A Pooled Blood Genome-Wide Association Study of Hypertension in Sindhi Families: Results from the DISFIN Study

Samika Kanaskar , Ashwini A Patel , Manisha T Jaisinghani , Kanchan V Pipal , Mangesh Kanaskar , Manju Mamtani , Hemant Kulkarni *

Posted Date: 10 March 2026

doi: 10.20944/preprints202603.0818.v1

Keywords: type 2 diabetes; genome-wide association study; ethnicity

 doi:10.20944/preprints202603.0818.v1

Article

A Pooled Blood Genome-Wide Association Study of Hypertension in Sindhi Families: Results from the DISFIN Study

Samika Kanaskar [1,†], Ashwini A Patel [2,†], Manisha T Jaisinghani [2], Kanchan V Pipal [2], Mangesh Kanaskar [3], Manju Mamtani [2,4] and Hemant Kulkarni [2,4,*]

[1] Department of Public Health & Health Professions, University of Florida, Gainesville, Florida, USA
[2] Lata Medical Research Foundation, Nagpur, India
[3] Precisely Software, Inc, Burlington, Massachusetts, USA
[4] M&H Research, LLC, San Antonio, Texas, USA
* Correspondence: hemant.kulkarni@mnhresearch.com
† Equal contribution as first author.

Abstract

Hypertension is an important target for primordial prevention of complex, noncommunicable diseases and its prevalence remains high across populations. Urban population in India is at a high risk of hypertension but the genetic basis of hypertension in this population remains poorly understood. We conducted a pooled whole-blood genome-wide association study of 28 pools representing 1,402 participants of the Diabetes In Sindhi Families In Nagpur (DISFIN) study which enrolled families of probands with type 2 diabetes (T2D). Genotyping was done using Illumina's Global Screening Array. From a total of 608,550 single nucleotide variants, 191 were found to be significantly associated with hypertension even after adjusting for metabolic comorbidities, batch effects, pooling error, kinship status and pooling variation. These variants mapped to 180 well-characterized genes that comprised 55 (31%) genes encoding long noncoding RNA (lncRNA). Many of the genes significantly associated with hypertension (including 35% of the lncRNAs) have also been reported by other studies. However, we identified novel genes (SBF2, ARHGAP12, EPAS1, CLEC16A and LRPPRC) to be associated with hypertension. The most significantly associated lncRNA gene was FLYWCH-AS1. Bioinformatic analyses indicated that these novel genes are likely to have functional importance in hypertension. Our study thus points to the potential candidate genes associated with hypertension in endogamous Sindhi families with T2D patients. The replicable and functional role of these candidate genes should be investigated in future studies.

Keywords: type 2 diabetes; genome-wide association study; ethnicity

1. Introduction

Hypertension continues to be a common primordial risk factor for several cardiometabolic conditions including diabetes, cardiovascular diseases, and chronic kidney disease. The World Global Report on hypertension estimated that the prevalence of hypertension was 33% in the age group of 30-79 years. Further, only 54% of those with hypertension are diagnosed, 42% are receiving treatment, and only 21% successfully control hypertension.[1] Interestingly, essential hypertension – where the cause of hypertension is unknown – is known to be influenced in part both by the genetic and environmental risk factors as well as by the interactions between genetic and environmental factors. Previous studies from various parts of the world have been elegantly summarized [2–4] to reveal that the estimated heritability of hypertension ranges between 30-60%. A common approach

 doi:10.20944/preprints202603.0818.v1

to understanding the genetic basis of complex diseases like hypertension is to conduct a genome-wide association study (GWAS) that aims to identify key genetic variants associated with this disease. To date, over 2000 genetic variants have been identified in diverse populations to be associated with hypertension.[5] However, despite the 708 non-interactive and 38 environment-interactive genetic variants detailed by Waken et al[4] a comprehensive understanding of the genetic drivers of hypertension remains elusive.

The estimated prevalence of hypertension in India is high – 24% in males and 20% in females.[6,7] Recent studies from the INDIGENIUS Consortium have demonstrated that within different ethnic backgrounds in India the heritability estimates for systolic and diastolic blood pressure traits range between 0.11-0.39 and 0.13-0.38, respectively, indicating a noticeable genetic component to blood pressure.[8] Despite this knowledge, the genetic and genomic studies of hypertension in India have been few and far between. These studies have attempted to quantify association between genetic variants and blood pressure traits[9] as a part of a larger study but dedicated Indian-population-specific GWAS studies on hypertension and blood pressure related traits are currently lacking.

We conducted a pooled blood genome-wide association study of hypertension in the pedigrees of Indian Sindhi families enrolled in the Diabetes In Sindhi Families In Nagpur (DISFIN) study.[10] The DISFIN Study was designed with a focus on the genetics of type 2 diabetes but the prevalence of essential hypertension in this study was as high as 53% and thus provided us with an opportunity to conduct a genome-wide association study of hypertension as well. This population has a substantial coexistence of metabolic comorbidities such as type 2 diabetes, dyslipidemia, general obesity and central obesity. Since our work was constrained by project costs, we conducted a pooled GWAS study of hypertension. We used innovative statistical approaches to account for these comorbidities. Here, we report the results of our study that identified interesting insights into the genetic basis of hypertension in the urban Indian Sindhi families.

2. Materials and Methods

2.1. Study Participants

We used the clinical and genetic data collected during the DISFIN study. Participant enrollment and blood sample collection took place between 1 March 2016 and 28 February 2017. Details of the enrollment protocol, inclusion and exclusion criteria and the overall design of the study have been described previously.[10] Summarily, we enrolled endogamous Sindhi families with at least one case of type 2 diabetes per family to construct family pedigrees. Additional inclusion criteria were a resident of the study area (Jaripatka, Mecosabag and Khamla areas of Nagpur where the Sindhi ethnic population mostly resides); self-reported Sindhi ethnicity and age ≥20 years. Pregnant or lactating women and patients with type 1 diabetes (known or suggested by serum C peptide) were excluded. Following investigator-administered, semi-structured interviews and clinical examination, a trained phlebotomist collected blood samples for laboratory assays.

2.2. Definitions of Metabolic Conditions

Our study was designed to conduct a genome-wide interrogation in the context of hypertension. Hypertension was defined [11] as self-reported hypertension or currently on anti-hypertensives or systolic blood pressure ≥ 130 mmHg or diastolic blood pressure ≥ 85 mmHg. Since metabolic comorbidities such as type 2 diabetes, obesity, and dyslipidemia commonly coexist with hypertension, we also investigated the study participants with respect to these conditions. Type 2 diabetes was defined [12] as one or more of the following: self-reported diabetes; currently on anti-diabetics; fasting plasma glucose ≥ 126 mg/dL; random blood glucose ≥ 200 mg/dL; or HbA1c concentration ≥ 6.5%. Central obesity was defined based on cutoffs for Indian population [13] as a waist circumference ≥ 90 cm for males and ≥85 cm for females. Dyslipidemia was defined [14] as presence of any of the following: serum triglycerides ≥ 150 mg/dL or serum high density lipoproteins < 40 mg/dL (for males) or <50 mg/dL (for females).

 doi:10.20944/preprints202603.0818.v1

2.3. Pool Definitions for GWAS

We conducted a pooled whole-blood genome-wide association study. This technique is now well established as an acceptable alternative to individual genome-wide genotyping. [15] We previously conducted a pooled blood, genome-wide investigation for type 2 diabetes and have extensively described the definitions of whole blood pools used in the study.[16] Based on a combination of the presence of hypertension, central obesity, type 2 diabetes and dyslipidemia, we first generated a total of 16 potential combinations and reduced the number of pools to 14 (by collapsing those pool categories that had a frequency < 1%). To ensure repeatability, we ran the genotyping analyses on duplicated aliquots from pools. Therefore, we had a total of 28 whole-blood pools. The pool construction, DNA extraction and genotyping protocols have been described previously. [16] We used the Infinium Global Screening Array (GSA) for genotyping. The methods used for blood sample collection, storage, extraction of DNA and genotyping have been described elsewhere.[16]

2.4. Statistical Analyses

For each included variant on the GSA array, we estimated 28 B allele frequencies – one for each pool. We accounted for potential confounding by adjusting for the following covariates: comorbidities, batch effects, intra-replicate correlation, within-pool degree of kinship and the random effects across pools. To model these relationships statistically, we used the mixed effects logistic regression format to estimate the T statistic from a Wald test to test the significance of association of a given variant with the risk of hypertension. Specifically, we used the following regression model to estimate the strength and significance of association:

$$\text{logit(hypertension)} = \beta_0 + \beta_s\ \text{BAF} + \beta_1\ \text{T2D} + \beta_2\ \text{COB} + \beta_3\ \text{DYL} + \beta_b\ \text{BAT} + \beta_r\ \text{REP} + \beta_k\ \text{PHI} + \text{RE(POOL)}$$

where, hypertension is an indicator variable for presence of hypertension; BAF is a continuous variable indicating the B allele frequency; T2D, COB and DYL represent the concomitant presence of type 2 diabetes, central obesity and dyslipidemia, respectively; BAT is an indicator variable for the chip identifier; REP represents the replicate ID; PHI is the within pool degree of kinship and RE(POOL) represents the random effects across the study pools. The regression coefficients in the equation were used to quantify the differential influence of the B allele frequency (β_s), influence of comorbidities (β_1–β_3), batch effect (β_b), pooling error (β_r) and kinship effect (β_k). All the models were weighted by the pool frequency. A rigid Bonferroni correction was used to account for multiple testing and the global type I error rate was thus adjusted to 8.216×10^{-8}. These analyses were conducted using dedicated scripts in R. Manhattan and QQ plots were generated using the qqman library [17] in R. Pooling error was estimated per MacGregor et al.[18] All R scripts used in this study are described in an annotated fashion in Supplementary Notes 1-3.

2.5. Functional Relevance of Strongly Associated Variants and Genes

For annotation of the variants and for a comprehensive, genomic understanding of their role in health and disease we used the SNPAnnotator R package.[19] In addition to the in-built abilities of SNPAnnotator, we also used the g:Profiler online tool (https://biit.cs.ut.ee/gprofiler/gost) to conduct gene set enrichment analyses. The results from both SNPAnnotator and g:Profiler were reported as p-values after controlling the false discovery rate (FDRp). Deleteriousness of variants was estimated using the Combined Annotation Dependent Depletion (CADD, https://cadd.gs.washington.edu) score.

3. Results

3.1. Study Participants and Pools

This study represents a secondary analysis of the data derived from the DISFIN study. As described elsewhere, the study included a total of 1,444 participants representing 112 endogamous

 doi:10.20944/preprints202603.0818.v1

Sindhi families in Nagpur, Maharashtra, India. Our study combined the whole-blood samples with clinical data (n=1,402) collected into 28 pools based on the presence or absence of four dichotomous clinical traits: hypertension, type 2 diabetes, dyslipidemia and obesity. For this study, two pools were derived from the 28 pools to compare the presence or absence of hypertension as the trait of interest. The pool with hypertension represented whole-blood pooling of 742 (52.92%) participants. Using the same inclusion criteria for genetic variants as described previously,[16] we included a total of 608,550 autosomal single nucleotide polymorphisms (SNP) with a minor allele frequency >0.1. It is noteworthy, that the genotyping error (as measured using the GenTrain score) and the pooling error estimates were acceptable.[16]

3.2. Heritability of Blood Pressure-Related Traits

We first estimated the heritability of blood pressure traits in the study population. We found that the heritability estimates for the continuous traits: systolic blood pressure, diastolic blood pressure, pulse pressure and mean arterial pressure were 0.24 (SE 0.08, $p = 0.0010$), 0.31 (SE 0.08, $p = 0.0004$), 0.18 (SE 0.08, $p = 0.0106$) and 0.15 (SE 0.08, $p = 0.0149$), respectively. The heritability of hypertension (estimated using the liability threshold approach for a dichotomous trait) was 0.44 (SE 0.14, $p = 0.0003$), indicating that all traits studied here in the context of blood pressure showed a statistically significant and clinically meaningful heritability. For the genome-wide association study, we focused on the dichotomous trait of hypertension.

3.3. Pooled GWAS Results at the Level of Variant

Of the 608,550 autosomal markers studied here, we found that a total of 191 variants were significantly associated with hypertension even after adjusting for the covariates listed in the Methods section and the multiple comparisons alluded to earlier. The fully annotated description of the significantly associated 191 variants is provided in Supplementary Table 1. The genome-wide association pattern observed in the present study is shown in Figure 1A. Also, as shown in the QQ plot depicted in Figure 1C, we found that the genomic inflation factor (λ) was below unity indicating that there was negligible genomic inflation during genotypic assays. Of note, only 9 variants were associated with a CADD score between 10 and 20 while three variants (rs28933396, rs74740987, and rs10075131) were associated with a CADD score above 20. This indicated that with respect to deleteriousness, majority of the variants were benign.

The top five most significant markers (highlighted in Figure 1A and 1B) were following single nucleotide polymorphisms: rs7200229, rs7167587, rs3098945, rs1316826, and rs1514414. Queries run through the SNPAnnotator package identified two of these five polymorphisms – rs7200229 and rs3098945. The rs7200229 SNP is a non-coding exon variant associated with the FLYWCH1-AS1 gene on chromosome 16. The rs3098945 polymorphism is an intronic variant in the ANKRD13B gene on chromosome 17. Interestingly, 23 of the 191 variants have been previously reported by other genome-wide association studies (as queried against the Human Phenotype Ontology database) indicating that our study could replicate several of the known associations in the context of GWAS. The subset of variants found in GWASCatalog to be associated with blood-pressure related traits (shown in Figure 2) was consistent with this finding. The observed network of SNPs and disease association from the GWASCatalog is shown in Supplementary Figure 1.

 doi:10.20944/preprints202603.0818.v1

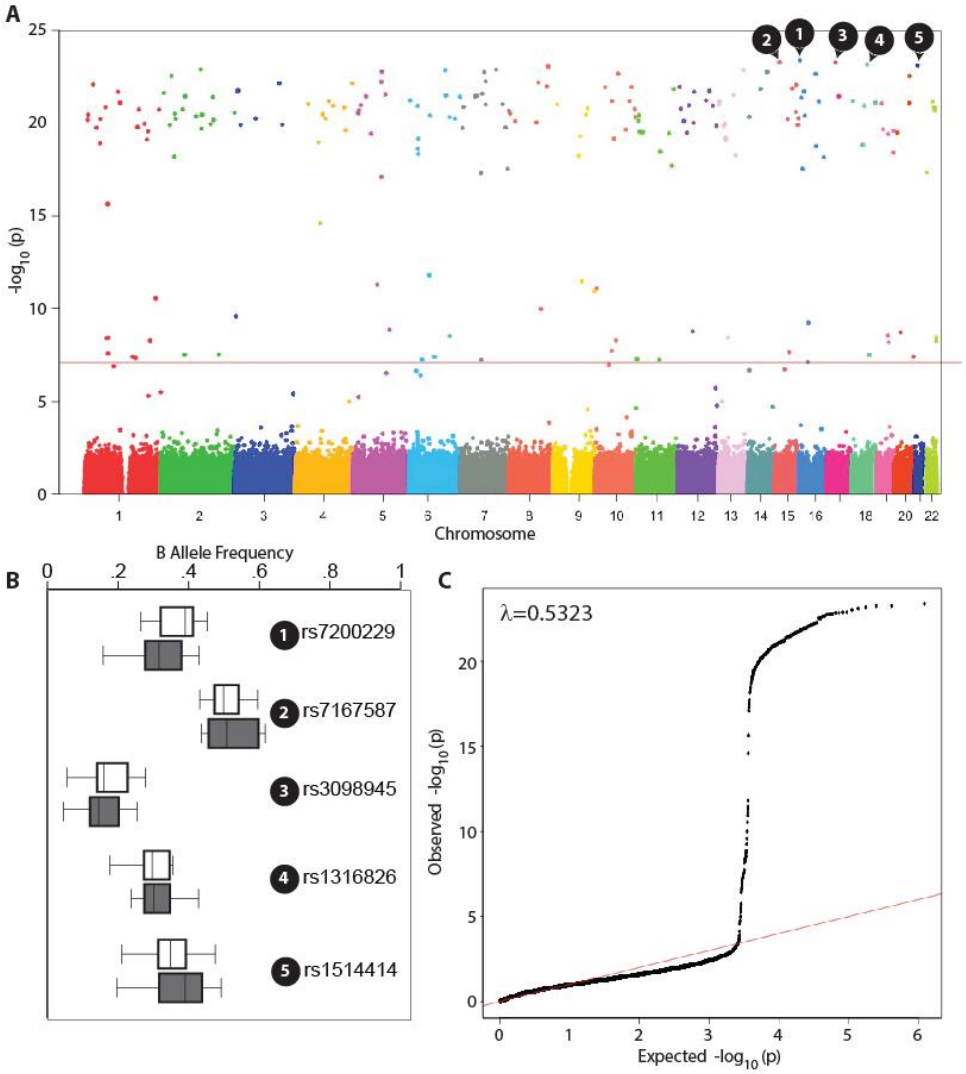

Figure 1. Results of the whole-blood, pooled genome-wide association study. (A) Manhattan plot. The points indicate the log-transformed, adjusted, corrected and statistically significant p-values. The five topmost significant associations are numbered as 1 through 5. (B) Box plots for the distribution of the top five significant SNP markers. Open boxes are for pools without hypertension and filled boxes are for pools with hypertension. The numbers indicated in black circles correspond to those in panel A. (C) QQ plot. The plot shows the relationship between observed and expected p-value distribution. The genomic inflation factor (λ) is shown at the top of the plot.

 doi:10.20944/preprints202603.0818.v1

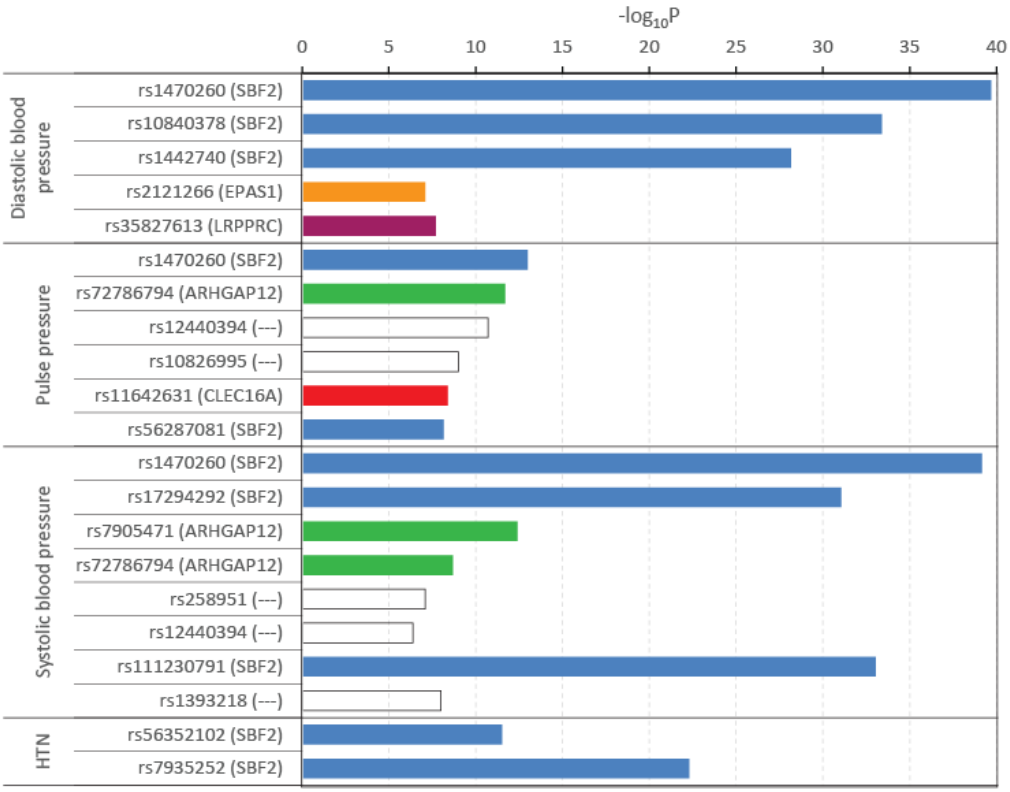

Figure 2. Association of significant variants with blood pressure-related traits from published studies in GWASCatalog. Names of the variants (gene) are shown on the vertical axis while the horizontal bars represent log-transformed P value. Colors of the bars represent different genes – blue, SBF; green, ARHGAP12; orange, EPAS1; brown, LRPPRC; and red, CLEC16A. Hollow bars represent variants from intergenic regions not mapping to any gene.

3.4. Pooled GWAS Results at the Level of Genes

The top 191 significantly associated SNPs mapped to 149 known and named genes (Supplementary Table 2). Of the 191 variants queried, the SNPAnnotator module could map 180 genes that included a total of 107 (59.44%) protein coding genes, 55 (30.56%) long noncoding RNA (lncRNA) genes, 1 small Cajal-body specific RNA gene, 1 small nuclear RNA gene and 16 pseudogenes. Of the 55 lncRNA genes, 37 (67.27%) were intergenic, 8 (14.54%) were intronic, 6 (10.91%) were anti-sense and 4 (7.28%) were sense lncRNAs. Full list of the variants associated with the lncRNAs is provided in Supplementary Table 3. Literature search revealed that 19 (34.54%) of the lncRNAs listed in Supplementary Table 3 have been previously reported to be associated with blood pressure related traits.

When the list of 149 named genes was queried against the Human Phenotype Ontology (HPO) terms, a total of 33 terms were significantly (FDR-corrected p <0.05) associated with the list (Supplementary Table 4). Strikingly, the list contained the following terms: blood pressure (FDRp = 0.0132), systolic blood pressure (FDRp = 0.0194), diabetes mellitus (FDRp = 0.0120), triglyceride measurement (FDRp = 0.0440) and body weight measurement (FDRp = 2.61×10^{-5}). These results affirmed a biological explanation and a strong plausibility of metabolic function to the observed association pattern. The genes associated with the terms blood pressure and systolic blood pressure included: GALNT18, SBF2, VIPR2, TENM4, SHROOM3, DUSP16, ZNF609, DGKH, ACMSD, GRM7, ZNF98, AGBL4, SIK3, CDH18, ALK, ZFPM2, RBFOX1, FTO, FGD4, PAFAH1B2, TRPC4, CSMD1, and LRP2. Comparatively, when gene enrichment analyses were conducted for the gene ontology terms using the g:Profiler tool, we found (Supplementary Figure 2) that there were six terms that were significantly associated with the gene set. These were: ion binding (FDRp = 0.0313), transmembrane

 doi:10.20944/preprints202603.0818.v1

transporter binding (FDRp = 0.0417), anatomical structure development (FDRp = 0.0023), biological regulation (FDRp = 0.0126), axon (FDRp = 0.0002) and juxtaparanode region of axon (FDRp = 0.0165).

In addition, novel associations found in relation to the topmost significant variants revealed some interesting patterns. For example, well-characterized genes related to the top 20 significant variants were associated with the gene FLYWCH1-AS1 (rs7200229), ANKRD13B (rs3098945), RNU6-976P (rs17258345), MAGI2 (rs12665877), DUSP29 (rs755228), COX6CP2 (rs8183309), MTHFD2P5 (rs7457005), and SDC2(rs2008026). Further, some long non-coding genes were also found to be associated with the top 20 significant variants and included ENSG00000294624, LINC01320, and ENSG00000249776. Indeed, results from Figure 2 indicated that variants in the genes SBF2, ARHGAP12, EPAS1, and CLEC16A were strongly associated with blood-pressure related traits in published GWAS studies. There were seven distinct variants in or around SBF2 gene which were associated with one or more pressure-related traits, making it a potential determinant of the risk of hypertension in the study population.

4. Discussion

Our study made the following critical observations. First, we found a striking concurrence of type 2 diabetes, prediabetes and hypertension in the urban Sindhi families. Second, we observed that whole-blood pooling and genotyping was able to identify interesting patterns of genetic variants that revealed known as well as novel genome-wide association. Third, there was a specificity of association of the genetic variants with hypertension such that several genomic hits observed by us have been reported by other genome-wide association studies previously. Fourth, we identified new variants and genes that were associated with hypertension in the study population. For example, the multipronged association of the SBF2 gene variants, the links between FLYWCH1-AS1 and ANKRD13B gene variants and hypertension and the identification of several long noncoding RNA genes as potential makers of hypertension have yielded additional insights into hypertension pathophysiology. Lastly, the data presented here have not been previously described in the context of the urban Sindhi population studies.

The findings of our study need to be considered in the light of an increasing burden of hypertension in India. The prevalence of hypertension in our study was alarmingly high (>50%). This prevalence is not reflective of the general population prevalence because of at least two factors. First, the study participants were ascertained on the presence of at least one known patient of type 2 diabetes in the family. Since hypertension is a risk factor for type 2 diabetes, a higher proportion of study participants (as compared to general population) are expected to have hypertension. In the 2015-16 data from the National Family Health Survey in India, the prevalence of hypertension in diabetic individuals was estimated to be 37%.[20] Second, this is a family study and therefore heritable traits are likely to cluster frequently in the study sample – a situation that can masquerade as high prevalence rate. Our study design precludes the use of available methods for family-based designs to estimate population prevalence of disease. Nevertheless, the estimated prevalence of hypertension in the study participants is indicative of a high prevalence of hypertension in Indian urban Sindhi population. For example, the well-conducted and nationally representative ICMR-INDIAB study estimated the prevalence of hypertension to be 35.5% in India. [21] Similarly, the Indian Society of Hypertension estimated the prevalence of hypertension to be 21% and 24% and the prevalence of pre-hypertension to be 39% and 49% in women and men, respectively.[22] Together, our study findings gain importance in the light of the increasing and high prevalence of hypertension in general and in individuals with metabolic comorbidities in particular. The novelty of our findings is further enhanced by the fact that, to our knowledge, this is the first study documenting a high prevalence of hypertension in the ethnically endogamous group of urban Sindhis in India.

We found variants related to five genes known to be associated with hypertension through well-recognized biological mechanisms. Of these five susceptibility genes, variants in and around the SBF2 gene were most common. The SBF2 gene (also called the MTMR13 gene) encodes a protein involved

 doi:10.20944/preprints202603.0818.v1

in phosphoinositide signaling.[23] This mechanism has been strongly implicated in the development of hypertension (https://maayanlab.cloud/Harmonizome/gene_set/Hypertension/GWAS+Catalog+SNP-Phenotype+Associations+2025, [24]). On the other hand, the ARHGAP12 gene interacts with the With No Lysine (K) pathway (WNK pathway) which is a key regulator of blood pressure.[25] Similarly, endothelial Epas1 (the protein product of the EPAS1 gene) has been implicated in renal damage, resulting in focal segmental glomerulosclerosis that manifests as hypertension.[26] Similarly, the CLEC16A gene is known to be involved in mitochondrial activity regulation that exercises renal control of blood pressure [27] as well as vascular stiffness[28]. Lastly, the LRPPRC gene also partakes in vascular tone control and oxidative stress through mitochondrial pathways[29,30] and can, thus, influence the risk of hypertension indirectly.

We also found some additional interesting genes as susceptibility loci for hypertension in the Sindhi families. Of note, the FLYCH1-AS1 which is a component of the Wnt signaling pathway is a long noncoding RNA. Derangements in Wnt signaling (which FLYWCH1 modulates and WNT2 participates in) cause vascular smooth muscle remodeling that may underlie hypertension and vascular diseases.[31,32] Similarly, the ANKRD13B gene has been found in genome-wide association studies to be linked with both coronary artery disease and blood pressure traits.[33] The MAGI2 gene is primarily involved in synaptic functions and has been associated with to blood pressure regulation in a large-scale study on 564,680 participants from diverse populations.[34] SDC2 which encodes syndecan-2 is functionally known to play a part in maintaining vascular endothelial integrity. Its close family member, syndecan-4, has been implicated in blood pressure regulation.[35]

Two more observations merit a mention. First, we found several lncRNA genes (approximately 31% of the significant genes) to be associated with hypertension. It is noteworthy that a recent transcriptome-wide association study[36] found that 30 transcripts of lncRNA (especially related to the UCP2 gene) were significantly associated with hypertension. Jiang and Ning[37] have summarized the potential of lncRNAs as potential mediators of blood pressure. Currently, the exact mechanism underlying the contribution is unknown and a matter of scientific interest.[38] Our study found several interesting lncRNAs that need further replicative confirmation and functional relevance assessment. Second, on the other extreme, we found that pseudogenes such as RNU-976P, COX6CP2 and MTHFD2P5 were associated with hypertension but the functional role of such associations is unknown.

In addition to the previously described advantages and strengths of this pooled GWAS approach,[16] our study has some limitations. First, the concept of whole-blood pooled GWAS is predicated on the assumption that the allele frequencies are faithfully captured by pooling. In the absence of individual-level GWAS data, the veracity of this assumption cannot be commented upon. Therefore, whether the results of this study will hold if compared to an individual-level GWAS on the same participants cannot be inferred. Second, our study did not have a replication cohort. Due to the ethnically endogamous and genetically related nature of the study population, it is practically very challenging to design another such cohort for validation purposes. Further, to our knowledge, GWAS data on such a cohort is not available, Thus, generalization of the observed associations is not possible. Third, biological explanation for the functional role of the susceptibility genes identified in this study is currently not available and cannot be inferred. Future studies need to specifically investigate the functional role of genes. We therefore compared the associations observed in this study with those reported in well-established repositories and in other published studies.

5. Conclusions

Notwithstanding these limitations, we conclude that the prevalence of hypertension in the endogamous Sindhi families studied here was high (52%) and this trait was highly heritable (h^2r = 0.44). Our pooled, whole-blood GWAS for hypertension in the families of type 2 diabetes (T2D) patients uncovered interesting candidate genes for future investigations. We identified several

 doi:10.20944/preprints202603.0818.v1

significant variants and genes that have been reported by other studies previously. Notably, nearly 31% of the genes identified in this study were related to lncRNAs which are being increasingly recognized as potential biomarkers of hypertension. Whether the genomic hits identified in this study are population-specific, whether there is a functional explanation for the role of these genomic variants in hypertension and whether some of the identified lncRNAs can be considered biomarkers in the study population are important questions for future research.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org, Supplementary Table 1: Annotated list of the significantly associated variants with hypertension in the DISFIN study participants; Supplementary Table 2: Annotation of the 149 named genes related to the 191 significantly associated variants; Supplementary Table 3: Long noncoding RNA genes associated with hypertension in the DISFIN Study; Supplementary Table 4: Human Phenotype Ontology terms related to significantly associated genes in the DISFIN Study; Supplementary Figure 1: Variant-phenotype network using data from GWASCatalog; Supplementary Figure 2: Gene set enrichment analyses using the g:Profiler tool; Supplementary Note 1: R code for conducting GWAS; Supplementary Note 2: R code for adjusting p-values; and Supplementary Note 3: R code for creating plots.

Author Contributions: Conceptualization, M.M. and H.K.; methodology, H.K.; R scripting, H.K. and M.K.; formal analysis, M.K., S.K. and H.K.; clinical and laboratory data collection; K.V.P; A.A.P; and M.T.J; data curation, K.V.P., A.A.P. and H.K.; writing—original draft preparation, S.K. and H.K.; writing—review and editing, all co-authors; supervision, M.M.; project administration, M.M.; funding acquisition, M.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by through the internal funding mechanism of Lata Medical Research Foundation, Nagpur, India. The APC was also funded by Lata Medical Research Foundation, Nagpur, India.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of University of Texas Rio Grande Valley (protocol code 2016-126-07 [940406], dated August 30, 2016) and by the Institutional Ethics Committee of the Lata Medical Research Foundation, Nagpur, India (protocol ID RPC21A, dated 7 May 2016).

Informed Consent Statement: Written, informed consent was obtained from all subjects involved in the DISFIN study.

Data Availability Statement: Lata Medical Research Foundation's Institutional Ethics Committee (LMRF-IEC) does not allow public data sharing to avoid potential identification. If data is requested for verification of results, we will seek permission from the LMRF-IEC before the requested data can be released. For further clarification of the LMRF-IEC's data access policy as well as for data access requests, please contact: Dr. Prabir Kumar Das, Member Secretary, Institutional Ethics Committee, Lata Medical Research Foundation, Kinkine Kutir, Vasant Nagar, Nagpur – 440022, Ph. No. 91-8805023450, Email: prabir_das23@rediffmail.com.

Acknowledgments: The authors are indebted to Dr. Archana Patel, Smita Puppalwar, Shilpa Pawar and Riddhi Tharani for administrative support; to Dr. Shantanu Sengupta, Sengupta Hospital and Research Institute, Nagpur, India for the laboratory support; to Neeti Chhangani for financial management; and to Dattaraj Gondane, Harshad Sonavane and Amar Bagde for logistics during the data collection phase.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. World_Health_Organization. *Global report on hypertension 2025: High stakes - turning evidence into action* World Health Organization: Geneva, Switzerland, 2025.

 doi:10.20944/preprints202603.0818.v1

2. Shih, P.A.; O'Connor, D.T. Hereditary determinants of human hypertension: strategies in the setting of genetic complexity. *Hypertension* **2008**, *51*, 1456-1464, doi:10.1161/HYPERTENSIONAHA.107.090480.
3. Kolifarhood, G.; Daneshpour, M.; Hadaegh, F.; Sabour, S.; Mozafar Saadati, H.; Akbar Haghdoust, A.; Akbarzadeh, M.; Sedaghati-Khayat, B.; Khosravi, N. Heritability of blood pressure traits in diverse populations: a systematic review and meta-analysis. *J Hum Hypertens* **2019**, *33*, 775-785, doi:10.1038/s41371-019-0253-4.
4. Waken, R.J.; de las Fuentes, L.; Rao, D.C. A Review of the Genetics of Hypertension with a Focus on Gene-Environment Interactions. *Curr Hypertens Rep* **2017**, *19*, 23, doi:10.1007/s11906-017-0718-1.
5. Keaton, J.M.; Kamali, Z.; Xie, T.; Vaez, A.; Williams, A.; Goleva, S.B.; Ani, A.; Evangelou, E.; Hellwege, J.N.; Yengo, L.; et al. Genome-wide analysis in over 1 million individuals of European ancestry yields improved polygenic risk scores for blood pressure traits. *Nat Genet* **2024**, *56*, 778-791, doi:10.1038/s41588-024-01714-w.
6. Gupta, R.; Ram, C.V.S. Hypertension epidemiology in India: emerging aspects. *Curr Opin Cardiol* **2019**, *34*, 331-341, doi:10.1097/HCO.0000000000000632.
7. Mohammad, R.; Bansod, D.W. Hypertension in India: a gender-based study of prevalence and associated risk factors. *BMC Public Health* **2024**, *24*, 2681, doi:10.1186/s12889-024-20097-5.
8. Venkatesan, V.; Lopez-Alvarenga, J.C.; Arya, R.; Ramu, D.; Koshy, T.; Ravichandran, U.; Ponnala, A.R.; Sharma, S.K.; Lodha, S.; Sharma, K.K.; et al. Burden of Type 2 Diabetes and Associated Cardiometabolic Traits and Their Heritability Estimates in Endogamous Ethnic Groups of India: Findings From the INDIGENIUS Consortium. *Front Endocrinol (Lausanne)* **2022**, *13*, 847692, doi:10.3389/fendo.2022.847692.
9. Newton-Cheh, C.; Johnson, T.; Gateva, V.; Tobin, M.D.; Bochud, M.; Coin, L.; Najjar, S.S.; Zhao, J.H.; Heath, S.C.; Eyheramendy, S.; et al. Genome-wide association study identifies eight loci associated with blood pressure. *Nat Genet* **2009**, *41*, 666-676, doi:10.1038/ng.361.
10. Mamtani, M.; Jaisinghani, M.T.; Jaiswal, S.G.; Pipal, K.V.; Patel, A.A.; Kulkarni, H. Genetic association of anthropometric traits with type 2 diabetes in ethnically endogamous Sindhi families. *PLoS One* **2021**, *16*, e0257390, doi:10.1371/journal.pone.0257390.
11. Whelton, P.K.; Carey, R.M. The 2017 American College of Cardiology/American Heart Association Clinical Practice Guideline for High Blood Pressure in Adults. *JAMA Cardiol* **2018**, *3*, 352-353, doi:10.1001/jamacardio.2018.0005.
12. Kumar, R.; Nandhini, L.P.; Kamalanathan, S.; Sahoo, J.; Vivekanadan, M. Evidence for current diagnostic criteria of diabetes mellitus. *World J Diabetes* **2016**, *7*, 396-405, doi:10.4239/wjd.v7.i17.396.
13. Snehalatha, C.; Viswanathan, V.; Ramachandran, A. Cutoff values for normal anthropometric variables in asian Indian adults. *Diabetes Care* **2003**, *26*, 1380-1384, doi:10.2337/diacare.26.5.1380.
14. Sharma, Y.K.; Prakash, N.; Gupta, A. Prevalence of metabolic syndrome as per the NCEP and IDF definitions vis-a-vis severity and duration of psoriasis in a semi-urban Maharashtrian population: A case control study. *Diabetes Metab Syndr* **2016**, *10*, S72-76, doi:10.1016/j.dsx.2016.01.033.
15. Craig, J.E.; Hewitt, A.W.; McMellon, A.E.; Henders, A.K.; Ma, L.; Wallace, L.; Sharma, S.; Burdon, K.P.; Visscher, P.M.; Montgomery, G.W.; et al. Rapid inexpensive genome-wide association using pooled whole blood. *Genome Res* **2009**, *19*, 2075-2080, doi:10.1101/gr.094680.109.
16. Pipal, K.V.; Mamtani, M.; Patel, A.A.; Jaiswal, S.G.; Jaisinghani, M.T.; Kulkarni, H. Susceptibility Loci for Type 2 Diabetes in the Ethnically Endogamous Indian Sindhi Population: A Pooled Blood Genome-Wide Association Study. *Genes (Basel)* **2022**, *13*, doi:10.3390/genes13081298.
17. Turner, S. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *The Journal of Open Source Software* **2018**, doi:doi:10.21105/joss.00731.
18. Macgregor, S.; Zhao, Z.Z.; Henders, A.; Nicholas, M.G.; Montgomery, G.W.; Visscher, P.M. Highly cost-efficient genome-wide association studies using DNA pools and dense SNP arrays. *Nucleic Acids Res* **2008**, *36*, e35, doi:10.1093/nar/gkm1060.
19. Ani, A.; Kamali, Z.; Vaez, A. SNPAnnotator: Automated Functionanl Annotation of Genetic Variants and Linked Proxies., 1.4.6; 2026.

 doi:10.20944/preprints202603.0818.v1

20. Santra, A.; Mallick, A. Prevalence of hypertension among individuals with diabetes and its determinants: evidences from the National Family Health Survey 2015-16, India. *Ann Hum Biol* **2022**, *49*, 133-144, doi:10.1080/03014460.2022.2072525.
21. Anjana, R.M.; Unnikrishnan, R.; Deepa, M.; Pradeepa, R.; Tandon, N.; Das, A.K.; Joshi, S.; Bajaj, S.; Jabbar, P.K.; Das, H.K.; et al. Metabolic non-communicable disease health report of India: the ICMR-INDIAB national cross-sectional study (ICMR-INDIAB-17). *Lancet Diabetes Endocrinol* **2023**, *11*, 474-489, doi:10.1016/S2213-8587(23)00119-5.
22. Maheshwari, A.; Gupta, R.; Verma, N.; Narasingan, S.N.; Singh, R.B.; Saboo, B.; Kumar, C.H.V.; Gupta, A.; Srivastava, M.K.; Gupta, A.; et al. Position statement on hypertension by Indian Society of Hypertension, 2023. *J Hum Hypertens* **2024**, *38*, 736-744, doi:10.1038/s41371-024-00960-z.
23. Senderek, J.; Bergmann, C.; Weber, S.; Ketelsen, U.P.; Schorle, H.; Rudnik-Schoneborn, S.; Buttner, R.; Buchheim, E.; Zerres, K. Mutation of the SBF2 gene, encoding a novel member of the myotubularin family, in Charcot-Marie-Tooth neuropathy type 4B2/11p15. *Hum Mol Genet* **2003**, *12*, 349-356, doi:10.1093/hmg/ddg030.
24. Wang, J.; Guo, W.; Wang, Q.; Yang, Y.; Sun, X. Recent advances of myotubularin-related (MTMR) protein family in cardiovascular diseases. *Front Cardiovasc Med* **2024**, *11*, 1364604, doi:10.3389/fcvm.2024.1364604.
25. Dbouk, H.A.; Huang, C.L.; Cobb, M.H. Hypertension: the missing WNKs. *Am J Physiol Renal Physiol* **2016**, *311*, F16-27, doi:10.1152/ajprenal.00358.2015.
26. Luque, Y.; Lenoir, O.; Bonnin, P.; Hardy, L.; Chipont, A.; Placier, S.; Vandermeersch, S.; Xu-Dubois, Y.C.; Robin, B.; Lazareth, H.; et al. Endothelial Epas1 Deficiency Is Sufficient To Promote Parietal Epithelial Cell Activation and FSGS in Experimental Hypertension. *J Am Soc Nephrol* **2017**, *28*, 3563-3578, doi:10.1681/ASN.2016090960.
27. Barry, A.; McNulty, M.T.; Jia, X.; Gupta, Y.; Debiec, H.; Luo, Y.; Nagano, C.; Horinouchi, T.; Jung, S.; Colucci, M.; et al. Multi-population genome-wide association study implicates both immune and non-immune factors in the etiology of pediatric steroid sensitive nephrotic syndrome. *medRxiv* **2022**, 2022.2009.2013.22279644, doi:10.1101/2022.09.13.22279644.
28. Proust, C.; Empana, J.P.; Boutouyrie, P.; Alivon, M.; Challande, P.; Danchin, N.; Escriou, G.; Esslinger, U.; Laurent, S.; Li, Z.; et al. Contribution of Rare and Common Genetic Variants to Plasma Lipid Levels and Carotid Stiffness and Geometry: A Substudy of the Paris Prospective Study 3. *Circ Cardiovasc Genet* **2015**, *8*, 628-636, doi:10.1161/CIRCGENETICS.114.000979.
29. Gao, G.; Chen, A.; Gong, J.; Lin, W.; Wu, W.; Mohammad Ismail Hajary, S.; Lian, G.; Luo, L.; Xie, L. Comprehensive analyses of m6A RNA methylation patterns and related immune microenvironment in idiopathic pulmonary arterial hypertension. *Front Genet* **2023**, *14*, 1222368, doi:10.3389/fgene.2023.1222368.
30. Andreassen, O.A.; McEvoy, L.K.; Thompson, W.K.; Wang, Y.; Reppe, S.; Schork, A.J.; Zuber, V.; Barrett-Connor, E.; Gautvik, K.; Aukrust, P.; et al. Identifying common genetic variants in blood pressure due to polygenic pleiotropy with associated phenotypes. *Hypertension* **2014**, *63*, 819-826, doi:10.1161/HYPERTENSIONAHA.113.02077.
31. Mani, N.; Nygaard, R.; Mancia, F. Gone with the Wnt(less): a mechanistic perspective on the journey of Wnt. *Biochem Soc Trans* **2022**, *50*, 1763-1772, doi:10.1042/BST20220634.
32. Williams, H.; Mill, C.A.; Monk, B.A.; Hulin-Curtis, S.; Johnson, J.L.; George, S.J. Wnt2 and WISP-1/CCN4 Induce Intimal Thickening via Promotion of Smooth Muscle Cell Migration. *Arterioscler Thromb Vasc Biol* **2016**, *36*, 1417-1424, doi:10.1161/ATVBAHA.116.307626.
33. 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, doi:10.1161/CIRCRESAHA.117.312086.
34. Lee, S.; Miller, C.L.; Bentley, A.M.; Brown, M.R.; Nagarajan, P.; Noordam, R.; Morrison, J.; Schwander, K.; Westerman, K.; Kho, M.; et al. A Large-Scale Genome-wide Association Study of Blood Pressure Accounting for Gene-Depressive Symptomatology Interactions in 564,680 Individuals from Diverse Populations. *Research Square (Preprint)* **2025**, doi:https://doi.org/10.21203/rs.3.rs-6025759/v1.
35. Kunnas, T.; Nikkari, S.T. Contribution of syndecan-4 genetic variants to hypertension, the TAMRISK study. *BMC Res Notes* **2014**, *7*, 815, doi:10.1186/1756-0500-7-815.

 doi:10.20944/preprints202603.0818.v1

36. Vargas, J.D.; Abbas, M.; Goodney, G.; Le, H.; Hinton, A.O.; Gaye, A. Regulatory Roles of Long Noncoding RNAs in Arterial Stiffness and Hypertension. *Hypertension* **2025**, *82*, 1195-1207, doi:10.1161/HYPERTENSIONAHA.124.23580.
37. Jiang, X.; Ning, Q. Long noncoding RNAs as novel players in the pathogenesis of hypertension. *Hypertens Res* **2020**, *43*, 597-608, doi:10.1038/s41440-020-0408-2.
38. Gluba-Sagr, A.; Franczyk, B.; Rysz-Gorzynska, A.; Olszewski, R.; Rysz, J. The Role of Selected lncRNAs in Lipid Metabolism and Cardiovascular Disease Risk. *Int J Mol Sci* **2024**, *25*, doi:10.3390/ijms25179244.