genotyped; the MAF was 0.01 in Aboriginals and 0.08 in Caucasians ($P=4.10^{-6}$) (Table 2). The T/T, T/C, and CC genotype frequencies were respectively 0.98, 0.02, and 0.00 in Aboriginals and 0.85, 0.14, and 0.01 in Caucasians. Interestingly, all carriers of the rs6232 variation also carried the rs6234/rs6235 variation cluster, but not the reverse, suggesting that the former occurred after the latter on the same *PCSK1* allele. The linkage remains to be confirmed by sequencing. The genotype frequencies associated with the rs6232 and rs6234 SNPs in the two Northern Ontario populations were consistent with Hardy–Weinberg equilibrium ($P>0.05$), excluding the possibility of a significant bias due to subject sampling, genotyping, or population dynamics. Through mechanisms yet to be elucidated, SNPs tend to cluster, and these clusters are non-randomly distributed in the human genome (Amos 2009). Our study is the first to demonstrate allelic clustering of the three non-synonymous *PCSK1* variations.

We have determined that the triple-variant zymogen specified by this allele undergoes post-translational proteolytic processing at its termini more efficiently than the common and the double-variant forms, when expressed ex vivo in a pituitary cell line. We have speculated that this accelerated processing might render this isoform unstable, causing a functional deficit of PC1/3 enzymatic activity (Mbikay et al. 2011).

Collectively, our data indicate that, in Northern Ontario, the rs6232, rs6234, and rs6235 SNPs are more frequent in Caucasians than in Aboriginals. MAFs for the rs6232 (0.08) and rs6235 (0.35) SNPs appear on average to be higher in Espanola Caucasians than those recently reported for various European population samples (0.052 to 0.064 and 0.22 to 0.47, respectively) (Benzinou et al. 2008; Goossens et al. 2009; Kilpelainen et al. 2009; Willer et al. 2009; Heni et al. 2010; Gjesing et al. 2011; Rouskas et al. 2011) (see also Electronic supplementary material Table S1). The sample size in the latter studies was far larger than that one used in our study, giving greater statistical power to their MAF

Table 1 Genotype and allele frequencies for the non-synonymous *PCSK1* rs6234 (G>C, Q665E) SNP

Ethnicity	Site	Genotype frequencies							Allele frequencies				
		G/G		G/C		C/C		Total count	G		C		Total count
		Freq	Count	Freq	Count	Freq	Count		Freq	Count	Freq	Count	
Aboriginals	Espanola	0.68	68	0.28	28	0.04	4	100	0.82	164	0.18	36	200
	Sioux Lookout	0.63	63	0.34	34	0.03	3	100	0.80	160	0.20	40	200
	Wikweminkong	0.62	62	0.33	33	0.05	5	100	0.785	157	0.215	43	200
	All	0.643	193	0.317	95	0.04	12	300	0.802	481	0.198	119*	600
Caucasians	Espanola	0.43	43	0.50	50	0.07	7	100	0.68	136	0.32	64	200

Freq frequency

*$P<0.001$ relative to Espanola Caucasians by Fisher's exact test

determinations. The DNA panel of this study was previously used to determine the allele frequencies at the highly polymorphic D18S535 short tandem-repeat locus. These frequencies differed not only between Aboriginal and Espanola Caucasians but also between the latter and European Caucasians (Currie et al. 2005). Historical immigration records indicate Northern Ontario Caucasians were predominantly of French and British origins. It is likely that the higher MAFs of *PCSK1* SNPs obtained with our limited sample of Northern Ontario Caucasians are due to genetic drift (e.g., founder effect) or adaptive evolution.

An association of the minor allele of rs6232 SNP with obesity parameters (body mass index, waist circumference, or waist–hip ratio) has been observed in three separate studies of European populations (Benzinou et al. 2008; Kilpelainen et al. 2009; Willer et al. 2009; Gjesing et al. 2011), but not two others (Heni et al. 2010; Rouskas et al. 2011). For the minor allele of rs6234 or rs6235 SNP, it was observed in three studies (Benzinou et al. 2008; Gjesing et al. 2011; Rouskas et al. 2011), but not in two others (Kilpelainen et al. 2009; Renstrom et al. 2009). In Han Chinese, the rs6234 SNP was associated with a combined phenotype of obesity and overweight in men, but not in women (Qi et al. 2010). In a study by Heni et al. (2010) of a German population sample, neither the rs6232 nor the rs6235 SNP minor allele was associated with weight-related traits; but they were both associated with impaired proinsulin processing and, for the rs6232 SNP, with insulin sensitivity. Interactions of these minor alleles with glucose homeostasis and incretin metabolism were also observed in a Danish population (Gjesing et al. 2011).

The relatively lower frequencies of these SNPs in Aboriginals calls for a systematic resequencing of the *PCSK1* locus to determine whether there are other common genetic variations that might affect body mass and glucose homeostasis in this population. This call is also applicable for sub-Saharan Africans as our and online genotyping data suggest that the SNPs examined in this study are even rarer in this population (Electronic supplementary material Table S1).

Table 2 Genotype and allele frequencies for non-synonymous *PCSK1* rs6232 (T>C, N221D) SNP

Ethnicity	Site	Genotype frequencies							Allele frequencies				
		T/T		T/C		C/C		Total count	T		C		Total count
		Freq	Count	Freq	Count	Freq	Count		Freq	Count	Freq	Count	
Aboriginals	Espanola	0.97	97	0.03	3	0	0	100	0.985	197	0.015	3	200
	Sioux Lookout	0.99	99	0.01	1	0	0	100	0.995	199	0.005	1	200
	Wikweminkong	0.98	98	0.02	2	0	0	100	0.99	198	0.010	2	200
	All	0.98	294	0.02	6	0	0	300	0.98	594	0.01	6*	600
Caucasians	Espanola	0.853	81	0.137	13	0.01	1	95	0.921	175	0.079	15	190

Freq frequency

*$P=4.10^{-6}$ relative to Espanola Caucasians by Fisher's exact test

Acknowledgments This study was supported by grants from the Canadian Institutes of Health Research and from the Strauss Foundation.

Conflict of interest The authors declare that they have no conflict of interest.

References

Amos W (2009) Even small SNP clusters are non-randomly distributed: is this evidence of mutational non-independence? Proc R Soc B 277:1443–1449

Benzinou M, Creemers JW, Choquet H et al (2008) Common non-synonymous variants in PCSK1 confer risk of obesity. Nat Genet 40:943–945

Currie KA, Michael P, Nkongolo KK (2005) Characterization of the D18S535 STR locus in northern Ontario European and Aboriginal populations for forensic purposes. Hum Biol 77:267–279

Didenko VV (2001) DNA probes using fluorescence resonance energy transfer (FRET): designs and applications. Biotechniques 31:1106–1121

Dong W, Seidel B, Marcinkiewicz M et al (1997) Cellular localization of the prohormone convertases in the hypothalamic paraventricular and supraoptic nuclei: selective regulation of PC1 in corticotrophin-releasing hormone parvocellular neurons mediated by glucocorticoids. J Neurosci 17:563–575

Farooqi IS, Volders K, Stanhope R et al (2007) Hyperphagia and early-onset obesity due to a novel homozygous missense mutation in prohormone convertase 1/3. J Clin Endocrinol Metab 92:3369–3373

Gjesing AP, Vestmar MA, Jorgensen T et al (2011) The effect of PCSK1 variants on waist, waist-hip ratio and glucose metabolism is modified by sex and glucose tolerance status. PLoS One 6: e23907

Goossens GH, Petersen L, Blaak EE et al (2009) Several obesity- and nutrient-related gene polymorphisms but not FTO and UCP variants modulate postabsorptive resting energy expenditure and fat-induced thermogenesis in obese individuals: the NUGENOB study. Int J Obes (Lond) 33:669–679

Heni M, Haupt A, Schafer SA et al (2010) Association of obesity risk SNPs in PCSK1 with insulin sensitivity and proinsulin conversion. BMC Med Genet 11:86

Jackson RS, Creemers JW, Farooqi IS et al (2003) Small-intestinal dysfunction accompanies the complex endocrinopathy of human proprotein convertase 1 deficiency. J Clin Invest 112:1550–1560

Jackson RS, Creemers JW, Ohagi S et al (1997) Obesity and impaired prohormone processing associated with mutations in the human prohormone convertase 1 gene. Nat Genet 16:303–306

Kilpelainen TO, Bingham SA, Khaw KT, Wareham NJ, Loos RJ (2009) Association of variants in the PCSK1 gene with obesity in the EPIC-Norfolk study. Hum Molec Genet 18:3496–3501

Liu J, Hanley AJ, Young TK, Harris SB, Zinman B (2006) Characteristics and prevalence of the metabolic syndrome among three ethnic groups in Canada. Int J Obes 30:669–676

Lloyd DJ, Bohan S, Gekakis N (2006) Obesity, hyperphagia and increased metabolic efficiency in PC1 mutant mice. Hum Mol Genet 15:1884–1893

Mbikay M, Croissandeau G, Sirois F et al (2007) A targeted deletion/insertion in the mouse Pcsk1 locus is associated with homozygous embryo preimplantation lethality, mutant allele preferential transmission and heterozygous female susceptibility to dietary fat. Dev Biol 306:584–598

Mbikay M, Sirois F, Nkongolo KK, Basak A, Chretien M (2011) Effects of rs6234/rs6235 and rs6232/rs6234/rs6235 PCSK1 single-nucleotide polymorphism clusters on proprotein convertase 1/3 biosynthesis and activity. Mol Genet Metab 104:682–687

Muller L, Lindberg I (1999) The cell biology of the prohormone convertases PC1 and PC2. Prog Nucleic Ac Res Mol Biol 63:69–108

Qi Q, Li H, Loos RJ et al (2010) Association of PCSK1 rs6234 with obesity and related traits in a Chinese Han population. PLoS One 5:e10590

Renstrom F, Payne F, Nordstrom A et al (2009) Replication and extension of genome-wide association study results for obesity in 4923 adults from northern Sweden. Hum Molec Genet 18:1489–1496

Rouskas K, Kouvatsi A, Paletas K et al (2011) Common variants in FTO, MC4R, TMEM18, PRL, AIF1, and PCSK1 show evidence of association with adult obesity in the Greek population. Obesity. doi:10.1038/oby.2011.177

Seidah NG, Chretien M (1999) Proprotein and prohormone convertases: a family of subtilases generating diverse bioactive polypeptides. Brain Res 848:45–62

Sirois F, Gbeha E, Sanni A et al (2008) Ethnic differences in the frequency of the cardioprotective C679X PCSK9 mutation in a West African population. Genet Testing 12:377–380

Willer CJ, Speliotes EK, Loos RJ et al (2009) Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nat Genet 41:25–34

Zhu X, Zhou A, Dey A et al (2002) Disruption of PC1/3 expression in mice causes dwarfism and multiple neuroendocrine peptide processing defects. Proc Nat Acad Sci U S A 99:10293–10298