Manuscript (Word document)
Click here to access/download;Manuscript (Word document);Witney et al Placental virome (FINAL) Revised.docx

RESEARCH ARTICLE

Metagenomic profiling of placental tissue suggests DNA virus infection of the placenta is rare

Adam A Witney[a#], Sean Aller[a] and Blair L Strang[a#]

Institute of Infection & Immunity, St George's, University of London, Cranmer Terrace, London, SW17 0RE

Running Title: Viral DNA sequences in the placenta

#Corresponding Authors
Email: awitney@sgul.ac.uk, bstrang@sgul.ac.uk

No. of words in Abstract: 142
No. of words in Manuscript: 3,512

Keywords: virus, placenta, metagenomic profiling, herpes

CONFIDENTIAL

ABSTRACT

It is widely recognized that pathogens can be transmitted across the placenta from mother to fetus. Recent reevaluation of metagenomic studies indicates that the placenta has no unique microbiome of commensal bacteria. However, viral transmission across the placenta, including transmission of DNA viruses such as the human herpesviruses, is possible. A fuller understanding of which DNA virus sequence can be found in the placenta is required. We employed a metagenomic analysis to identify viral DNA sequences in placental metagenomes from full term births (20 births), pre-term births (13 births), births from pregnancies associated with antenatal infections (12 births) or pre-term births with antenatal infections (3 births). Our analysis found only a small number of DNA sequences corresponding to the genomes of human herpesviruses in four of the forty-eight metagenomes analyzed. Therefore, our data suggests DNA virus infection of the placenta is rare and supports the concept that the placenta is largely free of pathogen infection.

CONFIDENTIAL

INTRODUCTION

Vertical transmission of pathogens from mother to child is a mechanism of pathogen dissemination within human populations (1). This can occur by several well-characterized mechanisms, including post-partum infection at birth and via breast milk (1). However, vertical transmission of pathogens from mother to fetus across the placenta *in utero* is very poorly understood and conflicting data regarding bacterial colonization of the placenta has led to recent reevaluation of the placenta as a sterile environment.

The womb is thought of as an environment free of bacteria as it is well documented that germ free neonates can be delivered by caesarean section (2, 3), an observation that can be recapitulated in animal models (3). However, there have been reports of bacteria found in placental tissue (4-12), amniotic fluid (7) and the fetus (13), leading to reports that the placenta harbors a unique microbiome of nonpathogenic commensal bacteria. Arguably the most prominent reports of a unique placental microbiome have been produced by Aagaard and colleagues (4-6), who reported the detection of a wide range of commensal bacterial DNA sequences in placental tissue using methods such as bacterial 16S sequencing and metagenomic DNA profiling.

There have been several challenges to the aforementioned reports. Notably, it has been observed that these studies reported only very low levels of bacteria were present and did not demonstrate that viable bacteria are present (14). It has been further suggested that the reported placental microbiome could

CONFIDENTIAL

be bacteria found in maternal blood within the placental tissue analyzed or come from another maternal site (14). However, there is unlikely to be sufficient maternal blood within placental tissue to justify this reasoning (15).

The strongest rebuttal to reports of a placenta microbiome have been offered by data that detection of bacterial DNA within placental tissue was likely due to contamination of samples during preparation of genomic DNA via the DNA purification kits used, or via contamination with bacteria found in the environment during sample collection and/or DNA purification (16-20). Similarly, the presence of bacteria in amniotic fluid (7) has been challenged by work reporting that the bacterial signals found in amniotic fluid samples were indistinguishable from background controls (21). Furthermore, a recent survey of the placentas from many hundreds of patients has indicated that the placenta has no obvious unique microbiome and that nearly all detection of bacteria in the survey was connected to contamination of reagents with bacterial DNA and acquisition of bacteria during labour and/or delivery (22).

The recent reevaluation of what bacteria are found in placental tissue stimulated us to evaluate what viruses can be found in the placenta. It is known that viruses can pass through the placenta, be transmitted to the fetus and cause disease (1). Some of the most prominent viruses vertically transmitted across the placenta are those with DNA genomes, such as the human herpesvirus (HHVs) herpes simplex virus (HSV), human cytomegalovirus (HCMV) and varicella-zoster virus (VZV) (1). However, a fuller understanding of the prevalence of herpesvirus infection of the placenta and how these viruses replicate and cause disease in the

CONFIDENTIAL

placenta or affect pregnancy outcomes is required (1, 23, 24). Plus, there have been contrasting reports of which herpesviruses can be found in placental tissue. It has been reported that HCMV and HSV can be found in full term placental and decidual tissue in the presence and absence of several common sexually transmitted pathogenic non-commensal bacteria (for example, chlamydia) (25). It is unknown what effect these bacterial infections could have on herpesvirus replication and pathogenesis. Moreover, it has been observed that more placental tissues contain HCMV and a pathogenic bacterial species than HCMV alone (25), suggesting pathogenic bacterial infection promotes HCMV infection. However, more recent studies have reported that DNA from HSV-1, HSV-2, or HCMV could not be found in placental DNA from full term pregnancies, pre-term pregnancies, pregnancies with preeclampsia or pregnancies with fetal growth restriction (26).

Also, other herpesviruses can utilize the placenta for vertical transmission, which leads to disease. For example, it has been demonstrated that HHV-6A, -7 and -8 can replicate in cells derived from placental tissue *in vitro* and HHV-6A, -6B and -7 could be involved in disease during pregnancy (26-33). Specifically, a recent report has demonstrated that the only viral RNA that could be reliably detected in cases of pre-eclampsia was from either HHV-6A or HHV-6B genomes and there was a strong association between the presence of chromosomally inherited HHV-6A or -6B and pre-eclampsia (26). However, it is unknown if HHV-6 viruses have an obvious impact on other pathologies during pregnancy or if there is any relationship between HHV-6 infection and bacterial infection during pregnancy.

CONFIDENTIAL

There has also been contrasting data presented on which DNA viruses other than herpesviruses may be found in the placenta. There are reports that adenovirus (34), adeno-associated virus (35), papillomavirus (36) and polyomavirus (37) can infect human or murine placental cells *in vitro*, suggesting that viruses from these families may be found in the placenta. However, a recent study has indicated that Adenovirus and Papillomavirus DNA could not be found in placental DNA extracts from full term pregnancies, pre-term pregnancies, pregnancies with preeclampsia or pregnancies with fetal growth restriction (26). Plus, it is as yet unknown if there are the links between placental infection by these viruses and disease in humans.

Therefore, it was possible that any of the aforementioned DNA viruses could be found in the placenta, alone or in combination, and that non-commensal bacterial infection could have a role in promoting virus infection. To understand which viruses with DNA genomes could be found in the placenta we performed an analysis of a published metagenomic dataset of genomic DNA isolated from the placental tissue of several patient groups, including those from preterm births and/or antenatal infections.

CONFIDENTIAL

MATERIALS AND METHODS

Microbiome data from public repositories.

Microbiome derived DNA sequences from a range of body sites, were selected (filtered by file type “FASTQ” and data “WGS”) and downloaded from either the National Institutes of Health (NIH) Human Microbiome Project (HMP) database or the NIH Integrative Human Microbiome Project (iHMP) database (38-40). Accession numbers of sequences analyzed can be found in Table S1. In addition, 52 metagenomes sequenced from samples from the upper reproductive tract were downloaded (41). Upper reproductive tract sequences are available from the EBI ENA database with project accession PRJEB24147 (41).

Analysis of DNA sequencing using metagenomic profiling tools.

Metagenomic sequence reads were trimmed using Trimmomatic (version 0.39) (42) in paired end mode to yield sets of high quality paired reads. Trimmed reads were filtered to remove any human sequences using Kraken2 (43) against a human genome sequence specific database. Taxonomic classifications were made against several databases using Kraken2; the databases included the full standard database and a viral genome specific database. The output of Kraken2 consists of the number of unique reads across multiple taxonomies. For the purposes of the present study, only reads categorized under Viruses by Kraken2 were analysed. Normalised read counts were generated by dividing by the sample read count and scaling to 10^6.

CONFIDENTIAL

Where indicated in the text, reads were also analyzed using Kraken2 by comparing reads to a database of human herpesvirus genomes created by downloading all human herpesvirus genomes from NCBI Viral Neighbor assemblies. A total of 1,379 genomes were analyzed comprising 70 HSV-1 genomes (5% of the genomes analyzed), 64 HSV-2 genomes (4.6% of the genomes analyzed), 202 VZV genomes (14.6% of the genomes analyzed), 659 EBV genomes (47.8% of the genomes analyzed), 333 HCMV genomes (24.1% of the genomes analyzed), 7 HHV-6A genomes (0.5% of the genomes analyzed), 9 HHV-6B genomes (0.65% of the genomes analyzed), 3 HHV-7 genomes (0.2% of the genomes analyzed), and 32 KSHV genomes (2.3% of the genomes analyzed).

Raw untrimmed reads were also converted to bam format using Picard, and classified using the GATK PathSeq pipeline using default settings (44). Analysis and plotting were performed using R (v4.0.4) with the tidyverse package (45).

Placental metagenome data.

Forty eight placental metagenomic sequences (5) were very kindly provided by Kjersti Aagaard and colleagues via the NCBI database of Genotype and Phenotype (request no. #65088-2). Please see Acknowledgments for further information. The interactions with patients, preparation of samples and DNA sequencing methodology used to generate this metagenomic data have been previously described in detail (5). Briefly, placental samples were collected immediately after delivery. Several tissue sections (between 4 to 6 sections) were taken from the same area of each placenta (with the maternal decidua and fetal

CONFIDENTIAL

chorion-amnion tissues removed), potentially gathering tissue from the placental arteries, placental vein, and villous. DNA was extracted from the combined 4 to 6 sections of each placenta using a MO-BIO PowerSoil DNA Isolation Kit (MO-BIO Laboratories), which should extract all DNA from the tissue, including human genomic DNA, viral genomic DNA, bacterial genomic DNA and phage genomic DNA. All DNA was sequenced using the Illumina HiSeq 2500 platform. The sequencing data from the combined sections of each placenta was assigned an accession number used in Figures 1, S1 and S2.

Where indicated in the text, the estimated genome coverage of sequence reads was calculated as: ((no. reads x read length x 2) / genome length).

CONFIDENTIAL

RESULTS

Investigation of viral DNA sequence reads in placental metagenomic DNA.

The largest public database of placental metagenomes available was generated from the aforementioned mentioned study by Aagaard *et al.* (5), which contained metagenomic data from different patient groups; those pregnancies which had delivered at full term (20 births), pregnancies that have delivered spontaneously preterm (less than 37 weeks) (13 births), full term pregnancies that had reported infections (antepartum sexually transmitted infections (for example, gonorrhea, chlamydia, or syphilis), urinary tract infections or systemic infections (for example, pneumonia)) (12 births) and pregnancies that were both preterm and had reported infections (3 births). Information on sample collection and preparation can be found in the placental metagenome data portion of the Materials and Methods section.

Based on previously reported factors such as accuracy and ease of use (46), a workflow based on Kraken2, a metagenomic classification tool that has previously been used to identify various viral DNA sequences in a range of tissue types (46-48), was developed. The workflow first included read trimming, then reads were filtered using a Kraken2 human genome specific database, and finally remaining reads were classified using the Kraken2 standard database containing the RefSeq complete genomes for archaea, bacteria, human, viruses, plus Univec core (a database of vector, adaptor, linker and primer sequences). Applying a minimum cutoff of greater than or equal to two or more reads, thus removing any

CONFIDENTIAL

single match classifications, only four placental metagenomes had reads classified as virus derived; all classifications were attributed to herpesviruses HHV-6A, HHV-6B and HHV-5 (HCMV) (Figure 1) and all HHV classified reads were found in metagenomic samples from full term births. Alignment to an HHV-6 reference genome showed that the small number of matching reads aligned along the entire HHV-6 genome rather than to one specific region (data not shown).

A second analysis classifying the human filtered reads against a Kraken2 virus genome only database identified a significant number of phiX174 reads in each metagenomic sequence (Supplementary Figure 1). These were likely introduced as a control in the sequencing protocol. Interestingly, these phiX174 matched reads were not identified in the initial standard database analysis as these reads were classified as bacterial derived phage sequences, and matches were filtered for virus only classified sequences in that analysis. Reads corresponding to phiX174 were not over represented in metagenomic samples from full term births. Therefore, the differences we observed in detection of reads corresponding to human viral genomes (Figure 1) was unlikely to be related to differences in sequencing of DNA from each placental metagenome.

To ensure that the reads we detected in Figure 1 were from human herpesvirus genomes, we used Kraken2 to compare reads from placental metagenomes with a database of human herpesvirus genomes that we created for this study (Supplementary Figure 2). We found near identical data to that shown in Figure 1.

CONFIDENTIAL

To confirm the data produced by the Kraken2 workflow an alternative method using the GATK PathSeq tool (44) was used. Results showed a very similar classification of the placental metagenomes (Supplementary Figure 3). One further low level match was found to Torque Teno virus which has previously been shown found in many human tissue types, although it is not thought to be transmitted across the placenta (49). Detection of Torque Teno virus reads by GATK, but not Kraken2, may reflect a difference in the content of the databases used by the two tools.

Therefore, using both Kraken2 and GATK workflows the only reads corresponding to viral DNA genomes that could be confidently identified in placental metagenomes were a small number of reads from the human herpesviruses in samples from full term pregnancies.

It was interesting to note that reads corresponding to HHV-6A could be detected (SRR1269152, SRR1269157, SRR1269162 (Fig. 1)). As chromosomal integration of HHV-6A has been reported in pregnant women (22, 26, 28, 32), we hypothesized that the HHV-6A sequences we detected could be from chromosomally integrated virus. We calculated the estimated HHV-6A genome coverage (relative abundance) of the matching reads in SRR1269157 to be 0.36 (Kraken2) or 0.46 (GATK PathSeq), compared to an estimated human genome coverage of 0.79 (assuming the complete read set were human derived). This possibly suggested that there was one integrated copy of the HHV-6A genome in the SRR1269157 genome. The relative abundance of HHV-6A reads in SRR1269152 was 0.03 or 0.05 (Kranken2 and GATK PathSeq analysis,

CONFIDENTIAL

respectively) and in SRR1269162 was 0.006 for both Kranken2 and GATK PathSeq analysis. However, the number of HHV-6A reads detected in SRR1269152 and SRR1269162 was very low. Thus, it was not possible to draw firm conclusions from these estimations as to whether integrated HHV-6A was or was not present in SRR1269152 and SRR1269162.

Validation of the metagenomic profiling tool by identifying viral DNA reads in human microbiome metagenomes.

Only reads corresponding to the genomes of human herpesviruses could be found in our analysis of placental metagenomes (Figure 1 and Supplementary Figure 1). To validate the use of Kraken2 in finding viral DNA reads from a range of viral genomes within metagenomic sequencing data, the Kraken 2 workflow was applied to human microbiome sequence data that had been deposited in a publicly available database; the Human Microbiome Project (38-40). 3646 metagenomic samples were analyzed. In addition, a study assessing 52 metagenomic profiles from the upper reproductive tract (49) was included.

Consistent with a pattern previously observed elsewhere (50), assessment of herpesvirus classifications (Figure 2) showed a pattern whereby oral samples were largely dominated by HHV-7 sequences and to a lesser extent HHV-4 (EBV) and HHV-6B. HCMV could be found in nares, upper reproductive tract and vaginal samples. Reads corresponding to HHV-1 and HHV-2 (HSV-1 and HSV-2, respectively) genomes were found in nares and oral samples or only oral samples,

CONFIDENTIAL

respectively. Therefore, Kraken2 could detect reads corresponding to genomes of many common human herpesviruses.

To further confirm the validity of the Kraken2 workflow, the distribution of papillomaviruses was assessed in the same metagenome sample set. At the genus level, metagenomes sequenced from the nares were dominated by Betapapillomaviruses and Gammapapillomaviruses (Figure 3), as has previously been described (51). Conversely samples isolated from the upper reproductive tract and vaginal sites were largely dominated by Alphapapillomaviruses.

Therefore, Kraken2 was capable of detecting reads corresponding to genomes from common human DNA viruses such as herpesviruses and papillomaviruses. This suggests that it was likely that Kraken2 could detect reads corresponding to a range of DNA virus genomes if they were present in placental metagenomes.

CONFIDENTIAL

DISCUSSION

We sought to understand which human DNA virus genomes could be found in placental tissue and if those DNA genomes differed between patient groups. For example, a difference between metagenomes from placentas of full-term births compared to those from full term births that had reported antenatal infection during pregnancy. We found that only reads corresponding to the genomes of HHV-6A, HHV-6B and HCMV were identified in four metagenomic samples from full term births.

The study herein was stimulated by recent reevaluation of the presence of bacteria in placental tissue. A weight of evidence indicated that reports of a unique placental microbiome of commensal bacteria may have been the result of technical issues surrounding preparation and execution of DNA sequencing studies or acquisition of bacteria during labor and/or delivery (16-20, 22). Thus, it may be possible that that placenta is not routinely colonized by bacteria. As we find very few reads corresponding to human DNA viruses in our analysis, it may be possible that the placenta is not routinely colonized by either bacteria or viruses with DNA genomes. Indeed, our observations are similar to those made elsewhere (26), in so far as HHV-6 sequences could be found in placental metagenomes.

Our data leads to questions surrounding why herpesviruses known to be vertically transmitted across the placenta (HSV, VZV) were not detected in our experiments and the relevance of detecting HHV-6 and HCMV in our analysis. It is possible that HSV and VZV viral genomes were present in the placental samples,

CONFIDENTIAL

but only at very low numbers that were undetectable at the level of sequencing performed or that these organisms were lost through the DNA extraction protocol. It is equally possible that human herpesvirus infection of the placenta is a rare occurrence and that considerably more placental metagenomes would have to be analyzed in order to detect reads corresponding to HSV and VZV genomes. It is interesting to speculate how common herpesvirus infection of the placenta might be. HCMV is perhaps the best studied in this regard. HCMV seroprevalence is high worldwide, with a seroprevalence ranging from 45% to 100% (52). However, congenital infection by HCMV occurs in only 0.2%-2% of live births (52). Therefore, HCMV infection of the placenta may not be common, in line with our observations made in our study and elsewhere (26).

It is interesting to contrast our observations with a previous report that HCMV and HSV genomes could be routinely observed in placental and decidual tissue obtained from both first trimester terminations of healthy pregnancies and in placental tissue from second trimester terminations of healthy pregnancies (HCMV in 28% and 26% of samples, respectively. HSV-1 in 3% and 3% of samples, respectively. HSV-2 in 6% and 13% of samples, respectively) (25)). It is possible that the differences in our observations are due to experimental differences that relate to the sensitivities of the assays used (polymerase chain reaction versus metagenomic profiling) or the tissue that was examined (In the aforementioned work (25), several tissue samples were taken from random sites in the placental and decidual tissue, whereas the DNA analyzed here was taken from the same region of each placenta). Furthermore, the DNA analyzed here was from near full

CONFIDENTIAL

term or full-term spontaneous births, whereas the tissue analyzed in the aforementioned study (25) was from first and second trimester terminations of healthy pregnancies. This implies that DNA virus infection of the placenta may be more common early in pregnancy, during the first and second trimesters, or that there are areas of the placenta enriched in virus infected cells.

We found reads corresponding to HHV-6 genomes in four of the placental metagenomic samples analyzed. We considered if the presence of reads corresponding to the HHV-6 genomes were due to contamination of the placental metagenomes we had analyzed, but know of no plausible route that would allow HHV DNA to contaminate the placental metagenomic data we analyzed. We know of no environmental source that would contain HHV-6 viruses. It is possible that HHV-6 viruses could have been introduced during preparation of DNA from placental tissue. HHV-6A can cause dermal infections, but given the precautions taken during metagenomic sample processing (5), it is unlikely that dermal shedding of HHV-6A into placental tissues is a route of contamination in this case. HHV-6A is not known to be transmitted in the air. HHV-6A is not known to be a contaminant found in DNA preparation kits, which was a likely source of bacterial contamination of placental tissue in studies elsewhere (16). Therefore, we submit that the presence of reads corresponding to HHV-6 genomes in our analysis are unlikely due to contamination of the metagenomic sequenced that we have analyzed. Furthermore, DNA sequences corresponding to HHV-6 genomes have been reported in placental genomes prepared using stringent isolation procedures (22, 26).

CONFIDENTIAL

Our observation that reads corresponding to HHV-6 were found in placental metagenomes suggests that these viruses can potentially replicate in the placenta. Several points support HHV-6A infection of the placenta. The aforementioned virus is found in the betaherpesvirus sub-family of the herpes viruses and, as such, are related to the vertically transmitted TORCH pathogen HCMV (53). It has been demonstrated that HHV-6A can replicate within placental tissue *ex vivo* (33), HHV-6 DNA can be found in fetal tissue, umbilical cord blood and villous tissue (30) and widespread infection of pregnant women with HHV-6A has been reported, which may be associated with disease (26-33, 54). Importantly, as HHV-6A can integrate into chromosomal DNA, it is possible that the presence of reads corresponding to HHV-6A genomes that we detect in placental metagenomes is not due to infection of the placenta, but from HHV-6A genomes integrated into the genomes of pregnant women, which has been reported elsewhere (22, 26, 28, 32).

A relationship between chromosomal integration of HHV-6A in placenta and preeclampsia has been established (26). We did not detect HHV-6A genomes in placental samples from either pre-term births or pregnancies with antenatal infection, suggesting no obvious link between HHV-6A infection and those patient groups. However, a larger study may be required to confirm this.

Another question we sought to answer was what breadth of DNA viruses that might be found in placental tissue. As outlined above, viruses from a number of DNA virus families can replicate in cells derived from placental tissue *in vitro* (34-37), although neither Adenovirus or Papillomavirus DNA has been detected in a range of pregnancies (26). In our study only reads corresponding to the HHV-

CONFIDENTIAL

6A/6B and HCMV genomes were found in placental metagenomes. Alternatively, it is possible that a number of other viral DNA genomes are in the placental metagenomes we analyzed, but present at levels that were undetectable at the level of sequencing performed. It is also possible that placental infection with DNA viruses other than herpesviruses is rare and further study of a range of placental tissue may reveal the presence of these viruses. Importantly, our data may indicate that replication of some DNA viruses in placental cells *in vitro* may not reflect the ability of those viruses to infect placental tissue *in vivo*.

Furthermore, we wished to examine was the potential relationship between the presence of non-commensal bacterial infection and virus infection of the placenta. We found no obvious relationship between the presence of any virus and antenatal infection during pregnancy. Similarly, we found no obvious relationship between virus infection and pre-term birth. It is possible that these links exist, but were not observed here due to the caveats we discuss in the preceding paragraphs, for example study size and levels of detection. That said, our data may imply that there is no obvious link between pre-term birth or antenatal infection and DNA virus infection.

Further study of virus infection of the placenta is warranted. As discussed here, to detect herpesviruses and other DNA viruses it is likely that investigation will have to be expanded to examine much larger numbers of tissue. Furthermore, based on the points we discuss here consideration should be given to examining tissue from both early trimesters of pregnancy and tissue from full term pregnancy. Finally, there are several important vertically transmitted viruses that have RNA

CONFIDENTIAL

genomes. For example, rubella virus and Zika virus. The methodology used here would not detect these RNA genomes. To fully understand virus infection of the placenta it will be necessary in future studies to consider viruses with both DNA and RNA genomes.

CONFIDENTIAL

AUTHOR CONTRIBUTIONS

AAW: Conceptualization, Methodology, Resources, Writing-Original Draft Preparation, Writing-Review and Editing, Supervision. SA: Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Data Curation, Writing-Review and Editing, Visualization. BLS: Conceptualization, Methodology, Writing-Original Draft Preparation, Writing-Review and Editing, Supervision, Project Administration, Funding.

CONFLICTS OF INTEREST

The authors declare there are no conflicts of interest.

CONFIDENTIAL

FUNDING INFORMATION

This work was supported by a Wellcome Trust Institutional Strategic Support Fund award to BLS [204809/Z/16/Z].

ACKNOWLEDGMENTS

We are grateful to Andrew Davison, Beth Holder and Guy Whitley for helpful comments on the manuscript. The authors gratefully acknowledge the data generously provided by Kjersti Aagaard (K.A.) and colleagues, who have requested the following statement be included here to ensure all colleagues and funders are correctly recognized:

A dataset used for the analyses described in this manuscript was obtained for the database of Genotype and Phenotype (dbGaP) found at http://www.ncbi.nlm.nih.gov/gap [phs000735.v1.p1]. Samples and associated data for the Placental Microbiome were provided by Kjersti Aagaard-Tillery, M.D., Ph.D. The provider of the dataset gratefully acknowledges support of the NIH Director's New Innovator Award (K.A. DP2 OD001500), The Burroughs Wellcome Fund Preterm Birth Initiative (K.A.) and the Human Microbiome Project funded by the NIH Director's Common Fund at the National Institutes of Health (as part of the NIH RoadMap 1.5). All sequencing and adaption of protocols for sequencing low biomass samples were performed at the Baylor College of Medicine Human Genome Sequencing Center (BCM-HGSC), which is funded by direct support from the National Human Genome Research Institute (NHGRI) at NIH [U54 HG004973(BCM), Dr. Richard Gibbs, P.I.]. The dataset provider thanks the Alkek Center of Metagenomics and Microbiome Research at the Baylor College of Medicine for their support through the microbiota discovery award (R.G., K.A.).

CONFIDENTIAL

The dataset provider also thanks the staff directly involved in clinical recruitment and specimen processing (Jia Chen, Debra Duncan, Irma Guerra, Claire Cook, Milena Pitashny-Hoffman) and the additional faculty who participated with K.A. and James Versalovic in development of the PeriBank protocol and resource building (Dr. Edwina Popek, Ignatia Van den Veyver, Tom Wheeler, and Michael Belfort). The dataset provider is grateful to Tulin Ayvas for assisting with protocol development for DNA isolation and sample processing.

CONFIDENTIAL

FIGURE LEGENDS

Figure 1 Number of reads corresponding to viral genomes in placental metagenomes. The sample accession of each metagenomic sample was noted on the x-axis. Metagenomic samples came from either full-term pregnancies, pre-term pregnancies, full term pregnancies that had reported antenatal infection or pre-term pregnancies that had reported antenatal infections (Pt + inf). Accession SRR1269153 was removed as it had a higher than expected level of phiX174 derived reads which were likely misclassified as other bacterial phage viruses (data not shown). HCMV is shown as human herpesvirus 5.

Figure 2 Number of reads corresponding to human herpesviruses in metagenomes from the Human Microbiome Project and a study of the upper reproductive tract. Read counts were divided by sample counts and normalized to $1x10^6$ reads. (A) bars show the classified read counts separated into four general body sites, total number of reads for each metagenome are plotted with dots above each accession. (B) The number of metagenomes where human herpesvirus sequences (>=2 reads) are found. In each figure, HSV-1 is shown as human herpesvirus 1, HSV-2 is shown as human herpesvirus 2, EBV is shown as human herpesvirus 4 and HCMV is shown as human herpesvirus 5.

Figure 3 Number of reads corresponding to papillomaviruses in metagenomes from the Human Microbiome Project and a study of the upper

CONFIDENTIAL

reproductive tract. Read counts were divided by sample counts and normalized to $1x10^6$ reads. (A) bars show the classified read counts separated into four general body sites, total number of reads for each metagenome are plotted with dots above each accession. (B) The number of metagenomes where papillomavirus sequences (>=2 reads) are found.

CONFIDENTIAL

REFERENCES

1. Coyne CB, Lazear HM. Zika virus - reigniting the TORCH. Nat Rev Microbiol. 2016;14(11):707-15. doi: 10.1038/nrmicro.2016.125

2. Barnes RD, Fairweather DV, Holliday J, Keane C, Piesowicz A, Soothill JF, et al. A germfree infant. Lancet. 1969;1(7587):168-71. doi: 10.1016/s0140-6736(69)91187-8

3. Perez-Munoz ME, Arrieta MC, Ramer-Tait AE, Walter J. A critical assessment of the "sterile womb" and "in utero colonization" hypotheses: implications for research on the pioneer infant microbiome. Microbiome. 2017;5(1):48. doi: 10.1186/s40168-017-0268-4

4. Antony KM, Ma J, Mitchell KB, Racusin DA, Versalovic J, Aagaard K. The preterm placental microbiome varies in association with excess maternal gestational weight gain. American journal of obstetrics and gynecology. 2015;212(5):653 e1-16. doi: 10.1016/j.ajog.2014.12.041

5. Aagaard K, Ma J, Antony KM, Ganu R, Petrosino J, Versalovic J. The placenta harbors a unique microbiome. Science translational medicine. 2014;6(237):237ra65. doi: 10.1126/scitranslmed.3008599

6. Prince AL, Ma J, Kannan PS, Alvarez M, Gisslen T, Harris RA, et al. The placental membrane microbiome is altered among subjects with spontaneous preterm birth with and without chorioamnionitis. American journal of obstetrics and gynecology. 2016;214(5):627 e1- e16. doi: 10.1016/j.ajog.2016.01.193

CONFIDENTIAL

7. Collado MC, Rautava S, Aakko J, Isolauri E, Salminen S. Human gut colonisation may be initiated in utero by distinct microbial communities in the placenta and amniotic fluid. Scientific reports. 2016;6:23129. doi: 10.1038/srep23129

8. Gomez-Arango LF, Barrett HL, McIntyre HD, Callaway LK, Morrison M, Nitert MD. Contributions of the maternal oral and gut microbiome to placental microbial colonization in overweight and obese pregnant women. Scientific reports. 2017;7(1):2860. doi: 10.1038/s41598-017-03066-4

9. Tomlinson MS, Bommarito PA, Martin EM, Smeester L, Fichorova RN, Onderdonk AB, et al. Microorganisms in the human placenta are associated with altered CpG methylation of immune and inflammation-related genes. PloS one. 2017;12(12):e0188664. doi: 10.1371/journal.pone.0188664

10. Zheng J, Xiao X, Zhang Q, Mao L, Yu M, Xu J, et al. The Placental Microbiota Is Altered among Subjects with Gestational Diabetes Mellitus: A Pilot Study. Frontiers in physiology. 2017;8:675. doi: 10.3389/fphys.2017.00675

11. Fischer LA, Demerath E, Bittner-Eddy P, Costalonga M. Placental colonization with periodontal pathogens: the potential missing link. American journal of obstetrics and gynecology. 2019. doi: 10.1016/j.ajog.2019.04.029

12. Onderdonk AB, Delaney ML, DuBois AM, Allred EN, Leviton A. Detection of bacteria in placental tissues obtained from extremely low gestational age neonates. American journal of obstetrics and gynecology. 2008;198(1):110 e1-7. doi: 10.1016/j.ajog.2007.05.044

CONFIDENTIAL

13. Martinez KA, 2nd, Romano-Keeler J, Zackular JP, Moore DJ, Brucker RM, Hooper C, et al. Bacterial DNA is present in the fetal intestine and overlaps with that in the placenta in mice. PloS one. 2018;13(5):e0197439. doi: 10.1371/journal.pone.0197439

14. Kliman HJ. Comment on "the placenta harbors a unique microbiome". Science translational medicine. 2014;6(254):254le4. doi: 10.1126/scitranslmed.3009864

15. Aagaard KM. Author response to comment on "the placenta harbors a unique microbiome". Science translational medicine. 2014;6(254):254lr3. doi: 10.1126/scitranslmed.3010007

16. Bushman FD. De-Discovery of the Placenta Microbiome. American journal of obstetrics and gynecology. 2019;220(3):213-4. doi: 10.1016/j.ajog.2018.11.1093

17. Leiby JS, McCormick K, Sherrill-Mix S, Clarke EL, Kessler LR, Taylor LJ, et al. Lack of detection of a human placenta microbiome in samples from preterm and term deliveries. Microbiome. 2018;6(1):196. doi: 10.1186/s40168-018-0575-4

18. Lauder AP, Roche AM, Sherrill-Mix S, Bailey A, Laughlin AL, Bittinger K, et al. Comparison of placenta samples with contamination controls does not provide evidence for a distinct placenta microbiota. Microbiome. 2016;4(1):29. doi: 10.1186/s40168-016-0172-3

19. Theis KR, Romero R, Winters AD, Greenberg JM, Gomez-Lopez N, Alhousseini A, et al. Does the human placenta delivered at term have a microbiota? Results of cultivation, quantitative real-time PCR, 16S rRNA gene

CONFIDENTIAL

sequencing, and metagenomics. American journal of obstetrics and gynecology. 2019;220(3):267 e1- e39. doi: 10.1016/j.ajog.2018.10.018

20. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol. 2014;12:87. doi: 10.1186/s12915-014-0087-z

21. Lim ES, Rodriguez C, Holtz LR. Amniotic fluid from healthy term pregnancies does not harbor a detectable microbial community. Microbiome. 2018;6(1):87. doi: 10.1186/s40168-018-0475-7

22. de Goffau MC, Lager S, Sovio U, Gaccioli F, Cook E, Peacock SJ, et al. Human placenta has no microbiome but can contain potential pathogens. Nature. 2019. doi: 10.1038/s41586-019-1451-5

23. Britt WJ. Maternal Immunity and the Natural History of Congenital Human Cytomegalovirus Infection. Viruses. 2018;10(8). doi: 10.3390/v10080405

24. Njue A, Coyne C, Margulis AV, Wang D, Marks MA, Russell K, et al. The Role of Congenital Cytomegalovirus Infection in Adverse Birth Outcomes: A Review of the Potential Mechanisms. Viruses. 2020;13(1). doi: 10.3390/v13010020

25. McDonagh S, Maidji E, Ma W, Chang HT, Fisher S, Pereira L. Viral and bacterial pathogens at the maternal-fetal interface. J Infect Dis. 2004;190(4):826-34. doi: 10.1086/422330

26. Gaccioli F, Lager S, de Goffau MC, Sovio U, Dopierala J, Gong S, et al. Fetal inheritance of chromosomally integrated human herpesvirus 6 predisposes

CONFIDENTIAL

the mother to pre-eclampsia. Nat Microbiol. 2020;5(7):901-8. doi: 10.1038/s41564-020-0711-3

27. Caserta MT, Hall CB, Schnabel K, Lofthus G, McDermott MP. Human herpesvirus (HHV)-6 and HHV-7 infections in pregnant women. The Journal of infectious diseases. 2007;196(9):1296-303. doi: 10.1086/522430

28. Hall CB, Caserta MT, Schnabel KC, Shelley LM, Carnahan JA, Marino AS, et al. Transplacental congenital human herpesvirus 6 infection caused by maternal chromosomally integrated virus. The Journal of infectious diseases. 2010;201(4):505-7. doi: 10.1086/650495

29. Ando Y, Kakimoto K, Ekuni Y, Ichijo M. HHV-6 infection during pregnancy and spontaneous abortion. Lancet. 1992;340(8830):1289. doi: 10.1016/0140-6736(92)92990-w

30. Revest M, Minjolle S, Veyer D, Lagathu G, Michelet C, Colimon R. Detection of HHV-6 in over a thousand samples: new types of infection revealed by an analysis of positive results. J Clin Virol. 2011;51(1):20-4. doi: 10.1016/j.jcv.2011.02.001

31. Drago F, Broccolo F, Javor S, Drago F, Rebora A, Parodi A. Evidence of human herpesvirus-6 and -7 reactivation in miscarrying women with pityriasis rosea. Journal of the American Academy of Dermatology. 2014;71(1):198-9. doi: 10.1016/j.jaad.2014.02.023

32. Das BB, Rakheja D, Lacelle C, Sedlak RH, Gulve N, Chowdhury SR, et al. Possible progesterone-induced gestational activation of chromosomally integrated human herpesvirus 6B and transplacental transmission of activated human

CONFIDENTIAL

herpesvirus 6B. The Journal of heart and lung transplantation : the official publication of the International Society for Heart Transplantation. 2016;35(11):1373-6. doi: 10.1016/j.healun.2016.08.009

33. Avanzi S, Leoni V, Rotola A, Alviano F, Solimando L, Lanzoni G, et al. Susceptibility of human placenta derived mesenchymal stromal/stem cells to human herpesviruses infection. PloS one. 2013;8(8):e71412. doi: 10.1371/journal.pone.0071412

34. Koi H, Zhang J, Parry S. The mechanisms of placental viral infection. Annals of the New York Academy of Sciences. 2001;943:148-56. doi:

35. Arechavaleta-Velasco F, Ma Y, Zhang J, McGrath CM, Parry S. Adeno-associated virus-2 (AAV-2) causes trophoblast dysfunction, and placental AAV-2 infection is associated with preeclampsia. The American journal of pathology. 2006;168(6):1951-9. doi: 10.2353/ajpath.2006.050781

36. Hermonat PL, Kechelava S, Lowery CL, Korourian S. Trophoblasts are the preferential target for human papilloma virus infection in spontaneously aborted products of conception. Human pathology. 1998;29(2):170-4. doi:

37. Tanaka K, Chowdhury K, Chang KS, Israel M, Ito Y. Isolation and characterization of polyoma virus mutants which grow in murine embryonal carcinoma and trophoblast cells. The EMBO journal. 1982;1(12):1521-7. doi:

38. iHMP Consortium. The Integrative Human Microbiome Project: dynamic analysis of microbiome-host omics profiles during periods of human health and disease. Cell host & microbe. 2014;16(3):276-89. doi: 10.1016/j.chom.2014.08.014

CONFIDENTIAL

39. Gevers D, Knight R, Petrosino JF, Huang K, McGuire AL, Birren BW, et al. The Human Microbiome Project: a community resource for the healthy human microbiome. PLoS biology. 2012;10(8):e1001377. doi: 10.1371/journal.pbio.1001377

40. Human Microbiome Project C. A framework for human microbiome research. Nature. 2012;486(7402):215-21. doi: 10.1038/nature11209

41. Li F, Chen C, Wei W, Wang Z, Dai J, Hao L, et al. The metagenome of the female upper reproductive tract. Gigascience. 2018;7(10). doi: 10.1093/gigascience/giy107

42. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-20. doi: 10.1093/bioinformatics/btu170

43. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome biology. 2014;15(3):R46. doi: 10.1186/gb-2014-15-3-r46

44. Walker MA, Pedamallu CS, Ojesina AI, Bullman S, Sharpe T, Whelan CW, et al. GATK PathSeq: a customizable computational tool for the discovery and identification of microbial sequences in libraries from eukaryotic hosts. Bioinformatics. 2018;34(24):4287-9. doi: 10.1093/bioinformatics/bty501

45. Wickham H, Averick M, Bryan J, Chang W, McGowan L, Francois R, et al. Welcome to the Tidyverse. Journal of Open Source Software. 2019;4(43):1686. doi: https://doi.org/10.21105/joss.01686

CONFIDENTIAL

46. Nooij S, Schmitz D, Vennema H, Kroneman A, Koopmans MPG. Overview of Virus Metagenomic Classification Methods and Their Biological Applications. Frontiers in microbiology. 2018;9:749. doi: 10.3389/fmicb.2018.00749

47. Li Z, Breitwieser FP, Lu J, Jun AS, Asnaghi L, Salzberg SL, et al. Identifying Corneal Infections in Formalin-Fixed Specimens Using Next Generation Sequencing. Investigative ophthalmology & visual science. 2018;59(1):280-8. doi: 10.1167/iovs.17-21617

48. Barrientos-Somarribas M, Messina DN, Pou C, Lysholm F, Bjerkner A, Allander T, et al. Discovering viral genomes in human metagenomic data by predicting unknown protein families. Scientific reports. 2018;8(1):28. doi: 10.1038/s41598-017-18341-7

49. Tyschik EA, Shcherbakova SM, Ibragimov RR, Rebrikov DV. Transplacental transmission of torque teno virus. Virol J. 2017;14(1):92. doi: 10.1186/s12985-017-0762-0

50. Wylie KM, Mihindukulasuriya KA, Zhou Y, Sodergren E, Storch GA, Weinstock GM. Metagenomic analysis of double-stranded DNA viruses in healthy adults. BMC Biol. 2014;12:71. doi: 10.1186/s12915-014-0071-7

51. Forslund O, Johansson H, Madsen KG, Kofoed K. The nasal mucosa contains a large spectrum of human papillomavirus types from the Betapapillomavirus and Gammapapillomavirus genera. J Infect Dis. 2013;208(8):1335-41. doi: 10.1093/infdis/jit326

CONFIDENTIAL

52. Sebghati M, Khalil A. New evidence on prognostic features, prevention and treatment of congenital Cytomegalovirus infection. Curr Opin Obstet Gynecol. 2020;32(5):342-50. doi: 10.1097/GCO.0000000000000651

53. Mocarski ES, Shenk T, Griffiths PD, Pass RF. Cytomegaloviruses In: Knipe DM, Howley PM, editors. Fields Virology. 2. 6th ed. New York, NY: Lippincott, Williams & Wilkins; 2015. p. 1960-2015.

54. Di Stefano M, Calabro ML, Di Gangi IM, Cantatore S, Barbierato M, Bergamo E, et al. In vitro and in vivo human herpesvirus 8 infection of placenta. PloS one. 2008;3(12):e4073. doi: 10.1371/journal.pone.0004073

CONFIDENTIAL

Figure 1

Click here to access/download;Figure;figure-1.tif

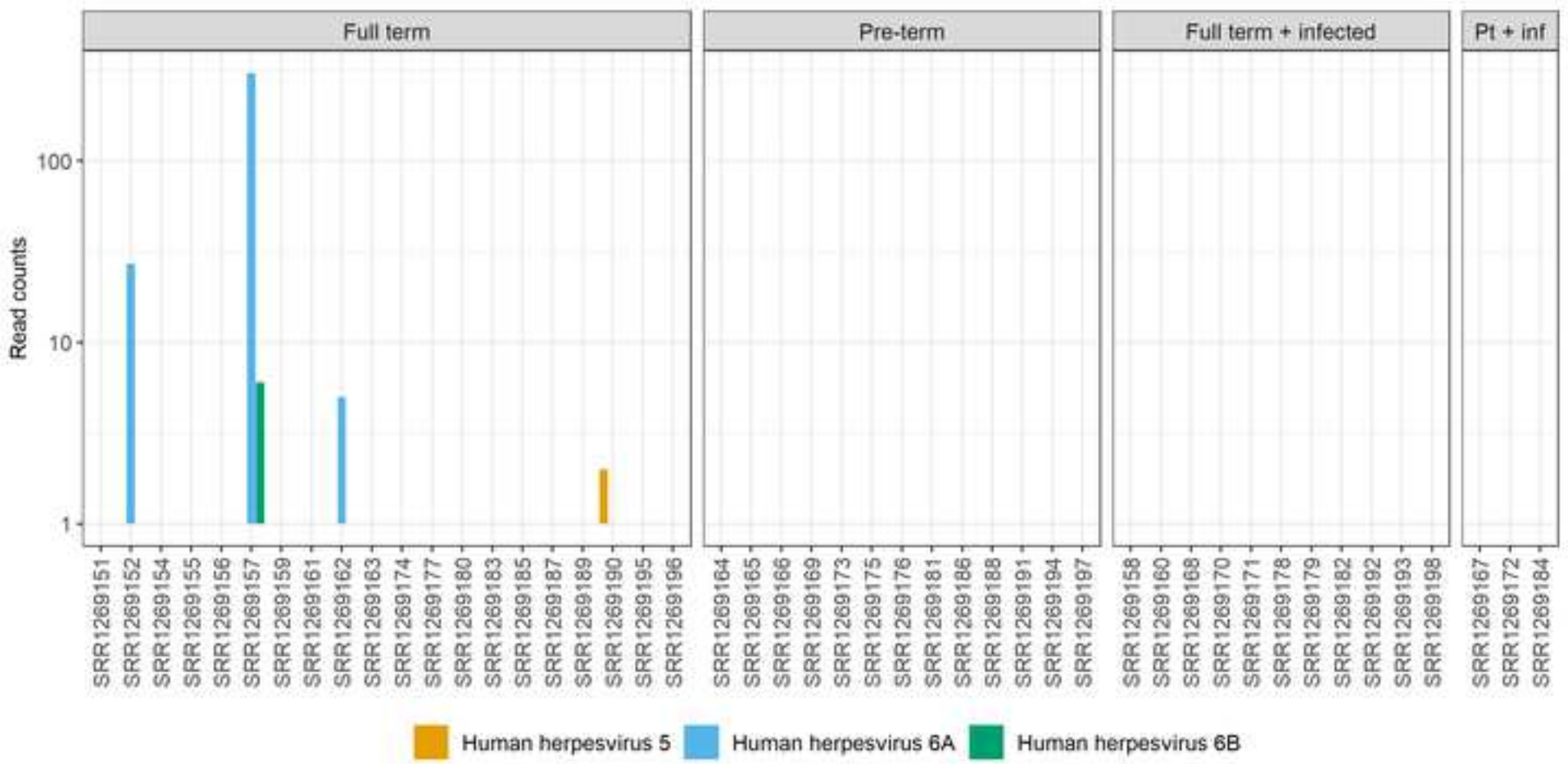


Figure 2

Click here to access/download;Figure;figure-2.tif

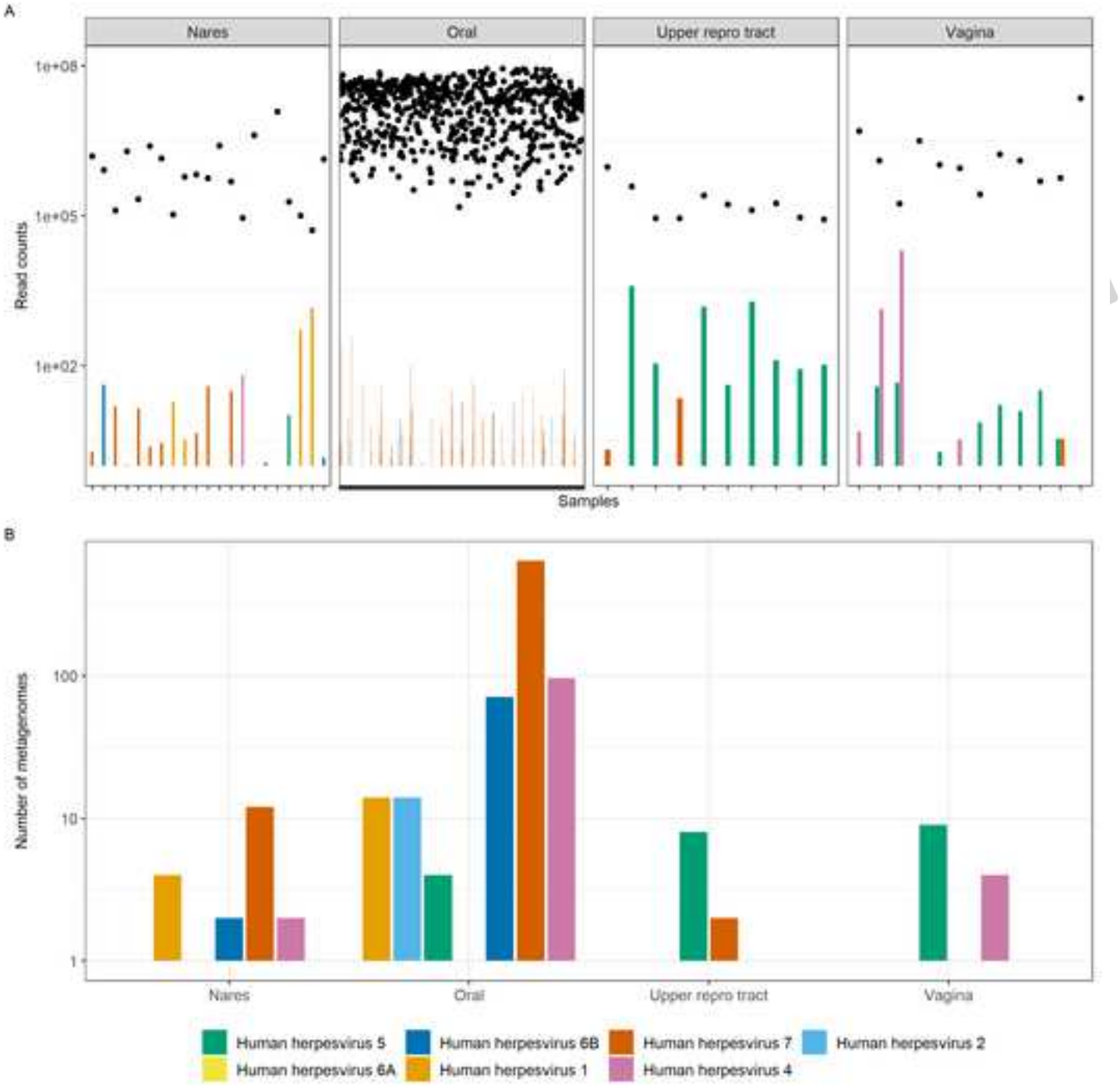
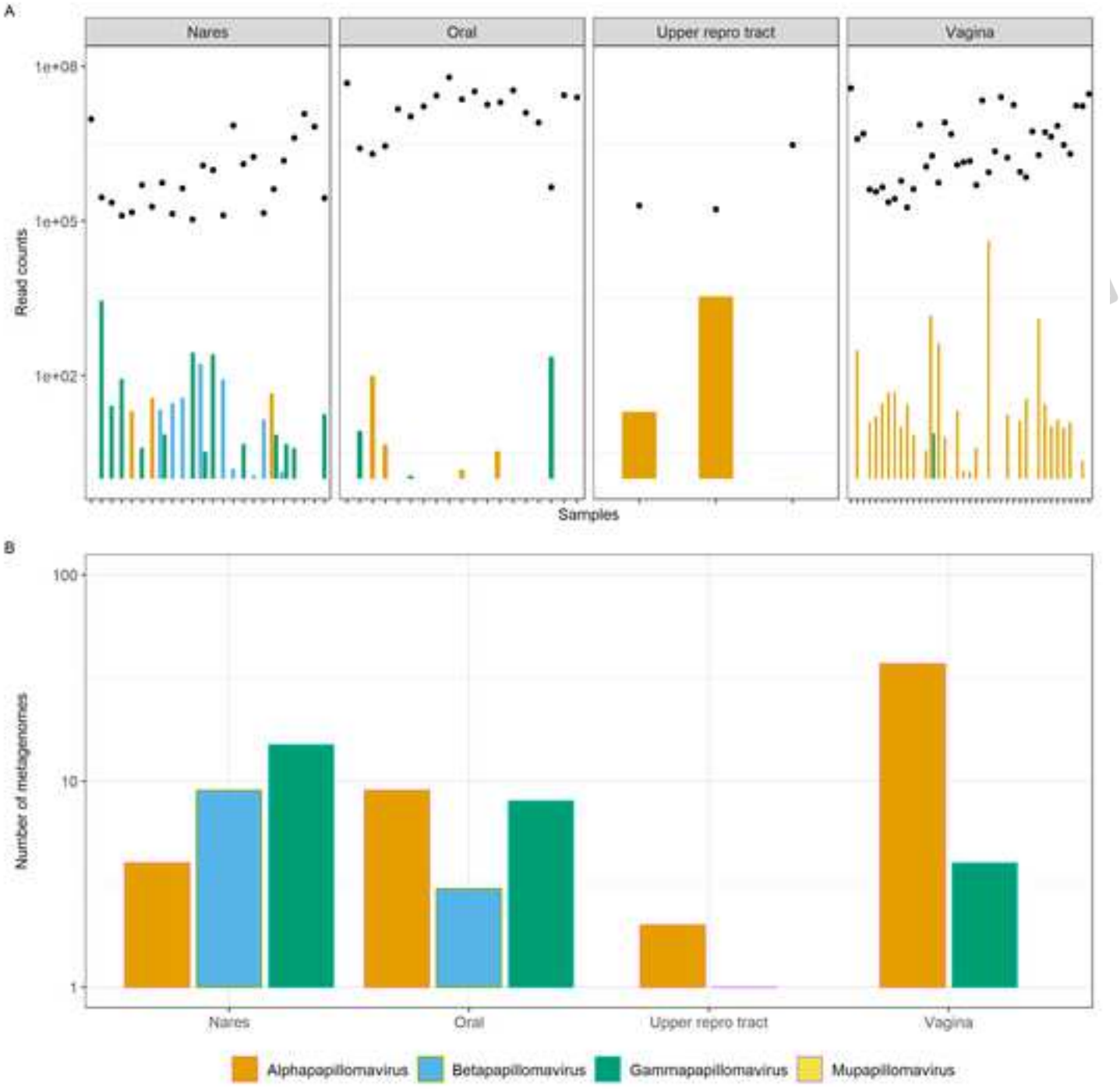


Figure 3

Click here to access/download;Figure;figure-3.tif



Click here to access/download;Supplementary Material;Supplementary Data.pdf

Supplementary Data

Metagenomic profiling of placental tissue suggests DNA virus infection of the placenta is rare

Adam A Witney, Sean Aller and Blair L Strang

CONFIDENTIAL

Supplementary Figure 1 Number of reads corresponding to phiX174 in the placental metagenomes. The number of reads corresponding to phiX174 in each placental metagenomic sample was determined using Kraken2. The sample accession of each metagenomic sample was noted on each x-axis. Metagenomic samples came from either full-term pregnancies, pre-term pregnancies, full term pregnancies that had reported antenatal infection or pre-term pregnancies that had reported antenatal infections (Pt + inf).

Supplementary Figure 2 Number of reads corresponding to human herpesvirus genomes in placental metagenomes using a human herpesvirus database. The number of reads corresponding to human herpesvirus genomes in each placental metagenomic sample was determined by comparing reads to a database of human herpesvirus genomes using Kraken2. The sample accession of each metagenomic sample was noted on the x-axis. Metagenomic samples came from either full-term pregnancies, pre-term pregnancies, full term pregnancies that had reported antenatal

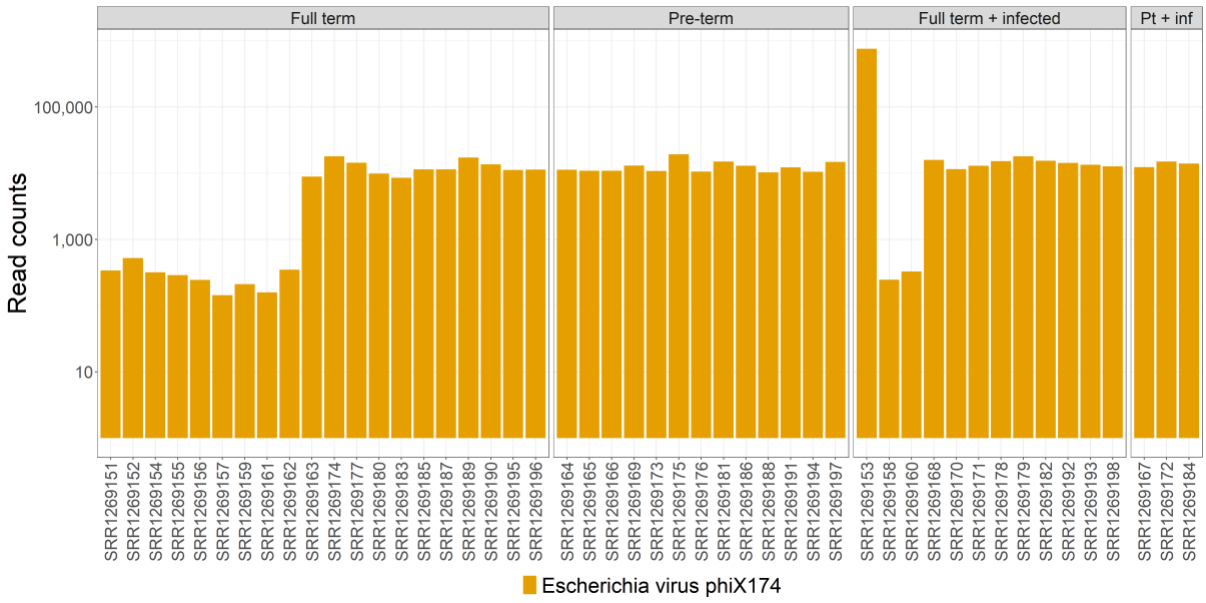
infection or pre-term pregnancies that had reported antenatal infections (Pt + inf). HCMV is shown as human herpesvirus 5.

Supplementary Figure 3 Number of reads corresponding to virus genomes in placental metagenomes using the GATK PathSeq tool. The number of reads corresponding to human viruses in each placental metagenomic sample was determined using GATK PathSeq. The sample accession of each metagenomic sample was noted on the x-axis. Metagenomic samples came from either full-term pregnancies, pre-term pregnancies, full term pregnancies that had reported antenatal infection or pre-term pregnancies that had reported antenatal infections (Pt + inf). HCMV is shown as human betaherpesvirus 5.

Table S1 Accession numbers for Human Microbiome Project data. See excel sheet the accompanies this manuscript.

CONFIDENTIAL

Supplementary figure 1

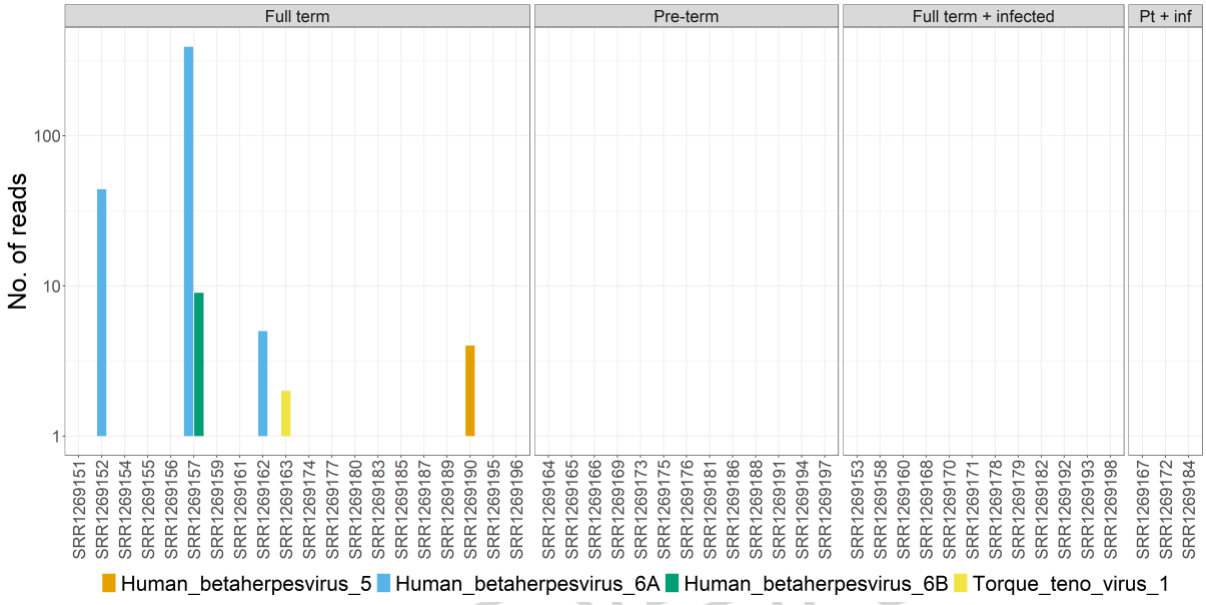


CONFIDENTIAL

Supplementary figure 2

CONFIDENTIAL

Supplementary figure 3



CONFIDENTIAL

CONFIDENTIAL

Click here to access/download

Supplementary Material - Excel file

Table S1 - Microbiome FINAL.xlsx