Endokrynologia Polska
DOI: 10.5603/EP.2015.0026
Tom/Volume 66; Numer/Number 3/2015
ISSN 0423–104X

Dense mapping of the region of insulin gene VNTR in polycystic ovary syndrome in a population of women from Central Europe

Gęste mapowanie regionu VNTR genu insuliny w zespole policystycznych jajników w populacji kobiet z Europy Środkowej

Redha Attaoua[1], Nelly Boeckler[1], Serban Radian[2], Samira Ait El Mkadem[1], Patrick Poucheret[3], Violaine Latapie[1], Monica Gheorghiu[2], Simona Fica[2], Madalina Vintila[2], Christophe Normand[1], Mihail Coculescu[2], Florin Grigorescu[1]

[1]*Molecular Endocrinology Laboratory, Montpellier, France*
[2]*Department of Endocrinology, University of Medicine and Pharmacy "Carol Davila", Bucharest, Romania*
[3]*Pharmacology and Experimental Physiology Laboratory, Faculty of Pharmacy, Montpellier France*

Abstract

Introduction: Insulin gene VNTR was associated with polycystic ovary syndrome (PCOS) in some studies but not in others. This could be due to the heterogeneity of the definition of PCOS and/or the use of inappropriate gene mapping strategies.
Material and methods: In this investigation, the association of VNTR with PCOS was explored in a population of women from Central Europe (377 cases and 105 controls) in whom PCOS was diagnosed according to Rotterdam criteria. Seven SNPs: rs3842756 (G/A), rs3842755 (G/T), rs3842754 (C/T), rs3842753 (A/C), rs3842752 (C/T), rs3842748 (G/C), and rs689 (T/A) were genotyped in a portion of the population (160 cases and 95 controls) by sequencing or by SSO-PCR. Analysis of linkage disequilibrium (LD) pattern allowed selecting three tagSNPs (rs3842754, rs3842748, and rs689), which were genotyped in the rest of the population by KASPar.
Results: Six haplotypes were reconstructed, among which three (h1, h2 and h6) were more frequent. Statistical analysis allowed observation of the association of the SNP rs3842748, through its GC genotype, with obesity in PCOS (P = 0.049; OR CI95% 1,59 [1.00–2.51]) and in classical PCOS (YPCOS) (P = 0.010), as well as the correlation of the SNP rs689 and the pair of haplotypes h1/h1 with higher levels of testosteronaemia in the PCOS group, although this was at the limit of significance (P = 0.054)
Conclusion: These results are in accordance with some studies in literature and highlight the role of insulin gene VNTR in complex metabolic disorders. **(Endokrynol Pol 2015; 66 (3): 198–206)**

Key words: *minisatellite; gene; PCOS; insulin; SNP*

Streszczenie

Wstęp. W niektórych badaniach, zmienna liczba powtórzeń tandemowych (VNTR) genu insuliny była związana z zespołem policystycznych jajników (PCOS), lecz w innych taki związek nie występował. Mogło tak być z powodu heterogeniczności definicji PCOS i/lub stosowania nieprawidłowych strategii mapowania genów.
Materiał i metody: W niniejszym badaniu, związek VNTR z PCOS zbadano w populacji kobiet pochodzących z Europy Środkowej (377 przypadków chorobowych oraz 105 osób kontrolnych), u których zdiagnozowano PCOS według kryteriów rotterdamskich. Siedem polimorfizmów pojedynczego nukleotydu (SNP): rs3842756 (G/A), rs3842755 (G/T), rs3842754 (C/T), rs3842753 (A/C), rs3842752 (C/T), rs3842748 (G/C), oraz rs689 (T/A) wytypowano w części populacji (160 przypadków chorobowych i 95 osób kontrolnych) poprzez sekwencjonowanie lub SSO-PCR. Analiza wzoru niezrównoważenia sprzężeń (LD) pozwoliła na selekcję trzech SNP znacznikowych (tagSNP) (rs3842754, rs3842748 i rs689), które wyselekcjonowano w pozostałej części populacji metodą KASPar.
Wyniki: Sześć haplotypów odtworzono, z których 3 (h1, h2 i h6) występowały częściej. Analiza statystyczna pozwoliła na obserwację związku SNP rs3842748, poprzez genotyp GC, z otyłością w PCOS (P = 0,049; OR CI 95% 1,59 [1,00–2,51]) i klasycznym PCOS (YPCOS) (P = 0,010), jak również korelacji SNP rs689 i pary haplotypów h1/h1 z wyższym stężeniem testosteronemii w grupie PCOS, chociaż wynik ten znajdował się na granicy istotności (P = 0,054).
Wnioski: Powyższe wyniki są zgodne z niektórymi badaniami w piśmiennictwie i podkreślają role VNTR genu insuliny w złożonych zaburzeniach metabolicznych. **(Endokrynol Pol 2015; 66 (3): 198–206)**

Słowa kluczowe: *minisatelita; gen; PCOS; insulina; SNP*

This work was supported by: Recurrent budget from University of Montpellier (France); Recurrent budget from University Medicine and Pharmacy "Carol Davila" in Bucharest (Romania). This work was also sustained by the grant MEDIGENE FP7-279171 of European Community (EC).

Florin Grigorescu M.D., Ph.D. Molecular Endocrinology Laboratory, UMR204 NUTRIPASS
Group Nutrition & Genomes, Institut Universitaire de Recherche Clinique (IURC)
641, avenue du Doyen Gaston Giraud, 34093 Montpellier France, tel.: +33 4 11 75 98 95, fax: +33 4 67 54 27 31, e-mail: florin.grigorescu@inserm.fr

Introduction

Polycystic ovary syndrome (PCOS) is a complex disease that affects 5 to 10% of women in reproductive age [1, 2]. PCOS is characterised by hyperandrogenism and chronic anovulation [2, 3], but also by a high frequency of insulin resistance, impaired glucose tolerance (IGT), insulin secretion abnormalities, and type 2 diabetes (T2D) [2, 4, 5].

The insulin gene minisatellite (VNTR, variable number of tandem repeats) was associated with PCOS [6, 7], although this has not been confirmed in other studies [8, 9]. VNTR of insulin gene was also associated with type 1 diabetes (T1D) [10, 11], T2D [12], childhood obesity [13], insulin sensitivity, insulin secretion in children [14], and insulin levels in plasma [15].

Insulin gene VNTR is localised 596 bp upstream from the transcription initiation site. It regulates the expression of both insulin and *IGF-2* (insulin-like growth factor-2) genes [16, 17]. Insulin gene VNTR is formed by a repetition of a consensus sequence of 14 to 15 bp. It involves two principle classes in the non-African populations: a short and a long one called class I and class III, respectively [10, 18]. VNTR of class I was associated with a higher risk of T1D [10] while VNTR of class III was associated with an increased risk of PCOS [6] and T2D [12]. The SNP (single nucleotide polymorphism) rs689 located upstream from exon 2 of the insulin gene was strongly associated with T1D [11]. This polymorphism has two alleles: A and T, in strong linkage disequilibrium (LD) with VNTR of class I and III, respectively, allowing for the use of this SNP as a surrogate of the minisatellite in association studies. However, it was reported that this polymorphism has a role in the regulation of insulin gene expression through an alternative splicing process [19], although the impact of the phenomenon is minor [20]. On the other hand, a high LD was observed between haplotypes of some SNPs in the insulin gene region and subclasses of the insulin gene VNTR in non-African populations [21]. Note that the subclasses IC and ID (VNTR of class I) were differently correlated with fasting insulin in childhood obesity in the French population [22], while the two other haplotypes in LD with the two subclasses IIIA and IIIB of the minisatellite were differently associated with the protective effect against T1D [23].

The role of the insulin gene and its minisatellite in PCOS and in the metabolic pattern of the disease remains controversial. This controversy was accentuated after the study carried out by Bouatia-Naji et al., who showed no familial association of insulin gene VNTR with childhood obesity in 1023 families or with fasting insulin levels or birth weight in obese children [24]. Moreover, genome-wide association studies (GWAS) screening hundred of thousands of SNPs associated the locus of insulin gene VNTR with prostate cancer [25], through the SNP rs7127900, but failed, to our knowledge, to find any association with metabolic disorders, which updates the question concerning loci found associated with complex disorders through the candidate gene approach and missed by GWAS studies.

In this context, the aim of the present study was to investigate the association of insulin gene VNTR with PCOS in a population of women from Central Europe based on haplotyping strategy. Our final aim was to compare the power of haplotypes with that of SNPs taken alone and to get a preliminary idea about the best strategy for studying the genetics of complex disorders.

Material and methods

Population recruitment

The study was approved by the Ethical Committee of The University of Bucharest (Romania), according to Helsinki Declaration (revised in 1983) and after informed consent. Participants in the investigation were not related and were of Caucasian origin (Region of Walachia, Romania). Type 1 genetic design [26] comparing PCOS with control lean healthy women was applied. Note that no patient in this study had diabetes.

Control population

The control population was composed of 105 lean healthy women recruited in the Obstetrics and Gynaecology Department of the University of Bucharest. These women had a normal reproductive status, regular menses, proven fertility (at least one child), and no clinical or biochemical hyperandrogenism.

PCOS population

A total of 377 women with PCOS were included. Diagnosis of PCOS was established according to Rotterdam criteria [27]. Individuals with Cushing Syndrome, non-classical deficiency of adrenal 21-hydroxylase, hyperprolactinaemia, and androgens secreting neoplasms were excluded from the study. PCOS women were stratified as lean (BMI < 30 kg/m^2) and obese (BMI > 30 kg/m^2). PCOS with more insulin resistance and *Acanthosis nigricans* (HA-IR-AN syndrome) was distinguished from classical PCOS (Y-PCOS), as previously reported [28].

Clinical and biochemical characteristics

The women were given a standard diet. Then, oral glucose tolerance test (OGTT) was applied to the group of PCOS women. IGT and impaired fasting glucose

(IFG) were defined according to American Diabetes Association (ADA) criteria [29]. Insulin resistance was assessed by $HOMA_{IR}$ (*homeostatic model assessment insulin resistance*) index, and insulin secretion by HOMA-β (*homeostatic model assessment* β*-cell function*). Note that $HOMA_{IR}$ (mmol/L) = (fasting glucose X fasting insulin)/22.5 while HOMA-β (mmol/L) = (fasting insulin X 20)/(fasting glucose — 3.5).

Measurement of lipid parameters and godanotropins (FSH and LH) in the plasma was carried out in the Department of Endocrinology in Bucharest. Clinical and biochemical characteristics of the population are shown in Table I.

Note that the first portion of the population (160 cases and 90 controls) was used for genotyping seven genetic markers (SNPs). The aim was to determine the LD pattern between these SNPs and select tag-SNPs to be genotyped in the rest of the population.

Genetic exploration

Description of the explored genomic region

The insulin gene (MIM: 176730) is located in the locus 11p15.5, downstream from its minisatellite (VNTR) by about 400 bp. The gene extends on 1431 bp and contains three exons. The region involving SNPs considered in the present study starts from the end of the first intron and extends beyond the 3'UTR side of the gene.

Genotyping of SNPs

All SNPs were genotyped in a portion of the population (160 women with PCOS and 95 controls) by classical sequencing or SSO-PCR. An overview of these techniques was previously reported [30]. Primers and PCR steps used for each SNP are reported in Table IIA. In the rest of the population (377 women with PCOS and 105 controls), only SNPs rs3842754, rs3842748, and rs689 were genotyped by KASPar technology [31], using primers and following the amplification steps reported in Table IIb.

Data and statistical analysis

The LD pattern of the explored genomic region was determined by Haploview 3.31 [32]. Haplotype reconstruction was carried out by PHASE 2.1 [33] and cladistics analysis by Arlequin 2.000 [34]. Concerning clinical and biochemical characteristics of the population, continuous variables (expressed as median [25% percentile — 75% percentile]) were analysed using non-parametrical tests. Significance when comparing between continuous variables in phenotypic groups was calculated by Mann-Whitney non-parametrical test (significance was attested for $P < 0.05$), while Fisher's exact test was used for comparing between percentages (significance was attested for $P(X^2) < 0.05$). The distribution of continuous parameters in the population was checked by Kolmogorov-Smirnov test. A non-normal distribution was attested for $P < 0.05$. Variation of SNPs was expressed as allelic frequency (q) or prevalence (%) of their genotypes, while variation of haplotypes was expressed as their prevalence or the prevalence of pairs of haplotypes. The genetic association was determined by logistic regression using StatView 5.0 and SAS (Abascus Concepts, Berkeley, CA), as previously described [35]. Results were expressed as nominal P, odds ratio (OR), and 95% confidence interval (CI). Genotype--phenotype correlations were assessed by Mann-Whitney test. The power of the case-control study was calculated by *Sampsize* software (http://sampsize.sourceforge.net/iface/s3.html).

Results

Phenotypic data

The explored population involved 482 women (377 with PCOS and 105 controls). Among PCOS women, 148 were obese and 71 were with severe insulin resistance (*hyperandrogenism-insulin resistance-Acanthosis nigricans* "HA-IR-AN" syndrome) (Table 1). Women with PCOS had higher BMI ($P < 0.0001$), higher LH/FSH ($P < 0.0001$), higher testosteronaemia ($P < 0.0001$), higher fasting insulin ($P < 0.0001$), and higher insulin secretion (HOMAβ) ($P < 0.0001$). On the other hand, *Acanthosis nigricans* was more often observed in PCOS obese women than in PCOS lean ($P < 0.0001$). Moreover, PCOS obese women had higher fasting insulin than lean PCOS ($P < 0.0001$), were more insulin resistant ($P < 0.0001$), and had higher insulin secretion (HOMAβ: $P < 0.0001$). HA-IR-AN women had higher BMI than YPCOS ($P < 0.0001$) and were more insulin resistant ($P < 0.0001$) and secretory ($P < 0.0095$). Note that no patient in this study had diabetes.

Note that the portion of the population where the seven SNPs were genotyped (160 women with PCOS and 95 controls) displayed the following characteristics: the median age of the PCOS group (25 and 75 percentiles) was 24.0 years (21.0–28.0) while it was 32.5 years (27.0 - 37.0) ($P < 0.0001$) for controls. PCOS women had higher BMI: 30.00 kg/m^2 (24.12–34.4) *vs.* 22.00 (20.00 – 23.85) ($P < 0.0001$) and were more insulin resistant: $HOMA_{IR}$: 3.31 (1.86–5.58) *vs.* 1.25 (0.90–1.68) ($P < 0.0001$). The test for the distribution of continuous parameters showed non-normal distribution of many parameters such as BMI ($P < 0.0001$), HOMA-β index, and plasmatic testosterone, while dehydroepiandrosterone sulphate (DHEAS) plasmatic levels displayed normal distribution (P).

Table I. *Phenotypic data of the female population from Central Europe used in the study. PS: One PCOS woman had no information concerning BMI and so was not included in calculations concerning lean/obese phenotypes. Continuous variables are reported as medians (25th percentile – 75th percentile). P of significance for continuous variables[(a)] was obtained by Mann-Whitney test while P of significance for nominal[(b)] variable was calculated by chi-square (X^2) test. The comparison was established between the group of women with PCOS and controls, between the group of PCOS obese women and PCOS lean women, or between women with HA-IR-AN syndrome and those with classical PCOS. The difference between groups is considered significant for P < 0.05*

Tabela I. *Dane fenotypiczne populacji kobiet z Europy Środkowej wykorzystane w badaniu. PS: Jedna kobieta cierpiąca na PCOS nie posiadała informacji dot. BMI i nie została włączona do obliczeń dot. fenotypów szczupłości/otyłości. Zmienne ciągłe są podane jako mediany (25 percentyl – 75 percentyl). P istotne dla zmiennych ciągłych[(a)] uzyskano poprzez test Manna-Whitneya, podczas gdy P istotne dla zmiennej nominalnej[(b)] zostało wyliczone testem chi-kwadrat (X^2). Porównanie zostało przeprowadzone pomiędzy grupami kobiet cierpiących na PCOS oraz osób kontrolnych, grupami otyłych kobiet cierpiących na PCOS oraz szczupłych kobiet cierpiących na PCOS, lub kobiet z zespołem HAIR-AN i kobiet z klasycznym PCOS. Różnica między grupami jest uznawana za istotną dla P < 0,05*

	Controls (n = 105)	PCOS (n = 377)	p	Lean PCOS (n = 228)	Obese PCOS (n = 148)	p	Y PCOS (306)	HA–IR–AN (n = 71)	p
Age (years)[a]	30.00 (25.75–37.00)	24.00 (21.00–28.00)	< 0.0001	23.00 (20.00–27.25)	26.00 (22.00–30.00)	< 0.0001	24.00 (21.00–28.00)	25.00 (21.00–28.75)	0.339
Acanthosis nigricans (%)[b]	0.0	19.5	< 0.0001	7.2	38.7	< 0.0001	0.0	100.0	< 0.0001
BMI [kg/m^2][a]	21.92 (19.42–23.75)	27.54 (22.032–33	< 0.0001	23.00 (20.82–26.17)	34.064 (31.90–38.00)	< 0.0001	25.87 (21.37–31.25)	34.00 (30.65–38.12)	< 0.0001
Obesity (%)[b]	0.0	40.3	< 0.0001	0.0	100.0	NA	31.1	79.7	< 0.0001
Waist/Hip ratio[a]	0.75 (0.69–0.785)	0.84 (0.78–0.9)	< 0.0001	0.8 (0.76–0.85)	0.88 (0.842–0.929)	< 0.0001	0.83 (0.77–0.89)	0.87 (0.832–0.933)	< 0.0001
LH/FSH	0.62 (0.50–0.90)	1.69 (1.06–2.53)	< 0.0001	1.90 (1.19–2.75)	1.41 (0.91–2.28)	0.0004	1.73 (1.03–2.55)	1.56 (1.17–2.49)	0.742
Testosterone [ng/dL][a]	42 (30–50)	75 (60–92)	< 0.0001	76.00 (61.62–92.00)	75.00 (55.00–91.5)	0.495	74.5 (60.0–90.0)	79.5 (59.0–102.0)	0.098
DHEAS [μg/dL][a]	194.5 (155.4–329.6)	254.9 (181.8–334)	0.212	254.7 (188.8–335.4)	257.75 (157–330.7)	0.757	251.7 (171.4–339.7)	283.9 (217.2–316.4)	0.250
Fasting glucose [mmol/L]	4.65 (4.37–5.10)	4.76 (4.44–5.15)	0.454	4.70 (4.42–5.04)	4.84 (4.50–5.26)	0.040	4.72 (4.42–5.10)	4.97 (4.55–5.32)	0.011
2–h glucose[a] [mmol/L]	5.88 (4.76–6.41)	5.71 (4.93–6.83)	0.879	5.46 (4.70–6.16)	6.30 (5.32–7.45)	< 0.0001	5.56 (4.84–6.50)	6.66 (5.55–7.67)	< 0.0001
Fasting insulin [mU/mL][a]	6.49 (4.49–8.61)	14.8 (8.87–23.32)	< 0.0001	11.87 (6.94–18.95)	21.25 (13.4–31.95)	< 0.0001	13.6 (8.1–21.6)	22.0 (15.0–31.3)	< 0.0001
HOMAIR[a]	1.33 (0.91–2.13)	3.24 (1.1.92–5.54)	< 0.0001	2.63 (1.48–4.20)	4.831 (3.0–7.59)	< 0.0001	3.05 (1.78–4.96)	4.93 (3.07–7.64)	< 0.0001
HOMAb[a]	120 (70.94–158.98)	240.2 (141.8–414.5)	< 0.0001	195.5 (122.0–345.8)	308.8 (184.8–621.2)	< 0.0001	224.7 (134.4–396.3)	274.8 (176.0–567.1)	0.0095
Total cholesterol [mmol/L][a]	4.91 (4.30–5.63)	4.94 (4.37–5.64)	0.742	4.91 (4.34–5.58)	4.98 (4.47–5.69)	0.106	4.86 (4.37–5.60)	5.20 (4.47–5.73)	0.064
Triglicerides [mmol/L][a]	0.92 (0.65–1.20)	1.02 (0.71–1.51)	0.124	0.84 (0.61–1.17)	1.33 (1.03–1.76)	< 0.0001	0.91 (0.67–1.33)	1.61 (1.14–1.94)	< 0.0001

Table II. *Techniques and primers used for genotyping SNPs:* **A.** *rs3842756, rs3842755, rs3842754, rs3842753, and rs3842752 in the first portion of the population (160 women with PCOS and 95 controls);* **B.** *rs3842754, rs3842748, and rs689 in the rest of the population (the whole population included 377 women with PCOS and 105 controls)*

Tabela 2. *Techniki i primery zastosowane w procesie genotypowania SNP:* **A.** *rs3842756, rs3842755, rs3842754, rs3842753 i rs3842752 w pierwszej części populacji (160 kobiet cierpiących na PCOS i 95 osób kontrolnych);* **B.** *rs3842754, rs3842748 i rs689 w pozostałej części populacji (cała populacja wynosiła 377 kobiet z PCOS i 105 kobiet w grupie kontrolnej)*

A.

SNPs	Technique of genotyping	Used primers	Steps of PCR
rs3842756 rs3842755 rs3842754 rs3842753 rs3842752	Sequencing (GoldStart® DNA polymerase)	Forward primer 5′-ACT GTG TCT CCC TGA CTG TGT C-3′ Reverse primer 5′- GAC TCC AAG AGT CCA GAG CTA C-3′	95°C (8′), 35 cycles of [94°C (15″), 60,8°C (10″), 72°C (45″)], and 72°C (6′)
rs3842748	Sequence specific oligonucleotide-PCR (SSO-PCR)	Common reverse primer 5′-CTC ACG GCA GCT CCA TAG TC-3′ Forward primers 5′-AGA GGG AGG GTC ACC CAC AC-3′ and 5′-AGA GGG AGG GTC ACC CAC AG-3′	95°C (8′), 35 cycles of [94°C (15″), 56.6°C (10″), 72°C (45″)], and 72°C (6′)
rs689	sequencing (GoldStart® DNA polymerase)	Forward primer 5′-CTT GGG TGT GTA GAA GAA GC-3′ Reverse primer 5′-AGC AGG TCT GTT CCA AGG-3′	95°C (8′), 35 cycles of [94°C (15″), 56,6°C (10″), 72°C (45″)], and 72°C (6′)
rs689	Sequence specific oligonucleotide-PCR (SSO-PCR)	Forward primers 5′-CAG AAG GAC AGT GAT CTG GGT-3′ and 5′-CAG AAG GAC AGT GAT CTG GGA-3′ Reverse primer 5′-TCC AGG ACA GGC TGC ATC AG-3′	95°C (8′), 35 cycles of [94°C (15″), 56,6°C (10″), 72°C (45″)], and 72°C (6′)

B.

SNPs	Technique of genotyping	Used primers	Steps of PCR
rs3842754	KASPar technology	Forward specific primers: 5′-CCC AAA GCG GCC ATG CCT GTT-3′ and 5′-CCA AAG CGG CCA TGC CTG TC-3′ Common reverse primer: 5′-CCA TAC TGG ACC CTG AGC CAC A-3′	94°C (15′), 10 cycles of [94°C (20″), 61°C (1′), drop annealing temperature 0.6°C per cycle to reach 55°C in the last cycle], 26 cycles of [94°C (10″), 55°C (1′)], and recycling [94°C (20″), 57°C (1′)] as many times as required
rs3842748	KASPar technology	Allele specific primers 5′-GCC TGT AGG TCC ACA CCC AC-3′ and 5′-GCC TGT AGG TCC ACA CCC AG-3′ Common reverse one 5′-GTT AGA GGG AGG GTC ACC CAC A-3′	94°C (15′), 10 cycles of [94°C (20″), 61°C (1′), drop annealing temperature 0.6°C per cycle to reach 55°C in the last cycle], 26 cycles of [94°C (10″), 55°C (1′)], and recycling [94°C (20″), 57°C (1′)] as many times as required
rs689	KASPar technology	Allele specific primers 5′-GCC TCA GCC CTG CCT GTC A-3′ and 5′-GCC TCA GCC CTG CCT GTC T-3′ Common reverse primer 5′-CAT CCA CAG GGC CAT GGC AGA A-3′	94°C (15′), 10 cycles of [94°C (20″), 61°C (1′), drop annealing temperature 0.6°C per cycle to reach 55°C in the last cycle], 26 cycles of [94°C (10″), 55°C (1′)], and recycling [94°C (20″), 57°C (1′)] as many times as required

Genotypic data

Polymorphism analysis in part of the population (160 women with PCOS and 95 controls) showed that SNPs rs3842756, rs3842755, rs3842754, rs3842753, rs3842752, rs3842748, and rs689 were in high LD according to the D′ index (Fig. 1). However, three groups of SNPs were observed according to the level of high LD between them (according to r^2 LD index): group 1 (rs3842756, rs3842755, rs3842754, rs3842752), group 2 (rs3842753 and rs689), and group 3 (rs3842748). Thus, we considered SNPs rs3842754, rs689, and rs3842748 to be representative for the first, the second, and the third group, respectively. These SNPs were genotyped in the rest of the population. Frequencies of SNPs rs3842754, rs3842748, and rs689 in the control group were 0.17, 0.83, and 0.75, respectively, which are very close to their frequencies in other Caucasian populations (http://www.ncbi.nlm.nih.gov/snp/). Reconstruction of haplotypes was done using the three SNPs allowed observing six haplotypes: h1 (CCA), h2 (CCT), h3 (CGA), h4 (CGT), h5 (TCA), and h6 (TGT) (Table III). It appeared that the three haplotypes h1, h2, and h3 were frequent while the

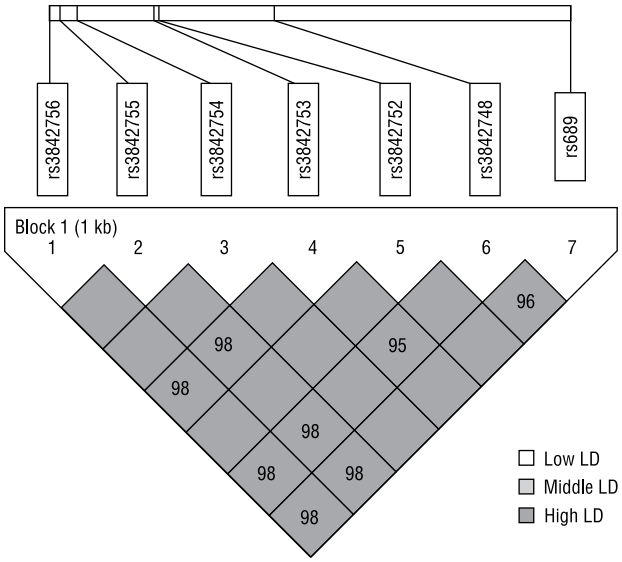

Figure 1. *Linkage disequilibrium (LD) pattern (estimated by D′ index) in the region of insulin gene between rs3842756 and rs689, in the female population (cases and controls) from Central Europe. The numbers in red squares indicate the D′ index (%). The squares without numbers indicate a perfect LD (D′ = 100% or 1) between the corresponding SNPs*

Rycina 1. *Wzór niezrównoważenia sprzężeń (LD) (wyliczony przez indeks D′) w regionie genu insuliny między rs3842756 i rs689 u populacji kobiet (przypadki chorobowe i osoby kontrolne) z Europy Centralnej. Liczby w czerwonych kwadratach oznaczają indeks D′ (%). Kwadraty bez liczb oznaczają idealne LD (D′ = 100% lub 1) między odpowiednimi SNP*

three others were rare. When we started from the root of the cladogram (h4), the mutation of rs3842754 from C to T allowed us to observe h6 (surrogate of VNTR of subclass IIIA), while mutation of rs3842748 from G to C allowed us to reach the rest of the cladogram (Fig. 2). Mutation of SNP rs689 from T to A led to h1 (VNTR of class I). This data is in accordance with literature [21].

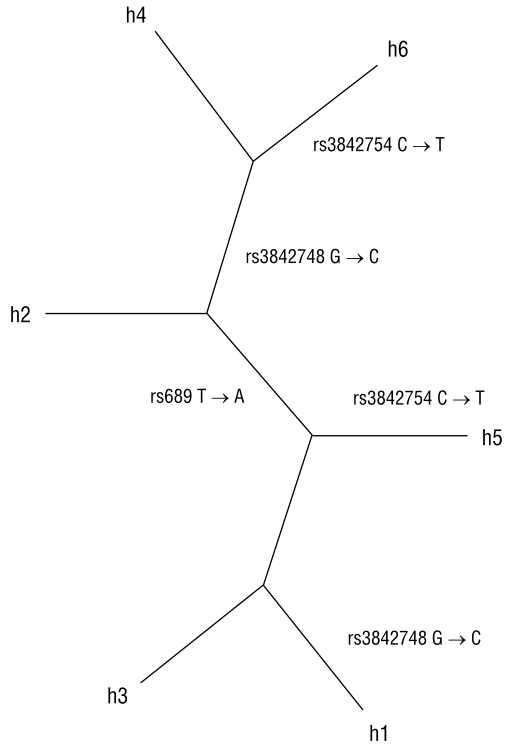

Figure 2. *Phylogenetic (cladistics) affiliation of SNPs rs3842754, rs3842748, and rs689 illustrating the historical evolution of haplotype diversity in the female population (cases and controls) from Central Europe used in the study. The ancestral (root) haplotype (CGT) is h4. The haplotype h6 (TGT) is due to the mutation (C → T) of the SNP rs3842754. The haplotype h2 (CCT) derived from h4 after the mutation (G → C) of the SNP rs3842748, while the haplotype h1 (CCA) is due to the mutation of both rs689 (T → A) and rs3842748 (G → C)*

Rycina 2. *Związek filogenetyczny (kladystyka) SNP: rs3842754, rs3842748 i rs689 ilustrujący ewolucję różnorodności haplotypów w populacji kobiet (przypadki chorobowe i osoby kontrolne) z Europy Centralnej, wykorzystane w badaniu. Haplotyp ancestralny (CGT) to h4. Haplotyp h6 (TGT) powstał w wyniku mutacji (C → T) SNP rs3842754. Haplotyp h2 (CCT) uzyskano z h4 po mutacji (G → C) SNP rs3842748, podczas gdy haplotyp h1 (CCA) powstał w wyniku mutacji zarówno rs689 (T → A), jak i rs3842748 (G → C)*

Table III. ***List of haplotypes resulting from combination of SNPs rs3842754, rs3842748, and rs689 in the population of women (cases and controls) from Central Europe used in the study, and their prevalence. Mutated alleles of SNPs in each haplotype are underlined. Ancestral alleles are those of chimpanzee (Pan troglodytes)***

Tabela III. ***Lista haplotypów otrzymanych z kombinacji SNP: rs3842754, rs3842748 i rs689 w populacji kobiet (przypadki chorobowe i osoby kontrolne) z Europy Centralnej, wykorzystane w badaniu, a także ich prewalencja. Zmutowane allele SNP zostały podkreślone w każdym z haplotypów. Allele ancestralne należą do szympansów (Pan troglodytes)***

Haplotype	rs3842754 (C/T)	rs3842748 (G/C)	rs689 (T/A)	Number of alleles	Prevalence (%)
h1	C	C	A	693	71.89
h2	C	C	T	96	9.96
h3	C	G	A	3	0.31
h4	C	G	T	3	0.31
h5	T	C	A	3	0.31
h6	T	G	T	166	17.22

Table IV. *Resume of positive genetic association between polymorphisms in the insulin gene and obesity in PCOS (all PCOS group or only women with classical PCOS "YPCOS") as well as the power of the study for each genetic association. Assumptions for calculating the power are: the size of control and cases group, the odd ratio, the percentage of the genotype among the control group, the ratio control group/cases group, and the alpha risk = 5%. The genetic association was attested as significant for P < 0.05*

Tabela IV. *Podsumowanie pozytywnego związku genetycznego między polimorfizmami w genie insuliny i otyłością w PCOS (cała grupa z PCOS lub tylko kobiety z klasycznym PCOS "YPCOS"), jak również siłą badania dla każdego powiązania genetycznego. Założenia dla wyliczenia siły są następujące: wielkość grupy kontrolnej i grupy przypadków chorobowych, współczynnik szans, procent genotypu w grupie kontrolnej, stosunek grupy kontrolnej do grupy przypadków chorobowych oraz współczynnik alfa = 5%. Związek genetyczny został potwierdzony jako istotny dla P < 0,05*

The genetic association	The phenotype associated with	Odds ratio 95% CI	p	Control group	Case group	% of the genotype among the control group	Power the association (%)
GC genotype of rs3842748	Obesity in YPCOS	2.01 (1.18–3.42)	0.010	YPCOS and lean (n = 212)	YPCOS with obesity (n = 93)	22.27	72.85
The pair of haplotypes h1/h6	Obesity in YPCOS	2.08 (1.20–3.61)	0.009	YPCOS and lean (n = 212)	YPCOS with obesity (n = 93)	19.34	73.95
GC genotype of rs3842748	Obesity in YPCOS	1.59 (1.00–2.51)	0.049	PCOS and lean (n = 228)	PCOS with obesity (n = 148)	23.50	51.19

We observed that the genotype GC of rs3842748 was associated with the obese phenotype in PCOS (P = 0.049) and in the classical PCOS (YPCOS) group (P = 0.010) (Table IV). The association with obesity in YPCOS remained after adjustment for age. The pair of haplotypes h1/h6 were associated with the obese phenotype in classical PCOS (P = 0.009) and remained also after adjustment for age. The data of association with obesity was supported by the correlation between the pair of haplotypes h1/h6 and higher BMI in the YPCOS group (h1/h6: P = 0.048, 28.5 ± 0.9 *vs.* 26.3 ± 0.4) and a trend for correlation between the GC genotype of rs3842748 and higher levels of BMI in the YPCOS group (P = 0.077).

SNPs and their haplotypes were also correlated with plasmatic levels of androgens in PCOS group. Indeed, we observed the correlation of rs689, through its AA genotype and the pair of haplotypes h1/h1, with higher levels of testosteronaemia, but this was at the limit of significance (rs689 (AA): P = 0.054, 87.1 ± 4.2 ng/dL *vs.* 74.9 ± 2.0 ng/dL; haplotype h1/h1: P = 0.054, 86.9 ± 4.1 ng/dL *vs.* 74.9 ± 2.0 ng/dL). Moreover, rs3842754 and rs3842748, through their alleles C and C, respectively, were correlated with lower plasmatic levels of DHEAS in PCOS women with obesity (rs3842754 (C): P = 0.041, 260.7 ± 16.9 μg/dL *vs.* 432.9 ± 61.6; rs3842748 (C): P = 0.041, 260.7 ± 16.9 *vs.* 432.9 ± 61.6) (Fig. 3). Note that the power of the study concerning each positive genetic association is reported in Table IV.

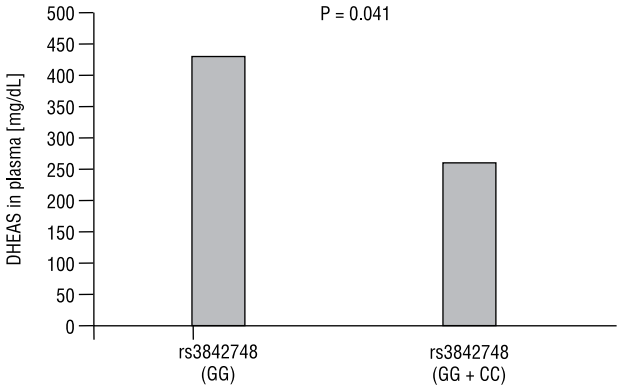

Figure 3. *Graphic comparing dehydroepiandrosterone sulphate (DHEAS) levels between individuals carrying the allele C of the SNP rs3842748 and the others. DHEAS levels are indicated by mean ± standard error. The significance of the difference is determined by the Mann-Whitney test. P is considered significant when < 0.05*

Rycina 3. *Graficzne porównanie stężenia siarczanu dehydroepiandrosteronu (DHEAS) pomiędzy nosicielami allela C SNP rs3842748 i innymi. Stężenie DHEAS oznaczono jako średnią ± błąd standardowy. Istotność różnicy statystycznejustalana jest na podstawie testu Manna-Whitneya. P uznano za istotne dla wartości P < 0,05*

Discussion and conclusions

PCOS is a complex disorder characterised by abnormalities related to reproduction and others with metabolism and glucose homeostasis, including insulin secretion abnormalities. Indeed, women with PCOS were reported as having a decrease in insulin secretion [5]. Insulin gene minisatellite (VNTR), which was implicated in the regulation of insulin gene expression, is a potential candidate in PCOS. Different investigations were carried out and gave controversial results [6–9]. Some of these studies focused on VNTR itself while others used rs689 as a surrogate since this SNP has two alleles in high LD with VNTR of Class I and III, respectively,

in non-African populations. Moreover, studies used different definitions for diagnosing PCOS. One could speculate that the heterogeneity between investigations is the cause of controversy concerning the association of insulin gene VNTR with PCOS.

In order to get more insight we investigated the association of insulin gene VNTR with PCOS using dense mapping of the genome and reconstruction of haplotypes. The Rotterdam Conference criteria [27] were used for diagnosing PCOS, giving an accurate and precise definition of the disease. All this accuracy was aimed to give an ideal context for determining the relation between the minisatellite and PCOS.

The SNPs taken into account in this study have been demonstrated to combine into haplotypes in high LD with classes and subclasses of VNTR, which allowed them to be used as surrogates. Their genotyping was carried-out in a portion of the population, with the aim of determining the LD pattern and selecting from them those that tag the region (*tag*SNP = representative of a group of SNPs in high LD). Three *tag*SNPs were then selected and genotyped in the rest of the population, including rs689. In addition to its ability to mimic the variation of the insulin gene minisatellite, this SNP was implicated in the regulation of insulin gene expression through an alternative splicing process [19, 20].

We observed the correlation of the genotype AA of rs689 and the pair of haplotypes h1/h1 with higher levels of testosteronaemia in PCOS and with lower plasmatic levels of DHEAS in PCOS with obesity, which highlights once again the close relation between the reproductive and the metabolic components in PCOS, as already reported many times in literature (Diamanti-Kandarakis. 2008).

Moreover, the correlation between the gene and lower plasmatic levels of DHEAS in our study is interesting and in accordance with literature. Indeed, previous studies reported negative correlation between DHEAS levels in plasma and $HOMA_{IR}$ index (reflects fasting insulin) in PCOS [37], as well as high levels of DHEAS being correlated with lower insulin levels and more favourable metabolic and cardiovascular parameters [38]. Our finding may give more insight into understanding the genetic background that controls variations of DHEAS.

Moreover, the genotypes GC of rs3842748 and h1/h6 (corresponding to the genotype I/III of VNTR) were associated with obesity and correlated with higher BMI in PCOS (especially PCOS with lower insulin resistance), which represents a novelty in our study. Note that this genotype has been previously associated in T1D with glutamic acid decarboxylase antibodies (GADA) [39] but never with T2D, PCOS, or obesity. At the same time, the minisatellite of class III, widely associated in literature, with PCOS and T2D [6, 12] displayed weak results in our investigation. Perhaps the situation is more complicated than expected. Indeed, locus involving insulin gene VNTR involves IGF-2 (MIM: 147470) and tyrosine hydroxylase (MIM: 191290) genes. These two genes were associated, as well as the minisatellite, with metabolic parameters [40]. Combining the variation of these three genes may ameliorate the genetic study and determine subgroups of VNTR of class III other than IIIA and IIIB (determined by our haplotypes). One could speculate that the heterozygote genotype I/III of VNTR in our population encompasses a subgroup of VNTR class III related to obesity. The association of VNTR I/III with obesity could be due also to VNTR of class I. Here also the context seems to be complex. Indeed, Le Stunff et al. observed an excess of paternal transmission of VNTR of class I to obese children [13] and other studies attested to the role of genetic imprinting occurring in the INS-IGF2 locus at the age of onset of monogenic diabetes caused by INS gene mutations [41]. Future studies concerning this locus should take into account this phenomenon.

Criticism could be addressed concerning the size of the explored population, especially in the context of current genetic association studies exploiting populations of thousands of individuals, with the aim of getting statistical power. Such studies may display weakness at the phenotypic level in the sense that large recruitment may lead to less stringency, contaminating the phenotypic pattern with inappropriate clinical backgrounds, which may lead to spurious associations. The power of our case-control study was acceptable (72.85, 73.95, 51.19% for the three positive associations obtained, respectively), and frequencies of SNPs in our control population were close to those in Caucasian populations [21], which attests to the reliability of the study.

The present investigation was able to detect association of insulin gene VNTR with metabolic components in PCOS, but it was not able to highlight any consistent advantage for using haplotypes instead of SNPs taken alone. This needs to be deepened by extending the explored region upstream and downstream through the screening of much more SNPs. This may allow us to observe that haplotypes carry more genetic information than SNPs taken alone. It is expected to carry out future studies in the context of the European Project MEDIGENE FP7 (FP7-279171-1) with deeper screening of the 50 kb locus involving VNTR of insulin gene, IGF-2, and TH genes for a better understanding of the relation between the locus and complex metabolic disorders. Indeed, we believe that genes reported as not associated with metabolic complex disorders by GWAS studies (as insulin gene VNTR) even though they were previously associated using the candidate gene approach is due to

a lack of genome characterisation by GWAS. Using hundred of thousands of SNPs (six KSNPs to about a million SNPs) for screening the whole genome is not enough, knowing the existence of at least 27 million SNPs and rare variants with potential effects. One could also criticise our study since we focused on the insulin secretion component without taking into account potential genes such as TCF7L2 (MIM: 602228). It is expected in future that genotyping, in our population, of other potential genes with a role in insulin secretion, notably TCF7L2 and study the interaction of the insulin gene locus with this panel of genes will take place.

Acknowledgments

This work was supported by the recurrent budget from the University of Montpellier (France), and University Medicine and Pharmacy "Carol Davila" in Bucharest (Romania). RA, FG, and MC are sustained by the grant MEDIGENE FP7-279171 of the European community (EC). Special thanks are also due to Dr. Mouna Houari-Barat and Mr. Bruno Dumont for their technical help and also to Prof. Jean-Pierre Daurès, Dr. Mohamed Belkacemi and Charles for their advice in statistics.

References

1. Polson DW, Adams J, Wadsworth J et al. Polycystic ovaries--a common finding in normal women. Lancet 1988; 1: 870–872.
2. Franks S. Polycystic ovary syndrome. N Engl J Med 1995; 333: 853–861.
3. Dunaif A, Givens JR, Haseltine FP et al. (eds.). Polycystic ovary syndrome. Oxford: Blackwell Science, 1992.
4. Dunaif A, Segal KR, Futterweit W et al. Profound peripheral insulin resistance, independent of obesity, in polycystic ovary syndrome. Diabetes 1989; 38: 1165–1174.
5. Dunaif A. Insulin resistance and the polycystic ovary syndrome: mechanism and implications for pathogenesis. Endocr Rev 1997; 18: 774–800.
6. Waterworth DM, Bennett ST, Gharani N et al. Linkage and association of insulin gene VNTR regulatory polymorphism with polycystic ovary syndrome. Lancet 1997; 349: 986–990.
7. Michelmore KF, Ong K, Mason S et al. Clinical features in women with polycystic ovaries: relationships to insulin sensitivity, insulin gene VNTR and birth weight. Clin Endocrinol 2001; 55: 439–446.
8. Powell BL, Haddad L, Bennett A et al. Analysis of multiple data sets reveals no association between the insulin gene variable number tandem repeat element and polycystic ovary syndrome or related traits. J Clin Endocrinol Metab 2005; 90: 2988–2993.
9. Vankova M, Vrbíková J, Hill M et al. Association of insulin gene VNTR polymorphism with polycystic ovary syndrome. Ann N Y Acad Sci 2002; 967: 558–565.
10. Bell GI, Horita S, Karam JH. A polymorphic locus near the human insulin gene is associated with insulin-dependent diabetes mellitus. Diabetes 1984; 33: 176–83.
11. Barratt BJ, Payne F, Lowe CE et al. Remapping the insulin gene/IDDM2 locus in type 1 diabetes. Diabetes. 2004; 53: 1884–1889.
12. Bennett ST, Todd JA. Human type 1 diabetes and the insulin gene: principles of mapping polygenes. Annu Rev Genet 1996; 30: 343–370.
13. Le Stunff C, Fallin D, Bougnères P. Paternal transmission of the very common class I INS VNTR alleles predisposes to childhood obesity. Nat Genet 2001; 29: 96–99.
14. Bazaes RA1, Petry CJ, Ong KK et al. Insulin gene VNTR genotype is associated with insulin sensitivity and secretion in infancy. Clin Endocrinol (Oxf) 2003; 59: 599–603.
15. Dos Santos C, Fallin D, Le Stunff C et al. INS VNTR is a QTL for the insulin response to oral glucose in obese children. Physiol Genomics 2004; 16: 309–313.
16. Kennedy GC, German MS, Rutter WJ. The minisatellite in the diabetes susceptibility locus IDDM2 regulates insulin transcription. Nat Genet 1995; 9: 293–298.
17. Paquette J, Giannoukakis N, Polychronakos C et al. The INS 5′ variable number of tandem repeats is associated with IGF2 expression in humans. J Biol Chem 1998; 273: 14158–14164.
18. Stead JD, Jeffreys AJ. Structural analysis of insulin minisatellite alleles reveals unusually large differences in diversity between Africans and non-Africans. Am J Hum Genet 2002; 71: 1273–1284.
19. Kralovicova J, Gaunt TR, Rodriguez S et al. Variants in the human insulin gene that affect pre-mRNA splicing: is -23HphI a functional single nucleotide polymorphism at IDDM2? Diabetes 2006; 55: 260–264.
20. Marchand L, Polychronakos C. Evaluation of polymorphic splicing in the mechanism of the association of the insulin gene with diabetes. Diabetes 2007; 56: 709–713.
21. Stead JD, Hurles ME, Jeffreys AJ. Global haplotype diversity in the human insulin gene region. Genome Res 2003; 13: 2101–2111.
22. Le Fur S, Auffray C, Letourneur F et al. Heterogeneity of class I INS VNTR allele association with insulin secretion in obese children. Physiol Genomics 2006; 25: 480–484
23. Bennett ST, Lucassen AM, Gough SC et al. Susceptibility to human type 1 diabetes at IDDM2 is determined by tandem repeat variation at the insulin gene minisatellite locus. Nat Genet 1995; 9: 284–292.
24. Bouatia-Naji N, De Graeve F, Brönner G et al. INS VNTR is not associated with childhood obesity in 1,023 families: a family-based study. Obesity (Silver Spring) 2008; 16: 1471–1475.
25. . Eeles RA, Kote-Jarai Z, Al Olama AA et al. Identification of seven new prostate cancer susceptibility loci through a genome-wide association study. Nat Genet 2009; 41: 1116–1121.
26. Collins A, Lonjou C, Morton NE. Genetic epidemiology of single-nucleotide polymorphisms. Proc Natl Acad Sci USA 1999; 96: 15173–15177.
27. Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. Fertil Steril 2004; 81: 19–25.
28. Hanzu FA, Radian S, Attaoua R et al. Association of insulin receptor genetic variants with polycystic ovary syndrome in a population of women from Central Europe. Fertil Steril 2010; 94: 2389–2392.
29. American Diabetes Association. Diagnosis and classification of diabetes mellitus. Diabetes Care 2006; Suppl. 1: S43–S48.
30. Attaoua R, Ait El Mkadem S, Radian S et al. FTO gene associates to metabolic syndrome in women with polycystic ovary syndrome. Biochem Biophys Res Commun 2008; 373: 230–234.
31. Cuppen E. Genotyping by Allele-Specific Amplification (KASPar). CSH Protoc: pdb.prot4841. 2007
32. Barrett JC, Fry B, Maller J et al. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005; 21: 263–265.
33. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 2001; 68: 978–989.
34. Schneider S, Roessli D, Excoffier L. Arlequin ver. 2.000: A software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland, 2000.
35. Lautier C, Ait El Mkadem SA, Renard E et al. Complex haplotypes of IRS2 gene are associated with severe obesity and reveal heterogeneity in the effect of Gly1057Asp mutation. Hum Genet 2003; 113: 34–43.
36. Diamanti-Kandarakis E. Polycystic ovarian syndrome: pathophysiology, molecular aspects and clinical implications. Expert Rev Mol Med 2008; 10: e3.
37. Brennan K, Huang A, Azziz R. Dehydroepiandrosterone sulfate and insulin resistance in patients with polycystic ovary syndrome. Fertil Steril 2009; 91: 1848–1852.
38. Carmina E, Lobo RA. Prevalence and metabolic characteristics of adrenal androgen excess in hyperandrogenic women with different phenotypes. J Endocrinol Invest 2007; 30: 111–116.
39. Rajasalu T1, Haller K, Salur L et al. Insulin VNTR I/III genotype is associated with autoantibodies against glutamic acid decarboxylase in newly diagnosed type 1 diabetes. Diabetes Metab Res Rev 2007; 23: 567–571.
40. Gu D, O'Dell SD, Chen XH et al. Evidence of multiple causal sites affecting weight in the IGF2-INS-TH region of human chromosome 11. Hum Genet 2002; 110: 173–178.
41. Fendler W, Borowiec M, Antosik K et al. Paternally inherited proinsulin mutations may result in earlier onset of monogenic diabetes mutation identity effect in monogenic diabetes. Diabetes Care 2011; 34: e9.