

Distinct Genetic Susceptibility Patterns of the Obese and Non-obese South Indian Women with Polycystic Ovary Syndrome

B. Mohan Reddy^{1,2#*} and Kumuda Irgam¹

¹Department of Genetics and Biotechnology, Osmania University, Hyderabad, India

²Molecular Anthropology Group, Indian Statistical Institute, Hyderabad, India

KEYWORDS Body Mass Index. Complex Disease. Genetic Susceptibility. Metabolic and Reproductive Pathways. Phenotypic Heterogeneity

ABSTRACT The researchers studied association of 92 SNPs of the metabolic and reproductive pathways genes in the obesity related cohorts of obese PCOS cases and non-obese controls, and non-obese PCOS cases and non-obese controls and observed eleven SNPs to be significantly associated ($p \leq 0.05$). While eight of those SNPs, five from FTO and one each from MTCH2, DENNDIA and THADA genes, were found to be associated in the first set involving obese PCOS cases, only three SNPs (IRS2-intronic, LOC107984901-intergenic and SUMO1P1-regulatory variants) showed association in the 2nd set involving non-obese cases. The SNPs associated in the two cohorts were distinct and mutually exclusive. However, all the SNPs associated in either obese or non-obese PCOS cohorts were not only risk-prone but also broadly represented a metabolic pathway, involving mostly obesity and T2DM related genes, prompting one to surmise if the recent spurt in PCOS prevalence is not driven by changes in the lifestyles.

INTRODUCTION

Concurrent to the association of Polycystic Ovary Syndrome (PCOS) with several co-morbidities, obesity has been recognised from the beginning as a common feature of this syndrome. Its consistent association with PCOS may suggest biological basis (Sam 2007), as implicated in the fact that obesity was observed to exacerbate many of the reproductive and metabolic abnormalities associated with PCOS (Kiddy et al. 1992; Holte et al. 1995; Taponen et al. 2003; Barber et al. 2006; Sam 2007). Therefore, it is possible that similar or interrelated genes of PCOS may also predispose obesity in the affected women (Barber et al. 2006). One way forward is to design and focus on the obesity specific association studies of PCOS and unfortunately there have not been any hitherto. In this context, it may be pertinent to mention that in a recent paper based on the results of a large scale

study (Irgam et al. 2019), the researchers have analysed a comprehensive set of 92 SNPs from a wide spectrum of variants belonging to reproductive and metabolic pathway genes among the PCOS women from Hyderabad and observed only thirteen of the 92 SNPs, representing six genes, to have been significantly associated with PCOS ($p < 0.05$). Out of those 13 SNPs, five belong to FTO gene and by far the most significant ($P = 0.002$) of the 13 SNPs was observed to be rs10838738 of MTCH2, both these genes have a putative role in high body mass index (BMI) and general obesity (Willer et al. 2009). Further, after adjusting for BMI as covariate, the researchers found twelve of the 13 SNPs excepting rs1421085 of FTO gene to lose significance, suggesting possible confounding nature of the effect of obesity as phenotype or through its quantitative trait loci. Therefore, in the process of identification of the primary causes of the PCOS and its underlying molecular mechanisms, it is imperative to re-examine the patterns of genetic association of these 92 SNPs in the obesity related cohorts of PCOS samples so that the patterns of genetic association of obese and non-obese PCOS women can be ascertained. The researchers present here the salient features of the findings based on the analyses of the above 92 SNPs in the obesity specific subsets of the pooled cohort of obese PCOS cases and non-

Address for correspondence:

B. Mohan Reddy, Ph.D.

Professor

ICMR Emeritus Scientist

Department of Genetics & Biotechnology

Osmania University

Hyderabad 500 007, India

Mobile: 9866289771,

E-mail: bmrissi@gmail.com

obese controls, and non-obese PCOS cases and non-obese controls.

MATERIAL AND METHODS

The current study is a part of the major project entitled, "Identification of susceptibility genes associated with PCOS among the South Indian women" initiated during 2007 by the corresponding author when he was in the Molecular Anthropology Group of Biological Anthropology Unit, Indian Statistical Institute, Hyderabad, for which 250 primary PCOS cases and 299 controls were enrolled during 2008-2009. The patients were recruited from the gynaecology clinic of the Osmania General Hospital and Anus Infertility Clinic from Hyderabad, India, as per the Rotterdam criteria (2003). Metabolic and hormonal parameters, epidemiological data and clinical information have been collected from each subject along with ~5 ml of intravenous blood samples, which were drawn by a trained laboratory technician. Blood samples and the background information were collected after obtaining informed written consent prior to their enrolment in the study. The study protocol was approved by the Indian Statistical Institute Review Committee for Protection of Research Risks to Humans. Further details on the sampling procedure and the inclusion and exclusion criteria for the cases and controls, laboratory methods of DNA isolation and genotyping along with the list of 92 SNPs and the primer sequences for each of those were already reported in a recent paper (Irgam et al. 2019).

Out of the total sample of 250 cases and 299 controls, the researchers had BMI data for 167 cases and 232 controls. Further, thirty-seven of the 232 controls were found to be obese, hence excluded from the analyses to avoid the possible effect of confounding. Finally, for the present study, the 167 cases and 195 controls were distributed into two subsets of obese and non-obese cases along with non-obese controls and all the relevant statistical analyses were performed on each of the two subsets. The statistical methods and packages used for the analyses of these subsets were the same as in our previous publication (Irgam et al. 2019).

RESULTS

The Nature of Allelic and Genotypic Association of SNPs in the Obese and Non-obese PCOS Subsets

After excluding the SNPs that showed minor allele frequency of less than one percent and/or deviated from Hardy Weinberg Equilibrium ($p < 0.001$), the researchers were left with 75 and 73 SNPs respectively, of the total 92 SNPs for further analyses of the two subsets. The results of logistic regression analyses of the alleles for the two subsets are furnished in Tables 1 and 2. The allelic association analyses of the obese (Table 1) and non-obese (Table 2) PCOS subsets resulted in only 8 and 3 SNPs respectively, as significantly associated with PCOS ($p \leq 0.05$), albeit none after correction for multiple testing suggesting minor nature of the effect of these variants/genes, adjusting for age and socioeconomic status as covariates did not alter the results of association (results not presented). However, it is pertinent to note that all the SNPs associated in both the subsets were risk prone and belong mostly to the metabolic pathway genes related to obesity (FTO, MTCH2), clathrin-mediated endocytic pathway (DENNDIA), apoptosis (THADA) and type 2 diabetes (IRS2), besides the other two with unknown function (LOC107984901, SUMO1P1). While all the eight associated SNPs in the obese subset and one of the 3 in the non-obese subset (IRS2 SNP) were from intronic regions of the respective genes, the remaining two SNPs are from the intergenic and regulatory regions. The results of the genotypic association analyses (Tables 3 and 4) are complementary to that of the allelic association patterns and based on the lowest AIC (Akaike Information Criteria) value, the log additive model was found to be the best fit in case of all the associated SNPs of the obese subset and in case of the non-obese subset, while the IRS2 SNP fits best under log additive model the other two SNPs show best fit under dominant model.

SNP-SNP Interactions, Linkage Disequilibrium and Haplotype Association

Despite relatively small sample sizes for the subsets the researchers have in an exploratory

Table 1: Allelic association of SNPs in the obese PCOS and non-obese controls with chi square and odds ratio values from logistic regression analysis

Gene / nature of SNP	SNP rs ID	Major/ Minor allele	Minor allele frequency		χ^2	OR (95% CI)	p-value	Function of the gene
			Obese PCOS (N=104)	Non-obese controls (N=195)				
MTCH2 Intron	rs10838738	A/G	0.42	0.29	9.94	1.75 (1.23-2.49)	0.001	Adipocyte differentiation
FTO	rs1421085	T/C	0.40	0.29	7.89	1.65 (1.16-2.34)	0.004	Involves in DNA damage reversal pathway and
Introns	rs8050136	C/A	0.37	0.27	6.68	1.59 (1.11-2.28)	0.009	FTO obesity variant
	rs17817449	T/G	0.37	0.27	6.33	1.57 (1.04-2.13)	0.011	
	rs9939609	T/A	0.37	0.28	4.90	1.49 (1.01-2.01)	0.026	
	rs9940128	G/A	0.46	0.37	4.23	1.42 (1.10-2.27)	0.039	
DENND1A Intron	rs2479106	A/G	0.36	0.26	6.28	1.58 (1.10-2.27)	0.012	Vesicle-mediated transport and RAB GEFs exchange GTP for GDP on RABs
THADA Intron	rs12478601	T/C	0.43	0.35	4.01	1.42 (1.0-2.0)	0.045	Apoptosis

Table 2: Allelic association of SNPs in the non-obese PCOS and non-obese controls with chi square and odds ratio values from logistic regression analysis

Gene / nature of SNP	SNP rs ID	Major/ Minor allele	Minor allele frequency		χ^2	OR (95% CI)	p-value	Function of the gene
			Non-obese PCOS (N=63)	Non-obese controls (n=195)				
IRS2 Intron	rs12584136	C/A	0.07	0.02	8.51	3.64 (1.44-9.19)	0.003	Insulin signaling
LOC 107984901 Intergenic	rs4784165	T/G	0.34	0.24	4.91	1.63 (1.05-2.52)	0.026	Unknown
SUMO1P1 Regulatory	rs6022786	G/A	0.51	0.40	4.76	1.56 (1.04-2.34)	0.029	Unknown

spirit performed pair wise SNP-SNP interaction analysis using PLINK, and GMDR to get all possible combinations of SNP interactions up to five. While the researchers observed no significant interactions between any pair of SNPs in any of the two subsets, the GMDR results also did not yield any combination of significant multiple SNP interactions (results not presented). The LD plots were obtained and the haplo-

type blocks identified as per the procedure of Gabriel (Gabriel et al. 2002). Haplotype analysis (results not presented) yielded only one haplotype (ACGAA, constituted by 5 SNPs of FTO gene: rs9940128,rs1421085,rs17817449,rs8050136 and rs9939609) being significantly associated in the cohort of obese PCOS cases and non-obese controls (OR= 1.7, p≤0.009).

Table 3: Genotypic association of SNPs in the cohort of obese PCOS and non-obese controls with odds ratios from logistic regression analysis

SNP	Model best fitted	Genotype	Frequency		OR (95% CI)	P value
			Obese PCOS (N=104)	Non- obese controls (N=195)		
rs10838738	log additive	-	-	-	1.73(1.22-2.46)	0.002
rs1421085	log additive	-	-	-	1.70(1.18-2.45)	0.004
rs8050136	log additive	-	-	-	1.59(1.11-2.27)	0.011
rs17817449	log additive	-	-	-	1.57(1.10-2.25)	0.013
rs9939609	log additive	-	-	-	1.49(1.04-2.14)	0.028
rs9940128	log additive	-	-	-	1.43(1.02-2.02)	0.039
rs2479106	log additive	-	-	-	1.58(1.10-2.27)	0.013
rs12478601	log additive	-	-	-	1.4(1.00-1.95)	0.051

Table 4: Genotypic association of SNPs in the cohort of non-obese PCOS and non-obese controls with odds ratios from logistic regression analysis

SNP	Model best fitted	Genotype	Frequency		OR (95% CI)	P value
			Non- obese PCOS (N=63)	Non- obese controls (N=195)		
rs12584136	log additive	-	-	-	3.43(1.36-8.64)	0.015
rs4784165	Dominant	G/T-G/G	0.62	0.43	2.15 (1.2-3.84)	0.009
		T/T	0.38	0.57	-	-
rs6022786	Dominant	G/A-A/A	0.77	0.63	2.05(1.06-3.97)	0.027
		G/G	0.23	0.37	-	-

Cumulative Risk Score Analysis

The combined genetic risk score for each individual was calculated to determine the effect of the risk of significant SNPs (8 SNPs in obese and 3 SNPs in non-obese). The weighted mean proportion of the associated SNPs was computed by considering 2 for two risk alleles, 1 for single risk allele and 0 when there is no risk allele with weights as relative log odds ratios with respect to different SNPs. By multiplying allelic risk score with the total number of SNPs, the cumulative risk score for an individual was obtained. The risk scores ranging from 0-13.9 were grouped into 13 risk categories in case of obese set and the limited risk score range of 0-5.9 into five categories in the non-obese set. By considering risk category 1 as reference, odds ratios for the remaining categories were calculated in each of the two subsets. The results of logistic regression analyses of the risk catego-

ries in the obese and non-obese PCOS cohorts are given in Tables 5 and 6. Although there was no systematic pattern of increase in the odds ratios with increasing risk score/category, the higher risk categories generally displayed a relatively greater proportion of cases as compared to the controls suggesting significant risk for PCOS. The area under the curve (AUC) yielded by the receiver operator characteristic (ROC) curve analysis is highly significant ($p \leq 0.0001$) for each of the two subsets (Fig. 1), possibly suggesting some degree of predictive value of the risk variants for PCOS.

DISCUSSION

The obese PCOS women are characterised by severe hyperinsulinemia and insulin resistance and lower SHBG levels when compared to non-obese PCOS women or controls. Further, the presence of obesity, especially the abdomi-

Table 5: The results of cumulative risk score analysis according to the risk categories in the cohort of obese PCOS and non-obese controls

<i>Risk category (Risk score)</i>	<i>% of individuals in obese PCOS (N=104)</i>	<i>% of individuals in non- obese controls (N=195)</i>	<i>OR (95% CI)</i>	<i>Z Score</i>	<i>P value</i>
1(0-0.9)	4.80	11.28	Reference	-	-
2(1-1.9)	2.88	13.8	0.47 (0.09-2.41)	0.90	0.36
3(2-2.9)	10.57	13.4	1.69 (0.44-6.46)	0.76	0.44
4(3-3.9)	13.46	7.17	4.08 (1.00-16.5)	1.97	0.04
5(4-4.9)	2.88	8.20	0.82 (0.15-4.50)	0.23	0.82
6(5-5.9)	4.80	9.74	1.10 (0.24-4.96)	0.12	0.90
7(6-6.9)	14.4	12.8	2.36 (0.64-8.68)	1.30	0.19
8(7-7.9)	10.18	5.64	4.40 (0.97-19.80)	1.92	0.05
9(8-8.9)	18.26	6.67	5.65 (1.43-22.20)	2.47	0.01
10(9-9.9)	3.84	3.07	2.93 (0.46-18.30)	1.15	0.24
11(10-10.9)	1.92	2.56	2.20 (0.23-20.30)	0.69	0.48
12(11-11.9)	7.69	3.58	5.86 (1.07-32.0)	2.04	0.04
13(12-13.9)	3.84	2.04	4.40 (0.59-32.5)	1.45	0.14

Table 6: The results of cumulative risk score analysis according to the risk categories in the cohort of non-obese PCOS and non-obese controls

<i>Risk category (Risk score)</i>	<i>% of individuals in non-obese PCOS (N=63)</i>	<i>% of individuals in non-obese controls (N=195)</i>	<i>OR (95% CI)</i>	<i>Z Score</i>	<i>P value</i>
1(0-0.1)	7.90	20	reference	-	-
2(0.2-0.9)	22.2	37.43	1.48 (0.56-3.94)	0.79	0.425
3(1-1.90)	55.5	39.4	3.52 (1.40-8.81)	2.69	0.007
4(2-2.90)	9.50	2.05	11.2 (1.97-63.9)	2.73	0.006
5(3-5.90)	4.60	1.02	12.5 (1.25-124.4)	2.15	0.031

nal phenotype, appears to increase the availability of active androgens and oestrogens, worsens hirsutism, disturbs menstrual cycle and reduces fertility rate in the PCOS women (Teede et al. 2001). Hyperinsulinemia contributes to anovulatory infertility through the increased ovarian androgen secretion (Dunaif 1997). Insulin enhances intraovarian steroidogenesis by interacting with luteinizing hormone (LH), leading to inappropriate granulosa cell functions and the arrest of the follicular development (Franks et al. 1996). These cascading effects of the obesity related disturbances exacerbate the reproductive and metabolic disorders of PCOS. The central question that the researchers wanted to address was how distinct are the patterns of association with reference to obese and non-obese PCOS cohorts. Although the two sets display

mutually exclusive sets of SNPs being associated, the obese set shows a much larger number of 8 significant SNPs as against only three in case of the non-obese subset. Six of the 8 SNPs associated in the cohort of obese cases and non-obese controls belong to obesity genes, FTO and MTCH2, indicating that these SNPs may contribute to the elevated BMI in PCOS. Further, all the variants of FTO and MTCH2 and other related genes associated in this subset were risk in nature and possibly responsible for triggering the manifestation of this syndrome through their indirect role of exacerbating the reproductive and metabolic disorders of PCOS. The researchers surmise if this phenomenon was not really responsible for the spurt in PCOS prevalence in recent times with increasing urbanisation and the change of lifestyles albeit precise

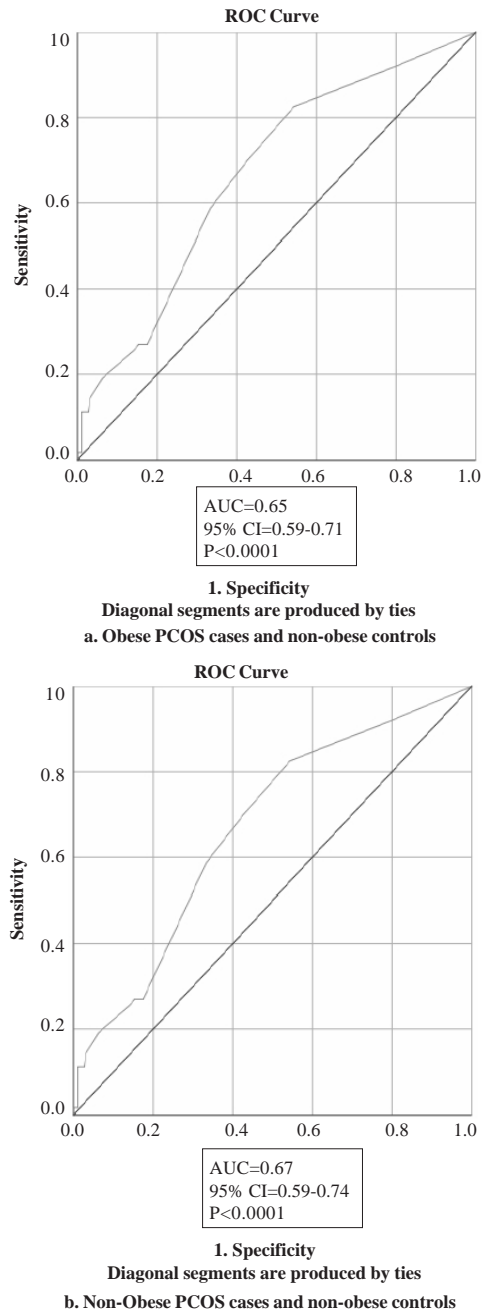


Fig. 1. Receiver Operating Characteristic (ROC) curve indicating the area under curve (AUC) characterizing the discriminative powers of risk scores of obese and non-obese PCOS subsets

quantitative data pertaining to the lifestyle parameters and the evidence of their impact on the body composition is required to be generated. This seems plausible particularly because none of the 31 SNPs from the reproductive pathway genes included in this study or their interactions with metabolic pathway genes were observed to be significantly associated with the syndrome either in obese or in non-obese subsets. Further, the SNPs-rs2479106 from DENND1A gene (a gene coding for modifier of guanine that is associated with multiple organ dysfunctions like ovaries, hypothalamus, pituitary and adrenal glands) and rs12478601 from THADA gene (a thyroid adenoma gene associated with disorders of glucose metabolism, hyper secretion of LH, hyperandrogenism and dyslipidemia) were also associated with PCOS in this subset. Nevertheless, it is pertinent to note that all these SNPs are from the intronic regions and their precise functional mechanisms cannot be readily ascertained. On the other hand, of the three associated SNPs in the non-obese subset (Table 2), one (from intronic region) belongs to IRS2 gene with a putative role in insulin resistance and the other two SNPs, a intergenic variant rs4784165 that belongs to an uncharacterised LOC107984901 and associated with DNA modification, and a regulatory variant rs6022786 that belongs to SUMO1P1 and involves in telomere dysfunction. Overall, the associated SNPs are from the metabolic pathway genes, none from the reproductive pathway. Yet, the cumulative risk scores and ROC curve analysis yielded highly significant AUC (0.66 – 0.67; $p \leq 0.0001$) for the risk variants in case of both obese and non-obese subsets. Nevertheless, all these risk variants lose significance after correction for multiple testing, and hence their role in the manifestation of the syndrome could be only of minor nature and cannot qualify for causative/biomarkers with predictive utility. Further, most of the significant SNPs observed in this study were found to be associated with BMI and diabetes related traits like insulin resistance and obesity in the PCOS women and may not directly contribute to the reproductive aspects of the syndrome albeit the role of interactions of these genes with that of the reproductive pathway in the manifestation of PCOS appears to be quite plausible.

CONCLUSION

The researchers observed distinct and mutually exclusive set of SNPs to be associated with the obese and non-obese PCOS cohorts in this study. Most of the significant SNPs observed in this study were found to be associated with BMI and diabetes related traits like insulin resistance and obesity in the PCOS women, and may not directly contribute to the reproductive aspects of the syndrome albeit the role of interactions of these genes with that of the reproductive pathway in the manifestation of PCOS appears to be a distinct possibility. Given that all the 11 risk variants lose significance after correction for multiple testing their role in the manifestation of the syndrome could be only minor in nature and cannot qualify as causative/ biomarkers with predictive utility.

RECOMMENDATIONS

The modest size of the samples for obesity specific cohorts did not yield any significant interactions and one should await specifically designed large scale obesity specific cohort studies for any clear-cut inferences to be drawn on the issues such as gene-gene and gene-environment interactions involving multivariate analysis, and to reflect on the leads provided by this study, in order to establish the role that obesity and/or obesity related genetic variants play in the manifestation of PCOS.

ACKNOWLEDGEMENTS

BMR is thankful to the Director General, ICMR, for granting him the Emeritus Medical Scientist position and providing assistance of an SRF, which made it possible to undertake this research work and to the Head of Department of Genetics, Osmania University, for providing logistics support and Dr. A. Sandhya for academic interactions.

The SNP and other background data used for the present study were generated when BMR was still Professor of the Indian Statistical Institute (ISI). BMR thanks Dr. Shilpi Dasgupta, former PhD student who assisted in collection

of blood samples and background data and in the isolation of DNA during her tenure as a research fellow of ISI at Hyderabad and the then Directors of ISI, Kolkata, for financial and logistics support.

REFERENCES

- Barber TM, McCarthy MI, Wass JA, Franks S 2006. Obesity and polycystic ovary syndrome. *Clin Endocrinol (Oxf)*, 65(2): 137-145.
- Dunaif A 1997. Insulin resistance and the polycystic ovary syndrome: Mechanism and implications for pathogenesis. *Endocr Rev*, 18: 774-800.
- Franks S, Robinson S, Willis D 1996. Nutrition, insulin and polycystic ovary syndrome. *Rev Reprod*, 1: 47- 53.
- Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B et al. 2002. The structure of haplotype blocks in the human genome. *Science*, 296: 2225-2229.
- Holte J, Bergh T, Berne C, Wide L, Lithell H 1995. Restored insulin sensitivity but persistently increased early insulin secretion after weight loss in obese women with polycystic ovary syndrome. *J Clin Endocrinol Metab*, 80: 2586-2593.
- Irgam K, Reddy BM, Annamaneni S, Rayabharapu P 2019. The genetic susceptibility profile of the South Indian women with polycystic ovary syndrome and the universality of the lack of association of type 2 diabetes genes. *Gene*, 701: 113-120.
- Kiddy DS, Hamilton-Fairley D, Bush A, Short F, Anyaoku V, Reed MJ et al. 1992. Improvement in endocrine and ovarian function during dietary treatment of obese women with polycystic ovary syndrome. *Clin Endocrinol (Oxf)*, 36: 105-111.
- Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group 2004. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. *Fertil Steril*, 81(1): 19-25.
- Sam S 2007. Obesity and polycystic ovary syndrome. *Obes Manag*, 3: 69-73.
- Taponen S, Martikainen H, Järvelin MR, Laitinen J, Pouta A, Hartikainen AL et al. 2003. Hormonal profile of women with self reported symptoms of oligomenorrhea and/or hirsutism: Northern Finland birth cohort 1966 study. *J Clin Endocrinol Metab*, 88: 141-147.
- Teede H, Deeks A, Moran L 2001. Polycystic ovary syndrome: A complex condition with psychological, reproductive and metabolic manifestations that impacts on health across the lifespan. *BMC Med*, 30: 41-51.
- Willer CJ, Speliotes EK, Loos RJ, Li S, Lindgren CM, Heid IM et al. 2009. Genetic investigation of anthropometric traits consortium- Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nat Genet*, 41: 25-34.

**Paper received for publication in December, 2019
Paper accepted for publication in January, 2021**