

Review Article

A Critical Review of Proteomic Studies in Gestational Diabetes Mellitus

Tao Zhou ¹, Lu Huang,² Min Wang ³, Daozhen Chen ¹, Zhong Chen ²,
and Shi-Wen Jiang^{1,3}

¹Research Institute for Reproductive Medicine and Genetic Diseases, The Affiliated Wuxi Maternity and Child Health Care Hospital of Nanjing Medical University, Wuxi 214002, China

²Department of Obstetrics, The Affiliated Wuxi Maternity and Child Health Care Hospital of Nanjing Medical University, Wuxi 214002, China

³Centre for Reproductive Medicine, The Affiliated Wuxi Maternity and Child Health Care Hospital of Nanjing Medical University, Wuxi 214002, China

Correspondence should be addressed to Tao Zhou; zhoutao@njmu.edu.cn and Zhong Chen; wxcz0812@163.com

Received 27 February 2020; Revised 18 June 2020; Accepted 30 June 2020; Published 16 July 2020

Academic Editor: Bernd Stratmann

Copyright © 2020 Tao Zhou et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Gestational diabetes mellitus is a progressive and complex pregnancy complication, which threatens both maternal and fetal health. It is urgent to screen for specific biomarkers for early diagnosis and precise treatment, as well as to identify key moleculars to better understand the pathogenic mechanisms. In the present review, we comprehensively summarized recent studies of gestational diabetes using mass spectrometry-based proteomic technologies. Focused on the entire experimental design and proteomic results, we showed that these studies have covered a broad range of research contents in terms of sampling time, sample types, and outcome associations. Although most of the studies only stayed in the stage of initial discovery, several proteins were further verified to be efficient for disease diagnosis. Functional analysis of all the combined significant proteins also showed that a small number of proteins are known to be involved in the regulation of insulin or indirect signaling pathways. However, many factors such as diagnostic criteria, sample processing, proteomic method, and statistical method can greatly affect the identification of reproducible and reliable protein candidates. Thus, we further provided constructive suggestions and recommendations for carrying out proteomic or follow-up studies of gestational diabetes or other pregnancy complications in the future.

1. Introduction

Gestational diabetes mellitus (GDM) is currently defined as a separate subcategory of diabetes, which is only accompanied by pregnancy [1, 2]. It is known that the high levels of hormones, such as progesterone and placental lactogen, may promote insulin resistance and result in hyperglycemia during pregnancy [3]. There are many pathogenic risk factors for GDM, including high body mass index (BMI), advanced maternal age, and family history of diabetes. Partially due to the boom of elderly parturient women and the transition of modern lifestyle, the prevalence of GDM was reported to be increasing over the past decades [4, 5]. Large cohort studies indicated that GDM is associated with

various pregnancy and delivery complications such as pre-eclampsia (PE), caesarean delivery, macrosomia, neonatal hypoglycemia, and large for gestational age [6–8]. Exposure to hyperglycemia during pregnancy may also result in long-term adverse effects on both postpartum women and newborns. For example, women with GDM were found to have an elevated risk of developing type 2 diabetes (T2D) and cardiovascular diseases [9, 10]. The offsprings may also be at a high risk of suffering from diabetes, obesity, hypertension, and dyslipidemia later on in their life [11]. Moreover, current treatments may have limited effects on early GDM (diagnosed in high-risk women at <24 weeks of gestation) [12]. A recent study also showed that the incidences of neonatal hypoglycemia and polyhydramnios remain high

after insulin treatment [13]. Thus, GDM has become a serious public health problem that increases health care costs during and post pregnancy [14].

There are several versions of screening and diagnostic criteria for GDM, as recommended by the International Association of the Diabetes and Pregnancy Study Groups (IADPSG) [15], the American Diabetes Association (ADA) [1], or the World Health Organization (WHO) [16]. Although the details of different criteria are frequently revised and still under debate, some common guidelines are accepted based on the Hyperglycemia and Adverse Pregnancy Outcome (HAPO) studies [17, 18]. First, GDM screening is usually suggested to be performed at 24–28 weeks of gestation (the second trimester), leaving only limited time for interventions or treatments. Second, laboratory measurements for traditional diabetes such as random plasma glucose (RPG), fasting plasma glucose (FPG), or hemoglobin A1c (HbA1c) are not directly applicable for GDM. Instead, one-step or two-step approaches based on the oral glucose tolerance test (OGTT) are widely used. However, there exists a portion of false detections and lacks precise cutoffs for predicting the outcomes using the current tests. Furthermore, most standards suggested to distinguish and diagnose overt diabetes or pre-existing pre-gestational diabetes in the early stage of pregnancy before the diagnosis of GDM. Hence, it is of vital importance to screen for specific biomarkers for early detection, differential diagnosis, outcome prognosis, and precise treatment of GDM, as well as to identify key molecular targets to better understand the basic mechanism of pathogenesis.

In clinical research, a biomarker is defined as any substance (such as RNA, proteins, or metabolites) that could be measured for the prediction of the incidence or outcome of a disease [19]. The predictive efficiency of a biomarker is evaluated by the area under the curve (AUC), as calculated based on the paired values of sensitivity and specificity for a biomarker. In general, a prediction model is considered to be acceptable when the AUC value is above 0.6 [20]. An AUC larger than 0.8 indicates a very good accuracy, while a value greater than 0.9 indicates an excellent diagnostic accuracy. As the major players of gene functions, proteins presented in various types of body fluids or tissues have been a focus for searching specific and sensitive disease biomarkers. Since GDM is believed to share some common pathogenic mechanisms with T2D, several studies have tested and evaluated a selection of protein biomarkers related to plasma glucose, insulin resistance, inflammatory pathway, or oxidative stress [21–23]. However, few proteins have shown to be efficient for the prediction of impending gestational diabetes and some results are still needed to be validated. Over the past decades, proteomic methods based on mass spectrometry (MS) are rapidly developing and have been applied in various fields of biomedical researches. Basically, MS-based proteomics can identify and quantify various proteins precisely based on the mass and intensity of fragmented peptides [24]. The development of biomarkers using a proteomic approach usually undergo three successive stages including discovery, verification, and validation, with an increasing number of enrolled subjects and a decreasing number of candidate biomarkers [25]. One of the main advantages of proteomic

approaches is that it can systematically evaluate the expression of the whole proteins of multiple samples in a single experiment. In addition, the recently developed targeted proteomic strategies, which do not require any antibodies, could replace traditional western blot or enzyme-linked immunosorbent assay (ELISA) as the new golden standard for protein verification and validation [26].

In recent years, more and more proteomic studies have been performed to construct differential expression profiling or to identify potential biomarkers for GDM and its complications. After an extensive PubMed search for proteomic studies of GDM using mass spectrometry, a total of 21 representative full-length research articles are included in this review [27–47]. The detailed information of the reviewed literatures are listed in Supplementary Data 1. Focused on the entire experimental design and proteomic results, the present review comprehensively summarizes and compares these studies. All of the datasets are also integrated to obtain an overview of functional pathways for those differentially expressed (DE) proteins between different groups. Finally, the limitations of current studies and possible improvements toward translational and precision proteomics are also discussed in depth.

2. Summary of Proteomic Studies of GDM

2.1. Experimental Designs and Objectives. In order to identify or quantify important proteins for GDM, most studies adopted a traditional disease-control design. Thus, GDM cases were usually treated as disease groups, while those cases with normal glucose tolerance (NGT) were regarded as control groups (Figure 1(a)). However, two studies are aimed at investigating adverse fetal outcomes, which selected GDM with macrosomia [33] or GDM with childhood obesity [35] as disease groups. Another study intended to identify biomarkers for one of the GDM-associated complications (early-onset PE) [40], which selected GDM as control groups. Additionally, one study is aimed at comparing GDM with NGT in obese pregnant women [43].

The research materials and sampling time further determine the research objectives. In clinical studies, the peripheral circulating blood is the most widely used specimen for searching biomarkers. Plasma or serum components can further be extracted from blood to remove cell parts or platelets. Among the ten studies using peripheral blood, five studies used plasma specimens (Figure 1(b)) [27, 32, 39, 44, 47], and the remaining studies used serum samples. Moreover, there is also a growing trend for seeking noninvasive testing strategies nowadays. Hence, one study used urine samples to identify biomarkers that could predict GDM [38].

In recent years, increasing attention is drawn to the utility of exosomes in monitoring various diseases. The concentration of circulating exosomes is found to be elevated in GDM cases [48]. Since circulating exosomes could be originated from the placenta and deliver to various types of cells, it is a novel and promising approach to elucidate the physiological mechanisms of GDM based on the proteins identified in the isolated plasma exosomes [41] or even urine exosomes [45]. It is generally accepted that the basic pathogenesis of

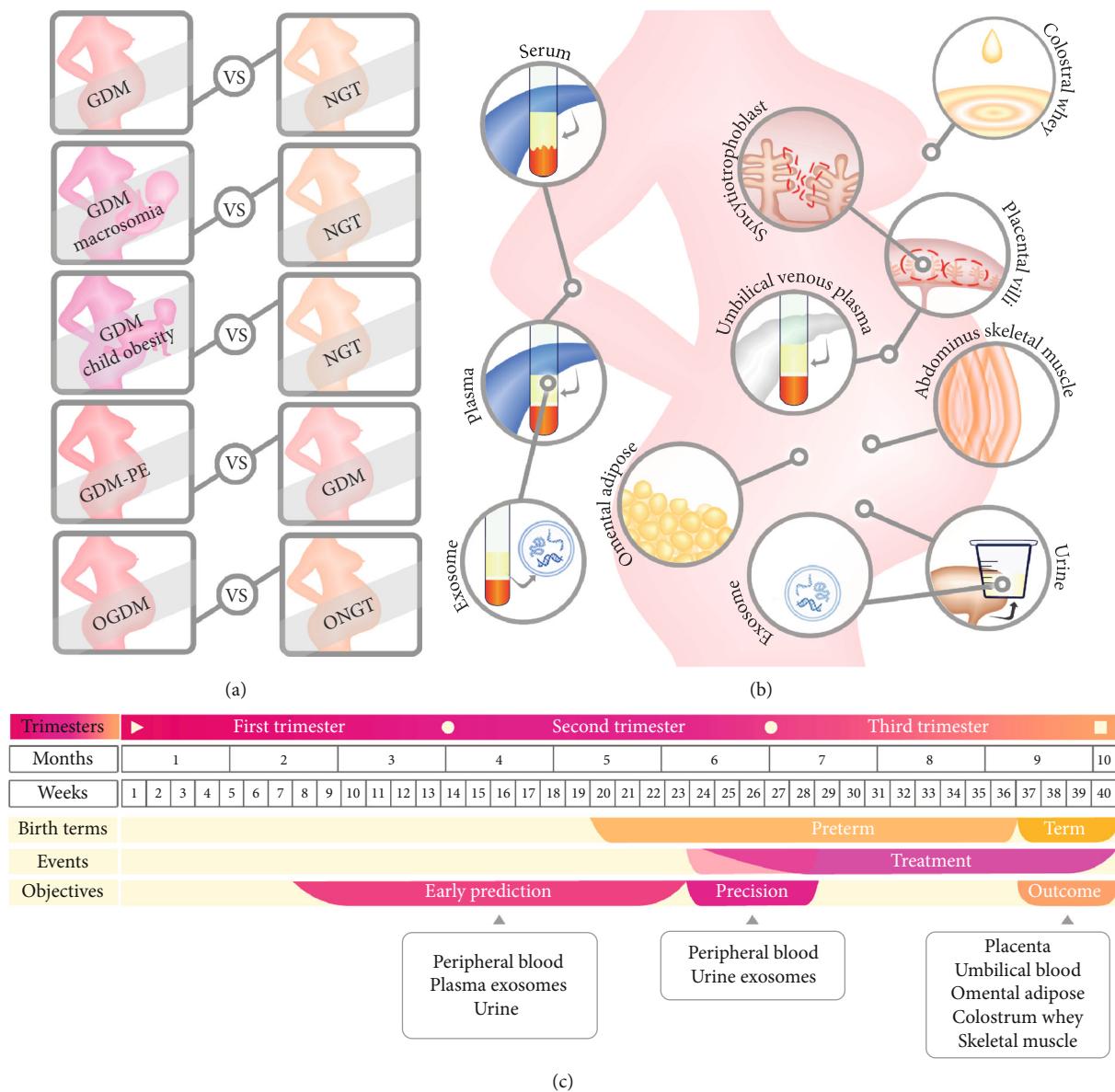


FIGURE 1: Summary of the experimental designs applied in the reviewed studies. (a) Five pairs of disease-control design for GDM and its adverse effects; (b) eleven types of samples for proteomic analyses; (c) a summarized timetable of pregnancy trimesters (the whole pregnancy period is divided into three trimesters), pregnancy time (in months or weeks), birth terms (preterm or term birth), clinical events for GDM (diagnosis time or treatment time), potential research objectives of these proteomic studies (early prediction, precision diagnosis, or outcome prognosis), and sample types.

GDM is insulin resistance. Thus, one study used omental adipose tissue, which is known to play important roles in metabolic disorders, to screen for key proteins involved in pregnancy-induced insulin resistance [34]. In addition, abdominus skeletal muscle tissue was used to investigate the metabolic consequences of GDM in obese pregnant women [43].

GDM is a complex complication affecting both maternal and fetal health. Several studies used specimens associated with fetomaternal communication, such as placental tissue (villi) [28, 36], syncytiotrophoblast [40], and umbilical venous plasma [33, 35, 39, 44]. These samples could provide key proteins for understanding the pathogenesis and progression of GDM, especially for GDM-related maternal complications and adverse fetal outcomes. In addition to the maternal-fetal

interface, breast milk delivers nutrition to breastfed infants. Thus, one study showed that colostrum whey proteins involved in immunity and nutrition are significantly changed in GDM [30].

Pregnancy is a progressive physiological process and lasts about 40 weeks. The gestational period is usually divided into three sections (trimesters). Each trimester lasts about three months or 13 weeks. As most of the hormones with diabetogenic potency were found to reach a peak around 24-28 weeks during pregnancy [49], the recent versions of guides (such as IADPSG 2010, WHO 2013, and ADA 2019) all recommended to perform screening and diagnosis of GDM at 24-28 weeks of gestation. After diagnosis, GDM patients are usually treated by dietary control or insulin intervention

to maintain their fasting plasma glucose levels. There are generally three periods of sampling time among the reviewed studies (Figure 1(c)). First, peripheral blood and urine samples were drawn before diagnosis and used to identify biomarkers for early prediction [27, 32, 37, 38, 42, 47]. Second, peripheral blood and urine exosome samples were also drawn at the same time with diagnosis and used to identify biomarkers for precision diagnosis or classification [29, 31, 41, 45]. Third, placental tissues, umbilical blood, omental adipose tissues, abdominal skeletal muscle tissue, and colostrum were obtained during or after delivery and used to identify markers for the pathogenesis of GDM or fetal outcomes [28, 30, 33–36, 39, 40, 43, 44, 46].

2.2. General Experimental Workflow. By combining all the reviewed studies, the general workflow of the proteomic-based discovery of biomarkers for GDM is shown in Figure 2. In brief, the whole workflow can be divided into three stages: initial discovery, expression verification, and final validation (for diagnosis or functional analysis). In the first stage, a proteome-wide quantitative approach was applied to identify differentially expressed proteins or peptides between GDM and the control groups. There are generally two types of proteomic quantitative strategies, which are based on gel image or spectrum intensity (label free or chemical labeling) separately. We further classified these technologies into two groups based on their throughput and identification coverage: the selective and comprehensive strategies. The classic proteomic methods such as two-dimensional gel electrophoresis (2DE), matrix-assisted laser desorption/ionization (MALDI), or surface-enhanced laser desorption/ionization (SELDI) are seen as low-throughput and selective approaches. In these approaches, only several differentially expressed spots or spectra are usually selected for the follow-up mass spectrometry analysis, which generate a selective list of protein identifications. Nowadays, more and more studies used the state-of-art proteomic technology, known as liquid chromatography-tandem mass spectrometry (LC-MS/MS), to comprehensively identify and quantify all the proteins in the initial discovery step. These approaches first generate a full list of protein identifications; then, a sublist of differentially expressed proteins are further identified based on label-free or labeling information. Although labeling methods based on isobaric tags for relative and absolute quantitation (iTRAQ) or tandem mass tag (TMT) are believed to be more reliable in accuracy, label-free-based methods could identify more proteins and cover a broad range of expression levels [50]. A previous study also showed that a label-free-based proteomic quantification approach could be used to quantify plasma biomarkers in clinic with high efficiency [51]. The traditional LC-MS/MS is running in a data-dependent acquisition (DDA) mode, which only allows the most intense ions to be analyzed. In contrast, the data-independent acquisition (DIA) mode records all the fragmented spectra of the entire isolated ions within a certain window [52]. The recently prevalent sequential window acquisition of all theoretical mass spectra (SWATH) technology, a representative strategy of DIA, is believed to be a promising tool for biomarker discovery and translational proteomics [53]. Among the reviewed studies, five of them used the selec-

tive approaches based on 2DE, MALDI, or SELDI quantification; 15 studies used the comprehensive LC-MS/MS methods based on label-free, iTRAQ, or TMT quantification; and one study applied the SWATH-MS technology for label-free quantification.

All the reviewed studies performed an initial screening of potential biomarkers or functional proteins. However, a few studies stopped in the first stage without the follow-up verification or validation analyses. The general workflow for developing clinic biomarkers usually undergoes multiple rounds of verification or validation to obtain the final list of proteins with enough diagnostic capacity. ELISA (enzyme-linked immunosorbent assay) is a classic and powerful tool for validating protein biomarkers. However, the sensitivity and specificity of ELISA largely depended on the quality of the corresponding antibodies. The development and production of antibodies take a lot of time, money, and resource and also have a high rate of failure. The recently emerging targeted proteomic technologies, such as multiple reaction monitoring (MRM) or parallel reaction monitoring (PRM), can quantify multiple proteins in a single experiment, providing a promising alternative for biomarker validation [54]. Another advantage of targeted proteomics is that it can handle protein isoforms and post-translational modifications (PTM) at the same time. Among the reviewed studies, eight of them used the ELISA method, while one recent study applied the MRM technology to validate the expression of selected candidates.

Besides identifying diagnostic biomarkers, several studies are also aimed at identifying key proteins that are involved in the pathogenesis of GDM and its outcomes. However, most of the reviewed studies only discussed the potential functional associations of the identified DE proteins theoretically or based on bioinformatics annotations. Very few studies have performed follow-up functional experiments. For example, one study performed RNA interference (RNAi) experiment on cultured cells (adipocytes) to analyze the *in vitro* functions of adipocyte plasma membrane-associated protein (APMAP), one of the downregulated proteins in GDM groups, and found that APMAP may play an important role in the impaired insulin signaling pathway [34]. In addition to clinical validation or functional verification, eight studies also performed expression verification experiments using immunodepletion, western blot, immunohistochemistry, real-time polymerase chain reaction, or bioinformatics database (the Human Protein Atlas) [55] to verify the existence, localization, and change trend of DE proteins.

2.3. Diagnostic Efficiency of Protein Biomarkers. As shown in Table 1, five studies evaluated the diagnostic efficiency of a selected list of significantly changed proteins [31, 37, 38, 42, 47]. The concentrations of target proteins were measured by the ELISA or MRM method in serum, plasma, or urine samples. One of the advantages of proteomic-driven biomarker discovery is that a list of protein candidates could be easily used to establish a model with multiple indicators. Compared to a single protein biomarker, a combination of multiple protein biomarkers can greatly improve diagnostic accuracy. For example, the AUC was increased from 0.81 to 0.85 combining two glycosylated proteins (fibronectin-

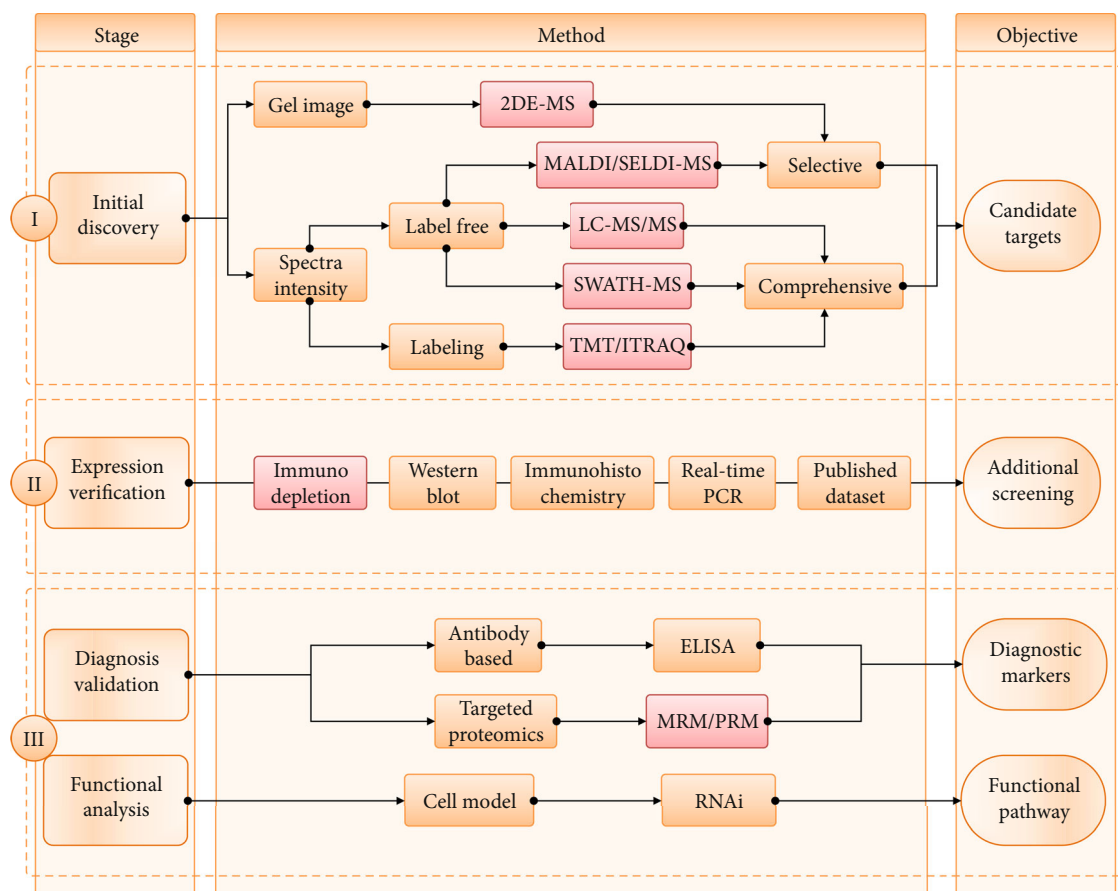


FIGURE 2: Schematic overview of proteomic-driven analyses of GDM. A summarized overview of stages, methods, and objectives for proteomic-driven analyses of GDM. Method boxes with a red border indicate the involvement of mass spectrometry.

TABLE 1: Summary of diagnostic models for GDM.

Cite No.	Sample size (GDM vs. NGT)	Sample type	Sampling time (weeks)	Markers in a model	AUC
31	15 vs. 14	Serum	24-28	(1) Fibronectin-SNA and PSG-AAL	0.85
				(2) CRP, adiponectin, SHBG, ratio of hCG to placental lactogen, fibronectin-SNA, and PSG-AAL	0.97
37	20 vs. 20	Serum	12-16	APOE, F9, FGA, and IGFBP5	0.985
38	40 vs. 40	Urine	15-20	CD59 and IL1RA	0.906
42	105 vs. 105	Serum	8-13	(1) Vitronectin	0.625
				(2) Maternal age, history of diabetes, and vitronectin	0.806
47	25 vs. 25	Plasma	11-13	(1) TSP-4	0.94
				(2) CNDP1	0.98

SNA: Sambucus nigra lectin; PSG: pregnancy-specific glycoprotein; AAL: Aleuria aurantia lectin; CRP: C-reactive protein; SHBG: sex-hormone-binding globulin; hCG: human chorionic gonadotropin; APOE: apolipoprotein E; F9: coagulation factor IX; FGA: fibrinogen alpha chain; IGFBP5: insulin-like growth factor-binding protein 5; CD59: CD59 glycoprotein; IL1RA: interleukin-1 receptor antagonist protein; TSP-4: thrombospondin-4; CNDP1: beta-alanin dipeptidase.

SNA and PSG-AAL) [31]. And an AUC of 0.97 can be reached further combining several biochemical indicators (CRP, adiponectin, SHBG, and ratio of hCG to placental lactogen) associated with GDM. Molecular biomarkers could also be used in conjunction with clinical characteristics to improve diagnosis. For example, vitronectin was validated to be signif-

icantly changed in GDM [42]. And the AUC was increased from 0.625 to 0.806 when vitronectin was used together with maternal age and history of diabetes.

It is also exciting to see several protein biomarkers that were identified in the early stages of pregnancy before GDM diagnosis. These markers could be used for early

prediction of GDM, leaving enough time for clinical interventions to prevent later adverse outcomes. Currently, most biomarkers were detected in serum or plasma. However, CD59 and IL1RA were tested using urine samples, providing a noninvasive way of diagnosis. As most biomarkers were evaluated using a small sample size, further rounds of expression validation and statistical assessment are required before these biomarkers are being applied in clinic.

2.4. Functional Annotation of Significant Proteins. In addition to identify potential biomarkers, the DE proteins could also help us to better understand the molecular mechanisms associated with the pathogenesis of GDM and its outcomes. To obtain a functional overview, we integrated and classified all the DE or validated proteins according to comparison design and sample type (Supplementary Data 2). The ToppGene suite was used for functional annotation based on Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways [56]. First, we systematically searched for known markers related to insulin functions. Among the 417 integrated genes, 45 of them are found to be directly associated with insulin secretion, binding, response, regulation, resistance, or other signaling pathways (Figure 3(a)). We also found that 38 of these genes are highly associated with each other using the STRING database [57] (Figure 3(b)).

As mentioned above, the reviewed studies could be firstly divided into five classes based on the research object: GDM versus NGT, GDM versus GDM with PE (maternal complication), OGDM versus ONGT, GDM with macrosomia versus NGT (fetal outcome), and GDM with childhood obesity versus NGT (long-term fetal effect). And eleven types of samples were used in these studies: peripheral plasma/serum, peripheral plasma exosomes, placental villi, syncytiotrophoblast, umbilical venous plasma, urine, urine exosomes, omental adipose tissues, abdominal skeletal muscle tissues, and colostrum whey. We divided the dataset into 12 groups with a combination of experimental comparison and sample type. We thus performed functional enrichment analysis to identify overrepresented KEGG pathways in different groups, which may provide novel insights into the complex pathogenesis of GDM. A P value of 0.05 was used as a cutoff for statistical significance. Only about 11.5% of the integrated proteins were overlapped with each other among the total 12 groups. However, it is interesting to see many of the top enriched pathways are presented in more than one grouped class (Figure 3(c); Supplementary Data 3). Although the DE proteins are largely different among various samples or groups, the same enriched pathways indicated that common pathogenic mechanisms may be identified. Several pathways including complement and coagulation cascades, platelet activation, ECM-receptor interaction, PI3K-Akt signaling pathway, and PPAR signaling pathway, which are already known to be associated with insulin resistance or T2D [58–62], are still significantly enriched after the applying of false discovery rate (FDR) correction. The samples of these studies were obtained at different gestational ages. Focused on the DE proteins from peripheral plasma or serum, we also compared the enriched pathways identified in different sampling times.

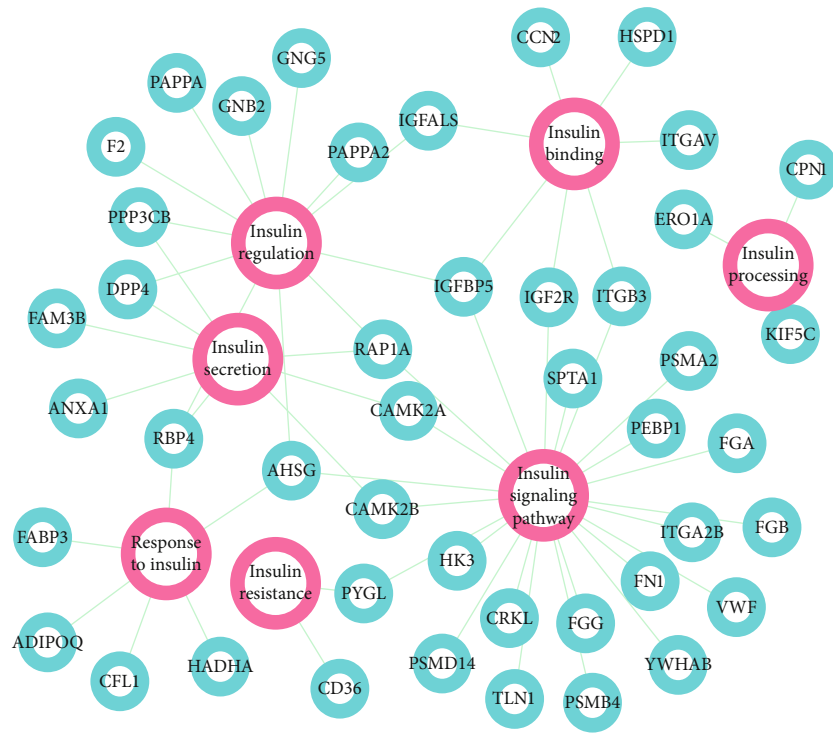
As expected, most of the enriched pathways in the combined peripheral blood proteins are repeatedly identified in the subclasses (Supplementary Data 4). And the top enriched pathways in different periods are slightly different. GDM is known to be a progressive disease; the differential distributions of the enriched pathways in different gestational weeks may reflect the corresponding stages of GDM. However, it still needs more experiments and data to prove and organize the results.

3. Bottlenecks and Limitations of Current Studies

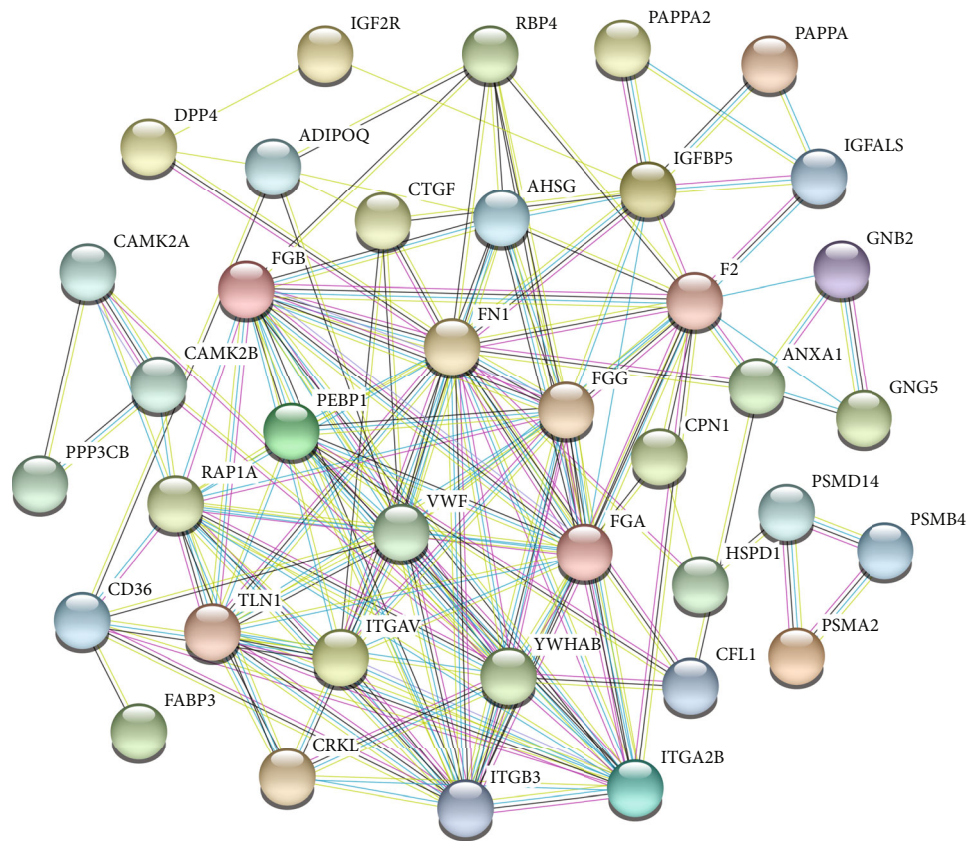
3.1. Variation of Proteomic Methods and Experimental Design. As described above, although many studies focused on the differences between GDM and NGT, few proteins are found to be overlapped. In addition, from DE proteins to potential biomarkers, there are also a low proportion of targets that could pass the validation testing. For example, one study identified 25 DE proteins in serum, while six of them can be used to develop a testing assay and only one protein was proved to be effective in the final predictive model [42]. There are many factors that may contribute to the high rates of variations and low rates of validations. One major factor that may result in variations is the difference of proteomic methods. One study compared the gel-based and gel-free methods using the same placenta samples and found that only 5 proteins were overlapped in a total of 42 DE proteins [36]. Since various proteomic methods such as 2DE, MALDI, SELDI, LC-MS/MS, and SWATH were used to discover significant proteins in the reviewed studies, there may be random biases in the detection of DE proteins.

Actually, because the detailed experimental design and procedure vary in different researches, most studies are not directly comparable. However, currently, there are many limitations and inconsistencies in the reviewed studies, which obviously reduced the reproducibility and reliability of the results. First, although most guides (including ADA and WHO) recommended to use the new consensus version of criterion (IADPSG 2010) [15] for GDM diagnosis, a few guides (including ACOG and ADA) also supported to use alternative criteria such as Carpenter-Coustan 1982 [63] and NDDG 1979 [64]. Thus, various versions of diagnostic criteria were used in these studies. Some of the key parameters, such as the application of an additional test, dosage of a glucose load, thresholds, and number of abnormal indicators for diagnosis, are different among these criteria (Table 2), which causes the pathological backgrounds of GDM groups to be different. Moreover, there are currently no specific standards to further classify GDM into subgroups. However, there is an increasing need to classify GDM patients into different degrees of severity or to identify patients for different treatments. Only one reviewed study divided GDM patients into two groups according to the glucose concentrations of 2 h OGTT [29]. And the GDM groups of many studies contain patients treated with dietary control or insulin, which may increase sample heterogeneity and decrease the power to identify reliable DE proteins.

In addition, other factors, such as the exclusion criteria, treatment strategies, and clinical characteristics (both maternal and fetal), also affect the constitutive portions of enrolled



(a)



(b)

FIGURE 3: Continued.

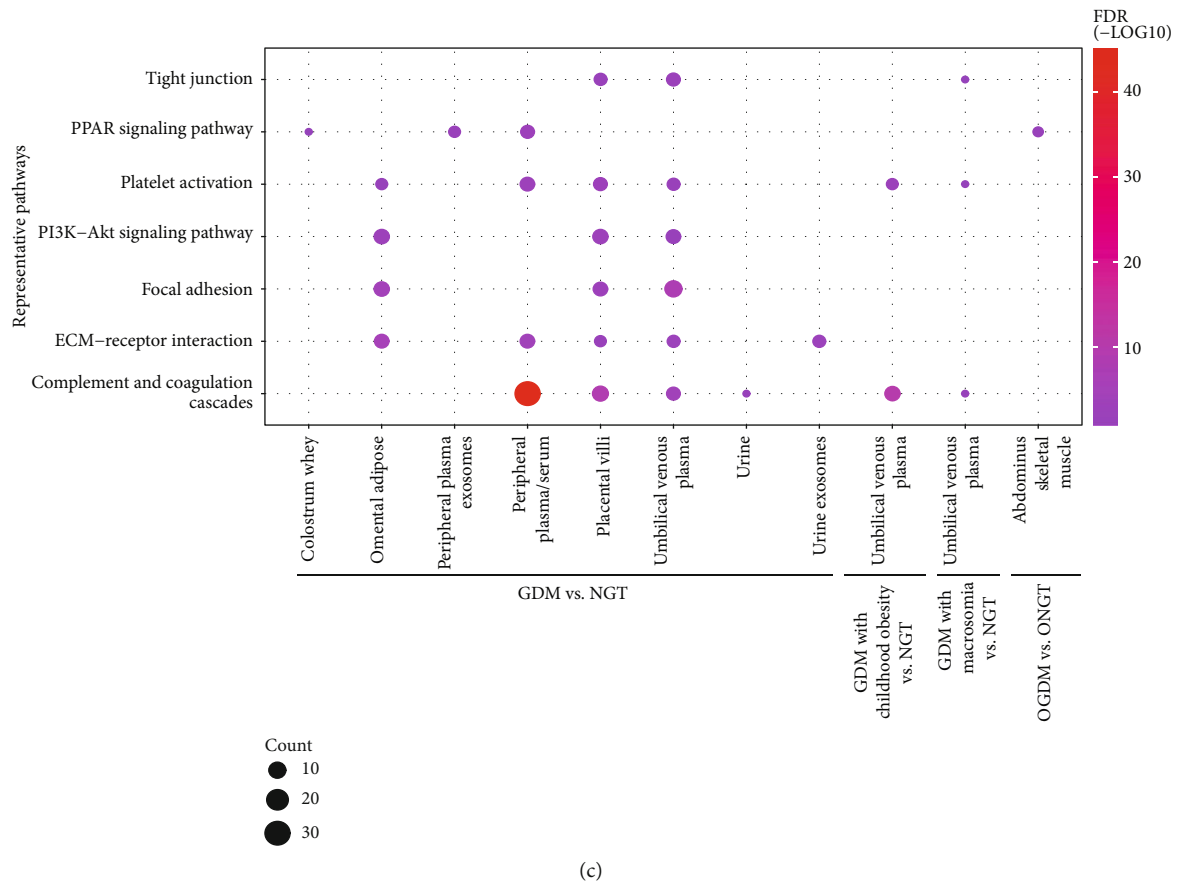


FIGURE 3: Functional annotation of the integrated differentially expressed proteins. (a) Relation network of differentially expressed proteins and insulin functions; (b) protein-protein relation network of insulin-associated DE genes; (c) representative enriched pathways for different groups of DE genes. Circle size is proportional to the number of gene count, while graduated color indicates the corresponding value of FDR.

subjects. Many risk factors (including age, geography, history of diabetes, and BMI) are known to be associated with GDM [65]. Thus, it is also important to consider matching-subjects for these factors between disease and normal groups. However, clinical characteristics such as gestational age, BMI, or fetal birth weight were found to be significantly different between GDM and NGT in several reviewed studies, which make it ambiguous to interpret the results. In addition, only one of the reviewed studies clearly described the ethnicity of the enrolled pregnant women. However, both maternal and paternal races were found to be associated with different rates of GDM [66]. Asians and Hispanics were found to have higher rates of GDM compared to whites and African Americans. With the increase of global migration and mixed race [67], it is suggested to adjust the race factor for multiracial samples.

3.2. Sample Size and Sample Processing Issues. When comparing different studies using the same method and sample type, there also exist great differences. For example, only two DE proteins in the placenta were the same in two studies using similar gel-based proteomic methods [28, 36]. This may be due to complex factors including but not limited to sample size, sample processing, and statistical method.

Sample size is an important issue that is easy to be overlooked. A small number of samples may lead to inaccurate results, while too many samples may also increase costs and waste of resources. Thus, the minimal sample size is usually required to be strictly calculated in clinical case-control and diagnostic test studies [68]. However, the considerations and formula for calculations are complex for most clinicians. Few studies actually applied a calculation of sample size according to a review of clinical diagnostic studies [69].

Sample processing can greatly affect the reproducibility of proteomic results in different laboratories. For example, placental tissue is a complex tissue with several heterogeneous cell types including cytotrophoblasts, syncytiotrophoblasts, mesenchymal cells, and fetal vascular cells. Although two reviewed studies claimed that the placental tissue was extensively washed to remove maternal and fetal blood, there are many more factors, such as delivery time, placental shape, placental weight, sampling region, sampling method, and storage condition, which are also needed to be carefully considered according to a proposed standard for placental sample collection [70]. Two reviewed studies used a peptidomic approach to identify endogenous peptides presented in peripheral serum or umbilical venous plasma. However, two different methods were used to enrich low-molecular-weight peptides: the weak cation exchange (WCX) magnetic bead kit

TABLE 2: Comparison of representative criteria for GDM diagnosis.

Criteria	First step	Glucose load	FPG	Thresholds (mmol/L)			Number of indicators
				1 h	2 h	3 h	
IADPSG 2010	NA	75 g	5.1	10	8.5	NA	≥1
WHO 1999	NA	75 g	7.0	NA	7.8	NA	≥1
ADIPS 1998	NA	75 g	5.5	NA	8	NA	≥1
C-C 1982	50 g GLT	100 g	5.3	10	8.6	7.8	≥2
NDDG 1979	50 g GLT	100 g	5.8	10.6	9.2	8.0	≥2

C-C: Carpenter-Coustan; NDDG: National Diabetes Data Group; ADIPS: Australian Diabetes in Pregnancy Society; NA: none; FPG: fasting plasma glucose; GLT: glucose load test.

and the molecular weight cutoff (MWCO) filter. The WCX magnetic bead purifies peptides based on physicochemical features, while the MWCO filter removes large proteins according to the molecular weight. Thus, the two methods may have different biases for peptide enrichment.

3.3. The Bottleneck of Statistical Methods. Finally, the statistical methods may be the bottleneck to identify bona fide targets. Most proteomic studies used the traditional *t* test or rank test methods to detect DE proteins [71]. However, these methods are usually only applicable to the data with a certain distribution and require large samples, which may be inefficient to dig out significant proteins in proteomic data. In addition, an empirical *P* value of 0.05 was usually used as a cutoff for statistical significance. For multiple comparisons of large dataset such as a microarray result, the calculation of adjusted *P* values was proposed to control the false discovery rate (FDR) [72]. However, due to the native features of proteomic technologies, applying this procedure may result in no significant proteins at all. For example, one study identifies 25 DE proteins based on the raw *P* values, while none of these proteins remained significant after the application of FDR corrections [42]. *P* values are usually used together with fold change to obtain more reliable results. It is known that the expression changes are underestimated due to the internal interferences of near isobaric ions in proteomic studies based on isobaric labeling [73]. Thus, low threshold values of 1.2 or 1.5 for fold change are also widely used in many studies.

4. Considerations and Recommendations for Future Studies

4.1. New Research Directions for Future Studies. The main application of proteomics in medicine is to develop biomarkers. The reviewed studies have already covered a broad range of research contents in terms of sampling time, sample types, and outcomes. As GDM is a progressive and complex complication, there are still some novel topics which are needed to be investigated in the future. First, a recent study established a pre-pregnancy model and found that a panel of four biomarkers, which was tested about 7 years before pregnancy, could be used to predict future GDM risks [74]. The result also indicated that GDM or prediabetes may already be developing long before pregnancy. It is known that interventions in early pregnancy have limited effects in pre-

venting GDM. Thus, it will be of great significance to identify potential protein biomarkers before conception. Second, the glucose levels of some GDM patients could be controlled by diet and healthy lifestyle, while other patients need insulin or metformin therapy. It is clear that GDM could be further classified into subgroups. However, there is a lack of efficient methods and common standards in this field; protein biomarkers could also be developed to solve this problem. Third, only a few maternal complications or fetal outcomes have been investigated in these studies. However, many conditions and complications, such as overweight, obese, hypertension, and hyperlipidemia, are known to be associated with metabolic disorders. It is also interesting to explore the cross-talk between GDM and these conditions at the protein level. Finally, many studies also found that the treatment of GDM patients did not significantly reduce some of the maternal and fetal outcomes such as subsequent diabetes, metabolic syndrome, perinatal death, neonatal hypoglycemia, and childhood obesity [13, 75–77]. Thus, it is necessary to develop prognostic markers for postpartum and long-term outcomes in treated patients.

It is also needed to pay more attention to some new directions about the types of biomarkers. First, one of the reviewed studies found that only the specific glycosylated forms of PSG and fibronectin are differentially expressed between GDM and NGT samples [31]. Protein phosphorylation and acetylation are also known to play important roles in regulating the signaling pathways of insulin [78, 79]. Thus, it may be more efficient to search for disease markers in proteins with various postmodifications. Instead of identifying protein biomarkers, two reviewed studies focused on active peptides which are endogenously cleaved [29, 33]. Recently, more and more evidences showed that small open frames are hidden in traditional transcripts of mRNA, lncRNA, microRNA, and circRNA and could encode functional polypeptides [80–82]. Short and low-abundance peptides are usually difficult to identify using the shotgun proteomic strategy. Thus, it is anticipated that more peptidomic studies will be performed to identify known or novel functional peptides. Moreover, MS technology could also help to clarify the microheterogeneity issue for complex components. For example, a protein complex, which includes transthyretin (TTR), serum retinol-binding protein (RBP4), and retinol (ROH), was presented in peripheral blood and known to be associated with the development of insulin resistance [83]. The concentrations, expression ratios, and

different types of isoforms (caused by various posttranslational modifications, amino acid change, and sequence truncation) are systematically analyzed to investigate the relationship between first-trimester maternal serum levels of the TTR-RBP4-ROH complex and the development of GDM. Specific ratios or variants are determined to be potential biomarkers for different types of GDM.

4.2. The Importance of Follow-Up Functional Studies. Besides identifying potential biomarkers for clinical usage, discovery proteomics also provided potential proteins that could explain the molecular mechanisms associated with the development of GDM and its outcomes. However, there are still two limitations for the initial discovery proteomics. First, only a list of protein names is generated; it is not known how these proteins interact with each other and what functional pathways are they involved in. Second, the DE proteins are prioritized based on statistical inferences. It is also not sure whether they have bona fide functions in the development of disease. Bioinformatics can partially solve the first problem, as it has been widely used in omic studies to translate the gene list into known functional categories [84, 85]. However, to clearly clarify these issues, functional proteomics and experimental studies are needed to verify the functions or explore the regulation relations of the identified DE proteins.

To avoid harming human bodies, various animal models and cell lines are used to mimic human diseases and to study the functions of individual genes. In addition, human xenograft and animal allograft are also widely used as preclinical models to study disease process and treatment effects [86]. Animal models for T2D were successfully developed with different levels of insulin resistance and beta cell dysfunction. It is also easy to establish rat models with gestational diabetes. However, human GDM is associated with complex maternal and fetal outcomes. Thus, it is still hard to fully reproduce the adverse outcomes in rat models [87]. Nevertheless, cell lines and animal models will promote to translate a protein list identified by proteomics into organized knowledge for the development of GDM.

5. Conclusions

In summary, GDM is a progressive and long-lasting pregnancy complication, it also has complex effects on both maternal and fetal health. The adoption of the new IADPSG guideline has greatly increased the number of diagnosis as well as the overall cost of health care. Thus, it is urgent to develop diagnostic biomarkers or to study the pathogenic mechanisms for the development of GDM and its outcomes. Over the decades, MS-based quantitative proteomic strategies have been shown to be efficient in discovering clinical biomarkers or pathogenic targets. In the present review, we systematically summarized the proteomic studies of GDM. Generally, the reviewed studies usually began with an initial discovery proteomic analysis to generate a list of significant proteins. Then, the protein list could be further used to identify potential biomarkers for diagnosis or key proteins for understanding the pathogenesis. For biomarker discovery, several proteins were shown to be effective in the diagnosis

of GDM as measured by the AUC. Functional annotation of all the integrated significant proteins also showed that a small number of proteins are known to be involved in the regulation of insulin. KEGG enrichment analysis indicated that the protein list could also be used to identify novel signaling pathways indirectly associated with the development of GDM. Thus, these studies provided a valuable resource for the study of GDM.

Apart from these achievements, there are also some limitations to these proteomic studies. First, the reviewed studies are largely different in many aspects of the experimental design and procedure. Many factors such as diagnostic criteria, sample processing, proteomic method, and statistical method can greatly affect the identification of reproducible and reliable protein candidates. Thus, we suggest to adopt the same diagnostic criteria and standardized procedure to reduce variations in and between different studies, as well as to improve the identification of verifiable markers. Second, most of these studies stopped in the stage of initial discovery or performed simple experiments to verify the expression of several selected DE proteins using a small number of samples. A serial of validation or functional experiments must be performed to further evaluate the diagnostic value of biomarker candidates or to explore the molecular mechanism for the development of GDM. Then, the protein list generated in discovery proteomics can be translated into clinical or biological significances. Finally, we also pointed out some new directions for the identification of biomarkers in GDM, such as pre-pregnancy biomarkers, subgroup classification, treatment assessment, and complex biomarkers. On the other hand, we also discussed the importance of functional studies based on animal models and cell lines to investigate the molecular mechanisms involved in the pathogenesis of GDM and its outcomes. Thus, although the reviewed studies covered a broad range of research contents, there is still much work to do to verify the results or to address unexplored issues. We believe that the presented review provides constructive suggestions and recommendations for carrying out proteomic or follow-up studies of GDM or other pregnancy complications in the future.

Conflicts of Interest

The authors have declared no conflict of interest.

Acknowledgments

This study was supported by the grants from the Jiangsu Provincial Young Medical Talents Program (QNRC2016168) and the Wuxi Young Medical Talents Program (QNRC078).

Supplementary Materials

Supplement Data 1: the detailed information for the reviewed literatures. Supplement Data 2: list of the integrated differentially expressed proteins. Supplement Data 3: list of enriched pathways in different sample groups. Supplement Data 4: list of enriched pathways in different sampling time. (*Supplementary Materials*)

References

- [1] American Diabetes Association, "2. Classification and diagnosis of Diabetes: Standards of medical care in diabetes-2019," *Diabetes Care*, vol. 42, Supplement 1, pp. S13–S28, 2018.
- [2] C. C. Thomas and L. H. Philipson, "Update on diabetes classification," *The Medical Clinics of North America*, vol. 99, no. 1, pp. 1–16, 2015.
- [3] J. F. Plows, J. L. Stanley, P. N. Baker, C. M. Reynolds, and M. H. Vickers, "The pathophysiology of gestational diabetes mellitus," *International Journal of Molecular Sciences*, vol. 19, no. 11, p. 3342, 2018.
- [4] A. Ferrara, "Increasing prevalence of gestational diabetes mellitus: a public health perspective," *Diabetes Care*, vol. 30, Supplement 2, pp. S141–S146, 2007.
- [5] X. Xu, Y. Liu, D. Liu et al., "Prevalence and determinants of gestational diabetes mellitus: a cross-sectional study in China," *International Journal of Environmental Research and Public Health*, vol. 14, no. 12, p. 1532, 2017.
- [6] C. Billionnet, D. Mitanchez, A. Weill et al., "Gestational diabetes and adverse perinatal outcomes from 716,152 births in France in 2012," *Diabetologia*, vol. 60, no. 4, pp. 636–644, 2017.
- [7] E. M. Wendland, M. R. Torloni, M. Falavigna et al., "Gestational diabetes and pregnancy outcomes—a systematic review of the World Health Organization (WHO) and the International Association of Diabetes in Pregnancy Study Groups (IADPSG) diagnostic criteria," *BMC Pregnancy and Childbirth*, vol. 12, no. 1, p. 23, 2012.
- [8] D. N. Voormolen, L. de Wit, B. B. van Rijn et al., "Neonatal hypoglycemia following diet-controlled and insulin-treated gestational diabetes mellitus," *Diabetes Care*, vol. 41, no. 7, pp. 1385–1390, 2018.
- [9] L. Bellamy, J. P. Casas, A. D. Hingorani, and D. Williams, "Type 2 diabetes mellitus after gestational diabetes: a systematic review and meta-analysis," *The Lancet*, vol. 373, no. 9677, pp. 1773–1779, 2009.
- [10] C. K. Kramer, S. Campbell, and R. Retnakaran, "Gestational diabetes and the risk of cardiovascular disease in women: a systematic review and meta-analysis," *Diabetologia*, vol. 62, no. 6, pp. 905–914, 2019.
- [11] D. Mitanchez, A. Burguet, and U. Simeoni, "Infants born to mothers with gestational diabetes mellitus: mild neonatal effects, a long-term threat to global health," *The Journal of Pediatrics*, vol. 164, no. 3, pp. 445–450, 2014.
- [12] A. N. Sweeting, G. P. Ross, J. Hyett et al., "Gestational diabetes mellitus in early pregnancy: evidence for poor pregnancy outcomes despite treatment," *Diabetes Care*, vol. 39, no. 1, pp. 75–81, 2015.
- [13] D. Bogdanet, A. M. Egan, C. Reddin et al., "ATLANTIC DIP: insulin therapy for women with IADPSG-diagnosed gestational diabetes mellitus. Does it work?," *The Journal of Clinical Endocrinology and Metabolism*, vol. 102, no. 3, pp. 849–857, 2017.
- [14] A. Danyliv, P. Gillespie, C. O'Neill et al., "Short- and long-term effects of gestational diabetes mellitus on healthcare cost: a cross-sectional comparative study in the ATLANTIC DIP cohort," *Diabetic Medicine*, vol. 32, no. 4, pp. 467–476, 2015.
- [15] B. E. Metzger, S. G. Gabbe, B. Persson et al., "International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy," *Diabetes Care*, vol. 33, no. 3, pp. 676–682, 2010.
- [16] "Diagnostic criteria and classification of hyperglycaemia first detected in pregnancy: a World Health Organization Guideline," *Diabetes Research and Clinical Practice*, vol. 103, no. 3, pp. 341–363, 2014.
- [17] H. D. McIntyre, S. Colagiuri, G. Roglic, and M. Hod, "Diagnosis of GDM: a suggested consensus," *Best Practice & Research. Clinical Obstetrics & Gynaecology*, vol. 29, no. 2, pp. 194–205, 2015.
- [18] B. E. Metzger, L. P. Lowe, A. R. Dyer et al., "Hyperglycemia and adverse pregnancy outcomes," *The New England Journal of Medicine*, vol. 358, no. 19, pp. 1991–2002, 2008.
- [19] H. Quezada, A. L. Guzman-Ortiz, H. Diaz-Sanchez, R. Valle-Rios, and J. Aguirre-Hernandez, "Omics-based biomarkers: current status and potential use in the clinic," *Boletín Médico del Hospital Infantil de México*, vol. 74, no. 3, pp. 219–226, 2017.
- [20] A. M. Simundic, "Measures of diagnostic accuracy: basic definitions," *EJIFCC*, vol. 19, no. 4, pp. 203–211, 2009.
- [21] H. M. Georgiou, M. Lappas, G. M. Georgiou et al., "Screening for biomarkers predictive of gestational diabetes mellitus," *Acta Diabetologica*, vol. 45, no. 3, pp. 157–165, 2008.
- [22] I. H. Odsaeter, A. Asberg, E. Vanky et al., "Hemoglobin A1c as screening for gestational diabetes mellitus in Nordic Caucasian women," *Diabetology and Metabolic Syndrome*, vol. 8, no. 1, p. 43, 2016.
- [23] G. S. Caglar, E. D. Ozdemir, S. D. Cengiz, and S. Demirtas, "Sex-hormone-binding globulin early in pregnancy for the prediction of severe gestational diabetes mellitus and related complications," *The Journal of Obstetrics and Gynaecology Research*, vol. 38, no. 11, pp. 1286–1293, 2012.
- [24] M. Mann and N. L. Kelleher, "Precision proteomics: the case for high resolution and high mass accuracy," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 47, pp. 18132–18138, 2008.
- [25] C. E. Parker and C. H. Borchers, "Mass spectrometry based biomarker discovery, verification, and validation—quality assurance and control of protein biomarker assays," *Molecular Oncology*, vol. 8, no. 4, pp. 840–858, 2014.
- [26] M. A. Gillette and S. A. Carr, "Quantitative analysis of peptides and proteins in biomedicine by targeted mass spectrometry," *Nature Methods*, vol. 10, no. 1, pp. 28–34, 2013.
- [27] S. M. Kim, J. S. Park, E. R. Norwitz et al., "Identification of proteomic biomarkers in maternal plasma in the early second trimester that predict the subsequent development of gestational diabetes," *Reproductive Sciences*, vol. 19, no. 2, pp. 202–209, 2012.
- [28] B. Liu, Y. Xu, C. Voss et al., "Altered protein expression in gestational diabetes mellitus placentas provides insight into insulin resistance and coagulation/fibrinolysis pathways," *PLoS One*, vol. 7, no. 9, article e44701, 2012.
- [29] T. Ai, F. Chen, S. Zhou et al., "Magnetic bead-based serum peptidome profiling in patients with gestational diabetes mellitus," *BioMed Research International*, vol. 2015, Article ID 586309, 8 pages, 2015.
- [30] D. Grapov, D. G. Lemay, D. Weber et al., "The human colostrum whey proteome is altered in gestational diabetes mellitus," *Journal of Proteome Research*, vol. 14, no. 1, pp. 512–520, 2015.

- [31] S. R. Nagalla, C. K. Snyder, J. E. Michaels et al., "Maternal serum biomarkers for risk assessment in gestational diabetes. A potential universal screening test to predict GDM status," *Indian Journal of Endocrinology and Metabolism*, vol. 19, no. 1, pp. 155–159, 2015.
- [32] C. Zhao, F. Wang, P. Wang, H. Ding, X. Huang, and Z. Shi, "Early second-trimester plasma protein profiling using multiplexed isobaric tandem mass tag (TMT) labeling predicts gestational diabetes mellitus," *Acta Diabetologica*, vol. 52, no. 6, pp. 1103–1112, 2015.
- [33] F. Liu, C. Zhao, L. Liu, H. Ding, R. Huo, and Z. Shi, "Peptidome profiling of umbilical cord plasma associated with gestational diabetes-induced fetal macrosomia," *Journal of Proteomics*, vol. 139, pp. 38–44, 2016.
- [34] Y. Ma, J. Gao, J. Yin et al., "Identification of a novel function of adipocyte plasma membrane-associated protein (APMAP) in gestational diabetes mellitus by proteomic analysis of omental adipose tissue," *Journal of Proteome Research*, vol. 15, no. 2, pp. 628–637, 2016.
- [35] Z. Miao, J. Wang, F. Wang, L. Liu, H. Ding, and Z. Shi, "Comparative proteomics of umbilical vein blood plasma from normal and gestational diabetes mellitus patients reveals differentially expressed proteins associated with childhood obesity," *Proteomics—Clinical Applications*, vol. 10, no. 11, pp. 1122–1131, 2016.
- [36] M. Roverso, M. Brioschi, C. Banfi et al., "A preliminary study on human placental tissue impaired by gestational diabetes: a comparison of gel-based versus gel-free proteomics approaches," *European Journal of Mass Spectrometry*, vol. 22, no. 2, pp. 71–82, 2016.
- [37] D. Zhao, L. Shen, Y. Wei et al., "Identification of candidate biomarkers for the prediction of gestational diabetes mellitus in the early stages of pregnancy using iTRAQ quantitative proteomics," *PROTEOMICS—Clinical Applications*, vol. 11, no. 7–8, 2017.
- [38] Y. Guo, Z. Han, L. Guo et al., "Identification of urinary biomarkers for the prediction of gestational diabetes mellitus in early second trimester of young gravidae based on iTRAQ quantitative proteomics," *Endocrine Journal*, vol. 65, no. 7, pp. 727–735, 2018.
- [39] Y. Liao, G. F. Xu, Y. Jiang et al., "Comparative proteomic analysis of maternal peripheral plasma and umbilical venous plasma from normal and gestational diabetes mellitus pregnancies," *Medicine*, vol. 97, no. 36, article e12232, 2018.
- [40] X. Sun, T. Qu, X. He et al., "Screening of differentially expressed proteins from syncytiotrophoblast for severe early-onset preeclampsia in women with gestational diabetes mellitus using tandem mass tag quantitative proteomics," *BMC Pregnancy Childbirth*, vol. 18, no. 1, p. 437, 2018.
- [41] N. Jayabalan, A. Lai, S. Nair et al., "Quantitative proteomics by SWATH-MS suggest an association between circulating exosomes and maternal metabolic changes in gestational diabetes mellitus," *Proteomics*, vol. 19, no. 1–2, article e1800164, 2019.
- [42] T. Ravnsborg, S. Svaneklink, L. L. T. Andersen, M. R. Larsen, D. M. Jensen, and M. Overgaard, "First-trimester proteomic profiling identifies novel predictors of gestational diabetes mellitus," *PLoS One*, vol. 14, no. 3, article e0214457, 2019.
- [43] K. E. Boyle, H. Hwang, R. C. Janssen et al., "Gestational diabetes is characterized by reduced mitochondrial protein expression and altered calcium signaling proteins in skeletal muscle," *PLoS One*, vol. 9, no. 9, article e106872, 2014.
- [44] I. Sreckovic, R. Birner-Gruenberger, C. Besenboeck et al., "Gestational diabetes mellitus modulates neonatal high-density lipoprotein composition and its functional heterogeneity," *Biochimica et Biophysica Acta*, vol. 1841, no. 11, pp. 1619–1627, 2014.
- [45] S. P. Ramachandrarao, A. A. Hamlin, L. Awdishu et al., "Proteomic analyses of urine exosomes reveal new biomarkers of diabetes in pregnancy," *Madridge Journal of Diabetes*, vol. 1, no. 1, pp. 11–22, 2016.
- [46] L. Shen, D. Zhao, Y. Chen et al., "Comparative proteomics analysis of serum proteins in gestational diabetes during early and middle stages of pregnancy," *PROTEOMICS—Clinical Applications*, vol. 13, no. 5, article e1800060, 2019.
- [47] D. Mavreli, N. Evangelinakis, N. Papantoniou, and A. Kolialexi, "Quantitative comparative proteomics reveals candidate biomarkers for the early prediction of gestational diabetes mellitus: a preliminary study," *In Vivo*, vol. 34, no. 2, pp. 517–525, 2020.
- [48] C. Salomon, K. Scholz-Romero, S. Sarker et al., "Gestational diabetes mellitus is associated with changes in the concentration and bioactivity of placenta-derived exosomes in maternal circulation across gestation," *Diabetes*, vol. 65, no. 3, pp. 598–609, 2016.
- [49] L. Jovanovic-Peterson and C. M. Peterson, "Review of gestational diabetes mellitus and low-calorie diet and physical exercise as therapy," *Diabetes/Metabolism Reviews*, vol. 12, no. 4, pp. 287–308, 1996.
- [50] L. Dayon and M. Affolter, "Progress and pitfalls of using isobaric mass tags for proteome profiling," *Expert Review of Proteomics*, vol. 17, no. 2, pp. 149–161, 2020.
- [51] P. E. Geyer, N. A. Kulak, G. Pichler, L. M. Holdt, D. Teupser, and M. Mann, "Plasma proteome profiling to assess human health and disease," *Cell Systems*, vol. 2, no. 3, pp. 185–195, 2016.
- [52] C. Ludwig, L. Gillet, G. Rosenberger, S. Amon, B. C. Collins, and R. Aebersold, "Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial," *Molecular Systems Biology*, vol. 14, no. 8, article e8126, 2018.
- [53] S. I. Anjo, C. Santa, and B. Manadas, "SWATH-MS as a tool for biomarker discovery: from basic research to clinical applications," *Proteomics*, vol. 17, no. 3–4, 2017.
- [54] H. A. Ebhardt, A. Root, C. Sander, and R. Aebersold, "Applications of targeted proteomics in systems biology and translational medicine," *Proteomics*, vol. 15, no. 18, pp. 3193–3208, 2015.
- [55] M. Uhlen, P. Oksvold, L. Fagerberg et al., "Towards a knowledge-based Human Protein Atlas," *Nature Biotechnology*, vol. 28, no. 12, pp. 1248–1250, 2010.
- [56] J. Chen, E. E. Bardes, B. J. Aronow, and A. G. Jegga, "ToppGene Suite for gene list enrichment analysis and candidate gene prioritization," *Nucleic Acids Research*, vol. 37, pp. W305–W311, 2009.
- [57] D. Szklarczyk, A. L. Gable, D. Lyon et al., "STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets," *Nucleic Acids Research*, vol. 47, no. D1, pp. D607–D613, 2019.
- [58] E. Hertle, M. M. van Greevenbroek, and C. D. Stehouwer, "Complement C3: an emerging risk factor in cardiometabolic disease," *Diabetologia*, vol. 55, no. 4, pp. 881–884, 2012.

- [59] P. Ferroni, S. Basili, A. Falco, and G. Davi, "Platelet activation in type 2 diabetes mellitus," *Journal of Thrombosis and Haemostasis*, vol. 2, no. 8, pp. 1282–1291, 2004.
- [60] A. S. Williams, L. Kang, and D. H. Wasserman, "The extracellular matrix and insulin resistance," *Trends in Endocrinology and Metabolism*, vol. 26, no. 7, pp. 357–366, 2015.
- [61] R. W. Mackenzie and B. T. Elliott, "Akt/PKB activation and insulin signaling: a novel insulin signaling pathway in the treatment of type 2 diabetes," *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy*, vol. 7, pp. 55–64, 2014.
- [62] M. Ahmadian, J. M. Suh, N. Hah et al., "PPAR γ signaling and metabolism: the good, the bad and the future," *Nature Medicine*, vol. 19, no. 5, pp. 557–566, 2013.
- [63] M. W. Carpenter and D. R. Coustan, "Criteria for screening tests for gestational diabetes," *American Journal of Obstetrics and Gynecology*, vol. 144, no. 7, pp. 768–773, 1982.
- [64] National Diabetes Data Group, "Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance," *Diabetes*, vol. 28, no. 12, pp. 1039–1057, 1979.
- [65] H. D. McIntyre, P. Catalano, C. Zhang, G. Desoye, E. R. Mathiesen, and P. Damm, "Gestational diabetes mellitus," *Nature Reviews Disease Primers*, vol. 5, no. 1, p. 47, 2019.
- [66] A. B. Caughey, Y. W. Cheng, N. E. Stotland, A. E. Washington, and G. J. Escobar, "Maternal and paternal race/ethnicity are both associated with gestational diabetes," *American Journal of Obstetrics and Gynecology*, vol. 202, no. 6, pp. 616.e1–616.e5, 2010.
- [67] U. A. Segal, "Globalization, migration, and ethnicity," *Public Health*, vol. 172, pp. 135–142, 2019.
- [68] J. Wittes, "Sample size calculations for randomized controlled trials," *Epidemiologic Reviews*, vol. 24, no. 1, pp. 39–53, 2002.
- [69] L. M. Bachmann, M. A. Puhan, G. ter Riet, and P. M. Bossuyt, "Sample sizes of studies on diagnostic accuracy: literature survey," *BMJ*, vol. 332, no. 7550, pp. 1127–1129, 2006.
- [70] G. J. Burton, N. J. Sebire, L. Myatt et al., "Optimising sample collection for placental research," *Placenta*, vol. 35, no. 1, pp. 9–22, 2014.
- [71] M. Bantscheff, M. Schirle, G. Sweetman, J. Rick, and B. Kuster, "Quantitative mass spectrometry in proteomics: a critical review," *Analytical and Bioanalytical Chemistry*, vol. 389, no. 4, pp. 1017–1031, 2007.
- [72] K. I. Kim and M. A. van de Wiel, "Effects of dependence in high-dimensional multiple testing problems," *BMC Bioinformatics*, vol. 9, no. 1, p. 114, 2008.
- [73] L. Ting, R. Rad, S. P. Gygi, and W. Haas, "MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics," *Nature Methods*, vol. 8, no. 11, pp. 937–940, 2011.
- [74] S. E. Badon, Y. Zhu, S. B. Sridhar et al., "A pre-pregnancy biomarker risk score improves prediction of future gestational diabetes," *Journal of the Endocrine Society*, vol. 2, no. 10, pp. 1158–1169, 2018.
- [75] B. M. Casey, M. M. Rice, M. B. Landon et al., "Effect of treatment of mild gestational diabetes on long-term maternal outcomes," *American Journal of Perinatology*, vol. 37, 2020.
- [76] M. B. Landon, M. M. Rice, M. W. Varner et al., "Mild gestational diabetes mellitus and long-term child health," *Diabetes Care*, vol. 38, no. 3, pp. 445–452, 2015.
- [77] M. B. Landon, C. Y. Spong, E. Thom et al., "A multicenter, randomized trial of treatment for mild gestational diabetes," *The New England Journal of Medicine*, vol. 361, no. 14, pp. 1339–1348, 2009.
- [78] Y. Zhang, F. Zhou, M. Bai et al., "The pivotal role of protein acetylation in linking glucose and fatty acid metabolism to β -cell function," *Cell Death & Disease*, vol. 10, no. 2, p. 66, 2019.
- [79] L. Golick, Y. Han, Y. Kim, and S. W. Park, "BRD7 regulates the insulin-signaling pathway by increasing phosphorylation of GSK3 β ," *Cellular and Molecular Life Sciences*, vol. 75, no. 10, pp. 1857–1869, 2018.
- [80] S. Prabakaran, M. Hemberg, R. Chauhan et al., "Quantitative profiling of peptides from RNAs classified as noncoding," *Nature Communications*, vol. 5, no. 1, p. 5429, 2014.
- [81] S. J. Andrews and J. A. Rothnagel, "Emerging evidence for functional peptides encoded by short open reading frames," *Nature Reviews. Genetics*, vol. 15, no. 3, pp. 193–204, 2014.
- [82] N. R. Pamudurti, O. Bartok, M. Jens et al., "Translation of CircRNAs," *Molecular Cell*, vol. 66, no. 1, pp. 9–21.e7, 2017.
- [83] A. Fruscalzo, A. P. Londero, L. Driul et al., "First trimester concentrations of the TTR-RBP4-retinol complex components as early markers of insulin-treated gestational diabetes mellitus," *Clinical Chemistry and Laboratory Medicine*, vol. 53, no. 10, pp. 1643–1651, 2015.
- [84] T. Zhou, Z. M. Zhou, and X. J. Guo, "Bioinformatics for spermatogenesis: annotation of male reproduction based on proteomics," *Asian Journal of Andrology*, vol. 15, no. 5, pp. 594–602, 2013.
- [85] J. Chen, B. J. Aronow, and A. G. Jegga, "Disease candidate gene identification and prioritization using protein interaction networks," *BMC Bioinformatics*, vol. 10, no. 1, p. 73, 2009.
- [86] T. Voskoglou-Nomikos, J. L. Pater, and L. Seymour, "Clinical predictive value of the in vitro cell line, human xenograft, and mouse allograft preclinical cancer models," *Clinical Cancer Research*, vol. 9, no. 11, pp. 4227–4239, 2003.
- [87] A. C. Kiss, P. H. Lima, Y. K. Sinzato et al., "Animal models for clinical and gestational diabetes: maternal and fetal outcomes," *Diabetology & Metabolic Syndrome*, vol. 1, no. 1, p. 21, 2009.