Int J Hum Genet, 17(3): 140-144 (2017)
DOI:10.31901/24566330.2017/17.03.07

Brief Communication

Genomics of Type 2 Diabetes Mellitus and Glycemic Traits

Vipin Gupta[1] and Gagandeep Kaur Walia[2]

[1]Department of Anthropology, University of Delhi, Delhi 110 007, India
[2]Public Health Foundation of India, Gurgaon, Haryana, India

KEYWORDS Genome-wide Association Study. Type 2 Diabetes. India. Glycemic Traits

ABSTRACT The increased burden of type 2 diabetes mellitus (T2DM) has accelerated the number of genetic studies in different human populations for detecting at least predictive biomarkers related to T2DM. Genome-wide association designs have successfully identified number of potential genetic variants associated with T2DM and explained the polygenic nature of T2DM inheritance using large sample sizes. Genetic epidemiologists have also explored the quantitative glycemic traits simultaneously with T2DM and provided comprehensive picture of overlapping findings through biological pathways. Indian studies on genetics of T2DM are generally limited to the candidate gene based association designs and efforts were primarily made to validate the European findings. This manuscript is a broad comparative account of efforts in genetic evaluation of T2DM and its related measures in Western and Indian populations.

INTRODUCTION

Type 2 diabetes mellitus (T2DM) is a multifactorial metabolic disease characterised by impaired insulin secretion followed by insulin resistance (Ramachandran et al. 2010). There are 415 million diabetic people worldwide reflecting the global rise in its prevalence of which 69.2 million are Indians and ranked 2nd in terms of number of people with diabetes (International Diabetes Federation 2015). The risk factors, like hypertension and obesity, due to changing dietary habits and preference for physical inactivity are major known environmental factors of T2DM. The high heritability of T2DM (0.69, Almgren et al. 2011) and glycemic traits (fasting glucose: 24-72%, (Mathias et al. 2009) has generated hope that a genotype might help to target those at greatest risk (Vassy et al. 2014). The underlying genomics of T2DM has been explored through two main approaches, like, linkage studies and association studies (that is, candidate gene and genome-wide association studies (GWASs)). The researchers' objective was to briefly review the current research developments on genetics of T2DM and related traits in Western population in comparison to India.

The earlier attempts of genetic dissection of T2DM with replication and validation studies were limited to *PPARG6* (Altshuler et al. 2000), *KCNJ11* (Gloyn et al. 2003) and *TCF7L2* (Grant et al. 2006). GWASs were driven by the "Common–Disease Common Variant" (CDCV) hypothesis (Reich and Lander 2001) and have substantially changed the landscape of genetic research in T2DM. The CDCV hypothesis lead to the agnostic survey of whole genome for detecting genetic variants associated with the trait under investigation known as genome-wide association studies. While conducting GWAS, researchers have adopted two-fold strategy for studying T2DM: first by exploring the genomics of T2DM as a disease, thus, expanded the list of common markers (Voight et al. 2010) and second by extending the GWAS approach in exploring the genomics of its diagnostic quantitative traits (Dupuis et al. 2010). Since 2007, series of GWASs have remarkably changed the known genomic profile of T2DM in different human populations. These large scale genomic investigations have largely covered western populations, and have confirmed genome-wide significant associations of 56 loci for type 2 diabetes (Morris et al. 2012) and 53 loci for glycemic traits (Scott et al. 2012), of which 33 have overlapping effects (Scott et al. 2012). These discoveries have considerably

Address for correspondence:
Vipin Gupta
Assistant Professor,
Department of Anthropology,
University of Delhi, Delhi, India
E-mail: udaiig@gmail.com
Mobile: 9899119242

increased the understanding of pathophysiology of type 2 diabetes and its related glycemic traits like levels of fasting glucose, insulin and Homeostatic Model Assessment for insulin secretion and resistance (HOMA-β and HOMA-IR). These known genetic polymorphisms discovered to date have explained about 10 percent of genetic variation underlying T2DM (Voight et al. 2010).

The genetic variants associated with T2DM and glycemic traits identified by independent genome-wide association studies have explained little of the variability. Apparent shortfall in "heritability explained" has prompted speculation about possible explanations (Maher 2008). The primary reasons were small effect sizes of common alleles, and requirement of large sample sizes for detecting relatively rare variants (Chapman et al. 2011; Evangelou and Ioannidis 2013). Therefore, after the first wave of GWASs, investigators started meta-analyses of available genome-wide data for detecting additional genetic variants associated with T2DM and glycemic traits (Zeggini et al. 2008; Dupuis et al. 2010; Cho et al. 2012; Scott et al. 2017). The merits of using this approach of meta-analyses of GWASs were to increase the sample size to several folds, raised the study power for detecting relatively rare genetic variants and reducing false-positive results (Evangelou and Ioannidis 2013) have successfully helped genetic epidemiologists involved in T2DM.

WESTERN SCENARIO

In European populations, GWASs have reached their limits for detecting common genetic variants (Vassy et al. 2014) related to T2DM and glycemic traits. GWASs have explained 4.8 percent and 1.2 percent of variation in fasting glucose and fasting insulin, respectively (Mahajan et al. 2015). Recently, a meta-analysis of genome-wide association studies has identified 13 additional variants associated with T2DM which were driven by common single nucleotide variants (Scott et al. 2017). Scientific argument was given based on simulation studies (Pritchard 2001) that the rare variants (minor allele frequency <5%) may have larger effects than common variants and greater potential to explain disease susceptibility. Therefore, investigators have started putting their energies in whole genome or exome based sequencing technologies in order to extend the explained genetic variance of T2DM and glycemic traits. A recent large attempt has tested the association of low frequency genetic variants using whole genome sequencing data on >2500 Europeans, exome sequencing of >12,000 individuals from different ancestries and genotyping on 111,548 subjects (Fuchsberger et al. 2016). Their findings were consistent with predictions under the 'neutral' common variant model, that is, common polygenic model where low frequency or rare variants explains only ~25% of T2DM heritability. Majority of genetic variants identified by them were common and located on previously detected GWAS loci. Moreover, they have also ruled out the possibility of synthetic association which was suggested by Goldstein (2011). Further, an exome-wide association study on non-diabetic European individuals was conducted in order to find the low frequency functional variants associated glycemic traits (Mahajan et al. 2015). They had found evidence of association of *GLP1R* (p.Ala316Thr) with fasting glucose levels and *URB2* (p.Glu594Val) with fasting insulin levels (Mahajan et al. 2015). In Finnish population, Manning et al. (2017) has explored the low frequency variants associated with glycemic traits and T2DM using exome sequencing techniques and identified a novel rare variant of *AKT2* (p.Pro50Thr). The locus discovered is fully penetrant and causal for monogenic glycemic disorders and can lead to 12 percent rise in fasting insulin levels in terms of effect size. There were consistent efforts in detecting universal biomarkers for T2DM to achieve easy translation into global public health therapeutics. Waters et al. (2010) conducted the first large attempt of trans-ethnic meta-analysis and reported 18 novel loci associated with T2DM. A trans-ethnic analysis has identified 23 shared genetic loci between the samples of European ancestry and African ancestry (Liu et al. 2016). They have also reported two new genetic variants, that is, *FAM133A* (rs213676) and *PELO* (rs6450057) in the replication sample of African ancestry (Liu et al. 2016). Interestingly, a trans-ancestral fine mapping of well- established four loci (*CDKAL1, CDKN2A-B, IGF2BP2* and *KCNQ1*) of T2DM were undertaken to localize the potential causal variants and showed that distinct association signals are shared across ancestries (that is, East Asian, European, South Asian, African American and Mexican American descent) (Horikoshi et al.

2016). In addition to above, researchers are also trying to disentangle the biological pathways underlying T2DM based on identified genetic polymorphisms, for example, a recent study evaluated four broad pathways, that is, 1. insulin sensitivity, 2. reduced insulin secretion and fasting hyperglycemia, 3. insulin processing, 4. insulin processing and secretion without a detectable change in fasting glucose levels (Dimas et al. 2014). The number of loci identified by GWASs related to impaired beta cell function is larger than loci associated with insulin resistance (Nolan et al. 2011). The detailed investigation of insulin secretion and sensitivity was conducted using >5000 individuals and discovered association of 30 genetic variants related to either T2DM or glycemic traits with first-phase insulin secretion (Wood et al. 2017). They have also found strong association of *PDX1* with quantitative first-phase insulin secretion that has no effect on T2DM, whereas, *TCF7L2* was associated with peak insulin response which suggests its role in hepatic insulin clearance or insulin processing (Wood et al. 2017). The high success of GWASs have only explained moderate amount of genetic variation (10%) of T2DM heritability (Manolio et al. 2009).

INDIAN SCENARIO

In India, majority of studies explored genetics of T2DM using candidate gene approach (Nair et al. 2010; Sugunam et al. 2010) or validating findings of GWASs conducted in European populations (Chandak et al. 2007; Sanghera et al. 2008; Gupta et al. 2010; Uma et al. 2013). Comparatively, the studies on quantitative traits of type 2 diabetes in India are few (Sanghera et al. 2009; Chauhan et al. 2010) and have investigated a limited number of SNPs. The researchers used population based study design of '*Indian Migration Study*' (Ebrahim et al. 2010) and had conducted India's first largest population based study on quantitative traits related to T2DM by genotyping over 6000 samples (Gupta et al. 2012). The researchers have identified associations of *CXCR4* (rs932206), *CDKAL1* (rs7756992) and *TCF7L2* (rs7903146, rs12255372) with fasting glucose; *NOTCH2* (rs10923931), *TCF-2* (rs757210), *ADAM30* (rs2641348) and *CDKN2A/B* (rs10811661) with fasting insulin; *TCF-2, ADAM30 and CDKN2A/B* with homeostatic model assessment-insulin resistance (HOMA-IR); and *ADAM30 and CDKN2A/B* with (homeostatic model assessment- β cell) (HOMA-β) (Gupta et al. 2012). Unfortunately, India's first genome-wide association study on T2DM was published quite late in 2013 by a group from Institute of Genomics and Integrative Biology (IGIB), Delhi, (Tabassum et al. 2013). They had used hospital based case-control design, performed genome-wide scan of 2465 individuals and only able to identify one new locus (that is, *TMEM163*) which might be due to falsely considering Indo-European linguistic division as single ethnic group. Institute of Genomics and Integrative Biology (IGIB) also conducted Indian genome variation database (2008) in which they concluded that the reality of Indian population structure is based on endogamy, which anthropologists were suggesting from several decades, but there seems to be no learning by the group working on T2DM in the same institute. Another study form south Indian population of relatively small sample size (578 cases and 578 controls) has validated genetic variants of *TCF7L2* and *SLC30A8* in association with T2DM (Phani et al. 2016). Similarly, Chidambaram et al. (2016) had validated single nucleotide polymorphisms in/nearby *TCF7L2* and *CDKN2A/2B* in south Indian population. Majority of Indian studies were plagued by small sample size, restricted to case-control study design and confounded by population stratification due to large ethnic diversity.

Genetic epidemiologists working in India sooner or later have to respect the Indian population diversity and ecological complexity while designing their genetic association studies. For instance, the researchers had conducted a small genetic study on T2DM using case-control design and collecting samples from single endogamous population and validated the role of *TCF7L2* in predisposing T2DM in this population with relatively large effect size (odds ratio of 2.0) in comparison to European populations. Interestingly, India's 2nd GWAS on T2DM also came in 2013, was conducted on Punjabi Sikhs (an ethnically valid sample). They had discovered another novel loci *SGCG* (rs9552911) and confirmed the role of previously identified loci: *TCF7L2* and *IGF2BP2* (Saxena et al. 2013). This study had highlighted the possibility of shared and unique genetic landscape of T2DM in India (Holliday 2013). A survey of confirmed genetic associations related to T2DM has shown the

highest level of directional genetic differentiation among the studied twelve common diseases (Chen et al. 2012) suggesting possible "*population-specific evolutionary adaptation to agricultural developments, dietary patterns, or food availability*" (Holliday 2013). This further strengthens the argument for appreciating the geography and endogamy for conducting genetics based association studies in India.

CONCLUSION

The era of GWASs has successfully generated vast amount of information on genetic architecture and pathophysiology of T2DM and their related glycemic traits. These discoveries demand validation in different human populations using appropriate study designs and optimum sample sizes. For culturally diverse country like India, we need India specific GWAS for detecting genetic variants related to T2DM and its measures.

RECOMMENDATIONS

There is a need to conduct population based genetic studies related to T2DM and glycemic traits with sample sizes representing the burden of the disease in India. In order to identify low frequency causal functional variants, large scale sequencing studies are required with careful study design fully representing ethnic diversity of India. It will be helpful in creating the T2DM knowledgebase that can be further used for developing and testing of super-specialized micro-level hypothesis.

REFERENCES

Almgren P, Lehtovirta M, Isomaa B et al. 2011. Heritability and familiality of type 2 diabetes and related quantitative traits in the Botnia Study. *Diabetologia,* 54: 2811-2819.

Altshuler D, Hirschhorn JN, Klannemark M et al. 2000. The common PPARã Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nature Genetics*, 26: 76–80.

Chandak GR, Janipalli CS, Bhaskar S et al. 2007. Common variants in the TCF7L2 gene are strongly associated with T2D mellitus in the Indian population. *Diabetologia,* 50: 63–67.

Chapman K, Ferreira T, Morris A et al. 2011. Defining the power limits of genome-wide association scan meta-analyses. *Genet Epidemiol,* 35: 781-789.

Chauhan G, Spurgeon CJ, Tabassum R et al. 2010. Impact of common variants of PPARG, KCNJ11, TCF7L2, SLC30A8, HHEX, CDKN2A, IGF2BP2, and CDKAL1 on the risk of type 2 diabetes in 5,164 Indians. *Diabetes,* 59: 2068-2074.

Chen R, Corona E, Sikora M et al. 2012. Type 2 diabetes risk alleles demonstrate extreme directional differentiation among human populations, compared to other diseases. *PLoS Genet,* 8: e1002621.

Chidambaram M, Liju S, Saboo B et al. 2016. Replication of genome-wide association signals in Asian Indians with early-onset type 2 diabetes. *Acta Diabetol,* 53(6): 915-923.

Cho YS, Chen CH, Hu C et al. 2012. Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in East Asians. *Nat Genet,* 44: 67-72.

Dimas AS, Lagou V, Barker A et al. 2014. Impact of type 2 diabetes susceptibility variants on quantitative glycemic traits reveals mechanistic heterogeneity. *Diabetes,* 63: 2158–2171.

Dupuis J, Langenberg C, Prokopenko I et al. 2010. New genetic loci implicated in fasting glucose homeostasis and their impact on T2D risk. *Nature Genetics,* 42: 105-116.

Ebrahim S, Kinra S, Bowen L et al. 2010. The effect of rural-tourban migration on obesity and diabetes in India: A cross-sectional study. *PLoS Med,* 7: e1000268.

Evangelou E, Ioannidis JPA 2013. Meta-analysis methods for genome-wide association studies and beyond. *Nature Reviews Genetics*, 14: 379–389.

Fuchsberger C, Flannick J, Teslovich TM et al. 2016. The genetic architecture of type 2 diabetes. *Nature,* 536: 41-47.

Gloyn AL, Weedon MN, Owen KR et al. 2003. Large-scale association studies of variants in genes encoding the pancreatic â-cell KATP channel subunits Kir6.2 (KCNJ11) and SUR1 (ABCC8) confirm that the KCNJ11 E23K variant is associated with type 2 diabetes. *Diabetes,* 52: 568–572.

Goldstein DB 2011. The importance of synthetic associations will only be resolved empirically. *PLoS Biol*. 9: e1001008.

Grant SF, Thorleifsson G, Reynisdottir I et al. 2006. Variant of transcription factor 7 like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nature Genetics,* 38: 320–323.

Gupta V, Khadgawat R, Ng HK et al. 2010. A validation study of type 2 diabetes-related variants of the TCF7L2, HHEX, KCNJ11, and ADIPOQ genes in one endogamous ethnic group of north India. *Ann Hum Genet,* 74: 361–368.

Gupta V, Vinay DG, Rafiq S et al. 2012. Association analysis of 31 common polymorphisms with type 2 diabetes and its related traits in Indian sib pairs. *Diabetologia,* 55: 349-357.

Holliday EG 2013. Hints of unique genetic effects for type 2 diabetes in India. *Diabetes,* 62: 1369-1370.

Horikoshi M, Pasquali L, Wiltshire S et al. 2016. Transancestral fine-mapping of four type 2 diabetes susceptibility l oci highlights potential causalregulatory mechanisms. *Hum Mol Genet,* 25(10): 2070-2081.

Indian Genome Variation Consortium 2008. Genetic landscape of the people of India: A canvas for disease gene exploration. *J Genet,* 87 (1): 3-20.

International Diabetes Federation 2015. IDF Diabetes Atlas. *International Diabetes Federation*. 7th Edition. Brussels, Belgium: International Diabetes Federation.

Liu CT, Raghavan S, Maruthur N et al. 2016. Trans-ethnic meta-analysis and functional annotation illuminates the genetic architecture of fasting glucose and insulin. *Am J Hum Genet,* 99(1): 56-75.

Mahajan A, Sim X, Ng HJ et al. 2015. Identification and functional characterization of G6PC2 coding variants influencing glycemic traitsdefine an effector transcript at the G6PC2-ABCB11 locus. *PLoS Genet,* 11(1): e1004876.

Maher B 2008. Personal genomes: The case of the missing heritability. *Nature,* 456: 18-21.

Manning A, Highland HM, Gasser J et al. 2017. A low-frequency inactivating AKT2 variant enriched in the Finnish population is associated with fasting insulin levels and type 2 diabetes risk. *Diabetes,* 66(7): 2019-2032.

Manolio TA, Collins FS, Cox NJ et al. 2009. Finding the missing heritability of complex diseases. *Nature,* 461(7265): 747-753.

Mathias RA, Deepabn M, Deepab R et al. 2009. Heritability of quantitative traits associated with type 2 diabetes mellitus in large multiplex families from South India. *Metabolism,* 58: 1439-1445.

Morris AP, Voight BF, Teslovich TM et al. 2012. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet,* 44: 981-990.

Nair AK, Sugunan D, Kumar H et al. 2010. Case–control analysis of SNPs in GLUT4, RBP4 and STRA6: association of SNPs in STRA6 with type 2 diabetes in a South Indian population. *PLoS One,* 5: e11444.

Nolan CJ, Damm P, Prentki M 2011. Type 2 diabetes across generations: From pathophysiology to prevention and management. *Lancet,* 378: 169-181.

Phani NM, Adhikari P, Nagri SK et al. 2016. Replication and relevance of mult iple susceptibility loci discovered f rom genome wideassociation studies for type 2 diabetes in an Indian population. *PLoS One,* 11(6): e0157364.

Pritchard JK 2001. Are rare variants responsible for susceptibility to complex diseases? *Am J Hum Genet,* 69: 124–137.

Ramachandran A, Das AK, Joshi SR et al. 2010. Current status of diabetes in India and need for novel therapeutic agents. *JAPI,* 58: 7-9.

Reich DE, Lander ES 2001. On the allelic spectrum of human disease. *Trends Genet,* 17: 502–510

Sanghera DK, Been L, Ortega L et al. 2009. Testing the association of novel meta-analysis-derived diabetes risk genes with type II diabetes and related metabolic traits in Asian Indian Sikhs. *J Hum Genet,* 54: 162–168.

Sanghera DK, Ortega L, Han S et al. 2008. Impact of nine common T2D risk polymorphisms in Asian Indian Sikhs: PPARG2 (Pro12Ala), IGF2BP2, TCF7L2 and FTO variants confer a significant risk. *BMC Med Genet,* 3(9): 59.

Saxena Richa, Saleheen Danish et al. 2013. Genome-wide association study identifies a novel locus contributing to type 2 diabetes susceptibility in Sikhs of Punjabi origin from India. *Diabetes,* 62(5): 1746–1755.

Scott RA, Lagou V, Welch RP et al. 2012. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nat Genet,* 44: 991-1005.

Scott RA, Scott LJ, Mägi R et al. 2017. An expanded genome-wide association study of type 2 diabetes in Europeans. *Diabetes*, pii: db161253.

Sugunan D, Nair AK, Kumar H et al. 2010. A case–control analysis of common variants in GIP with type 2 diabetes and related biochemical parameters in a South Indian population. *BMC Med Genet,* 11: 118. doi: 10.1186/1471-2350-11-118.

Tabassum R, Chauhan G, Dwivedi OP et al. 2013. Genome-wide association study for type 2 diabetes in Indians identifies a new susceptibility locus at 2q21. *Diabetes,* 62: 977–986.

Uma Jyothi K, Jayaraj M, Subburaj KS et al. 2013. Association of TCF7L2 gene polymorphisms with T2DM in the population of Hyderabad, India. *PLoS One,* 8(4): e60212.

Vassy JL, Hivert MF, Porneala B et al. 2014. Polygenic type 2 diabetes prediction at the limit of common variant detection. *Diabetes,* 63: 2172-2182.

Voight BF, Scott LJ, Steinthorsdottir V et al. 2010. Twelve T2D susceptibility loci identified through large-scale association analysis. *Nature Genetics,* 42: 579-589.

Waters KM, Stram DO, Hassanein MT et al. 2010. Consistent association of type 2 diabetes risk variants found in Europeans in diverse racial and ethnic groups. *PLoS Genet,* 6(8). pii: e1001078.

Wood AR, Jonsson A, Jackson AU et al. 2017. A genome-wide association study of IVGTT-based measures of first-phase insulin secretion refines the underlying physiology of type 2 diabetes variants. *Diabetes,* 66(8): 2296-2309.

Zeggini E, Scott LJ, Saxena R et al. 2008. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nature Genetics,* 40: 638–645.

Paper received for publication on April 2017
Paper accepted for publication on August 2017