EBioMedicine 65 (2021) 103253

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.elsevier.com/locate/ebiom

PIWI-interacting RNA sequencing profiles in maternal plasma-derived exosomes reveal novel non-invasive prenatal biomarkers for the early diagnosis of nonsyndromic cleft lip and palate

Shanshan Jia[a], Qiang Zhang[a,b], Yu Wang[a,c], Yanfu Wang[a], Dan Liu[a], Yiwen He[a], Xiaowei Wei[a], Hui Gu[a], Wei Ma[a], Wenting Luo[a], Zhengwei Yuan[a,*]

[a] Key Laboratory of Health Ministry for Congenital Malformation, Shengjing Hospital, China Medical University, Shenyang, PR China
[b] Department of Pulmonary and Critical Care Medicine, Shengjing Hospital, China Medical University, Shenyang, PR China
[c] Department of Ultrasound, Shengjing Hospital, China Medical University, Shenyang, PR China

ARTICLE INFO

Article History:
Received 19 August 2020
Revised 10 January 2021
Accepted 4 February 2021
Available online xxx

Keywords:
Biomarker
Congenital heart defect
Neural tube defect
Nonsyndromic cleft lip and palate
PIWI-interacting RNA
Prenatal diagnosis

ABSTRACT

Background: Congenital malformations are common birth defects with high neonatal morbidity and mortality. It is essential to find simpler and more efficient biomarkers for early prenatal diagnosis. Therefore, we investigated PIWI-interacting RNAs (piRNAs) as potential prenatal biomarkers in plasma-derived exosomes from pregnant women carrying foetuses with congenital malformations.

Methods: Small RNA sequencing was used to screen piRNA biomarkers in plasma-derived exosomes of five pregnant women carrying foetuses with nonsyndromic cleft lip and palate (nsCLP) and five women carrying normal foetuses. Differentially expressed piRNAs were verified in 270 pregnant women, including 111 paired women carrying foetuses with congenital malformations and normal foetuses (at 24 gestational weeks), 10 paired women carrying foetuses with nsCLP and normal foetuses (at 15–19 gestational weeks), and 28 women at different stages of normal pregnancy. piRNA biomarkers were also verified in placentas, umbilical cords, fetal medial calf muscles, and lip tissues of nsCLP and normal foetuses.

Findings: We identified a biomarker panel of three pregnancy-associated exosomal piRNAs (*hsa-piR-009228, hsa-piR-016659*, and *hsa-piR-020496*) could distinguish nsCLP foetuses from normal foetuses. These three piRNAs had better diagnostic accuracy for nsCLP at the early gestational stage, at which time typical malformations were not detected upon prenatal ultrasound screening, and had diagnostic value for neural tube defects (NTDs) and congenital heart defects (CHDs).

Interpretation: Our work revealed the potential clinical applications of piRNAs for predicting nsCLP, NTDs, and CHDs.

Funding: National Key Research and Development Program, National Natural Science Foundation of China, and LiaoNing Revitalization Talents Program .

1. Introduction

Congenital malformations are serious birth defects associated with substantial morbidity and mortality, causing a significant burden on affected families and society. The exact aetiology of congenital malformations is not known but involves a combination of genetic and environmental risk factors with a global birth prevalence of 2–3% [1]. A multidisciplinary team may be needed to perform several surgeries and provide professional care for children with congenital malformations from birth until adulthood [2]. Congenital heart defects (CHDs), nonsyndromic cleft lip and palate (nsCLP), and neural tube defects (NTDs) are common congenital structural anomalies, with prevalence rates of approximately 8/1000, 1.2/1000, and 0.91/1000 worldwild, respectively [3–5]. In particular, the highest birth nsCLP prevalence rates are reported in Asian and Amerindian populations, and are often as high as 1/500 [6]. Currently, clinical strategies to diagnose congenital malformations mostly depend on foetal ultrasound images during the second trimester of pregnancy. However, these methods have many disadvantages. This testing is only performed in large prenatal diagnosis centres, and the success of ultrasound examination is profoundly limited by equipment quality, technician skill, foetal position, and maternal obesity [7]. For acoustic shadowing, the palate is often not adequately visualized, resulting in misdiagnoses of the cleft palate, and appropriate diagnoses are often not established until after birth [8]. More importantly, congenital

* Corresponding author.
E-mail address: yuanzw@hotmail.com (Z. Yuan).

https://doi.org/10.1016/j.ebiom.2021.103253

Research in context

Evidence before this study

Current clinical strategies to diagnose congenital abnormalities prenatally rely mainly on ultrasound, often performed when typical structural defects have already formed; these strategies thus miss the optimal time for intervention and treatment. Exosome-derived PIWI-interacting RNA (piRNAs) may exhibit stability in clinical tests and could be closely related to embryonic development. However, no studies have reported the application of piRNAs as non-invasive prenatal biomarkers for diagnosing congenital malformations.

Added value of this study

Pregnancy-associated *hsa-piR-009228*, *hsa-piR-016659*, and *hsa-piR-020496* in maternal plasma-derived exosomes may be promising biomarkers for the non-invasive prenatal diagnosis of foetal nonsyndromic cleft lip and palate (nsCLP) or neural tube defects (NTDs). *hsa-piR-009228* and *hsa-piR-016659* may be potential biomarkers for non-invasive prenatal diagnosis of foetal congenital heart defects (CHDs). The expression levels of the three piRNAs gradually increased during early embryonic development, and the differentially expressed piRNAs (*hsa-piR-009228* and *hsa-piR-020496*) in exosomes were significantly downregulated in nsCLP fetal lip tissues. Moreover, the expression levels of the piRNA biomarkers were no statistical difference in placentas, umbilical cords, and fetal medial calf muscle between nsCLP foetuses and normal foetuses.

Implications of all the available evidence

This is the first study analysing the expression levels of circulating piRNAs in plasma exosome samples from pregnant women carrying foetuses with congenital malformations using RNA sequencing and subsequent real-time quantitative reverse transcription polymerase chain reaction validation. Pregnancy-associated *hsa-piR-009228*, *hsa-piR-016659*, and *hsa-piR-020496* in plasma exosomes from pregnant women provided high sensitivity/specificity for diagnosing foetal congenital abnormalities. These piRNAs also acted as promising early biomarkers for nsCLP before the clefts could be screened using ultrasound examination. These findings presented in this study are of potential clinical use, and may have implications in the studies of other prenatal diseases.

malformations can only be detected by ultrasound examination after complete defect formation, at which time irreversible damage to the foetus has occurred, and intervention is not feasible. It is necessary to find simpler and more efficient screening methods that can be widely used in primary hospitals.

Liquid biopsies are considered to be the gold standard for earlier, non-invasive detection of cancer [9]. Tumour-associated single nucleotide variants from postoperative liquid biopsy analysis can be used to diagnose a relapse of lung cancer 70 days earlier than computed tomography scans [10], indicating that abnormally expressed molecular biomarkers occur earlier than structural anomalies. Diagnoses of embryonic abnormalities mainly rely on conventional screening or diagnostic methods, and there is a lack of biomarkers with early predictive value for congenital malformations [11]. Alpha-fetoprotein (AFP) levels in the maternal serum are also used for the screening of NTDs. However, elevated AFP is non specific for NTDs and also be caused by other congenital malformations or pregnancy complications [12,13]. The development of novel serum screening tests with increased specificity for NTDs and with the potential to diagnose other congenital malformations would be a welcome addition to the currently available options for prenatal diagnosis. In addition, prenatal biomarker analysis can result in early diagnosis before the formation of typical structural abnormalities, which is vital for early intervention and treatment. In our previous studies, based on serum proteomic analyses, we showed that several proteins could serve as more accurate targets for non-invasive prenatal diagnosis of NTDs and CHDs [14–16]. Studies have confirmed that circulating microRNAs (miRNAs) in the peripheral blood of pregnant women can act as prenatal screening biomarkers for pregnancy complications and embryonic abnormalities. We used serum miRNA microarray analyses to discover maternal circulating biomarkers for the prenatal diagnosis of CHDs and NTDs [17,18]. However, there is still a lack of biomarkers for early diagnosis before the formation of malformations, particularly for nsCLP. Therefore, the identification of new biomarkers that can be used for early diagnosis is essential.

PIWI-interacting RNAs (piRNAs) are another class of short noncoding regulatory RNAs with an estimated size of 26–31 nucleotides [19]. These RNAs bind to PIWI proteins, which function in the differentiation of embryonic stem cells and in early embryonic development [20,21], piRNAs regulate gene expression through an miRNA-like base complementary mechanism that is essential for embryonic development [22]. Satellite repeats regulate global gene expression in trans via piRNA-mediated gene silencing, which is essential for embryonic development; indeed, the deletion of tapiR1 leads to embryonic stagnation in *Aedes* [23]. The piRNAs in body fluids are more resistant to oxidation and degradation than miRNAs because of the distinctive 2'-O-methyl modification at their 3′-ends; [24] similar to miRNAs, these molecules can cross cell membranes [25]. Among all noncoding RNAs, piRNAs are thought to be the most abundant and diverse small noncoding RNAs; these molecules are derived from all types of genomic sequences, and more than 30,000 piRNA species have been identified in the human genome to date [26]. Compared with other noncoding RNAs, piRNAs are likely to be the best biomarkers for early diagnosis of congenital malformations. However, no studies have reported piRNAs as prenatal biomarkers for birth defects.

In this study, we included different types of populations and evaluated the overall changes in piRNA expression profiles in plasma-derived exosomes from women carrying foetuses with nsCLP using piRNA sequencing analysis and real-time quantitative reverse transcription polymerase chain reaction (qRT-PCR). The potential piRNA biomarkers were validated again in early exosome samples from pregnant women from a birth cohort before the point in time at which structure abnormalities can be detected on foetal ultrasound images. We also examined piRNA profiles in plasma-derived exosomes from women carrying foetuses with CHDs or NTDs. Moreover, we verified the relative expressions of these piRNA biomarkers in placentas, umbilical cords, fetal medial calf muscles, and fetal lip tissues from nsCLP and normal groups. Overall, our findings provided insights into the use of piRNAs as potential biomarkers of major congenital malformations.

2. Methods

2.1. Study design

The study design included six phases. All thresholds were set to evaluate the potential analytical applicability of the piRNA biomarkers (Fig. 1). The first phase was a screening phase in which 10 plasma-derived exosome samples from five pregnant women carrying nsCLP foetuses and five pregnant women carrying normal foetuses (population 1) were collected and subjected to small RNA sequencing to explore differentially expressed piRNAs. piRNAs with the absolute value of $\log_2$ fold change greater than 1 and false discovery rate (FDR) value less than 0.001 were identified as differentially

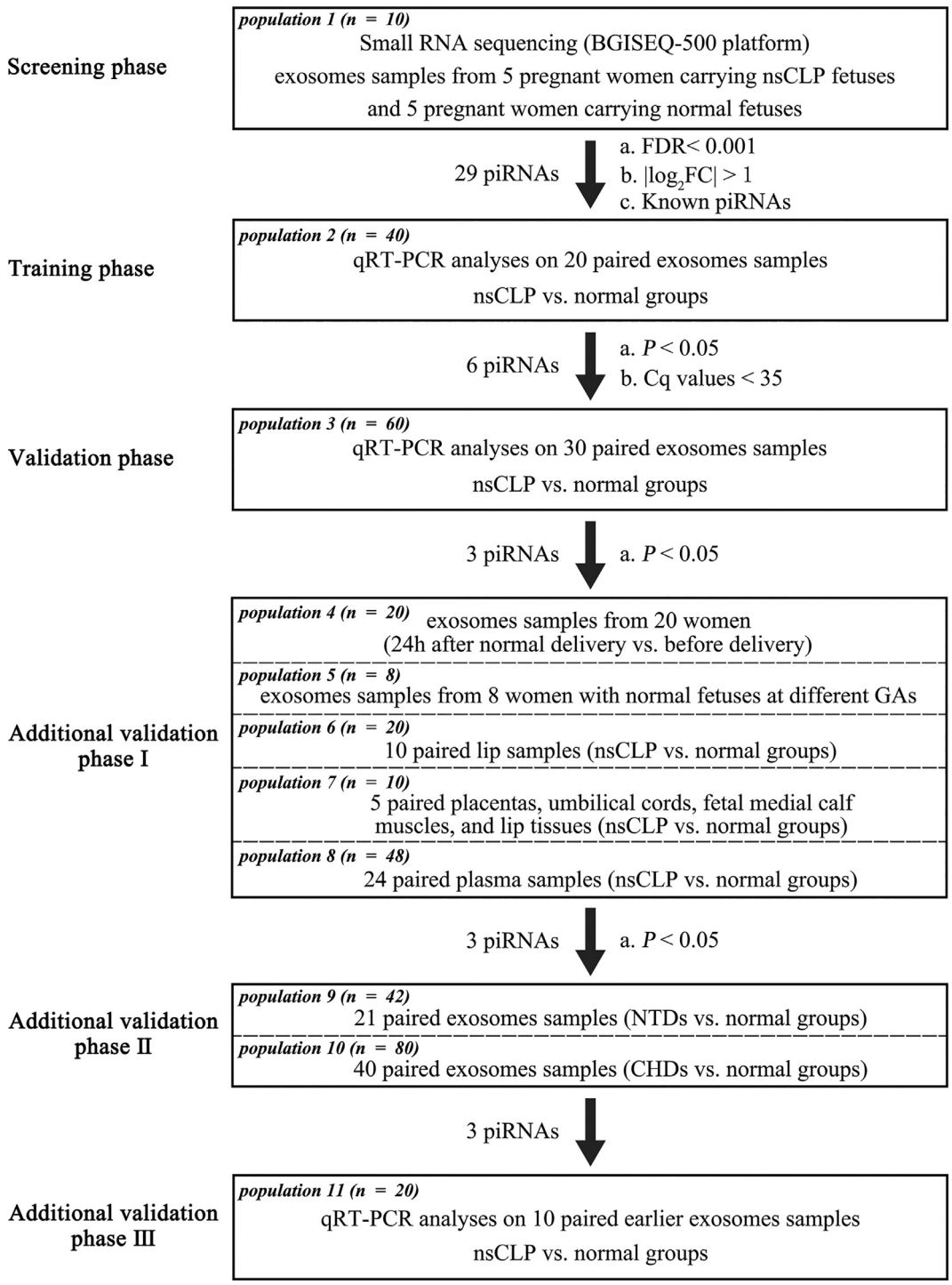

Fig. 1. Flow diagram of the study design. Participants from 11 different populations were divided into screening, training, validation, and additional validation phases of the study.

expressed piRNAs [27]. The second phase was a training phase in which the 29 differentially expressed piRNAs with known mature sequences identified in the screening phase were further tested in an independent cohort of 20 pregnant women carrying nsCLP foetuses and 20 pregnant women carrying normal foetuses (population 2) using qRT-PCR. piRNAs with quantification cycle (Cq) values of less than 35 in all samples and p values less than 0.05 were selected as potential biomarkers. The third phase was a validation phase in which the diagnostic performance of these six piRNAs was evaluated in 30 pregnant women carrying nsCLP foetuses and 30 pregnant women carrying normal foetuses (population 3). To verify whether these piRNA biomarkers were involved in the early embryonic development and acted as promising prenatal biomarkers for CHDs and NTDs, we performed analyses on the following three additional validation phases. First, in the additional validation phase I, three piRNAs with p values less than 0•05 were further analysed in 20 paired plasma exosome samples from women whose gestational ages were between 38 and 40 weeks and from women at 24 h after normal delivery (population 4). Thus, piRNAs showing significant ($p < 0.05$ by paired t-test) decreases in expression at 24 h after normal delivery compared with that before delivery were presumed to be related to pregnancy. To determine the relationships between piRNA biomarker

concentrations and pregnancy, the expression levels of piRNA biomarkers were also measured at eight consecutive gestational ages during normal pregnancy in eight pregnant women (population 5). To verify whether these screened biomarkers were associated with foetal pathological tissues, we performed qRT-PCR analyses on lip samples from 10 foetuses with nsCLP and 10 normal foetuses (population 6). piRNA biomarkers were also verified in placentas, umbilical cords, fetal medial calf muscles, and fetal lip tissues from 5 foetuses with nsCLP and 5 normal foetuses (population 7). piRNAs satisfying the above phase thresholds were confirmed as being differentially expressed in exosomes or whole plasma in paired plasma samples from 24 pregnant women carrying foetuses with nsCLP and 24 pregnant women carrying normal foetuses (population 8). Second, to verify whether these piRNAs could also act as prenatal biomarkers for other related congenital malformations, they were verified in 21 paired exosome samples from pregnant women carrying foetuses with NTDs and normal foetuses (population 9) and 40 paired exosome samples from pregnant women carrying foetuses with CHDs and normal foetuses (population 10) in the additional validation phase Ⅱ. Finally, in the additional validation phase Ⅲ, the three potential piRNAs were also verified in samples from an earlier sampling time (gestational age: 15−19 weeks), at which time the structural defects were not yet detectable by ultrasound examination in the Shengjing Birth Cohort (SJBC) study (population 11).

2.2. Study population

Fifty pregnant women carrying nsCLP foetuses and 50 pregnant women carrying normal foetuses were recruited in populations 1−3 for study of the top three phases. Twenty-one pregnant women carrying foetuses with NTDs and 21 pregnant women carrying normal foetuses as well as 40 pregnant women carrying foetuses with CHDs and 40 pregnant women carrying normal foetuses were recruited in populations 9 and 10, respectively, for analysis in the additional validation phase Ⅱ. Overall, 111 pregnant women carrying foetuses with congenital malformations were recruited after foetal ultrasound images were obtained at 22−26 weeks of pregnancy for clinical prenatal care at Shengjing Hospital of China Medical University. Additionally, 111 gestational age- and maternal age-matched pregnant women carrying normal foetuses were selected from the SJBC study. The SJBC study is an ongoing prospective cohort study of pregnant women, spouses, and their children residing in Northeast China, recruiting from April 2017. Its primary purpose is to determine the effects of environmental pollution, diet, lifestyle, and biochemical factors on the health of pregnant women and their children. During the first, second, and third trimesters of pregnancy, peripheral blood from pregnant women was collected, and questionnaires (demographic and clinical characteristics) were completed. The participants in population 4 were also selected from the SJBC study. The participants in population 5 were from a special cohort of 278 pregnant women used to explore dynamic changes in serum protein profiles during the entire pregnancy in the SJBC study. Blood samples and questionnaires were collected every 4 weeks during participation in the cohort within 7 weeks of pregnancy until about 42 days after delivery. Fifteen foetuses with nsCLP and 15 gestational age- and maternal age-matched foetuses that were inevitably aborted without an nsCLP clinical diagnosis were included in population 6 and 7 from our biobank. The pregnant women in population 8 included 24 pregnant women carrying foetuses with nsCLP and 24 pregnant women carrying normal foetuses selected from populations 1−3. In population 11, 10 pregnant women carrying foetuses with nsCLP at 15−19 weeks of pregnancy and 10 pregnant women carrying normal foetuses were selected from all 4117 participants in our SJBC study. All participants included in this study were matched for maternal and gestational ages (Table 1). Foetuses with chromosomal abnormalities, multiple malformations, or other comorbidities were excluded. Gynaecologists and foetal ultrasonologists at our hospital performed evaluations based on the codes of the International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) and ensured the accuracy of the final diagnosis. The participants in the control group were from the birth cohort and were followed up for at least 1 year after birth to ensure that their babies were healthy. Pregnant women with multiple pregnancies or pregnancy-related complications (such as abnormal gestational weight gain, gestational diabetes, gestational hypertension, preeclampsia, placental dysfunction) were excluded.

2.3. Ethics statement

Written informed consent was obtained from all pregnant women, and the study was performed in accordance with the principles of the Declaration of Helsinki and approved by the Ethics Committee of Shengjing Hospital Affiliated to China Medical University (approval no. 2017PS264K).

2.4. Sample processing and RNA extraction

The samples in the current study were collected from the biobank of SJBC. Ethylenediaminetetraacetic acid-coated tubes were used to collect 6 mL venous blood from pregnant women under fasting conditions. Venous blood samples were centrifuged at 1,600 × *g* for 10 min at 4 °C to obtain plasma. Then, plasma samples were centrifuged at 16,000 × *g* for 10 min at 4 °C to remove cell debris. The plasma samples were then transferred to RNase-free tubes and stored at −80 °C until analysis. A 600 μL aliquot of the plasma was utilized for exosome extraction. First, the plasma was filtered through a 0.22 μm filter (MILLEXGP; Millipore Express PES Membrane; Millipore, Billerica, MA, USA) into a fresh high-speed centrifuge tube; samples were then supplemented with ice-cold phosphate buffered saline (PBS) to a volume of 8 mL. After centrifugation using a Micro ultracentrifuge CS120FNX (Hitachi Koki Co., Ltd., Tokyo, Japan) at 10,000 × *g* at 4 °C for 1 h, the supernatant was moved to a new centrifuge tube and centrifuged again at 100,000 × *g* at 4 °C for 4 h. The supernatant was discarded, and exosomes were obtained by resuspension in 100 μL PBS. The collected exosomes were stored at −80 °C.

Total RNA containing piRNAs from plasma or plasma-derived exosomes was extracted using a Qiagen miRNeasy Serum/Plasma Kit (Qiagen, Valencia, CA, USA; Cat# 217184) according to the manufacturer's protocol. Lip specimens from foetuses with nsCLP were obtained from the medial edge of the cleft. The control specimens were collected from lip tissues without orofacial defects. Placentas, umbilical cords, fetal medial calf muscles, and fetal lip tissues were washed with ice-cold PBS to remove blood, divided into several pieces, and stored at −80 °C. Total RNA containing piRNAs from the above tissues was extracted using TRIzol reagent (Ambion, Foster City, CA, USA) according to the manufacturer's protocol. RNA quantity and quality were assessed using a Nanodrop ND-1000 (Thermo Fisher Scientific, Wilmington, DE, USA) and gel electrophoresis.

2.5. Exosome characteristics

According to the instructions given in Minimal Information for Studies of Extracellular Vesicles (MISEV) 2014 and their updates in MISEV2018, the yield and purity of exosomes were assessed by transmission electron microscopy (TEM), nanoparticle tracking analysis, and western blot analysis [28,29]. Twenty-microliter aliquots of exosomes suspended in PBS were applied to copper electrode materials grids and incubated for 3 min. The samples were then negatively stained with a 2% aqueous solution of uranium dioxide acetate for 5 min, dried, and examined using TEM (Tecnai G2 Spirit BioTWIN; The Netherlands). The size distribution profiles and concentrations of

Table 1
Clinical characteristics of the study populations.

Population 1. Five pregnant women carrying nsCLP foetuses and five pregnant women carrying normal foetuses (n = 10).								
Characteristics	Normal (n = 5)	nsCLP (n = 5)	*p*					
Gestational age (weeks)	25•00 ± 0•71[a]	24•80 ± 0•84	0•46					
Maternal age (years)	30•00 ± 2•74	29•00 ± 3•81	0•52					
Population 2.[b] Twenty pregnant women carrying nsCLP foetuses and 20 pregnant women carrying normal foetuses (n = 40).								
Characteristics	Normal (n = 20)	nsCLP (n = 20)	*p*					
Gestational age (weeks)	25•10 ± 1•29	24•65 ± 0•75	0•19					
Maternal age (years)	30•55 ± 2•78	28•40 ± 5•71	0•14					
Population 3. Thirty pregnant women carrying foetuses with nsCLP and 30 pregnant women carrying normal foetuses (n = 60).								
Characteristics	Normal (n = 30)	nsCLP (n = 30)	*p*					
Gestational age (weeks)	25•33 ± 1•40	24•77 ± 3•05	0•36					
Maternal age (years)	30•13 ± 2•98	28•93 ± 5•08	0•27					
Population 4. Twenty pregnant women carrying normal foetuses (n = 20).								
Characteristics	Before delivery (n = 20)				24 h after delivery (n = 20)			
Gestational age (weeks)	39.13 ± 0.80							
	-							
Population 5. Eight pregnant women carrying normal foetuses (n = 8).								
Characteristics	4–6 weeks	8–9 weeks	12–13 weeks	16–17 weeks	20–21 weeks	24–26 weeks	36–38 weeks	41–48 days after delivery
Gestational age (weeks)	4•88 ± 0•35	8•75 ± 0•46	12•75 ± 0•46	16•63 ± 0•52	20•63 ± 0•52	25•00 ± 0•53	37•13 ± 0•64	-
Population 6. Ten foetuses with nsCLP and ten normal foetuses (n = 20).								
Characteristics	Normal (n = 10)	nsCLP (n = 10)	*p*					
Gestational age (weeks)	24•20 ± 1•03	24•70 ± 1•06	0•30					
Maternal age (years)	27•20 ± 1•55	28•20 ± 2•49	0•30					
Population 7. Five foetuses with nsCLP and five normal foetuses (n = 10).								
Characteristics	Normal (n = 5)	nsCLP (n = 5)	*p*					
Gestational age (weeks)	23•20 ± 0•84	23•80 ± 1•48	0.45					
Maternal age (years)	30•00 ± 1•58	28•20 ± 4•49	0.42					
Population 8.[c] **Twenty-four pregnant women carrying foetuses with nsCLP and 24 pregnant women carrying normal foetuses (n = 48).**								
Characteristics	Normal (n = 24)	nsCLP (n = 24)	*p*					
Gestational age (weeks)	24•92 ± 1•47	24•46 ± 2•02	0•37					
Maternal age (years)	30•00 ± 2•78	28•29 ± 6•08	0•22					
Population 9. Twenty-one pregnant women carrying foetuses with NTDs and 21 pregnant women carrying normal foetuses (n = 42).								
Characteristics	Normal (n = 21)	NTDs (n = 21)	*p*					
Gestational age (weeks)	24•24 ± 1•95	23•38 ± 8•05	0•64					
Maternal age (years)	29•33 ± 2•50	28•14 ± 3•71	0•23					
Population 10. Forty pregnant women carrying foetuses with CHDs and 40 pregnant women carrying normal foetuses (n = 80).								
Characteristics	Normal (n = 40)	CHDs (n = 40)	*p*					
Gestational age (weeks)	25.68 ± 2•20	25.55 ± 3•30	0.84					
Maternal age (years)	30.55 ± 3•84	29.90 ± 4.97	0.51					
Population 11. Ten pregnant women carrying foetuses with nsCLP and 10 pregnant women carrying normal foetuses (n = 20).								
Characteristics	Normal (n = 10)	nsCLP (n = 10)	*p*					
Gestational age (weeks)	16.30 ± 1.16	16.40 ± 1.07	0.77					
Maternal age (years)	29.00 ± 2.27	29.30 ± 2.36	0.86					

[a] the data are expressed as means ± standard deviations
[b] 10 pregnant women in population 1 were included in population 2
[c] 48 pregnant women were selected from populations 2 and 3

exosomes were determined using a ZetaView instrument (Particle Metrix, PMX 110; Meerbusch, Germany). Data analysis was performed with ZetaView 8.04.02 SP2 in the auto mode. Exosomes were lysed in radioimmunoprecipitation assay buffer containing 1% phenylmethylsulfonyl fluoride, and protein concentrations were analysed by the Bradford method (Beyotime Institute of Biotechnology, Haimen, China). Proteins were electroblotted onto nitrocellulose membranes, and the membranes were then blocked with 5% fat-free milk and incubated with the following primary antibodies: anti-ALG-2-interacting protein X (ALIX; 1:1,000; Cat# 2171S; RRID: AB_2299455), anti-CD81 (1:1,000; Cat# 66866-1-Ig; RRID: AB_2882203), anti-CD9 (1:1,000; Cat# 13403S; RRID:AB_2732848), and anti-placental alkaline phosphatase (PLAP; 1:1,000; Cat# 18507-1-ap; RRID:AB_10644334) at 4°C overnight. The next day, membranes were incubated with secondary antibodies (goat anti-rabbit [Cat# A-11008; RRID:AB_143165] or goat anti-mouse [Cat# A-11001; RRID:AB_143165]; 1:5,000) at room temperature for 2 h. The immunoblots were visualized by enhanced chemiluminescence (ECL kit; Millipore) and scanned using Azure Biosystem C300 software. Then, the ImageJ software (National Institutes of Health, Bethesda, MD, USA) was used to calculate the integrated density values.

2.6. Small RNA sequencing and piRNA target gene prediction

Ten libraries were prepared from the plasma exosomal RNA from five pregnant women carrying foetuses with nsCLP and five pregnant women carrying normal foetuses. The libraries were quantified using an Agilent 2100 Bioanalyzer (Agilent Technologies), and sequencing was conducted on a BGISEQ-500 platform (BGI-Shenzhen). Small RNAs were ligated to adenylated 3′ adapters annealed to unique molecular identifiers, followed by ligation of 5′ adapters. The raw sequences were filtered to remove tags with low quality, with 5′ contamination, with poly A tails, without 3′ primer, or shorter than 18 nucleotides. The obtained clean reads were mapped against piRNA-Bank (http://pirnabank.ibab.ac.in/request.html).

The targeting genes of piRNAs were predicted using the RNAhybrid and miRanda databases. The targets of selected piRNAs were identified using RNAhybrid, with a mean free energy of maximum -25 kcal/mol and *p* value less than 0.05. The targets of selected piRNAs were also identified using miRanda against the RNA library of the human genome, with a mean free energy of maximum -20 kcal/mol and a very stringent score threshold of 140. In the current study, only genes identified by both approaches were selected for further

analysis. Gene Ontology (GO) functional module enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis for piRNA-targeted mRNAs were used to assess the potential functions of dysregulated piRNAs [30,31]. GO terms and KEGG pathways with p values less than 0•05 were considered significantly enriched functions.

2.7. Reverse transcription and qRT-PCR

Extracted piRNAs were reverse transcribed into cDNA using a miRNA First Strand cDNA Synthesis Tailing Reaction kit (Sangon Biotech, Shanghai, China). Experiments were carried out in 20 μL reaction volumes according to the manufacturer's instructions. The cDNA was then subjected to real-time PCR using a SYBR Premix Ex Taq kit (Takara, Ohtsu, Japan) on a 7500 FAST Real-time PCR system (ABI, Carlsbad, CA, USA). Thermocycling conditions were as follows: 95 °C for 30 s (enzyme activation), followed by 45 cycles at 95 °C for 5 s and 60 °C for 34 s. piRNA primers (Sangon Biotech) were designed based on the piRNA sequences (piRNABanK) and are listed in Supplementary Table S1. Relative expression levels of piRNAs were calculated by the $2^{-\triangle\triangle CT}$ method; U6 was used as an internal control [32].

2.8. Statistical analysis

The data presented in this study are expressed as means ± standard deviations (Gaussian distributions assumed) or median and interquartile ranges (Gaussian distributions not assumed). Differentially expressed piRNAs in RNA sequencing data were screened using the DEGseq R package. Statistical analyses of population characteristics and relative piRNA expression levels were performed using SPSS 25.0 statistical software (IBM, New York, NY, USA) or GraphPad Prism 5.0 software. Two groups of data with Gaussian distributions were compared using Student's t-tests (two−tailed). Otherwise, non-parametric Mann-Whitney tests were used. Statistical analyses for the relative expression levels of piRNAs at 38−40 weeks of pregnancy and 24 h after normal delivery were determined by paired t-test for normally distributed data, and Wilcoxon test of paired t test for non-parametric data. Statistical analyses of relative expression of piRNAs at eight consecutive gestational ages during normal pregnancy in eight pregnant women were performed using Friedman tests [33]. Hierarchical clustering analysis was performed by MeV 4.9.0. Differences were considered significant when the p value was less than 0.05. To assess the diagnostic potential, the area under the receiver operating characteristic (ROC) curve (AUC), sensitivity, and specificity were calculated for each piRNA. Differences in AUCs were assessed using Medcalc 19.0.7.

2.9. Role of funding source

Funders had no role in study design, data collection, data analyses, data interpretation, or writing of the report.

3. Results

3.1. Characteristics of exosomes and exosomal piRNA profile analysis

TEM images revealed that the majority of particles were approximately 125 nm in diameter, corresponding to the size of exosomes (Fig. 2a). Size distribution analysis showed that the mean particle diameter was approximately 129.2 nm (Fig. 2b), consistent with the TEM results. Finally, western blot analysis showed that the exosomes were positive for typical exosome markers, i.e., ALIX, CD9, and CD81, and the placenta-specific exosome marker PLAP (Fig. 2c). In total, 13,674 differentially expressed piRNAs were detected in nsCLP and control exosomes, including 6,519 upregulated piRNAs and 7,155

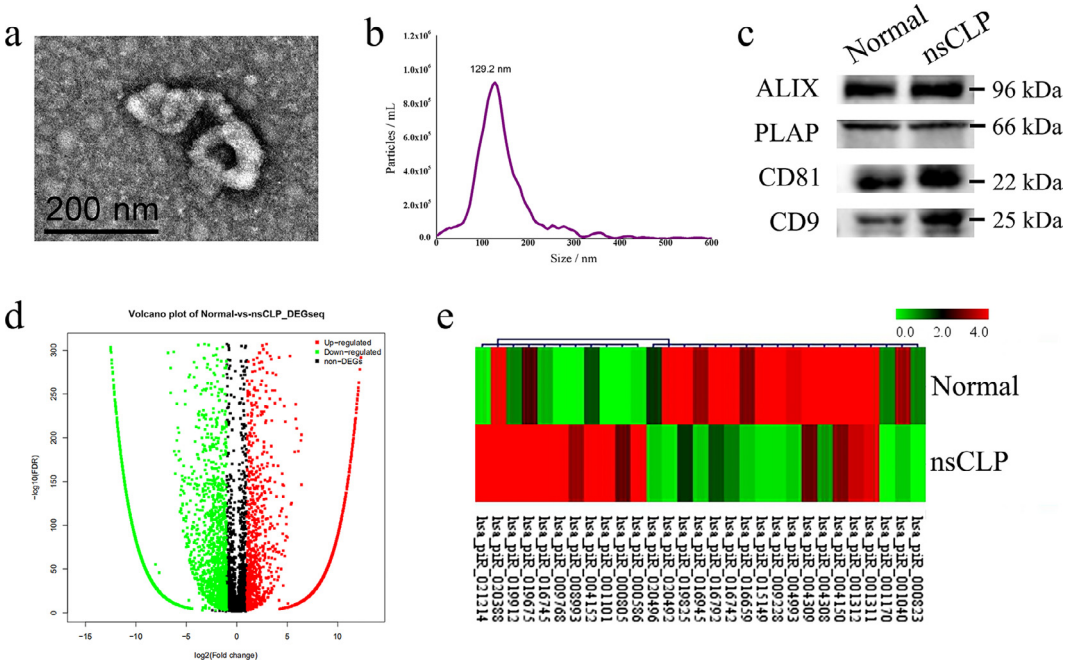

Fig. 2. Characteristics of exosomes and exosomal piRNA profile analysis. (a) Representative TEM images of plasma-derived exosomes (scale bar = 200 nm). (b) Quantification of exosomes isolated from the plasma of one healthy pregnant woman via nanoparticle tracking analysis. The mean diameter of particles was approximately 129.2 nm. (c) Western blot analysis of exosome markers (ALIX, CD81, and CD9) and the placental-expressed exosome marker PLAP in plasma exosomes isolated from one healthy woman and one pregnant woman carrying a foetus with nsCLP. The experiment was repeated in duplicate. (d) Volcano plot analysis of differentially expressed piRNAs in plasma-derived exosomes (Normal group, n = 5; nsCLP group, n = 5). (e) Hierarchical clustering analysis of 29 differentially expressed known mature piRNAs detected in all 10 samples (Normal group, n = 5; nsCLP group, n = 5). Upregulated piRNAs ($\log_2$ FC > 1; FDR < 0.001 by DEGseq R test) and downregulated piRNAs (log2 FC < -1; FDR < 0.001 by DEGseq R test) are shown in the cluster analysis.

downregulated piRNAs (Fig. 2d). Twenty-nine known mature piRNAs were differentially expressed (Supplementary Table S2), including 11 upregulated and 18 downregulated piRNAs. Hierarchical clustering analysis of known piRNAs revealed differential expression patterns in the nsCLP and control groups (Fig. 2e).

3.2. piRNA target gene prediction and GO and KEGG pathway analyses

Using RNAhybrid and miRanda and applying specific cut-offs, we identified gene targets for each piRNA. Target genes are listed in Supplementary Table S3. For upregulated piRNAs, the most enriched GO terms included gland morphogenesis, mammary gland morphogenesis, and cellular amino acid catabolic process in biological processes (BPs); cytosol, cytoplasm, and extracellular vesicle in cellular components (CCs); and benzodiazepine receptor binding, profiling binding, and maltose alpha-glucosidase activity in molecular functions (MFs; Fig. 3a). The top GO processes predicted by downregulated piRNAs were developmental process, anatomical structure development, and multicellular organism development in BPs; intracellular, intracellular organelle, and cytosol in CCs; and amide binding, peptide binding, and protein binding in MFs (Fig. 3b). KEGG pathway analysis identified 18 and 99 KEGG pathways in the upregulated and downregulated piRNAs, respectively. The top 10 pathways of upregulated and downregulated piRNAs are listed in Fig. 3c and 3d.

3.3. qRT-PCR analyses of 20 paired exosome samples in the training phase of the study

Twenty-nine differentially expressed piRNAs were verified in 20 paired plasma exosome samples in the training phase of the study. As shown in Fig. 4, six piRNAs (*hsa-piR-001101, hsa-piR-009228, hsa-piR-016659, hsa-piR-019912, hsa-piR-020388*, and *hsa-piR-020496*) were dysregulated in the nsCLP group and further evaluated in the validation phase. The differential expression levels of another 18 piRNAs (*hsa-piR-000805, hsa-piR-000823, hsa-piR-001040, hsa-piR-001311, hsa-piR-001312, hsa-piR-004150, hsa-piR-004152, hsa-piR-004309,*

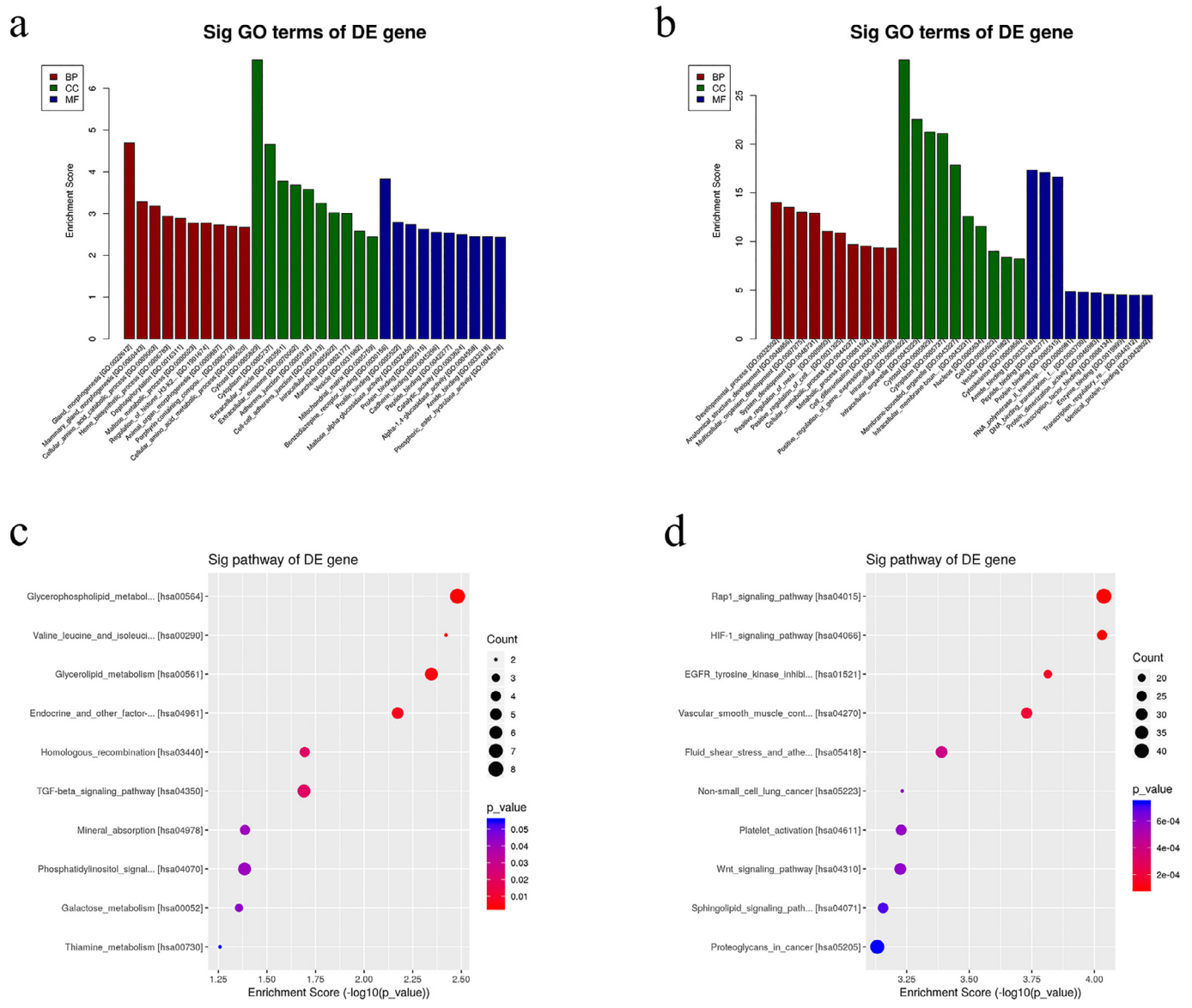

Fig. 3. GO terms and KEGG pathway analysis of the piRNA target genes. (a) GO annotation corresponding to the 11 upregulated piRNAs with top 10 enrichment scores covering BPs, CCs, and MFs. (b) GO annotations corresponding to the 18 downregulated piRNAs with top 10 enrichment scores covering BPs, CCs, and MFs. (c) Pathway enrichment analysis corresponding to 11 upregulated piRNAs with top 10 enrichment scores. (d) Pathway enrichment analysis corresponding to 18 downregulated piRNAs with top 10 enrichment scores.

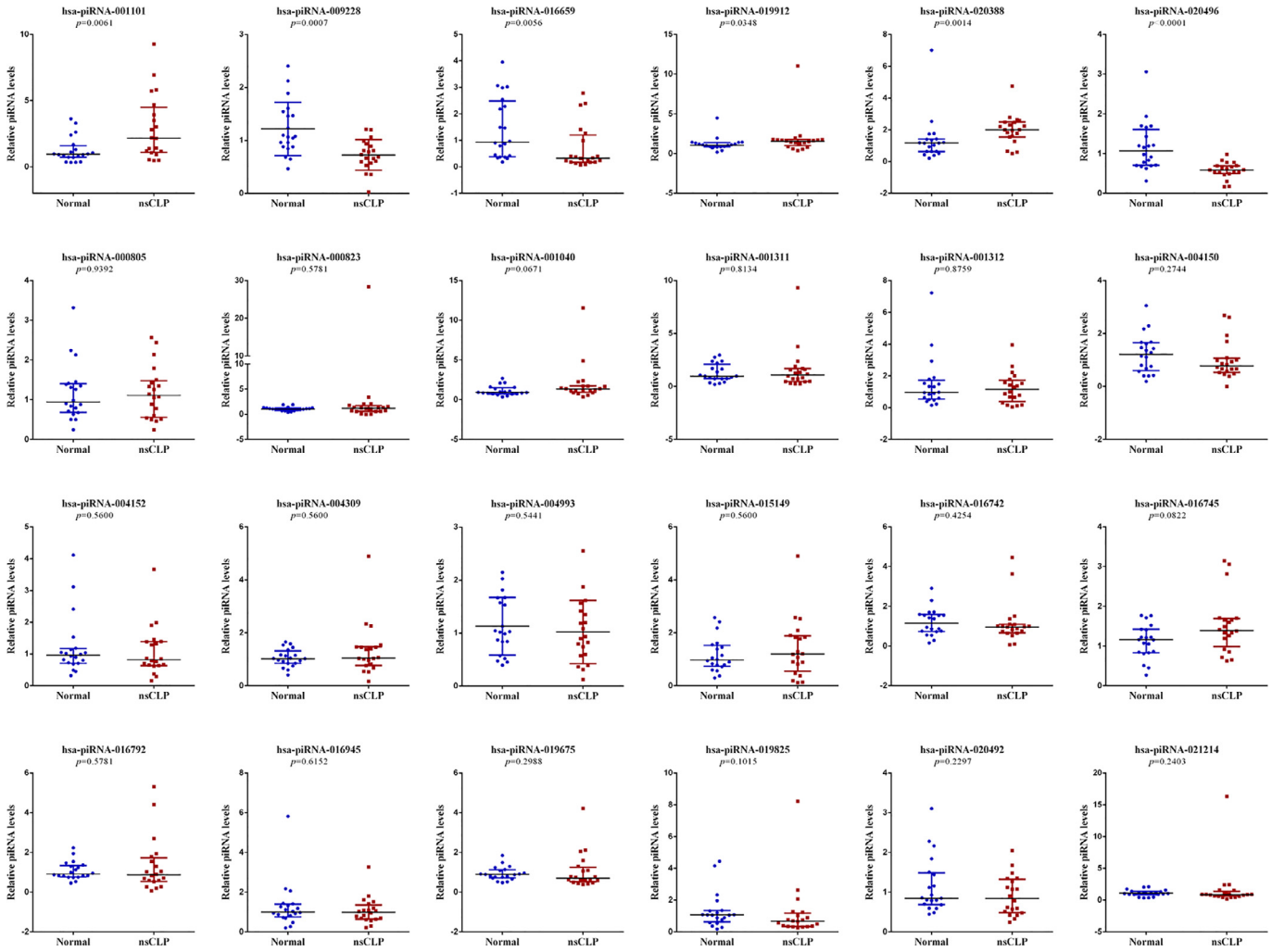

Fig. 4. Relative expression of the 24 indicated piRNAs in the training phase. qRT-PCR analysis was performed on the 20 paired plasma-derived exosome samples from pregnant women carrying foetuses with nsCLP (n = 20) and pregnant women carrying normal foetuses (n = 20). U6 was used as an internal control. *P* values were determined by unpaired t-test for normally distributed data, and Mann-Whitney test for non-parametric data. qRT-PCR measurements were performed in duplicate reactions. Data is expressed as mean ± SD (by unpaired t-test) or median and interquartile ranges (by Mann-Whitney test).

hsa-piR-004993, hsa-piR-015149, hsa-piR-016742, hsa-piR-016745, hsa-piR-016792, hsa-piR-016945, hsa-piR-019675, hsa-piR-019825, hsa-piR-020492, and *hsa-piR-021214*) were not verified in the training phases. In addition, because the Cq values of five piRNAs (*hsa-piR-000586, hsa-piR-001170, hsa-piR-004308, hsa-piR-008993*, and *hsa-piR-009768*) were higher than 35, and these piRNAs were not detected in all samples, they were eliminated from further analyses.

3.4. qRT-PCR analyses of 30 paired exosome samples in the validation phase of the study

Based on the criteria described in the Materials and Methods ($p < 0.05$; Cq value < 35), six piRNAs (*hsa-piR-001101, hsa-piR-009228, hsa-piR-016659, hsa-piR-019912, hsa-piR-020388*, and *hsa-piR-020496*) were further analysed in the validation phase in another 30 paired exosome samples. As shown in Fig. 5a, 5d, and 5e, *hsa-piRNA-001101, hsa-piR-019912*, and *hsa-piR-020388* were not confirmed in the validation phase. *hsa-piR-009228, hsa-piR-016659*, and *hsa-piR-020496* were significantly dysregulated in the plasma exosome samples of pregnant women carrying foetuses with nsCLP compared with those from women carrying normal foetuses ($p < 0.05$ by unpaired t-tests or Mann-Whitney tests; Fig, 5b, 5c, 5f). We further assessed the diagnostic value of *hsa-piR-009228, hsa-piR-016659*, and *hsa-piR-020496* for foetal nsCLP by ROC curve analysis for 30 case-control pairs and showed that these piRNAs had high diagnostic accuracy, with AUCs of 0•843, 0•711, and 0•820, respectively ($p < 0$•001; Fig. 5g). The combination of these three piRNAs generated a robustly increased AUC of 0•966 (95% confidence interval [CI], 0.883–0.990; $p < 0.0001$), with 87% sensitivity and 97% specificity (Table 2).

3.5. piRNA expression in different stages of normal pregnancy and tissues from foetuses with nsCLP

To observe whether the expression levels of the three candidate piRNAs were related to pregnancy, we verified plasma-derived exosomes via qRT-PCR in 20 pregnant women at two timepoints: 38−40 weeks of pregnancy and 24 h after normal delivery (Fig. 6a–6c). The expression levels of all the evaluated piRNAs decreased at 24 h after delivery. Hence, *hsa-piR-009228, hsa-piR-016659*, and *hsa-piR-020496* were assumed to be associated with pregnancy and were therefore selected for a more detailed analysis. To determine the changes in the expression of these three piRNAs at different stages of embryonic development, we measured plasma exosome piRNA levels at eight different gestational timepoints in eight women carrying normal foetuses. The levels of the three candidate piRNAs in the plasma exosomes of pregnant women were gradually elevated from gestational ages of 4−5 weeks with embryonic development, peaking

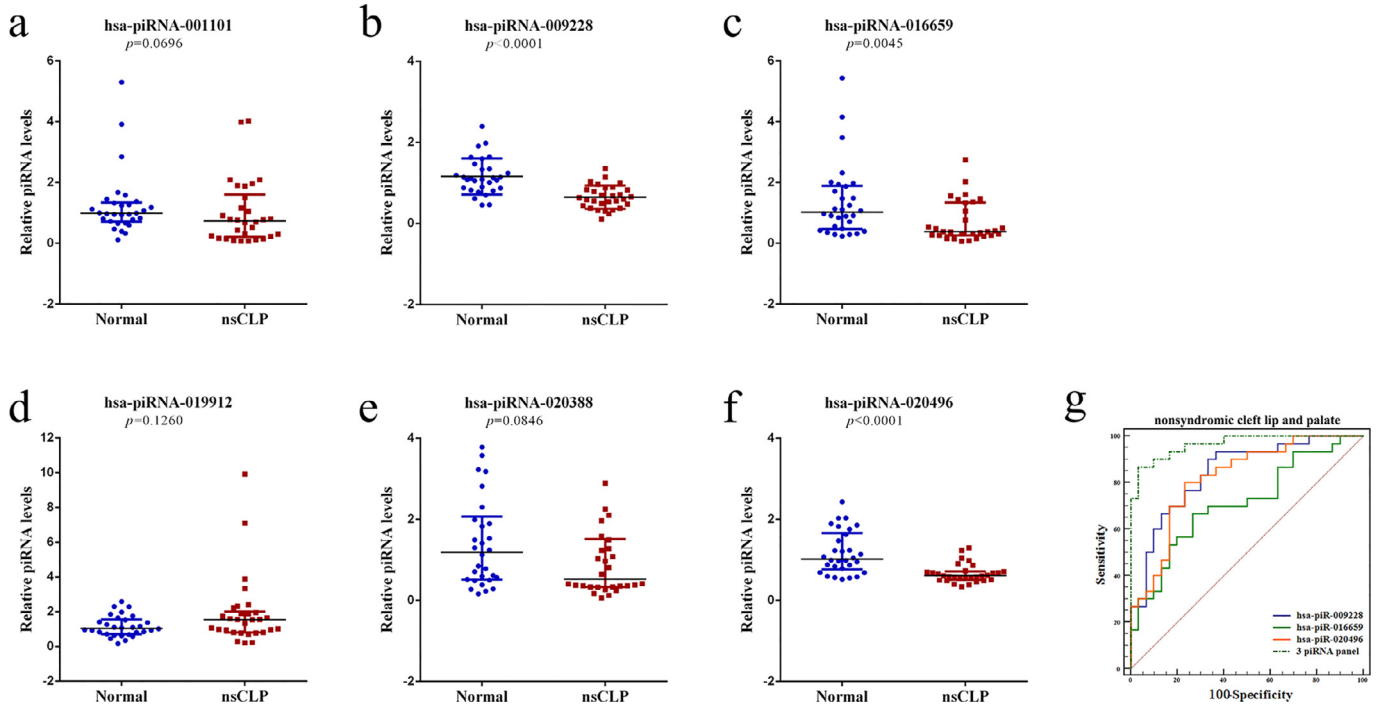

Fig. 5. Relative expression of the six indicated piRNAs in the validation phase. (a–f) qRT-PCR analysis was performed on the 30 paired plasma-derived exosomes from pregnant women carrying foetuses with nsCLP (n = 30) and pregnant women carrying normal foetuses (n = 30). U6 was used as an internal control. P values were determined by unpaired t-test for normally distributed data, and Mann-Whitney test for non-parametric data. qRT-PCR measurements were performed in duplicate reactions. Data is expressed as mean ± SD (by unpaired t-test) or median and interquartile ranges (by Mann-Whitney test). (g) The ROC curves of three differentially expressed piRNAs and a combination of the three piRNAs in plasma exosomes from the 30 case-control pairs (Normal group, n = 30; nsCLP group, n = 30).

at 8–9 weeks, and then decreased gradually until delivery (Fig. 6d–6f). The duration of high-level expression of these piRNA biomarkers was consistent with the time window of embryonic tissue and organ formation, implying that these piRNAs may be involved in embryogenesis of congenital malformations. Therefore, we examined the expression of these piRNAs in lip tissues from 10 pairs of foetuses with nsCLP and matched normal foetuses. As showed in Fig. 6g–6i, *hsa-piR-009228* and *hsa-piR-020496* were downregulated in lip tissues from foetuses with nsCLP, whereas the expression level of *hsa-piR-016659* was not statistically significant. The expression levels of these three piRNAs in placentas, umbilical cords, and fetal medial calf muscles were not statistically different between the nsCLP group and the normal group. *hsa-piR-009228* and *hsa-piR-020496* were downregulated in lip tissues from these 5 paired foetuses (Fig. 6j–6l). To further validate whether these significantly downregulated pregnancy-associated piRNAs were specific differentially expressed in exosomes, the expression levels of all three piRNAs were measured in 24 paired whole plasma samples from 24 pregnant women carrying foetuses with nsCLP and 24 plasma samples from women carrying normal foetuses. The expression levels of *hsa-piR-009228* and *hsa-piR-020496* were not significantly different in plasma from pregnant women carrying foetuses with nsCLP and those carrying healthy foetuses, indicating that these two piRNAs were specific differentially expressed in exosomes (Fig. 6m, 6o). The expression of *hsa-piR-016659* was significantly downregulated in both plasma-derived exosomes (p = 0.0045 by Mann-Whitney test) and whole plasma samples (p = 0.0002 by Mann-Whitney test; Fig. 6n).

3.6. Expression levels of piRNAs in exosome samples from women carrying foetuses with other related congenital malformations

The occurrence of CHDs, NTDs, and nsCLP is often related to the closure of the neural tube and the derivatives of neural crest cells. Therefore, qRT-PCR was performed to verify whether these three piRNAs could also be potential biomarkers for CHDs and NTDs. The expression levels of *hsa-piR-009228, hsa-piR-016659*, and *hsa-piR-020496* were significantly downregulated in plasma exosomes from pregnant women carrying foetuses with NTDs (Fig. 7a). *hsa-piR-009228* was downregulated and *hsa-piR-016659* was upregulated in plasma exosomes from pregnant women carrying foetuses with CHDs (Fig. 7c). ROC curve analysis was performed to determine the ability of the downregulated piRNAs to act as diagnostic biomarkers for NTDs or CHDs in the additional validation phase (Fig. 7c, 7d). As shown in Table 2, these dysregulated piRNAs could distinguish foetuses with CHDs or NTDs from normal foetuses with high sensitivity and specificity ($p < 0.05$). In 21 paired NTDs and normal exosome samples, *hsa-piR-009228, hsa-piR-016659*, and *hsa-piRNA-020496* had higher diagnostic accuracies, with AUCs of 0•989, 1•000, and 0•984, respectively . The combination of these three piRNAs generated a robustly increased AUC of 1•000 (95% CI, 0.916–1.000; $p < 0.0001$) with 100% sensitivity and 100% specificity. In 40 paired CHD and normal exosome samples, *hsa-piR-009228* and *hsa-piR-016659* had moderate diagnostic accuracies, with AUCs of 0•639 and 0•766, respectively, and the two-piRNA panel generated an AUC of 0•839 (95% CI, 0.740–0.911; $p < 0.0001$).

3.7. Expression levels of piRNAs in early plasma exosome samples

The ability to diagnose congenital malformations before the detection of typical structural abnormalities by ultrasound appears to be a unique advantage of prenatal biomarkers in the clinical setting. We collected samples from women carrying nsCLP foetuses in the early stages of pregnancy from the SJBC cohort. The results showed that the expression levels of *hsa-piR-009228, hsa-piR-016659*, and *hsa-piR-020496* were also downregulated in plasma exosomes from 10 pregnant women carrying nsCLP foetuses with gestational ages of 15–19 weeks (Fig. 8a–8c). To assess the diagnostic value of *hsa-piR-*

Table 2
ROC analysis of diagnostic ability of maternal plasma exosomal piRNAs to discriminate foetuses with congenital malformations and normal foetuses..

piRNA ID	nsCLP AUC, 95% CI	Sensitivity	Specificity	*p*	NTDs AUC, 95% CI	Sensitivity	Specificity	*p*	CHDs AUC, 95% CI	Sensitivity	Specificity	*p*
hsa-piR-009228	0.843, 0.726–0.924	90%	67%	< 0.0001	0.989, 0.895–1.000	100%	95%	< 0.0001	0.639, 0.524–0.744	60%	68%	0•0253
hsa-piR-016659	0.711, 0.585–0•821	67%	73%	0.0017	1.000, 0.916–1.000	100%	100%	< 0.0001	0.766, 0.658–0.854	60%	93%	< 0.0001
hsa-piR-020496	0•820, 0.699–0.907	80%	77%	< 0.0001	0•984, 0.887–1.000	100%	95%	< 0.0001	-	-	-	-
Three piRNA panel	0.966, 0•883–0•990	87%	97%	< 0.0001	1.000, 0•916–1•000	100%	100%	< 0.0001	0.839, 0.740–0.911	75%	85%	< 0.0001

AUC: area under the receiver operating characteristic curve

009228, hsa-piR-016659, and *hsa-piR-020496* for foetal nsCLP, ROC curve analysis for 10 case-control pairs was performed (Fig. 8d). These three piRNAs had high diagnostic accuracy, with AUCs of 0.930, 0.860, and 0.790, respectively. The combination of these three piRNAs generated a robustly increased AUC of 0•980 (95% CI, 0.797–1.000; $p < 0.0001$), with 100% sensitivity and 90% specificity (Table 3); these values were even better than those from later plasma exosome samples at about 24 weeks.

4. Discussion

Currently, early diagnosis of abnormal foetal development mainly relies on routine ultrasound examination. There is still a lack of specific molecular markers for diagnosing congenital malformations. In this study, we verified circulating piRNAs from plasma exosomes of pregnant women could be promising non-invasive prenatal biomarkers for congenital malformations. To the best of our knowledge, this is the first study analysing the expression levels of circulating piRNAs in plasma exosome samples from pregnant women carrying foetuses with nsCLP which are collected from an ongoing cohort study. Using RNA sequencing and subsequent qRT-PCR validation, 11 upregulated and 18 downregulated known mature piRNAs were detected. Then, we identified a panel of three piRNAs (*hsa-piR-009228, hsa-piR-016659*, and *hsa-piR-020496*) with robust performance for sensitivity and specificity in differentiating foetuses with congenital malformations from normal foetuses. Moreover, the diagnostic value of these piRNAs for nsCLP was better in early samples. The findings presented in our study are of potential clinical use and may have implications in the early intervention and treatment of congenital malformations.

Molecular biomarker-based non-invasive testing is already a successful clinical approach in the diagnosis of cancers [9], but has not been widely applied for pregnancy complications and foetal abnormalities, which cause high maternal and foetal mortality and morbidity. With the development of next-generation sequencing, cell-free noncoding RNAs are attracting much attention as promising molecular biomarkers [34]. miRNAs, long noncoding RNAs, and circular RNAs are the focus of many clinical studies aimed at detecting common pregnancy complications, such as gestational diabetes [35,36] and preeclampsia [37,38]. However, no studies have reported the use of piRNAs as potential prenatal biomarkers. In 2015, Yang and colleagues first described the presence of *piR-57125* in serum and plasma and showed that, similar to miRNAs, this piRNA remained extremely stable, regardless of repetitive freeze-thawing or long-term incubation at room temperature [39]. Additionally, previous studies on prenatal diagnosis have mainly focused on whole plasma or serum from pregnant women to identify potential biomarkers [40], and little attention has been paid to exosomes. Exosomes are vesicles measuring approximately 40–160 nm in diameter; these vesicles are rich in a large number of biologically active molecules and involved in transporting molecular signals between cells and are not easily degraded by enzymes in bodily fluids [41]. Studies have confirmed that exosomes can cross the placental barrier and infiltrate embryonic tissues, specifically reflecting the biological status of foetuses [11]. The concentration of plasma derived exosomes was more than 50-fold greater in pregnant women than in non-pregnant women [42]. It was reported that researchers constructed a Cyclic-recombinase-reporter mouse model to determine exosome communication and function during pregnancy, and demonstrated that total fetal exosomes in maternal plasma was about 35% [43]. Moreover, as noncoding RNAs, piRNAs are specifically involved in embryonic development. Accordingly, circulating piRNAs in plasma-derived exosomes may be a novel potential source of non-invasive biomarkers for prenatal diagnosis. Currently, studies on the screening of prenatal biomarkers for congenital malformations have all been case-control studies. The samples were obtained from pregnant women after

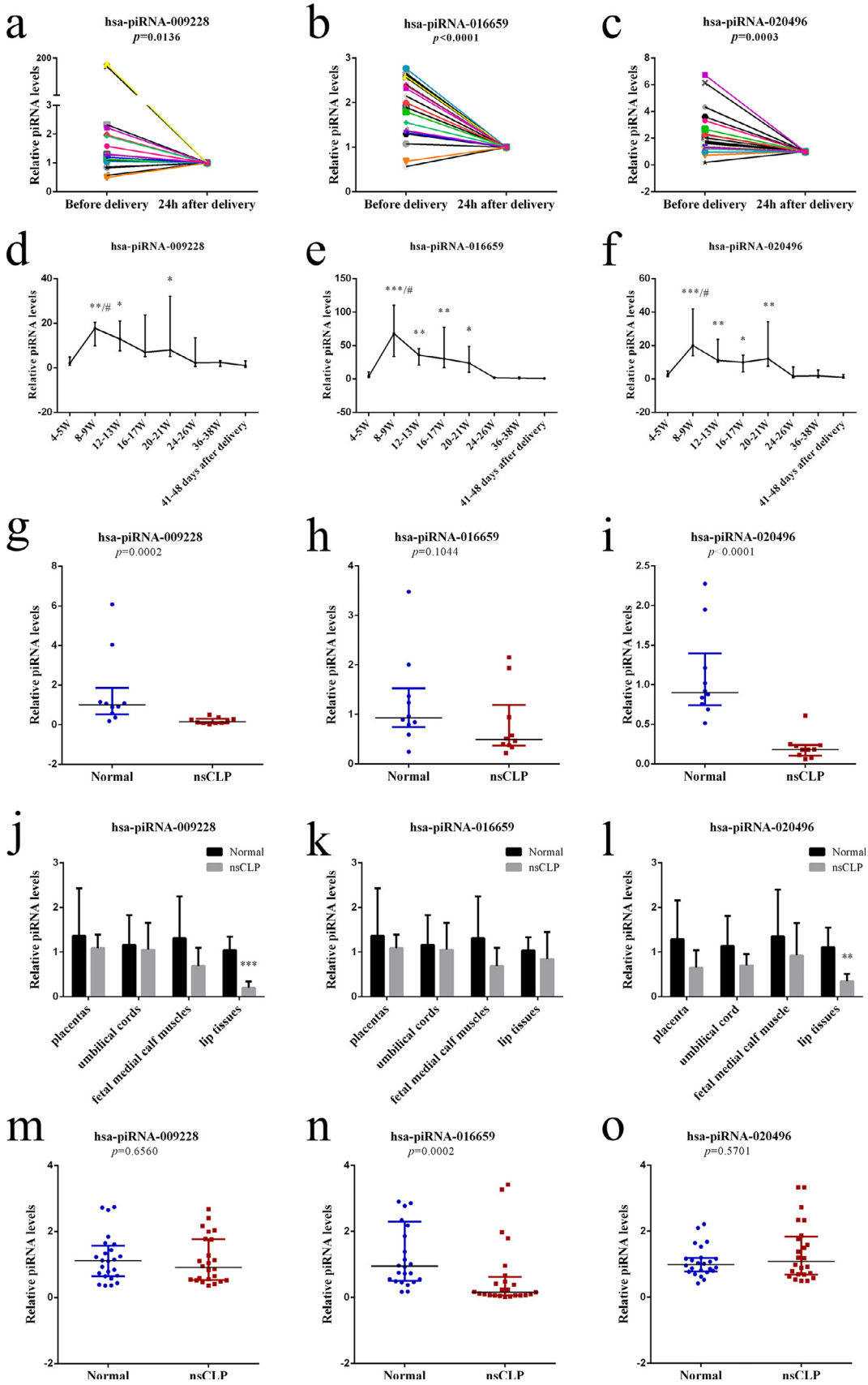

Fig. 6. Real-time PCR analysis of piRNAs in additional validation phase I. (a–c) Scatter plots of *hsa-piR-009228, hsa-piR-016659*, and *hsa-piR-020496*. Each line represents one plasma-derived exosome sample from one pregnant woman before and at 24 h after delivery ($n = 20$). *P* values were determined by paired t-test for normally distributed data, and Wilcoxon test of paired t test for non-parametric data. (d–f) piRNA levels at different gestational ages in eight women with normal foetuses were analysed by the Friedman test ($n = 8$). Error bars represent interquartile ranges of median. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ compared to values at 41–48 days after delivery; # $p < 0.05$ compared to values at the 4–5 weeks. (g–i) piRNA levels in lip tissues from foetuses with nsCLP ($n = 10$) and normal foetuses ($n = 10$). (j–l) piRNA levels in placentas, umbilical cords, fetal medial calf

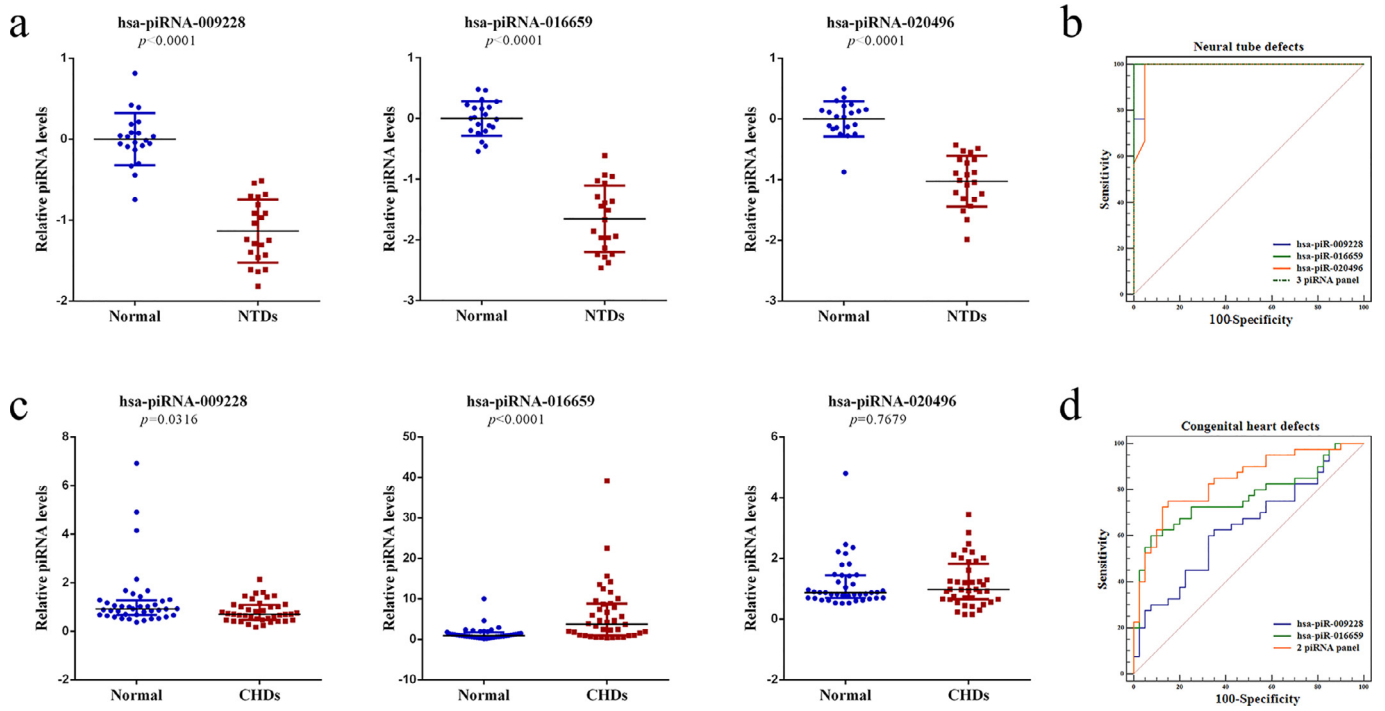

Fig. 7. Real-time PCR analysis of piRNAs in pregnant women carrying foetuses with NTDs/CHDs and pregnant women carrying normal foetuses. (a) Relative expression of *hsa-piR-009228, hsa-piR-016659*, and *hsa-piR-020496* in 21 paired plasma-derived exosomes from pregnant women carrying foetuses with NTDs ($n = 21$) and pregnant women carrying normal foetuses ($n = 21$). piRNA expression levels were normalized to U6 RNA (log10 scaled). *P* values were determined by unpaired t-test. (b) ROC curves of *hsa-piR-009228, hsa-piR-016659*, and *hsa-piR-020496* and the combination of these piRNAs for distinguishing pregnant women carrying foetuses with NTDs from healthy foetuses (Normal group, $n = 21$; NTDs group, $n = 21$). (c) Relative expression of *hsa-piR-009228, hsa-piR-016659*, and *hsa-piR-020496* in 40 paired plasma-derived exosomes from pregnant women carrying foetuses with CHDs ($n = 40$) and pregnant women carrying normal foetuses ($n = 40$). *P* values were determined by Mann-Whitney test. qRT-PCR measurements were performed in duplicate reactions. (d) ROC curves of *hsa-piR-009228, hsa-piR-016659*, and *hsa-piR-020496,* and the combination of these piRNAs for distinguishing pregnant women carrying foetuses with CHDs from healthy foetuses (Normal group, $n = 40$; CHDs group, $n = 40$).

undergoing foetal ultrasound examination at gestational ages of approximately 24 weeks. In this report, we collected gestational data during pregnancy and evaluated different types of samples from women who had babies with three major congenital malformations. In particular, the results showed that the diagnostic value was better in early samples from the birth cohort than in late samples from case-control studies. Our findings are expected to facilitate the development of predictive models of abnormal foetal developmental disorders and give insights into early intervention and treatment.

In this study, we found that three piRNAs (*hsa-piR-009228, hsa-piR-016659*, and *hsa-piR-020496*) showed good diagnostic accuracy to distinguish foetuses with nsCLP from normal foetuses. In addition to nsCLP, our results also revealed the diagnostic value of CHDs and NTDs and showed that the performance of these biomarkers for NTDs was particularly high. The following perspectives may explain why these piRNA biomarkers were associated with these congenital malformations. First, the occurrence of CHDs, nsCLP, and NTDs was often related to the closure of the neural tube and derivatives of neural crest cells. For example, in addition to the spinal ganglia, neural crest cells were shown to be involved in the building of craniofacial and cardiac regions [44]. Second, a review provided a uniquely comprehensive view of the teratogens linked to the development of the above congenital malformations and revealed several shared multifactorial causes, such as maternal folate deficiency, gestational diabetes, and retinoic acid [45]. Finally, it should be noted that CHDs and nsCLP are also often found together, which illustrates the variety of neurocristopathy associations. Moreover, in the case of orofacial clefts, there are several similarities with NTDs, including their occurrence at a similar time during embryogenesis, their involvement in the midline of the embryo, their near identical population characteristics, and their similar gene contributions [46]. Endoplasmic reticulum membrane protein complex (EMC1) variants have been identified in patients exhibiting combined neurodevelopmental, craniofacial malformations and CHDs via dysfunction of the neural crest cells [47]. Overall, these results suggested that the occurrence of these three congenital malformations may share a common molecular mechanism, which could also explain the rationality of abnormal expression of piRNAs as biomarkers for the diagnosis of the three congenital malformations.

GO and KEGG pathway analyses were performed to explore the biological functions and molecular mechanisms of related piRNAs. GO analysis of piRNA target genes revealed that cytosol/cytoplasm was over-represented in CCs. In fact, piRNA post-transcriptional regulation occurs in the cytoplasm. In addition, vesicles and extracellular exosomes were also identified in the category of CCs. For exosomes serving as carriers of placental barrier information, our results provided insights into identifying candidate biomarkers for foetal developmental disorders. Among the top 10 biological process terms found in this study, developmental process, anatomical structure development, multicellular organism development, and system development have been reported to be involved in nsCLP [48]. This finding indicates that dysregulated piRNAs may participate in organism development during nsCLP progression. Transforming growth factor-β, epidermal growth factor receptor, and Wnt signalling are

muscles, and fetal lip tissues in foetuses with nsCLP ($n = 5$) and normal foetuses ($n = 5$). Data is presented as mean ± SD. $^{**}p < 0.01$, $^{***}p < 0.001$ compared to values at normal groups by unpaired t-test. (m–o) Relative expression of *hsa-piR-009228, hsa-piR-016659*, and *hsa-piR-020496* in 24 paired case-control plasma samples (Normal group, $n = 24$; nsCLP group, $n = 24$). *P* values were determined by Mann-Whitney test. qRT-PCR measurements were performed in duplicate reactions. Data is expressed as median and interquartile ranges.

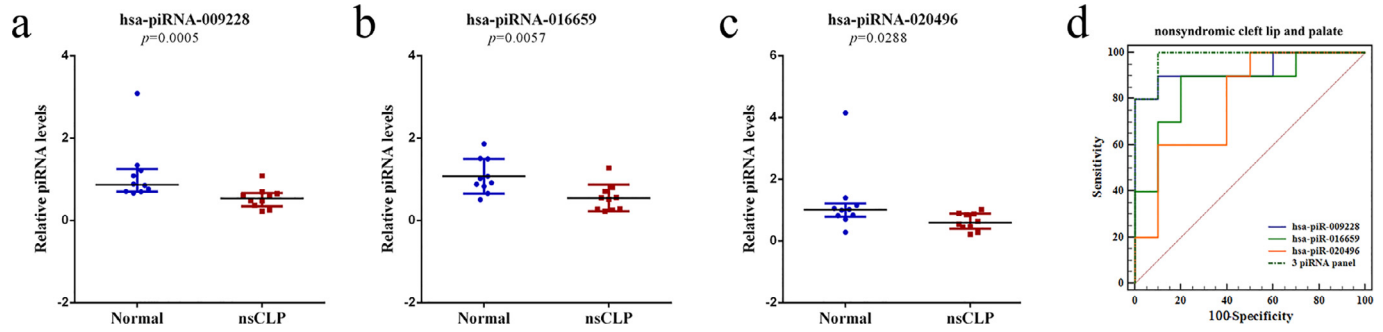

Fig. 8. Real-time PCR analysis of piRNAs in additional validation phase Ⅲ. (a–c) Relative expression of *hsa-piR-009228, hsa-piR-016659*, and *hsa-piR-020496* in 10 paired plasma-derived exosomes from pregnant women carrying foetuses with nsCLP (n = 10) and pregnant women carrying normal foetuses (n = 10) at 15–19 weeks of pregnancy. *P* values were determined by unpaired t-test for normally distributed data, and Mann-Whitney test for non-parametric data. qRT-PCR measurements were performed in duplicate reactions. (d) The ROC curves of *hsa-piR-009228, hsa-piR-016659*, and *hsa-piR-020496*, and the combination of these piRNAs for distinguishing pregnant women carrying foetuses with nsCLP from normal foetuses in this phase (Normal group, n = 10; nsCLP group, n = 10).

necessary for normal craniofacial development [49,50] and were identified in pathway enrichment analysis in this study. Consistent with this, altered functions of signalling pathways during early embryonic development contributed to nsCLP.

To the best of our knowledge, the functions of these three piRNAs have not been thoroughly studied. Potential target genes of *hsa-piR-020496*, such as *KRTAP8-1* and *KLK5* are involved in developmental biology signalling pathways. Some other potential target genes of *hsa-piR-020496* have been reported to be related to foetal abnormalities. For example, a homozygous mutation in the *ERCC6* gene was found to be related to cerebro-oculo-facio-skeletal syndrome, which is characterized by facial dysmorphism, microcephaly, congenital cataracts, neurogenic arthrogryposis, growth failure, and severe psychomotor retardation [51]. Reduced sulfate transporter SLC13A4 capacity in the placenta has been linked to disorders of craniofacial development [52]. The common target genes of *hsa-piR-020496* and *hsa-piR-009228*, including *MFRP, MAP4K5*, and *NF1*, are involved in the mitogen-activated protein kinase signalling pathway. These piRNA biomarkers may be involved in embryonic development of the neural tube and heart. A previous study also showed that *hsa-piR-009228* (DQ582496) was upregulated in cardiosphere-derived cells [53]. In our study, its expression was downregulated in the CHD group, suggesting that *hsa-piR-009228* may be necessary during the development of the foetal heart. Additionally, target genes of *hsa-piR-016659*, such as *KLHDC10, CAMTA2, CLIP2, NR2F2*, and *LMNA*, are related to foetal heart development [54–57]. Based on serum proteomic analysis for pregnant women carrying foetuses with CHDs in our previous study, *LMNA* was downregulated and showed better discriminatory power for prenatal biomarkers for foetal CHDs [15]. The lip, palate, and neural tube are all derived from the ectoderm, whereas the heart system develops from the mesoderm during embryonic development. These findings may explain why the three piRNAs identified in this study showed the same trends in expression in the nsCLP and NTD groups, but not the CHD group.

In this study, we found that three piRNAs (*hsa-piR-009228, hsa-piR-016659*, and *hsa-piR-020496*) showed good diagnostic accuracy to distinguish foetuses with nsCLP, NTDs, and CHDs from normal foetuses. As we know, the occurrence of fetal congenital malformations is caused by both maternal and fetal factors. The expression level of *hsa-piR-016659* was no statistical difference in lip tissues between nsCLP foetuses and normal foetuses, which indicated that the downregulation of *hsa-piR-016659* in maternal whole plasma and plasma derived exosomes may due to maternal factors. *hsa-piR-009228* and *hsa-piR-020496* were downregulated in nsCLP fetal lip tissues and plasma derived exosomes, but there were no statistical difference in placentas, umbilical cords, and fetal medial calf muscles between nsCLP foetuses and normal foetuses. These results indicated that dysregulated piRNAs from fetal lesion tissues might present characteristic changes in maternal peripheral blood via exosome vesicles pathway, and could be used as promising biomarkers for diagnosis of foetal abnormalities. Overall, the piRNAs from blood-derived exosomes of pregnant women can be used as prenatal biomarkers for identifying congenital malformations, and this emerging concept may provide insights into the diagnosis of other pregnancy complications, such as preeclampsia. However, there were some limitations in our study. First, the study is limited by the small sample size. In particular, only ten pregnant women carrying nsCLP foetuses at 15–19 weeks of pregnancy were selected from all 4117 participants in the SJBC study. Here, the incidence is little higher than reported due to more high-risk pregnant women in our birth cohort study, which may be as a potential confounder. piRNA biomarkers should be validated in more samples from early pregnancy from a larger cohort study. Multicentre studies with larger sample sizes are needed to evaluate the prenatal diagnostic potential of using piRNAs as biomarkers for diagnosing abnormal foetal developmental diseases for translation into clinical applications. Secondly, since the standardized methods of isolation and purification of human fetal exosomes have not been reported, only the total concentration of plasma derived exosomes in pregnant women are considered in our study. Therefore, the origin of piRNA biomarkers from different tissues, such as placenta and fetal lip tissues, are worthy of further research and exploration. Finally, the mechanisms through which piRNAs are related to the embryogenesis of CHDs, NTDs, and nsCLP are not yet clear, and further studies are needed to explore the potential functions of piRNAs and elucidate their associations with these congenital malformations.

In summary, in the current study, we identified three pregnancy-related piRNAs, i.e., *hsa-piRNA-009228, hsa-piRNA-016659*, and *hsa-piRNA-020496*, as novel promising biomarkers for foetal nsCLP, NTDs, and CHDs. *hsa-piRNA-009228* and *hsa-piRNA-020496* were specific to plasma-derived exosomes and downregulated in lip tissues of foetuses with nsCLP. In particular, we used the samples of early pregnancy from the birth cohort at which time the typical structural abnormalities were not screened by foetal ultrasound, and found that the diagnostic efficiency of nsCLP was better than the samples of late pregnancy. Taken together, these findings reveal the potential

Table 3
ROC analysis of the diagnostic ability of maternal plasma exosome piRNAs to discriminate foetuses with nsCLP from normal foetuses in the additional validation phase Ⅲ.

piRNA ID	AUC, 95% CI	Sensitivity	Specificity	*p*
hsa-piR-009228	0.930, 0.723–0.996	80%	100%	< 0.0001
hsa-piR-016659	0.860, 0.633–0.972	90%	80%	< 0.0001
hsa-piR-020496	0.790, 0.552–0.937	60%	90%	0.0064
Three piRNA panel	0.980, 0.797–1•000	100%	90%	< 0.0001

application value of exosome piRNAs as predictive biomarkers for congenital malformations. This is still a pilot study and multicentre studies with larger sample sizes are needed to translate this into clinical applications.

5. Contributions

SSJ and ZWY designed the experiments. SSJ, QZ, YW, and YFW prepared plasma and exosomes samples. SSJ, YWH, and WM collected data and performed experiments for the study. SSJ, XWW, and DL analysed the data. SSJ and ZWY wrote the first draft of the paper. SSJ, ZWY, HG, and WTL contributed to the writing of the paper. All authors read and approved the final manuscript.

6. Data sharing statement

The related data and materials are available for sharing upon request to Dr. Shanshan Jia and Prof. Zhengwei Yuan.

Research in context

Evidence before this study

Current clinical strategies to diagnose congenital abnormalities prenatally rely mainly on ultrasound, often performed when typical structural defects have already formed; these strategies thus miss the optimal time for intervention and treatment. Exosome-derived PIWI-interacting RNA (piRNAs) may exhibit stability in clinical tests and could be closely related to embryonic development. However, no studies have reported the application of piRNAs as non-invasive prenatal biomarkers for diagnosing congenital malformations.

Added value of this study

Pregnancy-associated *hsa-piR-009228, hsa-piR-016659*, and *hsa-piR-020496* in maternal plasma-derived exosomes may be promising biomarkers for the non-invasive prenatal diagnosis of foetal nonsyndromic cleft lip and palate (nsCLP) or neural tube defects (NTDs). *hsa-piR-009228* and *hsa-piR-016659* may be potential biomarkers for non-invasive prenatal diagnosis of foetal congenital heart defects (CHDs). The expression levels of the three piRNAs gradually increased during early embryonic development, and the differentially expressed piRNAs (*hsa-piR-009228* and *hsa-piR-020496*) in exosomes were significantly downregulated in nsCLP fetal lip tissues. Moreover, the expression levels of the piRNA biomarkers were no statistical difference in placentas, umbilical cords, and fetal medial calf muscle between nsCLP foetuses and normal foetuses.

Implications of all the available evidence

This is the first study analysing the expression levels of circulating piRNAs in plasma exosome samples from pregnant women carrying foetuses with congenital malformations using RNA sequencing and subsequent real-time quantitative reverse transcription polymerase chain reaction validation. Pregnancy-associated *hsa-piR-009228, hsa-piR-016659*, and *hsa-piR-020496* in plasma exosomes from pregnant women provided high sensitivity/specificity for diagnosing foetal congenital abnormalities. These piRNAs also acted as promising early biomarkers for nsCLP before the clefts could be screened using ultrasound examination. These findings presented in this study are of potential clinical use, and may have implications in the studies of other prenatal diseases.

Declaration of Competing Interest

The corresponding author has a pending patent related to this research work, no other authors have reported a potential conflict of interest relevant to this article.

Acknowledgements

This work was supported bythe National Key Research and Development Program (No. 2016YFC1000505),the National Natural Science Foundation of China (No. 81871219, 81671469), and the LiaoNing Revitalization Talents Program (No. XLYC1902099). The authors would like to thank the doctors from the Obstetrics and Ultrasound departments who have been involved in SJBC study. They carried out professional clinical examination of all participants of the SJBC study.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ebiom.2021.103253.

References

[1] Corsello G, Giuffre M. Congenital malformations. J Matern Fetal Neonatal Med 2012;25(Suppl 1):25–9.
[2] Stanier P, Moore GE. Genetics of cleft lip and palate: syndromic genes contribute to the incidence of non-syndromic clefts. Hum Mol Genet 2004;13:R73–81 Spec No 1.
[3] Marelli AJ, Ionescu-Ittu R, Mackie AS, Guo L, Dendukuri N, Kaouache M. Lifetime prevalence of congenital heart disease in the general population from 2000 to 2010. Circulation 2014;130:749–56.
[4] Rahimov F, Jugessur A, Murray JC. Genetics of nonsyndromic orofacial clefts. Cleft Palate Craniofac J 2012;49:73–91.
[5] Khoshnood B, Loane M, de Walle H, et al. Long term trends in prevalence of neural tube defects in Europe: population based study. BMJ 2015;351:h5949.
[6] Dixon MJ, Marazita ML, Beaty TH, Murray JC. Cleft lip and palate: understanding genetic and environmental influences. Nat Rev Genet 2011;12:167–78.
[7] Walker NJ, Podda S. Cleft Lip. StatPearls. Treasure Island (FL) 2020.
[8] Zheng W, Li B, Zou Y, Lou F. The prenatal diagnosis and classification of cleft palate: the role and value of magnetic resonance imaging. Eur Radiol 2019;29:5600–6.
[9] Bardelli A, Pantel K. Liquid Biopsies, What We Do Not Know (Yet). Cancer Cell 2017;31:172–9.
[10] Bardelli A. Medical research: Personalized test tracks cancer relapse. Nature 2017;545:417–8.
[11] Yang H, Ma Q, Wang Y, Tang Z. Clinical application of exosomes and circulating microRNAs in the diagnosis of pregnancy complications and foetal abnormalities. J Transl Med 2020;18:32.
[12] Krantz DA, Hallahan TW, Sherwin JE. Screening for open neural tube defects. Clin Lab Med 2010;30:721–5.
[13] Milunsky A, Jick SS, Bruell CL, et al. Predictive values, relative risks, and overall benefits of high and low maternal serum alpha-fetoprotein screening in singleton pregnancies: new epidemiologic data. Am J Obstet Gynecol 1989;161:291–7.
[14] Dong N, Gu H, Liu D, et al. Complement factors and alpha-fetoprotein as biomarkers for noninvasive prenatal diagnosis of neural tube defects. Ann N Y Acad Sci 2020;1478:75–91.
[15] Chen L, Gu H, Li J, et al. Comprehensive maternal serum proteomics identifies the cytoskeletal proteins as non-invasive biomarkers in prenatal diagnosis of congenital heart defects. Sci Rep 2016;6:19248.
[16] An D, Wei X, Li H, et al. Identification of PCSK9 as a novel serum biomarker for the prenatal diagnosis of neural tube defects using iTRAQ quantitative proteomics. Sci Rep 2015;5:17559.
[17] Gu H, Chen L, Xue J, et al. Expression profile of maternal circulating microRNAs as non-invasive biomarkers for prenatal diagnosis of congenital heart defects. *Biomed Pharmacother* 2019;109:823–30.
[18] Gu H, Li H, Zhang L, et al. Diagnostic role of microRNA expression profile in the serum of pregnant women with fetuses with neural tube defects. J Neurochem 2012;122:641–9.
[19] Aravin A, Gaidatzis D, Pfeffer S, et al. A novel class of small RNAs bind to MILI protein in mouse testes. Nature 2006;442:203–7.
[20] Lau NC, Seto AG, Kim J, et al. Characterization of the piRNA complex from rat testes. Science 2006;313:363–7.
[21] Li Y, Zeng A, Li G, et al. Dynamic regulation of small RNAome during the early stage of cardiac differentiation from pluripotent embryonic stem cells. Genom Data 2017;12:136–45.
[22] Rojas-Rios P, Simonelig M. piRNAs and PIWI proteins: regulators of gene expression in development and stem cells. Development 2018:145.
[23] Halbach R, Miesen P, Joosten J, et al. A satellite repeat-derived piRNA controls embryonic development of Aedes. Nature 2020;580:274–7.

[24] Kurth HM, Mochizuki K. 2′-O-methylation stabilizes Piwi-associated small RNAs and ensures DNA elimination in tetrahymena. RNA 2009;15:675–85.
[25] Mei Y, Clark D, Mao L. Novel dimensions of piRNAs in cancer. Cancer Lett 2013;336:46–52.
[26] Fonseca Cabral G, Azevedo Dos Santos Pinheiro J, Vidal AF, Santos S. piRNAs in gastric cancer: a new approach towards translational research. Int J Mol Sci 2020;21.
[27] Wang L, Feng Z, Wang X, Wang X, Zhang X. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. Bioinformatics 2010;26:136–8.
[28] Lotvall J, Hill AF, Hochberg F, et al. Minimal experimental requirements for definition of extracellular vesicles and their functions: a position statement from the international society for extracellular vesicles. J Extracell Vesicles 2014;3:26913.
[29] Thery C, Witwer KW, Aikawa E, et al. Minimal information for studies of extracellular vesicles 2018 (MISEV2018): a position statement of the international society for extracellular vesicles and update of the MISEV2014 guidelines. J Extracell Vesicles 2018;7:1535750.
[30] Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The gene ontology consortium. Nat Genet 2000;25:25–9.
[31] Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000;28:27–30.
[32] Weng W, Liu N, Toiyama Y, et al. Novel evidence for a PIWI-interacting RNA (piRNA) as an oncogenic mediator of disease progression, and a potential prognostic biomarker in colorectal cancer. Mol Cancer 2018;17:16.
[33] Luetkens JA, Voigt M, Faron A, et al. Influence of hydration status on cardiovascular magnetic resonance myocardial T1 and T2 relaxation time assessment: an intraindividual study in healthy subjects. J Cardiovasc Magn Reson 2020;22:63.
[34] Nagy B. Cell-free nucleic acids in prenatal diagnosis and pregnancy-associated diseases. EJIFCC 2019;30:215–23.
[35] Zhu Y, Tian F, Li H, Zhou Y, Lu J, Ge Q. Profiling maternal plasma microRNA expression in early pregnancy to predict gestational diabetes mellitus. Int J Gynaecol Obstet 2015;130:49–53.
[36] Wu H, Wu S, Zhu Y, et al. Hsa_circRNA_0054633 is highly expressed in gestational diabetes mellitus and closely related to glycosylation index. Clin Epigenetics 2019;11:22.
[37] Sun Y, Hou Y, Lv N, et al. Circulating lncRNA BC030099 increases in preeclampsia patients. Mol Ther Nucleic Acids 2019;14:562–6.
[38] Zhang YG, Yang HL, Long Y, Li WL. Circular RNA in blood corpuscles combined with plasma protein factor for early prediction of pre-eclampsia. BJOG 2016;123:2113–8.
[39] Yang X, Cheng Y, Lu Q, Wei J, Yang H, Gu M. Detection of stably expressed piRNAs in human blood. Int J Clin Exp Med 2015;8:13353–8.
[40] Carbone IF, Conforti A, Picarelli S, Morano D, Alviggi C, Farina A. Circulating nucleic acids in maternal plasma and serum in pregnancy complications: are they really useful in clinical practice? a systematic review. Mol Diagn Ther 2020;24:409–31.
[41] Kalluri R, LeBleu VS. The biology, function, and biomedical applications of exosomes. Science 2020;367.
[42] Salomon C, Torres MJ, Kobayashi M, et al. A gestational profile of placental exosomes in maternal plasma and their effects on endothelial cell migration. PLoS One 2014;9:e98667.
[43] Sheller-Miller S, Choi K, Choi C, Menon R. Cyclic-recombinase-reporter mouse model to determine exosome communication and function during pregnancy. Am J Obstet Gynecol 2019;221 502 e1- e12.
[44] Etchevers HC, Dupin E, Le Douarin NM. The diverse neural crest: from embryology to human pathology. Development 2019:146.
[45] Cerrizuela S, Vega-Lopez GA, Aybar MJ. The role of teratogens in neural crest development. Birth Defects Res 2020;112:584–632.
[46] De-Regil LM, Pena-Rosas JP, Fernandez-Gaxiola AC, Rayco-Solon P. Effects and safety of periconceptional oral folate supplementation for preventing birth defects. Cochrane Database Syst Rev 2015 CD007950.
[47] Marquez J, Criscione J, Charney RM, et al. Disrupted ER membrane protein complex-mediated topogenesis drives congenital neural crest defects. J Clin Invest 2020;130:813–26.
[48] Zou J, Li J, Li J, Ji C, Li Q, Guo X. Expression profile of plasma microRNAs in nonsyndromic cleft lip and their clinical significance as biomarkers. Biomed Pharmacother 2016;82:459–66.
[49] Miettinen PJ, Chin JR, Shum L, et al. Epidermal growth factor receptor function is necessary for normal craniofacial development and palate closure. Nat Genet 1999;22:69–73.
[50] Reynolds K, Kumari P, Sepulveda Rincon L, et al. WNT signaling in orofacial clefts: crosstalk, pathogenesis and models. Dis Model Mech 2019;12.
[51] Jaakkola E, Mustonen A, Olsen P, et al. ERCC6 founder mutation identified in Finnish patients with COFS syndrome. Clin Genet 2010;78:541–7.
[52] Rakoczy J, Zhang Z, Bowling FG, Dawson PA, Simmons DG. Loss of the sulfate transporter Slc13a4 in placenta causes severe fetal abnormalities and death in mice. Cell Res 2015;25:1273–6.
[53] Vella S, Gallo A, Lo Nigro A, et al. PIWI-interacting RNA (piRNA) signatures in human cardiac progenitor cells. Int J Biochem Cell Biol 2016;76:1–11.
[54] Siekierska A, Stamberger H, Deconinck T, et al. Biallelic VARS variants cause developmental encephalopathy with microcephaly that is recapitulated in vars knockout zebrafish. Nat Commun 2019;10:708.
[55] Kaufmann R, Straussberg R, Mandel H, et al. Infantile cerebral and cerebellar atrophy is associated with a mutation in the MED17 subunit of the transcription preinitiation mediator complex. Am J Hum Genet 2010;87:667–70.
[56] Stessman HA, Xiong B, Coe BP, et al. Targeted sequencing identifies 91 neurodevelopmental-disorder risk genes with autism and developmental-disability biases. Nat Genet 2017;49:515–26.
[57] Page DJ, Miossec MJ, Williams SG, et al. Whole exome sequencing reveals the major genetic contributors to nonsyndromic tetralogy of fallot. Circ Res 2019;124:553–63.