nature research

Corresponding author(s): S Behjati, GCS Smith, DS Charnock-Jones

Last updated by author(s): Feb 2, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was required to collect data.
Data analysis	BWA (0.7.17; https://github.com/lh3/bwa) CaVEMan (1.14.0; https://github.com/cancerit) Pindel (3.3.0; https://github.com/cancerit) ASCAT (4.3.3; https://github.com/cancerit) Battenberg (3.5.2; https://github.com/cancerit) BRASS (6.3.0; https://github.com/cancerit) MPBoot (1.1.0; https://github.com/diepthihoang/mpboot) HDP (0.1.5; https://github.com/nicolaroberts/hdp) mosdepth (0.3.0; https://github.com/brentp/mosdepth) R (3.6.1) deconstructSigs (1.8.0; CRAN) Bespoke R scripts used for analysis and visualisation in this study are available online from GitHub (https://github.com/TimCoorens/Placenta).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

DNA sequencing data are deposited in the European Genome-Phenome Archive (EGA) with accession code EGAD00001006337 (https://ega-archive.org/datasets/EGAD00001006337). Sample information and data on mutation burdens and signatures can be found in Supplementary Table 1. Further clinical information can be found in Supplementary Tables 2 and 3. Somatic mutations and embryonic mutations (bulk samples) can be found in Supplementary Tables 4 and 6, respectively. Calls of structural variants with associated copy number changes can be found in Supplementary Table 5.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample-size calculation was performed. Sample size was determined by the availability of tissue and cost of the experiment. There is no accepted method by which one could power an experiment principally performing phylogenetic analyses either.
Data exclusions	Some microdissected samples were excluded due to evidence of contamination by an unrelated patient's DNA during the library preparation process (PD42138b3_lo0001, PD42142b2_lo0003,PD45567e_lo0005). This exclusion criterion was pre-established. Outlier mesenchymal cores with an excessive mutation burden that showed high sharedness with trophoblast clusters when mapped onto the phylogenies were excluded as likely contaminated by the overlying trophoblast tissues (PD42138b4_lo0001, PD42138b_lo0005, PD42142b_lo0011). This exclusion criterion was not pre-established, as it only became evident after analysis of all mesenchymal core genomes.
Replication	The initial bulk sequencing experiment was conducted with placental biopsies from 6 healthy placentas. The microdissection experiment was first performed on 2 healthy placentas. The only discrepancy found between the initial and final groups in both of these experiments was that whole genomes sequenced on the HiSeq platform seemed to call fewer substitutions overall than those on the Novaseq platform although this does not affect the final conclusions drawn.
Randomization	Patients were allocated to groups based on previously validated clinical parameters (detailed in the extended data). Other covariates, such as basic demographics, were matched for, where possible, between clinical groups.
Blinding	The investigators were not blinded to group allocation. In doing so, we would have been unable to know which were the healthy placentas that we were to conduct the microdissection experiment with. Furthermore, the variant calling is group agnostic and, having established similar mutational profiles of samples across all groups, they are treated as one for the remaining analyses.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	Women of child-bearing age. Further clinical information is provided in Supplementary Tables 2-3.
Recruitment	The pregnancy outcome prediction study, from which these samples are taken, has previously published its recruitment sampling strategy has previously been described in papers published. Please see references within manuscript.
Ethics oversight	Ethical approval for this study was given by the Cambridgeshire 2 Research Ethics Committee (reference number 07/H0308/163) and all participants provided written informed consent.

Note that full information on the approval of the study protocol must also be provided in the manuscript.