Mitochondrial DNA variants segregate during human preimplantation development into genetically different cell lineages that are maintained postnatally

Joke Mertens 1, Marius Regin 1, Neelke De Munck 2, Edouard Couvreu de Deckersberg 1, Florence Belva 3, Karen Sermon 1, Herman Tournaye 2,4,5, Christophe Blockeel 3, Hilde Van de Velde 1,2,6 and Claudia Spits 1,*

1Research Group Reproduction and Genetics, Faculty of Medicine and Pharmacy, Vrije Universiteit Brussel (VUB), 1090 Brussels, Belgium
2Brussels IVF, Center for Reproductive Medicine, Universitair Ziekenhuis Brussel, 1090 Brussels, Belgium
3Center for Medical Genetics, Universitair Ziekenhuis Brussel, 1090 Brussels, Belgium
4Research Group Biology of the Testis, Faculty of Medicine and Pharmacy, Vrije Universiteit Brussel (VUB), 1090 Brussels, Belgium
5Department of Obstetrics, Gynaecology, Perinatology and Reproduction, Institute of Professional Education, Sechenov First Moscow State Medical University of the Ministry of Health of the Russian Federation, Moscow 119992, Russia
6Research Group Reproduction and Immunology, Faculty of Medicine and Pharmacy, Vrije Universiteit Brussel (VUB), 1090 Brussels, Belgium

*To whom correspondence should be addressed at: Laarbeeklaan 103, Building K, first floor, 1090 Brussels, Belgium. Tel. +32 (0)24774638; Fax: not applicable; Email: claudia.spits@vub.be

Abstract

Humans present remarkable diversity in their mitochondrial DNA (mtDNA) in terms of variants across individuals as well as across tissues and even cells within one person. We have investigated the timing of the first appearance of this variant-driven mosaicism. For this, we deep-sequenced the mtDNA of 254 oocytes from 85 donors, 158 single blastomeres of 25 day-3 embryos, 17 inner cell mass and trophectoderm samples of 7 day-5 blastocysts, 142 bulk DNA and 68 single cells of different adult tissues. We found that day-3 embryos present blastomeres that carry variants only detected in that cell, showing that mtDNA mosaicism arises very early in human development. We classified the mtDNA variants based on their recurrence or uniqueness across different samples. Recurring variants had higher heteroplasmic loads and more frequently resulted in synonymous changes or were located in non-coding regions than variants unique to one oocyte or single embryonic cell. These differences were maintained through development, suggesting that the mtDNA mosaicism arising in the embryo is maintained into adulthood. We observed a decline in potentially pathogenic variants between day 3 and day 5 of development, suggesting early selection. We propose a model in which closely clustered mitochondria carrying specific mtDNA variants in the ooplasm are asymmetrically distributed throughout the cell divisions of the preimplantation embryo, resulting in the earliest form of mtDNA mosaicism in human development.

Introduction

Most studies in human genetics have been performed on bulk DNA, extracted from peripheral blood or other tissues. In recent years, it has become obvious that the cells of our body are not as genetically homogeneous as previously thought. Next to the well-known classical cases of cellular mosaicism such as the variation in the somatic rearrangements of immunoglobulin and T-cell receptor in lymphocytes, a whole new dimension of diversity is being uncovered due to the emergence of single-cell comprehensive genome analysis (1). The mutations driving somatic mosaicism probably occur at all stages of development, from early preimplantation development, as seen for chromosomal abnormalities (2), to the ageing individual (1).

The mitochondrial DNA (mtDNA) is particularly diverse. It has been known for long that a given inherited mtDNA mutation can be present at different loads in different individuals of the same family and in different tissues of the same individual (3), while somatic cellular heterogeneity has been shown in blood cells (4,5), neurons, glia (6) and single muscle fibers (7). In recent years, the advent of massive parallel sequencing has had a profound impact on the field of mitochondrial genetics. Deep sequencing of the relatively small mtDNA has facilitated the simultaneous detection of all variants within this genome and their individual heteroplasmic loads (8,9). This type of work has shown, for instance, that pathogenic mtDNA variants are commonly present in healthy individuals, with a mean load of 2%, and cross-generation studies have shown that these variants are heritable (8–10).

The composition and heteroplasmic load of the variants in the mtDNA of an individual can vary during development, as it goes through various bottlenecks. The first bottleneck occurs during oogenesis, where...
the few mtDNA copies in the primordial germ cells replicate rapidly, to multiply a 1000-fold when reaching the mature oocyte stage. During this process, low-load heteroplasmic variants in early primordial germ cells can increase to much higher loads in the late primordial germ cells, where selection mechanisms will eliminate mitochondria with variants affecting their functionality (11). This selection process circumvents Muller’s ratchet, an evolutionary process where deleterious variants can accumulate rapidly over generations in an irreversible manner due to uniparental inheritance and lack of genomic recombination of the mtDNA (12). This first bottleneck is also responsible for the diversity in heteroplasmic loads of the same variant across children of the same mother. The second bottleneck occurs after fertilization, where the mtDNA copy number per cell declines transiently during early embryonic development due to halted replication, which is resumed when the embryonic cells initiate differentiation (13,14). Finally, later in development, the mtDNA becomes more susceptible to somatic mutagenesis due to errors of the polymerase gamma, its proximity to reactive oxygen species and a very low protection against mutagenesis by repair mechanisms and histones (15–17), leading to ageing-related somatic mosaicism (reviewed by van den Ameele et al. 18).

Several studies have demonstrated mtDNA mosaicism for both disease and non-disease-causing variants, both at the single-cell level and across tissues of one individual (19–22). While some of this variation has been attributed to ageing (20), the recurrence of other mosaic variants across multiple cells of the same individual suggests that the variant emerged very early in development, after or during the formation of the three germ layers. Other variants appear to be tissue-specific, leading to the suggestion that mtDNA variant composition can be cell-type specific (21,22). Overall, these mosaic mtDNA variants can be present at a very high load in one cell, while at a very low load or absent in other cells. These differences across cells of the same individual could be explained through clonal expansion by random genetic drift. Because this process is assumed to be relatively slow, the introduction of a mosaic variant should occur early in development, but the exact timing and mechanism remain to be elucidated (23).

Currently, our limited knowledge on the segregation of mtDNA variants during human preimplantation development is mainly based on the study of inherited pathogenic mtDNA mutations, mostly in the context of preimplantation genetic testing. A number of studies have explored the possibility of quantifying pathogenic heteroplasmic variants in polar bodies, single blastomeres biopsied at the cleavage stage or trophoderm biopsies at the blastocyst stage and have investigated if the results of these biopsies are representative of the rest of the embryo. The segregation of mutations causing mitochondrial diseases such as Leigh syndrome (24,25), neuropathy, ataxia and retinitis pigmentosa (NARP) (26), Leber’s disease (27) and MERRF/MELAS (myoclonic epilepsy with ragged red fibers/mitochondrial encephalopathy, lactic acidosis and stroke-like episodes) (28–31) have been reported. With some exceptions (31), the different groups have found a good consistency in heteroplasmic loads between samples of the same embryo, showing that these variants are homogeneously distributed across the mitochondria in the oocyte, and homogeneously partitioned during the early developmental cell divisions. Conversely, there is only limited knowledge on the appearance and segregation of non-disease-causing mtDNA variants during early human development.

In this study, we aimed at determining the timing of appearance of mosaicism for non-disease associated mtDNA variants during human preimplantation development. We studied to which extent individual embryonic cells differed from each other and identified different types of mtDNA variants depending on their recurrence across samples of the same donor. Comparison of the patterns of mtDNA variants in oocytes, day-3 and day-5 embryos, and adult-stage tissues and single cells revealed that mtDNA mosaicism appears as early as day 3 of human preimplantation development, and is maintained through development, resulting in genetically diverse cell lineages in the adult.

**Results**

**Mitochondrial DNA mosaicism occurs as early as day 3 of human preimplantation development**

We deep-sequenced the mtDNA of 254 oocytes from 85 donors, 158 single blastomeres from 25 day-3 embryos obtained from nine donors, 17 samples from 7 day-5 blastocysts (7 inner cell mass (ICM) and 10 trophoderm (TE)), 142 adult DNA samples from bulk tissue (59 buccal, 57 blood and 26 urine samples) and 68 single cells from buccal swab (N = 37) and urine (N = 31) of three donors. For the oocytes, we collected at least two samples per donor (either two or more oocytes of the same donor or oocytes and somatic tissues), and for the adult DNA samples, we also collected at least two tissues per donor.

We identified heteroplasmic variants in 58.7% of the oocytes, 86.1% of blastomeres of day-3 embryos, all of the ICM/TE samples of day-5 blastocysts, 47.9% of the adult bulk tissues and 95.6% of adult single cells (Fig. 1A). The average heteroplasmic variant load for all variants in all samples in the cohort was under 10% for the oocytes and embryos (oocytes: 7.6 ± 13.8%, day-3 blastomeres: 6.1 ± 11.2% and day-5 ICM/TE biopsies: 5.1 ± 5.5%), 20.1 ± 27.6% for the adult bulk tissues and 12.1 ± 13.6% in adult single cells (Fig. 1B).

Further analysis of the variants showed that while some of the variants recurred across samples of the same donor, other variants were unique to one sample. Therefore, in the further downstream analysis of our dataset, we categorized the heteroplasmic variants as ‘recurrent’ when they were present above the thresholds in two or multiple tissues/cells of one individual, two or multiple cells of one day-3 embryo, different biopsies of a...
Figure 1. Overview of the heteroplasmic variants found in the different cohorts. (A) Percentage of samples with heteroplasmic variants in oocytes, day-3 embryos and day-5 blastocysts, adult bulk tissues and adult single cells. (B) The average load and standard deviation of the heteroplasmic variants found in the oocytes, day-3 embryos and day-5 blastocysts, adult bulk tissues and adult single cells. (C–E) Examples of variants and their respective load found in oocytes from the same donor (C), blastomeres from the same day-3 embryo (D) and biopsies from the same day-5 blastocyst (E). Variants that recur across cells or samples are shown in black and in red are variants that are unique to one sample. (F) Heteroplasmic load of recurrent variants found in the bulk DNA sample (red line) and in the single cells from the same tissue (black dots). (G) Variants found in the adult bulk DNA and in 10 single cells of urine and buccal samples that were recurrent across bulk and single cells (black dots for the urine samples and black crosses for the buccal samples) and variants that were unique to one cell (red dots for the urine cells and red crosses for the buccal cells).

In Figure 1D, you find an example on the detected variants in the different cells of one day-3 embryo. This plot shows that the heteroplasmic loads of recurrent variants are very consistent across different cells of the same embryo. It is worth noting that if we found a day-5 blastocyst or in two or multiple oocytes of one donor, and as ‘unique’ when they were present in only one sample and under the detection limit in other samples of the same donor. An example of the variant composition of different sample types is shown in Figure 1C–G.
variant in at least two cells of the same embryo, we would detect it in all other cells, even if the variant was present at a low heteroplasmic load. The full datasets can be found in Supplementary Material, Tables S1–S4. We found recurrent mtDNA variants in 26.8% of oocytes, 76.0% of day-3 embryos, 85.7% of day-5 blastocysts, 38.7% of adult bulk DNA samples and 72.1% of adult single cells. Unique variants were found in 48.0% of oocytes, 88.0% of day-3 embryos, 85.7% of day-5 blastocysts (detection limit for oocytes and embryos: >2%), 16.9% of bulk DNA samples (detection limit: >1.5%) and 86.8% of single cells (detection limit: >5%).

Overall, the data revealed considerable mtDNA variation across oocytes of the same donor, showing that germ line mosaicism for non–disease-associated mtDNA variants is exceedingly common. In the embryos, we found that somatic mtDNA mosaicism occurs already on day 3 of development and is maintained in day-5 blastocysts. This first type of mosaicism is due to the appearance of unique variants in individual cells of the embryo. At this stage of development, variants that recur across cells of the embryo do so consistently at very similar loads. These recurrent variants are the source of a second type of somatic mosaicism that emerges later in development and that is evidenced by the results of the adult tissues and single cells. Here, we find that the heteroplasmic load of variants measured in a DNA sample extracted from a given tissue represent the average of the widely variable loads found in the individual cells in that tissue, ranging from a homoplasmic state to absent in some cells (Fig. 1F). Finally, a third type of mosaicism is present in the adult single cells, in the form of numerous unique variants that have most likely originated by somatic mutagenesis related to ageing. Interestingly, we have identified four such variants that appear in a tissue-specific manner in different donors: m.215A > G and m.152 T > C were found in buccal samples of seven donors, m.16211 T > C in blood and buccal samples of two donors and m.72 T > C in urine samples of two donors.

Sibling oocytes carry both recurrent and unique variants that differ in location, type and heteroplasmic load

In 26.8% of the oocytes, we found variants that recurred across two or more of the oocytes of the same donor. These variants were evenly distributed between the non-coding and protein-coding regions, and none were detected in the ribosomal RNA/transfer RNA (rRNA/trRNA) coding sequences. A non-synonymous change was induced by 16.7% of the variants (Fig. 2A). We calculated the mutation rate per base as the number of variants found in each mtDNA locus, divided by the number of sequenced base pairs of that locus. In the oocytes, the highest mutation rate per base was in the hypervariable region (Fig. 2B). Of recurrent variants in the oocyte, 61.1% had heteroplasmic loads over 20% (Fig. 2C), while their load was independent from their location or type of change (Fig. 2D).

We sequenced somatic tissues (buccal swabs, blood and urine) from a subgroup of 25 oocyte donors. In these, nine variants were present in at least one somatic tissue and were transmitted to the oocytes (referred to as ‘transmitted’) and three variants were only present in the oocytes and not in the somatic tissues (referred to as ‘germline-specific’). Both types of variants were similarly distributed across non-coding and protein-coding regions with a distribution of 55.6 versus 44.4% for the transmitted variants, respectively, and 66.7 versus 33.3% for the germline-specific variants, respectively. However, the transmitted protein-coding variants were exclusively synonymous while the germline-specific protein-coding variants were all non-synonymous (Fig. 2E). Of the transmitted variants, 77.8% were present at loads >20%, while this was only 33.3% for the germline-specific variants, although this difference was not statistically significant (Fisher’s exact test, \( P = 0.2364 \), Fig. 2F). Lastly, the higher the load in the somatic tissues, the higher the likelihood that the variant would be present in the majority of the oocytes of the donor (correlation \( R = 0.65 \), \( P < 0.0001 \)), with variants with loads as low as 3% in the somatic tissues also being identified in the oocytes. Conversely, the heteroplasmic load of the variant could significantly differ from oocyte to oocyte, in extreme cases going from undetectable levels in one oocyte to homoplasm in another (Fig. 2G).

Nearly half (48.0%) of oocytes carried variants that were unique to one oocyte in a cohort, and these were remarkably different in their location and heteroplasmic load from the recurrent variants. Of the unique variants, 16.3% were located in the rRNA/trRNA regions (versus 0% of the recurrent variants, Fisher’s exact test, \( P = 0.0814 \), Fig. 2A) and 21.1% of the unique protein-coding variants induced a non-synonymous change (versus 16.7% of the recurrent variants, Fisher’s exact test, \( P = 1 \), Fig. 2A). On a per-gene base, mitochondrially encoded NADH:ubiquinone oxidoreductase core subunit 6 (MT-ND6) had lower mutation rates per base in the unique than in the recurrent variants; however, after correcting for multiple testing using the Bonferroni method, this difference was not statistically significant (Fisher’s exact test uncorrected, \( P = 0.038 \), Fig. 2B). The unique variants had lower heteroplasmic loads than the recurrent ones, with 80.0% of the unique variants having loads below 5% (versus 16.7% of recurrent variants with loads <5%, Chi-square test, \( P < 0.0001 \), Fig. 2C). In the unique variants, we found that the rRNA and tRNA variants were exclusively found at loads <5% (Fig. 2H) and that the incidence of non-synonymous protein-coding variants with loads >10% was slightly higher than that of synonymous variants (Fisher’s exact test, \( P = 0.04349 \), Fig. 2H).

Overall, these results show that half of the oocytes differ from their siblings within a cohort due to the presence of variants unique to them. These unique variants...
Figure 2. Sibling oocytes carry both recurrent and unique variants that differ in location, type and heteroplasmic load. (A) Distribution of recurrent and unique variants in oocytes based on their location or type (in case of a variant in the protein-coding regions) in the mitochondrial genome (non-coding, protein-coding synonymous, protein-coding non-synonymous or rRNA/tRNA-coding regions). No significant differences were found in the distribution of recurrent and unique variants (Fisher’s exact test). (B) Mutation rate per base for the recurrent and unique variants in the oocytes. No significant differences across loci were found between recurrent and unique variants (Fisher’s exact test with Bonferroni correction for multiple testing). (C) Recurrent and unique variants in the oocytes categorized for their variant load. Both types of variants differed significantly in variant load (Chi-square test, $P < 0.0001$). (D) Recurrent variants in the oocytes categorized for their load and distribution in the mtDNA. (E) Distribution in the mtDNA based on their location or type of the recurrent variants in the oocytes categorized whether they were present in the somatic tissues of the donor (“Transmitted”) or only present in the oocytes of the donor and not in the somatic tissues (“Germline-specific”). (F) Recurrent transmitted and germline-specific variants in oocytes categorized for their variant load. No significant differences were found in the distribution of transmitted and germline specific variants (Fisher’s exact test). (G) Example of transmitted variants where the average load in the somatic tissues is plotted against the load of the same variant in the oocytes of the respective donor. The black lines indicate the mean load of the variant in the oocytes. (H) Unique variants in the oocytes categorized for their load and distribution in the mtDNA. The incidence of unique variants inducing non-synonymous changes was slightly higher than the unique synonymous variants at loads > 10% (Fisher’s exact test, $P = 0.04549$).
differ from the recurrent variants in their location and load, being more frequently located in the rRNA and tRNA genes, and more often resulting in a non-synonymous change. However, their pathogenic potential may be limited by their low heteroplasmic load, which in most cases was below 5%.

The differences between unique and recurrent variants in oocytes are maintained during preimplantation development

Figure 3A shows an example of the variants found in each of the single blastomeres of two sibling embryos. The variants were categorized as recurrent across sibling embryos, recurrent across blastomeres (but not across the siblings) and unique to each blastomere. We reasoned that the variants that are recurrent across oocytes of one donor would have become variants recurring across sibling embryos further in development. Six variants of this type were found in nine embryos from the four sets of sibling day-3 embryos, with a similar distribution and heteroplasmic load as the recurrent variants we found in the oocytes. Half were located in the non-coding region and the other half in the protein-coding region, 16.7% being non-synonymous changes, the highest mutation rate per base being in the hypervariable region and 66.7% of variants having a heteroplasmic load over 20% (Fig. 3B, D and F).

The second type of recurrent variant, which appears in multiple cells of the same day-3 embryo, but not between embryos of the same cohort, is found in 60% of embryos. These variants were located in 61.9% of cases in the non-coding region, 38.1% were in protein-coding sequences, 23.8% inducing non-synonymous changes and none were found in the rRNA/tRNA regions (Fig. 3B). The mutation rate per base was highest for the hypervariable region (Fig. 3D). This distribution was maintained in the day-5 blastocysts, although a slight shift was seen with an increase in variants in the non-coding region (71.4%) and a decrease in protein-coding variants, none in the rRNA/tRNA, and with the highest mutation rate per base in the hypervariable region (Fig. 3C and E). On day 3, 9.5% of recurrent variants had loads >20% and 71.4% < 5%, of which most were present in the non-coding regions (Fig. 3F and H). Noticeably, the recurrent protein-coding variants inducing non-synonymous changes were exclusively seen at loads <10%, while synonymous variants often showed loads >10% (Fig. 3H). This distribution changed slightly on day 5, with 28.6% of recurrent variants having loads >20% and 57.1% < 5%, the majority of which located in the non-coding regions (Fig. 3G and J). Remarkably, in both stages of development, these recurrent variants showed heteroplasmic loads that differed in average 3.5% points across cells or biopsies, and maximally in 13.4% points. This consistency is in line with previous reports on inherited pathogenic mtDNA variants detected during preimplantation genetic testing (32).

This shows that during the very early stages of development, these recurrent variants are homogeneously distributed to the daughter cells, suggesting that they were evenly distributed in the cytoplasm of the original oocyte. These variants restricted to one embryo differ from the ones found recurrently across embryos of the same cohort in their heteroplasmic load (71.4% of the recurrent variants across blastomeres have loads <5% versus 66.7% of the recurrent variants across siblings have loads >20%) but not in their location.

With regard to the unique variants, the difference in location was more prominent in the day-3 embryos, where 41.0% of the unique variants were located in the rRNA/tRNA regions as compared with none of the recurrent variants (Fisher’s exact test, P < 0.0001, Fig. 3B). This was also seen when looking at the mutation rate per base, where the mutation rate of the hypervariable region was significantly higher in the recurrent variants than in the unique (Fisher’s exact test uncorrected, P < 0.0001, Fig. 3D) and the rRNA regions showed higher rates in the unique variants (Fisher’s exact test uncorrected, P = 0.0003, Fig. 3D). On day 3, 83.1% of the unique variants and 71.4% of the recurrent variants had a heteroplasmic load <5% (Chi-square test, P < 0.0001, Fig. 3F). Most of the variants with a load < 5% were protein-coding variants with a similar distribution between synonymous and non-synonymous variants (Fig. 3I).

On day 5 of development, the variants in the hypervariable region were more likely to be recurrent, though not statistically significant after correcting for multiple testing (Fisher’s exact test uncorrected, P = 0.0047, Fig. 3E) while the unique rRNA/tRNA variants represented only 20.0% of the variants (Fisher’s exact test, P = 0.5453, Fig. 3C). The heteroplasmic loads were still different though not statistically significant due to the limited sample size, with 85.0% of the unique variants and 57.1% of the recurrent variants having loads <5% (Fisher’s exact test, P = 0.2896, Fig. 3G), the majority of which located in the protein-coding regions inducing non-synonymous changes (Fig. 3K).

The cells carrying unique mtDNA variants in preimplantation embryos likely give rise to stable lineages in adult individuals

In 38.7% of the bulk adult tissues we found variants that recurred between at least two samples of the same individual. These variants resembled the recurrent variants found in the oocytes and across sibling embryos, with 45.5% located in the non-coding regions, 48.5% in the protein-coding sequences and only a small proportion in the rRNA/tRNA coding regions (6.1%) (Fig. 4A). The mutation rate per base was highest in the hypervariable region (Fig. 4B), and most frequently had loads >20% (Fig. 4C). Further, there were no differences in the type of variants in function of their variant load (Fig. 4D). Also, 16.9% of the samples carried unique variants. Of these,
Figure 3. The differences in unique and recurrent variants are further maintained during preimplantation development. (A) Example of two sibling embryos carrying recurrent variants across siblings (green dots), recurrent variants across blastomeres of the same embryo but not recurring across siblings (black dots) and unique variants that are unique to one blastomere (red symbols). (B) Distribution of recurrent across siblings, recurrent across blastomeres and unique variants in day-3 embryos based on their location or type. Unique variants were more frequently located in the rRNA/tRNA regions when compared with the recurrent variants across blastomeres (Fisher’s exact test, \( P < 0.0001 \)). (C) Distribution of recurrent and unique variants in day-5 blastocysts based on their location or type. Recurrent variants were more frequently located in the hypervariable regions when compared with the unique variants across biopsies (Fisher’s exact test, \( P = 0.0114 \)). (D) Mutation rate per base for the recurrent across siblings, recurrent across blastomeres and unique variants in the day-3 embryos. Variants in the hypervariable regions were more likely to be recurrent across blastomeres (Fisher’s exact test with Bonferroni correction for multiple testing, \( P < 0.0001 \)) and variants in the rRNA regions were more likely to be unique (Fisher’s exact test with Bonferroni correction for multiple testing, \( P = 0.0003 \)). (E) Mutation rate per base for the recurrent and unique variants in the day-5 blastocysts. Variants in the hypervariable regions were more likely to be recurrent, although not statistically significant after correcting for multiple testing (Fisher’s exact test with Bonferroni correction for multiple testing, \( P = 0.0047 \)).

66.7% were found in the non-coding regions while no variants were located in the rRNA/tRNA regions (Fig. 4A), and the mutation rate per base was as expected highest in the hypervariable region (Fig. 4B). Of the unique variants, 85.2% had loads <5%, which was significantly different compared with 24.2% of the recurrent variants (Fisher’s exact test, \( P < 0.0001 \), Fig. 4C), of which 65.2% were located in the non-coding regions (Fig. 4E). In sum,
the main difference between the recurrent and unique variants in the bulk DNA samples was their heteroplasmic load.

Next, we studied single cells of two tissues of three individuals. In this part, we set the threshold for variant calling at >5% to ensure a more conservative calling. The main reason for this is the higher number of polymerase chain reaction (PCR) cycles required to amplify the mtDNA of these cells, which could result in an increase in false positives. Of the adult single cells, 72.1% carried recurrent mtDNA variants (found in at least two different cells from one individual, irrespective of the tissue of origin), while unique variants appeared in 86.8% of cells. In total, we identified 18 recurrent variants and 642 unique variants in the 68 single cells. Furthermore, not all recurrent variants in the single cells were identified in the bulk samples of the same tissues, as the average of the loads of all cells could drop below the sequencing detection limit.

The location of the recurrent and unique variants in the adult single cells was very similar, with most variants located in protein-coding and tRNA/rRNA loci, and this is in contrast with the variants identified in the bulk tissues. Protein-coding variants were more frequent at the single-cell level than in the bulk tissues (recurrent bulk 48.5% versus single-cell 72.2%, Chi-square, \( P < 0.0001 \), unique bulk 33.3% versus single-cell 83.3%, Chi-square, \( P < 0.0001 \)) reminiscent of the unique variants in the day-3 and day-5 embryos, suggesting that the cells carrying unique variants in the preimplantation embryos can give rise to stable lineages in terms of mtDNA variants in the adult individual. In the same line, there were no differences between the recurrent and unique single-cell variants in mutation rate per base across the different loci (Fig. 4G) but these were notably different to their counterparts in bulk tissues, where the hypervariable region had the highest mutation rate per base. As in the bulk tissues, recurrent variants tended to have higher variant loads than the unique variants (Chi-square test, \( P < 0.0001 \), Fig. 4H), and no differences were found in location according to heteroplasmic load (Fig. 4I and J).

Finally, we controlled the recurrent and unique single-cell variants for the type of base pair change they induce and found no statistically significant differences in the incidence of transitions, transversions or insertions and deletions (Supplementary Material, Fig. S1).

Discussion

In this study we deep-sequenced the mtDNA of a large cohort of human oocytes, blastomeres of day-3 embryos, small groups of cells of day-5 blastocysts, and single cells as well as bulk DNA of different somatic adult tissues. We identified three different types of mtDNA variants based on their recurrence or uniqueness across siblings, single cells or samples from the same individuals. These types of variants are differently distributed throughout the regions of the mtDNA and show differences in variant load (summarized in Fig. 5). In early development, the similarities between the low-load unique variants in the oocytes and the unique variants in the day-3 embryos and day-5 blastocysts on the one hand and between the recurrent variants in the oocytes and recurrent across sibling embryos on the other hand lead us to propose a model of segregation for each type of variant (illustrated in Fig. 6). The recurrent variants across oocytes are the equivalent of the recurrent variants across sibling embryos (Fig. 6A), while the variants that are recurrent restricted to one embryo are the equivalent of the unique variants in oocytes that are found at higher heteroplasmic loads. We propose that both these variants are distributed evenly in the cytoplasm, and therefore homogeneously segregated in the blastomeres throughout the cleavage divisions (Fig. 6B).

Finally, we hypothesize that variants that are unique to a single blastomere would originate from the low-load unique variants in the oocytes (Fig. 6C). This can be explained if these variants are found in a low number of mitochondria that physically cluster together in the cytoplasm. It is likely that these low-load unique variants in the oocyte arise postnatally during folliculogenesis due to replication of a subpopulation of mtDNA molecules, as demonstrated in a mouse model, while the high-load unique variants were already present in the primordial germ cells after the first bottleneck (33). This would explain why the low-load unique variants in the oocytes (and unique across blastomeres) are more frequently non-synonymous and in rRNA/tRNA loci than the high-load unique variants in oocytes (and recurrent across blastomeres) since the latter were subjected to the selective effect of the germline bottleneck (11).

During the first embryonic cleavage divisions, this colocalization would result in an asymmetric distribution of mitochondria containing the variant, leading to an embryo that already presents mosaicism as early as day 3 of development. This same pattern would continue through to the blastocyst stage, where sister cells tend to remain in proximity of each other. These lineages persist throughout development, and in the adult individual as supported by the study of Lee et al. (34) and first suggested in mammals in 1988 (35).

It is plausible that this type of segregation of mtDNA variants is a common event during the cleavage-stages of mammalian development, as it was previously shown in mice and Rhesus monkeys during the preimplantation period (34,36). Also, work on human embryos has shown that asymmetrical mitochondrial distribution can result in a proportion of blastomeres with a reduced mitochondrial pool (37), which could contribute to the fixing of mtDNA variants in the clonally derived blastomeres.

While we have identified the persistence of unique variants throughout preimplantation development, it is noteworthy that there is a decline in unique rRNA and tRNA variants from day 3 to day 5 of embryonic development. This suggests the existence of selection mechanisms during preimplantation development...
Figure 4. The cells carrying unique mtDNA variants in embryos give rise to stable lineages in adult individuals. (A) Distribution of recurrent and unique variants in adult bulk tissues based on their location or type. No statistically significant differences were observed (Fisher’s exact test). (B) Mutation rate per base for the recurrent and unique variants in adult bulk tissues. No statistically significant differences were observed (Fisher’s exact test). (C) Recurrent and unique variants in adult bulk tissues categorized for their variant load. Unique variants had more often loads < 5% (Fisher’s exact test, P < 0.0001). (D) Recurrent variants in adult bulk tissues categorized for their load and distribution in the mtDNA. (E) Unique variants in adult bulk tissues categorized for their load and distribution in the mtDNA. (F) Distribution of recurrent and unique variants in adult single cells based on their location or type. Unique variants were more frequently located in protein-coding regions (Chi-square test, P = 0.0006). (G) Mutation rate per base for the recurrent and unique variants in adult single cells. No statistically significant differences were observed (Fisher’s exact test). (H) Recurrent and unique variants in adult single cells categorized for their variant load. Unique variants more often showed loads < 10% (Chi-square test, P < 0.0001). (I) Recurrent variants in adult single cells categorized for their load and distribution in the mtDNA. (J) Unique variants in adult single cells categorized for their load and distribution in the mtDNA.
that filter out pathogenic variants. This selection has been observed later in the mouse ontogenesis, where deleterious variants present in the oocyte and in the embryo were found to be selected against by unknown mechanisms occurring postimplantation and postnatally (38).

Next to possible selection mechanisms, the somatic bottleneck can cause a diverse population of cells harboring different mtDNA heteroplasmic variants and at different loads. We observed that the variant load in adult single cells varies widely among cells of the same tissue. The number of cells harboring variants at higher loads could exceed a certain tissue threshold and could increase the risk of pathologies (39). This was already described in muscle fibers (40) and the heart (41).

Furthermore, our study confirmed the presence of tissue-specific variants. These variants were located in the non-coding regions and were in close proximity to regions that regulate the mtDNA replication. This is in line with the suggestion of Samuels et al. (22) that these variants could have a beneficial effect on the mtDNA replication in the given tissue. For instance, variant m.72 T > C was already described by two other research groups in liver and kidney tissue (21, 22).

A limitation of this study is that we required to use different detection limits in the bulk DNA (1.5%), oocytes and embryos (2%) and somatic cells (5%). It is also important to bear in mind that a unique variant was defined as a variant that is only detected in a certain cell above a certain threshold and that potential recurrent variants below that said threshold were not included in this study. We controlled for a possible bias where the detection limit would change the outcome of this study using the data set of the day-3 embryos, since this data set gave us the most complete insight on recurrent and unique variants. Reassuringly, we found that changing the detection limit would not change the outcomes of the study (Supplementary Material, Fig. S2).

In conclusion, our work is the first to comprehensively describe mtDNA mosaicism in human preimplantation development and identifies a subgroup of low-load variants that may give rise to stable lineages of genetically diverse cells in the adult. We propose that these lineages appear due to asymmetrical distribution of mitochondria carrying mtDNA variants in the oocyte, which possibly appeared during folliculogenesis. Finally, future research will give us more insight on the mechanisms behind the asymmetrical distribution of variants in the oocyte and on the potential implications of this type of mosaicism in health and disease.

**Material and Methods**

**Sample and single-cell collection**

All buccal (N = 59), blood (N = 57) and urine (N = 26) samples and oocytes (N = 254) and embryos on day 3 (N = 25) or at the blastocyst stage (N = 7) were obtained after signed informed consent from the donors at the Center
Figure 6. Proposed segregation model of recurrent and unique variants in the preimplantation development. The figure shows the development of three oocytes of the same donor, resulting in three sibling blastocysts. Mitochondria in red (A–C) are carrying variants that are recurrent across sibling oocytes and embryos. The mitochondria in blue (B) carry unique variants at higher loads in the oocyte and are homogenously distributed together with the red mitochondria in the ooplasm. They remain in a similar distribution during development and become the variants found as unique in one individual but present in all their tissues. The mitochondria in green (C) carry unique variants at lower loads and cluster in the ooplasm. These mitochondria will remain in close proximity to each other during the cleavage stages and will be present in only one blastomere. This cell will then give rise to a lineage of cells that in the adult individual manifests as a subpopulation of rare cells carrying this same variant, potentially across different germ layers. This figure was created with BioRender.

for Medical Genetics and the Brussels IVF Center for Reproductive Medicine of the Universitair Ziekenhuis Brussel (Table 1 and Supplementary Material, Table S5). Supernumerary oocytes were donated after oocyte pick-up, while preimplantation embryos were donated for research after the legally determined cryostorage period of 5 years passed (42). Prior to the start of the study, approval was acquired from the Local Ethical Committee of the Vrije Universiteit Brussel and the Universitair Ziekenhuis Brussel, and by the Federal Ethical Committee on Medical and Scientific Research on embryos in vitro.

Day-3 embryos were warmed using the Vitrification Thaw kit (Vit Kit-Thaw, Irvine Scientific, USA) according to manufacturer’s instructions. Subsequently, they were left to recover in 25 μL droplets of Origio blastocyst medium (Origio, The Netherlands) for 3 h in an incubator at 37°C with 89% N₂, 6% CO₂ and 5% O₂. The day-3 embryos and oocytes were freed from their zona pellucida by incubating them in a droplet of pronase (100 mg/100 μL human tubal fluid) and by gently pipetting them up and down. The ICM and TE samples were biopsied from day-5 blastocysts that were diagnosed as affected by a monogenic disease after preimplantation genetic testing, as previously described (19). The oocytes, embryos, ICM and TE samples and adult single cells were washed three times in Ca²⁺ - and Mg²⁺-free medium. The individual blastomeres obtained from cleavage stage embryos dissociated in the Ca²⁺ - and Mg²⁺-free medium were washed three additional times before collecting in 2.5 μL ALB (alkaline lysis buffer) as described before (43). The samples were kept at −20°C until further processing. The bulk DNA of the somatic tissues was extracted using a kit according to manufacturer’s instructions (DNeasy Blood and Tissue, Qiagen, Catalogue number 69504).

mtDNA enrichment and massive parallel sequencing

Before PCR, the single oocytes and blastomeres, ICM/TE biopsies from the day-5 blastocysts and adult cells were incubated at 65°C for 10 min to ensure full lysis of the cells. Long-range PCR (44) was performed using a primer set to generate amplicons of 13 Kbp (5042f – 1424r). Previously, this protocol was controlled for the
possible amplification of nuclear mtDNA sequences (NUMTs) by performing the PCR on mtDNA-lacking RhoZero cells and confirming the absence of PCR amplicons. The primer sequences for the primer set were 5’-AGGAGTTCTACCGTACAACC-3’ (forward) and 5’-ATCCACTTTGACCCTTAAG-3’ (reverse). The amplification was done in a total volume of 50 μL per sample containing 10 μL of LongAmp buffer, 2 μL of Taq DNA polymerase (LongAmp Taq DNA Polymerase kit, New England Biolabs), 7.5 μL of deoxynucleotide triphosphates (dNTPs) (dNTP set, Illustra™), 2 μL of each primer (10 μM), 2.5 μL Tricine (200 mM, Sigma-Aldrich) and 21.5 μL H2O. The PCR protocol started with an initiation step of 30 s at 94°C followed by a touchdown of eight cycles of 15 s at 94°C, 30 s at 64°C (−0.4°C per cycle) and 11 min at 61°C, 29–37 cycles were added of 15 s at 94°C, 30 s at 61°C and 11 min at 65°C and completed with a final elongation step of 11 min at 65°C. Successful amplification was confirmed using agarose gel electrophoresis (1.5%). After PCR purification with AMPure beads (Beckmann Coulter), library preparation as described in Mertens et al. (44) was performed using the TruSeq DNA PCR-free Library Preparation kit (Illumina, Eindhoven, The Netherlands). The amplicons were sheared using a Covaris™ M220 sonicator (Life Technologies), following instrument specification to generate fragments of ±100 base pairs. The detection of the nucleotide sequence was done on the Illumina NovaSeq6000 platform using the according kit (Illumina).

### Data analysis and bioinformatics processing

After extracting 1.5 million reads at random (this was to prevent computational errors in the downstream analysis), the generated fastq files were aligned to the reference genome (NC_012920.1) with BWA-MEM, generating bam files. These bam files were then realigned for insertions and deletions and base recalibration using GATK v3.3. and uploaded to mtDNA server (45) (v1.1.3) which detected the homoplasmic (>98.5% frequency) and heteroplasmic (<98.5% frequency) single nucleotide variants (SNV) as well as the haplogroup and possible contaminations. Small insertions and deletions and SNVs were confirmed by Mutect2 (46) (GATK v3.6 Mutect2). The threshold for reliably detecting SNVs was 5% for single adult cells, 2% for single oocytes, blastomeres and ICM/TE biopsies and 1.5% for bulk DNA. If mtDNA server and Mutect2 gave discrepancies regarding the variant load, the load was calculated from the coverage file using following formula: the sum of the variant coverage forward and variant coverage reverse divided by the total coverage of the base position. Variants on base positions with a coverage lower than 1000 were excluded from further analysis. The average coverage per base was 8497 reads (see Supplemental Microsoft Excel file for average coverage per sample). The annotation of the variants was done using MitoWheel (www.mitowheel.org) and possible amino-acid changes were identified using MutPred2 (47). Variants with a possible or proven pathogenicity were included in the study but were not categorized differently. A more detailed protocol of the bioinformatic processing and the validation of the full sequencing setup can be found in our previously published work (19,44,48).

### Statistics

Statistics were performed using the two-tailed Fisher’s exact with Bonferroni correction for multiple testing (P < 0.0029 were considered significant) or using the Chi-square test (P-values < 0.05 were considered significant).

### Supplementary Material

Supplementary Material is available at HMG/ online.

### Funding

This work was supported by the Scientific Fund Willy Gepts of the Universitair Ziekenhuis Brussel (Wetenschappelijk Fonds Willy Gepts) the Methusalem Grant of Karen Sermon of the Vrije Universiteit Brussel the Research Foundation Flanders (Fonds voor Wetenschappelijk Onderzoek Vlaanderen, FWO 1518418N). M.R. and E.C.D.D. are doctoral fellows at the Fonds voor Wetenschappelijk Onderzoek Vlaanderen.

Conflict of Interest statement. The authors declare no conflicts of interest.

### References

1. Forsberg, L.A., Gisselsson, D. and Dumanski, J.P. (2016) Mosaicism in health and disease — clones picking up speed. Nat. Rev. Genet., 18, 128–142.
2. Starostik, M.R., Sosina, O.A. and McCoy, R.C. (2020) Single-cell analysis of human embryos reveals diverse patterns of aneuploidy and mosaicism. Genome Res., 30, 814–825.
3. Stewart, J.B. and Chinnery, P.F. (2015) The dynamics of mitochondrial DNA heteroplasmacy: implications for human health and disease. Nat. Rev. Genet., 16, 530–542.
4. Yao, Y., Ogasawara, Y., Kajigaya, S., Molldrem, J.J., Falcão, R.P., Pintão, M.-C., McCoy, J.P., Rizzatti, E.G. and Young, N.S. (2007) Mitochondrial DNA sequence variation in single cells from leukemia patients. Blood, 109, 756–762.

5. Ogasawara, Y., Nakayama, K., Tarnowka, M., McCoy, J.P., Kajigaya, S., Levin, B.C. and Young, N.S. (2005) Mitochondrial DNA spectra of single human CD3+4+ cells, T cells, B cells, and granulocytes. Blood, 106, 3271–3284.

6. Cantuts-Castelvetri, I., Lin, M.T., Zheng, K., Keller-McGandy, C.E., Betensky, R.A., Johns, D.R., Beal, M.F., Standaert, D.G. and Simon, D.K. (2005) Somatic mitochondrial DNA mutations in single neurons and glia. Neurobiol. Aging, 26, 1343–1355.

7. Payne, B.A.I., Wilson, I.J., Hatelye, C.A., Horvath, R., Santibanez-Koref, M., Samuels, D.C., Price, D.A. and Chinnery, F.P. (2011) Mitochondrial ageing is accelerated by anti-retroviral therapy through the clonal expansion of mtDNA mutations. Nat. Genet., 43, 806–810.

8. Payne, B.A.I., Wilson, I.J., Yu-Wai-Man, P., Coxhead, J., Deehan, D., Horvath, R., Taylor, R.W., Samuels, D.C., Santibanez-Koref, M. and Chinnery, F.P. (2013) Universal heteroplasmy of human mitochondrial DNA. Hum. Mol. Genet., 22, 384–390.

9. Guo, Y., Li, C.-I., Sheng, Q., Winther, J.F., Cai, Q., Boice, J.D. and Shyr, Y. (2013) Very low-level heteroplasmy mtDNA variations are inherited in humans. J. Genet. Genomics, 40, 607–615.

10. Ye, K., Lu, J., Ma, F., Keinan, A. and Gu, Z. (2014) Extensive pathogenicity of mitochondrial heteroplasmy in healthy human individuals. Proc. Natl. Acad. Sci. U.S.A., 111, 10654–10659.

11. Flores, V.I., Pyle, A., Dietmann, S., Wei, W., Tang, W.W.C., Irie, N., Payne, B., Capalbo, A., Noli, L., Coxhead, J. et al. (2018) Segregation of mitochondrial DNA heteroplasmy through a developmental genetic bottleneck in human embryos. Nat. Cell Biol., 20, 144–151.

12. Chou, J.-Y. and Leu, J.-Y. (2015) The red queen in mitochondria: cyto-nuclear co-evolution, hybrid breakdown and human diseases. Front. Genet., 6, 187.

13. St. John, J.C., Facuccho-Oliveira, J., Jiang, Y., Kelly, R. and Salah, R. (2010) Mitochondrial DNA transmission, replication and inheritance: a journey from the gamete through the embryo and into offspring and embryonic stem cells. Hum. Reprod. Update, 16, 488–509.

14. Hashimoto, S., Morimoto, N., Yamanaka, M., Matsumoto, H., Yamochi, T., Goto, H., Inoue, M., Nakao, K., Shibahara, H. and Morimoto, Y. (2017) Quantitative and qualitative changes of mitochondria in human preimplantation embryos. J. Assist. Reprod. Genet., 34, 573–580.

15. Wallace, D. (2010) Mitochondrial DNA mutations in disease and ageing. Environ. Mol. Mutagen., 51, 440–450.

16. Trifunovic, A., Wredenberg, A. and Falkenberg, M. (2004) Premature ageing in mice expressing defective mitochondrial DNA polymerase. Nature, 429, 417–423.

17. Fosslien, E. (2001) Mitochondrial medicine - molecular pathology of defective oxidative phosphorylation. Ann. Clin. Lab. Sci., 31, 25–67.

18. Van den Armeele, J., Li, A.Y.Z., Ma, H. and Chinnery, P.F. (2020) Mitochondrial heteroplasmy beyond the oocyte bottleneck. Semin. Cell Dev. Biol., 97, 156–166.

19. Zambelli, F., Mertens, J., Dziedzicka, D., Sterckx, J., Markouli, C., Keller, A., Tropel, P., Jung, L., Viville, S., Van De Velde, H. et al. (2018) Random mutagenesis, clonal events and embryonic or somatic origin determine the mitochondrial DNA variant type and load in human pluripotent stem cells. Stem Cell Rep., 11, 102–114.

20. Kang, E., Wang, X., Tippner-Hedges, R., Ma, H., Holmes, C.D.L., Gutierrez, N.M., Lee, Y., Van Dyken, C., Ahmed, R., Li, Y. et al. (2016) Age-related accumulation of somatic mitochondrial DNA mutations in adult-derived human iPSCs. Cell Stem Cell, 18, 625–636.

21. Li, M., Schröder, R., Ni, S., Madea, B. and Stoneking, M. (2015) Extensive tissue-related and allele-related mtDNA heteroplasmy suggests positive selection for somatic mutations. Proc. Natl. Acad. Sci. U.S.A., 112, 2491–2496.

22. Samuels, D.C., Li, C., Li, B., Song, Z., Torstenson, E., Boyd Clay, H., Rokas, A., Thornton-Wells, T.A., Moore, J.H., Hughes, T.M. et al. (2013) Recurrent tissue-specific mtDNA mutations are common in humans. PLoS Genet., 9, e1003929.

23. Lawless, C., Greaves, L., Reeve, A.K., Turnbull, D.M. and Vincent, A.E. (2020) The rise and rise of mitochondrial DNA mutations. Open Biol., 10, 200061.

24. Tajima, H., Sueoka, K., Moon, S.Y., Nakabayashi, A., Sakurai, T., Murakoshi, Y., Watanabe, H., Iwata, S., Hashiba, T., Kato, S. et al. (2007) The development of novel quantification assay for mitochondrial DNA heteroplasmy aimed at preimplantation genetic diagnosis of Leigh encephalopathy. J. Assist. Reprod. Genet., 24, 227–232.

25. Salleveit, S.C.E.H., Dreesen, J.C.F.M., Druesedau, M., Hellebrekers, D.M.E.I., Paulussen, A.D.C., Coonen, E., Van Golde, R.J.T., Geraedts, J.P.M., Gianaroli, L., Magli, M.C. et al. (2017) PGD for the m.14487 T>C mitochondrial DNA mutation resulted in the birth of a healthy boy. Hum. Reprod., 32, 698–703.

26. Steffann, J., Frydman, N., Gigarel, N., Burlet, P., Ray, P.F., Fanchin, R., Feyereisen, E., Kerbrat, V., Tachdjian, G., Bonnefont, J.-P. et al. (2006) Analysis of mtDNA variant segregation during early human embryonic development: a tool for successful NARP preimplantation diagnosis. J. Med. Genet., 43, 244–247.

27. Salleveit, S.C.E.H., Dreesen, J.C.F.M., Druesedau, M., Spierts, I., Coonen, E., Van Tienen, F.H.J., Van Golde, R.J.T., De Coo, I.F.M., Geraedts, J.P.M., De Die-Smulders, C.E.M. et al. (2013) Preimplantation genetic diagnosis in mitochondrial DNA disorders: challenge and success. J. Med. Genet., 50, 125–132.

28. Trefl, N.R., Campos, J., Tao, X., Levy, B., Ferry, K.M. and Scott, R.T. (2012) Blastocyst preimplantation genetic diagnosis (PGD) of a mitochondrial DNA disorder. Fertil. Steril., 98, 1236–1240.

29. Gigarel, N., Hesters, L., Samuels, D.C., Monnot, S., Burlet, P., Kerbrat, V., Lamazou, F., Benachi, A., Frydman, R., Feingold, J. et al. (2011) Poor correlations in the levels of pathogenic mitochondrial DNA mutations in polar bodies versus oocytes and blastomeres in humans. Am. J. Hum. Genet., 88, 494–498.

30. Heindryckx, B., Neupane, J., Vandewoestyne, M., Christodoulou, C., Jackers, Y., Gerris, J., Van Den Abbeel, E., Van Coster, R., Deforce, D. and De Sutter, P. (2014) Mutation-free baby born from a mitochondrial encephalopathy, lactic acidosis and stroke-like syndrome carrier after blastocyst trophectoderm preimplantation genetic diagnosis. Mitochondrion, 18, 12–17.

31. Vandewoestyne, M., Heindryckx, B., De Gheselle, S., Lepeiz, T., Neupane, J., Gerris, J., Van Coster, R., De Sutter, P. and Deforce, D. (2012) Poor correlation between polar bodies and blastomere mutation load in a patient with m.3243A>G tRNALEU(UUR) point mutation. Mitochondrion, 12, 477–479.
32. Sallevelt, S.C.E.H., Dreesen, J.C.F.M., Coonen, E., Paulussen, A.D.C., Hellebrekers, D.M.E.I., De Die-Smulders, C.E.M., Smeets, H.J.M. and Lindsey, P. (2017) Preimplantation genetic diagnosis for mitochondrial DNA mutations: analysis of one blastomere suffices. J. Med. Genet., 54, 693–697.

33. Wai, T., Teoli, D. and Shoubridge, E.A. (2008) The mitochondrial DNA genetic bottleneck results from replication of a subpopulation of genomes. Nature, 40, 1484–1488.

34. Lee, H.S., Ma, H., Janes, R.C., Tachibana, M., Sparman, M., Woodward, J., Ramsey, C., Xu, J., Kang, E.J., Amato, P. et al. (2012) Rapid mitochondrial DNA segregation in primate preimplantation embryos precedes somatic and germline bottleneck. Cell Rep., 1, 506–515.

35. Laipis, P.J., Van De Walle, M.J. and Hauswirth, W.W. (1988) Unequal partitioning of bovine mitochondrial genotypes among siblings. Proc. Natl. Acad. Sci. U.S.A., 85, 8107–8110.

36. Cree, L.M., Samuels, D.C., De Sousa Lopes, S.C., Rajasimha, H.K., Wonnapinij, P., Mann, J.R., Dahl, H.H. and Chinnery, P.F. (2008) A reduction of mitochondrial DNA molecules during embryogenesis explains the rapid segregation of genotypes. Nat. Genet., 40, 249–254.

37. Johnston, I.G. and Burgstaller, JP (2019) Evolving mtDNA populations within cells. Biochem. Soc. Trans., 47, 1367–1382.

38. Wanagat, J., Cao, Z., Pathare, P. and Aiken, J.M. (2001) Mitochondrial DNA deletion mutations colocalize with segmental electron transport system abnormalities, muscle fiber atrophy, fiber splitting, and oxidative damage in sarcopenia. FASEB J., 15, 322–323.

39. Baris, O.R., Ederer, S., Neuhaus, J.F.G., von Kleist-Retzow, J.-C., Wunderlich, C.M., Pal, M., Wunderlich, F.T., Peeva, V., Zsurka, G., Kunz, W.S. et al. (2015) Mosaic deficiency in mitochondrial oxidative metabolism promotes cardiac arrhythmia during ageing. Cell Metab., 21, 667–677.

40. Spits, C., Le Caignec, C., De Rycke, M., Van Haute, L., Van Steirteghem, A., Liebaers, I. and Sermon, K. (2006) Whole-genome multiple displacement amplification from single cells. Nat. Protoc., 1, 2015–2016.

41. Venter, M., Malan, L., Van Dyk, E., Elson, J.L. and Van Der Westhuizen, F.H. (2017) Using MutPred derived mtDNA load scores to evaluate mtDNA variation in hypertension and diabetes in a two-population cohort: the SABPA study. J. Genet. Genomics, 44, 139–149.

42. Zambelli, F., Vancampenhout, K., Daneels, D., Brown, D., Mertens, J., Van Dooren, S., Caljon, B., Gianaroli, L., Sermon, K., Voet, T. et al. (2017) Accurate and comprehensive analysis of single nucleotide variants and large deletions of the human mitochondrial genome in DNA and single cells. Eur. J. Hum. Genet., 25, 1229–1236.