Name of Journal: World Journal of Clinical Cases

Manuscript NO: 74561

Manuscript Type: MINIREVIEWS

Diabetes mellitus susceptibility with varied diseased phenotypes and it’s comparison with phenome interactome networks

Madhusmita Rout, Bhumandeep Kour, Sugunakar Vuree, Sajitha S. Lulu, Krishna Mohan Medicherla, Prashanth Suravajhala

Abstract
An emerging area of interest in understanding disease phenotypes is systems genomics. Complex diseases such as diabetes have played an important role towards understanding the susceptible genes and mutations. A wide number of methods have been employed and strategies such as polygenic risk score, allele frequencies have been useful but understanding the candidate genes harboring those mutations is an unmet goal. In this perspective, using systems genomic approaches, we highlight the application of phenome-interactome networks in diabetes and provide deep insights. A LINC01128 which we previously described as candidate for diabetes is shown as an example to discuss the approach.

INTRODUCTION

Introduction
Diabetes mellitus occurs as a result of insufficient insulin production or impaired insulin sensitivity, and it has become a serious threat to people's health [1,2]. It is a heterogeneous problem with numerous aetiologies comprising three main types, viz. type 1 diabetes (T1D), type 2 diabetes (T2D) and gestational diabetes mellitus (GDM). Understanding the biological mechanisms associated would allow us to identify
candidate proteins and genes [3]. The emergence of genome-wide association studies (GWASs) has substantially enhanced our understanding of the genetic basis of disease risk in the past few years. Prior to the introduction of GWASs in 2006, very little information was available about the genes that influence common complicated or multifactorial diseases and quantitative traits. These research findings imply that susceptibility to prevalent diseases is influenced by a variety of genetic topologies, including common genetic variants with minimal effects and uncommon variants with substantial impact sizes [4, 5, 6]. Nevertheless, the combination of candidate T2D genes discovered using GWASs does not fully confirm established features of disease pathogenesis. Several system-level approaches have been used to bridge the gap between genome and phenome correlation [7]. Computational analyses of disease linked genes using interactome and toxicogenomic data help us to connect T2D candidate genes found in GWAS with disease pathophysiology, including abnormal pancreatic cell formation and function, and insulin sensitivity. On the other hand, computational predictions of potential proteins/genes are less expensive and time-saving than experimental methods [8, 9]. In order to unravel the genetic roots of common disorders, it is necessary to understand the complexity of the gene-phenotype connection. Recent research employing the human interactome and phenome has uncovered not just common phenotypic and genetic overlap between diseases, but also a modular architecture of the genetic landscape of human diseases, opening up new avenues for reducing the complexity of human diseases [10, 11]. Because diseases are rarely caused by the malfunction of a single protein, a more comprehensive and robust interactome is essential for identifying groups of interconnected proteins associated with disease aetiology [12].

**Phenome interaction networks**

The phenome interaction networks are used to study the wide range of phenotypic traits based on the analysis of the complete genome; it follows a genotypic to phenotypic approach in order to analyse the phenotypic traits [13]. The diseases with
overlapping clinical signs can be predicted because of the mutation in different genes which are playing a role in similar functions. More recently the studies on humans as well as model organisms has revealed that the primary or secondary association between proteins can also be one of the reasons of the same phenotype that means the mutation in particular protein along with its direct or indirect association with a single or multiple proteins can be responsible for overlapping of the clinical manifestations. The opposite scenario can also be analysed using a phenome interactome network, in case of pleiotropy, the cases in which a single gene is responsible for different phenotypic traits. The protein-protein interaction (PPI) network models are used to analyse the phenomic traits which in turn is helpful in understanding cell signalling, drug development, in the diseased as well as normal cell physiology, basically it is important to understand almost every process of the cell. Protein-protein interaction networks are the mathematical representation of physical interaction between similar or different proteins for the analysis of phenomes. The mathematical representation of interaction among different proteins in PPIs is based upon graph theory where the proteins are represented as nodes and edges to depict the type of interaction between two different interacting proteins. Protein-protein interaction networks help to find the genes for a particular disease with a huge accuracy and when PPIs are implemented on the large datasets, it could lead to prediction of novel gene candidates. The phenome interaction networks are quite important to understand and mine the gene associated with a particular disease. The genes that are responsible for similar functions have a higher chance of having same phenotypes therefore; understanding phenotypic as well as genotypic data is a must in order to understand the origination and development of a disease at the systems biology level for the better treatment. The origin and cause of several complex diseases including cancer, diabetes and obesity can be understood by protein-protein interaction network analysis.

**Gestational diabetes mellitus (GDM)**

Gestational diabetes mellitus (GDM) is categorised as insulin resistance leading to hyperglycemia during pregnancy, which mostly retracts after parturition. According to
WHO the prevalence rate is 15.8% accounting to about 20.4 million live births with the majority of cases in pregnant women above the age of 35 years. The International Diabetes Federation in 2019 estimated a prevalence of 28.5% in India with incidence varying in each state due to challenges in screening strategies and paucity of consensus among physicians and healthcare providers in prepartum and postpartum management of GDM [19]. The diagnostic criteria may differ worldwide, understanding the pathophysiology is crucial as it affects both the mother and the fetus during gestation, delivery and later stages of life making them susceptible to diabetes, obesity and cardiovascular complications in the long [47]. Major challenges that have governed this disease are the guidelines for screening and diagnosis. The testing criteria is different with varying forms of oral glucose tolerance test (OGTT) being followed worldwide [20]. Management of GDM is another challenge as both mother and fetus are at risk in their current milieu. Studies have highlighted the importance of treating GDM, reducing the risk of perinatal morbidity and improving post-delivery outcomes [21]. Glucose intolerance leads to the manifestation of the disease hence the benchmark of GDM treatment should be glycaemic control which is achieved through lifestyle intervention such as diet and exercise, pharmacological intervention such as insulin, oral drugs, herbal medicines and finally postnatal management [22].

Pregnant women with GDM have an inherent risk of developing Type 2 diabetes mellitus (T2DM) post-delivery or later on in life. The offspring is also susceptible to any form of diabetes postnatal or long term. The genetic factors responsible for GDM and future risk of developing T2DM through epidemiological and physiological studies reveal commonality in susceptibility loci which implies that most of the diabetes genes are involved in causing GDM. The list of few key genes that share common variants are 
\textit{KCNJ11, GCK, HNF1A, TCF7L2, CDKAL1, KCNQ1, CDKN2A, MTNR1B, SRR, HHEX, TCF2, SLC30A8 and IGF2BP2} [23, 24]. Genetic similarities between Type 1 diabetes mellitus (T1DM) and GDM is less studied, study among Asian Indian women with GDM showed the presence of pancreatic autoantibodies like GAD which is a biomarker
for T1DM \cite{25}. Maturity onset diabetes of young (MODY) are of different types and each type is characterised by a single gene and few studies have shown mutations in HNF1A and HNF4A are MODY genes which are prone to GDM \cite{26}.

Integrating phenotypic data with genotypic data through a computationally created high-confidence interaction network to analyse human diseases concurrently defines a phenome-interactome network \cite{14}. An organized study on genes expressed in thigh subcutaneous adipose tissue of Asian Indian Type 2 Diabetes Mellitus (AIT2DM) revealed evidence of "sick thigh fat" as a causative disease. The phenome-interactome network had significant correlation of differentially expressed genes (DEGs) and hub proteins with its phenotypic traits obtained from clinical, biochemical, and radiological, cellular and molecular, thus enumerating their role in T2DM, T1DM and obesity \cite{27}. RNA-Seq analysis enables identification of differentially expressed genes and its role in a disease. The depth of literature available on RNA-Seq analysis performed on pregnant ladies with GDM is negligible. The GDM is a condition in which the intrauterine milieu, especially the placenta, plays a central role in altering the course of the fetus. Hence having an understanding of the key genes regulated in the placenta is paramount for the disease diagnosis. Most of the literature available on RNA-Seq analysis is centred on identifying DEGs in the placenta, umbilical cord and amniocytes \cite{28, 29, 30, 31}. Studies have identified non-coding RNAs such as long non-coding (Inc)RNAs, microRNAs, and circular RNAs play a central role in GDM pathogenesis. MicroRNAs have been identified as non-invasive early diagnostic biomarkers for GDM \cite{32}. LncRNA-associated feed-forward loops (Lnc-FFLs) network had a strong correlation between dysregulated glucose metabolism and hormone regulation in GDM cases \cite{33}. The mechanism governing the pathophysiology of the disease is still not clear and the studies available are limited. Hence the current problem is to understand the genetic background that affects both the maternal and fetal with changes in the intrauterine environment and thus identify early diagnostic biomarkers. GDM is associated with a number of comorbidities due to the multifactorial nature of the disease. A study to
identify key genes involved in GDM maternal and placental milieu revealed associations with T2DM, T1DM, obesity, hyperglycaemia, preeclampsia, neonatal diabetes, maturity onset diabetes of young, neurological disorders, cardiovascular disease, preeclampsia, hepatitis C, rheumatoid arthritis and neoplasms [34]. Hence the need to identify genes governing this disease, the variations that might affect the phenotype needs to be understood.

**Prostate cancer (PCa) and Diabetes, LINCO1128**

As glucose level in the body is regulated by insulin, a hormone (peptide) which increases the glucose up taking and its assimilation. However, insulin resistance is stated when it becomes unable to perform this function in a diabetic patient. On the other hand, the beta cell continuously secretes insulin to make up and maintain balance but it results in hyperinsulinemia [35]. This increased level will trigger the production of IGF-1 from liver cells. IGF-1 will then bind to its tyrosine kinase receptor IGF-1R, α, to and stimulate various metabolic and mitogenic signalling pathways to control processes like cancer cell proliferation, differentiation, and apoptosis. Later some downstream targets like PI3KB and rat sarcoma-mitogen-activated protein kinase/extracellular signal regulated kinase signaling pathways get stimulated. PI3KB signaling has a role in cancer cell survival and migration. While rat sarcoma mitogen-activated protein kinase/extracellular signal regulated kinase signaling pathway to control cancer cell proliferation and metabolism [36]. Hence, patients who have diabetes show increased levels of IGF-1 bringing in them more susceptibility towards higher risk of different cancer development like breast, prostate and colorectal cancer [37]. However, the growth factor IGF-II which shares locus with IncH19 (IGF-II/H19) forms an imprinted gene. This silencing is found disrupted in different cancers including prostate cancer. The association of adipose tissue and obesity is a known risk factor for both T2DM and prostate cancer by disturbing cellular environments. As a result, hyperglycaemia or inflammatory metabolic situations are hypothesized to be the cause of this loss of imprinting (LOI) [38]. Differentially expressed IncRNA (LINC01128) is already known to
increase the rate of cervical cancer progression and is also predicted as a biomarker of gestational hypertension [30, 40]. Similarly, Pradeep Tiwari et al, in 2019 suggested that LINC01128 could serve as a biomarker for diabetes diagnosis and prognosis (Figure 1). Metformin, an antidiabetic drug from several studies, proved that not only effects on glucose metabolism but also shows interactions with androgen receptors (AR). It plays a role in stabilizing PSA levels [41]. In certain therapy, another commonly used method for T2DM, reported that Glucagon-like peptide-1 receptor expression plays an anti-prostate cancer effect. It is helping in attenuating cell cycle progression. So, its forceful activation to express can be a potential therapeutic approach [42]. Therefore, both metformin and certain therapies help in blocking cell cycle progression by reducing mTOR activity [43]. Hypogonadism (decrease in level of testosterone) is also found associated with both diabetes and PCa. Fall in its level serum is capable of causing high graded PCa. Hence diabetes (T2DM) is suggested to be a crucial predictor of high graded PCa especially with BPH [44]. For early possible detection prostate specific antigens (PSA) levels are broadly used, but its concentration shows variation due to several other comorbidities, age and lifestyle which make it to demand more précised analysis of test results. Based on a linear aggression analysis, fall in PSA of those patients who are taking antidiabetics and by hemodilution in obese people. This establishes an inverse relationship between diabetes obesity and PSA level. Such study suggests to deliberately checking the PSA level especially the diabetic and obese patients [45]. Both PCa and DM incidence is rising parallel with age. Despite the fact diabetes mellitus reduces the risk of PCa; DM can also increase its mortality [45]. Understanding of association between DM and PCa is still insufficient. Moreover, obesity makes its pathophysiology a more complex situation [46].

LINC01128:
In a study, GEO datasets of osteosarcoma (OS) were analysed for LINC01128 expression to clear its oncogenic role. It revealed that its increased expression in OS patients is accompanied with their shorter survival. However, its knockdown turned down the
proliferation, migration and invasions. In OS LINC01128 is identified working as a sponge in triggering Wnt/β-Catenin signaling by promoting MMP2 expression through miR-299-3p \cite{55,56}. In promoting cervical cancer development again, it functions as a sponge for miR-383-5p \cite{49}. In cervical cancer tissues expression of LINC01128 is found significantly high and its fall suggests that it might lower the SFN (stratifin) at both mRNA and protein level. SFN, a known potential biomarker in cervical cancer and is also majorly expressed in the early stage of lung adenocarcinomas. It clearly explains how LINC01128 could accelerate cell processes like cell proliferation, its migration, and invasion and even can inhibit the apoptosis through SFN up regulation and its release by binding miR-383-5p and also working as its antagonist \cite{50,51}. miR-383 is under regulation of LINC01128. However, overexpression of miR-383 in T2DM serum reverses the cell apoptosis under high glucose in mice β cells by TLR4 and APOC3 suppression \cite{52}. Also, LINC01128 is seen high in Stage III-IV of CRC and mediates there a PRMT5 function which is a mediator of methylation of proteins \cite{53}. In pancreatic cancer it is found as EMT-LPS (epithelial mesenchymal transition related IncRNA prognostic signature) molecule \cite{54}.

**CONCLUSION**

Conclusions and outlook

The phenotype-interactome networks have been a powerful approach to understand and characterize networks. There is a greater scope of relevance underlying the pathophysiology of above mentioned

1. To fully comprehend the importance of phenotype interactome networks and diabetes associated metabolism, it is vital to ensure there is a healthy diet regimen followed which also addresses the clinical implications of its absorption, bioavailability, and human health benefits. Integrated systems approaches can be used to discover the novel genes and pathways with an emphasis on the molecular physiological insights gained
through systems/nutrigenomic modules and thereby candidate differentially expressed genes (DEGs) could be detected.

2. Standard operating procedures (SOP), recommendations and guidelines in consideration of the aforementioned Diabetes phenotypes for better dissemination of phenome interactome predictions will help avoid the risk of over/under treatment.

3. Post next generation sequencing, a large focus nowadays should be on the development of NGS/genotyping panels which can set a precedent for a global consortium effort, community health bridging the gap between the nutritional deficiency diseases and Diabetes

ACKNOWLEDGEMENTS
The authors gratefully acknowledge Arvinpreet Kaur, Mehak Chopra, Berenice, Kiran Telukunta, Anshu Bharadwaj, Harpreet Singh and Purnima Sharma for subtle scientific deliberations.
<table>
<thead>
<tr>
<th></th>
<th>Primary Source</th>
<th>Originality</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><a href="http://www.ncbi.nlm.nih.gov">www.ncbi.nlm.nih.gov</a></td>
<td>86 words — 3%</td>
</tr>
<tr>
<td>2</td>
<td>Shalini Rajagopal, Ayam Gupta, Rabia Parveen, Nidhi Shukla et al. &quot;Vitamin K in human health and metabolism: A nutri-genomics review&quot;, Trends in Food Science &amp; Technology, 2022</td>
<td>36 words — 1%</td>
</tr>
<tr>
<td>3</td>
<td><a href="http://www.hindawi.com">www.hindawi.com</a></td>
<td>14 words — 1%</td>
</tr>
<tr>
<td>4</td>
<td><a href="http://www.wjgnet.com">www.wjgnet.com</a></td>
<td>13 words — &lt; 1%</td>
</tr>
</tbody>
</table>