Chapter 5

Universal mosaicism of the human placenta

5.1 Introduction

In the previous chapters, I have established that somatic mutations serve as a readout of the developmental history of the cells they are present in. I have shown the recapitulation of early embryonic dynamics in normal human development, as well as early clonal expansions as precursors to Wilms tumours. However, all research up until this point was performed on tissues and cells derived from the embryo proper, a logical consequence of confinement to postnatal sampling. Many of the early lineage commitments during embryogenesis will pertain to extraembryonic tissues, derived from the trophectoderm, hypoblast or amnion. In fact, our current explanation for the observed asymmetric contribution in the first cell division is the cell allocation to inner cell mass and trophectoderm. Therefore, the final study presented in this dissertation aims to further elucidate cell dynamics in embryogenesis by sequencing placental biopsies in conjunction with the umbilical cord.

Besides answering questions about early embryogenesis, lineage tracing within the placenta might also reveal its natural course of development. The human placenta is a temporary organ whose dysfunction contributes substantially to the global burden of disease (Brosens et al., 2011). Among its many peculiarities is the occurrence of chromosomal aberrations confined to the placenta, which are absent from the new-born infant (Kalousek and Dill, 1983). Confined placental mosaicism affects one to two percent of pregnancies (Hahnemann and Vejerslev, 1997). This mosaicism may pervade both components of placental villi, the trophectoderm or the inner cell mass-derived mesenchyme, alone or in combination.

The genetic segregation of placental biopsies in confined placental mosaicism suggests that bottlenecks may exist in early development that provide opportunities for genetically separating placental and fetal lineages. It is conceivable that these are physiological genetic bottlenecks underlying the normal somatic development of placental tissue. Alternatively, genetic segregation may represent pathological perturbation of the normal clonal dynamics of early embryonic lineages. For example, it has been suggested that confined placental mosaicism represents a depletion from the fetus-forming inner cell mass of cytogenetically abnormal cells, commonly found in early embryos (Los et al., 1998).

Definition of clinical group	**Number of placentas (bulk)**	**Number of placentas (LCM)**
Birth weight <3rd percentile. Low first trimester PAPP-A.	3	0
Birth weight <3rd percentile. Top decile uterine artery PI (plus bilateral notches if possible).	3	0
Birth weight <3rd. Top decile umbilical artery PI.	3	0
Birth weight <3rd. Lowest decile ACGV.	3	0
Birth weight >97th. Highest decile ACGV. No gestational diabetes.	3	0
Severe preeclampsia and normal birth weight (40-60th percentile), plus sFLT1:PlGF >38 at 36 weeks.	3	0
Any preeclampsia and birth weight <3rd percentile.	3	0
Healthy controls. Birth weight from 40 to 60th percentile, normal PAPP-A and Dopplers, normal ACGV. No preeclampsia.	13	5
Placenta >97th percentile and birth weight range from 39th to 74th percentile	3	0
Total	37	5

Table 5.1 Overview of different clinical groups in the placenta cohort. ACGV=abdominal circumference growth velocity, PAPP-A=Pregnancy Associated Plasma Protein-A, PI=pulsality index, PIGF=Placental growth factor, sFLT1=soluble fms-like tyrosine kinase 1

Lastly, the temporary nature of the placenta might be reflected in the level of mutagenesis that manifests in the organ. The short lifespan of the placenta means mutations have a very limited time to have a phenotypic effect and cause neoplastic lesions of consequence. Concordantly, the rates of placental cancer (choriocarcinoma) are extremely low (Smith et al., 2003). Because of the lack of a selective pressure against a high mutation rate, it is plausible

to find many more somatic mutations in placenta than expected in embryonic tissues at the equivalent age.

Here, I studied the development of the placenta and the split between the inner cell mass and trophectoderm by whole-genome sequencing of bulk placenta samples, as well as laser capture microdissections of distinct components of placental villi. In addition, I reconstructed phylogenies of microdissected clusters of trophoblast and mesenchymal cores (**Fig. 5.1**). All of the tissues had been curated by the Pregnancy Outcome Prediction study, a prospective collection of placental tissue and extensive clinical data, including histological assessment of individual biopsies (Pasupathy et al., 2008; Sovio et al., 2017). I included placentas from normal pregnancies and from pregnancies associated with a range of abnormal parameters, such as preeclampsia or low birth weight (**Table 5.1**).

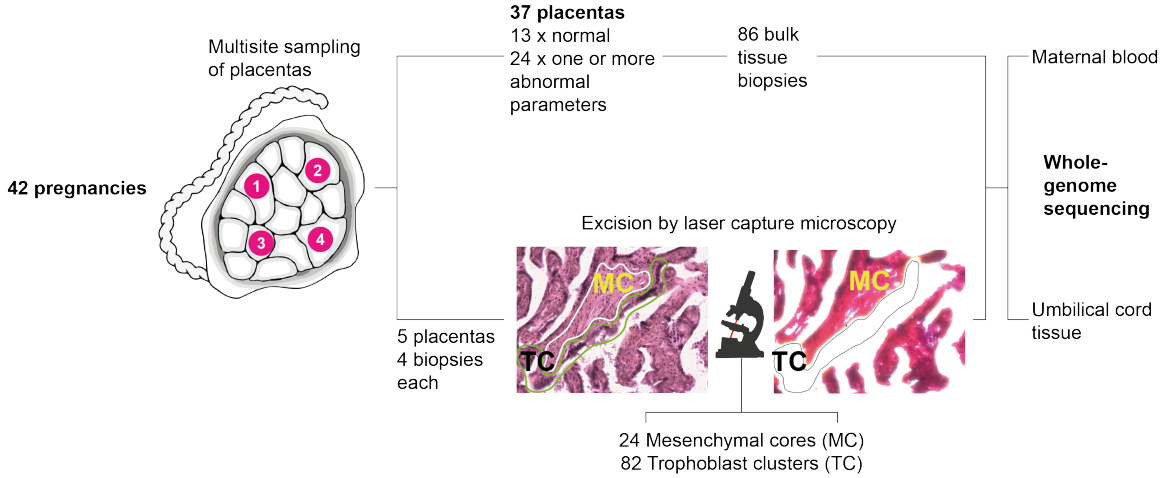

Figure 5.1 Workflow detailing the experimental design with photomicrograph demonstrating microdissection of trophoblast components. TC=trophoblast clusters, MC=mesenchymal core.

5.2 Placental biopsies contain clonal populations

The starting point of this study were whole genome sequences of 86 placental biopsies, obtained from 37 term placentas along with inner cell mass-derived umbilical cord tissue and maternal blood. Placental and umbilical cord biopsies were washed in phosphate-buffered saline to remove maternal blood. Placental biopsy contamination with maternal blood was excluded by screening placental genome sequences for low-level maternal germline variants (see Chapter 2, section 2.5.3).

SNVs called in these placental biopsies revealed a high burden and an unusual degree of clonality (**Fig. 5.2**). On average, each placental bulk sample harboured 145 SNVs, with

the burden ranging from 38 to 259. In addition, the VAF profile of these samples generally revealed the presence of a large clone residing in these tissues. The median VAF within these samples was 0.24 on average, ranging from 0.15 to 0.44. This indicates that generally these large clones account for approximately half of the cells in these biopsies. As demonstrated in previous chapters, solid organ biopsies are consistently polyclonal and only harbour a handful of mutations acquired during the earliest cell divisions of life. Clonal nephrogenesis in Wilms tumour patients represents an exception to this. However, in those cases, only a few mutations signal the presence of the early expansion. This falls short of the high burden and clonality in bulk placenta samples.

Neither the observed median VAF nor the mutation burden differed between normal placentas and those with an abnormal feature, such as growth restriction. This indicates that the dynamics generating these clones are a general feature of physiological placental development.

Analysis of other classes of somatic mutations, indels and copy number changes confirmed the clonal composition of biopsies. Remarkably, 41 out of 86 biopsies harboured at least one copy number change (gain or loss; median size per unique segment, 73.6 kb). However, only one copy number variant, a trisomy of chromosome 10, would have been detectable by clinical karyotyping of chorionic villi. Comparing somatic changes between

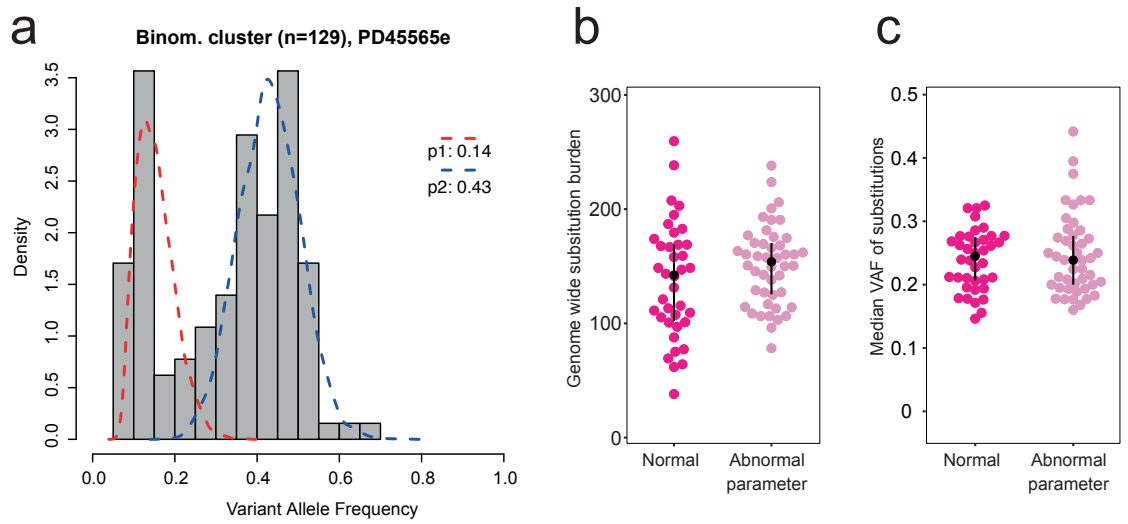

Figure 5.2 (a) VAF distribution within one placental bulk sample (PD45565e). Red and blue dashed lines indicate clonal decomposition through a binomial mixture model, with the estimated VAF per cluster indicated in the legend. (b) SNV burden of each placental biopsy, adjusted for coverage and median VAF. An abnormal pregnancy is defined by the deviation of one or more clinically validated markers from their normal range over the course of pregnancy. (c) Median variant allele frequency of SNVs in each placental biopsy.

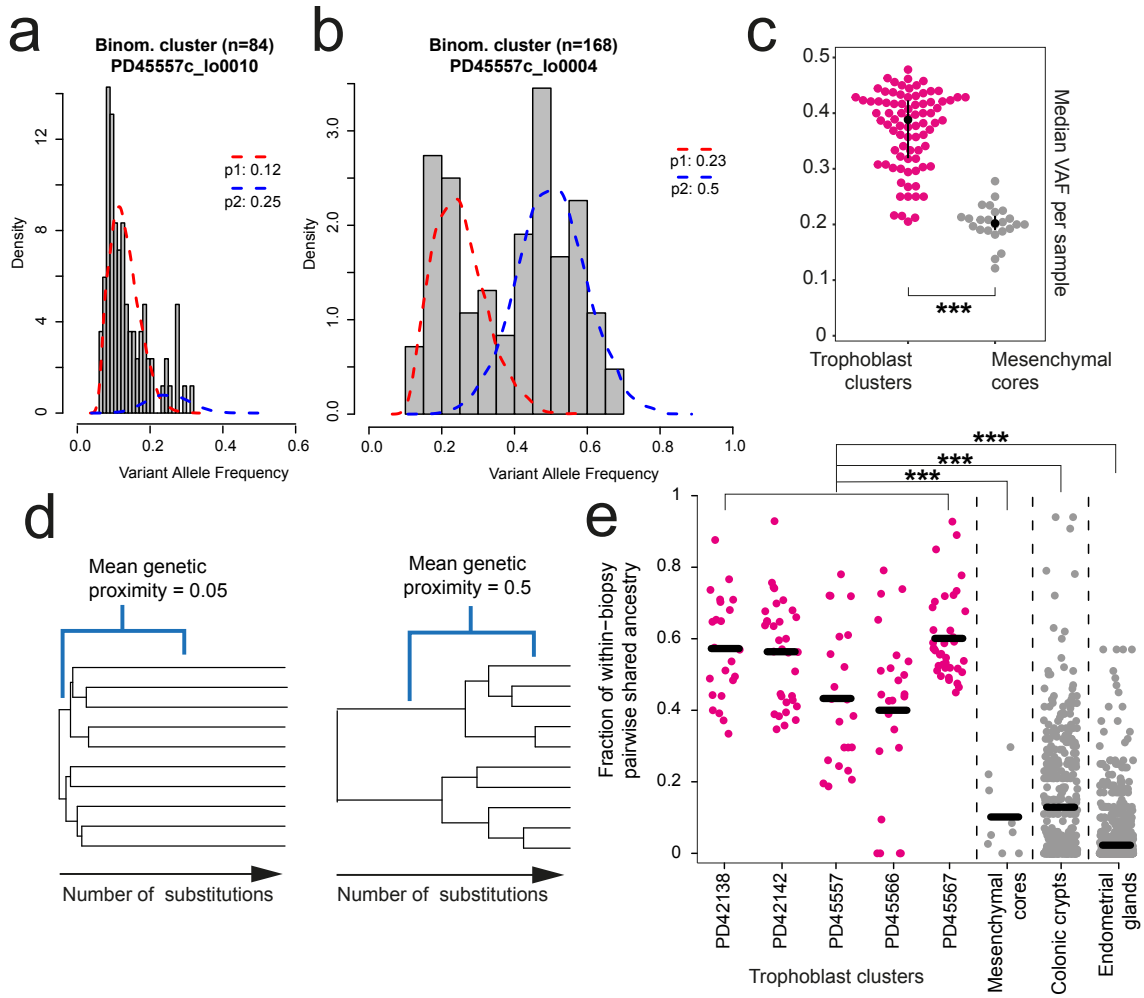

Figure 5.3 Histograms of VAF distribution, overlaid with components as identified by a binomial mixture model for a mesenchymal core sample (a) and a trophoblast cluster (b). 'p' indicates the estimated peak VAF of the components. (c) Comparison of the median substitution VAF between microdissected trophoblast and mesenchymal cores. (*** indicates p<0.001, Wilcoxon rank-sum test) (d) Genetic proximity scores were calculated as the fraction of shared mutations of a pair of samples divided by their mean total mutation burden. For example, a mean score of 0.05 conveys little sharing, while 0.5 signifies a longer shared development. (e). Genetic proximities across trophoblast clusters and mesenchymal cores from the same placental biopsies and data from colonic crypts (Lee-Six et al., 2019) and endometrial glands (Moore et al., 2020). Each dot represents the comparison of two of the same histological unit (e.g., two trophoblast clusters) from the same biopsy. To avoid including adult clonal expansions, bifurcations in phylogenies after 100 post-zygotic mutations were not considered for colon and endometrium. (*** indicates p<0.001, Wilcoxon-rank sum test)

multiple biopsies from the same placenta showed that the majority were unique to the given sample. This suggests that each biopsy represents a genetically independent unit. Of note, placental biopsies had been obtained from separate quadrants of the placenta, several centimetres apart, and therefore they represent distinct lobules in the organ. Therefore, these observations indicate that placental biopsies inherently possessed confined, mosaic genetic alterations.

5.3 Trophoblast clusters are closely related clonal units

The large clones residing in placental tissues could be due to one of the two main components in chorionic villi, namely the trophoblast or the inner cell mass-derived mesenchymal cores. To investigate the cellular origin of these clones in placental biopsies, 82 trophoblast clusters and 24 mesenchymal cores were excised using LCM from the term placentas of five normal pregnancies. These were then subjected to low-input library preparation and whole-genome sequencing.

Calling SNVs revealed that mesenchymal cores generally exhibited a polyclonal VAF distribution (**Fig. 5.3a**). In contrast, clusters of trophoblast exhibited more elevated VAF distributions, with a distinct monoclonal architecture (**Fig. 5.3b**). Concordantly, the median VAF observed per LCM cut was significantly higher in trophoblast clusters compared to mesenchymal cores ($p<0.001$; Wilcoxon rank-sum test) (**Fig. 5.3c**). Hence, the mosaic clonal architecture observed in bulk sample most likely emanates from the trophoblast. In other words, the SNVs identified in placental biopsies are largely accumulated by the trophectodermal lineage, rather than the inner cell mass lineage.

This conclusion is further reinforced by studying the genetic relationship between trophoblast derivatives and mesenchymal cores from the same biopsies. From the reconstructed phylogenetic trees a pairwise genetic proximity scores of microdissections of the two components was calculated. This score was defined as the fraction of shared mutations out of the total mutation burden of the pair (**Fig. 5.3d**; see Chapter 2, section 2.5.3). A low genetic proximity score for pairs of trophoblast clusters or of mesenchymal cores from the same biopsy would indicate that the pool of precursor cells forming these diverged early in development. Conversely, a high score would suggest that histological units within each patch of tissue arose from only a few precursor cells with a relatively long shared ancestry. This analysis revealed a significant difference in the developmental clonal composition between trophoblast clusters and mesenchymal cores ($p < 0.001$; Wilcoxon rank-sum test) (**Fig. 5.3e**). On average, within each biopsy, pairs of trophoblast clusters shared 53% of somatic mutations, indicating a long, joint developmental path of these cells. In contrast,

pairs of mesenchymal cores from the same biopsy exhibited a mean genetic proximity of 10% and thus a short, shared phylogeny, in line with other inner cell mass-derived tissues, such as colon and endometrium (**Fig. 5.3e**). These observations suggest that large expansions of single trophoblastic progenitors underpin the normal clonality and observed confined mosaicism of placental biopsies.

The monoclonal organisation of trophoblast clusters provided the opportunity to examine mutational processes that forged placental tissue in detail. Examining the burden of SNVs of individual trophoblast clusters further, I found an average of 192 variants per cluster.

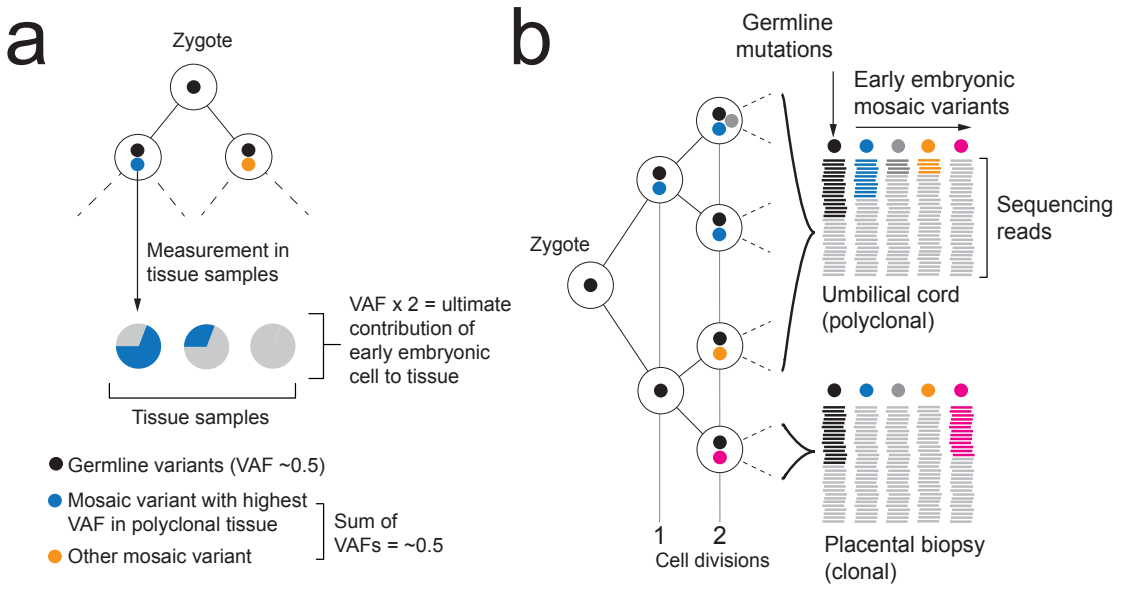

Figure 5.4 (a) Schematic depicting the detection of the earliest post-zygotic mutations and the estimation of contribution to samples from their variant allele frequencies. (b) Hypothetical lineage tree of the early embryo showing how measurements of VAF may relate to cell divisions.

5.4 Biases in cell allocation to trophectoderm and inner cell mass

So far, I have established that the placenta encompasses large clonal populations of trophectodermal cells. It is most likely that these clones are seeded by single cells very early in development. After the initial seeding there appears to be little to no migration of trophectodermal cells along in the developing placenta. Thus, post-zygotic acquisitions of an aneuploidy in this lineage can result in clinically detectable, trophoblastic confined placental

mosaicism. In addition to the committed lineage of trophoblast progenitors, it is possible to investigate the patterns of early, pre-blastulation cell differentiation decisions by identifying mutations shared with the umbilical cord, a tissue entirely derived from the inner cell mass (**Fig. 5.4**).

The most straightforward assessment of this early split is to reconstruct phylogenies of trophoblast clusters. By interrogating the presence of variants in the phylogeny in the umbilical cord genomes I found three different configurations. (1) In PD45566 and PD45567, the trophoblast clusters and the umbilical cord share a most recent common ancestor, the root of the tree, which is most likely the zygote (**Fig. 5.5a**). The umbilical cord exhibits an early asymmetry in line with the expectation, of approximately 2:1 (Behjati et al., 2014; Ju et al., 2017; Lee-Six et al., 2018). (2) In PD45557, one lineage of trophoblast clusters does not share any SNVs with the umbilical cord (**Fig. 5.5b**). However, not all cell divisions are resolved as the tree starts with a trifurcation, obscuring definitive conclusions about the most recent common ancestor of either lineage. The umbilical cord of PD45557 still exhibits the expected asymmetry of 2:1. (3) In PD42138 and PD42142, there was a complete separation between trophectodermal samples and the umbilical cord, without any shared SNVs between them (**Fig. 5.5c**). For PD42138, this indicates that the most recent common ancestor for all trophoblast clusters and for the umbilical cord are two distinct cells later than the zygote, presumably the first generation of daughter cells. The phylogeny of PD42142 starts with a trifurcation, but nevertheless, the most recent common ancestor for the umbilical cord must be a cell later than the zygote.

These three patterns are recapitulated in the bulk genomes as well (**Fig. 5.5d**). In about half of pregnancies (17/37), the earliest post-zygotic mutation exhibited an asymmetric VAF across inner cell mass and trophectoderm lineages, without genetically segregated placental samples in this configuration. In about a quarter of pregnancies (11/37), I found that one placental biopsy did not harbour the early embryonic mutations shared between umbilical cord and other placental biopsies. In other words, that trophectodermal lineage shared no post-zygotic genetic ancestry with the inner cell mass. In the remaining quarter (9/37) of pregnancies, the early cell allocation generated a complete separation of all placental tissues from umbilical cord samples (**Fig. 5.5e**). Taken together, this data suggests that in about half of placentas, there is evidence for a trophectodermal lineage sharing no mutations with the inner cell mass. Consequently, genomic alterations that pre-exist in the zygote, or arise within the first few cell divisions, may segregate between the placenta and fetal lineages by virtue of the natural separation between these lineages.

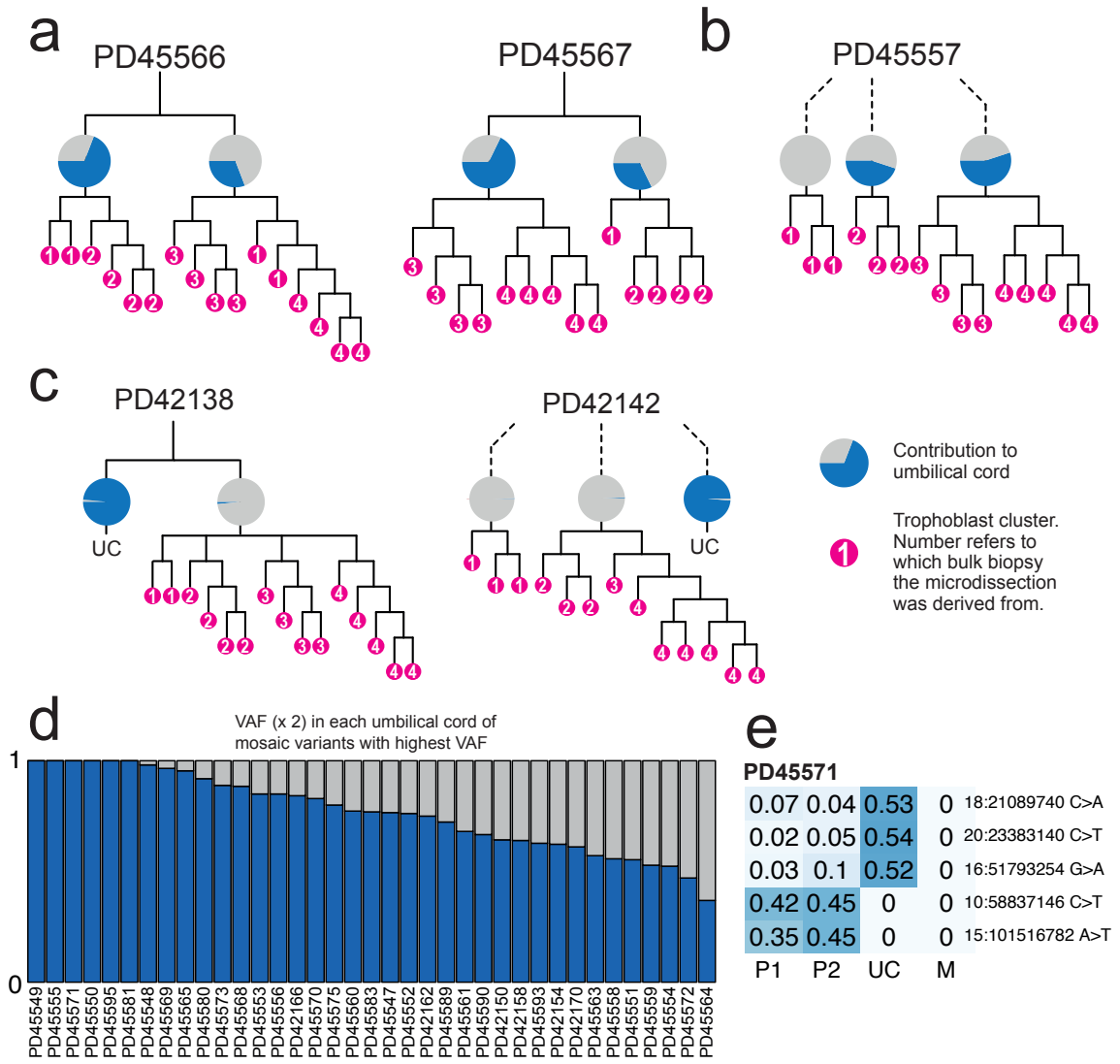

Figure 5.5 (a) Early trees of trophoblast clusters of PD45566 and PD45567, with the contribution of lineages to the umbilical cord colored in blue in pie charts. The umbilical cord exhibits an asymmetric contribution of the daughter cells of the zygote. (b) Early cellular contribution in PD45557 shows separation of one placental lineage. (c) In PD42138 and PD42142 the placental and umbilical cord lineages do not share any early embryonic mutations. (d) The contribution of the major lineage to the umbilical cord as calculated from the embryonic mutation with the highest VAF. (e) Heatmap of early embryonic variants in PD45571 showing a complete split between placenta (P1 and P2) and umbilical cord (UC). SNVs are absent from the maternal blood (M).

5.5 A reversal of trisomy 10

This potential for a full segregation between trophectodermal and inner cell mass-derived lineages is most strikingly exemplified by PD45581. One of the two placental biopsies

sequenced for this patient harbours a trisomy of chromosome 10 (PD45581c). This trisomy is absent in the other placental biopsy (PD45581e), as well as the umbilical cord (PD45581e), both of which displayed a regular disomic pattern for chromosome 10 (**Fig. 5.6a**). From interrogating common SNP sites on this chromosome across all samples from this patient and the whole-genome data of maternal blood, it became apparent that the trisomy consisted of two maternal copies and one paternal one. However, the disomic samples did not carry one paternal and one maternal copy as expected, but rather two non-identical maternal copies. This was evident from SNPs present in the trisomy but absent from the disomy, as well as SNPs homozygous in the disomy, but less than homozygous in the trisomy (**Fig. 5.6b**). If the disomy had evolved into a trisomy by a conventional gain of a whole chromosome, neither of these patterns would have manifested.

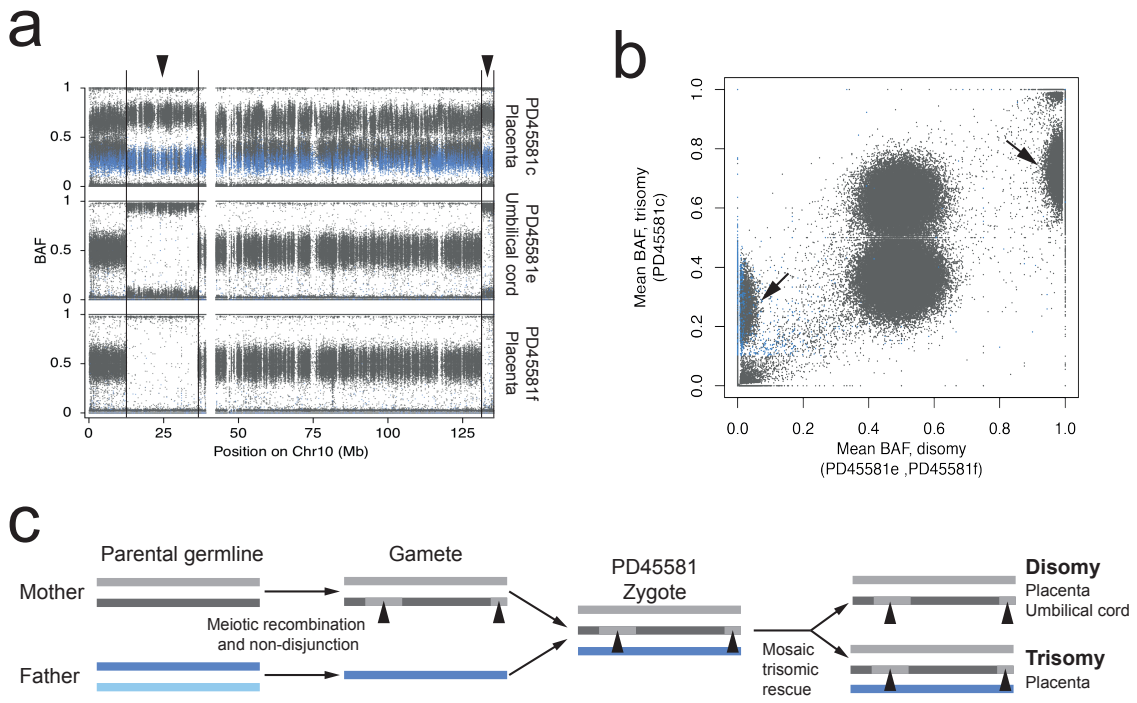

Figure 5.6 (a) B-allele frequency (BAF) of germline SNPs on chromosome 10 in PD45581, showing a trisomy in PD45581c. SNPs absent from mother are colored in blue. (b) Mean BAF of samples with trisomy 10 (PD44581c) versus the mean BAF in disomic samples (PD45581e, PD45581f). SNPs absent from mother are colored in blue. The clusters of variants absent or homozygously present in disomic samples, but distributed around 0.33 and 0.66 (arrows) in the trisomy refute the possibility of a somatic duplication. (c) Overview of genomic events in PD45581 and parents leading to the observed mosaic trisomic rescue. The arrowheads highlight areas of two genotypes in PD45581c due to meiotic recombination in the mother.

Instead, all evidence points to the zygote starting out with a trisomy consisting of one paternal and two maternal copies (**Fig. 5.6c**). Subsequently, the paternal copy was mosaically lost, resulting in a subset of cells with a normal disomic profile. Hence, this is a direct observation of a trisomic rescue. The origin of the trisomy as maternal meiotic non-disjunction is further reinforced by two large regions of homologous recombination, which appear fully homozygous in disomic samples (**Fig. 5.6a**).

A full trisomy 10 is lethal during embryogenesis (Hassold et al., 1996). It stands to reason that the embryo was only able to develop because a subset of cells in the inner cell mass-derived lineages had reverted back to disomy. The low number of embryonic mutations that were found at a clonal level in the umbilical cord hint that the initial trisomic rescue event must have happened within the first few cell divisions of life. This is consistent with the initial event occurring during the cleavage stage of embryogenesis. The reportedly high level of genomic instability due to rapid divisions at this time might aid aneuploid embryonic cells to self-correct.

The trophectodermal line, however, does not seem to harbour the same strong selective pressure against aneuploidies as the embryo proper. Because of this, cells containing the trisomy of chromosome 10 are still abundant in the placenta. It is plausible that many clinical cases of confined placental mosaicism represent trisomic rescues.

5.6 Processes and impact of placental mutagenesis

By deconvoluting the mutational signatures that contribute to the profile of SNVs seen in bulk biopsies and trophoblast microdissections, it is possible to determine the mutagenic processes operating in the human placenta. Three different reference signatures contributed to the mutagenesis in placenta: signatures 1, 5, and 18 (Alexandrov et al., 2020). Signature 1 consists of C>T mutations at CpG sites, and corresponds to spontaneous deamination of methylated cytosines. Signature 5 is a rather flat signature, without prominent features. Both signatures 1 and 5 are ubiquitous in human tissues and accumulate throughout life (Alexandrov et al., 2015). In contrast, signature 18 has been identified infrequently in normal tissues. It is characterised by C>A variants and has been associated with reactive oxygen species and oxidative stress (Poetsch, 2020). In placental biopsies and microdissected trophoblast clusters, signature 18 contributed 43% of all SNVs (**Fig. 5.7a,b**). In comparison, in normal human colorectal crypts, the normal tissue with the highest prevalence of signature 18 mutations described to date, it contributed an average of 13% of substitutions (Lee-Six et al., 2019). Within the confines of limited sampling size of this cohort, I did not observe a

significant difference in signature composition between normal placentas and those with an abnormal parameter.

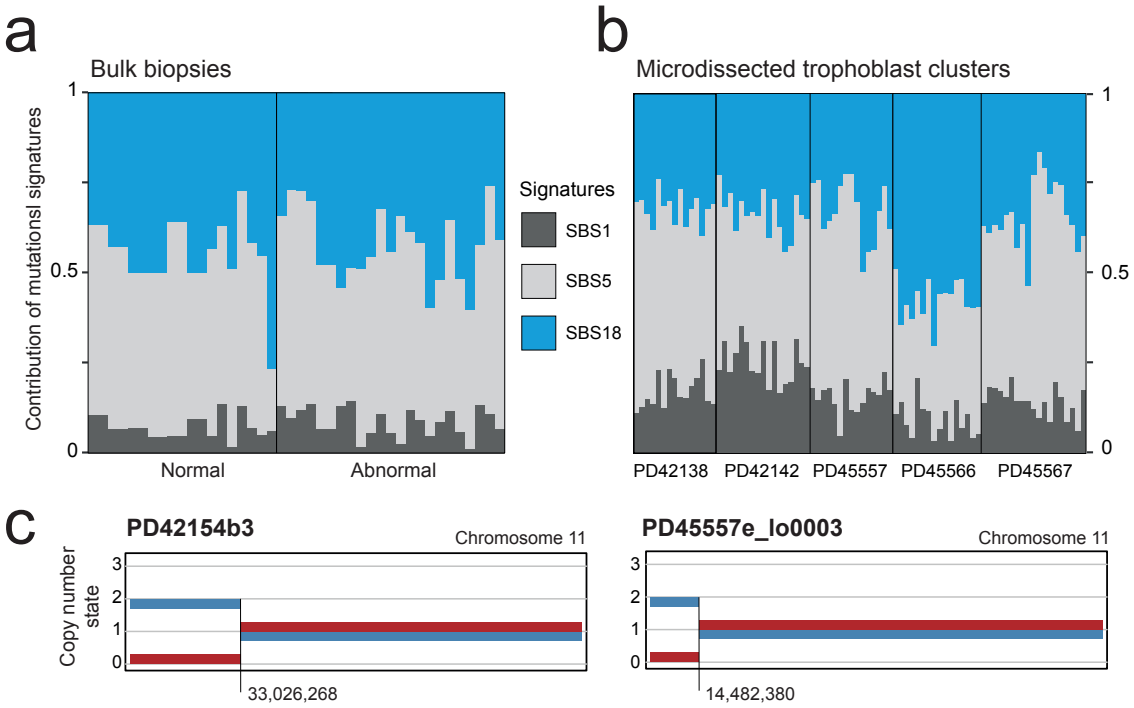

Figure 5.7 Proportion of mutational signatures 1, 5 and 18 attributed to placental bulk samples (a) and microdissected trophoblast clusters (b). (c) Amplification of the paternal allele through copy number neutral loss of heterozygosity of 11p in one bulk placental sample (PD42154b3) and one trophoblast cut (PD45557e_lo0003).

It is conceivable that the high exposure to signature 18 in the human placenta is entirely due to its unique function as the supplier of oxygen *in utero* and the rapid cell division required for full development of the organ. An alternative explanation stems from the temporary nature of the placenta. Somatic mutations arising in the trophectodermal lineage have a very limited timeframe to have a negative impact on the fitness of the offspring, as the very low incidence of malignant neoplasia illustrates. It stands to reason that the placenta tolerates most somatic mutations, in line with its tolerance of aneuploidies. Therefore, it might have a lower activity of the DNA repair machinery than tissues derived from the embryo proper, where mutations can impact phenotype for an entire lifetime. In line with this, the mutation burden attributable to signatures 1 and 5 in trophoblasts is at least 109 SNVs[1] for the duration of a pregnancy, corresponding to a mutation rate of approximately 146 a year. This is several times higher than any mutation rate reported for an adult normal

[1]The mean burden of trophoblast clusters is 192, the mean attribution to signatures 1 and 5 is 57%, so the burden due to signatures 1 and 5 is their product.

tissue, the highest of which is colonic epithelium with a rate of 43.6 SNVs a year (Lee-Six et al., 2019).

Annotating functional consequences of all somatic variants found in bulk biopsies and trophoblast samples, indicated that most changes were unlikely to have any impact on phenotype. The majority (42 of 81 unique variants) of copy number changes lay within fragile sites. Interestingly, two placentas out of 42 harboured copy number neutral loss of heterozygosity of chromosome 11p15. In both cases, this constituted uniparental disomy of the paternal allele (**Fig. 5.7c**). The 11p15 region contains the imprinted loci, among which the *H19* locus discussed at length in the previous chapter. Paternal amplification of this region underpins the overgrowth syndrome, Beckwith-Wiedemann, while maternal amplification would result in Silver-Russell syndrome. Both have been implicated in placental disease, mainly in pregnancy-induced hypertension and intrauterine growth restriction, respectively (Angiolini et al., 2011; Bourque et al., 2010; Fowden et al., 2006). Since these imprinted regions can be affected by a loss of imprinting rather than a loss of heterozygosity, it is likely that a proportion of placental cells might harbour an abnormal methylation landscape of the 11p15 imprinted sites.

5.7 Conclusion

In this exploration of the somatic genomes of human placentas, I identified genetic bottlenecks at different developmental stages that confined placental tissues genetically. Most prominently, every placental biopsy that I examined represented an independent clonal trophoblast unit, suggesting that mosaicism represents the inherent trophectodermal clonal architecture of human placentas. During the first few divisions of life, the prospective split between trophectoderm and inner cells segregates placental tissues from the embryo proper, genetically isolating trophoblast lineages. Together, these bottlenecks may represent developmental pathways through which cytogenetically abnormal cells phylogenetically and spatially separate. This renders them detectable by genomic assays utilized in the clinical assessment of chorionic villi. The findings thus provide plausible, physiological developmental routes through which confined placental trophoblast mosaicism may arise. I suspect that as our understanding of the clonal dynamics of human embryonic lineages grows, it is possible to find additional bottlenecks that account for placental mosaicism affecting mesenchymal lineages also.

The landscape of somatic mutations in placental biopsies and at the level of individual trophoblast units was unusual compared to other normal human tissues studied to date. In particular, placental tissue was an outlier with regards to the abundance of signature

18 mutations. It is possible that these somatic genetic peculiarities represent the specific challenges that trophoblast lineages undergo during placental growth, such as the approximate threefold rise in changes in the local oxygen tension of blood surrounding the villi between eight and twelve week's gestation (Jauniaux et al., 2000). Furthermore, it may be conceivable that as a temporary, ultimately redundant organ, some of the mechanisms protecting the somatic genome elsewhere do not operate in placental trophoblasts.

It is possible that genomic placental alterations contribute to the pathogenesis of placental dysfunction, which is a key determinant of the "Great Obstetrical Syndromes", such as preeclampsia, fetal growth restriction and stillbirth. Previous studies associating confined placental mosaicism with these syndromes have yielded conflicting results (Amor et al., 2006; Baffero et al., 2012; Grati et al., 2020; Jauniaux et al., 2000; Kalousek et al., 1991; Toutain et al., 2018). This study may explain these discrepancies, as the genomic alterations I observed were not uniformly distributed across multiple biopsies from the same placenta. Larger scale systematic studies of the genomic architecture of the human placenta in health and disease might establish the role of placental genomic aberrations in driving placenta-related complications of human pregnancy.