


Article

Genome-Wide Association Study of Metabolic Syndrome Reveals Primary Genetic Variants at *CETP* Locus in Indians

Gauri Prasad ^{1,2}, Khushdeep Bandesh ^{1,2}, Anil K. Giri ^{1,2} , Yasmeen Kauser ^{1,2}, Prakriti Chanda ³, Vaisak Parekatt ¹, INDICO [†], Sandeep Mathur ⁴, Sri Venkata Madhu ⁵, Pradeep Venkatesh ⁶, Anil Bhansali ⁷, Raman K. Marwaha ⁸, Analabha Basu ⁹, Nikhil Tandon ^{10,*} and Dwaipayan Bharadwaj ^{2,3,*}

- ¹ Genomics and Molecular Medicine Unit, CSIR-Institute of Genomics and Integrative Biology, New Delhi 110025, India
 - ² Academy of Scientific and Innovative Research, CSIR-Institute of Genomics and Integrative Biology Campus, New Delhi 110020, India
 - ³ Systems Genomics Laboratory, School of Biotechnology, Jawaharlal Nehru University, New Delhi 110067, India
 - ⁴ Department of Endocrinology, S.M.S. Medical College, Jaipur, Rajasthan 302004, India
 - ⁵ Department of Endocrinology, Centre for Diabetes Endocrinology & Metabolism, University College of Medical Sciences (University of Delhi) & GTB Hospital, New Delhi 110095, India
 - ⁶ Dr Rajendra Prasad Centre for Ophthalmic Sciences, All India Institute of Medical Sciences, New Delhi 110029, India
 - ⁷ Department of Endocrinology, Post Graduate Institute of Medical Education and Research, Sector-12, Chandigarh 160012, India
 - ⁸ Department of Endocrinology, International Life Sciences Institute, New Delhi 110024, India
 - ⁹ National Institute of Biomedical Genomics, P.O.: Netaji Subhas Sanatorium, Kalyani 741251, West Bengal, India
 - ¹⁰ Department of Endocrinology, All India Institute of Medical Sciences, New Delhi 110029, India
- * Correspondence: nikhil_tandon@hotmail.com (N.T.); db@jnu.ac.in or db@mail.jnu.ac.in (D.B.); Tel.: +91-98-18211663 (N.T.); +91-98-11677141 (D.B.); Fax: +91-11-26741561 (N.T. & D.B.)
- † Membership of INDICO consortium has been listed in back matter.

Received: 29 June 2019; Accepted: 30 July 2019; Published: 30 July 2019



Abstract: Indians, a rapidly growing population, constitute vast genetic heterogeneity to that of Western population; however they have become a sedentary population in past decades due to rapid urbanization ensuing in the amplified prevalence of metabolic syndrome (MetS). We performed a genome-wide association study (GWAS) of MetS in 10,093 Indian individuals (6617 MetS and 3476 controls) of Indo-European origin, that belong to our previous biorepository of The Indian Diabetes Consortium (INDICO). The study was conducted in two stages—discovery phase ($N = 2158$) and replication phase ($N = 7935$). We discovered two variants within/near the *CETP* gene—rs1800775 and rs3816117—associated with MetS at genome-wide significance level during replication phase in Indians. Additional *CETP* loci rs7205804, rs1532624, rs3764261, rs247617, and rs173539 also cropped up as modest signals in Indians. Haplotype association analysis revealed GCCCAGC as the strongest haplotype within the *CETP* locus constituting all seven *CETP* signals. In combined analysis, we perceived a novel and functionally relevant sub-GWAS significant locus—rs16890462 in the vicinity of *SFRP1* gene. Overlaying gene regulatory data from ENCODE database revealed that single nucleotide polymorphism (SNP) rs16890462 resides in repressive chromatin in human subcutaneous adipose tissue as characterized by the enrichment of H3K27me3 and CTCF marks (repressive gene marks) and diminished H3K36me3 marks (activation gene marks). The variant displayed active DNA methylation marks in adipose tissue, suggesting its likely regulatory activity.

Further, the variant also disrupts a potential binding site of a key transcription factor, NRF2, which is known for involvement in obesity and metabolic syndrome.

Keywords: genome wide association study; metabolic syndrome; genetic variants; gene regulation

1. Introduction

Metabolic syndrome (MetS) refers to a complex pathophysiological state attained by conjugation of a set of cardiometabolic risk components in an individual that include—central obesity, dyslipidemia, elevated blood pressure, and fasting plasma glucose [1]. MetS aggravates the risk for various health complications including type 2 diabetes (T2D), cardiovascular disease (CVD), cancer, and mortality from all causes [2–4]. The prevalence of MetS in India was observed to be ranging from 11–41% on the basis of geographical region, socioeconomic status, urban–rural environment, age, sex, ethnicity of the individuals, and the definition used [5–7]. Given the high prevalence of MetS in Indians, with its prompt spread in younger adults [7] and adolescents [8], effective plans for its early detection and intervention are critically desirable to mitigate the burden of associated diseases.

Genetic and environmental factors and their cumulative gene–environment interactions contribute to pathophysiology of MetS [9–12]. The genetic heritability is liable up to 50% for some individual metabolic components and 13–30% for collective MetS phenotype [9–12]. Several large-scale genetic studies have been performed to identify MetS related single nucleotide polymorphisms (SNPs) considering independent components of MetS as a quantitative trait [13–17]. In view of MetS as a binary phenotype, several other genome-wide studies worldwide identified numerous loci influencing the combined metabolic syndrome outcome. For instance, genetic loci in *BUD13*, *ZNF259*, *APOA5*, *LPL*, and *CETP* in Europeans [18], *TCF7L2*, *APOA5*, *LPL*, *CETP*, *APOE*, and *APOC1* in African Americans [19], *CA10* and *CTNNA3* in Africans [20], *APOA1/C3/A4/A5* gene cluster region in Finnish [21], *APOA5*, *BUD13*, and *ALDH2* in Han Chinese [22], *LPL*, *MYL2*, *CCDC63*, and *CETP* in Koreans [23], and *APOA* and *COLEC12* in Taiwanese [24] have been attributed for strong association with MetS. These studies highlight multiple shared lipid metabolism pathway genes across diverse populations as well as novel population-specific genes, which require the need for additional population-wide genetic studies to delineate remaining genetic heritability of MetS across varied ethnicities [25].

Indians represent a unique population with a distinct genetic make-up, food habit, and lifestyle compared to other world populations [26,27]. Moreover, Indians display comparatively higher atherogenic dyslipidemia, glucose intolerance, subclinical inflammation, thrombotic propensity, and endothelial abnormality compared to Caucasians [28,29]. Many of such metabolic deregulations are extremely severe and have an earlier age of onset in Indians than Caucasians [28,29]. Besides, Indians have an increased body fat mass, a greater truncal, intra-abdominal subcutaneous adipose tissue with ectopic fat buildup compared to Caucasians, resulting in an enhanced risk for metabolic syndrome and CVD [28,30]. These reasons underscore the possibility of population-specific genetic risk towards MetS phenotype in India. We ourselves have shown this population-specificity of genetic loci for a few metabolic phenotypes in previous reports [15,27], but this is yet unknown for compound MetS phenotype.

Previously, a genome-wide association study was conducted in Indian Asian men ($N = 4794$) of The London Life Sciences Population (LOLIPOP) cohort who were living in West London, United Kingdom (UK) at the time of sample collection [31]. The study lacked a homogenous Indian population as majority of Indian individuals were of mixed ethnicities (Indo-Europeans, Dravidians etc.) in the LOLIPOP cohort that were not appropriately segregated. The study was also limited by only considering Indians living in UK, who differ considerably from native Indians in terms of food habit and lifestyle, which may have an influence on differential genetic architecture for MetS between the two groups. Moreover, the study did not identify genome-wide association study (GWAS) level association

($p < 5 \times 10^{-8}$) of any genetic variant for compound metabolic syndrome phenotype in representative Indian individuals of London, UK.

The present two-stage genome-wide association study was intended to identify genetic variants governing compound MetS phenotype in 10,093 native Indians (6617 MetS and 3476 controls) speaking Indo-European language.

2. Materials and Methods

2.1. Ethical Approval

Ethical approval for the study was obtained from Human Ethical Committees of All India Institute of Medical Sciences, New Delhi, India and CSIR-Institute of Genomics and Integrative Biology, New Delhi, India, following principles of Helsinki Declarations (BSC0122, NIDDK GRANT NUMBER: UOO DK085545). All the individuals included in the study were well informed about objectives of study and written consent was taken from each one of them before their involvement in the study.

2.2. Study Subjects

All the study subjects were Indo-European speakers from Northern India primarily from neighborhoods of Delhi. In India, genetic ancestry and language are strongly confounded. Moreover, among the Indo-European speakers, the geographical origin of individuals is the essential correlate of genetic variability [32].

The study individuals included in the present study were part of a type 2 diabetes GWAS conducted previously in our laboratory [27] and were members of The Indian Diabetes Consortium (INDICO) [33]. Non-diabetic control subjects who served as controls during GWAS study were enrolled through diabetes alertness camps conducted across different zones of Delhi and adjoining areas [27,33].

Discovery phase T2D subjects were registered from the Department of Endocrinology, All India Institute of Medical Sciences (New Delhi), who joined the clinic before September 2008. Replication phase T2D subjects were registered from departments of collaborator hospitals: All India Institute of Medical Sciences (New Delhi), Guru Teg Bahadur Hospital (New Delhi), and Sawai Man Singh Hospital (Jaipur). T2D patients enrolled from All India Institute of Medical Sciences for replication phase were the patients who joined the clinic post September 2008. In addition, patients with self-known diabetes or under prescription for diabetes and recently diagnosed were also recruited from Diabetes Alertness Camps. T2D subjects were identified as per WHO criteria as described previously [27]. Pregnant females, children, teenagers, and those with type 1 diabetes were excluded from the study.

Blood samples were collected from subjects after an overnight fast, and their DNA was isolated from peripheral blood through salt precipitation protocol. All the study individuals underwent detailed measurements of biochemical and anthropometric measures as described earlier [32]. Waist circumference (WC), fasting glucose (FG), high density lipoprotein cholesterol (HDL-C), triglycerides (TG), and systolic and diastolic blood pressure (SBP and DBP) were measured using standard procedure as described previously [33].

2.3. Phenotype Definition (MetS)

We defined MetS using modified National Cholesterol Education Program (NCEP) adult treatment panel (ATP III) measure for Asian populations [34], as used earlier in our previous study [35]. Subjects were classified as MetS cases who attained three or more of these metabolic measures: (1) WC ≥ 90 cm in men or ≥ 80 cm in women, (2) FG ≥ 100 mg/dL or on medication, (3) HDL-C < 40 mg/dL in men or < 50 mg/dL in women or on medication, (4) TG ≥ 150 mg/dL or on medication, and (5) SBP ≥ 130 mm Hg or DBP ≥ 85 mm Hg or taking medication for blood pressure control. Subjects having ≤ 2 number of MetS components were classified as MetS controls.

2.4. Genome-Wide Association Study

2.4.1. Discovery Phase

DNA samples were genotyped genome-wide using Illumina Human610-Quad Beadchips (Illumina Inc., San Diego, CA, USA) as part of GWAS studies conducted for T2D and interrelated quantitative metabolic phenotypes earlier in our laboratory [14,15,27,36–38]. The GenCall algorithm employed in GenomeStudio software (Illumina, Inc., San Diego, CA, USA) was used to compute genotype calls. The genotype calls were further exported in PLINK v1.07 for downstream analysis [39].

Brief analysis steps and quality control pipeline employed in the study have been summarized in Supplemental Figures S1 and S2. Samples with less than 95% call rate, sex discrepancy, and extremely low or high heterozygosity (mean \pm 3 SD) were removed. Further, SNPs with less than 99% call rate, MAF $<$ 0.01, and Hardy Weinberg equilibrium (HWE) $p < 1 \times 10^{-7}$ were excluded. Identity by descent (IBD) analysis was performed to identify related and duplicated individuals in the data and those with pi hat score \geq 0.1875 were excluded. Principal component analysis (PCA) was implemented to spot population outliers. Linkage disequilibrium (LD) pruning of markers was carried in autosomal SNPs using the `-indep-pairwise` command provided in PLINK v1.07 using r^2 of 0.2 and window dimension of 50 SNPs [39]. Analysis of the initial ten principal components detected 41 subjects as potential population outliers (mean \pm 6 SD) that were expelled.

Followed by quality control (QC), a total of 519,607 SNPs and 2158 individuals (1596 MetS and 562 controls) remained that were followed for logistic regression analysis assuming an additive model adjusted for age, sex, and the first two principal components in PLINK. Median χ^2 statistics were applied to infer genomic inflation factor λ . Manhattan and quantile-quantile (QQ) plots were created using qqman package in R (<http://www.r-project.org/>) [40].

2.4.2. Replication Phase and Meta-Analysis

The current study is part of a large-scale genetic study to discover genetic variants influencing type 2 diabetes pathophysiology and levels of related quantitative metabolic phenotypes in Indian population [14,15,27,36–38]. Markers with discovery phase $p < 10^{-4}$ for MetS and other traits, besides earlier known signals for MetS and other traits, were genotyped in replication phase using GoldenGate assay (Illumina, San Diego, CA, USA). In total, 930 samples (11.72%) were genotyped as technical replicates and an error rate of $<$ 0.01% was observed between them.

Samples with less than 90% call rate were removed from the analysis. Further, SNPs with genotype confidence score $<$ 0.25, GenTran score $<$ 0.60, cluster separation score $<$ 0.4 and call rate $<$ 90% were expelled. SNPs with MAF $<$ 0.01 were also discarded. From SNPs with MAF $>$ 0.01, those with HWE $p < 1 \times 10^{-7}$ were excluded. After rigorous QC, we retained 2699 SNPs and 7935 individuals (5021 MetS and 2914 controls) in the replication phase that were tested for logistic regression under an additive model adjusted for age and sex.

PLINK [39] was employed for meta-analysis of summary association statistics of discovery and validation phases under a fixed-effect inverse variance model.

Previous associations of identified variants and genes were obtained from the GWAS catalog, GWAS atlas, and Type 2 Diabetes Knowledge Portal [41–43]. Regional association plot within \pm 1 Mb of lead signal was plotted using locuszoom [44].

2.5. Statistical Power of the Study

We calculated the power of the study using Quanto software (Department of Preventive Medicine, University of Southern California, Los Angeles, CA, USA) [45]. Log-additive model of inheritance for allele frequencies in the range from 0.01–0.5 and odds ratios (OR) in the range from 0.63 to 2.08 derived from literature were used. Prevalence of disease was taken as 11% at significance level of 0.05.

2.6. Conditional and Haplotype Association Analysis

Conditional analysis for seven variants identified in *CETP* locus was performed in replication phase data employing logistic regression model. Age, sex, and identified SNP genotypes were used as covariates in the model using PLINK.

Haplotype-based association analysis of *CETP* locus was carried out using a logistic regression model adjusting for age and sex at 10,000 permutations in replication phase data using PLINK.

2.7. Imputation Analysis

Imputation analysis of novel sub-GWAS loci near *SFRP1* in the discovery phase dataset was performed as detailed earlier [27]. For reference population, 1000 Genome phase 3 panel was used. Prephasing of chromosome 8 was performed with SHAPEIT [46]. A total genomic region of 2 Mb (1 Mb each towards 5' and 3' end of the variant) was imputed utilizing IMPUTE 2 [47] which covered the entire LD block of the variant. After imputation, SNPs underwent stringent QC. Imputed SNPs with certainty score <0.90, info score <0.5, and MAF < 0.01 were removed. Further, the QC qualified SNPs were used in association test with compound MetS phenotype using logistic regression adjusting for age, sex, PC1, and PC2 as covariates in the model using PLINK.

2.8. Overlaying Gene Regulatory Features

For identification of potential functional relevance of novel sub-GWAS loci near *SFRP1*, we used several publicly available gene regulatory databases. Tissue-wide gene expression profiles of *SFRP1* were downloaded from GTEx-portal-v7 (<https://www.gtexportal.org/home/>) [48]. ATAC-seq data for human subcutaneous adipose tissue was obtained from an adult female of 53 years of age from ENCODE [49]. ChIP-seq data for regulatory histone marks H3K36me3 or H3K27me3 had been derived from subcutaneous adipose tissue of 5 adult females aged 25, 41, 49, 59, and 81 years, and was obtained from ENCODE [49]. ChIP-seq data for CTCF binding was derived from subcutaneous adipose tissue of 2 adult females aged 51 and 53 years from ENCODE [49]. Whole genome bisulphite sequencing data (WGBS) for adipose tissue was derived from a male adult subject aged 34 years from ENCODE dataset [49]. Predicted sites for transcription factor (TF) binding were obtained from JASPAR portal [50]. UCSC browser was used for visualization of genome regulatory features of sub-GWAS loci [51].

3. Results and Discussion

The present study was the first genome-wide association study that identifies a common genetic basis of compound MetS phenotype in Indians of Indo-European origin living in India. Our study was robustly powered to detect loci with similar odds ratios as identified in previous GWAS studies for MetS in literature (>98%) (Supplemental Figure S3). Further, the QQ plot displayed good agreement of calculated *p*-values with theoretical *p*-values under the null hypothesis (Supplemental Figure S4). The genomic inflation factor (λ) was observed to be 1.06 reflecting a homogenous study population. Characteristics of the study population are presented in Supplemental Table S1.

3.1. Genome-Wide Association Analysis of MetS

In the discovery phase, variant rs11108860, which was located within a long intergenic non-coding RNA gene *RP11-541G9.1*, was the strongest signal ($p = 8.72 \times 10^{-7}$) [Figure 1]. Though, in the replication phase, association of *RP11-541G9.1* was not sustained ($p = 0.96$).

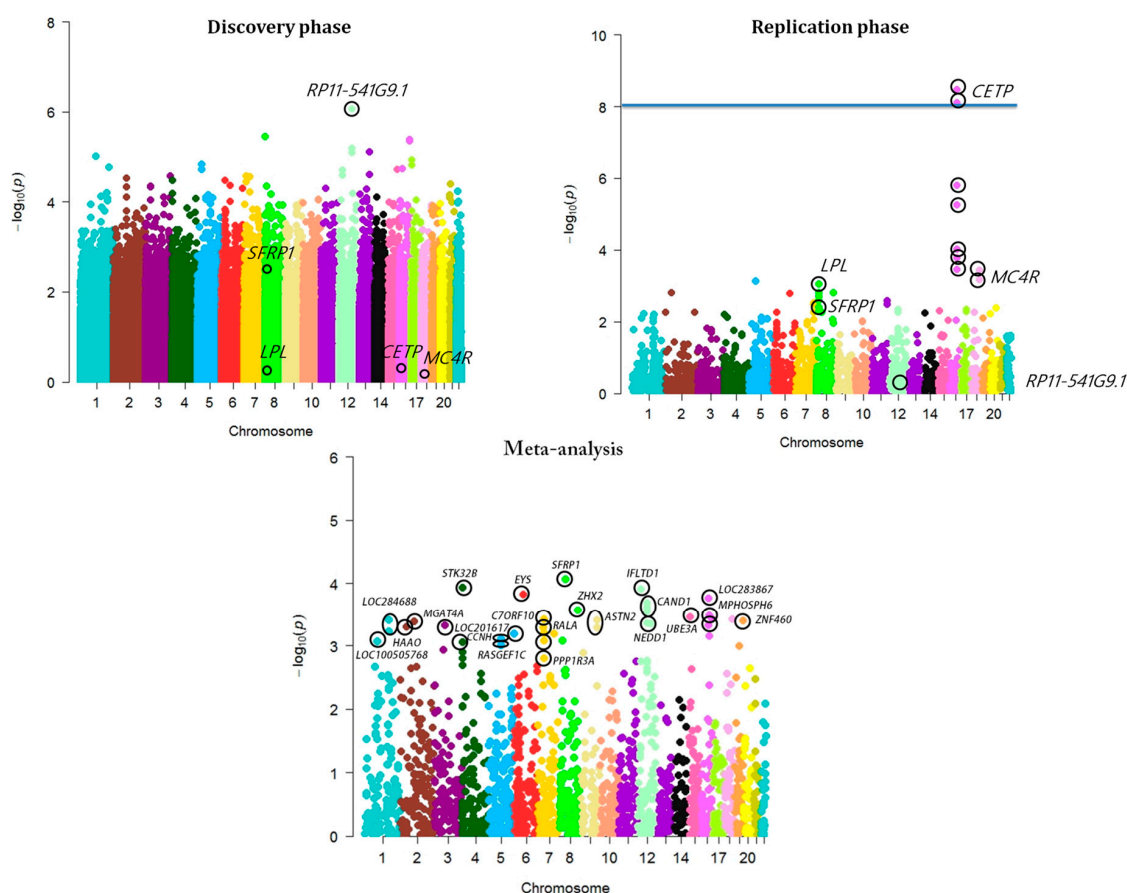


Figure 1. Manhattan plot of association p-values for metabolic syndrome (MetS) in discovery phase, replication phase, and meta-analysis. The $-\log_{10} p$ -values of genotyped single nucleotide polymorphisms (SNPs) calculated from association analysis have been presented with respect to SNP positions across autosomes (National Center for Biotechnology Information Build 37).

Interestingly, amid the earlier known gene regions for MetS and related metabolic traits that were genotyped in the replication phase, variants-rs1800775 ($p = 3.48 \times 10^{-9}$) and rs3816117 ($p = 7.71 \times 10^{-9}$) within/near the *CETP* gene were associated with MetS at genome-wide significance levels in Indians (Table 1, Supplemental Table S2, Figure 1). This was followed by modest associations of an additional five *CETP* loci-rs7205804 ($p = 1.58 \times 10^{-6}$), rs1532624 ($p = 5.57 \times 10^{-6}$), rs3764261 ($p = 9.52 \times 10^{-5}$), rs247617 ($p = 1.67 \times 10^{-4}$), and rs173539 ($p = 3.48 \times 10^{-4}$) (Table 1, Supplemental Table S2, Figure 1).

Table 1. Association status of known metabolic traits associated gene regions with MetS in replication phase in Indians ($p < 1 \times 10^{-3}$).

Replication Phase					
Gene Region	CHR	Start Base Pair (hg19)	Number of SNPs	p-Value (Lead SNP)	N
<i>CETP</i>	16	56988044	7	3.48×10^{-9}	4671
<i>MC4R</i>	18	57851097	2	3.66×10^{-4}	4666
<i>LPL</i>	8	19919655	1	8.82×10^{-4}	4650

Association analysis with compound MetS phenotype, adjusted for age and sex as covariates. Start base pair is position of farthest 5' SNP in context of gene. CHR: Chromosome; N: Number of non-missing individuals.

The identified GWAS variant near *CETP* in Indians—rs1800775 (C allele)—has been recently demonstrated to have a nominal association with risk of MetS and its individual components (high FG and low HDL-C levels) in the Uyghur ethnic group of China [52]. Besides, another GWAS locus within *CETP*-rs3816117, though not documented in association with MetS until now, is in strong LD with

rs1800775 in Africans, Americans, Asians, and Europeans (both r^2 and $D' > 0.9$) [53], including Indians ($r^2 = 0.88$, $D' = 0.95$) [Supplemental Figure S5], suggesting its potential association with MetS and related components in certain ethnic origins. Supplemental Figure S6 depicts a regional association plot of signals within/near ± 1 Mb of *CETP* gene region in discovery phase, replication phase and meta-analysis in Indians.

CETP gene (cholesteryl ester transfer protein) mRNA is primarily expressed by the spleen, adipose tissue, kidney, liver, lungs, and thyroid [48]. CETP protein plays an important role in the net transport of neutral lipids such as cholesteryl esters and triglycerides [54]. It transfers cholesteryl esters of high-density lipoproteins (HDL) to very low-density lipoproteins (VLDL), in exchange for equimolar amounts of triglyceride from VLDL or chylomicrons to HDL [54]. CETP is also a central protein that maintains the reverse cholesterol transport pathway, wherein overloaded cholesterol is taken from peripheral tissues and restored to the liver for removal from the body [54].

The *CETP* gene is vastly polymorphic. Our identified *CETP* loci serve as strong cis-expression quantitative trait signals (cis-eQTL) in various human tissues including liver stomach, aorta artery, pancreas, and subcutaneous adipose tissue etc., thereby affecting the expression of their occupied gene [Supplemental Table S3]. These variants also confer risk for various cardiometabolic and mental health diseases including dyslipidemia, coronary artery disease, hypertension, obesity, type 2 diabetes, depression, bipolar disorder, and schizophrenia [Supplemental Table S3].

It has been found that high CETP protein activity lowers the concentration of HDL-C [55]. A few functional genetic variants in this gene also present lower plasma protein levels and activity, with parallel increases in HDL-C levels [56,57]. For instance, our identified GWAS locus for MetS—rs1800775-A (protective allele), located in promoter region of the *CETP* gene—has been reported to affect the promoter activity and thereby lowers the gene expression [58], which may have further influence on encoded protein function to maintain lipid levels in the body. This was further demonstrated in another interesting report where subjects with rs1800775-A displayed reduced CETP protein level and activity, and higher HDL-C and apolipoprotein A-I concentrations [59].

On similar lines, rs1800775-A has been shown to confer protection for coronary artery disease (CAD) in a recent large-scale GWAS conducted in individuals from the United Kingdom [60]. In addition to rs1800775, other *CETP* variants, rs247616 and rs1532624, have also been linked to influencing the risk of CAD in Polish populations [61], and rs173539 with coronary artery calcification in a GWAS conducted with Finnish individuals [62].

Further, two variants near *MC4R*-rs17782313 ($p = 3.66 \times 10^{-4}$) and rs12970134 ($p = 6.75 \times 10^{-4}$), and one variant near *LPL*-rs4128744 ($p = 8.82 \times 10^{-4}$) were also modestly associated with MetS in the replication phase in Indians (Table 1, Supplemental Table S2, Figure 1). Discovered *MC4R* locus is a previously documented signal for obesity and type 2 diabetes (rs17782313 and rs12970134) in multiple ethnic cohorts [41]. Moreover, one variant within *ZNF259*-rs964184 ($p = 2.61 \times 10^{-3}$), an earlier reported variant for MetS [21], also featured a nominal association with MetS in Indians (Supplemental Table S4).

Genetic variance is essentially governed by multiple SNPs of little effect size that are frequently neglected due to strict GWAS p -value limits and multiple testing corrections [63]. We did not detect any genome-wide significant signal in combined analysis during meta-analysis of the discovery and replication phases (Table 2). We only observed a novel sub-GWAS level association ($p < 1 \times 10^{-4}$) of locus rs16890462 ($p = 8.75 \times 10^{-5}$), that is 23kb 5' of *SFRP1* locus, for the first time in association with MetS in Indians (Table 2, Figure 1). This was followed by modest associations with other signals in *STK32B*, *IFLTD1*, *EYS*, *CAND1*, *ZHX2*, *LOC283867*, *UBE3A*, *MPHOSPH6*, *C7orf10*, *ASTN2*, *LOC284688*, *ZNF460*, *MGAT4A*, *NEDD1*, *LOC201617*, *HAAO*, *RALA*, *PPP1R3A*, *RASGEF1C*, *CCNH*, and *LOC100505768* genes (Table 2, Figure 1).

Table 2. Novel signals associated with MetS (p -value $< 1 \times 10^{-3}$) in meta-analysis in Indians.

SNP	CHR	Base Position	Nearby Gene	SNP Location	Alleles (Effect/Other)	MAF	Discovery Phase			Replication Phase			Meta-Analysis				
							N	p -Value	OR	N	p -Value	OR	p -Value	OR	Dir	I	Q
rs16890462	8	41309355	<i>SFRP1</i>	Intergenic	A/G	0.21	2156	5.48×10^{-3}	1.29	3259	5.51×10^{-3}	1.25	8.75×10^{-5}	1.26	++	0	0.78
rs1530611	4	5206254	<i>STK32B</i>	intronic	A/G	0.29	2148	6.69×10^{-3}	1.23	3257	6.37×10^{-3}	1.21	1.18×10^{-4}	1.22	++	0	0.89
rs11048180	12	25735148	<i>IFLTD1</i>	intergenic	A/G	0.08	2157	5.46×10^{-4}	0.66	3259	0.04	0.8	1.25×10^{-4}	0.73	-	31.56	0.23
rs16896746	6	66289412	<i>EYS</i>	intronic	G/A	0.08	2156	4.32×10^{-5}	0.62	3259	0.1	0.86	1.51×10^{-4}	0.73	-	75.4	0.04
rs710630	12	65983583	<i>CAND1</i>	intronic	G/A	0.46	2158	0.01	1.18	3259	4.61×10^{-3}	1.19	2.08×10^{-4}	1.19	++	0	0.91
rs710628	12	65943747	<i>CAND1</i>	intergenic	A/G	0.46	2157	0.01	1.18	3258	5.53×10^{-3}	1.19	2.47×10^{-4}	1.18	++	0	0.93
rs7005211	8	123538147	<i>ZHX2</i>	intergenic	G/A	0.47	2155	6.69×10^{-3}	0.83	4654	8.90×10^{-3}	0.89	2.72×10^{-4}	0.87	-	0	0.33
rs1060350	12	65992732	<i>CAND1</i>	synonymous	G/A	0.48	2157	6.15×10^{-3}	1.21	3254	0.01	1.16	2.85×10^{-4}	1.18	++	0	0.67
rs1152877	12	65989452	<i>CAND1</i>	intronic	G/A	0.48	2155	5.50×10^{-3}	1.22	3220	0.01	1.16	2.89×10^{-4}	1.18	++	0	0.62
rs564210	16	64314818	<i>LOC283867</i>	intergenic	A/G	0.26	2156	3.55×10^{-4}	0.76	3251	0.1	0.89	3.30×10^{-4}	0.83	-	60.29	0.11
rs12595506	15	25744981	<i>UBE3A</i>	intergenic	G/A	0.39	2156	0.01	1.19	7930	4.99×10^{-3}	1.1	3.53×10^{-4}	1.11	++	0.22	0.32
rs2967379	16	80770811	<i>MPHOSPH6</i>	intergenic	G/A	0.49	2156	4.52×10^{-6}	0.73	3258	0.4	0.96	3.69×10^{-4}	0.84	-	88.69	2×10^{-3}
rs10499618	7	40787165	<i>C7orf10</i>	intronic	G/A	0.11	2158	3.98×10^{-3}	1.41	3257	0.03	1.26	3.82×10^{-4}	1.32	++	0	0.48
rs1337212	9	119239170	<i>ASTN2</i>	intergenic	A/G	0.11	2157	4.53×10^{-4}	1.55	3259	0.09	1.18	3.93×10^{-4}	1.31	++	62.74	0.1
rs7554931	1	170365623	<i>LOC284688</i>	intergenic	A/G	0.41	2157	2.91×10^{-3}	0.81	4673	0.01	0.9	3.98×10^{-4}	0.87	-	44.4	0.18
rs3746228	19	57804362	<i>ZNF460</i>	3'-UTR	A/G	0.18	2158	7.18×10^{-4}	1.38	4529	0.03	1.13	4.02×10^{-4}	1.18	++	70.48	0.06
rs885036	2	98671225	<i>MGAT4A</i>	intronic	G/A	0.49	2157	2.96×10^{-5}	1.35	3254	0.3	1.07	4.18×10^{-4}	1.18	++	83.49	0.01
rs1066396	12	66005634	<i>CAND1</i>	intergenic	G/A	0.48	2157	6.14×10^{-3}	1.21	3254	0.02	1.15	4.37×10^{-4}	1.17	++	0	0.59
rs11108860	12	96081536	<i>NEDD1</i>	intergenic	G/A	0.04	2157	8.72×10^{-7}	0.45	3259	0.9	1.01	4.52×10^{-4}	0.66	-+	91.59	6×10^{-4}
rs4677119	3	72291958	<i>LOC201617</i>	intergenic	A/G	0.33	2158	5.76×10^{-3}	0.81	4653	0.01	0.89	4.80×10^{-4}	0.87	-	22.01	0.26
rs9309089	2	43028132	<i>HAAO</i>	intergenic	A/G	0.35	2158	0.1	1.12	4623	1.54×10^{-3}	1.15	5×10^{-4}	1.14	++	0	0.72
rs6948816	7	39661439	<i>RALA</i>	intronic	A/G	0.03	2158	5.36×10^{-3}	0.62	3257	0.03	0.69	5.13×10^{-4}	0.65	-	0	0.62
rs10983653	9	119237233	<i>ASTN2</i>	intergenic	A/G	0.11	2158	6.15×10^{-4}	1.53	3259	0.09	1.19	5.14×10^{-4}	1.31	++	60.43	0.11
rs1333144	1	170364401	<i>LOC284688</i>	intergenic	G/A	0.41	2157	5.40×10^{-3}	0.82	4654	0.02	0.9	5.92×10^{-4}	0.88	-	27.92	0.24
rs17530234	7	40783104	<i>C7orf10</i>	intronic	G/A	0.11	2155	5.49×10^{-3}	1.38	3258	0.03	1.24	6.25×10^{-4}	1.3	++	0	0.49
rs2462683	7	112971889	<i>PPP1R3A</i>	intergenic	G/A	0.34	2158	6.34×10^{-4}	1.3	3257	0.1	1.11	6.51×10^{-4}	1.18	++	59.06	0.12
rs11749727	5	179540965	<i>RASGEF1C</i>	intronic	G/A	0.44	2156	6.75×10^{-3}	1.21	3258	0.03	1.15	6.59×10^{-4}	1.17	++	0	0.55
rs475479	16	64324819	<i>LOC283867</i>	intergenic	G/A	0.26	2158	5.41×10^{-4}	0.76	3259	0.1	0.9	7.12×10^{-4}	0.83	-	61.71	0.12

Table 2. Cont.

SNP	CHR	Base Position	Nearby Gene	SNP Location	Alleles (Effect/Other)	MAF	Discovery Phase			Replication Phase			Meta-Analysis				
							N	p-Value	OR	N	p-Value	OR	p-Value	OR	Dir	I	Q
rs35814902	5	86835416	CCNH	intergenic	A/G	0.32	2158	5.04×10^{-3}	1.24	4573	0.02	1.11	8.01×10^{-4}	1.13	++	38.95	0.2
rs17529882	7	40761636	C7orf10	intronic	G/A	0.09	2157	2.10×10^{-3}	1.48	3259	0.08	1.22	8.28×10^{-4}	1.32	++	26.33	0.24
rs17456070	1	87599332	LOC100505768	intronic	G/A	0.30	2155	0.05	0.86	3257	5.98×10^{-3}	0.83	8.56×10^{-4}	0.84	-	0	0.74
rs12650617	4	5238437	STK32B	intronic	A/G	0.22	2157	3.35×10^{-5}	1.45	3245	0.3	1.07	8.81×10^{-4}	1.21	++	85.49	8.70×10^{-3}
rs13177543	5	86842168	CCNH	intergenic	A/G	0.32	2158	5.04×10^{-3}	1.24	4658	0.03	1.1	9.93×10^{-4}	1.13	++	42.89	0.18

Novel loci associated with compound MetS phenotype in meta-analysis, adjusted for age, sex, and first two principal components in discovery phase, and age and sex in replication phase. Meta-analysis was done using fixed effect inverse variance method in PLINK. SNP location is position of SNP in context of gene. CHR: Chromosome; MAF: Minor allele frequency; N: Sample number; OR: Odds ratio; Dir: Direction; I: I2 heterogeneity index (0–100); Q: p-value for Cochran's Q statistic. Direction ++/- features a concordance between the discovery and replication phase.

3.2. Conditional Analysis of *CETP* Locus

To classify robust independent variants within the *CETP* locus, conditional analysis was performed using a logistic regression model. Identified GWAS significant *CETP* loci—rs1800775 and rs3816117—were the primary leading signals in the MetS-associated *CETP* region (Supplemental Table S5). Association of other variants with MetS were lost upon adjusting for genotypes of rs1800775 and rs3816117.

However, the remaining five *CETP* loci—rs7205804, rs1532624, rs3764261, rs247617, and rs173539—also had prominent influence on leading GWAS loci (rs1800775 and rs3816117), as adjustment with genotypes of these five loci resulted in loss of genome-wide significance of GWAS loci, though nominal significance was maintained (Supplemental Table S5).

All the identified *CETP* SNPs are in moderate–high LD with each other in Indians ($r^2 = 0.32$ – 0.99 , $D' = 0.75$ – 1) (Supplemental Figure S5).

3.3. Haplotype Association Analysis

Haplotype analysis revealed a stronger haplotype within *CETP* locus (OR = 1.26, $p = 7.97 \times 10^{-8}$ for GCCCAGC haplotype) harboring risk alleles of SNPs rs173539, rs247617, rs3764261, rs1800775, rs3816117, rs7205804, and rs1532624 respectively for association with MetS in Indians (Supplemental Table S6).

3.4. *SFRP1*—A Novel Sub-GWAS Locus for MetS in Indians

We searched for all reported genetic variants within/near our novel sub-GWAS significant *SFRP1* locus. The region has never been reported for metabolic syndrome [43]. However, an intergenic signal-rs973441 near *SFRP1* has been robustly associated ($p = 3.4 \times 10^{-8}$) with Type 2 diabetes in a GWAS conducted in Europeans [43]. Indeed, the *SFRP1* locus has also been modestly associated with all MetS component phenotypes including—waist circumference, fasting glucose, HDL cholesterol, triglycerides, and systolic and diastolic blood pressure, as evident from literature [43].

3.5. Imputation Analysis of Novel Locus

For *SFRP1*, we perceived few variants that exhibited greater significance with MetS than index variant rs16890462 (Supplemental Table S7). Imputation analysis pipeline has been briefly described in Supplemental Figure S7. A few imputed variants (rs57208963, rs73628732, rs58109926, rs11986767, and rs76305295) were positioned in key regulatory elements and modulated strong binding sites for several key transcription factors like RREB1, Nkx2-5, SREBF2, HNF4G, and EBF1, respectively, that have been already implicated earlier in the context of metabolic disease and associated complications. The identified variants may affect the binding of these transcription factors to *SFRP1* genic regions, and thereby alter the transcriptional and translational levels of *SFRP1* mRNA and protein levels.

3.6. *SFRP1*, a Biologically Relevant Locus

Interestingly, the *SFRP1* gene was found to be considerably expressed in human subcutaneous/visceral adipose tissue and kidney-cortex, and weakly in liver, skeletal muscle, pancreas, and whole blood [Supplemental Figure S8]. This gene encodes SFRP1 protein (secreted frizzled related protein 1), which is a soluble inhibitor of Wnt/ β -catenin signaling pathway [54], a key pathway that maintains adipocyte differentiation [64,65]. It has been shown that adding recombinant SFRP1 protein to 3T3-L1 adipocyte cells hampers the antiadipogenic Wnt/ β -catenin pathway and stimulates preadipocyte differentiation [66]. Further, another interesting study found *SFRP1* deficient mice to display augmented adiposity, deregulated glucose homeostasis, and elevated inflammation in response to diet induced obesity [67].

In humans, both RNA and protein levels of SFRP1 are increased in slightly obese subjects that get lowered in morbidly obese individuals due to extreme body weight [66]. Moreover, increased

obesity results in elevated proinflammatory cytokine secretion and macrophage-infiltration where SFRP1 and Wnt5a are speculated to modulate the inflammatory response. For instance, Wnt5a, which is secreted by antigen presenting cells in joints of rheumatoid arthritis patients, facilitates production of cytokines, like Interleukins (IL-1, IL-6 and IL-8) via Fzd5-CamKII non-canonical Wnt signaling [68]. SFRP1 has been demonstrated to hinder this process [69] and also inhibits the activation of leukocytes and cytokine production in vitro [70], in addition to reducing infiltration of neutrophils in ischemic tissue in vivo [71]. These studies indicate an important role of *SFRP1* in fine tuning the adipogenesis, glucose metabolism and, inflammatory response.

Besides, adipose tissue, *SFRP1* is also considerably expressed in the kidney cortex (Supplemental Figure S8). The Wnt/ β -catenin pathway, which is also regulated by SFRP1 protein, is a critical regulator of various important cellular functions including maintenance of homeostatic state, embryonic development, and tissue injury [72]. This pathway also gets activated during kidney development and renal injury besides adipocyte differentiation [64,65,73].

A study found increased SFRP1 protein expression in mice models of kidney injury [74]. In addition, they also found that in these models, kidneys of *SFRP1* knock-out mice showed enhanced renal fibrosis, suggesting it as a protective factor to inhibit renal fibrosis, an initial stage for successive renal diseases [74]. This is in agreement with *SFRP1*'s protective role for influencing a few metabolic syndrome component phenotypes as well [67]. Also, individuals with metabolic syndrome are under greater risk for microalbuminuria/chronic kidney diseases in later stages, depending on the number of aggregated components of metabolic syndrome [75]. In some instances, hypertension, which is one of the components of metabolic syndrome, is considered a prime risk factor for kidney related complications [76]. Thus, metabolic syndrome represents an initial precursor for renal complications where *SFRP1* may serve as an essential protective factor linking metabolic syndrome with renal complications.

Further, to decode the likely functional role of our identified *SFRP1* variant, rs16890462, we explored the open chromatin features, active and repressive histone modifications (H3K36me3 and H3K27me3 respectively), transcription factor (TF), and CTCF binding sites in human subcutaneous adipose tissue. ATAC-seq and histone marks data suggested the desired region under closed chromatin state in subcutaneous adipose tissue, displayed by substantially higher enrichment for H3K27me3 and CTCF marks and an absence of H3K36me3 marks, which symbolizes repressive chromatin state (Figure 2). Further, dynamic peaks of CTCF binding and hypermethylation (from WGBS data) at variant regions in adipose tissue indicates its potential regulatory activity [Figure 2]. Intriguingly, the variant represented strong binding sites for Nfe212 (also known as NRF2), a crucial transcription factor implicated in obesity and metabolic syndrome [77].

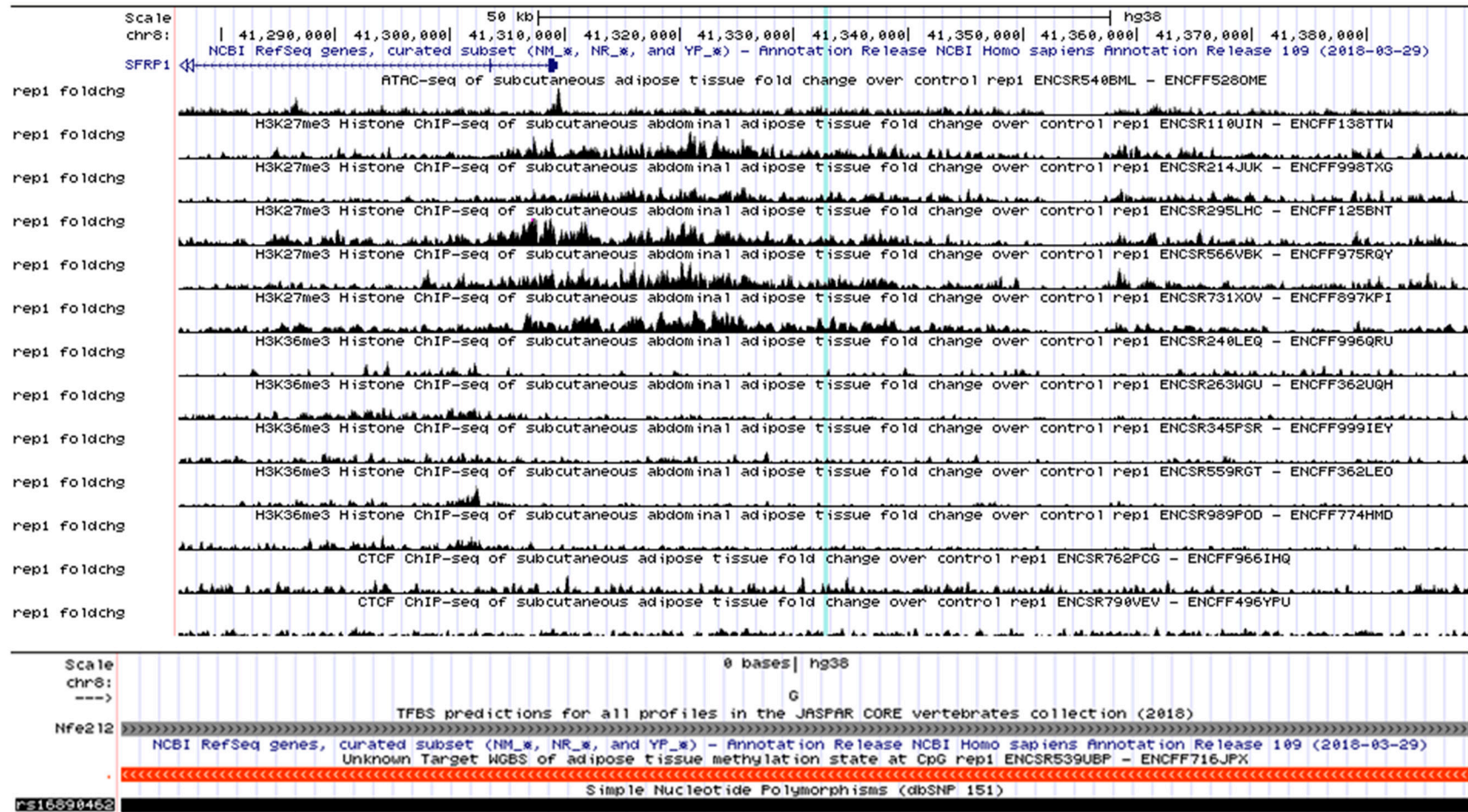


Figure 2. Regulatory features of *SFRP1* locus - rs16890462. Gene regulatory signatures in human subcutaneous adipose tissue. H3K36me3: Active transcription; H3K27me3: Repressed transcription; ATAC-seq peaks: Open chromatin; CTCF: TF that is enriched in repressed genic regions [ENCODE data]. Predicted binding sites for TFs where gray scale denotes enrichment of TF [JASPAR data].

4. Conclusions

Lipid metabolism is closely coupled with energy balance cycle and physiological homeostasis. Lipid molecules are involved in diverse biological functions ranging from acting as long-term energy depots to a variety of signaling functions. Deregulated lipid metabolism is increasingly being implicated in plentiful metabolic diseases including obesity, cardiovascular disease, and diabetes. *CETP* as a quantitative trait locus for lipid levels in human body is already established. Here, our study reveals association of *CETP* loci with MetS, which is a precursor for numerous complex disease phenotypes. So, in future, *CETP* may serve as a potential drug target for MetS.

In conclusion, our study assigns *CETP*, a known gene controlling lipid homeostasis, as a major locus for regulating MetS pathophysiology in Indians. We also discovered a novel sub-GWAS locus in *SFRP1*, which has already been functionally tested in mice and humans to regulate a few individual components of MetS, including obesity and glucose metabolism.

Supplementary Materials: The following are available online. Figure S1: Brief data analysis pipeline employed in the study; Figure S2: Data quality control executed in discovery and replication phase of the study; Figure S3: Statistical power of the study; Figure S4: Quantile-Quantile plot (QQ plot) between calculated and theoretical distribution of *p*-values in discovery phase; Figure S5: Pairwise linkage disequilibrium (LD) between the 7 *CETP* variants associated with MetS in the present study; Figure S6: Regional association plots of *CETP* signals in discovery phase, replication phase, and meta-analysis; Figure S7: Imputation flow chart for signal near *SFRP1*; Figure S8: Gene expression summary of *SFRP1* in major human tissues related to MetS; Table S1: Characteristics of study population; Table S2: Association status of known metabolic traits associated GWAS variants with MetS in replication phase in Indians ($p < 1 \times 10^{-3}$); Table S3: Reported eQTL and trait/disease associations of *CETP* region; Table S4: Association status of known metabolic traits associated GWAS variants with MetS in replication phase in Indians ($1 \times 10^{-3} < p > 0.05$); Table S5: Conditional analysis of *CETP* variants in Indians; Table S6: Haplotype association analysis of *CETP* locus for rs173539, rs247617, rs3764261, rs1800775, rs3816117, rs7205804, and rs1532624 respectively; Table S7: Association analysis of novel imputed *SFRP1* variants with MetS in Indians.

Author Contributions: Conceptualization, N.T. and D.B.; methodology, G.P., K.B., A.K.G., Y.K., P.C., V.P., INDICO, S.M., S.V.M., P.V., A.B. (PGIMER-Chnadigarh), R.K.M., A.B. (NIBMG-Kalyani), N.T. and D.B.; formal analysis, G.P.; data curation, G.P.; writing—original draft preparation, G.P.; writing—review and editing, G.P., A.B. (NIBMG-Kalyani), N.T. and D.B.; supervision, D.B.; project administration, N.T. and D.B.; funding acquisition, D.B.

Funding: This work was supported by Council of Scientific and Industrial Research [CSIR], Government of India through Centre for Cardiovascular and Metabolic Disease Research [CARDIOMED] grant [Grant No.: BSC0122] given to CSIR-Institute of Genomics and Integrative Biology, New Delhi, India. This study was also partially funded by Department of Science & Technology, Government of India through PURSE II CDST/SR/PURSE PHASE II/11 provided to Jawaharlal Nehru University, New Delhi, India.

Acknowledgments: The authors are greatly thankful to all the study participants. We acknowledge the support and participation of members of the INDICO consortium in data generation. GP and AKG acknowledge University Grants Commission (UGC), Government of India for Senior Research Fellowship. KB acknowledges Council of Scientific and Industrial Research (CSIR), Government of India for Senior Research Fellowship.

Conflicts of Interest: The authors declare no conflict of interest.

Members of INDICO: Viswanathan Mohan, Saurabh Ghosh, Abhay Sharma, Rubina Tabassum, Ganesh Chauhan, Anubha Mahajan, Om Prakash Dwivedi, Lakshmi Ramakrishnan, Radha Venkatesan, M. Chidambaram, D. Prabhakaran, K.S Reddy, Monisha Banerjee, Madhukar Saxena, Sandeep Mathur, Anil Bhansali, Viral Shah, S.V. Madhu, R.K. Marwaha, Pradeep Venkatesh, S.K. Aggarwal, Shantanu SenGupta, Sreenivas Chavali, Amitabh Sharma, Analabha Basu, Khushdeep Bandesh, Anil K Giri, Shraddha Chakraborty, Yasmeen Kauser, B. Abitha, Aditya Undru, Donaka Rajashekar, Gauri Prasad, Punam Jha, Priyanka Banerjee, Vaisak Parekatt, Suki Roy, Anjali Singh, Nikhil Tandon and Dwaipayana Bharadwaj (Coordinator).

References

1. Alberti, K.G.M.M.; Zimmet, P.; Shaw, J. Metabolic syndrome—A new world-wide definition. A consensus statement from the International Diabetes Federation. *Diabet. Med.* **2006**, *23*, 469–480. [[CrossRef](#)] [[PubMed](#)]
2. Mottillo, S.; Filion, K.B.; Genest, J.; Joseph, L.; Pilote, L.; Poirier, P.; Rinfret, S.; Schiffrin, E.L.; Eisenberg, M.J. The metabolic syndrome and cardiovascular risk: A systematic review and meta-analysis. *J. Am. Coll. Cardiol.* **2010**, *56*, 1113–1132. [[CrossRef](#)] [[PubMed](#)]

3. Esposito, K.; Chiodini, P.; Colao, A.; Lenzi, A.; Giugliano, D. Metabolic syndrome and risk of cancer: A systematic review and meta-analysis. *Diabetes Care*. **2012**, *35*, 2402–2411. [[CrossRef](#)] [[PubMed](#)]
4. Hui, W.S.; Liu, Z.; Ho, S.C. Metabolic syndrome and all-cause mortality: A meta-analysis of prospective cohort studies. *Eur. J. Epidemiol.* **2010**, *25*, 375–384. [[CrossRef](#)] [[PubMed](#)]
5. Misra, A.; Khurana, L. Obesity and the metabolic syndrome in developing countries. *J. Clin. Endocrinol. Metab.* **2008**, *93*, S9–S30. [[CrossRef](#)] [[PubMed](#)]
6. Prasad, D.S.; Kabir, Z.; Dash, A.K.; Das, B.C. Prevalence and risk factors for metabolic syndrome in Asian Indians: A community study from urban Eastern India. *J. Cardiovasc. Dis. Res.* **2012**, *3*, 204–211. [[CrossRef](#)]
7. Khan, Y.; Lalchandani, A.; Gupta, A.; Khadanga, S.; Kumar, S. Prevalence of metabolic syndrome crossing 40% in Northern India: Time to act fast before it runs out of proportions. *J. Fam. Med. Prim. Care* **2018**, *7*, 118–123. [[CrossRef](#)]
8. Tandon, N.; Garg, M.K.; Singh, Y.; Marwaha, R.K. Prevalence of metabolic syndrome among urban Indian adolescents and its relation with insulin resistance (HOMA-IR). *J. Pediatr. Endocrinol. Metab.* **2013**, *26*, 1123–1130. [[CrossRef](#)]
9. Joy, T.; Hegele, R.A. Genetics of Metabolic syndrome: Is there a role for phenomics? *Curr. Atheroscler. Rep.* **2008**, *10*, 201–208. [[CrossRef](#)]
10. Henneman, P.; Aulchenko, Y.S.; Frants, R.R.; Zorkoltseva, I.V.; Zillikens, M.C.; Frolich, M.; Oostra, B.A.; Van Dijk, K.W.; Van Duijn, C.M. Genetic architecture of plasma adiponectin overlaps with the genetics of metabolic syndrome-related traits. *Diabetes Care* **2010**, *33*, 908–913. [[CrossRef](#)]
11. Terán-García, M.; Bouchard, C. Genetics of the metabolic syndrome. *Appl. Physiol. Nutr. Metab.* **2007**, *32*, 89–114. [[CrossRef](#)]
12. Watanabe, R.M.; Valle, T.; Hauser, E.R.; Ghosh, S.; Eriksson, J.; Kohtamaki, K.; Ehnholm, C.; Tuomilehto, J.; Collins, F.S.; Bergman, R.N.; et al. Familiality of quantitative metabolic traits in Finnish families with non-insulin-dependent diabetes mellitus. Finland-United States Investigation of NIDDM Genetics (FUSION) Study investigators. *Hum. Hered.* **1999**, *49*, 159–168. [[CrossRef](#)]
13. Willer, C.J.; Schmidt, E.M.; Sengupta, S.; Peloso, G.M.; Gustafsson, S.; Kanoni, S.; Ganna, A.; Chen, J.; Buchkovich, M.L.; Mora, S.; et al. Discovery and refinement of loci associated with lipid levels. *Nat. Genet.* **2013**, *45*, 1274–1283. [[CrossRef](#)]
14. Bandesh, K.; Prasad, G.; Giri, A.K.; Kauser, Y.; Upadhyay, M.; Basu, A.; Tandon, N.; Bharadwaj, D. Genome-wide association study of blood lipids in Indians confirms universality of established variants. *J. Hum. Genet.* **2019**, *64*, 573–587. [[CrossRef](#)]
15. Giri, A.K.; Prasad, G.; Bandesh, K.; Parekatt, V.; Mahajan, A.; Banerjee, P.; Kauser, Y.; Chakraborty, S.; Rajashekar, D.; INDICO; et al. Multifaceted genome-wide study identifies novel regulatory loci for body mass index in Indians. *BioRxiv* **2019**. [[CrossRef](#)]
16. Dupuis, J.; Langenberg, C.; Prokopenko, I.; Saxena, R.; Soranzo, N.; Jackson, A.U.; Wheeler, E.; Glazer, N.L.; Bouatia-Naji, N.; Gloyn, A.L.; et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat. Genet.* **2010**, *42*, 105–116. [[CrossRef](#)]
17. Giri, A.; Hellwege, J.N.; Keaton, J.M.; Park, J.; Qiu, C.; Warren, H.R.; Torstenson, E.S.; Kovesdy, C.P.; Sun, Y.V.; Wilson, O.D.; et al. Trans-ethnic association study of blood pressure determinants in over 750,000 individuals. *Nat. Genet.* **2019**, *51*, 51–62. [[CrossRef](#)]
18. Kraja, A.T.; Vaidya, D.; Pankow, J.S.; Goodarzi, M.O.; Assimes, T.L.; Kullo, I.J.; Sovio, U.; Mathias, R.A.; Sun, Y.V.; Franceschini, N.; et al. A bivariate genome-wide approach to metabolic syndrome: STAMPEED Consortium. *Diabetes* **2011**, *60*, 1329–1339. [[CrossRef](#)]
19. Carty, C.L.; Bhattacharjee, S.; Haessler, J.; Cheng, I.; Hindorff, L.A.; Aroda, V.; Carlson, C.S.; Hsu, C.-N.; Wilkens, L.; Liu, S.; et al. Comparative Analysis of Metabolic Syndrome Components in over 15,000 African Americans Identifies Pleiotropic Variants: Results from the PAGE Study. *Circ. Cardiovasc. Genet.* **2014**, *7*, 505–513. [[CrossRef](#)]
20. Tekola-Ayele, F.; Doumatey, A.P.; Shriner, D.; Bentley, A.R.; Chen, G.; Zhou, J.; Fasanmade, O.; Johnson, T.; Oli, J.; Okafor, G.; et al. Genome-wide association study identifies African-ancestry specific variants for metabolic syndrome. *Mol. Genet. Metab.* **2015**, *116*, 305–313. [[CrossRef](#)]

21. Kristiansson, K.; Perola, M.; Tikkanen, E.; Kettunen, J.; Surakka, I.; Havulinna, A.S.; Stančáková, A.; Barnes, C.; Widen, E.; Kajantie, E.; et al. Genome-wide screen for metabolic syndrome susceptibility loci reveals strong lipid gene contribution but no evidence for common genetic basis for clustering of metabolic syndrome traits. *Circ. Cardiovasc. Genet.* **2012**, *5*, 242–249. [[CrossRef](#)]
22. Zhu, Y.; Zhang, D.; Zhou, D.; Li, Z.; Li, Z.; Fang, L.; Yang, M.; Shan, Z.; Li, H.; Chen, J.; et al. Susceptibility loci for metabolic syndrome and metabolic components identified in Han Chinese: A multi-stage genome-wide association study. *J. Cell. Mol. Med.* **2017**, *21*, 1106–1116. [[CrossRef](#)]
23. Lee, H.S.; Kim, Y.; Park, T. New Common and Rare Variants Influencing Metabolic Syndrome and Its Individual Components in a Korean Population. *Sci. Rep.* **2018**, *8*, 5701. [[CrossRef](#)]
24. Lin, E.; Kuo, P.-H.; Liu, Y.-L.; Yang, A.C.; Tsai, S.-J. Detection of susceptibility loci on APOA5 and COLEC12 associated with metabolic syndrome using a genome-wide association study in a Taiwanese population. *Oncotarget* **2017**, *8*, 93349–93359. [[CrossRef](#)]
25. Vattikuti, S.; Guo, J.; Chow, C.C. Heritability and genetic correlations explained by common SNPs for metabolic syndrome traits. *PLoS Genet.* **2012**, *8*, e1002637. [[CrossRef](#)]
26. Indian Genome Variation Consortium. The Indian Genome Variation database (IGVdb): A project overview. *Hum Genet.* **2005**, *118*, 1–11. [[CrossRef](#)]
27. Tabassum, R.; Chauhan, G.; Dwivedi, O.P.; Mahajan, A.; Jaiswal, A.; Kaur, I.; Bandesh, K.; Singh, T.; Mathai, B.J.; Pandey, Y.; et al. Genome-wide association study for type 2 diabetes in Indians identifies a new susceptibility locus at 2q21. *Diabetes* **2013**, *62*, 977–986. [[CrossRef](#)]
28. Mohan, V.; Rao, G.H.R. *Type 2 Diabetes in South Asians*, 1st ed.; South Asian Society on Atherosclerosis and Thrombosis: New Delhi, India, 2007.
29. Rao, G.H.R.; Thanickachalam, S. *Coronary Artery Disease: Risk Promoters, Pathophysiology and Prevention*, 1st ed.; South Asian Society on Atherosclerosis and Thrombosis: New Delhi, India, 2005.
30. Prasad, D.S.; Kabir, Z.; Dash, A.K.; Das, B.C. Abdominal obesity, an independent cardiovascular risk factor in Indian subcontinent: A clinico epidemiological evidence summary. *J. Cardiovasc. Dis. Res.* **2011**, *2*, 199–205. [[CrossRef](#)]
31. Zabaneh, D.; Balding, D.J. A genome-wide association study of the metabolic syndrome in Indian Asian men. *PLoS ONE* **2010**, *5*, e11961. [[CrossRef](#)]
32. Basu, A.; Sarkar-Roy, N.; Majumder, P.P. Genomic reconstruction of the history of extant populations of India reveals five distinct ancestral components and a complex structure. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 1594–1599. [[CrossRef](#)]
33. Indian Diabetes Consortium. INDICO: The development of a resource for epigenomic study of Indians undergoing socioeconomic transition. *Hugo J.* **2011**, *5*, 65–69. [[CrossRef](#)]
34. Grundy, S.M.; Cleeman, J.I.; Daniels, S.R.; Donato, K.A.; Eckel, R.H.; Franklin, B.A.; Gordon, D.J.; Krauss, R.M.; Savage, P.J.; Smith, S.C., Jr.; et al. Diagnosis and management of the metabolic syndrome: An American Heart Association/National Heart, Lung, and Blood Institute scientific statement: Executive Summary. *Crit. Pathw. Cardiol.* **2005**, *4*, 198–203. [[CrossRef](#)]
35. Mahajan, A.; Jaiswal, A.; Tabassum, R.; Podder, A.; Ghosh, S.; Madhu, S.V.; Mathur, S.K.; Tandon, N.; Bharadwaj, D. Elevated levels of C-reactive protein as a risk factor for Metabolic Syndrome in Indians. *Atherosclerosis* **2012**, *220*, 275–281. [[CrossRef](#)]
36. Giri, A.K.; Banerjee, P.; Chakraborty, S.; Kauser, Y.; Undru, A.; Roy, S.; Parekatt, V.; Ghosh, S.; Tandon, N.; Bharadwaj, D. Genome wide association study of uric acid in Indian population and interaction of identified variants with Type 2 diabetes. *Sci. Rep.* **2016**, *6*, 21440. [[CrossRef](#)]
37. Prasad, G.; Giri, A.K.; Basu, A.; Tandon, N.; Bharadwaj, D. Genomewide association study for C-reactive protein in Indians replicates known associations of common variants. *J. Genet.* **2019**, *98*, 20. [[CrossRef](#)]
38. Bandesh, K.; Prasad, G.; Giri, A.K.; Voruganti, V.S.; Butte, N.F.; Cole, S.A.; Comuzzie, A.G.; INDICO; Tandon, N.; Bharadwaj, D. Genomewide association study of C-peptide surfaces key regulatory genes in Indians. *J. Genet.* **2019**, *98*, 8. [[CrossRef](#)]
39. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [[CrossRef](#)]
40. Turner, S.D. qqman: An R package for visualizing GWAS results using Q-Q and manhattan plots. *J. Open Source Softw.* **2014**. [[CrossRef](#)]

41. Buniello, A.; MacArthur, J.A.L.; Cerezo, M.; Harris, L.W.; Hayhurst, J.; Malangone, C.; McMahon, A.; Morales, J.; Mountjoy, E.; Sollis, E.; et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics. *Nucleic Acids Res.* **2019**, *47*, D1005–D1012. [[CrossRef](#)]
42. Watanabe, K.; Stringer, S.; Frei, O.; Mirkov, M.U.; Polderman, T.J.C.; van der Sluis, S.; Andreassen, O.A.; Neale, B.M.; Posthuma, D. A global overview of pleiotropy and genetic architecture in complex traits. *BioRxiv* **2018**. [[CrossRef](#)]
43. Type 2 Diabetes Knowledge Portal. Available online: <http://www.type2diabetesgenetics.org/> (accessed on 25 May 2019).
44. Pruim, R.J.; Welch, R.P.; Sanna, S.; Teslovich, T.M.; Chines, P.S.; Gliedt, T.P.; Boehnke, M.; Abecasis, G.R.; Willer, C.J.; Frishman, D. LocusZoom: Regional visualization of genome-wide association scan results. *Proc. Bioinform.* **2011**, *26*, 2336–2337. [[CrossRef](#)]
45. Gauderman, W.J. Sample size requirements for association studies of gene-gene interaction. *Am. J. Epidemiol.* **2002**, *5*, 478–484. [[CrossRef](#)]
46. Delaneau, O.; Zagury, J.-F.; Marchini, J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat. Methods* **2013**, *10*, 5–6. [[CrossRef](#)]
47. Howie, B.N.; Donnelly, P.; Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **2009**, *5*, e1000529. [[CrossRef](#)]
48. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **2013**, *45*, 580–585. [[CrossRef](#)]
49. Davis, C.A.; Hitz, B.C.; Sloan, C.A.; Chan, E.T.; Davidson, J.M.; Gabdank, I.; Hilton, J.A.; Jain, K.; Baymuradov, U.K.; Narayanan, A.K.; et al. The Encyclopedia of DNA elements (ENCODE): Data portal update. *Nucleic Acids Res.* **2018**, *46*, D794–D801. [[CrossRef](#)]
50. Khan, A.; Fornes, O.; Stigliani, A.; Gheorghe, M.; Castro-Mondragon, J.A.; Van Der Lee, R.; Bessy, A.; Chèneby, J.; Kulkarni, S.R.; Tan, G.; et al. JASPAR 2018: Update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res.* **2018**, *46*, D260–D266. [[CrossRef](#)]
51. Casper, J.; Zweig, A.S.; Villarreal, C.; Tyner, C.; Speir, M.L.; Rosenbloom, K.R.; Raney, B.J.; Lee, C.M.; Lee, B.T.; Karolchik, D.; et al. The UCSC Genome Browser database: 2018 update. *Nucleic Acids Res.* **2018**, *46*, D762–D769. [[CrossRef](#)]
52. Hou, H.; Ma, R.; Guo, H.; He, J.; Hu, Y.; Mu, L.; Yan, Y.; Ma, J.; Li, S.; Zhang, J.; et al. Association between six CETP polymorphisms and metabolic syndrome in uyghur adults from Xinjiang, China. *Int. J. Environ. Res. Public Health* **2017**, *14*, E653. [[CrossRef](#)]
53. Machiela, M.J.; Chanock, S.J. LDlink: A web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* **2015**, *31*, 3555–3557. [[CrossRef](#)]
54. Rebhan, M.; Chalifa-Caspi, V.; Prilusky, J.; Lancet, D. GeneCards: Integrating information about genes, proteins and diseases. *Trends Genet.* **1997**, *13*, 163. [[CrossRef](#)]
55. Papp, A.C.; Pinsonneault, J.K.; Wang, D.; Newman, L.C.; Gong, Y.; Johnson, J.A.; Pepine, C.J.; Kumari, M.; Hingorani, A.D.; Talmud, P.J.; et al. Cholesteryl ester transfer protein (CETP) polymorphisms affect mRNA splicing, HDL levels, and sex-dependent cardiovascular risk. *PLoS ONE* **2012**, *7*, e31930. [[CrossRef](#)]
56. Lu, H.; Inazu, A.; Moriyama, Y.; Higashikata, T.; Kawashiri, M.A.; Yu, W.; Huang, Z.; Okamura, T.; Mabuchi, H. Haplotype analyses of cholesteryl ester transfer protein gene promoter: A clue to an unsolved mystery of TaqIB polymorphism. *J. Mol. Med.* **2003**, *81*, 246–255. [[CrossRef](#)]
57. Frisdal, E.; Klerkx, A.H.E.M.; Le Goff, W.; Tanck, M.W.T.; Lagarde, J.P.; Jukema, J.W.; Kastelein, J.J.P.; Chapman, M.J.; Guerin, M. Functional interaction between -629C/A, -971G/A and -1337C/T polymorphisms in the CETP gene is a major determinant of promoter activity and plasma CETP concentration in the REGRESS Study. *Hum. Mol. Genet.* **2005**, *14*, 2607–2618. [[CrossRef](#)]
58. Datchet, C.; Poirier, O.; Cambien, F.; Chapman, J.; Rouis, M. New functional promoter polymorphism, CETP/-629, in cholesteryl ester transfer protein (CETP) gene related to CETP mass and high density lipoprotein cholesterol levels: Role of Sp1/Sp3 in transcriptional regulation. *Arterioscler. Thromb. Vasc. Biol.* **2000**, *20*, 507–515. [[CrossRef](#)]
59. Thompson, A.; Di Angelantonio, E.; Sarwar, N.; Erqou, S.; Saleheen, D.; Dullaart, R.P.F.; Keavney, B.; Ye, Z.; Danesh, J. Association of cholesteryl ester transfer protein genotypes with CETP mass and activity, lipid levels, and coronary risk. *JAMA J. Am. Med. Assoc.* **2008**, *299*, 2777–2788. [[CrossRef](#)]

60. van der Harst, P.; Verweij, N. Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. *Circ. Res.* **2018**, *122*, 433–443. [[CrossRef](#)]
61. Iwanicka, J.; Iwanicki, T.; Niemiec, P.; Balcerzyk, A.; Krauze, J.; Górczyńska-Kosiorz, S.; Ochalska-Tyka, A.; Grzeszczak, W.; Żak, I. Relationship between CETP gene polymorphisms with coronary artery disease in Polish population. *Mol. Biol. Rep.* **2018**, *45*, 1929–1935. [[CrossRef](#)]
62. Inouye, M.; Ripatti, S.; Kettunen, J.; Lyytikäinen, L.P.; Oksala, N.; Laurila, P.P.; Kangas, A.J.; Soininen, P.; Savolainen, M.J.; Viikari, J.; et al. Novel Loci for Metabolic Networks and Multi-Tissue Expression Studies Reveal Genes for Atherosclerosis. *PLoS Genet.* **2012**, *8*, e1002907. [[CrossRef](#)]
63. Yang, J.; Benyamin, B.; McEvoy, B.P.; Gordon, S.; Henders, A.K.; Nyholt, D.R.; Madden, P.A.; Heath, A.C.; Martin, N.G.; Montgomery, G.W.; et al. Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **2010**, *42*, 565–569. [[CrossRef](#)]
64. Ross, S.E.; Hemati, N.; Longo, K.A.; Bennett, C.N.; Lucas, P.C.; Erickson, R.L.; MacDougald, O.A. Inhibition of adipogenesis by Wnt signaling. *Science* **2000**, *289*, 950–953. [[CrossRef](#)]
65. Bennett, C.N.; Hodge, C.L.; MacDougald, O.A.; Schwartz, J. Role of Wnt10b and C/EBP α in spontaneous adipogenesis of 243 cells. *Biochem. Biophys. Res. Commun.* **2003**, *302*, 12–16. [[CrossRef](#)]
66. Lagathu, C.; Christodoulides, C.; Tan, C.Y.; Virtue, S.; Laudes, M.; Campbell, M.; Ishikawa, K.; Ortega, F.; Tinahones, F.J.; Fernández-Real, J.M.; et al. Secreted frizzled-related protein 1 regulates adipose tissue expansion and is dysregulated in severe obesity. *Int. J. Obes.* **2010**, *34*, 1695–1705. [[CrossRef](#)]
67. Gauger, K.J.; Bassa, L.M.; Henchey, E.M.; Wyman, J.; Bentley, B.; Brown, M.; Shimono, A.; Schneider, S.S. Mice deficient in Sfrp1 exhibit increased adiposity, dysregulated glucose metabolism, and enhanced macrophage infiltration. *PLoS ONE* **2013**, *8*, e78320. [[CrossRef](#)]
68. Blumenthal, A.; Ehlers, S.; Lauber, J.; Buer, J.; Lange, C.; Goldmann, T.; Heine, H.; Brandt, E.; Reiling, N. The Wingless homolog WNT5A and its receptor Frizzled-5 regulate inflammatory responses of human mononuclear cells induced by microbial stimulation. *Blood* **2006**, *108*, 965–973. [[CrossRef](#)]
69. Pereira, C.; Schaer, D.J.; Bachli, E.B.; Kurrer, M.O.; Schoedon, G. Wnt5A/CaMKII signaling contributes to the inflammatory response of macrophages and is a target for the antiinflammatory action of activated protein C and interleukin-10. *Arterioscler. Thromb. Vasc. Biol.* **2008**, *28*, 504–510. [[CrossRef](#)]
70. Barandon, L.; Casassus, F.; Leroux, L.; Moreau, C.; Allières, C.; Lamazière, J.M.D.; Dufourcq, P.; Couffignal, T.; Duplâa, C. Secreted frizzled-related protein-1 improves postinfarction scar formation through a modulation of inflammatory response. *Arterioscler. Thromb. Vasc. Biol.* **2011**, *31*, e80–e87. [[CrossRef](#)]
71. Barandon, L.; Couffignal, T.; Ezan, J.; Dufourcq, P.; Costet, P.; Alzieu, P.; Leroux, L.; Moreau, C.; Dare, D.; Duplâa, C. Reduction of Infarct Size and Prevention of Cardiac Rupture in Transgenic Mice Overexpressing FrzA. *Circulation* **2003**, *108*, 2282–2289. [[CrossRef](#)]
72. Moon, R.T.; Kohn, A.D.; De Ferrari, G.V.; Kaykas, A. WNT and β -catenin signalling: Diseases and therapies. *Nat. Rev. Genet.* **2004**, *5*, 691–701. [[CrossRef](#)]
73. DiRocco, D.P.; Kobayashi, A.; Taketo, M.M.; McMahon, A.P.; Humphreys, B.D. Wnt4/ β -Catenin Signaling in Medullary Kidney Myofibroblasts. *J. Am. Soc. Nephrol.* **2013**, *24*, 1399–1412. [[CrossRef](#)]
74. Matsuyama, M.; Nomori, A.; Nakakuni, K.; Shimono, A.; Fukushima, M. Secreted frizzled-related protein 1 (Sfrp1) regulates the progression of renal fibrosis in a mouse model of obstructive nephropathy. *J. Biol. Chem.* **2014**, *289*, 31526–31533. [[CrossRef](#)]
75. Locatelli, F.; Pozzoni, P.; Del Vecchio, L. Renal manifestations in the metabolic syndrome. *J. Am. Soc. Nephrol.* **2006**, *17*, S81–S85. [[CrossRef](#)]
76. Johnson, R.J.; Segal, M.S.; Srinivas, T.; Ejaz, A.; Mu, W.; Roncal, C.; Sánchez-Lozada, L.G.; Gersch, M.; Rodriguez-Iturbe, B.; Kang, D.-H.; et al. Essential Hypertension, Progressive Renal Disease, and Uric Acid: A Pathogenetic Link? *J. Am. Soc. Nephrol.* **2005**, *16*, 1909–1919. [[CrossRef](#)]
77. Zhang, Z.; Zhou, S.; Jiang, X.; Wang, Y.H.; Li, F.; Wang, Y.G.; Zheng, Y.; Cai, L. The role of the Nrf2/Keap1 pathway in obesity and metabolic syndrome. *Rev. Endocr. Metab. Disord.* **2015**, *16*, 35–45. [[CrossRef](#)]

