



# OPEN Profiling the cell-specific small non-coding RNA transcriptome of the human placenta

Nikita Telkar<sup>1,2,3</sup>, Desmond Hui<sup>1</sup>, Maria S. Peñaherrera<sup>1,2</sup>, Victor Yuan<sup>1</sup>, Victor D. Martinez<sup>4,5,6</sup>, Greg L. Stewart<sup>3</sup>, Alexander G. Beristain<sup>1,7</sup>, Wan L. Lam<sup>3,8,9</sup>✉ & Wendy P. Robinson<sup>1,2,9</sup>✉

The human placenta is the composite of multiple cell types, each which contributes uniquely to placental function. Small non-coding RNAs (sncRNAs) are regulators of gene expression and can be cell-specific. The sncRNA transcriptome of individual placental cell types has not yet been investigated due to difficulties in their procurement and isolation. Using a custom sequencing method, we explored the expression of seven sncRNA species (miRNA, piRNA, rRNA, scaRNA, snRNA, snoRNA, tRNA) from whole chorionic villi and four major sample-matched FACS-sorted cell type (cytotrophoblast, stromal, endothelial, Hofbauer) samples from 9 first trimester and 17 term placentas. After normalization for technical variables, samples clustered primarily by cell type lineage. No sncRNAs were uniquely expressed by cell type, however, mean expression differed by cell type for 115 sncRNAs. Known placentally-expressed sncRNAs showed differing expression by cell type and trimester. Expression of few sncRNAs varied by sex. Lastly, sample-matched sncRNA expression and DNA methylation correlation was not significant, although high correlation ( $> R^2 \pm 0.6$ ) was observed for some sncRNA-CpG pairs. This study represents the first exploration of the sncRNA transcriptome of bulk placental villi and placental cell types, informing about the expression and regulatory patterns underlying human placental development.

**Keywords** Placenta, Small non-coding RNA, MiRNA, Trophoblast, Chorionic villi

Gene expression and cellular differentiation in the human placenta are dynamically regulated to ensure maintenance and progression of a healthy pregnancy. This temporary, highly complex organ facilitates all communication between the mother and fetus, and aberrant functioning can lead to severe consequences in both mother and baby. The human placenta is derived from the conceptus and is thus, in most cases, genetically identical to the fetus. It is anchored to the maternal uterus and connected to the fetus by the umbilical cord, serving as the interface for nutrient, gas, and waste exchange between the mother and fetus. Even though the placenta is invaluable in its role in sustaining life, it remains under-investigated<sup>1</sup>.

The bulk of the human placenta consists of the chorionic villi (CV)<sup>2</sup>, which are branched, finger-like projections arising from the chorionic plate (the fetal side of the placenta). The CV are composed of several placental cell types, each with their own distinct developmental origins and functions<sup>3</sup>. The trophoblast layer of the blastocyst differentiates into cytotrophoblast (CTB) cells, which form the inner lining of the CV. The outer layer of the CV (which are in direct contact with the maternal blood) arises from the fusion of these CTBs into a multi-nucleated syncytium, and are termed as syncytiotrophoblast (SCT) – the major cell type component of the CV<sup>4</sup>. CTBs proliferate through the SCT layer to form CTB columns, attaching the CV to the maternal decidua. At the ends of these columns, CTBs further differentiate into extravillous trophoblasts (EVT), which invade and remodel the maternal spiral arteries, facilitating maternal blood flow into intervillous space (IVS). The chorionic villous core contains mainly placental stromal cells (SC), endothelial cells (EC), and Hofbauer cells (HBC). The

<sup>1</sup>British Columbia Children's Hospital Research Institute, Vancouver, BC V5Z 4H4, Canada. <sup>2</sup>Department of Medical Genetics, University of British Columbia, Vancouver, BC V6H 3N1, Canada. <sup>3</sup>British Columbia Cancer Research Institute, Vancouver, BC V5Z 1L3, Canada. <sup>4</sup>Department of Pathology and Laboratory Medicine, IWK Health Centre, Halifax, NS B3K 6R8, Canada. <sup>5</sup>Department of Pathology, Faculty of Medicine, Dalhousie University, Halifax, NS B3K 6R8, Canada. <sup>6</sup>Beatrice Hunter Cancer Research Institute, Halifax, NS B3H 4R2, Canada. <sup>7</sup>Department of Obstetrics & Gynecology, University of British Columbia, Vancouver, BC V6T 1Z3, Canada. <sup>8</sup>Department of Pathology, University of British Columbia, Vancouver, BC V6T 1Z7, Canada. <sup>9</sup>Wan L. Lam and Wendy P. Robinson contributed equally to this work. ✉email: wanlam@bccrc.ca; wrobinson@bcchr.ca

SCs are primarily fibroblast cells which maintain structural morphology and tissue homeostasis. ECs, which line the placental capillaries and are involved in cell-cell communication, play an essential role in the expansion of the placental vasculature and control blood flow. HBCs are M2-like placental macrophages derived from fetal monocytes that have several functions in the placenta including promoting vasculogenesis, angiogenesis, immune regulation, and transfer of ions and serum proteins<sup>5</sup>. Functional coordination of these various placental cell types is vital for proper structural maintenance, cellular communication, immune response, and vascular remodelling in the placenta.

Regulation of gene expression is governed by numerous factors, and small non-coding RNAs (sncRNAs) are one such post-transcriptional group of regulators. SncRNAs typically function by recruiting and binding to Argonaute and other proteins<sup>6</sup> to form the RNA-induced silencing complex (RISC)<sup>7</sup>, leading to RNA interference (RNAi) via repression of messenger RNA (mRNA), modifications on ribosomal RNA (rRNA), and even sponging of other ncRNAs<sup>8</sup>. Several species of sncRNAs have been discovered, with each species designated by their transcription origination and the factors that they regulate. NcRNAs are transcribed genome-wide in all organs, and show moderate conservation across various species<sup>9</sup>. Due to their smaller size, sncRNAs are relatively stable and therefore less prone to degradation after tissue sampling than mRNA, and have thus been proposed to be more reliable biomarkers of health<sup>10</sup>.

The most widely studied sncRNA species are microRNAs (miRNAs), which are the smallest sncRNAs in length spanning 19–22 nucleotides, and known to bind to and repress mRNA through the RISC. Several studies have investigated miRNA expression in placental pathologies<sup>11–14</sup>, particularly for two placentally-expressed imprinted miRNA clusters on chromosome 14 and chromosome 19<sup>15,16</sup>. miRNAs can bind mRNAs with imperfect complementarity at all nucleotide positions other than at the miRNA seed sequence (nucleotides 2–8), meaning that a single miRNA targets several different mRNAs<sup>17</sup>. Piwi-interacting RNAs (piRNAs) primarily suppress repetitive elements and are critically important in mammalian sperm cells<sup>18</sup>, and we previously observed that their expression profiles are unique in the placenta relative to other organs<sup>19</sup>. The biogenesis of ribosomal RNA (rRNA) subunits, which interact with transfer RNA (tRNA) to translate mRNA into protein, is complex – erroneous rRNA sequences have been suggested to generate antisense sequences which may interfere with translation of normal rRNA itself<sup>20</sup>. Small nuclear RNAs (snRNAs) and small nucleolar RNAs (snoRNAs) modify pre-mRNA processing<sup>21,22</sup>. Finally, tRNAs are thought to engage with mRNA, along with speculated roles in regulating cell death<sup>23,24</sup>. While we have previously described the overarching expression landscape of snRNAs, snoRNAs, and tRNAs in the human placenta<sup>25</sup>, in-depth characterization of bulk placental tissue and placental cell type sncRNAs is still limited<sup>26–29</sup>.

To date, sncRNA studies in the placenta have focused on either bulk placental CV tissue or on the investigation of miRNAs, and in limited cell types (usually CTBs)<sup>30,31</sup>. Bulk tissue profiling gives an approximate measure of the average of all cell types, but the proportions of each placental cell type in the sampled tissue may bias the observed overall expression, and can output an inaccurate representation of the expression landscape<sup>32–35</sup>. With all placental cell types performing their own distinct roles in the growth and functioning of the placenta, elucidating the cell-specific sncRNA profiles would be an invaluable resource.

To advance our understanding of the factors that regulate gene expression in placental cell types, we explored the sncRNA transcriptome of four major fluorescence activated cell (FAC)-sorted placental cell types - CTBs, SCs, ECs, and HBCs and their matched bulk villi samples, in first trimester and term placental samples. We assessed the differences and similarities in the expression of sncRNA species within these cell types, identified how expression varies in association with biological variables such as gestational age, trimester, and sex, and assessed the correlation of sncRNA expression with sample-matched DNA methylation data. To our knowledge, this is the first exploratory resource of placental cell-specific sncRNAs, and our findings highlight the importance of considering the contributions of placental cell types when assessing bulk placental tissue.

## Methods

### Sample acquisition and processing

Placental tissue samples were obtained with approval from The University of British Columbia / Children's and Women's Health Centre of British Columbia Research Ethics Board (H16-02280) as previously described<sup>36</sup>. All experiments were performed in accordance with relevant guidelines and regulations. In brief, for the 17 term (36–40 weeks) samples, pregnant individuals scheduled for a C-section at >36 weeks gestation with a singleton pregnancy and no pregnancy complications were recruited. Samples and data were collected by the BC Children's Hospital BioBank (Vancouver, BC). Additionally, 9 de-identified first trimester (6–13 weeks) placental samples from elective terminations were obtained, for which no gross pathologies were detected. All samples were confirmed to be chromosomally normal using copy number variant (CNV) calling on the Illumina Infinium MethylationEPIC array<sup>37</sup>.

For bulk CV term samples ( $n = 17$ ), three to four sites were sampled from below the surface of the fetal-facing side of the placental disc to avoid maternal contamination. These samples were preserved in RNAlater and pooled for processing. The first trimester placentas ( $n = 9$ ) are physically smaller in size, therefore a small singular tissue sample was preserved in RNAlater for all first trimester placentas. From each of the 26 pooled or single CV samples, 30 mg of tissue was processed, and excess RNAlater was removed by blotting. Samples were homogenized in the Next Advance Bullet Blender Tissue Homogenizer (Next Advance, USA), using 3.2 mm stainless steel beads (Next Advance, USA) with 1 ml of TRIzol reagent (ThermoFisher Scientific, USA). Samples were incubated for five minutes at room temperature (RT), and then centrifuged at 4,700  $\times$  g for three minutes. 200  $\mu$ l of chloroform (ThermoFisher Scientific, USA) was added to the supernatant, which was thoroughly mixed by inversion, and incubated for five minutes at RT. Samples were centrifuged at 4  $^{\circ}$ C at 7,800  $\times$  g for 20 min, and 500  $\mu$ l of isopropanol (ThermoFisher Scientific, USA) was added to the aqueous phase mixed by inversion, and incubated for 10 min at RT. Samples were centrifuged at 4  $^{\circ}$ C at 7,800  $\times$  g for 15 min. Supernatant

was discarded and the RNA pellet was washed in 75% ethanol by gentle inversion. Samples were centrifuged at 4 °C at 7,800 x g for 10 min. Supernatant was discarded, and the RNA pellet was air-dried for five minutes at RT. RNA was dissolved in 50 µl of ultrapure distilled RNase/DNase free water (Gibco-LifeTechnologies, USA). Genomic DNA removal was carried out using the RNase-Free DNase Set (Qiagen, Germany). RNA concentration was measured on a Nanodrop 2000 (ThermoFisher Scientific, USA), and RNA quality was assayed on an Agilent Bioanalyzer 2100 (Agilent, USA). CV samples were extracted by two separate people using the same protocol (extraction type 3, listed in Results).

Sample processing and cell isolation has been described previously<sup>36</sup>. Briefly, for the FAC-sorted placental cells (*n* = 104, representing four cell types per each of the 26 placentas), fresh CV samples were dissected and processed into a single-cell suspension, and then stained for various markers to be sorted by FACS using the BC Children's Hospital Research Institute FACS Core Equipment. Dead cells and red blood cells were first gated out using 7AAD-A + and CD235a+, respectively. Specific cell populations were then identified using various surface markers: CTBs as CD34-CD45-CD14-EGFR+, SCs as CD34-CD45-CD14-EGFR-, ECs as CD34+ CD45-, and HBCs as CD14+ CD34-. Cells were collected directly in DNA/RNA Shield 2x concentrate or Zymo RNA Lysis Buffer, stored at -80 °C, and total RNA was extracted from each sorted-cell population using the ZYMO Quick-RNA MiniPrep kit according to manufacturer's instructions (extraction type 1, listed in Results). Two samples were extracted using TRIzol (ThermoFisher Scientific, USA) (extraction type 2, listed in Results). All samples were DNased using the QIAGEN RNase-Free DNase Set before sequencing (QIAGEN, Germany).

As an initial quality control step, maternal contamination of all bulk and sorted cell samples was assessed, as described previously<sup>36</sup>. Samples predicted to have > 35% contamination based on Illumina DNA methylation array data and those with poor RNA concentration (< 1 ng/ul) were excluded. 28 of the 130 samples (130 samples = 26 CV + 104 sorted cells) showing high maternal contamination were excluded, of which the majority were first trimester (excluded: 8/26 CTB, 0/26 SC, 9/26 EC, 9/26 HBC, 2/26 CV), retaining 102 samples (130 total samples minus 28 high maternal contamination samples) for RNA sequencing (Table 1).

RNA sequencing

RNA sequencing was performed at Canada's Michael Smith Genome Sciences Centre (GSC) in Vancouver, Canada. The sequence read length profile reflects the size distribution of the isolated RNA (Supplementary Fig. 1), with the protocol being specific for miRNAs (fragments between 18 and 24 nucleotides in length). However, partial matches to other longer sncRNA species were also possible. Samples were sequenced based on a low-input RNA protocol described by Hagemann-Jensen et al.<sup>38</sup>. Briefly, 5.8 S rRNA was masked during adapter ligation by a complementary oligonucleotide, while miRNAs underwent 3' adapter ligation, digestion of unligated adapters, and 5' adapter ligation sequentially. The 5' adapter contains a unique molecular identifier (UMI), followed by two nucleotides (CA) prior to the miRNA sequence. Upon conversion to cDNA by reverse transcription, the cDNA was amplified by PCR to include the sequences required for Illumina cluster generation and sample indexing developed by the GSC. The number of PCR cycles was dependent on the amount of input RNA; samples with low and high input were subjected to 20 and 15 cycles of PCR, respectively. Quality of libraries was assessed using the Agilent Bioanalyzer 2100 HS DNA chip Assay, (Agilent, USA) after which size selection for the miRNA fraction was conducted using the Blue Pippin 3% agarose gel cassette (Sage Science Inc., USA). Samples were pooled for single-end 75 bp (SE75) sequencing on the Illumina NextSeq 500 (Illumina, USA) on two plates. Ten samples were sequenced initially (Batch 1) to estimate the minimum sample concentration required for subsequent sequencing (10ng/ul for sorted cells, 50 ng/µl for CV), and for the remaining 92 samples (Batch 2), 4 pools of 24 libraries each were aggregated and sequenced.

Data processing and analysis

Raw sequencing reads were trimmed to remove the 10 nucleotide UMI consisting of the eight nucleotides and CA dinucleotide from the 5' end, the 20 nucleotide Illumina TruSeq 3' adapter sequence (TGGAATTCTCGGG TGCCAAG), and all nucleotides after the 3' adapter. FASTQC<sup>39</sup>.

	First trimester Placentas = 8 (9) Samples = 21 (28)	Term Placentas = 17 (17) Samples = 71 (74)	Total Placentas = 25 (26) Samples = 92 (102)
Sex			
Female	5 (6)	9 (9)	14 (15)
Male	3 (3)	8 (8)	11 (11)
Cell type			
Cytotrophoblast (CTB)	4 (4)	12 (14)	16 (18)
Stromal cells (SC)	7 (9)	16 (17)	23 (26)
Endothelial cells (EC)	3 (5)	12 (12)	15 (17)
Hofbauer cells (HBC)	2 (3)	14 (14)	16 (17)
Whole chorionic Villi (CV)	5 (7)	17 (17)	22 (24)

**Table 1.** Sample distribution. The total 92 sorted-cell and matching chorionic villi samples that passed quality control and were included for analysis. The numbers in parentheses are the total samples of the 102 that were originally sequenced.

was used to assess sequencing quality metrics of samples. Trimmed reads were uploaded to the online miRMaster2 (v 2.0.0) tool<sup>40</sup> to discard sequences with less than 17 nucleotides, perform quality filtering using Phred  $\geq 20$  over a four-nucleotide sliding window, and to align and count raw reads relative to the human GRCh37 reference genome using STAR<sup>41</sup>. miRMaster2 utilizes sequence data from miRBase (22.1), Ensembl ncRNA (v 100), RNACentral piRNA (v 15), GtRNAdb (v 18.1), and NONCODE (v 5), and trimmed reads were aligned for seven sncRNA species using these databases: microRNAs (miRNAs), Piwi-interacting RNAs (piRNAs), ribosomal RNAs (rRNAs), small Cajal body-specific RNAs (scaRNAs), small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs), and transfer RNAs (tRNAs). Manual parameters for alignment were: 3' Adapter= TGGAATTCTCGGGTGCCAAGTCGACGTACGATC, 5' Barcode length=0 (as already trimmed beforehand), Adapter barcode length=0 (as already trimmed beforehand), UMI length=0 (as already trimmed beforehand), Minimum read length=17 (minimum length for miRNAs), Alignment tool=STAR. All other parameters selected were default settings (after consulting with the author of the software that the default parameters were suitable for use with the specific sequencing method that this study used).

miRMaster2<sup>40</sup> has been trained specifically for the alignment and annotation of non-coding RNAs and includes 72 successive steps in aligning reads to the genome, with extensive user-defined input parameters available. All non-coding RNAs thus had high confidence of alignment and subsequent read counts; a known caveat is that higher read counts can lead to a higher overall percentage of reads aligned. For sncRNAs other than miRNAs which have read lengths of  $> 22$ , those with perfect partial matches are also aligned and annotated by miRMaster2. These sncRNAs were explored with caution as the sequencing protocol for this study was specific towards miRNAs. Longer sncRNAs inherently have a greater probability of alignment (as longer genes have several regions that a read can potentially map to); miRMaster2 counts each read alignment as one count, however, this bias is present across all samples for longer sncRNAs, and thus is effectively nullified between cross-sample comparisons (as confirmed in personal correspondence with the author of the software). The non-normalized, raw count matrices for the above seven sncRNA species were downloaded to be processed for further analysis.

All analyses were carried out in R (v4.3.1)<sup>42</sup>. To assess sequencing quality, a variable termed 'sample quality' was constructed by combining two sequencing metrics: the total number of reads sequenced and percent of reads mapping to miRNAs (as this protocol was geared towards selecting for miRNAs). Samples with either (a) total number of reads less than one million or (b) less than 1% of reads mapping to miRNAs were flagged as 'red'; 1–30% of total reads mapping to miRNAs as 'yellow'; and either of (a) total number of reads more than one million or (b) more than 30% of reads mapping to miRNAs as 'green'. In cases where the total reads or total reads mapped to miRNAs did not have the same sample quality colour, the lower sample quality was assigned. The numerical cutoffs for each category were based on the expected sequencing quality as provided by the GSC. Five samples with sample quality designated 'red', two sorted-cell samples extracted by a different extraction method than the rest (extraction type 2 mentioned above, i.e., TRIzol), two EC samples that clustered with the CV samples along with having a 'yellow' sample quality, and one first trimester CV sample which had both its matched sorted-cell samples (SC and EC) assigned as outlier samples were removed from further processing (total samples removed = 10). Hence, the total samples that passed quality control were 92 of the original 102, as described in Table 1.

Data was batch-corrected using COMBAT-seq<sup>43</sup> for the two sequencing batches (Batch 1 and 2), normalized using the Relative Log Expression (RLE) method, and counts per million scaled using the 'edgeR' (v4.0.16)<sup>44</sup> and 'DESeq2' (v1.42.1)<sup>45</sup> packages. Principal Component Analysis (PCA) was conducted using the `prcomp` function from the base R 'stats' package (v4.3.1)<sup>42</sup> and the strength of association between metadata variables and principal components was estimated using linear models implemented with the `plomics` package (v 0.2.0)<sup>46</sup>.

Sequence data of all samples were split by cell type (CTB, SC, EC, HBC) for all sncRNAs that passed quality control. sncRNAs that had a normalized count  $\geq 2$  in at least 75% of term samples per cell type were retained. The mean expression of each of these sncRNAs was compared amongst all four cell types using the Kruskal-Wallis test. If the mean expression of a sncRNA in one cell type was significantly different than in all other cell types at a Benjamini-Hochberg (BH) corrected  $p$ -value  $\leq 0.05$ , the sncRNA was designated as cell type-associated. Several databases and software tools were used to mine information about sncRNA species including miRBase v22.1<sup>47</sup>, piRBase v3.0<sup>48</sup>, RNACentral v22<sup>49</sup>, GeneCards<sup>50</sup> and GeneCaRNA<sup>51</sup>, HUGO Gene Nomenclature Committee (HGNC)<sup>52</sup>, GtRNAdb<sup>53</sup>, UCSC Genome Browser<sup>54</sup>, and Ensembl<sup>55</sup>, as each database provides non-overlapping information.

For targeted analyses, we specifically focused on several large miRNA clusters, the human C14MC, C19MC, and miR-371-3, as they have been widely studied in the literature and highly expressed in the placenta. The miRNAs and their genomic locations belonging to the human C14MC (14q32.3), C19MC (19q13.41), and miR-371-3 (19q13.42) clusters were sourced and compiled from previous publications<sup>15,56</sup>.

Linear regression implemented with the 'limma' package<sup>57</sup> was used to identify sncRNAs that showed differing expression by sex. Sex differences in expression were assessed only in the term samples and separately within each cell type for sncRNAs on (a) all autosomes and (b) only the X and Y sex chromosomes, using the formula:

$$\text{Normalized Reads} \sim \text{Sex} + \text{Plate} + \text{Extraction Type} + \epsilon.$$

where, Plate stands for the physical sequencing plate ID. BH multiple testing correction was applied.

We did not use linear regression to observe cell type expression differences to avoid overfitting of linear regression models, due to the small sample size. For sex, as the independent variable was binary (XX or XY), and the confounding variables were categorical, linear regression estimates showed higher confidence.

Gene and pathway enrichment for sncRNAs of interest was carried out using the Molecular Signatures Database (MSigDB)<sup>58,59</sup>, pathDIP v5.0.32.3<sup>60</sup>, DIANA-miRPath v4.0<sup>61</sup>, and miRPathDB v2.0<sup>62</sup>, where all databases provide different types of miRNA-to-function predictions. MSigDB contains manually-curated lists



of genes that are involved in specific biological functions and diseases, whereas pathDIP, DIANA-miRPath, and miRPathDB all focus on pathways in which the miRNAs have previously been identified/associated.

### SncRNA expression correlation with DNA methylation

Cell type sample-matched DNA methylation normalized beta ( $\beta$ ) values were downloaded for the samples in this study from the GEO accession GSE159526 dataset. Non-variable CpG sites were removed<sup>63,64</sup>. CpGs up to 10 kilobases (kb) upstream or downstream of sncRNA sequences that passed quality control in the first part of this study were selected: 586 CpG sites passed this criterion, corresponding to 77 sncRNAs, with some sncRNAs having multiple probes located within the 20 kb window. Multi-omics integration tools require data to be split into training (70%) and testing (30%) datasets, however, as the number of term samples per cell type were less than 20, we applied Spearman Rank Correlation to test the relationship between sncRNA expression and DNA methylation beta values. A total of 702 sncRNA-CpG pairs were assessed for correlation, and adjusted for multiple testing correction using the BH method at  $p$ -value  $\leq 0.05$ .

## Results

### Landscape of placentally-expressed sncRNAs

As the first exploratory study of the sncRNA transcriptome of sample-matched placental cells and chorionic villi (CV), we first undertook an in-depth examination of sample quality (Supplementary Methods). Sample quality was the main driver of expression variation in raw reads before quality control and normalization, with its effect reducing after data processing (Supplementary Methods Fig. 1). After the QC and normalisation steps, cell type was the highest driver of expression variation.

To explore sncRNAs that may have broad roles in cell and placental maintenance, sncRNAs that were consistently highly expressed in all term samples in all four cell types at RPM  $\geq 1000$  were examined. As the cell-sorted samples were lower quality than bulk CV, a threshold of 1000 reads was chosen instead of the commonly-used threshold of 10,000 reads in similar profiling studies. 188 sncRNAs (1% of the 9485 sncRNAs that passed quality control) satisfied this criterion; of these, 133 also had RPM  $\geq 1000$  in bulk CV (16 miRNAs, 32 piRNA, 7 rRNA, 10 scaRNA, 24 snoRNA, 24 snRNA, and 20 tRNA [Supplementary Table 1, Supplementary Fig. 2]). Expression of these 133 sncRNAs was varied across cell types - when expression for each individual sncRNA was compared between the four cell types, CTB and SC showed the most similar expression (Wilcoxon-Rank-Sum FDR  $> 0.05$  for 100/133 sncRNAs), whereas CTB and HBC had the most dissimilar expression (FDR  $> 0.05$  for 52/133). Interestingly, when only the 16 miRNAs were considered, 7/16 miRNAs had similar expression (FDR  $> 0.05$ ) in CTB and EC, with SC and HBC having the most dissimilar (2/16 with FDR  $> 0.05$ ). Pathway enrichment of the 16 miRNAs highlighted several critical developmental signalling and homeostasis pathways such as fatty acid biosynthesis, proteoglycan functioning, and Hippo signalling (FDR  $\leq 0.05$  using the DIANA-miRPath 3.0 software's DIANA-microT)<sup>65</sup> (Supplementary Table 2), and GO annotations output were for angiogenesis, cell migration, and apoptosis. Surprisingly, hierarchical clustering of these 16 miRNAs (Supplementary Fig. 2) grouped samples by cell type, further underscoring that even for fundamental functional pathways, cell types can exhibit significant expression differences - and in the context of this study, differing levels of gene regulation that can be exerted.

### Known placentally-relevant sncRNAs

We next examined the expression of sncRNA species previously reported to be expressed and relevant in the human placenta. To date, published studies examining roles of sncRNA species in the placenta have been largely limited to the study of miRNAs<sup>12,26,31</sup>; for example, mature miRNAs (most commonly designated with a suffix of either -3p or -5p after the name of the precursor miRNA) of the precursors hsa-miR-100, hsa-miR-16, hsa-miR-155, hsa-miR-193b, hsa-miR-21, hsa-miR-210, and hsa-miR-675 have been previously identified for showing altered expression in placentas from pregnancies of pre-eclampsia (PE), fetal growth restriction (FGR), and/or intrauterine growth restriction (IUGR)<sup>66,67</sup>. Indeed, we found that all these miRNA species showed expression differences between cell types, and even between first trimester and term samples for the same cell type - except for hsa-miR-100-3p, hsa-miR-155-3p, and hsa-miR-16-3p, which were largely undetected in our samples (Fig. 1a).

Three well-studied imprinted miRNA clusters located on chromosome 14 (C14MC), chromosome 19 (C19MC), and the miR-371-3 cluster have been implicated in placental function<sup>31,68</sup>. C14MC miRNAs are most highly expressed during the first trimester of pregnancy, while C19MC miRNAs have the opposite pattern with expression levels peaking in the later stages of pregnancy. Among the 161 total miRNAs lying within these clusters, we found that all 89 C14MC miRNAs, 59 of 64 C19MC miRNAs, and all 8 miRNAs from the miR-371-3 cluster had non-zero expression in most cell types. The exceptions were hsa-miR-371b-3p which had no expression in CTBs and SCs, hsa-miR-371b-5p with zero expression in CTBs, and hsa-miR-372-5p with expression only in HBCs. Interestingly, while the expected expression pattern (slightly higher in term samples compared to the first trimester) was observed for the C19MC miRNAs, the mean expression levels of the C14MC miRNAs showed no difference between trimesters (Fig. 1b); this discrepancy may be attributed to the comparatively lower sample numbers and potentially reduced sample quality of the first trimester samples relative to term. Moreover, when we examined the expression levels per miRNA individually within these clusters (Supplementary Fig. 3), all 89 C14MC had the highest expression in SCs (Wilcoxon-Rank-Sum FDR  $< 0.05$ ). Conversely, for the C19MC miRNAs and the miR-371-2 cluster miRNAs, CTBs and ECs exhibited the highest expression, respectively. For all C19MC miRNAs, HBCs also showed the lowest expression. This difference in expression levels within the cell types and their comparison to the CV expression underscores the importance of cell type consideration, also highlighting how expression within these different cell types changes with gestation.



**Fig. 1.** Expression of known placentially-expressed miRNAs. All reads are per million. **(a)** The expression of 7 miRNAs known to be associated with placental/pregnancy disorders, where the lighter coloured boxplots on the left are first trimester samples, and the darker colour boxplots are term samples. The thick horizontal line in the middle of each boxplot represents the median of the data (Quartile 2). The bottom box represents the first quartile (Q1) i.e., 25% of the data falls below this value, and the top box represents the third quartile (Q3), i.e., 75% of the data falls below this value. The total box represents the interquartile range (IQR), which is the middle 50% of the data (from Q1 to Q3). The vertical line extending from the middle of the boxplot represents the total spread of the data, with outliers as independent points. **(b)** Combined expression of the 89 C14MC, 59 C19MC, and 8 miR-371-3 cluster miRNAs in first trimester and term samples in the different cell types and matched whole villi.

The average expression across all cell types for these miRNA clusters roughly matched the average expression of the CV samples; when expression of all cell types was combined and compared with that of villi, no statistically significant difference in expression was observed overall (hsa-miR-100-3p and hsa-miR-100-5p being the exception, where expression in CV was higher). CTB expression did not match closely with that of CV, but it should be noted that the major component of the CV samples are STBs, and that we were not able to isolate these cells by FACS due to their presence as a continuous syncytium. It is possible that STBs have a very different sncRNA profile than their CTB progenitors, despite their DNAm profile being quite similar. This further highlights the need for and importance of considering cell composition when studying the placenta, as sampled placental tissue may in fact not represent the true expression landscape.

Placental cell type-associated sncRNA profiles

To determine if the placental sorted-cells had cell type-specific expression for sncRNAs, we investigated each cell type independently. As we had fewer first trimester samples ( $n < 8$  for each cell type), and to maintain consistency, we restricted this analysis to only term samples ( $n = 71$ ).

No sncRNAs showed unique expression in a single cell type. We thus sought to investigate cell type-associated expression by considering sncRNAs with a read count of  $\geq 2$  in at least 75% of term samples (i.e., expressed in at least 9/12 CTB = 1,836 sncRNAs; 12/16 SC = 1,879 sncRNAs; 9/12 EC = 1,825 sncRNA; 10/14 HBC = 2,081 sncRNAs) for each cell type. A proportion of the above selected sncRNAs – 115 sncRNAs – showed significantly different mean expression in one specific cell type when compared to the mean expression of the other cell types (Kruskal-Wallis BH  $p$ -value  $\leq 0.05$ ), and were classified as cell type-associated sncRNAs (Table 2; Fig. 2a,b, Supplementary Table 3).

These cell type-associated sncRNAs were spread across the genome (Table 2; Fig. 2c). Surprisingly, none of the 103 cell type-associated miRNAs we identified belonged to either the placental-associated miRNA clusters on chromosomes 14 or 19. Of the 103 cell type-associated miRNAs, 39 have previously been reported in the literature as having some relation (pathogenic or not) with the human placenta; the remaining 64 miRNAs and 12 non-miRNAs are, to our knowledge, novel associations (Supplementary Table 4). We also did not find any piRNAs that overlapped between this and our previous study<sup>19</sup> which investigated piRNAs in bulk placental tissue; however, the placental extraction process and analysis methodology differs between these two studies. Furthermore, the lower quality of the sorted-cell data may have missed species that are expressed at lower levels; however, our filtering thresholds were generous in attempt to circumvent this limitation.

HBC sncRNA expression was the most distinct as compared to all other cell types, attesting to the unique embryonic origin of HBCs as compared to the other cell types (Fig. 2a). Seven of the 78 HBC-associated miRNAs mapped to the X chromosome (hsa-miR-20b-3p, hsa-miR-1277-5p, hsa-miR-652-5p, hsa-miR-676-3p, hsa-miR-1468-5p, hsa-miR-548ax, hsa-miR-651-5p), and these were the only cell-associated sncRNAs that were detected on the X chromosome (Fig. 2c). Given their presence on the X chromosome, we also compared the expression between sexes (placentas from male and female fetuses) for these seven miRNAs, however none showed significant expression differences by sex.

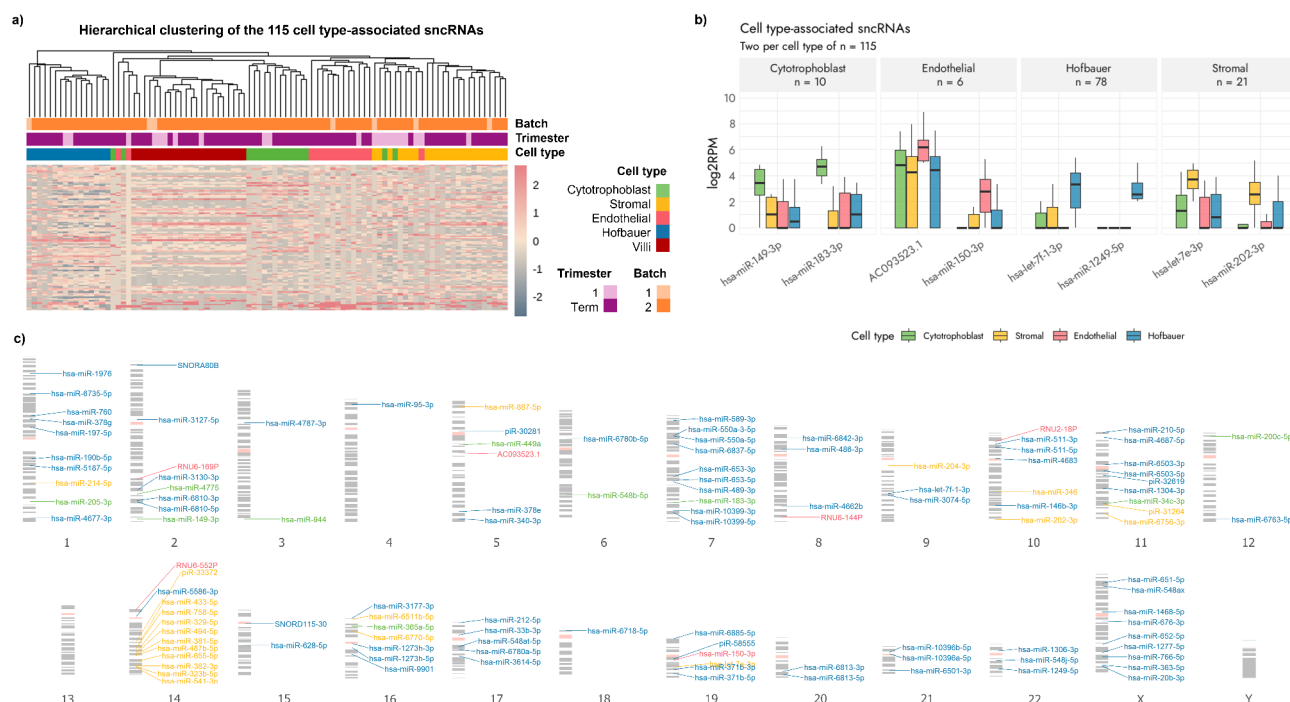
Shared miRNAs between placental cytotrophoblasts and cancer

Gene expression of both coding genes and small RNAs in the placenta has been shown to overlap that of cancer, likely due to the shared invasive and stem-like properties of CTBs and cancer cells<sup>69</sup>. We thus wanted to examine whether any of the CTB-associated sncRNAs that we discovered had previously been identified in cancer.

Interestingly, all 10 CTB-associated sncRNAs (all miRNAs) (hsa-miR-149-3p, hsa-miR-183-3p, hsa-miR-200c-5p, hsa-miR-205-3p, hsa-miR-34c-3p, hsa-miR-365a-5p, hsa-miR-449a, hsa-miR-4775, hsa-miR-548b-5p, hsa-miR-944) have been associated with cancer, most commonly lung, breast, and colorectal cancer<sup>70–72</sup>. All 10 miRNAs were also significant ( $FDR \leq 0.05$ ) for being included in 13 of the 50 Hallmark Gene Sets within the Molecular Signatures Database (MSigDB)<sup>58,59</sup>, a database that contains manually curated gene set lists for several disease conditions (Fig. 3). Furthermore, when examined in The Cancer Genome Atlas (TCGA) cohorts, all 10 miRNAs showed expression differences (both up- and downregulated depending upon the miRNA) between malignant and non-malignant samples (Supplementary Fig. 4). Expression of oncogenic miRNAs is generally upregulated in cancers, where high expression of these oncomiRs leads to suppression of key target tumour suppressor genes<sup>73</sup>.

Total cell type-associated sncRNAs ( $n = 115$ )	Cytotrophoblast $n = 10$	Stromal $n = 21$	Endothelial $n = 6$	Hofbauer $n = 78$
miRNA ( $n = 103$ )	10	19	1	73
piRNA ( $n = 5$ )	0	2	0	3
snRNA ( $n = 4$ )	0	0	4	0
snoRNA ( $n = 3$ )	0	0	1	2
Genomic position of cell type-associated sncRNAs				
Exonic ( $n = 17$ )	2	4	0	11
Intronic ( $n = 39$ )	4	11	5	19
Intergenic ( $n = 59$ )	4	6	1	48

**Table 2.** Cell type-associated placental SncRNAs. Distribution of the cell type-associated sncRNAs( $n = 115$ ) by species and genomic position in all term placenta cell type samples.



**Fig. 2.** Cell type-associated sncRNA expression. **(a)** Clustering heatmap (Z-scores of log<sub>2</sub>-transformed normalized expression counts) of the 115 cell type-associated sncRNAs. **(b)** Expression of two examples each of cell type-associated sncRNAs detected per cell type. *n* represents the number of sncRNAs designated as cell type-associated for that specific cell type. The thick horizontal line in the middle of each boxplot represents the median of the data (Quartile 2). The bottom box represents the first quartile (Q1) i.e., 25% of the data falls below this value, and the top box represents the third quartile (Q3), i.e., 75% of the data falls below this value. The total box represents the interquartile range (IQR), which is the middle 50% of the data (from Q1 to Q3). The vertical line extending from the middle of the boxplot represents the total spread of the data, with outliers as independent points. **(c)** Chromosomal location of each of the 115 cell type-associated sncRNAs.

## Sex-influenced profiles of placental cell types

Sex chromosome complement (XX or XY) can also contribute towards inter-individual variability, contributing towards expression differences in genes located on both the autosomes as well as the sex chromosomes<sup>74–76</sup>. We investigated whether placental sex was associated with sncRNA expression in term placental cell type samples.

When only CV samples were considered, no autosomal or sex-linked sncRNAs showed significantly differing expression by sex. Within the sorted-cell samples at a BH *p*-value  $\leq 0.05$ , three piRNAs (piR-31637, piR-31638, piR-35551) showed higher expression in XX as compared to XY CTB samples, while higher expression was observed for one snoRNA (SNORD116-22) in XY SC samples and one miRNA (hsa-miR-1249-3p) in XY EC samples (Fig. 4a) on the autosomes. Hsa-miR-1249-3p has previously been identified as contributing towards angiogenesis<sup>77</sup>.

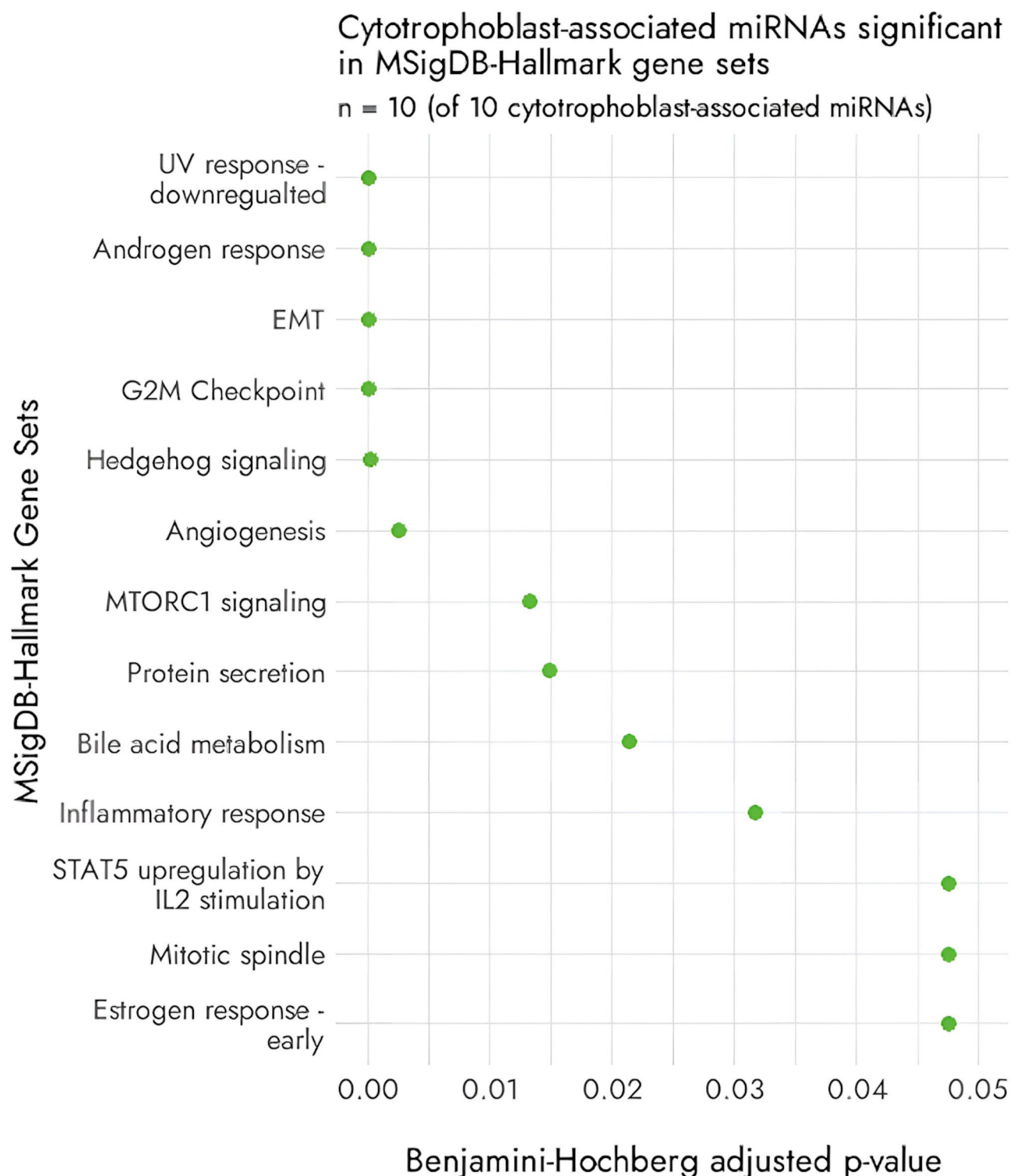
When only the X and Y chromosomes were considered, three X-linked miRNAs (hsa-miR-221-5p, hsa-miR-222-3p, hsa-miR-424-3p) had higher expression in XY samples as compared to XX and one snoRNA (RNU6-854P) had higher expression in XX CTB samples (BH *p*-value  $\leq 0.05$ ) (Fig. 4b). Of the three X-linked miRNAs, hsa-miR-221 and hsa-miR-424 have previously been reported in the context of placental and/or pregnancy disorders<sup>78–80</sup>. None of the four Y-linked miRNAs (hsa-miR-3690-2, hsa-miR-6089-2, hsa-miR-9985, hsa-miR-12120) that had passed quality control showed differing expression by sex.

## Assessing correlation between cell type-associated sncRNAs and DNA methylation

We previously comprehensively cataloged the DNA methylation profiles of the same sorted-cell samples used in this study with the Illumina MethylationEPIC array<sup>36</sup>. DNA methylation, similar to sncRNAs, plays an essential role in the epigenetic regulation of gene expression. In many instances, these two regulatory mechanisms work synergistically to influence downstream gene expression<sup>81,82</sup>.

To assess any correlation that may exist between sncRNA expression levels and DNA methylation, previously processed sample-matched DNA methylation data was considered<sup>36</sup>. The expression of the 115 cell type-associated sncRNAs was measured against CpGs that fell within 10Kb upstream of the 5' end and 10Kb downstream of the 3' end; this basepair window differs by study<sup>83,84</sup> and we chose 10 kb to assess any cis-interactions. 400 CpG probes fell within these 20Kb windows for 77 of the 115 cell type-associated sncRNAs, and for some sncRNAs there was more than one sncRNA-CpG pair (there were multiple CpGs present within the 20Kb window for 63 miRNAs). When normalized sncRNA RPM expression was compared with normalized DNA methylation beta values, no sncRNA-CpG pairs were significantly correlated at a BH *p*-value  $\leq 0.05$  in

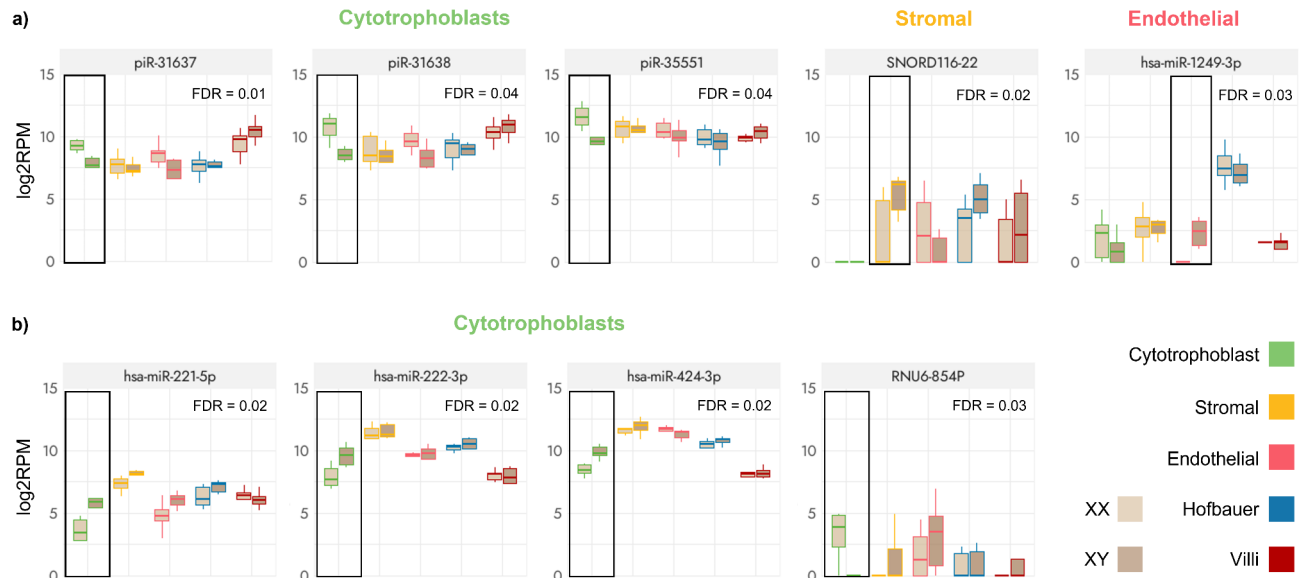




**Fig. 3.** Pathways predicted for the 10 cytotrophoblast cell-associated miRNAs. Gene sets for which all 10 cytotrophoblast cell-associated miRNAs were significant at a BH  $p$ -value  $\leq 0.05$ .

term samples. Albeit not reaching the established threshold of multiple-testing correction significance, we did observe high correlation ( $R^2 > 0.6$  or  $R^2 < -0.6$ ) between six CTB, two SC, and two HBC sncRNA-CpG pairs ( $n = 10$ ) (Supplementary Fig. 5, Supplementary Table 5). We did not observe correlation between global DNAm and piRNA values when analysed per cell type (data not shown in manuscript).

When only term CV samples were considered, one sncRNA-CpG pair had  $R^2 > 0.8$  (BH  $p$ -value = 0.65): hsa-miR-512-3p (intergenic) and cg27369447 (3' UTR of ZNF665) on chromosome 19 (Supplementary Fig. 5). Hsa-



**Fig. 4.** SncRNAs showing variable expression by sex in term sorted-cell placental samples. **(a)** Autosomal sncRNAs which showed differing expression by sex in each cell type. The lighter coloured boxplots on the left are XX samples and the darker coloured boxplots are XY samples. The thick horizontal line in the middle of each boxplot represents the median of the data (Quartile 2). The bottom box represents the first quartile (Q1) i.e., 25% of the data falls below this value, and the top box represents the third quartile (Q3), i.e., 75% of the data falls below this value. The total box represents the interquartile range (IQR), which is the middle 50% of the data (from Q1 to Q3). The vertical line extending from the middle of the boxplot represents the total spread of the data, with outliers as independent points. **(b)** X-linked sncRNAs with differing expression by sex in cytotrophoblast sorted-cell samples.

miR-512-3p belongs to the placental-exclusive, paternally-expressed C19MC<sup>85,86</sup>. When CpGs and miRNAs lying only within the C14MC (14q32.31) and C19MC (19q13.42) regions were separately examined in term villi samples, only one pair within C19MC passed the significance threshold (BH  $p$ -value=0.00) and had  $R^2 < 0.8$  - hsa-miR-523-3p and cg24815529 (Supplementary Fig. 5). Hsa-miR-523-3p has been previously shown to be downregulated in PE<sup>87</sup>. Lastly, despite not reaching multiple testing significance, 28 miRNA-CpG pairs on 14q32.31 and 10 pairs on 19q13.42 had high correlation values ( $R^2 > 0.8$  or  $R^2 < -0.8$ ) (Supplementary Table 5).

## Discussion

We characterized the small non-coding RNA profiles of four major placental cell types (cytotrophoblasts, stromal cells, endothelial cells, and Hofbauer cells) and their matching chorionic villi samples in first trimester and term placentas. To our knowledge, this is the first genome-wide exploratory profiling of placental cell-specific sncRNAs. Given sample quality limitations and the variable nature of the data, we provide a more descriptive analysis of sncRNA expression in the placenta, with efforts to correct for this variation where possible. While no sncRNAs showed unique expression in just one cell type, certain sncRNAs showed significantly different mean expression in one cell type as compared to the other cell types, which we termed as 'cell type-associated sncRNAs'. These 115 cell type-associated sncRNAs were distributed across chromosomes, and differences in their expression profiles were also indicated by PCA and hierarchical clustering.

HBCs - macrophages that play a role in both immune response and angiogenesis<sup>5,88</sup> - showed the highest number of cell type-associated sncRNAs ( $n=78$ ), and target prediction for these HBC-associated sncRNAs identified similar functional pathways (Supplementary Fig. 6). Interestingly, the expression levels of all 78 HBC-associated sncRNAs was distinct from the other cell types, indicative of their unique developmental lineage; HBC samples also grouped in their own top-level cluster separate from CVs, CTBs, SCs and ECs (Fig. 2a). This is similar to DNAm patterns observed for these same cells and reflect the embryonic origin of HBCs<sup>36</sup>.

The sncRNA profile of CTBs most closely matched that of bulk placental villi, as is true of DNAm profiles, reflecting the large TB contribution to bulk villi<sup>36,84</sup>. A limitation of our study is that it did not investigate STB which is derived from CTBs (and represents the major cell type of the bulk placenta) due to the inability to sort this cell type by FACS as its composition of a continuous syncytium. We thus cannot assess whether the differences between CTB and CV were due to the exclusion of this cell type or due to technical factors, i.e. differences in sample quality. The CTB-associated sncRNAs we identified were significantly enriched for cancer-related pathways; trophoblasts fundamentally invade into the maternal endometrium to anchor the placenta and establish blood flow, and this invasive and proliferative nature is often likened to cancer cells<sup>69</sup>. These 10 CTB-associated sncRNAs, which were all miRNAs, have been previously reported to be differentially expressed in several placental disorders<sup>89,90</sup> and cancer subtypes<sup>91–94</sup>, providing further support in the shared characteristics of placental growth and cancer.

ECs and SCs are both derived from the extraembryonic mesoderm and have shown the most similar DNA methylation patterns than other cell types<sup>36</sup>; however, we did not find sncRNA expression to be more closely related in these relative to the other cell types. SCs are known to be important in structural maintenance and immunomodulation, and the predicted biological pathways for the 21 SC-associated sncRNAs we discovered reported the above functions (Supplementary Fig. 6), with pathways such as *MTORC1* and cell cycle checkpoint signalling<sup>95,96</sup>. Interestingly, half (11/21) of the SC-associated sncRNAs were located on chromosome 14, with the majority of these being intergenic (9/11). The six EC-associated sncRNAs (here, sampled from small chorionic villi vessels that are found in the mesenchymal core) were enriched for pathways significant for cell communication and signalling (Supplementary Fig. 6), also consistent with previous findings<sup>97,98</sup>.

Three known placental miRNA clusters, the imprinted C14MC, C19MC, and miR-371-3 cluster, each showed differing expression by cell type, confirming previous reports<sup>85,99</sup>. In all 89 C14MC miRNAs, SCs had the highest expression, whereas CTBs had higher expression for most of the 58 C19MC miRNAs, and ECs had the highest expression for all eight miRNAs belonging to the miR-371-3 cluster (the expression difference between these cell types and others was not substantial for them to be classified as cell type-associated, however). Surprisingly, even though the C19MC (19q13.41) and miR-371-3 (19q13.42) clusters only lie about 30 kb apart, the cell-specific expression was distinct in these two clusters, as alluded in previous studies<sup>99</sup>. While the exact functions of these clusters is not fully known, their expression has markedly distinguished the placenta from other organs<sup>100</sup>, and has often observed to differ in cancer<sup>101,102</sup>. Altered levels of C19MC miRNAs have shown to impact trophoblast development<sup>79,85,103,104</sup>, and C14MC miRNAs have shown high expression in CV and chorionic plate-derived SCs as compared to decidual basal plate-derived cells<sup>99</sup>. Placental miRNAs are also known to enter the maternal bloodstream encased in extracellular vesicles (EVs)<sup>105</sup>, and C19MC miRNAs have been reported to be the predominant miRNAs in placenta-secreted (predominantly CTB-secreted) exosomes in maternal serum<sup>106</sup>. They may thus have potential utility as specialized serum biomarkers for pregnancy complications (such as gestational diabetes mellitus (GDM), gestational hypertension, PE, FGR) should their expression patterns be capable of distinguishing healthy from pathogenic samples associated with these conditions<sup>107,108</sup>.

We also compared the expression between XX and XY placentas, separately in the autosomes and sex chromosomes. Five autosomal sncRNAs (three piRNAs in CTBs, one snoRNA in SCs, and one miRNA in ECs) and four X-linked sncRNAs (three miRNAs and one snRNA, all in CTBs), had significantly differing expression between the sexes. That differences exist in placentas carrying male fetuses compared to female fetuses is known; on average, placentas from male fetuses weigh more and show higher expression of immune-related genes, and carrying male fetuses has an increased risk for placental complications<sup>109</sup>. While still limited, our results build on the evidence in expression differences observed between the sexes<sup>110</sup>.

While we did not find a strong correlation between sncRNA expression and DNA methylation profiles within our matched samples, crosstalk between the two features does occur<sup>81,84,111</sup>. The lower sample size and power of this study may have been a reason for none of the sncRNA-CpG pairs passing multiple-testing correction. Given the rarity of acquiring isolated placental cells, the use of instruments and methodologies that incorporate multi-omic measurements in future placental studies would allow for maximal data generation from these scarce tissues (for example, using technology such as the Oxford Nanopore MinION, which allows for sequencing and DNA methylation measurement simultaneously).

Advancements in approaches to the placenta are steadily growing. Placental tissue modelling using organ-on-a-chip platforms and organoids have been useful to better understand placental growth and cell type differences<sup>112–114</sup>. However, due to the complexity of this tissue and the dynamic interactions between maternal, placental, and fetal cell types, these models have certain drawbacks. While organ-on-a-chip has limitations in the 3D recapitulation and cell-type spatial interactions, organoid models are currently derived from CTBs and SCT, with ongoing efforts to incorporate other cell type. As opposed to the above methods which use cultured tissue for in-vitro and in-vivo modelling (growth and expression of which is highly dependent on the artificial culture conditions), our study investigated primary, non-cultured tissue. Our results provide data which would be of benefit for future cell-type studies in both bulk tissue and organoid models. However, validating our findings and evaluating effects of gestational age and perinatal parameters in an independent cohort will be needed.

The novelty of this study is three-fold: first - we assessed the expression of sncRNAs in four major placental cell types, second - we concurrently examined expression in sample-matched placental cell and whole villi samples, and third - we profiled several understudied sncRNA species in the placenta. Our results show that inter-placental and inter-cell type variability exists for several sncRNAs (as show in Figs. 2a, 4a), in particular for known placental-expressed miRNAs. While studies of sncRNA species in one cell type (for example, miRNA expression in TBs) have been highly informative in discovering sncRNAs associated with specific conditions (such as hsa-miR-16, hsa-miR-21 and hsa-miR-210 in PE and hsa-miR-100b in IUGR<sup>115,116</sup>), the magnitude (level of expression) and direction of expression (over or under expression) of the sncRNA profiles between and within placental cell types has generally not been analyzable in the majority of studies. Furthermore, gene expression differences are also known to exist between placental and other extraembryonic tissues such as the umbilical cord, amnion, and decidua<sup>117,118</sup>, providing a future avenue of investigation.

A limitation of our study was that the cell-sorting procedure involved several time-intensive processing steps which may have impacted cell quality (i.e., due to stress that the samples may have incurred). Also, the low overall sample yield after cell-sorting could have produced a higher number of artefact sequences. However, all sorted-cells were processed sequentially from the same sample, and care was taken to be consistent for all samples. To furthermore alleviate low sample quality, our designation of cell type-associated sncRNAs was based on non-stringent selection, in an effort to perform a generous profiling estimation. The lower number of first trimester and the non-matching number of samples for each cell type also led to unbalanced sample numbers, however we tried alleviating this by examining each cell type, and first trimester and term samples separately. The sequencing and alignment protocol was also geared towards miRNAs, and as such, all sncRNA species were

analysed per species to ensure that species-dependant variability was controlled for. Lastly, a very low percentage of previous sncRNA associations have proven to be reproducible to date<sup>28</sup> due to small sample sizes, varied processing procedures, the downstream effects of non-standardized genome alignment, data quality control, and analysis pipelines - we highlight how technical variables also influence expression levels and the need in controlling for their effects for reliable expression profiling. Other biological variables, such as maternal BMI and age at delivery, may also contribute towards the variation in expression observed, but we were not able to assess any effects that these additional variables may exert due to lack of information.

## Conclusion

To our knowledge, this is the first exploratory study of the small non-coding RNA transcriptome of the human placenta and its cell types. This study provides insight into the sncRNA profiles of four isolated placental cell types and whole chorionic villi in sample-matched placentas from both the first trimester and term samples and is meant to provide a preliminary resource to the placental community. Our results summarize that certain sncRNAs are expressed more distinctly in one cell type in comparison to others, and build on the proposed functions of these cell types from previous studies. Overall, the profile of bulk placental villi was distinct from the individual cell types, underscoring that whole tissue expression is not representative of individual cell type expression. This data is a novel resource for furthering our understanding of placental cell type heterogeneity and its contributions to regulatory mechanisms in the healthy placenta.

## Data availability

Raw and processed expression data and sample metadata can be accessed on GEO at accession number GSE288435. Upon a reasonable request, the corresponding authors of this article will provide unrestricted access to the original data.

Received: 3 February 2025; Accepted: 15 April 2025

Published online: 26 April 2025

## References

- Mercuri, N. D. & Cox, B. J. The need for more research into reproductive health and disease. *eLife* **11**, e75061 (2022).
- Turco, M. Y. & Moffett, A. Development of the human placenta. *Dev. Camb. Engl.* **146**, dev163428 (2019).
- Wang, Y. & Zhao, S. Cell types of the placenta. In *Vascular Biology of the Placenta* (Morgan & Claypool Life Sciences, 2010).
- Martin et al. Human placenta and trophoblast development: key molecular mechanisms and model systems. *Cell. Mol. Life Sci.* **76**, 3479–3496 (2019).
- Reyes, L. & Golos, T. G. Hofbauer cells: their role in healthy and complicated pregnancy. *Front. Immunol.* **9**, 2628 (2018).
- Czech, B., Hannon, G. J. & Small RNA sorting: matchmaking for argonautes. *Nat. Rev. Genet.* **12**, 19–31 (2011).
- Pratt, A. J. & MacRae, I. J. The RNA-induced Silencing complex: A versatile Gene-silencing machine. *J. Biol. Chem.* **284**, 17897–17901 (2009).
- Esteller, M. Non-coding RNAs in human disease. *Nat. Rev. Genet.* **12**, 861–874 (2011).
- Qu, Z. & Adelson, D. L. Evolutionary conservation and functional roles of NcRNA. *Front. Genet.* **3**, 205 (2012).
- Watson, C. N., Belli, A. & Di Pietro, V. Small Non-coding RNAs: new class of biomarkers and potential therapeutic targets in neurodegenerative disease. *Front. Genet.* **10**, 364 (2019).
- Higashijima, A. et al. Characterization of placenta-specific MicroRNAs in fetal growth restriction pregnancy. *Prenat Diagn.* **33**, 214–222 (2013).
- Awamleh, Z., Gloor, G. B. & Han, V. K. M. Placental MicroRNAs in pregnancies with early onset intrauterine growth restriction and preeclampsia: potential impact on gene expression and pathophysiology. *BMC Med. Genomics.* **12**, 91 (2019).
- Inno, R., Kikas, T., Lillepea, K. & Laan, M. Coordinated expressional landscape of the human placental mirnome and transcriptome. *Front. Cell. Dev. Biol.* **9**, 697947 (2021).
- Timofeeva, A. V. et al. Diagnostic role of Cell-Free MiRNAs in identifying placenta accreta spectrum during First-Trimester screening. *Int. J. Mol. Sci.* **25**, 871 (2024).
- Morales-Prieto, D. M., Ospina-Prieto, S., Chaiwangyen, W., Schoenleben, M. & Markert, U. R. Pregnancy-associated miRNA-clusters. *J. Reprod. Immunol.* **97**, 51–61 (2013).
- Paquette, A. G. et al. Distinct communication patterns of trophoblastic MiRNA among the maternal-placental-fetal compartments. *Placenta* **72–73**, 28–35 (2018).
- Shang, R., Lee, S., Senavirathne, G. & Lai, E. C. MicroRNAs in action: biogenesis, function and regulation. *Nat. Rev. Genet.* **24**, 816–833 (2023).
- Czech, B. et al. piRNA-Guided genome defense: from biogenesis to Silencing. *Annu. Rev. Genet.* **52**, 131–157 (2018).
- Martinez, V. D. et al. Human placental piwi-interacting RNA transcriptome is characterized by expression from the DLK1-DIO3 imprinted region. *Sci. Rep.* **11**, 14981 (2021).
- Zhu, C. et al. Erroneous ribosomal RNAs promote the generation of antisense ribosomal siRNA. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 10082–10087 (2018).
- Huang, Z. H., Du, Y. P., Wen, J. T., Lu, B. F. & Zhao, Y. SnoRNAs: functions and mechanisms in biological processes, and roles in tumor pathophysiology. *Cell. Death Discov.* **8**, 259 (2022).
- Morais, P., Adachi, H. & Yu, Y. T. Spliceosomal SnRNA epitranscriptomics. *Front. Genet.* **12**, 652129 (2021).
- Berg, M. D. & Brandl, C. J. Transfer RNAs: diversity in form and function. *RNA Biol.* **18**, 316–339.
- Raina, M. & Ibba, M. tRNAs as regulators of biological processes. *Front. Genet.* **5**, 171 (2014).
- Martinez, V. D. et al. Profiling the small non-coding RNA transcriptome of the human placenta. *Sci. Data.* **8**, 166 (2021).
- Gong, S. et al. The RNA landscape of the human placenta in health and disease. *Nat. Commun.* **12**, 2639 (2021).
- Telkar, N. et al. Small Non-Coding RNAs in the human placenta: regulatory roles and clinical utility. *Front. Genet.* **13**, 868598 (2022).
- Yong, H. E. J. & Chan, S. Y. Current approaches and developments in transcript profiling of the human placenta. *Hum. Reprod. Update.* **26**, 799–840 (2020).
- Liu, Z. et al. Resolving the gene expression maps of human first-trimester chorionic villi with Spatial transcriptome. *Front. Cell. Dev. Biol.* **10**, 1060298 (2022).
- Sadovsky, Y., Mouillet, J. F., Ouyang, Y., Bayer, A. & Coyne, C. B. The function of trophomirs and other MicroRNAs in the human placenta. *Cold Spring Harb Perspect. Med.* **5**, a023036 (2015).
- Morales-Prieto, D. M. et al. MicroRNA expression profiles of trophoblastic cells. *Placenta* **33**, 725–734 (2012).



32. Yu, X., Abbas-Aghababazadeh, F., Chen, Y. A. & Fridley, B. L. Statistical and bioinformatics analysis of data from bulk and Single-Cell RNA sequencing experiments. *Methods Mol. Biol. Clifton NJ*. **2194**, 143–175 (2021).
33. Campbell, K. A. et al. Placental cell type Deconvolution reveals that cell proportions drive preeclampsia gene expression differences. *Commun. Biol.* **6**, 264 (2023).
34. Pavličev, M. et al. Single-cell transcriptomics of the human placenta: inferring the cell communication network of the maternal-fetal interface. *Genome Res.* **27**, 349–361 (2017).
35. Wang, Q. et al. Single-cell transcriptional profiling reveals cellular and molecular divergence in human maternal-fetal interface. *Sci. Rep.* **12**, 10892 (2022).
36. Yuan, V. et al. Cell-specific characterization of the placental methylome. *BMC Genom.* **22**, 6 (2021).
37. Daenekas, B. et al. Conumee 2.0: enhanced copy-number variation analysis from DNA methylation arrays for humans and mice. *Bioinforma Oxf. Engl.* **40**, btac029 (2024).
38. Hagemann-Jensen, M., Abdullayev, I., Sandberg, R. & Faridani, O. R. Small-seq for single-cell small-RNA sequencing. *Nat. Protoc.* **13**, 2407–2424 (2018).
39. Andrews, S. & FastQC: A quality control tool for high throughput sequence data. (2010).
40. Fehlmann, T. et al. MiRMaster 2.0: multi-species non-coding RNA sequencing analyses at scale. *Nucleic Acids Res.* **49**, W397–W408 (2021).
41. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinforma Oxf. Engl.* **29**, 15–21 (2013).
42. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2021).
43. Zhang, Y., Parmigiani, G. & Johnson, W. E. ComBat-seq: batch effect adjustment for RNA-seq count data. *NAR Genomics Bioinforma.* **2**, lqaa078 (2020).
44. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. EdgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinforma Oxf. Engl.* **26**, 139–140 (2010).
45. Love, M. I., Huber, W. & Anders, S. Moderated Estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
46. Yuan, V. plomics: Miscellaneous functions, mainly for DName analysis. (2022).
47. Kozomara, A., Birgaoanu, M. & Griffiths-Jones, S. MiRBase: from MicroRNA sequences to function. *Nucleic Acids Res.* **47**, D155–D162 (2019).
48. Wang, J. et al. PiRBase: integrating PiRNA annotation in all aspects. *Nucleic Acids Res.* **50**, D265–D272 (2022).
49. The RNAcentral Consortium. RNAcentral: a hub of information for non-coding RNA sequences. *Nucleic Acids Res.* **47**, D221–D229 (2019).
50. Stelzer, G. et al. The genecards suite: from gene data mining to disease genome sequence analyses. *Curr. Protoc. Bioinforma.* **54**, 1301–13033 (2016).
51. Barshir, R. et al. GeneCaRNA: A comprehensive Gene-centric database of human Non-coding RNAs in the genecards suite. *J. Mol. Biol.* **433**, 166913 (2021).
52. Seal, R. L. et al. Genenames.org: the HGNC resources in 2023. *Nucleic Acids Res.* **51**, D1003–D1009 (2023).
53. Chan, P. P. & Lowe, T. M. GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Res.* **44**, D184–189 (2016).
54. Kent, W. J. et al. The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
55. Martin, F. J. et al. Ensembl 2023. *Nucleic Acids Res.* **51**, D933–D941 (2023).
56. Smith, M. D. et al. Large-scale transcriptome-wide profiling of MicroRNAs in human placenta and maternal plasma at early to mid gestation. *RNA Biol.* **0**, 1–14 (2021).
57. Ritchie, M. E. et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
58. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U S A.* **102**, 15545–15550 (2005).
59. Liberzon, A. et al. The molecular signatures database (MSigDB) hallmark gene set collection. *Cell. Syst.* **1**, 417–425 (2015).
60. Rahmati, S. et al. PathDIP 4: an extended pathway annotations and enrichment analysis resource for human, model organisms and domesticated species. *Nucleic Acids Res.* **48**, D479–D488 (2020).
61. Tastsoglou, S. et al. DIANA-miRPath v4.0: expanding target-based MiRNA functional analysis in cell-type and tissue contexts. *Nucleic Acids Res.* **51**, W154–W159 (2023).
62. Kehl, T. et al. MiRPathDB 2.0: a novel release of the MiRNA pathway dictionary database. *Nucleic Acids Res.* **48**, D142–D147 (2020).
63. Edgar, R. D., Jones, M. J., Robinson, W. P. & Kobor, M. S. An empirically driven data reduction method on the human 450K methylation array to remove tissue specific non-variable CpGs. *Clin. Epigenetics.* **9**, 11 (2017).
64. Dieckmann, L. et al. Reliability of a novel approach for reference-based cell type Estimation in human placental DNA methylation studies. *Cell. Mol. Life Sci. CMLS.* **79**, 115 (2022).
65. Tastsoglou, S. et al. DIANA-microT 2023: including predicted targets of virally encoded MiRNAs. *Nucleic Acids Res.* **51**, W148–W153 (2023).
66. Ali, A. et al. MicroRNA-mRNA networks in pregnancy complications: A comprehensive downstream analysis of potential biomarkers. *Int. J. Mol. Sci.* **22**, 2313 (2021).
67. Tsochandaridis, M., Nasca, L., Toga, C. & Levy-Mozziconacci, A. Circulating microRNAs as clinical biomarkers in the predictions of pregnancy complications. *BioMed Res. Int.* **2015** 294954. (2015).
68. Légaré, C. et al. Human plasma pregnancy-associated MiRNAs and their Temporal variation within the first trimester of pregnancy. *Reprod. Biol. Endocrinol. RBE.* **20**, 14 (2022).
69. Holtan, S. G., Creedon, D. J., Haluska, P. & Markovic, S. N. Cancer and pregnancy: parallels in growth, invasion, and immune modulation and implications for Cancer therapeutic agents. *Mayo Clin. Proc.* **84**, 985–1000 (2009).
70. Sherafatian, M. & Arjmand, F. Decision tree-based classifiers for lung cancer diagnosis and subtyping using TCGA MiRNA expression data. *Oncol. Lett.* **18**, 2125–2131 (2019).
71. Naoem, L. D., Muthaiyan, M. & Venkatesan, A. Identification of dysregulated MiRNAs in triple negative breast cancer: A meta-analysis approach. *J. Cell. Physiol.* **234**, 11768–11779 (2019).
72. Zhao, B. et al. A novel miRNA-based signature as predictive tool of survival outcome of colorectal cancer patients. *Chem. Biol. Drug Des.* **102**, 1024–1033 (2023).
73. Frixa, T., Donzelli, S. & Blandino, G. Oncogenic MicroRNAs: key players in malignant transformation. *Cancers* **7**, 2466–2485 (2015).
74. Oliva, M. et al. The impact of sex on gene expression across human tissues. *Science* **369**, eaba3066 (2020).
75. Gershoni, M. & Pietrokovski, S. The landscape of sex-differential transcriptome and its consequent selection in human adults. *BMC Biol.* **15**, 7 (2017).
76. Lopes-Ramos, C. M. et al. Sex differences in gene expression and regulatory networks across 29 human tissues. *Cell. Rep.* **31**, 107795 (2020).
77. Chen, X. et al. P53-induced miR-1249 inhibits tumor growth, metastasis, and angiogenesis by targeting VEGFA and HMGA2. *Cell. Death Dis.* **10**, 131 (2019).

78. Yang, Y. et al. MiR-221-3p is down-regulated in preeclampsia and affects trophoblast growth, invasion and migration partly via targeting thrombospondin 2. *Biomed. Pharmacother. Biomedicine Pharmacother.* **109**, 127–134 (2019).
79. Mouillet, J. F. et al. Transgenic expression of human C19MC MiRNAs impacts placental morphogenesis. *Placenta* **101**, 208–214 (2020).
80. Wang, W. J. et al. Exploring fetal sex dimorphism in the risk factors of gestational diabetes Mellitus-A prospective cohort study. *Front. Endocrinol.* **10**, 848 (2019).
81. Glaich, O. et al. DNA methylation directs MicroRNA biogenesis in mammalian cells. *Nat. Commun.* **10**, 5657 (2019).
82. Saito, J. et al. Association between DNA methylation in the miR-328 5'-flanking region and inter-individual differences in miR-328 and BCRP expression in human placenta. *PloS One.* **8**, e72906 (2013).
83. Varghese, R. S. et al. Integrative analysis of DNA methylation and MicroRNA expression reveals mechanisms of Racial heterogeneity in hepatocellular carcinoma. *Front. Genet.* **12**, 708326 (2021).
84. Saviana, M. et al. Crosstalk between MiRNAs and DNA methylation in Cancer. *Genes* **14**, 1075 (2023).
85. Malnou, E. C., Umlauf, D., Mouysset, M. & Cavaillé, J. Imprinted MicroRNA gene clusters in the evolution, development, and functions of mammalian placenta. *Front. Genet.* **9**, 706 (2018).
86. Noguer-Dance, M. et al. The primate-specific MicroRNA gene cluster (C19MC) is imprinted in the placenta. *Hum. Mol. Genet.* **19**, 3566–3582 (2010).
87. Ning, H. & Tao, H. Small RNA sequencing of Exosomal MicroRNAs reveals differential expression of MicroRNAs in preeclampsia. *Med. (Baltim).* **102**, e35597 (2023).
88. Zulu, M. Z., Martinez, F. O., Gordon, S. & Gray, C. M. The elusive role of placental macrophages: the Hofbauer cell. *J. Innate Immun.* **11**, 447–456 (2019).
89. Chim, S. S. C. et al. Detection and characterization of placental MicroRNAs in maternal plasma. *Clin. Chem.* **54**, 482–490 (2008).
90. Mayor-Lynn, K., Toloubeydokhti, T., Cruz, A. C. & Chegini, N. Expression profile of MicroRNAs and mRNAs in human placentas from pregnancies complicated by preeclampsia and preterm labor. *Reprod. Sci. Thousand Oaks Calif.* **18**, 46–56 (2011).
91. He, Y. et al. miR-149 in human cancer: A systemic review. *J. Cancer.* **9**, 375–388 (2018).
92. Mutlu, M. et al. miR-200c: a versatile watchdog in cancer progression, EMT, and drug resistance. *J. Mol. Med. Berl Ger.* **94**, 629–644 (2016).
93. Chauhan, N., Dhasmana, A., Jaggi, M., Chauhan, S. C. & Yallapu, M. M. miR-205: A potential biomedicine for Cancer therapy. *Cells* **9**, 1957 (2020).
94. Li, S. et al. The comprehensive landscape of miR-34a in cancer research. *Cancer Metastasis Rev.* **40**, 925–948 (2021).
95. Ma, C. et al. mTOR hypoactivity leads to trophoblast cell failure by enhancing lysosomal activation and disrupting the cytoskeleton in preimplantation embryo. *Cell. Biosci.* **13**, 219 (2023).
96. Beetch, M. & Alejandro, E. U. Placental mTOR signaling and sexual dimorphism in metabolic health across the lifespan of offspring. *Child. Basel Switz.* **8**, 970 (2021).
97. Félétou, M. *The Endothelium: Part 1: Multiple Functions of the Endothelial Cells—Focus on Endothelium-Derived Vasoactive Mediators* (Morgan & Claypool Life Sciences, 2011).
98. Li, H. et al. Human placental endothelial cell and trophoblast heterogeneity and differentiation revealed by Single-Cell RNA sequencing. *Cells* **12**, 87 (2022).
99. Fuchi, N., Miura, K., Doi, H., Li, T. S. & Masuzaki, H. Feasibility of placenta-derived mesenchymal stem cells as a tool for studying pregnancy-related disorders. *Sci. Rep.* **7**, 46220 (2017).
100. Mouillet, J. F., Ouyang, Y., Coyne, C. B. & Sadovsky, Y. MicroRNAs in placental health and disease. *Am. J. Obstet. Gynecol.* **213**, S163–172 (2015).
101. Pan, B. et al. MicroRNA-371-3 cluster as biomarkers for the diagnosis and prognosis of cancers. *Cancer Manag Res.* **11**, 5437–5457 (2019).
102. McCarthy, E. C. & Dwyer, R. M. Emerging evidence of the functional impact of the miR379/miR656 cluster (C14MC) in breast Cancer. *Biomedicines* **9**, 827 (2021).
103. Xie, L. et al. C19MC MicroRNAs regulate the migration of human trophoblasts. *Endocrinology* **155**, 4975–4985 (2014).
104. Mong, E. F. et al. Chromosome 19 MicroRNA cluster enhances cell reprogramming by inhibiting epithelial-to-mesenchymal transition. *Sci. Rep.* **10**, 3029 (2020).
105. Clark, J. et al. Comparing the predictivity of human placental gene, MicroRNA, and CpG methylation signatures in relation to perinatal outcomes. *Toxicol. Sci.* **183**, 269–284 (2021).
106. Donker, R. B. et al. The expression profile of C19MC MicroRNAs in primary human trophoblast cells and exosomes. *Mol. Hum. Reprod.* **18**, 417–424 (2012).
107. Liu, L., Zhang, J. & Liu, Y. MicroRNA-1323 serves as a biomarker in gestational diabetes mellitus and aggravates high glucose-induced Inhibition of trophoblast cell viability by suppressing TP53INP1. *Exp. Ther. Med.* **21**, 230 (2021).
108. Hromadnikova, I. et al. Expression profile of C19MC MicroRNAs in placental tissue in pregnancy-related complications. *DNA Cell. Biol.* **34**, 437–457 (2015).
109. Meakin, A. S., Cuffe, J. S. M., Darby, J. R. T., Morrison, J. L. & Clifton, V. L. Let's talk about placental sex, baby: Understanding mechanisms that drive Female- and Male-Specific fetal growth and developmental outcomes. *Int. J. Mol. Sci.* **22**, 6386 (2021).
110. Flowers, A. E. et al. Sex differences in MicroRNA expression in first and third trimester human placenta†. *Biol. Reprod.* **106**, 551–567 (2022).
111. Aure, M. R. et al. Crosstalk between MicroRNA expression and DNA methylation drives the hormone-dependent phenotype of breast cancer. *Genome Med.* **13**, 72 (2021).
112. Cherubini, M., Erickson, S. & Haase, K. Modelling the human placental interface in Vitro-A review. *Micromachines* **12**, 884 (2021).
113. Li, X., Li, Z. H., Wang, Y. X. & Liu T.-H. A comprehensive review of human trophoblast fusion models: recent developments and challenges. *Cell. Death Discov.* **9**, 372 (2023).
114. Keenen, M. M. et al. Comparative analysis of the syncytiotrophoblast in placenta tissue and trophoblast organoids using snRNA sequencing. *BioRxiv Prepr. Serv. Biol.* 07 (2024).
115. Maccani, M. A., Padbury, J. F. & Marsit, C. J. miR-16 and miR-21 expression in the placenta is associated with fetal growth. *PloS One.* **6**, e21210 (2011).
116. Zhu, Y. et al. MicroRNA-16 inhibits feto-maternal angiogenesis and causes recurrent spontaneous abortion by targeting vascular endothelial growth factor. *Sci. Rep.* **6**, 35536 (2016).
117. Sood, R., Zehnder, J. L., Druzin, M. L. & Brown, P. O. Gene expression patterns in human placenta. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 5478–5483 (2006).
118. Wu, M. et al. Comparison of the biological characteristics of mesenchymal stem cells derived from the human placenta and umbilical cord. *Sci. Rep.* **8**, 5014 (2018).

## Acknowledgements

We gratefully acknowledge the study participants who kindly donated their placentas to the BCCHB. We gratefully acknowledge the work and staff of the BC Children's Hospital BioBank for their assistance with patient recruitment and sample and data acquisition. Thank you to Dr. Lisa Xu, BCCHR Flow Core Facility Manager for help with the cell-sorting, and Miwa Suzuki and May Zhang for the sorted-cell RNA extractions. We thank

Pawan Pandoh, Dr. Yongjun Zhao, Dr. Andy Mungall, and the entire Library Tech Development Group at Canada's Michael Smith Genome Sciences Centre at BC Cancer.

### Author contributions

NT conducted the data analysis, drafted, and revised the manuscript. WPR, WLL, AGB conceived this study, obtained funding and were responsible for the project design, input on data collection, and analysis. MSP contributed to study design, data collection, sample processing. VY contributed to data collection and analysis. DH contributed to data collection. VDM and GLS provided feedback on data analysis and revised the manuscript. All authors read, edited, and approved the final manuscript.

### Funding

Research reported in this publication was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health (NIH) under Award Number 5R01HD089713, and the Canadian Institutes of Health Research (CIHR) grants JT169131 and GSK171375 to WPR, and FRN-143345 and 183775 to WLL. NT was supported by a 4 Year Doctoral Fellowship from The University of British Columbia. WPR receives salary support through an investigatorship award from the BC Children's Hospital Research Institute.

### Declarations

### Competing interests

The authors declare no competing interests.

### Ethics approval and consent to participate

Written informed consent was obtained from all participants including a tissue banking consent. This study was approved by The University of British Columbia and BC Women's and Children's Hospital research ethics board in Vancouver, BC, Canada (certificate number: H16-02280).

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-98939-4>.

**Correspondence** and requests for materials should be addressed to W.L.L. or W.P.R.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025