



Review

Overview of genomics and post-genomics research on type 2 diabetes mellitus: Future perspectives and a framework for further studies

BATTINI MOHAN REDDY*, RAYABARAPU PRANAVCHAND and S A A LATHEEF

Department of Genetics, Osmania University, Hyderabad, India

**Corresponding author (Email, bmrasi@gmail.com)*

MS received 4 January 2018; accepted 6 October 2018; published online 7 February 2019

In this review, we briefly outlined salient features of pathophysiology and results of the genetic association studies hitherto conducted on type 2 diabetes. Primarily focusing on the current status of genomic research, we briefly discussed the limited progress made during the post-genomic era and tried to identify the limitations of the post-genomic research strategies. We suggested reanalysis of the existing genomic data through advanced statistical and computational methods and recommended integrated genomics-metabolomics approaches for future studies to facilitate understanding of the gene-environment interactions in the manifestation of the disease. We also propose a framework for research that may be apt for determining the effects of urbanization and changing lifestyles in the manifestation of complex genetic disorders like type 2 diabetes in the Indian populations and offset the confounding effects of both genetic and environmental factors in the natural way.

Keywords. Complex genetic disorders; genomics; Indian scenario; metabolomics; proteomics; transcriptomics; type 2 diabetes mellitus

1. Introduction

Diabetes is one of the most prevalent complex disorders with type 2 diabetes accounting for more than 90% of all diabetic cases. Hyperglycemia is the characteristic feature of this syndrome, which results from defective insulin secretion or action. The disease itself may not lead to death of the affected individual but being the major risk factor of macrovascular complications like coronary artery disease, cerebrovascular events and peripheral vascular disease, diabetes is an indirect cause of deaths due to such diseases. It is also responsible for disabilities such as diabetic nephropathy, diabetic neuropathy, diabetic retinopathy, skin complications, eye complications as well as mental illness. The International Diabetes Federation (IDF) 2015 reported an estimate of 415 million adults (20–79 years of age) worldwide to have diabetes in the year 2015, which is projected to reach 642 million by the year 2040. Diabetes has been a major public health concern in the 21st century (IDF 2015) among the worldwide countries/territories, particularly in China, India and USA, which show the alarmingly increasing prevalence (figure 1). India, in particular, is expected to have doubled its prevalence by 2040.

2. The pathophysiological processes leading to type 2 diabetes

Glucose, a monosaccharide, is the key carbohydrate of energy metabolism. The three major sources of circulating glucose in the human body are intestinal absorption, gluconeogenesis and glycogenolysis. Blood glucose homeostasis is regulated by gluco-regulatory hormones such as insulin, glucagon, amylin, glucagon-like peptide 1, glucose-dependent insulinotropic peptide, epinephrine, cortisol and growth hormone (Stephen *et al.* 2004). Insulin is the key regulatory hormone of blood glucose homeostasis with its excitatory action of stimulating glucose uptake and inhibitory actions on gluconeogenesis, glycogenolysis, proteolysis, lipolysis and ketogenesis (Sonksen and Sonksen 2000). Ever since the role of insulin in glucose homeostasis is understood, it has been the primary therapeutic target in type 2 diabetes patients (Tibaldi 2013). The major pathological mechanisms of type 2 diabetes are the defective insulin secretion due to dysfunctional pancreatic β -cells and impaired insulin action through insulin resistance (Lin and Sun 2010; Ashcroft and Rorsman 2012).

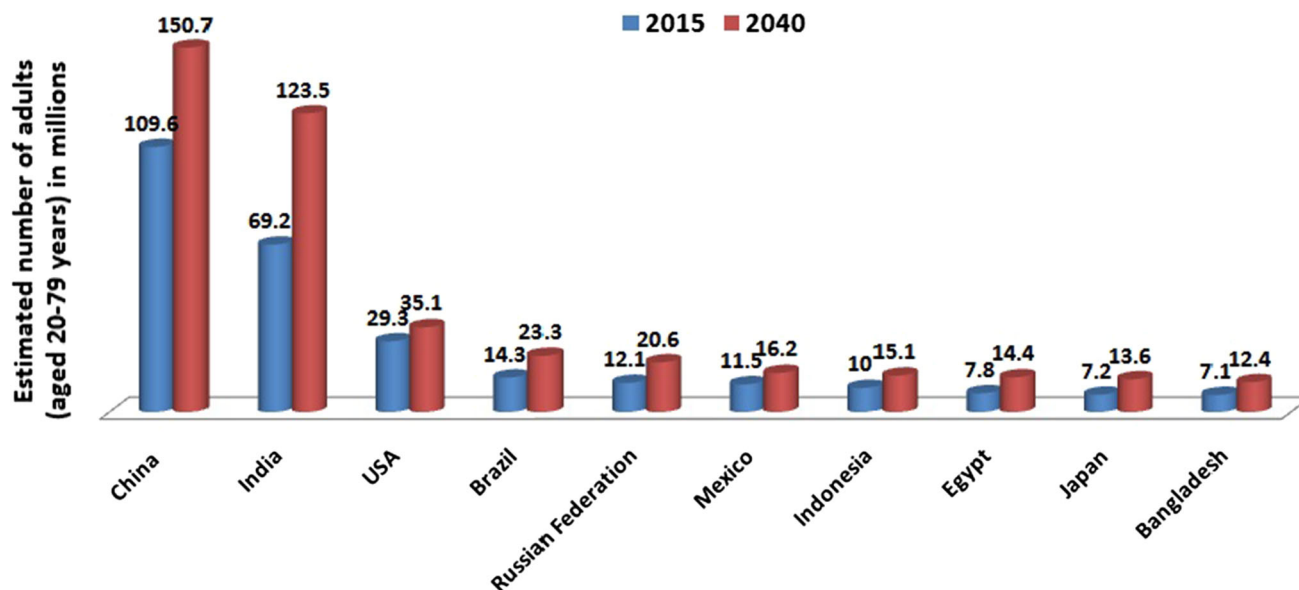


Figure 1. Top ten countries for estimated number (millions) of adults (aged 20–79 years) with type 2 diabetes in the years 2015 and 2040 (Data taken from IDF Diabetes Atlas, 7th Edition, 2015).

3. Genetic susceptibility

The complex genetic disorders in general are hypothesized to result from the action of a large number of genes, each with a small effect, besides the epistatic interactions among them as well as the interaction of each one of them with the environmental factors. The genetic susceptibility towards type 2 diabetes is well established from twin and family based studies. Mutations in the genes that are involved in the regulation of plasma glucose levels and synthesis/action of gluco-regulatory hormones can increase the risk for type 2 diabetes, implying its polygenic nature. However, in an overwhelmingly large proportion of type 2 diabetes cases, the disease mechanisms are triggered through an interaction between genetic and environmental factors, as reflected through the association of processed and carbohydrate rich diet, physical inactivity, stress, smoking and alcohol consumption with the disease (Kommoju and Reddy 2011; Olokoba *et al.* 2012). The variation in genetic architecture of individuals/populations and the intensity of environmental risk factors might have contributed to the variation in geographical prevalence and ethnic susceptibility of the disease. Pleiotropy, a phenomenon of multiple effects of the same genetic loci, has been the characteristic feature of prominent susceptible genes of complex diseases. For example, the *TCF7L2*, which is so far considered to be the most prominent gene associated with type 2 diabetes is found to be involved in multiple metabolic pathways and associated with other diseases like obesity and abnormal lipid traits (Delgado-Lista *et al.* 2011). The observed pleiotropic disease causing effects of these genes might be by virtue of their location in the regulatory genomic domains. All these features make type 2 diabetes a challenging complex genetic

disorder to study. In this review we shall briefly outline the status of understanding on the genetic etiology of type 2 diabetes, with focus on the post-genomic strategies and briefly deal with the relative progress in the transcriptomics, proteomics and metabolomics research as well. We also outline the status of research on the Indian populations and propose an appropriate framework that may be best suited for the Indian population scenario and help in proper understanding of the role of gene–environment interactions in the manifestation of the disease, which may help in devising preventive measures.

4. Broad overview of research on type 2 diabetes

4.1 Genomics

During the past few decades, candidate gene approach with case-control study design has been most successful in understanding the genetic etiology of any complex disease. The method begins with selection of putative candidate gene based on its functional role in disease related metabolic pathway, followed by prioritizing single nucleotide polymorphisms (SNPs) that have functional consequences either by affecting the gene regulation or its product. Finally, the prioritized SNPs/variants are genotyped in a random sample of cases and controls and tested for their association with the trait. So far, a total of 1874 unique markers that belong to 421 genes were identified as associated with type 2 diabetes through this approach (Lim *et al.* 2010). However, an overwhelming inconsistency is observed in the patterns of their association with the disease, with exception to the polymorphisms that belong to *TCF7L2*, *CAPN10*, *PPARG*,

KCNJ11, *ABCC8*, *HNF1A*, *HNF4A*, *GCK*, *PC-1/ENPPI*, *IRS*, *PTPNI*, and *LMNA* genes which showed much greater degree of consistency (Kommoju and Reddy 2011; Ali 2013). Not being satisfied with this approach, researchers shifted the focus to genome wide association studies (GWAS), which is an agnostic method of testing for association of all the SNPs identified in human genome project with a particular disease through chip based microarray technologies such as Illumina and Affymetrics. A large number of cases and controls are screened through this method and the SNPs with strong signal/high significance ($p \leq 10^{-08}$) are considered to be disease susceptible/causing. Only these SNPs are further evaluated for their functional consequences. Through this approach, numerous polymorphisms have been identified as associated with type 2 diabetes and the SNPs of *TCF7L2*, *HHEX*, *CDKN2A/2B*, *IGF2BP2*, *SLC30A8*, *CDKAL1*, *HMG2*, *KCNQ11*, and *NOTCHADAM30* genes being the most replicated ones (www.genome.gov/gwastudies). The search results for type 2 diabetes associated genetic variants yielded 388 significant SNPs from 58 GWAS studies. However, many of these type 2 diabetes associated variants need to be functionally validated before attempting to understand their prospective clinical benefits. The *TCF7L2* is the only gene which is hitherto functionally characterized as key transcription factor coding gene and involved in regulating the glucose homeostasis (Savic *et al.* 2011; Boj *et al.* 2012). As a key component of WNT signaling pathway, it is involved in pancreatic β -cell proliferation and in turn insulin secretion and action (Gupta *et al.* 2008). It was initially identified as associated with the disease through a genetic linkage study on the Icelandic population (Grant *et al.* 2006) and subsequently replicated in Danish (Grant *et al.* 2006), European (Scott *et al.* 2006) and US cohorts (Zhang *et al.* 2006) and currently known to be associated across the ethnic groups worldwide (Kommoju and Reddy 2011). Additionally, a 4kb haplotype block at 9p21.3 chromosomal region was found specific to and associated with type 2 diabetes (Silander *et al.* 2009). Harboring *CDKN2A/CDKN2B* genes with functional implications in cell proliferation pathway, this chromosomal region was observed to be associated with multiple complex diseases and needs detailed exploration for its potential as a therapeutic target in general and particularly with type 2 diabetes. However, the variants identified by GWAS were found to explain only 10% of variation in type 2 diabetes and most of those (more than 90%) are located in the non-coding region (Grarup *et al.* 2014; Scott *et al.* 2016). The search for rare variants with larger penetrance and functional significance is on through next generation and exome sequencing strategies (Jenkinson *et al.* 2016).

4.2 Transcriptomics

The complete set of products in particular tissue or population of cells transcribed from genome may be called

transcriptomics. The advantages of transcriptomics include bridging the gap between genotype and phenotype and detecting genes of phenotypes due to proximity and measured effect of environment on transcription (Jenkinson *et al.* 2016). Transcriptomics research has shown that *IGF2B2* (insulin like growth factor 2) is associated with glucose and insulin homeostasis (Chen *et al.* 2016), decreased *ADH1B* expression with increased obesity, insulin resistance of the whole body, liver, skeletal muscle and adipose tissue and reducing β cell function (Winnier *et al.* 2015), upregulation of *GPD2* and down regulation of *FXYD2* in the state of lower β cell mass (Segerstolpe *et al.* 2016), resemblance of transcriptomic profile of α and β cells of pancreas of type 2 diabetics with that of children (Wang *et al.* 2016) and reduced expression of *BCAT2* and *BCKDHB* and increased methylation in *BCKDH* (Hernández-Alvarez *et al.* 2017), altered expression of genes involved in pathway of adrenergic signaling in cardiomyocytes in association with insulin resistance (Matone *et al.* 2017) and upregulation of *ERAF*, *ALAS2*, *OSBP2*, *CA1*, *STYK1*, and *ZIC2a* and down regulation of *GOS2*, *TEP1*, *PTGS2*, *IL4*, *IL8*, *IFI27*, *IFI3*, *IFIT2*, *NFAIP6*, *RSAD2*, *APOBEC3A*, *ABGL4*, *ABCA1*, *EPST11*, *EPHA6* and *LRRN3* genes in white blood cells of type 2 diabetes patients when compared to the controls (Mao *et al.* 2011). In a study using peripheral blood mononuclear transcriptome, a tissue specific interactome (T2Di) was generated and identified 420 molecular signatures associated with type 2 diabetes comorbidity and symptoms. It was observed that at a novel locus near GWAS loci AchE, upstream of *SRRT*, interacts with *JAZF1*, a type 2 diabetes GWAS gene involved in beta cell biology through chromatin regulation and miRNA. The tissue specific interactome (T2Di) identified drug targets *PPARD*, *MAOB* and druggable targets *NCOR2*, *PDGFR* for type 2 diabetes (Li *et al.* 2016). The drawbacks of transcriptomics include failure to link nucleotides and clinical traits, dependence on the limited source of tissue and the lack of feasibility of detecting genetic mechanisms that are not mediated through transcript or protein level (Jenkinson *et al.* 2016).

4.3 Proteomics

The objective of proteomics is to study mechanisms and find out novel drug targets and prognostic markers for early detection of the disease (Garbis *et al.* 2005). Proteomic research on type 2 diabetes revealed downregulation of apolipo-protein A-I and apolipoprotein E and up regulation of leptin, apolipoprotein J and C-reactive protein in the serum as compared to the controls (Riaz *et al.* 2010; Trougakos *et al.* 2002). Stronger spots of Immunoglobulin (Ig), α 1-antitrypsin, α 2-HS-glycoprotein, and complement C fragment were observed in the vitreous humour in the patients of diabetic retinopathy when compared to the macular hole. The pigment epithelium-derived factor was clearly detected in the vitreous humour patients with diabetes

(Nakanishi *et al.* 2002). In the urine samples of patients of type 2 diabetes with nephropathy, seven proteins such as α 1-B glycoprotein, zinc- α 2-glycoprotein, α 2-HS-glycoprotein, vitamin D binding protein, calgranulin B, α 1-antitrypsin, and hemopexin were found to be upregulated (Rao *et al.* 2007). In the saliva of type 2 diabetes patients, 487 unique proteins were identified out of which 65 showed two-fold difference with the controls. Majority of the identified proteins are related to pathways regulating metabolism of trypsin and immune response (Rao *et al.* 2009). In the plasma samples of type 2 diabetes patients, plasma prolactin-induced protein (PIP), thrombospondin-2 (THBS2), L1 cell adhesion molecule (L1CAM) and neutrophil gelatinase-associated lipocalin (NGAL) levels were upregulated. Similarly, the PIP, THBS2 and NGAL in the type 2 diabetes patients with nephropathy (albuminuria) and L1CAM levels in those with retinopathy were also upregulated (Yeh *et al.* 2016). In the urine samples of the diabetics with normal albumin excretion, diabetic pattern of polypeptide excretion was observed whereas in diabetics with high albumin excretion, polypeptide pattern indicating renal damage was observed (Mischak *et al.* 2004). Another study comparing type 2 diabetes with normal and/or macroalbuminuria showed difference of 113 polypeptides suggesting renal damage pattern (Rossing *et al.* 2005). Using the urine samples of microalbuminuria positive type 2 diabetes patients, albumin, zinc alpha-2-glycoprotein, alpha-1-acid glycoprotein, alpha-1-microglobulin and IgG were identified using mass spectrometry and validated by the western blot method (Jain *et al.* 2005).

The limitations of proteomics include complexity in analysis, lack of standardization in sample processing, risk of high false positivity and dynamic range of sample limits the estimation of low abundance of proteins, failure in validation of biomarkers in larger number of patients due to lack of antibodies, inaccessibility of softwares for analysis due to their propriety, difficulty in establishing threshold between signal and noise, low reproducibility, non-application of stringent statistics, ignorance of protein hindrance, failure to homogenize a differentiated tissue, suppression of signal by expression of specialized tissue, no possibility of detecting organelle location of proteins and metabolites of lysate if the sample was homogenized by ante mass spectrometric analysis, influence of age, sex, medication and disease state, failure to extract subcellular organelle from frozen sample, contamination of subfractions and organelle, inefficient separation of organelle, gel to gel variation, failure to detect proteins solubilized in detergent, and acid proteins with pI values lower than 4 and proteins outside the range 10 to 120 kDa (Garbis *et al.* 2005; Sidoli *et al.* 2017; Lasonder 2017).

The primary goal of genomics and proteomic studies is to offer the molecular basis of understanding of the disease, thereby improving the diagnosis, treatment and prevention of diseases. Today, genetic testing is widely applied in medical fields, including newborn screening for highly penetrant disorders and diagnostic and carrier testing for

inherited disorders. On the other hand, the technological improvements have led to the development of methods for predictive and pre-symptomatic testing for the adult-onset complex genetic disorders and pharmacogenetic testing to guide individuals' drug dosage, selection and response (www.acpm.org). The genetic testing registry is the centralized database that provides information on about 49509 tests, 10734 conditions, and 16223 genes with 492 labs offering these tests (www.ncbi.nlm.nih.gov/gtr/). Recently, a workgroup consisting of American college of Medical Genetics and Genomics (ACMG), Association for Molecular Pathology (AMP), and College of American Pathologists (CAP) and the members representing clinical laboratory directors and clinicians provided recommendations and guidelines for interpretative classification of genetic variants, albeit are applicable only in the case of variants that belong to monogenic/polygenic conditions and follow the Mendelian inheritance pattern. Even though genomics is a powerful means of identifying hereditary elements which interact with environmental factors leading to diseases, there is no appropriate regulatory agency that recommends the clinical utility and validity of genetic testing. Moreover, there is a need for training of clinical and genetic professionals in the proper understanding and usage of the terms, definitions and criteria made by ACMG (Richards *et al.* 2015). There is a slow and steady progress in this science pertaining to complex genetic diseases and might take some time to develop such guidelines for determining the causes of these diseases.

4.4 Metabolomics

Metabolome deals with small molecules of less than 1500 Da in cells or body fluids (Abu Bakar *et al.* 2015) and represents the metabolite constituents that include proteins along with other biomolecules of the organism. The metabolite composition of a cell type, tissue or biological fluid are influenced by genetic variants, epigenetic factors, changes in the gene expression and environmental factors and therefore are the most informative molecules of biochemical activity of an organism (Kretowski *et al.* 2016). Nuclear magnetic resonance (NMR) and mass spectrometry (MS) are the two well-known approaches used in generating metabolomics data (Alonso *et al.* 2015). The NMR detects the metabolites based on chemical shift in the resonance spectrum of protons whereas MS characterizes metabolites by retention time and mass-charge ratio. Compared to NMR, MS is more sensitive, cost-effective and has wider availability. Two approaches are employed in both the techniques and they are the untargeted and targeted metabolomics. While the former emphasizes on the unknown and global profiling of metabolites (Zheng and Hu 2015), the latter approach focuses on the quantification of selective known metabolites. Isoleucine, leucine, valine, tyrosine and phenylalanine were found to be the predictors of future

diabetes (Lu *et al.* 2013). Increased serum levels of acylcarnitine and decreased levels of glycine and lysophosphatidylcholine were found to predict the development of impaired glucose tolerance (Wang-Sattler *et al.* 2012). The plasma glutamine, glutamate and glutamine to glutamate ratio were found to be associated with insulin resistance and elevated levels of plasma 2-aminoadipic acid with increased risk of type 2 diabetes (Cheng *et al.* 2012; Wang *et al.* 2013). Ha *et al.* in 2012 have identified significant differences in circulating levels of glucose, triglyceride, oxidized low-density lipoprotein (LDL), high-sensitivity C-reactive protein, interleukin-6, tumor necrosis factor- α and urinary 8-epi-prostaglandin $F_{2\alpha}$ between diabetic and nondiabetic men. Further, a recent study suggested 39 metabolites that belong to amino acid, lipoprotein, carbohydrate and nucleotide metabolisms as commonly associated with diabetes and high body mass index (BMI) providing metabolomic evidence of link between these conditions (Park *et al.* 2015). Reduced levels of organic anion transporters, OAT1 and OAT3, were observed in the renal biopsies of patients with diabetes (Guan *et al.* 2013). Seventeen biomarkers such as Isoleucine, valine, isopropanol, alanine, leucine, acetate, proline, glutamine, arginine, trans-aconitate, creatine, creatinine, glucose, glycine, threonine, tyrosine and 3-methylhistidine were identified in the plasma samples of type 2 diabetes patients with CHD which showed sensitivity and specificity of 92.9% and 93.3%, suggesting high predictive value for type 2 diabetes-CHD (Liu *et al.* 2016). Increased levels of plasma branched chain amino acids, aromatic amino acids and α -hydroxybutyrate and decreased levels of glycine and lysophosphatidylcholine were found to be the predictive markers for the development of type 2 diabetes (Klein and Shearer 2016). In another study, plasma sorbitol, galacticol, mannose, galactose, uric acid, oxalic acid, glucaric acid-1,4-lactone, 3-methyl-2-oxopentanoic acid and 2-hydroxybutyric acid were found to be positively associated with IGT and T2D (Savolainen *et al.* 2017). In the Chinese population plasma alanine, phenylalanine, tyrosine and palmitoylcarnitine were found to be predictive biomarkers of type 2 diabetes (Qiu *et al.* 2016).

One the other hand, nutritional systems biology deals with the interaction between endogenous molecular entities and dietary nutrients to focus on the disease modifying effect of nutritional molecules (Zhao *et al.* 2015). In addition to the diet, smoking, alcohol intake, physical inactivity and obesity are significant risk factors for type 2 diabetes (Chen *et al.* 2012). Diet has been shown to influence insulin sensitivity and other glycemic traits (Zhao *et al.* 2015). Most of the studies were primarily based on animal models, which need to be validated in humans. Studies involving the effect of dietary components are limited by complex diet composition, tissue specific response and dynamic dietary response (Zhao *et al.* 2015). Synder group at Harvard University has developed integrated personal omic profiling system (iPOP) to study the molecular changes under the influence of life style and diet to predict the health consequences. Integration

of nutritional system biology with iPOP can provide insights on the role of diet in progressing or reversing the disease condition (Zhao *et al.* 2015). Further, nutrिमicrobiomics studies revealed low abundance of butyrate-producing bacteria (*Clostridiales* sp. SS3/4, *Eubacteriumrectale*, *Faecalibacterium prausnitzii*, *Roseburia intestinalis*, and *Roseburia inulinivorans*) and high abundance of opportunistic pathogens (*Bacteroides caccae*, *Clostridium hathewayi*, *Clostridium ramosum*, *Clostridium symbiosum*, *Eggerthella lenta*, and *Escherichia coli*) in type 2 diabetes patients (Hartstra *et al.* 2015).

The limitations of metabolomics/nutritional systems biology are the heterogeneity of compounds, dynamic range in concentrations, influence of medications, food and gut microbiota, metabolites involved in multiple pathways and depiction of metabolite levels by quartiles, lack of feasibility to analyze all metabolites simultaneously, an incomplete metabolome, lack of metabolite annotations in the search databases and the low statistical power for enrichment analysis. The limitations with reference to 1. Gas Chromatography-MS: Need for proper derivatization method for analysis of analytes of interest, limited mass range and instrument based variability, 2. Liquid Chromatography-MS: the ion trap used for analysis of ions unable to perform multiple reaction monitoring measurements, 3. MS: availability of only proprietary standalone software, lack of standards for characterization of metabolome due to its diversity, variation in the concentration and lack of software tools to translate the data into biologically meaningful evidence and 4. NMR: which can only measure low abundant metabolites (Sas *et al.* 2015; Aretz and Meierhofer 2016; Scalbert *et al.* 2009).

In a nutshell, genomic and post-genomic approaches identified a large number of biomarkers to ponder over and explore further but we are yet to identify universally accepted biomarker which can be used for the successful management and prevention of type 2 diabetes. In order to understand environment related modifications of genetic susceptibility, it may be prudent to conduct studies with integrated genomic-metabolomic approach. It is also imperative to gather existing molecular genetic data and curate it into uniform format and analyze the same for understanding the present status of research. A few attempts were, however, made to develop type 2 diabetes informative databases. While the databases T2DGADB and T2D-DB are only a collection of publications related to type 2 diabetes genetic association studies, protein-protein interactions and expression studies, T2D@ZJU is a comprehensive collection of pathway databases, protein-protein interaction databases, and literature (Yang *et al.* 2013). Further, T2D@ZJU is a user-friendly interface database that provides graphical output of information organized in networks. These attempts may provide basis for studying type 2 diabetes utilizing systems biology, which is a better approach for understanding complex genetic diseases.

5. Indian scenario

With greater than 69 million of its adult population being diabetic (figure 1), India is considered to be the second diabetic capital of the world (IDF 2015); although China was leading in terms of the absolute number of adults with type 2 diabetes (20–79 years) in 2015, the projections are that India would be doubling its diabetic population in 2040 and outnumber the Chinese by the year 2045. There is a dramatic change in the dietary patterns of Indians over the past decades, which resulted in their unique Asian Indian phenotype, characterized by relatively low BMI, greater abdominal fat, high insulin resistance and high CRP levels, low levels of adiponectin and atherogenic dyslipidemia. Given the characteristic Asian Indian phenotype of the Indian populations, several candidate gene and GWAS variants were studied among them for their association patterns with the disease in the genomic era. The salient features of the results of these studies are presented in table 1. Chandak *et al.* (2007) have studied three SNPs (rs7903146, rs12255372 and rs4506565) of *TCF7L2* and suggested their risk conferring nature in a sample of western Indians. Subsequently, rs7903146 and rs12255372 of the same gene were observed to show similar association in two independent studies among south Indians (Bodhini *et al.* 2007; Uma Jyothi *et al.* 2013). In a sample of Punjabi Sikhs, Sanghera *et al.* (2008a, b) have identified polymorphisms that belong to *TCF7L2*, *IGF2BP2* and *FTO* genes to be significantly risk conferring and *PPARG2* risk reducing in nature towards type 2 diabetes. In another study of 45 most replicated variants of GWAS (Chidambaram *et al.* 2010) has identified rs7756992, rs7754840 and rs6931514 of *CDKAL1* and rs7923837 of *HHEX* to confer risk towards the disease, while rs7020996 and rs12056034 of *CDKN2A/2B* and *BAZ1B*, respectively, to be protective in nature among the south Indians. Similarly, Ali *et al.* (2013) studied 91 most replicated variants of GWAS that belong to 55 genes and identified eight SNPs from five genes (*TCF7L2*, *HHEX*, *ENPP1*, *IDE* and *FTO*) as susceptible genetic factors of type 2 diabetes among three different ethnic groups- Kashmiris, Punjabis and Oriyas. While Chauhan *et al.* (2010) have identified eight most replicated variants from GWAS that belong to *PPARG*, *KCNJ11*, *TCF7L2*, *SLC30A8*, *HHEX*, *CDKN2A*, *IGF2BP2*, and *CDKAL1* genes as risk conferring among the northern and western Indian cohorts, Uma Jyothi *et al.* (2013); Uma Jyothi and Reddy (2015); Kommoju *et al.* (2014 and 2016) have studied fifteen SNPs that belong to the nine genes (above eight genes and *IRS1*) and identified only 6 SNPs, two each from *TCF7L2* (rs7903146, rs12255372) and *CDKAL1* (rs7754840, rs7756992) and one each from *CAPN10* (rs3792267) and *IRS-1* (rs1801278) genes as significantly associated with type 2 diabetes in the population of Hyderabad. Besides suggesting cumulative effect of the risk variants and the high discriminative power of the risk scores, these studies have identified *IGF2BP2*, *SLC30A8*, *HHEX*, *CDKN2A/B*, *PPARG* genes as significantly interacting among them. These studies also suggested environmental factors

such as BMI, alcohol and smoking to be significantly interacting with *TCF7L2*, *CDKAL1* and *CAPN10* genes, providing support for gene–gene and gene–environment interactions in the manifestation of type 2 diabetes in this population. Overall, while the most prominent genes of GWAS such as *TCF7L2* and *CDKN2A/CDKN2B* were validated across the ethnic groups of India, genes such as *IGF2BP2* and *SLC30A8* were found to be associated only among the north Indians (Uma Jyothi *et al.* 2013; Kommoju *et al.* 2013). On the other hand, couple of GWAS on Indians identified novel type 2 diabetes susceptible variants that belong to *SCGC* and *TMEM163* genes (Saxena *et al.* 2013; Tabassum *et al.* 2013). Further, Indian genetic association studies suggest that there is a genetic predisposition of these populations to diabetes under certain environmental triggers (Uma Jyothi and Reddy 2015; Reddy 2013). Particularly, adoption of high fat rich and high carbohydrate diet might be triggering the molecular mechanisms leading to their characteristic Asian Indian phenotype. Unnikrishnan *et al.* (2014) have hypothesized that mechanisms such as impaired non-oxidative glucose disposal and nutrient-sensing mammalian target of rapamycin pathway might be leading to the increasing prevalence of diabetes among Indians. However, these molecular mechanisms need to be elucidated through more extensive studies with appropriate frameworks. Further, studies on post-genomic strategies such as transcriptomics, proteomics, metabolomics and systems biology on type 2 diabetes patients in the Indian context are limited and further studies are needed to understand the disease better for effective management and/or prevention. The hitherto undertaken transcriptomics analysis of visceral adipose tissue of obese diabetics and healthy controls, matched for age and BMI, revealed decreased biosynthesis of unsaturated fatty acids and increased natural killer cell mediated cytotoxicity in obese type 2 diabetes. Proteomic studies carried out in India dealt with saliva and urinary proteome and found that in saliva, 65 proteins showed two fold increase in patients when compared to controls (Rao *et al.* 2009) and in urine samples, albumin, zinc alpha-2-glycoprotein, alpha-1-acid glycoprotein, alpha-1-microglobulin and IgG in microalbuminuria positive type 2 diabetes (Jain *et al.* 2005) and in type 2 diabetes with nephropathy, α 1-B glycoprotein, zinc- α 2-glycoprotein, α 2-HS-glycoprotein, vitamin D binding protein, calgranulin B, α 1-antitrypsin and hemopexin proteins were detected (Rao *et al.* 2007). Metabolomic study showed that elevated levels of saturated fatty acids and amino acids (leucine, isoleucine, lysine, proline, threonine, valine, glutamine, phenylalanine and histidine), lactic acid, 3-hydroxybutyric acid, choline, 3,7-dimethyluric acid, pantothenic acid, myoinositol, sorbitol, glycerol, and glucose were observed in type 2 diabetes with high BMI when compared to the healthy control with low BMI (Gogna *et al.* 2015). Systems biology approach revealed that genes identified for type 2 diabetes through GWAS correlate with insulin secretion and by interacting with other genes, are related to insulin resistance (Jain *et al.* 2013).

Table 1. Salient features of type 2 diabetes genetic association studies in India during the genomic era

Study	Geographical location of the Population(s)	No. of Cases	No. of Controls	No of SNPs/genes studied	SNPs associated	Gene Associated	Nature of Association
Chandak <i>et al.</i> (2007)	Pune, Western India	955	399	3 (<i>TCF7L2</i>)	rs7903146 rs12255372 rs4506565	<i>TCF7L2</i>	Risk Risk Risk
Bodhini <i>et al.</i> (2007)	Chennai, Southern India	1031	1038	2 (<i>TCF7L2</i>)	rs7903146 rs12255372	<i>TCF7L2</i>	Risk Risk
Sanghera <i>et al.</i> (2008a)	Khatri Sikhs, Northern India	556	537	6 (<i>TCF7L2</i>)	rs7903146 rs11196205 rs10885409 rs4918789	<i>TCF7L2</i>	Risk Risk Risk Risk
Sanghera <i>et al.</i> (2008b)	Khatri Sikhs, Northern India	532	386	9 (<i>PPARG2, IGF2BP2, CDK5, SLC30A8, CDKN2A, HHEX, TCF7L2, KCNJ11, FTO</i>)	rs1801282 rs4402960 rs10885409 rs9939609	<i>PPARG2</i> <i>IGF2BP2</i> <i>TCF7L2</i> <i>FTO</i>	Protective Risk Risk Risk
Chauhan <i>et al.</i> (2010)**	Pooled	2486	2678	8 (<i>PPARG, KCNJ11, TCF7L2, SLC30A8, HHEX, CDKN2A, IGF2BP2, CDKALI</i>)	rs1801282 rs5219 rs7903146 rs13266634	<i>PPARG</i> <i>KCNJ11</i> <i>TCF7L2</i> <i>SLC30A8</i>	Risk Risk Risk Risk
	Delhi, North India	1019	1006		rs1111875 rs10811661 rs4402960 rs10946398*	<i>HHEX</i> <i>CDKN2A</i> <i>IGF2BP2</i> <i>CDKALI</i>	Risk Risk Risk Risk
	Pune, Western India	1467	1672		rs7756992 rs7754840 rs6931514 rs7020996	<i>CDKALI</i> <i>CDKALI</i> <i>CDKALI</i> <i>CDKN2A/2B</i>	Risk Risk Risk Protective
Chidambaram <i>et al.</i> (2010)	Chennai, southern India	926	812	45 (<i>AH11, BAZ1B, CDKALI, EXT2, HHEX, IGF2BP2, LOC387761, LOC441171, MC4R, MLXIPL, SLC30A8, STK32C, PPARG, WFS1, and 12unannotated loci</i>)	rs7923837 rs12056034 rs7903146 rs11196205 rs12255372 rs4402960 (NS) rs1470579 (NS) rs13266634 (NS)	<i>HHEX</i> <i>BAZ1B</i> <i>TCF7L2</i>	Risk Protective Risk Risk Risk Risk NS NS NS
Umajyothi <i>et al.</i> (2013)	Hyderabad, Southern India	758	621	3 (<i>TCF7L2</i>)	rs7903146 rs11196205 rs12255372	<i>TCF7L2</i>	Risk Risk Risk
Umajyothi <i>et al.</i> (2015)	Hyderabad, Southern India	758	621	3 (<i>IGF2BP2, SLC30A8</i>)	rs4402960 (NS) rs1470579 (NS) rs13266634 (NS)	NS NS NS	NS NS NS
Ali <i>et al.</i> (2013)**	Pooled Kashmir Punjab Orissa	1583 507 649 427	1317 600 300 417	91 SNPs belong to 55 genes, identified through GWAS mostly in European population and replicated in other populations	rs7903146 rs12255372 rs1887922 [¶] rs1111875 [¶] rs5015480 [¶] rs1044498 [¶] rs9939609 [#] rs3751812 [#] rs1801278	<i>TCF7L2</i> <i>TCF7L2</i> <i>IDE</i> <i>HHEX</i> <i>HHEX</i> <i>ENPP1</i> <i>FTO</i> <i>FTO</i> <i>IRS1</i>	Risk Risk Risk Risk Risk Risk Risk Risk
Kommoju <i>et al.</i> (2014)				4 (<i>IRS1, CAPN10, PPARG</i>)	rs3792267 rs5030952 rs1801282 rs7754840	<i>CAPN10</i> <i>CAPN10</i> <i>PPARG2</i> <i>CDKALI</i>	Risk Risk NS Risk
Kommoju <i>et al.</i> (2016)	Hyderabad, Southern India	758	621	5 (<i>CDKALI, CDKN2A/2B, HHEX</i>)	rs7756992 rs10811661 rs1111875 rs7923837	<i>CDKALI</i> <i>CDKALI</i> <i>CDKN2A/2B</i> <i>HHEX</i> <i>HHEX</i>	Risk Risk Risk NS NS

**Results presented for pooled cohort; *Not significant in the Pune cohort when two cohorts were analyzed separately; [#]Not significant in Kashmir cohort, [¶]Not Significant in Orissa cohort; NS – Not significant.

6. A framework for future Indian studies

To gauge the gene–environment interactions, several statistical methods have been developed where statistical software packages of R-program, PLINK (Purcell *et al.* 2007) and GMDR (Chen *et al.* 2011) are widely used for this purpose in the genetic association studies. Kaput and Dawson (2007) suggested a nutri-genomics approach which is a study of how food affects the gene expression and how genetic makeup affects the metabolism in response to nutrients. However, in order to ascertain that the manifestation of complex diseases is a result of changed food habits and urbanized life styles, it is apt to study groups of individuals that represent the transitional food habits but with genetic homogeneity (Reddy 2013). Indian population structure is unique and characterized by the division of its population into strictly defined hierarchical castes, tribes and religious groups of diverse ethnic and cultural backgrounds, which practice strict endogamy and therefore highly substructured. Each of these hierarchical castes, tribes and religious groups are known to be subdivided into number of endogamous subunits.

The impact of urbanization is fast spreading across all the social and economic strata of the Indian population and has even percolated into remotest of the tribal areas of the country. Consequently, the tribal populations are also affected by the urbanization to varying degrees resulting in certain proportion of their population moving to nearby towns and cities. This process has created many experimental situations in which a single genetically homogenous tribe presents its population into transitional groups – urban, semi-urban and rural – with different degrees of urbanization and graded life styles, which may provide excellent study frame to assess the role of specific environmental factors in the manifestation of complex diseases, particularly type 2

diabetes (Reddy 2013). In this context, as discussed in the previous section, metabolomics and systems biology approaches would be the right way forward to get appropriate understanding of these mechanisms.

The present scenario of at least urban India is such that almost every second adult person in the 40+ year's age group seems to be reporting as diabetic and the rural population may not be lagging far behind. This prompts one to surmise if we would not find diabetic population to outnumber the normal adult population in the foreseeable near future, qualifying type 2 diabetes as the 'normal phenotype'! Given this scenario, Reddy (2013) has hypothesized that the putative genetic variants would be uniformly prevalent across the populations and it is the environmental triggers that make the difference between those who are affected (cases) and those who are not (controls). This could be more aptly reflected in the tribal populations of India with transitional groups, which may reflect genetic homogeneity by presenting similar frequencies of susceptible/protective genetic variants among them but with environmental heterogeneity as reflected in the increasing urbanization and graded lifestyles, hence increased prevalence of type 2 diabetes, from rural to urban areas (figure 2). The outcome of testing this hypothesis among the genetically homogenous and environmentally graded subunits of a single caste/tribe in India would naturally nullify the problems of confounding factors in the association studies and help in precisely identifying the role of changing lifestyles/urbanization in the manifestation of complex diseases and their increasing prevalence. We have initiated such a study among the three transitional tribes of undivided AP with a considerable proportion of their population settled in semi-urban and urban locales. Such an effort among ethnically and geographical heterogeneous populations across India would help comprehensive understanding of the genetic and environ-

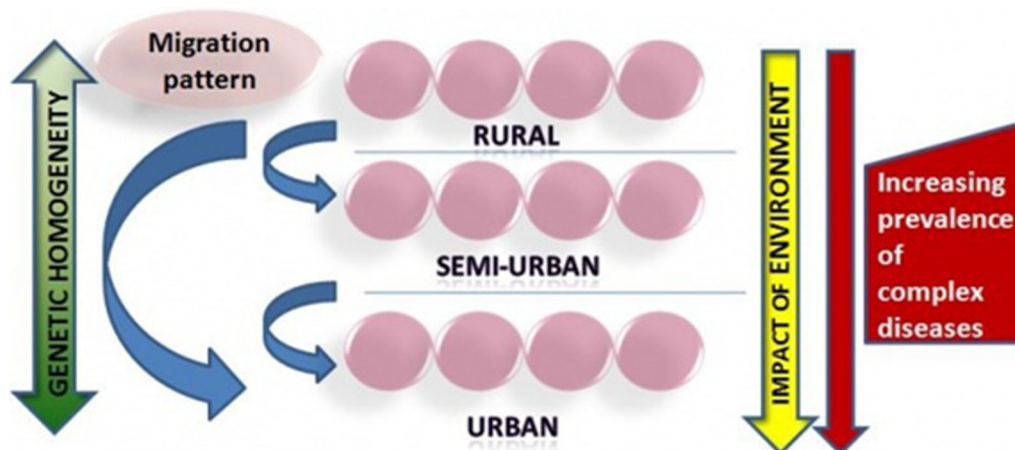


Figure 2. Schematic representation of the transitional groups within genetically homogenous castes/tribes during the process of urbanization leading to increasing prevalence of complex diseases.

mental triggers behind the fast emerging epidemic of diabetes and other common diseases such as CVDs, cancers and obesity in India.

7. Conclusions

In view of the overwhelming inconsistency observed in the results of genetic association studies of type 2 diabetes across the globe, it is pertinent to design the future studies in a way that neutralizes the confounding factors and provides useful results. It is equally important to curate the existing data and reanalyze it through advanced computational methods in the era of systems biology. Further, we need functional studies that complement the pace of genomic research. The post-genomic strategies are perplexed with practical difficulties; yet it is imperative to overcome those and conduct integrated genomic-metabolomic studies to derive meaningful outcomes of practical utility. These approaches may provide better insights into understanding the molecular mechanisms operating in the manifestation of the disease and may help in devising methods for prevention and/or treatment.

Given the geographical and ethnic variability in the prevalence of type 2 diabetes, prior knowledge on ethnic/cultural background of the study population may help in devising an apt population specific study framework, as for example, the one proposed by us (Reddy 2013) among the AP tribes and discussed in the previous section. Particularly, a detailed survey of the dietary intake and physical activity patterns that play a major role in manifestation of type 2 diabetes needs to be studied in the above frame work and appropriate quantitative measures/indices of energy intake and expenditure derived at the individual level before taking plunge into nutri-genomics and/metabolomics components of the disease manifestation.

References

- Abu Bakar MH, Sarmidi MR, Cheng KK, Ali Khan A, Suan CL, ZamanHuri H and Yaakob H 2015 Metabolomics - the complementary field in systems biology a review on obesity and type 2 diabetes. *Mol. Biosyst.* **11** 1742–1774
- Ali O 2013 Genetics of type 2 diabetes. *World J. Diabetes* **4** 114–123
- Ali S, Chopra R, Manvati S, Singh YP, Kaul N, Behura A, Mahajan A, Sehajpal P *et al.* 2013 Replication of type 2 diabetes candidate genes variations in three geographically unrelated Indian population groups. *PLoS ONE* **8** e58881
- Alonso A, Marsal S and Julià A 2015 Analytical methods in untargeted metabolomics state of the art in 2015. *Front. Bioeng. Biotechnol.* **3** 23.
- Aretz I and Meierhofer D 2016 Advantages and pitfalls of mass spectrometry based metabolome profiling in systems Biology. *Int. J. Mol. Sci.* **17** 632
- Ashcroft FM and Rorsman P 2012 Diabetes mellitus and the β cell the last ten years. *Cell* **148** 1160–1171
- Bodhini D, Radha V, Dhar M, Narayani N and Mohan V 2007 The rs12255372(G/T) and rs7903146(C/T) polymorphisms of the TCF7L2 gene are associated with type 2 diabetes mellitus in Asian Indians. *Metabolism* **56** 1174–1178
- Boj SF, van Es JH, Huch M, Li VS, José A, Hatzis P *et al.* 2012 Diabetes risk gene and Wnt effector TCF7L2/TCF4 controls hepatic response to perinatal and adult metabolic demand. *Cell* **151** 1595–1607
- Chandak GR, Janipalli CS, Bhaskar S, Kulkarni SR, Mohankrishna P, Hattersley AT, *et al.* 2007 Common variants in the TCF7L2 gene are strongly associated with type 2 diabetes mellitus in the Indian population. *Diabetologia* **50** 63–67
- Chauhan G, Spurgeon CJ, Tabassum R, Bhaskar S, Kulkarni SR, Mahajan A, *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 BH, Hivert MF, Peters MJ, Pilling LC, Hogan JD, Pham LM, Harries LW, Fox CS, *et al.* 2016 Peripheral blood transcriptomic signatures of fasting glucose and insulin concentrations. *Diabetes* **65** 3794–3804
- Chen GB, Xu Y, Xu HN, Li MD, Zhu J, Lou XY 2011 Practical and theoretical considerations in study design for detecting gene-gene interactions using MDR and GMDR approaches. *PLoS ONE* **6** e16981
- Chen L, Magliano DJ and Zimmet PZ 2012 The worldwide epidemiology of type 2 diabetes mellitus-present and future perspectives. *Nat. Rev. Endocrinol.* **8** 228–236
- Cheng S, Rhee EP, Larson MG, Lewis GD, McCabe EL, Shen D, *et al.* 2012 Metabolite profiling identifies pathways associated with metabolic risk in humans. *Circulation* **125** 2222–2231
- Chidambaram M, Radha V and Mohan V 2010 Replication of recently described type 2 diabetes gene variants in a south Indian population. *Metabolism* **59** 1760–1766
- Delgado-Lista J, Perez-Martinez P, García-Rios A, Phillips CM, Williams CM, Gulseth HL *et al.* 2011 Pleiotropic effects of TCF7L2 gene variants and its modulation in the metabolic syndrome from the LIPGENE study. *Atherosclerosis* **214** 110–116
- Garbis S, Lubec G and Fountoulakis M 2005 Limitations of current proteomics technologies. *J. Chromatogr. A* **1077** 1–18
- Genetic testing time tool 2017 A resource from American College of Preventive Medicine www.acpm.org
- Genetic testing registry 2017 <https://www.ncbi.nlm.nih.gov/gtr/>
- Gogna N, Krishna M, Oommen AM, Dorai K 2015 Investigating correlations in the altered metabolic profiles of obese and diabetic subjects in a south Indian Asian population using an NMR-based metabolomics approach. *Mol. Biosyst.* **11** 505–606
- Grarup N, Sandholt CH, Hansen T and Pedersen O 2014 Genetic susceptibility to type 2 diabetes and obesity from genome-wide association studies to rare variants and beyond. *Diabetologia* **57** 1528–1541
- Grant SF, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, Sainz J *et al.* 2006 Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat. Genet.* **38** 320–323
- Guan M, Xie L, Diao C, *et al.* 2013 Systemic perturbations of key metabolites in diabetic rats during the evolution of diabetes studied by urine metabolomics. *PLoS ONE* **8** e60409

- Gupta V, Khadgawat R, Saraswathy KN, Sachdeva MP and Kalla AK 2008 Emergence of TCF7L2 as a most promising gene in predisposition of diabetes type II. *Int. J. Hum. Genet.* **8** 199–215
- Ha CY, Kim JY, Paik JK, Kim OY, Paik YH, Lee EJ and Lee JH. 2012 The association of specific metabolites of lipid metabolism with markers of oxidative stress, inflammation and arterial stiffness in men with newly diagnosed type 2 diabetes. *Clin. Endocrinol.* **76** 674–682
- Hartstra AV, Bouter KEC, Backhed F and Nieuwdorp M. 2015 Insights into the role of the microbiome in obesity and type 2 diabetes. *Diabetes Care* **38** 159–165
- Hernández-Alvarez MI, Diaz-Ramos A, Berdasco M, Cobb J, Planet E, Cooper D, Pazderska A, Wanic K et al. 2017 Early-onset and classical forms of type 2 diabetes show impaired expression of genes involved in muscle branched-chain amino acids metabolism. *Sci. Rep.* **7** 13850
- International Diabetes Federation 2015 *IDF Diabetes Atlas* 7th edition (Brussels, International Diabetes Federation)
- Jain S, Rajput A, Kumar Y, Uppuluri N, Arvind AS and Tatu U 2005 Proteomic analysis of urinary protein markers for accurate prediction of diabetic kidney disorder. *J. Assoc. Physicians India* **53** 513–520
- Jenkinson CP, Göring HHH, Arya R, Blangero J, Duggirala R and DeFronzo RA 2016 Transcriptomics in type 2 diabetes bridging the gap between genotype and phenotype. *Genomics Data* **8** 25–36
- Kaput J and Dawson K 2007 Complexity of type 2 diabetes mellitus data sets emerging from nutrigenomic research a case for dimensionality reduction? *Mutat. Res.* **622** 19–32
- Klein MS and Shearer J 2016 Metabolomics and type 2 diabetes translating basic research into clinical application. *J. Diabetes Res.* 2016. <http://dx.doi.org/https://doi.org/10.1155/2016/3898502>
- Kommoju UJ and Reddy BM 2011 Genetic etiology of type 2 diabetes mellitus A review. *Int. J. Diabetes Dev. Countries* **31** 51–64
- Kommoju UJ, Maruda J, Subbaraj SK, Irgam K, Kotla JP, Velaga L and Reddy BM 2013 No detectable association of IGF2BP2 and SLC30A8 genes with type 2 diabetes in the population of Hyderabad, India. *Meta Gene* **1** 15–23
- Kommoju UJ, Maruda J, Subbaraj SK, Irgam K, Kotla JP and Reddy BM 2014 Association of IRS1, CAPN10, and PPARG gene polymorphisms with type 2 diabetes mellitus in the high-risk population of Hyderabad, *Indian J. Diabetes* **6** 564–573
- Kommoju UJ, Subbaraj SK, Maruda J, Irgam K, Kotla JP, Velaga L and Reddy BM 2016 Association of CDKAL1, CDKN2A/B & HHEX gene polymorphisms with type 2 diabetes mellitus in the population of Hyderabad, India. *Indian J. Med. Res.* **143** 455–63
- Kretowski A, Francisco J, Ruperez and Michal Ciborowski 2016 Genomics and metabolomics in obesity and type 2 diabetes. *J. Diabetes Res.* <https://doi.org/10.1155/2016/9415645>
- Lasonder E 2017 Clinical proteomics from biological sample to clinical exploitation. *Proteomes* **5** 10
- Li JW, Lee HM, Wang Y, Tong AH, Yip KY, Tsui SK, Lok S, Ozaki R et al. 2016 Interactome-transcriptome analysis discovers signatures complementary to GWAS Loci of Type 2 diabetes. *Sci. Rep.* **18** 35228
- Lim JE, Hong K-W, Jin H-S, Kim YS, Park HK and Oh B 2010. Type 2 diabetes genetic association database manually curated for the study design and odds ratio. *BMC Med. Inform. Decis. Mak.* **10** 76
- Lin Y and Sun Z 2010 Current views on type 2 diabetes. *J. Endocrinol* **204** 1–11
- Liu X, Gao J, Chen J, Wang Z, Shi Q, Man H, Guo S, Wang Y, Li Z and Wang W 2016 Identification of metabolic biomarkers in patients with type 2 diabetic coronary heart diseases based on metabolomic approach. *Sci. Rep.* **29** 30785
- Lu J, Xie G, Jia W and Jia W 2013 Metabolomics in human type 2 diabetes research. *Front. Med.* **7** 4–13
- Mao J, Ai J, Zhou X, Zhou X, Shenwu M, Ong M, Blue M, Washington JT, Wang X and Deng Y 2011 Transcriptomic profiles of peripheral white blood cells in type II diabetes and racial differences in expression profiles. *BMC Genomics* **12** S12
- Matone A, Derlindati E, Marchetti L, Spigoni V, Dei Cas A, Montanini B, Ardigo D, Zavaroni I et al. 2017 Identification of an early transcriptomic signature of insulin resistance and related diseases in lymphomonocytes of healthy subjects. *PLoS ONE* **12** e0182559
- Mischak H, Kaiser T, Walden M, Hillmann M, Wittke S, Herrmann A, Knueppel S, Haller H and Fliser D 2004 Proteomic analysis for the assessment of diabetic renal damage in humans. *Clin. Sci.* **107** 485–495
- Nakanishi T, Koyama R, Ikeda T and Shimizu A 2002 Catalogue of soluble proteins in the human vitreous humor comparison between diabetic retinopathy and macular hole. *J. Chromatogr. B* **776** 89–100
- Olokoba AB, Obateru OA and Olokoba LB 2012 Type 2 diabetes mellitus: A review of current trends. *Oman Med. J.* **27** 269–273
- Park S, Sadanala KC and Kim E-K 2015 A metabolomic approach to understanding the metabolic link between obesity and diabetes. *Mol. Cells* **38** 587–596
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P et al. 2007 PLINK a tool set for whole-genome association and population-based linkage analysis. *Am. J. Hum. Genet.* **81** 559–575
- Qiu G, Zheng Y, Wang H, Sun J, Ma H, Xiao Y, Li Y, Yuan Y et al. 2016 Plasma metabolomics identified novel metabolites associated with risk of type 2 diabetes in two prospective cohorts of Chinese adults. *Int. J. Epidemiol.* **45** 1507–1516
- Rao PV, Lu X, Standley M, Pattee P, Neelima G, Girishesh G, Dakshinamurthy KV, Roberts CT Jr and Nagalla SR. 2007 Proteomic identification of urinary biomarkers of diabetic nephropathy. *Diabetes Care* **30** 629–637
- Rao PV, Reddy AP, Lu X, Dasari S, Krishnaprasad A, Biggs E, Roberts CT and Nagalla SR. 2009 Proteomic identification of salivary biomarkers of type-2 diabetes. *J. Proteome Res.* **8** 239–245
- Reddy BM 2013 Genetic etiology of complex genetic disorders A framework for future Indian studies; In Proceedings of UGC Sponsored National Seminar on *Incidence and Prevalence of Mendelian Traits & Diseases in People of Odisha* (ed) B Das (Cuttack, Agraduta, Banka Bazar) pp 96–119
- Riaz S, Alam SS and Akhtar MW 2010 Proteomic identification of human serum biomarkers in diabetes mellitus type2. *J. Pharm. Biomed. Anal.* **51** 1103–1107
- Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M et al 2015 ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants a joint consensus recommendation of the

- American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **17** 405–424
- Rossing K, Mischak H, Parving HH, Christensen PK, Walden M, Hillmann M and Kaiser T 2005 Impact of diabetic nephropathy and angiotensin II receptor blockade on urinary polypeptide patterns. *Kidney Int.* **68** 193–205
- Sanghera DK, Nath SK, Ortega L, Gambarelli M, Kim-Howard X, Singh JR, Ralhan SK, Wander GS *et al.* 2008a TCF7L2 polymorphisms are associated with type 2 diabetes in Khatri Sikhs from North India: genetic variation affects lipid levels. *Ann. Hum. Genet.* **72** 499–509
- Sanghera DK, Ortega L, Han S, Singh J, Ralhan SK, Wander GS, Mehra NK, Mulvihill *et al.* 2008b Impact of nine common type 2 diabetes risk polymorphisms in Asian Indian Sikhs: PPARG2 Pro12Ala), IGF2BP2, TCF7L2 and FTO variants confer a significant risk. *BMC Med. Genet.* **9** 59
- Sas KM, Karnovsky A, Michailidis G and Pennathur S 2015 Metabolomics and diabetes analytical and computational approaches. *Diabetes*: **64** 718–732
- Savolainen O, Fagerberg B, Vendelbo Lind M, Sandberg A-S, Ross AB and Bergstrom G 2017 Biomarkers for predicting type 2 diabetes development can metabolomics improve on existing biomarkers? *PLoS ONE* **12** e0177738
- Savic D, Ye H, Aneas I, Park SY, Bell GI and Nobrega M A 2011. Alterations in TCF7L2 expression define its role as a key regulator of glucose metabolism. *Genome Res.* **21** 1417–1425
- Saxena R, Saleheen D, Been LF, Garavito ML, Braun T, Bjorntjes A, Young R, Ho WK *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** 1746–1755
- Scalbert A, Brennan L, Fiehn O, Hankemeier T, Kristal BB, van Ommen B, Pujos-Guillot E, Verheij E *et al.* 2009 Mass-spectrometry-based metabolomics limitations and recommendations for future progress with particular focus on nutrition research. *Metabolomics* **5** 435–458
- Scott LJ, Bonnycastle LL, Willer CJ, Sprau AG, Jackson AU, Narisu N, Duren WL, Chines PS *et al.* 2006 Association of transcription factor 7-like 2 (TCF7L2) variants with Type 2 diabetes in a Finnish sample. *Diabetes* **55** 2649–2653
- Scott LJ, Erdos MR, Huyghe JR, Welch RP, Beck AT, Wolford BN, Chines PS and Didion JP *et al.* 2016 The genetic regulatory signature of type 2 diabetes in human skeletal muscle. *Nat. Commun.* **29** 11764
- Segerstolpe Å, Palasantza A, Eliasson P, Andersson EM, Andréasson AC, Sun X, Picelli S and Sabirsh A *et al.* 2016 Single-cell transcriptome profiling of human pancreatic islets in health and type 2 diabetes. *Cell Metab.* **24** 593–607
- Sidoli S, Kulej K and Garcia BA 2017 Why proteomics is not the new genomics and the future of mass spectrometry in cell biology. *J. Cell Biol.* **216** 21–24
- Silander K, Tang H, Myles S, Jakkula E, Timpson NJ, Cavalliforza L and Peltonen L 2009 Worldwide patterns of haplotype diversity at 9p21.3, a locus associated with type 2 diabetes and coronary heart disease. *Genome Med.* **1** 51
- Sonksen P and Sonksen J 2000 Insulin understanding its action in health and disease. *Br. J. Anaesth.* **85** 69–79
- Stephen LA, Kathy B, Barb S and Laura W 2004 Glucose metabolism and regulation beyond insulin and glucagon. *Diabetes Spec.* **17** 183–190
- Tabassum R, Chauhan G, Dwivedi OP, Mahajan A, Jaiswal A, Kaur I, Bandesh K, Singh T, *et al.* 2013 Genome-wide association study for type 2 diabetes in Indians identifies a new susceptibility locus at 2q21. *Diabetes* **62** 977–86
- Tibaldi JM 2013 The future of insulin therapy for patients with type 2 diabetes mellitus. *J. Am. Osteopath. Assoc.* **113** S29–S39
- Trougakos IP, Poulakou M, Stathatos M, Chalikiab A, Melidonisb A and Gonosa FS 2002 Serum levels of the senescence biomarker clusterin/apolipoprotein J increase significantly in diabetes type II and during development of coronary heart disease or at myocardial infarction. *Exp. Gerontol.* **37** 1175–1187
- Uma Jyothi K, Jayaraj M, Subburaj KS, Prasad KJ, Kumuda I, Lakshmi V and Reddy BM 2013 Association of TCF7L2 gene polymorphisms with T2DM in the population of Hyderabad, India. *PLoS ONE* **8** e60212
- Uma Jyothi K and Reddy BM 2015 Gene–gene and gene–environment interactions in the etiology of type 2 diabetes mellitus in the population of Hyderabad, India. *Meta Gene* **5** 9–20
- Unnikrishnan R, Anjana RM and Mohan V 2014 Diabetes in South Asians: is the phenotype different? *Diabetes* **63** 53–55
- Wang TJ, Ngo D, Psychogios N, Dejam A, Larson MG, Vasan RS, Ghorbani A, O’Sullivan J *et al.* 2013 2-Aminoadipic acid is a biomarker for diabetes risk. *J. Clin. Invest.* **123** 4309 – 4317
- Wang YJ, Schug J, Won KJ, Liu C, Naji A, Avrahami D, Golson ML and Kaestner KH 2016 Single-cell transcriptomics of the human endocrine pancreas. *Diabetes* **65** 3028–38
- Wang-Sattler R, Yu Z, Herder C, Messias AC, Floegel A, He Y, Heim K, Campillos M *et al.* 2012 Novel biomarkers for prediabetes identified by metabolomics. *Mol. Syst. Biol.* **8** 615
- Winnier DA, Fourcaudot M, Norton L, Abdul-Ghani MA, Hu SL, Farook VS, Coletta DK, Kumar S *et al.* 2015 Transcriptomic identification of ADH1B as a novel candidate gene for obesity and insulin resistance in human adipose tissue in Mexican Americans from the Veterans Administration Genetic Epidemiology Study (VAGES). *PLoS One* **10** e0119941
- Yang Z, Yang J, Liu w, Wu L, Zing L, Wang Y, Fan X and Cheng Y 2013 T2D@ZJU A knowledgebase integrating heterogeneous connections associated with type 2 diabetes mellitus. *Database* 2013 bat052. <https://doi.org/10.1093/database/bat052>
- Yeh S-H, Chang W-C, Chuang H, Huang H-C, Liu R-T and Yang KD 2016 Differentiation of type 2 diabetes mellitus with different complications by proteomic analysis of plasma low abundance proteins. *J. Diabetes Metab. Dis.* **15** 24
- Zhang C, Qi L, Hunter DJ, Meigs ZB, Manson JAE, van Dam RM, Hu FB 2006 Variant of transcription factor 7-like 2 (TCFL2) gene and the risk of type 2 diabetes in large cohorts of US women and men. *Diabetes* **55** 2645–2648
- Zhao Y, Barrere-Cain RE and Yang X 2015 Nutritional systems biology of type 2 diabetes. *Genes Nutr.* **10** 31
- Zheng Y and Hu FB 2015 Comprehensive metabolomic profiling of type 2 diabetes. *Clin. Chem.* **61** 45