

In vitro fertilization does not increase the incidence of de novo copy number alterations in fetal and placental lineages

Masoud Zamani Esteki^{1,2,3,16,17*}, Triin Viltrop^{4,16}, Olga Tšuiiko^{3,4,16}, Airi Tiirats⁵, Mariann Koel⁶, Margit Nõukas^{7,8}, Olga Žilina⁷, Katre Teearu⁷, Heidi Marjonen⁹, Hanna Kahila¹⁰, Jeroen Meekels¹, Viveca Söderström-Anttila¹¹, Anne-Maria Suikkari¹¹, Aila Tiitinen¹⁰, Reedik Mägi⁸, Sulev Kõks^{12,13}, Nina Kaminen-Ahola⁹, Ants Kurg⁷, Thierry Voet^{10,3,14,17*}, Joris Robert Vermeesch^{10,3,17*} and Andres Salumets^{10,5,8,10,15,17*}

Although chromosomal instability (CIN) is a common phenomenon in cleavage-stage embryogenesis following in vitro fertilization (IVF)^{1–3}, its rate in naturally conceived human embryos is unknown. CIN leads to mosaic embryos that contain a combination of genetically normal and abnormal cells, and is significantly higher in in vitro-produced preimplantation embryos as compared to in vivo-conceived preimplantation embryos⁴. Even though embryos with CIN-derived complex aneuploidies may arrest between the cleavage and blastocyst stages of embryogenesis^{5,6}, a high number of embryos containing abnormal cells can pass this strong selection barrier^{7,8}. However, neither the prevalence nor extent of CIN during prenatal development and at birth, following IVF treatment, is well understood. Here we profiled the genomic landscape of fetal and placental tissues postpartum from both IVF and naturally conceived children, to investigate the prevalence and persistence of large genetic aberrations that probably arose from IVF-related CIN. We demonstrate that CIN is not preserved at later stages of prenatal development, and that de novo numerical aberrations or large structural DNA imbalances occur at similar rates in IVF and naturally conceived live-born neonates. Our findings affirm that human IVF treatment has no detrimental effect on the chromosomal constitution of fetal and placental lineages.

CIN is a common event in human cleavage-stage embryogenesis, resulting in a high frequency of chromosomal mosaicism in IVF embryos^{1–3}, and probably provides an explanation for the low success rate per embryo transfer of IVF⁹. Strikingly, most IVF embryos are mixtures of euploid and aneuploid cells and only a low fraction

of IVF embryos contains a normal karyotype in all blastomeres^{1,10}. Between the cleavage and blastocyst stages of embryogenesis, CIN may cause complex aneuploidies and the arrest of many embryos^{5,6}. Even after this substantial selection, genetically aberrant cells can be preserved at high frequency in the developing blastocysts, such that about 40% of blastocysts have been reported to be uniformly euploid^{7,8}. Despite the high prevalence of human IVF embryos with aberrant chromosomal configurations, its clinical implications are poorly understood. Recently, we established *Bos taurus* as a model organism to study CIN at cleavage-stage embryogenesis¹¹ and showed that CIN in in vitro-produced embryos is significantly higher in prevalence when compared to in vivo-conceived embryos⁴, raising alarming concern about the effect of suboptimal in vitro environment on embryonic genetic integrity. However, the fate of aberrant cells after cleavage-stage embryogenesis is not clear and it remains impossible to investigate naturally conceived human embryos in vivo.

Cell lineages containing genetic aberrations in mosaic cleavage-stage embryos have been speculated to survive in the trophectoderm, having little or no influence on the composition of the inner cell mass and its subsequent fetal development¹². This possibility is in accordance with the confined placental mosaicism found in approximately 1–2% of ongoing natural pregnancies¹³ and the low prevalence of genetic mosaicism in postnatal and adult DNA samples¹⁴. Furthermore, euploid blastomeres may outgrow blastomeres with chromosomal aberrations¹⁵, allowing normal embryonic development and birth. This view of preferential selection of normal cells is supported by recent studies reporting the birth of healthy babies after intrauterine transfer of genetically mosaic IVF

¹Department of Clinical Genetics, Maastricht University Medical Centre, Maastricht, the Netherlands. ²Department of Genetics and Cell Biology, GROW School for Oncology and Developmental Biology, Maastricht University, Maastricht, the Netherlands. ³Centre for Human Genetics, University Hospital Leuven, Department of Human Genetics, KU Leuven, Leuven, Belgium. ⁴Department of Biomedicine, Institute of Biomedicine and Translational Medicine, University of Tartu, Tartu, Estonia. ⁵Department of Obstetrics and Gynecology, Institute of Clinical Medicine, University of Tartu, Tartu, Estonia. ⁶Department of Cell Biology, Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia. ⁷Department of Biotechnology, Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia. ⁸Estonian Genome Center, Institute of Genomics, University of Tartu, Tartu, Estonia. ⁹Department of Medical and Clinical Genetics, Medicum, University of Helsinki, Helsinki, Finland. ¹⁰Department of Obstetrics and Gynecology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland. ¹¹Väestöliitto Fertility Clinic, Helsinki, Finland. ¹²Center for Molecular Medicine and Innovative Therapeutics, Murdoch University, Murdoch, Western Australia, Australia. ¹³The Perron Institute for Neurological and Translational Science, Nedlands, Western Australia, Australia. ¹⁴Sanger-EBI Single Cell Genomics Centre, Wellcome Trust Sanger Institute, Cambridge, UK. ¹⁵Competence Centre on Health Technologies, Tartu, Estonia. ¹⁶These authors contributed equally: Masoud Zamani Esteki, Triin Viltrop, Olga Tšuiiko. ¹⁷These authors jointly supervised this work: Masoud Zamani Esteki, Thierry Voet, Joris Robert Vermeesch, Andres Salumets. *e-mail: masoud.zamaniesteki@mumc.nl; thierry.voet@kuleuven.be; joris.vermeesch@kuleuven.be; andres.salumets@ccht.ee

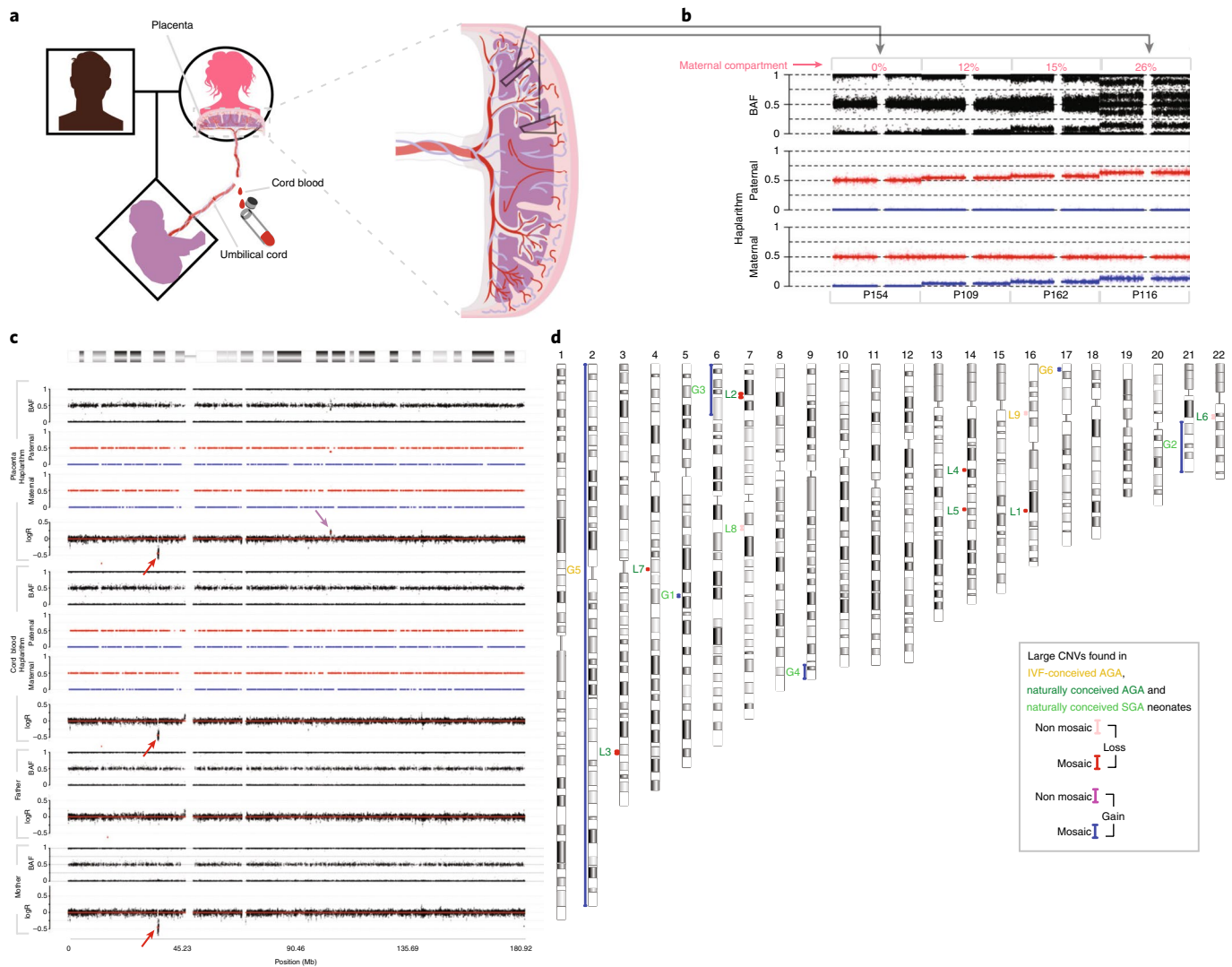


Fig. 1 | Haplithmis reveals genetic mosaicism with parent-of-origin information. **a**, Schematic representation of a quartet. **b**, Haplithmis can determine the maternal DNA contribution to bulk placental DNA sample (see Methods). **c**, A maternally inherited CNV (red arrows) and placental de novo CNV with paternal origin found on Chr 5 of P070 (purple arrow). BAF (B allele frequency) of a SNP is the signal intensity of allele B over signal intensities of both expected A and B alleles for that SNP ($BAF = \frac{B}{A+B}$), and logR (relative copy number) of that SNP is the base 2 logarithm of the summed normalized signal intensities of both alleles over expected signal intensity values ($\log R = \frac{R_{\text{sample}}}{R_{\text{expected}}}$). **d**, Overview of de novo CNVs found in this study (see also Source data). G, gains; L, losses (as shown in Table 1).

embryos^{16,17}. In addition, a recent report described lineage-specific fate in chimeric mouse embryos. That report demonstrated that different mechanisms act on aneuploid cells in fetal or placenta lineages such that apoptotic depletion of aneuploid cells occurs in fetal lineage, whereas aneuploid cells in placenta persist but slow down their divisions¹⁸.

To investigate the persistence of IVF-related CIN, we profiled de novo DNA copy number alterations in DNA samples extracted from cord blood cells and placental tissue, which are derived from embryonic and extra-embryonic cell lineages, respectively, of both IVF and naturally conceived neonates using haplithmis³ and molecular karyotyping. Importantly, the transferred embryos of IVF newborns were neither screened nor counterselected for chromosomal aneuploidy status before intrauterine transfer. Following live birth, we scrutinized genome-wide single nucleotide polymorphism (SNP) profiles in DNA from the father, mother, placenta and neonate umbilical cord blood of 111 families (quartets; Fig. 1a),

of which 49 and 62 quartets were from IVF and natural pregnancies, respectively. In the IVF cohort, the infertile patients underwent fresh or frozen IVF ($n=32$) or intracytoplasmic sperm injection (ICSI, $n=17$) embryo transfer, with about four out of five embryos transferred on days 2/3 ($n=38$, 77.5%) and the remainder on days 5/6 at the blastocyst stage ($n=11$, 22.5%). These quartets were further subdivided based on birth weight at delivery, with 48 IVF and 50 natural conception families categorized as ‘appropriate for gestational age’ (AGA), and one IVF and 12 natural conception families as ‘small for gestational age’ (SGA) (Supplementary Tables 1 and 2). DNA copy number variants (CNVs) were classified by whether they were found exclusively in a single sample or a few to all samples within a quartet. This approach empowered our analysis and allowed us to differentiate inherited CNVs (present in either or both parents, umbilical cord blood and placenta), de novo CNVs (present in placenta and umbilical cord blood only) and placental de novo CNVs (present in placenta only). CNVs were detected by

Table 1 | De novo CNVs detected by haplarithmis

Cohort	Family	Genome coordinates	Length (bp)	SNP probes (n)	O/T ^a	CN	Origin ^b	Mosaicism (%) ^c	Biopsy (CNV/bps) ^d
Naturally conceived: appropriate birth weight for gestational age	P017	Chr 16:78670954-78867839	196,885	37	OP/III/V	L1	Maternal	65	1/3
	P073	Chr 7:13238496-13348832	110,336	26	OP/III/V	L2	Paternal	60	1/3
	P080	Chr 3:174684649-174831994	147,345	32	OP/III/V	L3	Maternal	60	1/3
	P142	Chr 14:47624206-47746980	122,774	16	OP/III/V	L4	Paternal	70	1/3
	P162	Chr 14:64554517-64953726	399,209	189	OP/III/V	L5	Paternal	40	1/3
	F351	Chr 22:23250737-23644794	394,057	99	B/I	L6	Maternal	-	NA ^e
	F351	Chr 4:91709484-91910883	201,399	24	OP/III/V	L7	Maternal	70	NA ^e
Naturally conceived: small birth weight for gestational age	P070	Chr 5:103670162-104157752	487,590	45	OP/III/V	G1	Paternal	65	1/3
	P153	Chr 7:72305671-74115258	1,809,587	359	B/I	L8	Maternal	-	3/3
	P172	Chr 21:26058246-48084247	22,026,001	5,796	OP/III/V	G2	Paternal	25	1/5
	P172	Chr 6:204072-22569968	22,365,896	4,106	OP/III/V	G3	Paternal	10	1/5
	P172	Chr 9:135277529-141031439	5,753,910	2,915	OP/III/V	G4	Paternal	5	1/5
IVF: appropriate birth weight for gestational age	P091	Chr 16:21605180-21966869	361,689	41	B/I	L9	Maternal	-	3/3
	P106	Chr 2:1243048760	243,015,748	44,591	OP/III/V	G5	Maternal	10	5/5
	F1021	Chr 17:2832092-3070476	238,384	71	B/II	G6	Maternal	40	NA ^e

^a O/T denotes occurrence and type. Occurrence can be B, de novo copy number (CN) alteration occurring in both placenta and umbilical cord DNA samples; OP, de novo copy number alteration occurring in placental DNA sample only; and type can be in one of the stratified CNV types illustrated in Fig. 3. ^b Parent-of-origin is determined based on the parental allele affected; for example, maternal loss or gain means that the maternal allele is deleted or duplicated, respectively; for nonmosaic deletions present in both placenta and cord blood, we used the parent-of-origin module of siCHILD to determine parental origin. ^c Percentage of abnormal cells. ^d CNV/bps denotes number of CNVs found in spatially different placenta biopsies (bps) (see also Extended Data Figs. 2–4). ^e Further biopsies from these placentae were not available.

the combined use of relative copy number (logR) values, B allele frequency (BAF) and parent-of-origin haplotypes (haplarithms). The latter analysis allowed us to account for maternal DNA contribution in placental tissue (Fig. 1b and Extended Data Fig. 1), to identify the parental origin of de novo CNVs (Fig. 1c and Table 1) and to more accurately estimate the level of mosaicism of CNVs within the DNA sample (Fig. 2a,b and Table 1).

We detected de novo large CNVs (>100 kilobase pairs (kb)) in placental genomes of 12 families; of the 15 detected aberrations, four were also present in newborns and 11 were exclusively present in the placenta. We did not detect any CNVs present only in umbilical cord blood DNA, which would indicate a genomic alteration in the neonate only. The mosaic ($n = 12$) de novo CNVs were 5–70% mosaic (Table 1) and were, together with de novo nonmosaic CNVs ($n = 3$), randomly located across the genome (Fig. 1d). Our analysis showed comparable prevalence ($P = 0.327$ odds ratio test or $P = 0.775$ Pearson's chi-square test) of de novo CNVs in both cohorts with AGA: six of the 50 families with naturally conceived neonates (12.0%) and in three of the 48 families with IVF neonates (6.3%) (Table 1). When IVF with fresh and frozen embryo transfers was compared to ICSI fresh/frozen embryo transfers, similar proportions of mosaic de novo CNV cases were found (6.3%, 2/32 and 5.9%, 1/17, respectively, $P = 0.96$ odds ratio test), indicating that the modes of fertilization (IVF versus ICSI)

and embryo freezing are unlikely to influence the rate of large CNVs in fetal and placental lineages.

We found three partial trisomies with different levels of mosaicism in family P172 from the naturally conceived cohort, and a mosaic full-chromosome trisomy in family P106 from the IVF cohort (Fig. 2a,b, Table 1 and Extended Data Figs. 2 and 3), showing mosaicism of 5–25%. We obtained biopsies from two other spatially distinct locations, applied haplarithmis and confirmed the results using droplet digital PCR (ddPCR) (Fig. 2c and Supplementary Table 3). The partial mosaic trisomies on chromosome 6 (Chr6, 22.4Mb), Chr9 (5.8Mb) and Chr21 (22.0Mb) of P172 were not present in the other placental biopsies (Fig. 2c and Extended Data Fig. 2). In contrast, the full Chr2 mosaic trisomy was present in all P106 placental biopsies (Fig. 2c and Extended Data Fig. 3). This result indicates that the mosaic Chr2 trisomy arose as CIN in the early cleavage-stage embryo but segregated into the placental lineage only (Fig. 3 and Table 1), because it was present across different parts of the placenta with the same degree of mosaicism. The presence of the low-degree mosaic Chr2 trisomy reinforces an earlier observation in mouse embryos, which suggested a progressive reduction of aneuploid cell lineage in placenta¹⁸ because only 10% of the cells carried the anomaly. However, the segmental trisomies in family P172 were present in only one of the placental locations, indicating that these genomic alterations either emerged

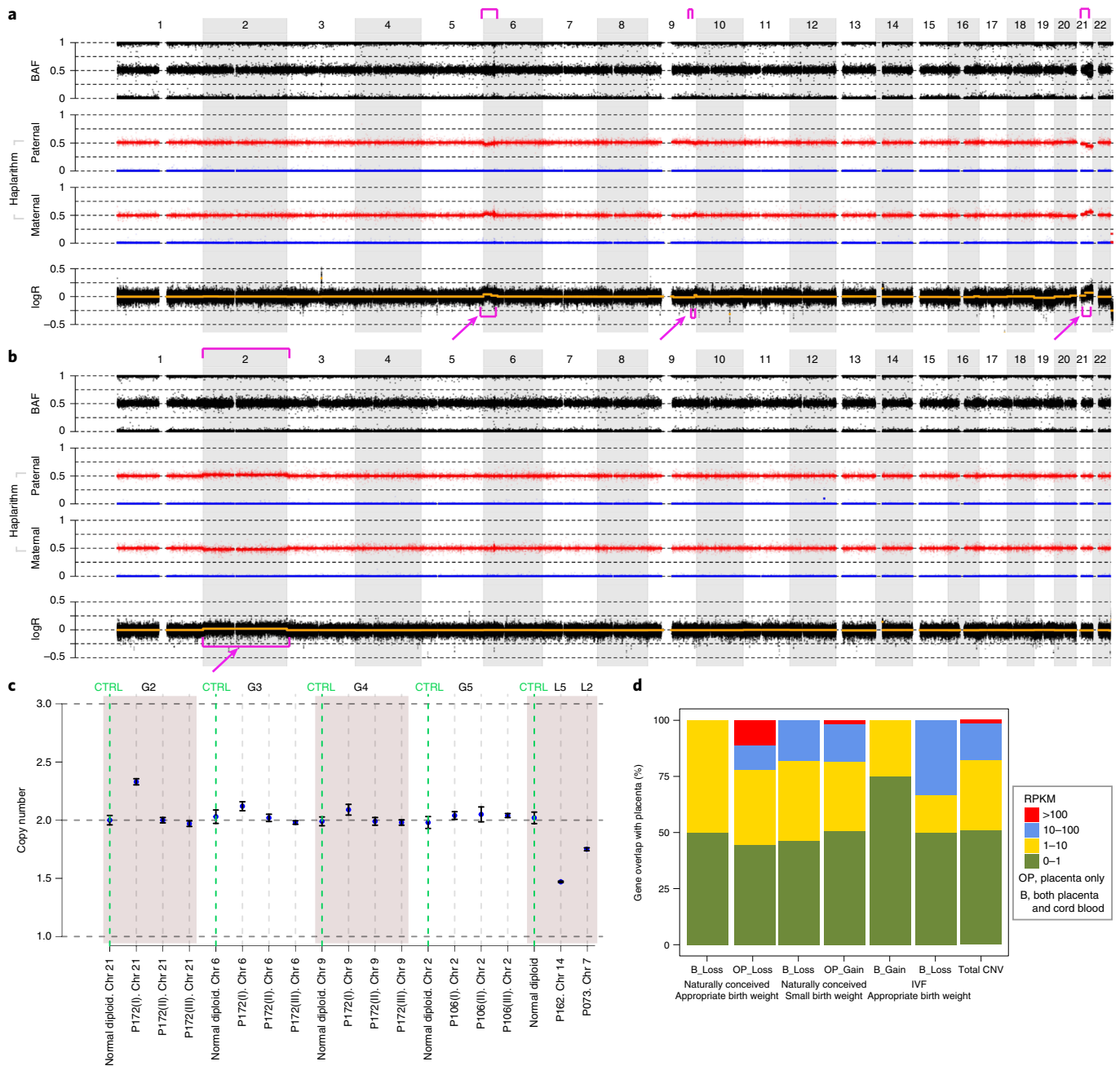


Fig. 2 | Mosaic de novo CNVs and overlap with placental transcriptome. a, b, Haplithms revealed de novo mosaic partial trisomies in P172 placenta of paternal origin (see also Extended Data Fig. 2) (**a**) and full-chromosome trisomy in P106 placenta of maternal origin (see also Extended Data Fig. 3) (**b**). These aberrations are indicated by purple arrows. **c,** Validation of detected mosaic de novo CNVs by ddPCR (each circle and error bar indicate mean and s.d., respectively, of four independent measurements). **d,** Placental expression of genes detected in de novo CNV regions, expressed as reads per kilobase per million mapped reads (RPKM). CTRL, control.

later in gestation through local clonal expansion of aberrant cells in the placenta or appeared early in development, contributed to the trophoblast and were largely overgrown by normal cells in placental tissue. The presence of multiple de novo DNA copy number variants has recently been shown to lead to genomic disorders¹⁹. The finding that multiple CNVs are present with a different degree of mosaicism, albeit in the placenta genome only, is indicative that they occurred in a cumulative fashion²⁰.

To further investigate CNV heterogeneity across the placenta and reaffirm the (non)mosaic nature of the detected de novo CNVs, we further analyzed 20 DNA samples derived from spatially different

locations of all but two placentae (F351 and F1021) with de novo CNVs (Table 1). As expected, de novo nonmosaic CNVs found in both placenta and cord blood were consistently present in all biopsies, whereas de novo mosaic CNVs were not present in the other biopsies (Extended Data Fig. 4). Importantly the full Chr2 trisomy was found in all biopsies (Extended Data Fig. 3), indicating its early embryonic origin (see above).

We identified two clinically notable nonmosaic CNVs in the SGA cohort, with genetic aberrations found in both umbilical cord and placental DNA. P153 carried a de novo deletion on Chr7q11.23. Postpartum follow-up revealed clinical features of Williams–Beuren

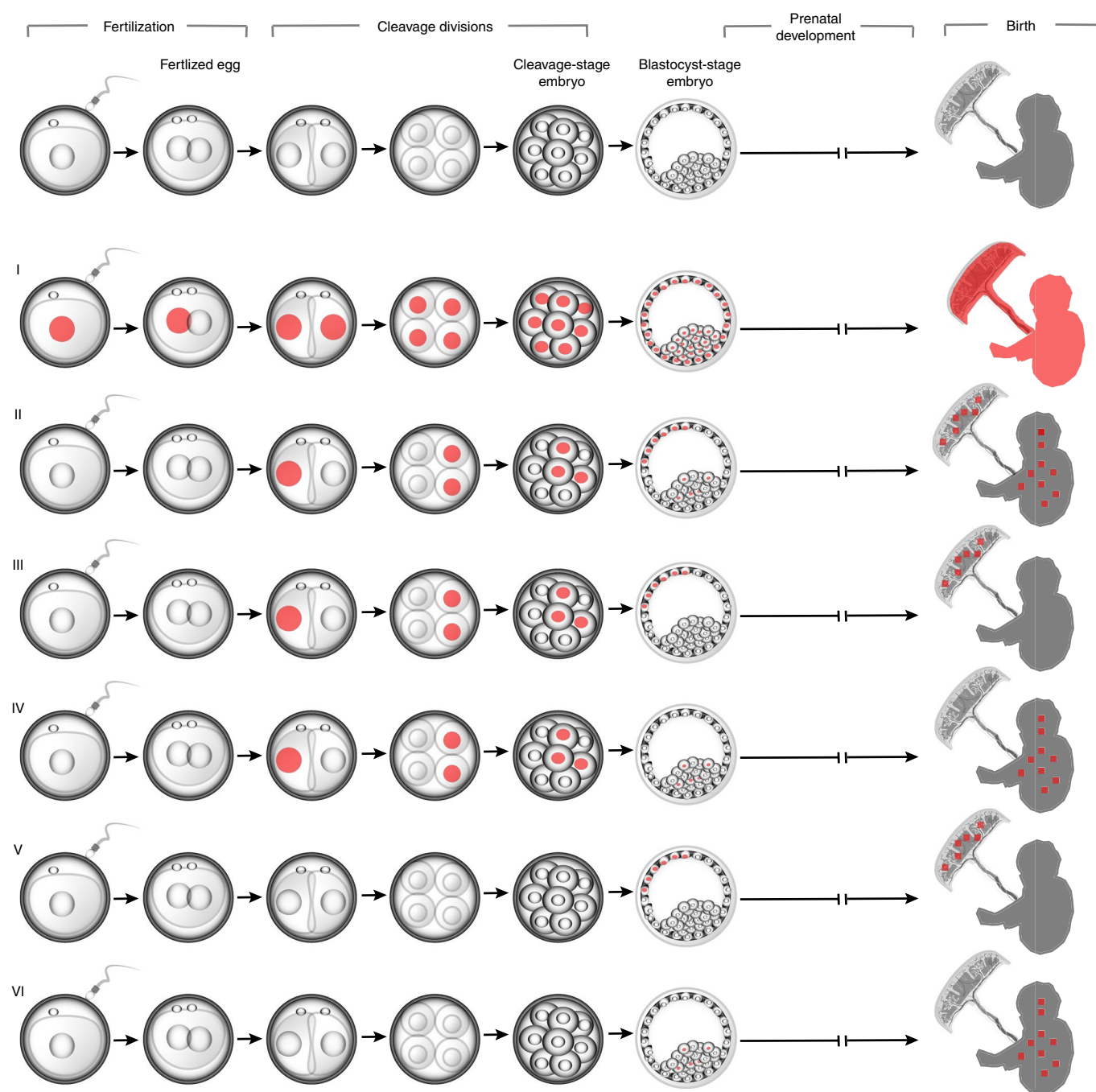


Fig. 3 | Schematic representation of plausible occurrence and segregation of de novo CNVs into fetal or placental lineages found in this study. De novo CNVs can occur before conception in the germ cells or during conception (I), during the cleavage divisions of early development (II–IV) or later during development (for example, blastocyst-stage embryogenesis (V–VI)). Note that, for simplicity in the CIN-derived mosaic CNVs, we merely show CNV burst in two-cell-stage embryos (types II–IV) and blastocyst-stage embryos (types V–VI) (see also Table 1). Red circles indicate cells carrying de novo CNV.

syndrome²¹. P177 carried a paternally inherited duplication on Chr 15 with features of Chr 15q13.3 duplication syndrome^{22,23}. The low birth weight of these neonates might be indicative of the underlying chromosomal aberration; therefore, SGA neonates should be subjected to submicroscopic chromosomal analysis.

Although the prevalence of de novo CNVs in SGA (25.0%, 3/12) and AGA (12.0%, 6/50) groups of the cohort of natural conceptions is not significantly different ($P=0.44$, odds ratio test), the current sample size for SGA ($n=12$) provides limited statistical power (<0.50). A study with increased SGA sample size ($n>44$; see Methods) is required to evaluate the effect of birth weight on

the prevalence of de novo CNVs. Furthermore, our data indicate that the current sample size ($n=50$ per AGA cohort) provides limited statistical power for detection of CNVs with small effect size ($w=0.1-0.2$) between the two cohorts. We estimated that future studies with smaller effect size would require a dataset of around tenfold larger²⁴.

The human placenta is characterized by deep invasion of trophoblasts into the endometrium. Trophoblasts have properties similar to cancer cells including rapid proliferation, migration and invasiveness. As CIN is a hallmark of both tumorigenesis and early embryogenesis, a recent study hypothesized that the placenta

resembles tumor tissue by demonstrating an extensive load of de novo genomic gains, harboring excessively expressed genes required for the invasiveness of trophoblasts²⁵. Our results challenge this view, because we found that placental de novo CNVs were present in only 12 of the 111 families (10.8%) (Fig. 1d and Table 1), and placental heterogeneity was found in 85.7% (6/7) of the studied 'placenta-only' de novo mosaic CNVs. To determine whether genes located in the de novo CNV regions ($n = 475$) had specific placental biology-related functions, we analyzed gene expression in full-term placentae following live birth (see Methods). Only 18.1% ($n = 86$) of these genes had moderate or high expression according to whole-RNA sequencing of placental tissue (Fig. 2d and Supplementary Table 4), and genes from regions with de novo gains (18.0%, 77/428) and losses (19.1%, 9/47, $P = 0.84$ odds ratio test) demonstrated similar proportion of moderate-to-high-expression genes in placental tissue. Moreover, enrichment analyses of genes located in de novo CNV regions did not reveal any molecular pathway related to placental biology. All this indicates that these new CNVs are probably sporadic and do not play any role in placental biology.

In this study, our main outcome measure was the extent and prevalence of de novo large CNVs (>100 kb) in live-born IVF-conceived neonates when compared to naturally conceived ones. Although we have not studied the embryonic loss with or without aneuploidies in IVF or in naturally conceived pregnancies, we speculate that when all cells of day-3 (cleavage-stage) embryos are investigated for the presence of (segmental) aneuploidies, on average 80% of the embryos present with at least one aneuploid cell (range, 70–90%)^{1,3,10}. The IVF success rate per embryo transfer of day-3 embryos is about 35% in the Helsinki and Tartu IVF centers. Assuming that all euploid day-3 embryos survive, these would constitute, at most, 20% of the total number of embryos transferred resulting in pregnancy/delivery, but indicating that at least 15% of all embryos did contain some aneuploid cells. This would amount to 60 and 40% of the live births resulting from, respectively, euploid and mosaic euploid/aneuploid cleavage-stage embryos. Conversely, about 20% (15/80) of mosaic embryos survive and result in healthy live births whereas 80% (65/80) are selected against. Since it is unlikely that all euploid embryos make it to term, the latter is a lower estimate.

Mosaicism rates in human IVF embryos are still debated and vary across studies, primarily because of the definition of mosaicism used, the technology applied and the variance in assisted reproduction protocols used. First, important differences in the definition of a mosaic embryo exist among studies. Embryos having both euploid and aneuploid cells were defined as mosaic in 1993 (ref. ²⁶) but, since then, an arbitrary classification of embryos as mosaic or diploid has influenced the mosaicism rate reported by different studies. For instance, some studies^{27–29} considered an embryo mosaic only if >50% of blastomeres within the embryo were aneuploid. The reason for considering an embryo having <50% abnormal cells as diploid was that they are probably viable, and a low percentage of abnormal cells in those embryos was believed clinically irrelevant. Second, since the discovery of aneuploidies in human preimplantation embryos^{30,31}, different technologies have been used to profile their chromosomal constitution. Depending on the technology used, the rate of mosaicism was reported to range from 15% (ref. ³²) to >90% (refs. ^{1,3}). However, in a meta-analysis of all technologies included and carried out on 815 embryos from different developmental stages, 73% were classified as mosaic¹⁰. High variability occurs because different technologies have different sensitivities and specificities for profiling the ploidy state of a selected chromosome (for example, FISH³²) or the entire genome (microarray^{1,3} or next-generation sequencing^{3,33}). The foregoing, and the number of cells per embryo analyzed with those technologies, can affect the reported embryo mosaicism rates in studies, with those studies analyzing all cells from a single embryo showing higher rates of mosaicism¹⁰. In the current study, we applied the same wet- and

dry-lab technologies as in our preimplantation embryo studies^{3,34,35} and further validated our approach using sensitive ddPCR and standard copy number analysis methods^{36,37}, allowing us to rule out technology-driven bias. Third, the mosaicism rate is reported to be influenced by the fertility clinic where the IVF procedure is carried out³⁸; one key component that can explain differences in the reported mosaicism rate is embryo culture conditions³⁹. However, in this study, the IVF procedure, including culture conditions, was similar throughout the participating IVF clinics.

Taken together, our findings suggest that the high level of chromosomal mosaicism in IVF cleavage-stage embryos is not preserved at later stages of prenatal development. Although we found a low-level mosaic trisomy in one IVF placental genome, it was not present in the fetal lineage of the same pregnancy, suggesting a proliferative defect of trisomic cells. The sporadic de novo genomic aberrations found in approximately 10.8% of pregnancies were scattered across the genome and represented random and rare events of embryonic CIN, without any observed functional consequence on placental biology or fetal health. Thus, our findings confirm that the IVF procedure has no detrimental effect on large genomic rearrangements in fetal and placental lineages of live-born neonates, which helps to remove a major health concern in regard to IVF neonates.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41591-019-0620-2>.

Received: 14 February 2019; Accepted: 20 September 2019;

Published online: 04 November 2019

References

- Vanneste, E. et al. Chromosome instability is common in human cleavage-stage embryos. *Nat. Med.* **15**, 577–583 (2009).
- Chavez, S. L. et al. Dynamic blastomere behaviour reflects human embryo ploidy by the four-cell stage. *Nat. Commun.* **3**, 1251 (2012).
- Zamani Esteki, M. et al. Concurrent whole-genome haplotyping and copy-number profiling of single cells. *Am. J. Hum. Genet.* **96**, 894–912 (2015).
- Tsuiko, O. et al. Genome stability of bovine in vivo-conceived cleavage-stage embryos is higher compared to in vitro-produced embryos. *Hum. Reprod.* **32**, 2348–2357 (2017).
- McCoy, R. C. et al. Evidence of selection against complex mitotic-origin aneuploidy during preimplantation development. *PLoS Genet.* **11**, e1005601 (2015).
- McCoy, R. C. et al. Common variants spanning PLK4 are associated with mitotic-origin aneuploidy in human embryos. *Science* **348**, 235–238 (2015).
- Fragouli, E. et al. Cytogenetic analysis of human blastocysts with the use of FISH, CGH and aCGH: scientific data and technical evaluation. *Hum. Reprod.* **26**, 480–490 (2011).
- Popovic, M. et al. Chromosomal mosaicism in human blastocysts: the ultimate challenge of preimplantation genetic testing? *Hum. Reprod.* **33**, 1342–1354 (2018).
- Vanneste, E. et al. What next for preimplantation genetic screening? High mitotic chromosome instability rate provides the biological basis for the low success rate. *Hum. Reprod.* **24**, 2679–2682 (2009).
- van Echten-Arends, J. et al. Chromosomal mosaicism in human preimplantation embryos: a systematic review. *Hum. Reprod. Update* **17**, 620–627 (2011).
- Destouni, A. et al. Zygotes segregate entire parental genomes in distinct blastomere lineages causing cleavage-stage chimerism and mixoploidy. *Genome Res.* **26**, 567–578 (2016).
- Ledbetter, D. H. Chaos in the embryo. *Nat. Med.* **15**, 490–491 (2009).
- Kalousek, D. K. & Vekemans, M. Confined placental mosaicism. *J. Med. Genet.* **33**, 529–533 (1996).
- Biesecker, L. G. & Spinner, N. B. A genomic view of mosaicism and human disease. *Nat. Rev. Genet.* **14**, 307–320 (2013).
- Santos, M. A. et al. The fate of the mosaic embryo: chromosomal constitution and development of day 4, 5 and 8 human embryos. *Hum. Reprod.* **25**, 1916–1926 (2010).

16. Greco, E., Minasi, M. G. & Fiorentino, F. Healthy babies after intrauterine transfer of mosaic aneuploid blastocysts. *N. Engl. J. Med.* **373**, 2089–2090 (2015).
17. Dimitriadou, E. et al. Principles guiding embryo selection following genome-wide haplotyping of preimplantation embryos. *Hum. Reprod.* **32**, 687–697 (2017).
18. Bolton, H. et al. Mouse model of chromosome mosaicism reveals lineage-specific depletion of aneuploid cells and normal developmental potential. *Nat. Commun.* **7**, 11165 (2016).
19. Liu, P. et al. An organismal CNV mutator phenotype restricted to early human development. *Cell* **168**, 830–842 e837 (2017).
20. Voet, T. & Vermeesch, J. R. Mutational processes shaping the genome in early human embryos. *Cell* **168**, 751–753 (2017).
21. Merla, G., Brunetti-Pierri, N., Micale, L. & Fusco, C. Copy number variants at Williams–Beuren syndrome 7q11.23 region. *Hum. Genet.* **128**, 3–26 (2010).
22. Szafranski, P. et al. Structures and molecular mechanisms for common 15q13.3 microduplications involving CHRNA7: benign or pathological? *Hum. Mutat.* **31**, 840–850 (2010).
23. van Bon, B. W. et al. Further delineation of the 15q13 microdeletion and duplication syndromes: a clinical spectrum varying from non-pathogenic to a severe outcome. *J. Med. Genet.* **46**, 511–523 (2009).
24. Chow S.-C., Shao, J. & Wang, H. *Sample Size Calculation in Clinical Research* (Marcel Dekker, 2003).
25. Kasak, L., Rull, K., Vaas, P., Teesalu, P. & Laan, M. Extensive load of somatic CNVs in the human placenta. *Sci. Rep.* **5**, 8342 (2015).
26. Delhanty, J. D. et al. Detection of aneuploidy and chromosomal mosaicism in human embryos during preimplantation sex determination by fluorescent in situ hybridisation (FISH). *Hum. Mol. Genet.* **2**, 1183–1185 (1993).
27. Munne, S. et al. Assessment of numeric abnormalities of X, Y, 18, and 16 chromosomes in preimplantation human embryos before transfer. *Am. J. Obstet. Gynecol.* **172**, 1191–1199 (1995), discussion 1199–1201.
28. Ziebe, S. et al. FISH analysis for chromosomes 13, 16, 18, 21, 22, X and Y in all blastomeres of IVF pre-embryos from 144 randomly selected donated human oocytes and impact on pre-embryo morphology. *Hum. Reprod.* **18**, 2575–2581 (2003).
29. Baart, E. B. et al. Preimplantation genetic screening reveals a high incidence of aneuploidy and mosaicism in embryos from young women undergoing IVF. *Hum. Reprod.* **21**, 223–233 (2006).
30. Steptoe, P. C. & Edwards, R. G. Birth after the reimplantation of a human embryo. *Lancet* **2**, 366 (1978).
31. Angell, R. R., Aitken, R. J., van Look, P. F., Lumsden, M. A. & Templeton, A. A. Chromosome abnormalities in human embryos after in vitro fertilization. *Nature* **303**, 336–338 (1983).
32. Harper, J. C. et al. Mosaicism of autosomes and sex chromosomes in morphologically normal, monospermic preimplantation human embryos. *Prenat. Diagn.* **15**, 41–49 (1995).
33. Voet, T. et al. Single-cell paired-end genome sequencing reveals structural variation per cell cycle. *Nucleic Acids Res.* **41**, 6119–6138 (2013).
34. Dimitriadou, E., Zamani Esteki, M. & Vermeesch, J. R. in *Methods in Molecular Biology* (ed. Kroneis, T.) 197–219 (Springer, 2015).
35. Destouni, A. et al. Genome-wide haplotyping embryos developing from OPN and 1PN zygotes increases transferrable embryos in PGT-M. *Hum. Reprod.* **33**, 2302–2311 (2018).
36. Wang, K. et al. PennCNV: an integrated hidden markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res.* **17**, 1665–1674 (2007).
37. Colella, S. et al. QuantiSNP: an objective Bayes hidden-Markov model to detect and accurately map copy number variation using SNP genotyping data. *Nucleic Acids Res.* **35**, 2013–2025 (2007).
38. Munne, S. et al. Euploidy rates in donor egg cycles significantly differ between fertility centers. *Hum. Reprod.* **32**, 743–749 (2017).
39. Munne, S. et al. Detailed investigation into the cytogenetic constitution and pregnancy outcome of replacing mosaic blastocysts detected with the use of high-resolution next-generation sequencing. *Fertil. Steril.* **108**, 62–71 e68 (2017).

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

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2019

Methods

Ethical approval. The study was approved by the local Ethics Committee of the University of Tartu (no. 213/T-21) and the Ethics Committee of Helsinki University Central Hospital (no. 285/13/03/03/2013). All couples signed the appropriate informed consent form.

Selection of participating families. In total, 111 couples participated in this study (Supplementary Tables 1 and 2), including 62 couples who delivered neonates after spontaneous conception (natural cohort) and 49 couples who delivered after IVF (IVF cohort). No previous genetic predisposition was identified in any of the couples participating in this study. Clinical data and biological material were collected from singleton pregnancies at birth from vaginal delivery or cesarean section (Supplementary Table 2). For each family, we isolated DNA from the father and mother, as well as from the placenta and neonate umbilical cord blood, following live birth (quartet DNA sample set). The natural cohort was divided into two subgroups based on neonate birth weight: (1) AGA neonates with a birth weight between the tenth and 90th percentiles ($n = 50$ families), and (2) SGA neonates with a birth weight below the tenth percentile ($n = 12$ families). The IVF cohort consisted of 48 families with AGA neonates and one family with an SGA neonate. Birth weight percentiles were calculated using data from the Estonian Medical Birth Registry⁴⁰ and the Finnish National Institute for Health and Welfare⁴¹.

Sample collection and DNA extraction. Full-thickness placental blocks of around 1 cm were taken from placenta and stored at -80°C for subsequent DNA extraction. To investigate placental CNV heterogeneity, we analyzed between three and five biopsies from spatially distinct locations of ten placentae (Table 1 and Extended Data Figs. 2–4). All samples were collected by the same medical personnel. Maturity and health of term placentae were confirmed by histological examination. A total of 9 ml peripheral venous blood from parents and 4 ml umbilical cord blood were collected into EDTA-containing tubes (BD Vacutainer, Becton, Dickinson and Co.). In Estonia, DNA was isolated from blood and placental tissue using the PureLink Genomic DNA Kit (Invitrogen, Life Technologies). DNA samples were eluted with 200 μl PureLink Genomic Elution Buffer (10 mM Tris-HCl pH 9.0, 0.1 mM EDTA) and stored at -20°C until analysis. Finnish parental DNA from blood samples was extracted by NucleoSpin Blood XL Kit (Macherey-Nagel), and DNA from umbilical cord blood by NucleoSpin Tissue Kit (Macherey-Nagel). NanoDrop 2000 (Thermo Fisher Scientific, Inc.) was used to assess DNA concentration and purity.

SNP genotyping. For SNP genotyping, the Infinium PsychArray-24 v.1.1 (Illumina, no. GEO: GPL22819) and PsychArray-24 v.1.3 BeadChip SNP arrays (Illumina, no. GEO: GPL26750) were used, which contain approximately 590,000 SNP markers with a median marker spacing of 5.5 kb. Genotype calls, SNP B-allele frequency values and logR values of all samples were computed using Illumina GenomeStudio software.

Haplithrithmis. We applied haplithrithmis³ on each quartet DNA sample. Discrete SNP calls derived from umbilical cord blood were used as seeds to phase the parental genotypes. Parental haplithrithmis were used to infer DNA copy number state, parent-of-origin and level of mosaicism of de novo copy number alterations in placental/newborn DNA. Levels of mosaicism were calculated as previously described⁴².

Haplithrithmis can be used to estimate the contributions of maternal and fetal genomes in placental DNA samples. We carried out a simulation analysis by blending maternal and fetal alleles in different proportions (from 1%_{mother}/99%_{child} to 99%_{mother}/1%_{child}) and applied haplithrithmis (Extended Data Fig. 5). This process allowed accurate determination of maternal and fetal contributions to placental tissue (Supplementary Table 5 and Extended Data Fig. 1).

Copy number profiling. We applied both PennCNV³⁶ (with 0.2% false-positive rate and 86% sensitivity for constitutional CNVs) and QuantiSNP³⁷ (with 0.1% false-positive rate for constitutional CNVs) algorithms, as well as the copy number module of siCHILD³ on the raw logR values, with minor modifications. Raw logR values were smoothed using a moving average window of five consecutive SNP probes, wave-corrected for GC% bias by a Lowess fit and normalized to a trimmed mean of normal diploid chromosomes. Normalized logR values were segmented by piecewise-constant fitting⁴³ ($\gamma = 14$). A copy number alteration was considered reliable if (1) more than six consecutive SNP probes supported the copy number change, (2) its length was >33 kb (sixfold 5.5-kb median distance between two probes) and (3) it was supported by raw or segmented parent-of-origin specific haplotypes, coined haplithrithmis. As a sanity check, we investigated whether there were signals detected only by haplithrithmis (Supplementary Table 6) or logR analysis (Supplementary Table 7). We did not include those in the analysis.

Copy number analysis was empowered by our quartet sample set and haplithrithmis, as we used all of the samples within a quartet to call a reliable genomic change. For each of the detected DNA copy number alterations in one sample (for example, placenta), we checked whether the same alteration was present in other samples of the same quartet. This approach allowed us to determine the parental origin of inherited CNVs, to detect de novo CNVs

(that is, in placental samples or in both placental and umbilical cord samples), to find CNVs exclusive to any of the samples (for example, present solely in the placental genome) and to reveal genotyping artifacts.

Other statistical and computational analysis. Since both predictor (with or without IVF treatment) and outcome (with or without de novo CNVs) are dichotomous, chi-squared is the most suitable test⁴⁴. To determine appropriate sample size, we performed a power analysis using chi-squared test with confidence level of at least 95% and feeding the conventional effect size (w) values of 0.1, 0.3 and 0.5, which correspond to small, medium and large effect sizes, respectively⁴⁵. With a medium effect size, at least 40 samples per cohort are required to reach the power of 0.8 (ref. ³⁴). Given the sample size of about 50 per cohort (50 naturally conceived quartets and 48 IVF quartets with appropriate birth weight for gestational age), this study has $>80\%$ power ($\alpha = 0.05$, middle effect size = 0.3; case-to-control ratio ≈ 1) to detect a significant difference (<0.05).

We also compared large CNVs with mosaic or nonmosaic nature separately. The rates of mosaic large CNVs found in naturally versus IVF-conceived AGA groups were not significantly different ($P = 0.157$, Pearson's chi-square test). Similarly, the rates of nonmosaic CNVs found in naturally versus IVF-conceived AGA groups were not significantly different ($P = 0.977$, Pearson's chi-square test).

We applied Hotelling's T^2 multivariate statistical test on the numerical demographic features of AGA neonates of both IVF and natural cohorts, including birth weight, birth length, placenta weight, parity and gestational age. This analysis showed no significant difference in these cohorts ($P = 0.22$, Hotelling's two-sample T^2 -test).

For statistical analyses we applied R (<http://www.r-project.org/>). For visualization purposes we applied siCHILD³ and R (<http://www.r-project.org/>).

RNA sequencing and comparison to placental transcriptome. We performed RNA sequencing as described previously⁴⁶ (Supplementary Table 4). Briefly, 12 placental samples were collected from singleton pregnancies at term (range, 38–42 weeks of gestational age). Placental samples from different families were used for RNA sequencing and CNV analysis. Full-thickness placental blocks of around 2–3 cm were taken from three different regions of placenta. Total RNA was extracted with the mirVana miRNA Isolation Kit (Ambion, Life Technologies), and RNA samples were pooled. DNA contamination was removed with the DNA-free Kit (Invitrogen, Life Technologies). Agilent 2100 Bioanalyzer and RNA 6000 Nano Kit (Agilent Technologies) were used to assess the integrity and concentration of total RNA samples.

For RNA-seq library preparation, 45 ng total RNA was amplified with the Ovation RNA-seq System V2 Kit (NuGen). The SOLiD 5500 System DNA fragment library (Life Technologies) was used for library preparation. Pooled samples were analyzed on three FlowChip lanes, and approximately 40 million mappable paired-end reads were obtained per sample. DESeq (v.3.0.2) was used to calculate reads per kilobase per million (RPKM). All genes from placenta were grouped into four classes according to their RPKM value (0–1, 1–10, 10–100 and >100). Genes with moderate to high expression were defined by more than ten mapped reads per RPKM and treated as placenta-expressed genes.

Protein-coding genes from de novo CNV regions were extracted from the University of California, Santa Cruz browser (<https://genome.ucsc.edu/>). Human genome build GRC37/hg19 was used as a reference. Genes found at de novo CNV regions were studied for gene expression in placental tissue from 12 individuals. Enrichment analyses for Gene Ontology terms were performed for genes present at de novo CNVs, using g:Profiler (v.Ensembl 90, Ensembl Genomes 37, rev 1741, build date 19 October 2017)⁴⁷ with default settings. Briefly, this tool performs gene set enrichment analyses for Gene Ontology, Human Phenotype Ontology and Kyoto Encyclopedia of Gene and Genomes gene sets. Enrichment analyses were conducted by hypergeometric test, and resulting enrichment P values were further adjusted for multiple testing by the g:SCS (Set Counts and Sizes) method, developed to address the nonindependent structure of tested Gene Ontology terms⁴⁷. Separate enrichment analyses were conducted for different groups of genes based on cohort, CNV type and expression in placenta (all genes and placenta-expressed genes). Results were filtered hierarchically, and only the most significant ontology per parent group is shown. De novo CNVs with gains and losses were analyzed separately.

Validation by ddPCR. We applied ddPCR to validate low-level mosaic aberrations (Fig. 2c and Supplementary Table 3). We used a proof-of-principle assay to evaluate the sensitivity of TaqMan-based ddPCR. A DNA sample from a trisomy 21 (copy number, 3) cell line was mixed with a DNA sample from a normal diploid cell line (copy number, 2) at different ratios, creating admixture series of DNA samples with 100, 75, 50, 25, 10–15 and 0% abnormal alleles (Extended Data Fig. 6). The cell lines of Epstein–Barr virus-immortalized lymphocytes had been established in-house using blood samples from patients at the Centre for Human Genetics, UZ Leuven, Leuven, Belgium, and were authenticated by conventional karyotyping. Mycoplasma PCR was performed to ensure that the cell lines were not contaminated⁴⁸. To evaluate the level of Chr21 mosaic gain, the FAM-fluorescent TaqMan copy number assay probe for the *RUNX1* gene locus (dHsaCP1000308) was used as a target probe. The HEX-fluorescent TaqMan copy number reference

assay *AP3B1* probe (dHsaCP2500348) was used as an internal reference control (Bio-Rad).

The ddPCR system was operated according to the manufacturer's instructions. Briefly, before droplet generation, restriction enzyme HaeIII (New England Biolabs, Inc.) was added directly to the ddPCR reaction mixture and incubated for 20 min at room temperature. Subsequently, droplets were generated in a DG8 disposable droplet generator cartridge using a QX100 Droplet Generator (Bio-Rad). PCR amplification was performed using the following conditions: 95 °C for 10 min, followed by 40 cycles of 94 °C for 30 s and 59 °C for 1 min and, finally, 98 °C for 10 min. Droplets were counted using the QX100 Droplet Reader. Data were analyzed using QuantaSoft Software (Bio-Rad).

Mosaic ratio was calculated based on the total positive signal counts of *RUNX1* normalized against reference *AP3B*-positive signal counts. All analyzed samples were run in four-replicate reactions. Mosaic placental DNA samples were processed in the same manner by performing initial DNA digestion, followed by droplet generation and PCR amplification. All DNA samples were quantified using the TaqMan copy number assay target probes for the specific gene in the region of interest. The *AP3B1* probe was used as an internal reference control (Supplementary Table 3).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All SNP array data generated in this study were deposited in the NCBI Gene Expression Omnibus under accession no. [GSE93353](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE93353).

Code availability

Custom code is available from the author upon reasonable request.

References

- Sildver, K., Veerus, P. & Lang, K. Birth weight percentiles and factors associated with birth weight: a registry-based study in Estonia. *Eesti Arst*. **94**, 465–470 (2015).
- Sankilampi, U., Hannila, M. L., Saari, A., Gissler, M. & Dunkel, L. New population-based references for birth weight, length, and head circumference in singletons and twins from 23 to 43 gestation weeks. *Ann. Med.* **45**, 446–454 (2013).
- Conlin, L. K. et al. Mechanisms of mosaicism, chimerism and uniparental disomy identified by single nucleotide polymorphism array analysis. *Hum. Mol. Genet.* **19**, 1263–1275 (2010).
- Nilsen, G. et al. Copynumber: efficient algorithms for single- and multi-track copy number segmentation. *BMC Genomics* **13**, 591 (2012).
- Hulley, S. B., Cummings, S. R., Browner, W. S., Grady, D. G. & Newman, T. B. *Designing Clinical Research* (Lippincott Williams & Wilkins, 2015).
- Cohen, J. *Statistical Power Analysis for the Behavioral Sciences* (Lawrence Erlbaum Associates, 1988).
- Metsalu, T. et al. Using RNA sequencing for identifying gene imprinting and random monoallelic expression in human placenta. *Epigenetics* **9**, 1397–1409 (2014).
- Reimand, J. et al. g:Profiler—a web server for functional interpretation of gene lists (2016 update). *Nucleic Acids Res.* **44**, W83–W89 (2016).
- Spaepen, M., Angulo, A. E., Marynen, P. & Cassiman, J. J. Detection of bacterial and mycoplasma contamination in cell cultures by polymerase chain reaction. *FEMS Microbiol. Lett.* **78**, 89–94 (1992).

Acknowledgements

We gratefully thank all families that participated in this study in Estonia and Finland. This research was funded by an institutional research grant from the Estonian Ministry of Education and Research (no. IUT34-16 to A.S.); Enterprise Estonia (grant no. EU48695 to A.S.); the Horizon 2020 innovation (WIDENLIFE) (grant no. EU692065 to A.K.); the European Union's FP7 Marie Curie Industry-Academia Partnerships and Pathways (grant no. EU324509 to A.S.); the Helsinki University Hospital fund (to A.Tiitinen); the Faculty of Medicine, University of Helsinki fund (to N.K.-A.); the EVA (Erfelijkheid Voortplanting & Aanleg) specialty program fund of Maastricht University Medical Centre (MUMC+) (to M.Z.E.); the Estonian Research Council (grant nos. IUT20-60 and IUT24-6); the European Union through the European Regional Development Fund Project (nos. 2014-2020.4.01.15-0012 GENTRANSMED and 2014-2020.4.01.16-0125 to R.M.); and KU Leuven funding (no. C1/018) and FWO grant (no. G.0392.14N to J.R.V. and T.Voet). We thank B. de Greef, A. van Montfoort and N. Davarzani for statistical consultations.

Author contributions

M.Z.E., A.K., T.Voet, J.R.V. and A.S. conceived the study and designed the experiments. M.Z.E., T.Viltrop, O.T., J.M., T.Voet, J.R.V. and A.S. analyzed and interpreted the data. T.Viltrop, O.T., A.Tiitinen, H.M., H.K., V.S.-A., A.-M.S., A.Tiirats, N.K.-A. and S.K. carried out sample collection. O.T. and M.Z.E. performed ddPCR assays. M.K. carried out RNA sequencing analysis. M.N., K.T., O.Z. and R.M. performed PennCNV and QuantiSNP analyses. M.Z.E. drafted the initial version of the manuscript. M.Z.E., T.Viltrop, M.K., A.K., T.Voet, J.R.V. and A.S. wrote and edited the manuscript. M.Z.E., T.Voet, J.R.V. and A.S. jointly supervised this study. All the authors read and approved the manuscript for submission.

Competing interests

M.Z.E., J.R.V. and T.Voet are co-inventors on patent application ZL913096-PCT/EP2014/068315-WO/2015/028576, 'Haplotyping and copy-number typing using polymorphic variant allelic frequencies'.

Additional information

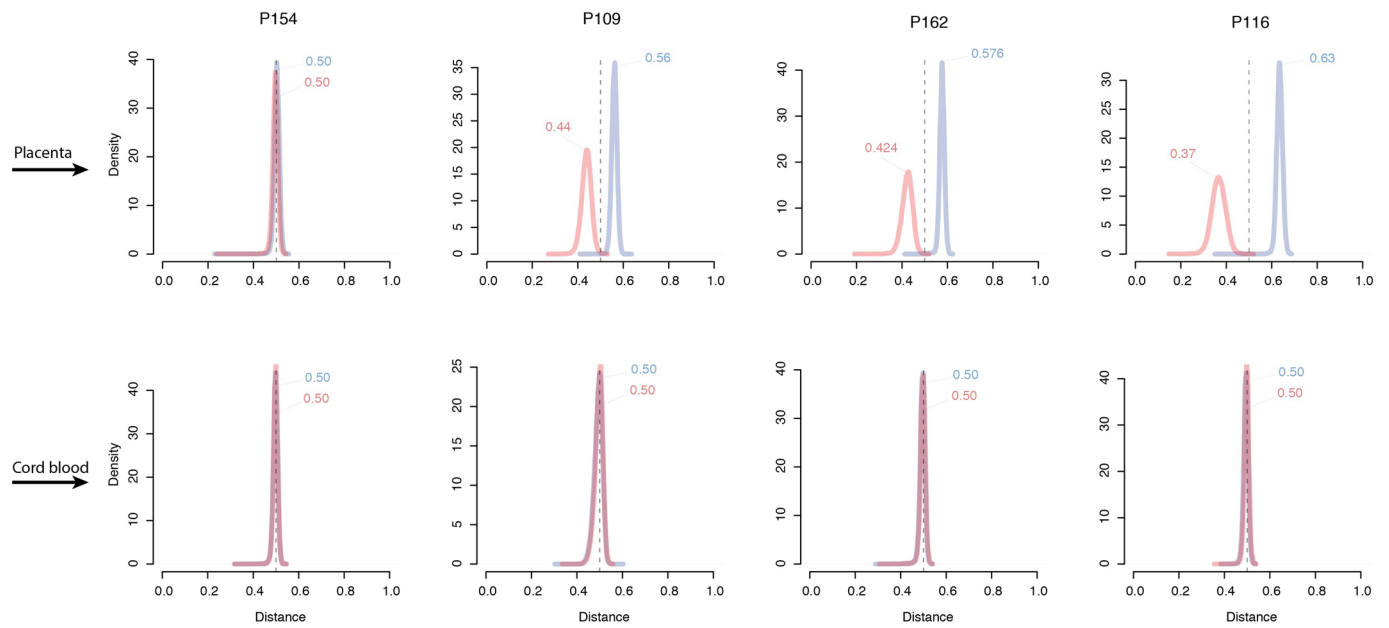
Extended data is available for this paper at <https://doi.org/10.1038/s41591-019-0620-2>.

Supplementary information is available for this paper at <https://doi.org/10.1038/s41591-019-0620-2>.

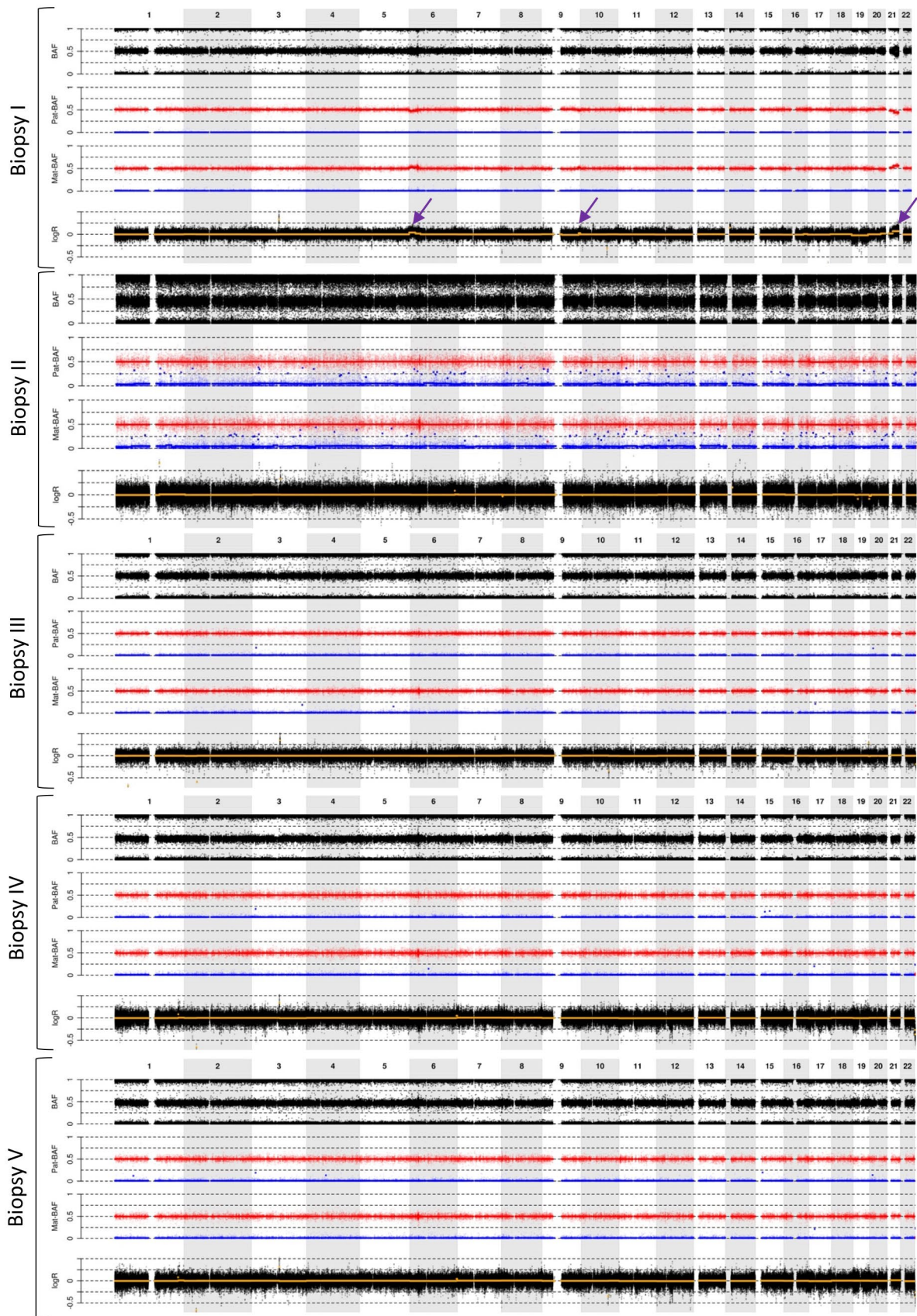
Correspondence and requests for materials should be addressed to M.Z., T.V., J.R.V. or A.S.

Peer review information Brett Benedetti was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

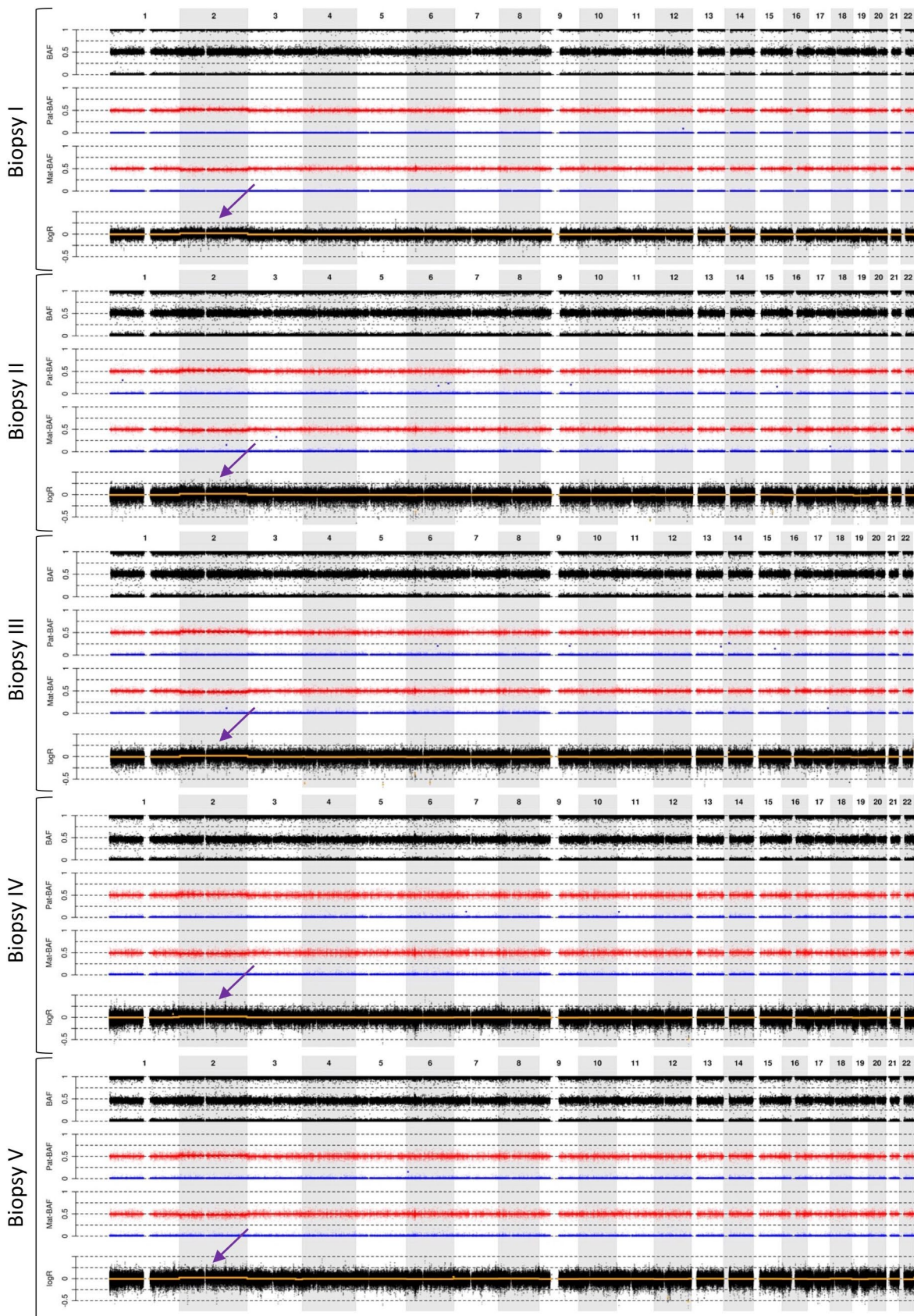
Reprints and permissions information is available at www.nature.com/reprints.



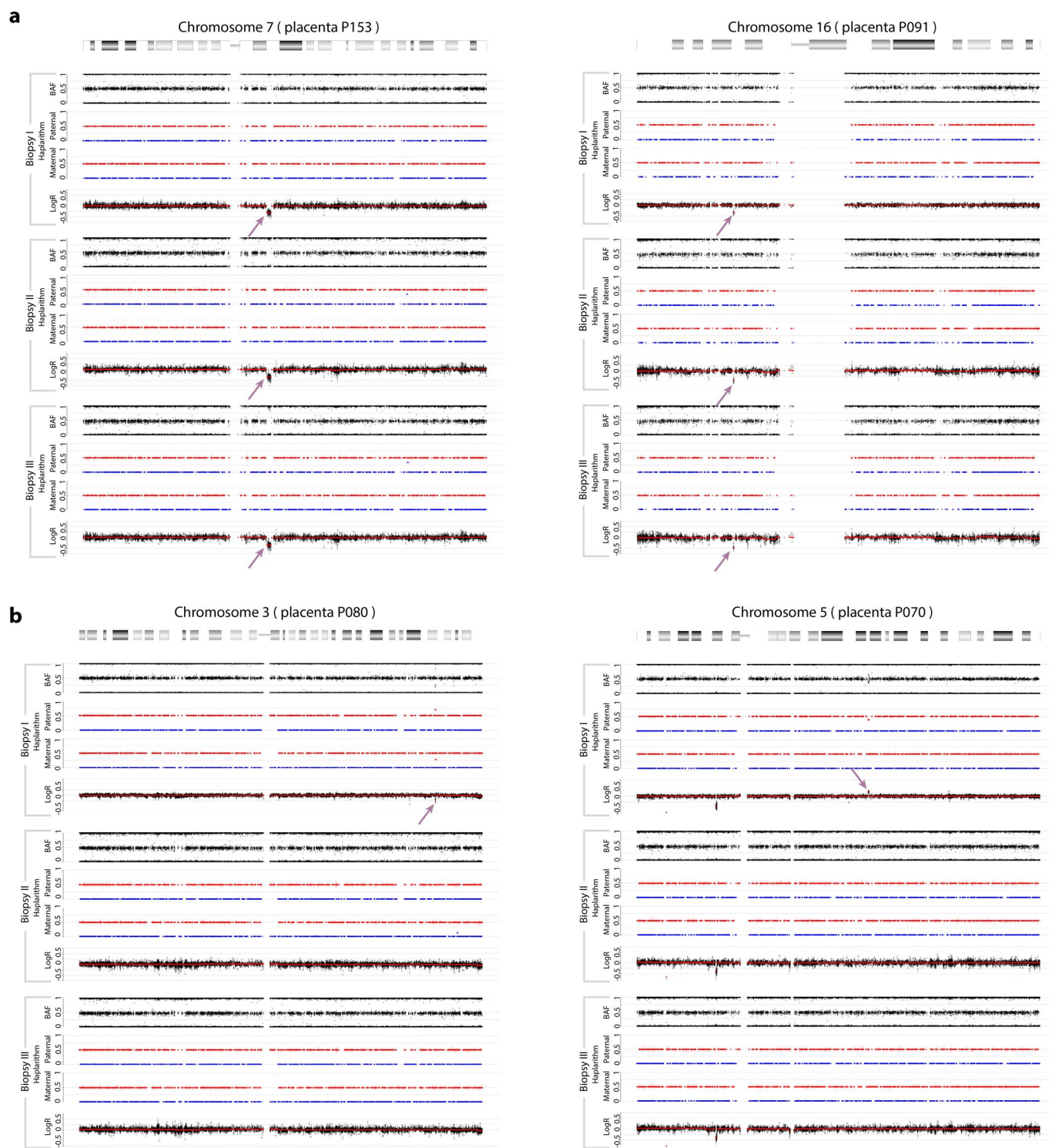
Extended Data Fig. 1 | Haplithmis revealed maternal and fetal contributions to placental DNA samples. We show density of paternal (in blue) and maternal (in red) distances computed from paternal and maternal haplithmis, respectively, of the placenta and cord blood samples of P154, P109, P162 and P116 (see also Fig. 1b and **Source Data**).



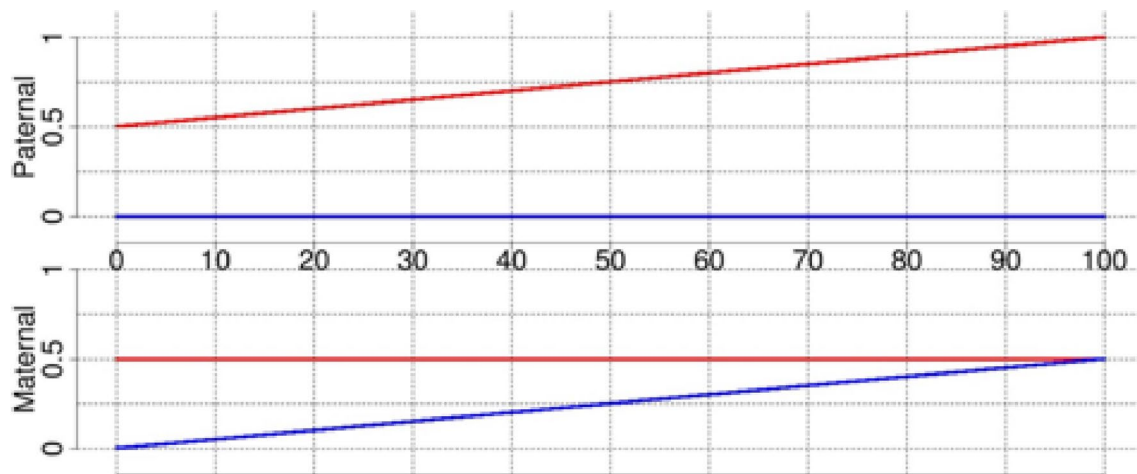
Extended Data Fig. 2 | The mosaic partial trisomies are not preserved across the P172 placenta. The mosaic partial trisomies (purple arrows) on Chr6, Chr9 and Chr21 are only present in one biopsy (Biopsy I) out of all the spatially different biopsies of P172 placenta.



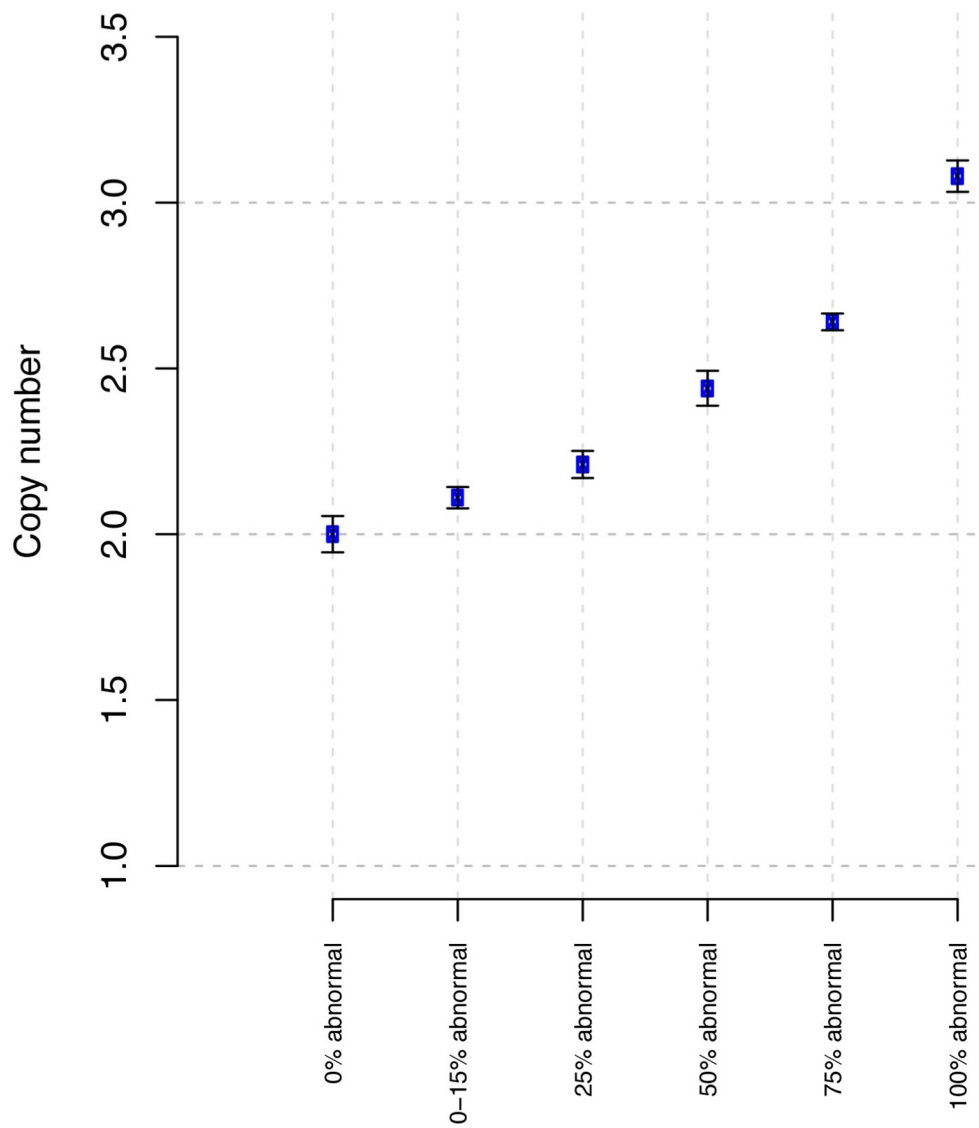
Extended Data Fig. 3 | The full Chr 2 mosaic trisomy is persistently present across the P106 placenta. The full Chr 2 mosaic trisomy (purple arrows) is persistently present in all the spatially different biopsies of P106 placenta.



Extended Data Fig. 4 | Placenta CNV heterogeneity. We analyzed DNA samples from spatially distinct biopsies across the placentas. **a**, the *de novo* non-mosaic CNVs were consistently present in all the biopsies (P153 and P091). **b**, the mosaic CNVs were present in one biopsy (P080 and P070), indicating placental mosaic CNV heterogeneity (see also Extended Data Fig. 2 and **Source Data**).



Extended Data Fig. 5 | Determining fetal and maternal compartments in placenta DNA-samples using haplarithmis. We performed an *in silico* simulation by combining genotypes of the child and the mother with different proportions (from 1%Mother : 99%Child to 99%Mother : 1%Child) and deduced haplarithm profiles for each of these combinations, representing fetal and maternal compartments in placenta DNA samples (see also **Source Data**).



Extended Data Fig. 6 | Proof-of-concept assay for the detection of mosaic aberrations using droplet digital PCR. We mixed up a DNA sample from a trisomy 21 (copy number, CN=3) cell line with a DNA sample derived from a normal diploid cell line (CN=2) at different ratios, creating admixture series of DNA samples with 100%, 75%, 50%, 25%, 10-15% and 0% of abnormal alleles. Mosaic DNA samples were normalized to the number of fully diploid control (i.e. 0% abnormal). Each circle and error bar indicate mean and standard deviation, respectively, of four independent measurements.

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 [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. 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
- An indication of 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 statistics 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
- Clearly defined error bars
State explicitly what error bars represent (e.g. SD, SE, CI)

Our web collection on [statistics for biologists](#) may be useful.

Software and code

Policy information about [availability of computer code](#)

Data collection

We have generated new data in this study. Specifically, following birth, we scrutinized genome-wide single nucleotide polymorphism (SNP) profiles in DNAs from the father, mother, placenta and neonate umbilical cord blood of 111 families (i.e. 111 quartets). For SNP genotyping, the Infinium PsychArray-24 v1.1 (Illumina; GEO: GPL22819) and PsychArray-24 v1.3 BeadChip SNP arrays (Illumina; GEO: GPL26750) were used, which contains approximately 590,000 SNP markers with a median marker spacing of 5.5 kb. No software has been used for data collection.

Data analysis

We developed a web platform that is freely available for research only (see the web platform at: <https://hiva.esat.kuleuven.be/>). Custom code for detecting fetal and maternal compartments is available upon request. Genotype calls, SNP B-allele frequency values and logR values of all samples were computed by the Illumina GenomeStudio 2.0 software. PennCNV and QuantiSNP were used for standard copy-number analysis and comparison with our detected CNVs. Haplarithmisis of siCHILD was used to compute parental haplarithms. Copy number module of siCHILD on the raw logR values was used with minor modifications. Raw logR-values were smoothed by using a moving average window of five consecutive SNP probes, wave-corrected for GC% bias by a Lowess fit and normalized to a trimmed mean of normal diploid chromosomes. Normalized logR-values were segmented by piecewise-constant fitting ($\gamma = 14$). For statistical analyses we applied R version 3.6.0, including packages stats, fmsb (v0.6.3), ICSNP (v1.1-1), ICS (v1.3-1) and TrialSize (v1.3). g:profiler release 2017-10-19 were used for enrichment analyses.

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/reviewers upon request. 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

All SNP array data generated in this study were deposited in the NCBI Gene Expression Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo/>) under accession number GEO: GSE93353.

Field-specific reporting

Please select 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/authors/policies/ReportingSummary-flat.pdf](https://www.nature.com/authors/policies/ReportingSummary-flat.pdf)

Life sciences study design

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

Sample size	Since both predictor (with or without IVF treatment) and outcome (with or without de novo CNVs) are dichotomous, chi-squared test is the most suitable test. To determine appropriate sample size, we performed a power analysis using chi-squared test with confidence level of at least 95% and feeding the conventional effect size (w) values of 0.1, 0.3 and 0.5, which correspond to small, medium and large effect sizes, respectively. With a medium effect size (w=0.3) at least 40 samples per cohort is required to reach to the power of 0.8.
Data exclusions	No data were excluded from the analysis.
Replication	For 10 placenta samples we had 3-5 biological replicates, i.e. from spatially different locations of the placenta (see Table 1). This was to further investigate CNV heterogeneity across the placenta and reaffirm the (non-)mosaic nature of the detected de novo CNVs. For that, we further analysed 20 DNA samples derived from spatially different locations of all but two placentas (F351 and F1021) with de novo CNVs (Table 1). As expected, de novo non-mosaic CNVs found in both placenta and cord blood were consistently present in all the biopsies, whereas de novo mosaic CNVs were not present in the other biopsies (Extended Data Fig. 4).
Randomization	The study participants were not randomized, but they were rather divided to IVF and naturally conceived families.
Blinding	Prior to data analyses all the quartets were anonymized.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Unique biological materials
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Unique biological materials

Policy information about [availability of materials](#)

Obtaining unique materials The majority of unique materials, including DNA samples from father, mother, placenta and umbilical cord (i.e. quartets) used, are readily available from the authors.

Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	In total, 111 couples participated in this study (Supplementary Table 1 and 2), including 62 couples, who delivered neonates after spontaneous conception (natural cohort) and 49 couples who delivered after IVF (IVF cohort).
Recruitment	No prior genetic predisposition was identified in any of the couples participated in this study. In addition, the transferred embryos of IVF newborns were not screened nor counter selected for chromosomal aneuploidy status before intrauterine transfer. The families enrolled in the current study were consecutively recruited.