The dynamics of mitochondrial DNA heteroplasmy: implications for human health and disease

James B. Stewart[1] and Patrick F. Chinnery[2]

Abstract | Common genetic variants of mitochondrial DNA (mtDNA) increase the risk of developing several of the major health issues facing the western world, including neurodegenerative diseases. In this Review, we consider how these mtDNA variants arose and how they spread from their origin on one single molecule in a single cell to be present at high levels throughout a specific organ and, ultimately, to contribute to the population risk of common age-related disorders. mtDNA persists in all aerobic eukaryotes, despite a high substitution rate, clonal propagation and little evidence of recombination. Recent studies have found that *de novo* mtDNA mutations are suppressed in the female germ line; despite this, mtDNA heteroplasmy is remarkably common. The demonstration of a mammalian mtDNA genetic bottleneck explains how new germline variants can increase to high levels within a generation, and the ultimate fixation of less-severe mutations that escape germline selection explains how they can contribute to the risk of late-onset disorders.

Proteobacterium
A major phylum of Gram-negative bacteria.

[1]*Max Planck Institute for Biology of Ageing, Cologne 50931, Germany.*
[2]*Wellcome Trust Centre for Mitochondrial Research, Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne NE1 1BZ, UK.*
Correspondence to P.F.C.
e-mail: patrick.chinnery@ncl.ac.uk
doi:10.1038/nrg3966

Since their discovery ~25 years ago[1,2], inherited mitochondrial DNA (mtDNA) mutations have emerged as a major cause of human disease. Originally thought to be extremely rare, mtDNA disorders are now known to affect ~1 in 4,300 of the population[3,4]. Hundreds of different point mutations and deletions of mtDNA have been associated with a wide range of overlapping clinical phenotypes. These disorders can present from birth to old age, and often affect the nervous system (recently reviewed in REFS 5,6). In parallel, somatic mutations of mtDNA have been shown to accumulate with age and to preferentially collect in organs affected in several common late-onset diseases[7]. Their role in neurodegenerative diseases is particularly compelling (recently reviewed in REFS 8,9), bolstered by recent evidence that inherited and acquired mtDNA mutations can interact and cause ageing phenotypes (see REF. 10, and accompanying commentary REF. 11). Finally, independent genetic association studies have shown that several late-onset degenerative disorders are associated with common inherited polymorphisms of mtDNA[12] (reviewed in REFS 7,13). These polymorphisms were thought to not have any phenotypic effects, but emerging evidence has shown that they can alter mitochondrial function[14] and thereby alter the risk of developing common diseases affecting organs dependent on mitochondrial energy metabolism. It has not been clear how these seemingly disconnected observations fit together, but recent observations have cast light on the mechanisms linking them.

In this Review, we describe a common thread linking the origin of mtDNA mutations from single molecules within an organelle to the characterization of mtDNA variation across human populations. We explain why the mechanisms responsible for mtDNA mutations, their inheritance and their subsequent fixation in the maternal germ line are important for our understanding of human disease and ageing. These findings challenge our understanding of the complex relationship between the cell nucleus and the mitochondrion, and raise questions relevant for human evolution.

Mitochondrial biogenesis and mtDNA

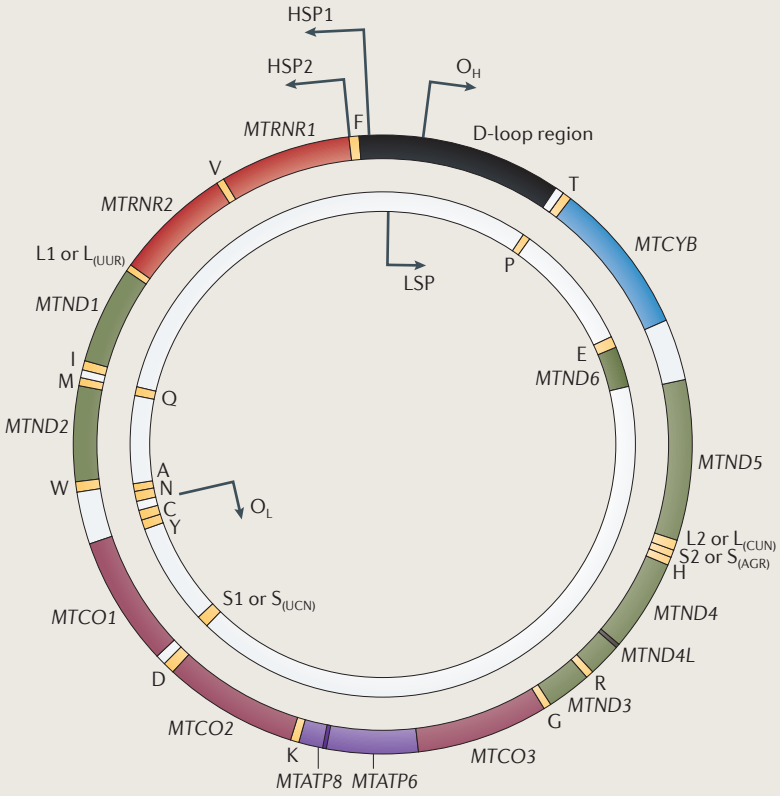
Approximately 2 billion years ago, a primitive proteobacterium became associated with a prokaryotic cell, leading to one of the most enduring symbiotic relationships in biology[15]. After endocytosis by the other cell, the proteobacterium became a double-membraned cytoplasmic organelle, and gradually passed genetic material into what was the prokaryotic cell nucleus. Over millions of years, this process reduced the overall

Box 1 | **Human mtDNA**

The human mitochondrial DNA (mtDNA) genome comprises 16,569 base pairs, which form an inner 'light' (L) strand and an outer 'heavy' (H) strand (the names reflect strand guanine content, which affects molecular weight) (see the figure). mtDNA genes encoding structural subunits of the mitochondrial respiratory chain include: mitochondrially encoded NADH dehydrogenase 1 (*MTND1*)–*MTND6* and *MTND4L* (complex I); mitochondrially encoded cytochrome *b* (*MTCYB*; complex III); mitochondrially encoded cytochrome c oxidase I (*MTCO1*)–*MTCO3* (complex IV); and mitochondrially encoded ATP synthase 6 (*MTATP6*) and *MTATP8* (complex V). The 22 tRNA and 2 rRNA genes are interspersed between the peptide-encoding genes.

mtDNA replication is carried out by a multi-protein replisome comprising mtDNA polymerase-γ (Pol γ), a helicase (Twinkle; also known as PEO1), topoisomerase I, mtDNA single-stranded DNA-binding protein and others. Two models of mtDNA replication have been proposed. In the 'strand-displacement' or 'asynchronous' model[111], replication is initiated by transcription within the non-coding mtDNA displacement (D) loop and proceeds clockwise from the origin of heavy-strand replication (O_H; also known as OriH) until the origin of light-strand replication (O_L) is exposed, allowing light-strand synthesis to proceed clockwise until the entire molecule is copied. Alternatively, symmetric strand-coupled replication might occur in certain circumstances[112,113]; here, replication is initiated from multiple origins distributed across a 4 kb fragment 3′ from the D loop and proceeds in both directions in replication 'bubbles' (REF. 114). Replication is arrested at O_H, allowing the remainder of the molecule to be copied in one direction.

Human mtDNA has three proposed transcription promoters: heavy strand promoter 1 (HSP1) enables transcription of the two ribosomal RNAs (but see REF. 115); HSP2 promotes transcription of the rest of the heavy strand as a large polycystronic transcript, which is subsequently spliced into functional tRNA, rRNA and mRNAs; and LSP promotes transcription of the light strand as either one long transcript or as several small primer transcripts by the RNase mitochondrial RNA processing (MRP). These primers link transcription to mtDNA replication. Transcription initiation involves several proteins, including the mitochondrial RNA polymerase and the mitochondrial transcription factors A and B2 (REF. 116).



size of the intra-organellar genome and consolidated the symbiotic relationship between the organelle, which subsequently became the mitochondrion, and the nucleus of all eukaryotic cells.

Incorporating ~10% of the cellular proteome[16], mitochondria carry out a central role in cellular energy metabolism, are important in intracellular calcium signalling and in the biogenesis of iron–sulfur clusters, and are involved in apoptotic cell death[17]. Human mitochondria have retained a 16.5 kb molecule of double-stranded DNA — mtDNA — which encodes 13 core peptide subunits of the oxidative phosphorylation system and 24 RNAs, which are essential for intra-mitochondrial protein synthesis[18,19] (BOX 1). However, the majority of mitochondrial proteins, the total number of which is currently estimated to be ~1,500 (REF. 20), are transcribed from genes present in the nucleus, and are translated in the cytosol before their delivery across the mitochondrial membrane. The optimal functioning of the human mitochondrion is therefore crucially dependent on these two distinct genetic systems; mutations that enhance or diminish the function of the mitochondrion could occur in either the mitochondrial or the nuclear genes, and would be subject to selection.

Since the time of the original symbiotic event, several differences have arisen between mtDNA and nuclear DNA. The mtDNA genetic code differs from the genetic code of the nuclear genome — for vertebrate mtDNA, codons AUA and AUG both code for methionine, UGA codes for tryptophan (not a stop codon, as in the nuclear genome), and AGA and AGG are read as stop codons (not arginine, as in the nuclear genome). In addition, mtDNA typically maintains a circular and compact coding arrangement and, instead of paired chromosomes, there are many independently proliferating copies of mtDNA present within each cell. In human cells, the number of mtDNA copies ranges from ~100 copies in sperm to hundreds of thousands in an unfertilized oocyte. The copy number seems to be tightly regulated and can vary in some cell types over time (for example, the number can increase in muscle cells with exercise). mtDNA is packaged together with proteins into nucleoids, and each nucleoid contains one or two mtDNA molecules[21]. Finally, although the relevant machinery is present within the mitochondrion, mtDNA repair is thought to be less efficient than nuclear DNA repair, in part because the high copy number buffers against the deleterious effects of stochastic mutation events.

Given that there are multiple copies of mtDNA within each cell, mutations can either affect all of the molecules (this is termed homoplasmy) or a proportion of the molecules (termed heteroplasmy) (FIG. 1). The level of heteroplasmy can vary between cells in the same tissue or organ, from organ to organ within the same person, and between individuals in the same family. As mtDNA is inherited exclusively down the maternal line, it undergoes negligible intermolecular recombination that is manifest at the population level[22], so mtDNA lineages are largely considered to be clonal.

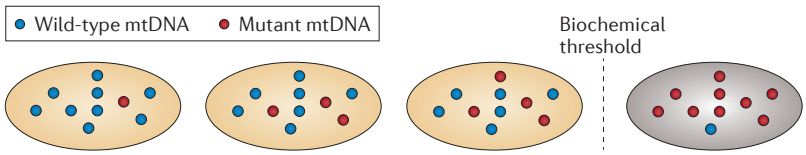


Figure 1 | **mtDNA heteroplasmy and the threshold effect.** Mitochondrial DNA (mtDNA) mutations that have occurred within approximately three human generations are usually heteroplasmic, and the same cell can contain varying proportions of mutated and wild-type mtDNA. If a mutation is pathogenic, the cell can usually tolerate a high percentage level of this variant before the biochemical threshold is exceeded and a defect in the respiratory chain is detected. Typically, this threshold level is >80%, suggesting that most mtDNA mutations are haploinsufficient or recessive[64].

Monogenic human disease

The first clear evidence linking mtDNA mutations with human disease came from studies of patients with unexplained neurological and neuromuscular diseases in the late 1980s, when parallel works led to the identification of the canonical pathogenic mutations of mtDNA. Homoplasmic point mutations in the region encoding subunit 1 of respiratory chain complex I (mitochondrially encoded NADH dehydrogenase 4 (*MTND4*, m.11778G>A)) were shown to be segregated with an inherited predisposition to blindness (Leber hereditary optic neuropathy (LHON)) in large maternal pedigrees[1]; and in sporadic cases of a biochemically defined mitochondrial myopathy patients were found to have large heteroplasmic deletions of mtDNA[2].

With advancing molecular techniques, the next decade saw the description of several hundred different point mutations and deletions of mtDNA in patients displaying a wide variety of clinical phenotypes (BOX 2, and reviewed in REF. 23). Despite the clinical and genetic heterogeneity, several trends have emerged. mtDNA point mutations are usually inherited down the maternal lineage, with multiple individuals in the same family being affected. By contrast, mtDNA deletions are rarely inherited (and possibly only through an intermediary duplication rearrangement) and are never homoplasmic[24]. Homoplasmic mtDNA point mutations usually cause a relatively mild biochemical defect that typically affects only one organ or tissue (for example, LHON, deafness or cardiomyopathy, although there are notable exceptions in which homoplasmic mutations cause severe multisystem diseases[25–27]). Conversely, heteroplasmic mutations affect multiple organ systems, particularly the brain and spinal cord, muscle, peripheral nerve, heart and endocrine organs. The level of heteroplasmy correlates with the extent of organ involvement — and therefore with the degree of severity of the clinical phenotype — and the biochemical defect is usually severe in affected tissues[5].

For heteroplasmic mutations, *in vitro* studies using cybrid cell lines[28] and studies using single cells isolated from tissue-biopsy samples[29] have shown that the proportion of mutated mtDNA must exceed a critical threshold level (typically 60–80%) before the biochemical defect can be detected using established laboratory techniques. The precise threshold varies from mutation to mutation and from tissue to tissue. Differences in the levels of heteroplasmy and the precise thresholds are thought to contribute to the characteristic patterns of organ vulnerability observed for different mitochondrial diseases and the clinical heterogeneity observed in patients harbouring the same mtDNA mutation.

Although some emerging evidence suggests that environmental and epistatic factors might influence the clinical manifestation of these disorders, particularly for homoplasmic mtDNA mutations[30], the primary mtDNA mutation is essential for the phenotype, so these disorders are considered to be primarily monogenic diseases (BOX 2). Epidemiological studies have shown that these primary mtDNA diseases affect ~1 in 4,300 of the population[4], making them among the most common inherited metabolic disorders. Intriguingly, the 10 most common pathogenic mtDNA point mutations were subsequently found to occur in ~1 in 200 healthy individuals, albeit generally at low levels of heteroplasmy[31]. The techniques used at the time could reliably detect only specific mtDNA alleles and levels of heteroplasmy that were >1%. However, these observations raised questions about the true frequency of potentially pathogenic mtDNA mutations in humans, how these mutations occurred and how levels of heteroplasmy might increase from just one molecule to eventually reach the high-percentage levels known to cause diseases.

Somatic mutations of mtDNA and disease

In patients with primary mtDNA disease, all of the mutated molecules have precisely the same mutation. However, tissue biopsies taken from elderly but otherwise healthy humans harbour very low levels of several different mtDNA mutations, and both mtDNA deletions and point mutations have been detected in the same healthy elderly individuals, typically in post-mitotic (non-dividing) tissues[32–35]. Although the overall mutational burden on a tissue or organ is low, the individual mutations can accumulate to higher levels within individual cells, and thus cause a biochemical defect[33]. The proportion of biochemically deficient cells subsequently increases over time, raising the possibility that these defects contribute to the ageing process; the mechanisms responsible are discussed below. A higher frequency of biochemically deficient cells has been observed in organs affected by several late-onset degenerative diseases, including Parkinson disease[36,37] and Alzheimer disease[35], implicating the somatic mutations in the pathogenesis of these common human disorders. The mechanistic involvement of these mutations is supported by data from animal models, in which an excess of inherited or *de novo* mutations can lead to classical age-related phenotypes[10,38,39].

mtDNA mutations have also been observed in patients with common cancers[40,41], leading to the suggestion that they might contribute to oncogenesis or metastatic spread. This concept is appealing because of the known shift from oxidative metabolism to glycolysis that occurs in malignancy (known as the Warburg effect), and is supported by the apparent observation of specific mtDNA mutations in different types of cancer. However,

Iron–sulfur clusters
Associated iron and sulfur molecules forming part of an iron–sulfur protein.

Homoplasmy
A situation in which all the mitochondrial DNA molecules within a cell or organism are identical.

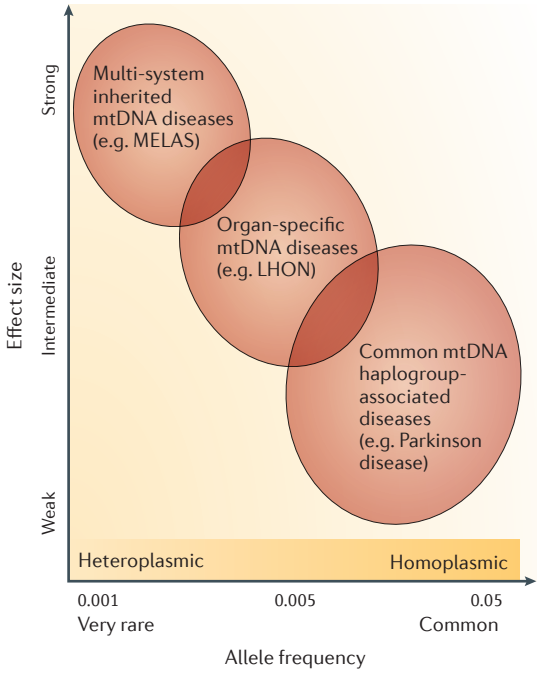
Heteroplasmy
A mixture of wild-type and mutant mitochondrial DNA within the cell. The percentage of mtDNA containing mutations can vary from 1% to 99%.

Cybrid
A cytoplasmic hybrid cell line made by fusing a whole cell with a cytoplast and usually used as a cellular model system to study the effects of mitochondrial DNA variants or mutations on the same nuclear genetic background.

Warburg effect
The shift from aerobic to anaerobic (glycolytic) metabolism that characterizes rapidly dividing cancer cell lineages.

Box 2 | **The range of mtDNA variation and its influence on human disease**

Mitochondrial DNA (mtDNA) variants are likely to fall within a continuous range, from those that are causal for monogenic disorders to those functioning as a risk allele for common complex diseases, with epistatic and environmental factors influencing clinical penetrance to a variable extent depending on the context (see the figure). Very rare highly penetrant mutations of mtDNA are the primary cause of inherited multi-system mitochondrial disorders. Examples include mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes (MELAS), which is usually caused by the m.3243A>G mutation; myoclonic ataxia and epilepsy with ragged red fibres (MERRF), which is usually caused by the m.8344A>G mutation; and Kearns–Sayre syndrome, which is usually caused by a large 4,977-bp deletion of mtDNA (other examples are discussed in REF. 5). These mutations usually occur *de novo* within two or three human generations and, for deletions, tend to affect only one generation. They are therefore heteroplasmic. At the other extreme, common population polymorphisms have a small effect on the risk of developing common human diseases. The most compelling case is for Parkinson disease, and the results of several large studies implicate common mtDNA variants in the pathogenesis of this disease[106]. The same variants might also contribute to the risk of developing several human diseases that share common mitochondrial mechanisms[12]. These variants are often ancient (>10,000 years old) and are therefore homoplasmic. Some variants lie between these two extremes, and cause disease only when present in combination with other genetic or environmental factors. For example, for the m.11778A>G, m.3460A>G and m.14484T>C mtDNA mutations that cause Leber hereditary optic neuropathy (LHON), the mtDNA background haplotype, nuclear genetic background and environmental factors, such as cigarette smoking, all influence the clinical penetrance of the disorder in pedigrees[117,118]. These mutations are usually, but not exclusively[119], homoplasmic.



evidence from the largest study to date characterizing mtDNA mutations in human cancer does not support this concept[42]. The unusual pattern of mtDNA mutations observed in 1,675 tumours showed a strong bias for the leading strand of mtDNA replication, similar to the pattern of homoplasmic mutations seen in mtDNA at the level of the human population. This suggested that the mutations were neither caused by nor contributed to the malignant process[42], but merely reflected accumulating errors during the replication of mtDNA molecules within the intensely proliferating malignant cells.

Recently, next-generation sequencing methods have been used to simultaneously identify and quantify mtDNA mutations[43] (FIG. 2). However, even short-read sequencing technologies have a high intrinsic error rate when applied at the very high depth required to detect and measure low-level heteroplasmy[44]. Long-read technologies have even greater intrinsic error rates. Using either short- or long-read sequencing technologies, the data can be contaminated with nuclear mitochondrial DNA sequences (NUMTs)[45,46] or DNA base damage, which can appear as real sequence variants. There is evidence of individual-level variation in NUMTs[46], so when using short-read next-generation sequencing data, these polymorphic NUMTs can be very difficult to filter out *in silico*. Long-amplicon methods and capture-enrichment methods can be used but are unable to completely eliminate NUMTs from the data[47]. Density-gradient centrifugation methods have been used to enrich for mtDNA, but they leave surprisingly high levels of nuclear DNA contamination[48]. Despite the technical difficulties using these approaches, mtDNA heteroplasmy has been found in almost every healthy individual who has been studied, albeit at a very low level[49]. These heteroplasmic variants can also be passed down the maternal line, raising the possibility that some of the ostensibly somatic mutations measured in somatic tissues (or tumours) late in life are actually low-level heteroplasmies that have been inherited down the maternal line. Studies in a mouse model have shown that low-level heteroplasmy can interact with *de novo* somatic mutations to exacerbate ageing and neurological phenotypes[10].

Lifetime changes in heteroplasmy

All mtDNA mutations originate within a single molecule within a single organelle, within a single cell. Given that each cell contains thousands of mtDNA molecules, a single molecule containing a mutation lies well below the mutation-detection threshold for most current technologies. Here, we consider possible mechanisms by which a mutation affecting one molecule can eventually reach a high frequency within a population of cells and thereby be detected using current technologies (FIG. 3).

Dividing cells. When a cell divides, the mtDNA molecules pass to the two daughter cells. If there is a heteroplasmic variant present then, by chance, each daughter cell can receive a different proportion of the mutated molecules. This process is called vegetative segregation and provides an explanation for observed changes in the level of heteroplasmy in dividing cells both *in vivo* and *in vitro*. For neutral genetic variants — that is, those variants that are not thought to directly affect fitness — the overall mutation level will remain constant in an infinite population; but, in a finite pool of dividing cells, the mutation level can drift up or down through random genetic drift. However, if a mutation has a severe effect on cell function and thus prevents further replication of

Nuclear mitochondrial DNA sequences
(NUMTs). Pseudogene sequences in nuclear DNA that are derived from mitochondrial DNA.

Vegetative segregation
Changes in heteroplasmy that occur owing to the unequal partitioning of different mitochondrial genotypes during cell division.

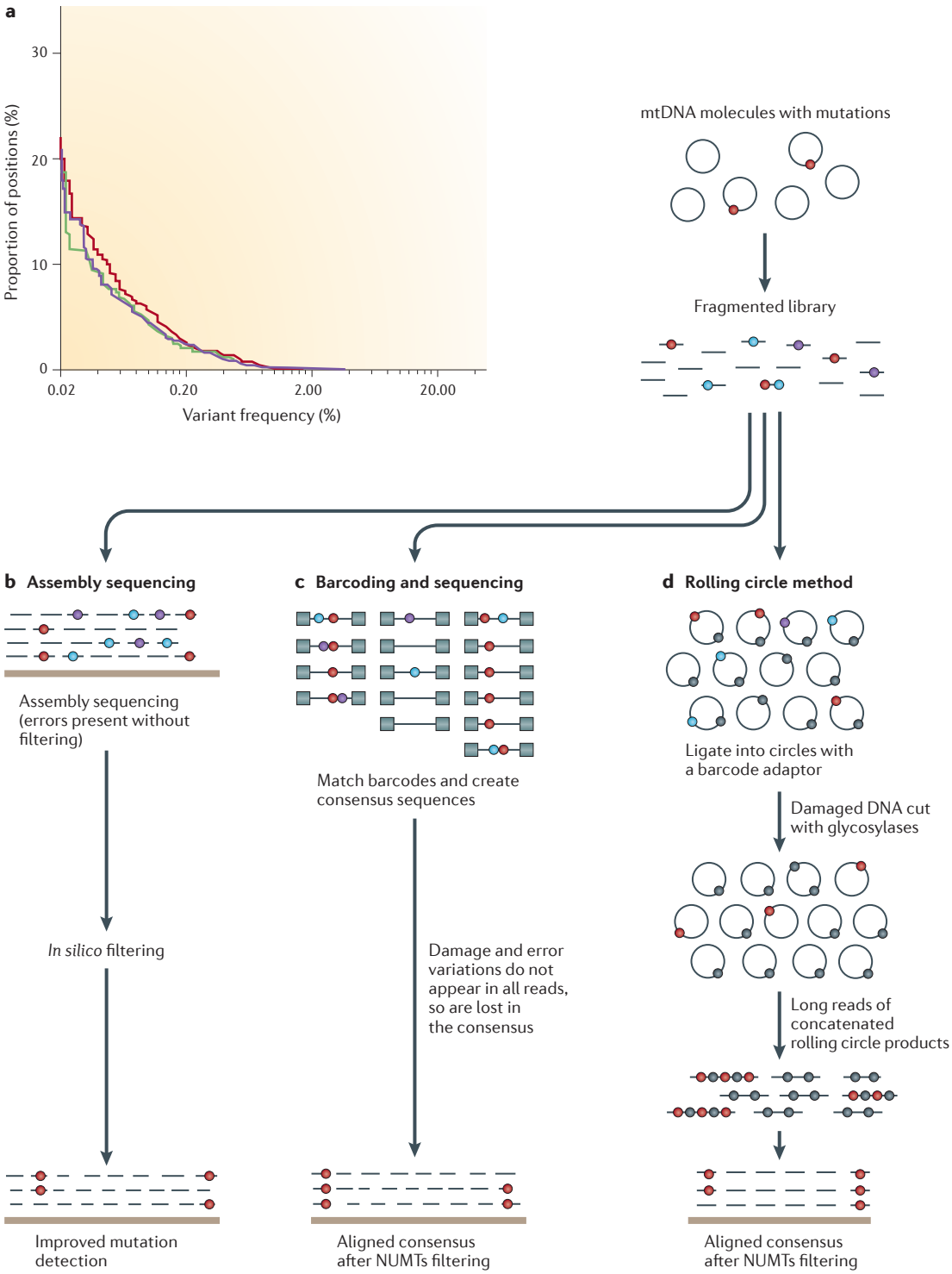


Figure 2 | **New technologies to analyse mtDNA.** **a** | The frequency distribution (*y* axis) of mitochondrial DNA (mtDNA) variants detected at different levels of heteroplasmy (*x* axis) in a typical deep-sequencing experiment. The mtDNA variants present *in vivo* and artefacts generated by the sequencing can be indistinguishable when a single subject is sequenced. The rarer variants are more likely to be sequencing errors or nuclear mitochondrial DNA sequences (NUMTs), but they could also be rare variants present in the original DNA sample. **b** | Appropriate computational filtering of the data is critical to identifying true mtDNA variants from the various artefacts in the data[47]. Applying stringent standards, heteroplasmies as infrequent as 0.2% have been identified[49]. **c**,**d** | Other approaches include the use of a barcode-adapted library[120]; and rolling circle amplification of circularized and barcoded DNA fragments with enzymes such as φ29 (REF. 121). DNA that has been oxidized or otherwise damaged during extraction can also be eliminated by the addition of appropriate DNA glycosylases to cleave DNA containing damaged bases prior to amplification[122].

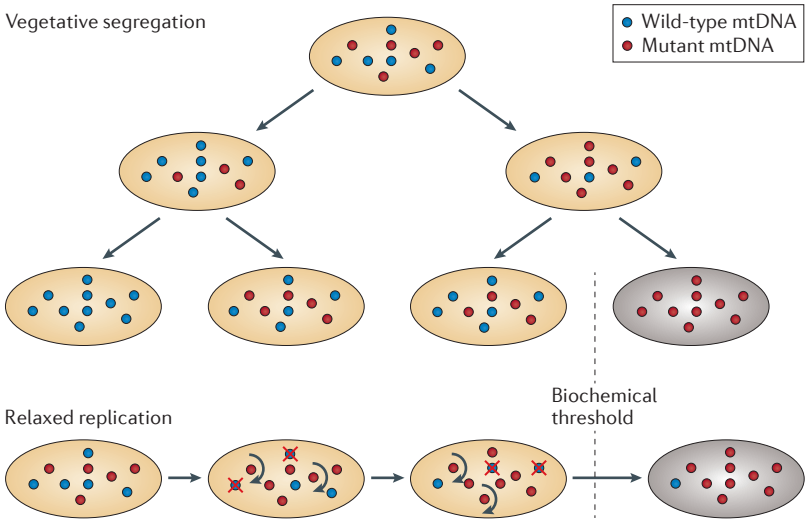


Figure 3 | **mtDNA heteroplasmy can change throughout the lifetime of an individual.** When cells divide, the proportion of mutated (red) and wild-type (blue) mitochondrial DNA (mtDNA) can differ in the daughter cells, leading to a shift in heteroplasmy levels (either up or down) over time through vegetative segregation. Furthermore, mtDNA is continuously turned over at all points in the cell cycle, even in non-dividing cells (crosses represent destruction, whereas curved arrows represent replication). Molecules are selected for replication at random, independently of the cell cycle, in a process known as relaxed replication. If one particular variant is copied more frequently than another, changes in the proportion of mutated and wild-type molecules occur over time, either up or down. Computational models predict that these two mechanisms will lead to changes in heteroplasmy in human cells throughout an individual's lifetime. The two mechanisms can occur simultaneously in some tissues and organs. Selection for or against a particular variant of mtDNA will influence the speed and direction of the heteroplasmic shift, and thus alter mutation levels over time.

the cell, then selection against the mutation will occur at the cellular level. This negative selection is thought to be the reason that levels of the common pathogenic mutation m.3243A>G in the gene mitochondrially encoded tRNA leucine 1 (UUA/G) (*MTTL1*) decrease exponentially in blood during life[50]. However, in cell culture studies, different mtDNA mutations can drift in either direction or the overall heteroplasmy level can remain stable[51]. It is not clear why the same mtDNA mutation can drift up, down or remain stable in different cell lines, but it is possible that the nuclear genes influence the segregation pattern[52].

Non-dividing (postmitotic) cells. Unlike the situation for nuclear DNA, no cellular mechanism exists to ensure that each mtDNA molecule is copied in equal number; furthermore, mtDNA is continuously being destroyed and replicated, even in non-dividing cells, through a process referred to as relaxed replication. In a heteroplasmic cell, it is possible that, by chance, a variant molecule might be replicated more frequently than the wild-type molecule, leading to a change in the overall heteroplasmy level within the cell. *In silico* statistical modelling studies have shown that, in a fixed population of cells, this random process can lead to significant changes in heteroplasmy levels in some cells; this provides one explanation for the apparent clonal expansion of specific mtDNA molecules during human life[53]. This random process could explain how a *de novo* mutation occurring in early life or if acquired from the oocyte leads to a biochemical defect in an older individual[54].

However, if there is a bias in favour of (or against) a particular variant molecule, this would influence the direction of the relaxed replication and accelerate the accumulation of one of the molecular variants. For example, mtDNA molecules with a pathogenic deletion are several kilobases smaller than wild-type mtDNA molecules. The smaller mtDNA molecules can replicate faster and are thus available for further replication more often. This could lead to a selective advantage favouring the smaller deleted molecules, explaining why deletions accumulate over time. This occurrence has been observed in human cell lines[55] and animal models. For example, studies in the nematode *Caenorhabditis briggsae* showed that mitochondria carrying mtDNA deletions had a replicative advantage and did preferentially accumulate in these animals, despite the eventual negative consequences of the mutations on mitochondrial function[56].

Studies in humans using massively parallel sequencing suggest that specific mtDNA variants with mutations in the displacement (D)-loop region preferentially accumulate in somatic cells over time[47,57]. Work with mice carrying artificially mixed mtDNA haplogroups shows similar results, in which a particular mtDNA genotype preferentially accumulates in specific tissues or organs. This phenomenon was first observed as a tissue-specific effect in mice with mixed mitochondria from the BALB/c and NZB mouse strains[58]. Other highly divergent mtDNA haplotypes have since been found to preferentially accumulate within tissues over time[59,60]. Studies in mice expressing an error-prone variant of mtDNA polymerase[38] found that the small origin of light strand replication[61] and specific segments of the D-loop region[62] are less prone to harbouring mutations than the other coding regions of the mitochondrial genome. Complementary *in vitro* replication assays of the small origin of light strand replication were able to demonstrate that mutations decreased replication efficiency, implying that these missing mutations were likely to have been copied less, creating the illusion of a reduced mutation rate in the mouse[61]. However, the analysis of these naturally occurring human mutations must be confirmed by additional studies, especially in light of the individual-by-individual variation in the number of copies of mitochondrial pseudogenes (NUMTs[45]) found within the nuclear DNA[46], which may confound short-read deep-coverage analyses.

Molecular selection is not the only means by which levels of a mutation can increase over time in non-dividing tissues. One of the defining pathological features of heteroplasmic mtDNA diseases is the proliferation of mitochondria within affected cells, over and above the normal mitochondrial turnover. In skeletal muscle cells, this occurs just below the plasma membrane, giving individual muscle fibres a 'ragged' appearance (leading to the presence of visible 'ragged red fibres' when using

Relaxed replication
DNA replication that occurs continuously throughout the cell cycle and is independent of nuclear division, as with mitochondrial DNA.

Replicative advantage
A situation in which one molecule is preferentially copied over another.

Haplogroups
Groups of similar haplotypes (groups of genes) that share several polymorphisms. Single-nucleotide variants acquired during human history have subdivided the mitochondrial DNA phylogeny into several major haplogroups, typically present at >5% in the population.

the Gömöri trichrome stain). This proliferation implies a compensatory mechanism, which can be measured at the DNA level, leading to the 'maintenance of wild type' hypothesis[53,63]. This hypothesis is based on the assumption that a cell contains an optimal amount of wild-type mtDNA; if one of the molecules is mutated then the cell responds by replicating the entire mtDNA content to redress the balance. Making the reasonable assumption that the nuclear signal initiating the replication cannot distinguish mutated from wild-type mtDNA, this would also result in the proliferation of all mtDNA within the cell. In the short term, this will ensure that there is an adequate amount of wild-type mtDNA, but the number of mutant mtDNA molecules will also increase and might eventually exceed 50%. After this point, the same corrective mechanism will lead to the preferential replication of mutated molecules (at the expense of wild-type ones), leading to wild-type mtDNA depletion. As the proliferation will preferentially occur in cells containing mutant mtDNA, the net effect for the tissue will be to increase the overall level of the mutation over time. This increase has been observed in patients with mtDNA diseases[64].

Although presented as separate mechanisms, vegetative segregation and relaxed replication are not mutually exclusive, and both could influence the level of heteroplasmy *in vivo* — explaining how mtDNA mutations populate a tumour by 'hitch-hiking' on a clonally expanding malignant cell line[42,65] and also within the developing germ line[66]. Finally, there is emerging evidence that the segregation of mtDNA mutations *in vivo* might be under tissue-specific nuclear genetic control[52,67,68], adding further complexity to the situation.

The mtDNA genetic bottleneck hypothesis. For more than 30 years, it has been known that the level of heteroplasmy can change markedly during germline transmission. Early work on Holstein cows transmitting heteroplasmic polymorphic variants showed that a *de novo* heteroplasmic mutation could become fixed and homoplasmic within two or three generations[69,70]. This led to the mitochondrial (or mtDNA) genetic bottleneck hypothesis, which postulates that only a small proportion of the total number of mitochondrial genomes are passed on from mother to offspring.

The mtDNA bottleneck theory, and early modelling-based studies[71–73], was based on analogy with population genetics[74], in which the variance among descendants is inversely related to the effective population size (N_e). N_e is a mathematical representation of a genetic bottleneck and, when small, dramatically increases the variation in individual allele frequencies observed within daughter populations. In simple terms, this resembles a statistical sampling effect, in which small random samples taken from a larger population will contain varying proportions of a particular allele, and the difference between each sample will be greater if the size of the sample is smaller (that is, a narrow bottleneck). If such a bottleneck were present in the germline, this could lead to differences in the level of heteroplasmy between offspring and, thus, to differences in the level of heteroplasmy transmitted from one mother to her different offspring.

Studies of pathogenic heteroplasmic mtDNA mutations in humans show similar, albeit less dramatic, shifts in the level of heteroplasmy, again supporting the presence of a germline genetic bottleneck in humans[75]. In this context, differences in the levels of transmitted mutations provide an explanation for differences in the severity of the clinical phenotype in families with heteroplasmic mtDNA diseases[76,77].

Studying the proposed mtDNA bottleneck directly in humans is difficult, but studies in mice have provided supportive evidence[66,78,79]. The precise mechanism has yet to be defined (FIG. 4), although several mechanisms have been proposed on the basis of the existing data. Two laboratories have measured a dramatic reduction in the amount of mtDNA present in the first discernible germ cells (the primordial germ cells) that appear shortly after implantation of the female embryo[66,78]. Moreover, *in silico* statistical modelling studies have shown that this restriction in mtDNA number is sufficient to cause the changes in heteroplasmy observed in mice transmitting heteroplasmic mtDNA variants[10]. Crossing heteroplasmic mice with mice expressing fluorescent germ-cell markers allows the isolation of individual cells at different stages in development, and this approach showed that the majority of heteroplasmy variance generated during germ-cell proliferation and migration, and the level within individual oocytes, is largely determined before birth[80]. Although the random partitioning of a limited number of mtDNA molecules could lead to the bottleneck effect[10], other mechanisms might come into play, including the preferential replication of a subpopulation of molecules either before or after birth[40], or some form of compartmentalization of mtDNA molecules[79], accelerating the rate of genetic drift. Direct experimental evidence in favour of one or all of these mechanisms is currently lacking, and a combination of these might actually occur *in vivo*[81].

Molecular selection during germline transmission. Initial observations in humans[82] and mice[83] suggested that the primary mechanism underpinning the changing levels of heteroplasmy was random genetic drift, and data from the analysis of human oocytes[84] and mouse ovaries[83] are consistent with this observation. From the perspective of the mitochondrion, the mutations seem to be effectively neutral variants and are segregating without a selective bias. Such a mechanism would suggest that the fitness effects of mtDNA mutations are hidden below a certain threshold, and selection would act only at the level of the whole organism, in the individuals who have sufficient levels of the mutation to show detrimental physiological effects. Indeed, selection in a mouse with a structure-disrupting mutation in the mitochondrial tRNA methionine gene *MTTM* seems to behave in such a manner[80]. In the developing female germ line, the behaviour of the mutation is consistent with neutral segregation; however, pups born to these mothers show a strong bias against very high levels of the mutation, implying the existence of a post-fertilization selection mechanism to limit offspring with excess levels of the mutation[80]. Pedigree analysis suggests that negative selection

Bottleneck
A restriction in the number of molecules (for example, mitochondrial DNA) or groups of molecules arranged in segregating units.

occurs during the process of transmission to the next generation in mice[85,86], humans[87] and even *Drosophila melanogaster*[88,89]. But what do these observations mean for our understanding of the origins and fixation of mtDNA mutations in humans, at the level of the cell, organism and population?

Bearing in mind important differences in development between the two species, work in *D. melanogaster* provides an intriguing insight into how the molecular selection might occur in humans. Researchers used a temperature-sensitive heteroplasmic mutation in the gene mitochondrially encoded cytochrome *c* oxidase I (*mtco1*)[88] that is lethal to the flies after 4 days at 29 °C in the homoplasmic state. Heteroplasmic flies are not killed at the high temperature[88] but, 3 days into the heat exposure, the researchers observed a strong bias against transmission of the mutation to the offspring, suggesting that the molecular selection occurred at the late germarium stage of oocyte development of the fly[89]. The timing of the initiation of selection correlated with an impairment of mtDNA replication in the developing oocytes. But how is *mtco1* — a core component of the respiratory chain — linked to the physically distinct act of mtDNA replication? The answer might lie in an underappreciated aspect of mitochondrial biology. As mentioned earlier, nearly all mitochondrial proteins, including all replication-associated proteins, are imported from the nucleus. It is well established that the import of most mitochondrial proteins is dependent on an active membrane potential[90], which is a direct consequence of a well-functioning set of proteins expressed from the mtDNA. In this way, a cell can maintain control of the mitochondria by supplying gene products to only those with a robust mitochondrial membrane potential — which is indicative of an active respiratory chain and fit mtDNA.

Studies in mammalian systems are required to confirm whether this respiratory chain–replication connection is conserved, but the underlying mechanism is probably linked to the mtDNA genetic bottleneck. In the early mouse embryo, until about 7.5 days into development, mtDNA does not undergo gainful replication[91]; however, after this point, it rapidly increases in cellular copy number, especially in the developing female germ line[66,78,79]. Thus, the point at which replication is re-initiated occurs when the total number of genomes in the cell is at its lowest. At this point, single mtDNA molecules harbouring a mutation will have their greatest influence on the cell, potentially compromising mtDNA replication and thus impairing their transmission. However, evidence indicates that this mechanism is far from perfect. Mutations that lack immediate obvious physiological consequences, or that arise in regions of the mitochondrial genomes that are less sensitive to sequence variation, can clearly pass through this filter[92]. Additionally, the timing of this event in early embryogenesis means that mtDNA mutations that arise

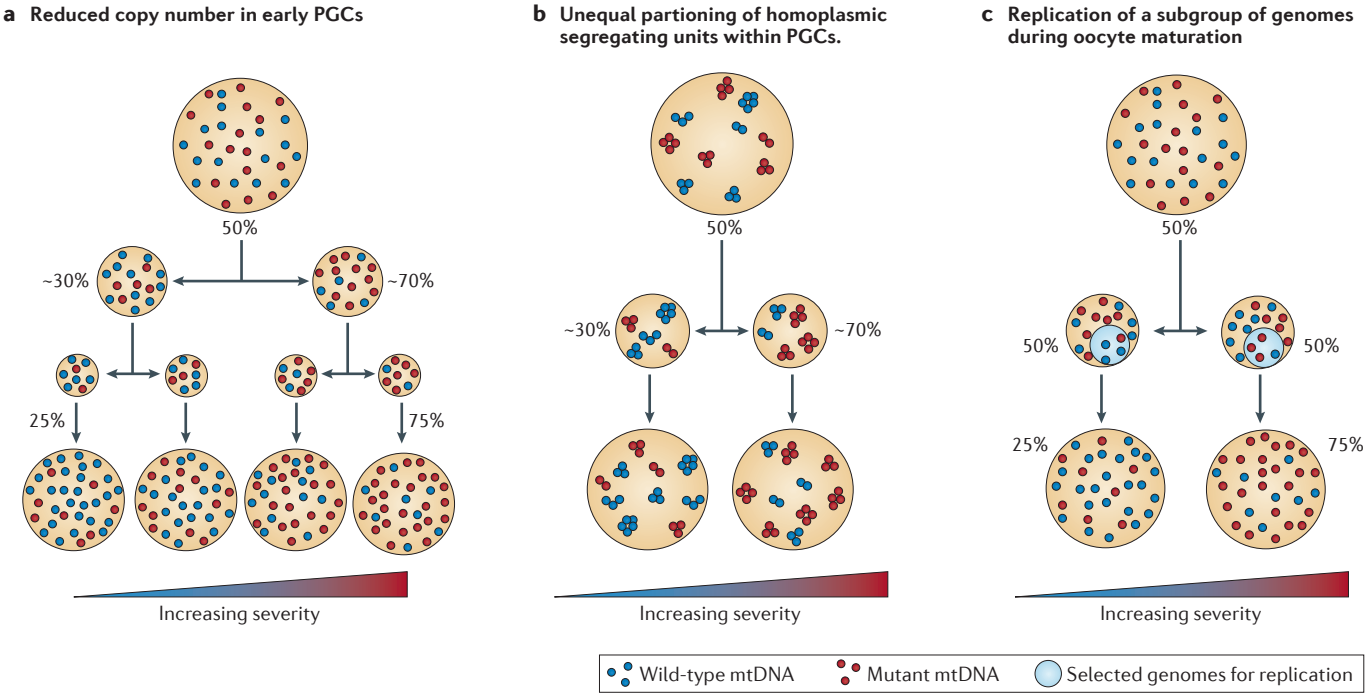


Figure 4 | **Models for the mtDNA genetic bottleneck. a** | The rate of vegetative segregation is faster if the mitochondrial DNA (mtDNA) content of the cells is smaller. Unequal partitioning (segregation) of mutated (red) and wild-type (blue) genotypes will lead to shifts in heteroplasmy in oocytes as a consequence of the reduction in mtDNA copy number immediately prior to expansion of the primordial germ cell (PGC) population, as observed in mice[66,78]. **b** | If mtDNA molecules are packaged into homoplasmic nucleoids, or simply kept separate within discrete homoplasmic mitochondria, this reduces the number of segregating units and accelerates the rate of drift, as proposed in REFS 79,123. **c** | The replication of a subpopulation of mitochondrial genomes would also lead to a bottleneck effect. There is some evidence that this might occur during postnatal oocyte maturation[78]. Adapted with permission from REF. 124, Elsevier.

subsequently will not be exposed to the same sensitive selection process because cells containing thousands of mtDNA molecules are unlikely to be affected by single mutated molecules. This will lead to the accumulation of mutations in somatic tissues during embryonic development, and postmitotic tissues will bear the brunt of the damage caused by these mutations. This is partly because relaxed replication, on its own, is less effective than vegetative segregation at removing mutations from an organ or tissue; it might also be because the affected cells strive to maintain wild-type mtDNA levels according to the maintenance of wild-type hypothesis described above.

Finally, it has been suggested that selection against deleterious mtDNA mutations might be linked to the exponential loss of primary oocytes (follicular atresia) that begins in females *in utero* and persists to the menopause. This could provide an explanation for the finite reproductive capacity of women, and in some instances, lead to premature ovarian failure[93]. A high mtDNA mutation burden could also contribute to the high frequency of early pregnancy miscarriages in humans[94]. Although these are both intriguing possibilities, definitive evidence is currently lacking.

mtDNA in human populations

African origin. With the expansion and migration of the human population, new mutations arose (and continue to arise) within different maternal lineages. Ultimately, these mutations must have arisen from within a germ cell and been heteroplasmic, then have survived the mtDNA genetic bottleneck over several generations and eventually have become fixed (homoplasmic) within a maternal line. The uniparental inheritance of mtDNA and associated lack of intermolecular recombination mean that these variants have remained restricted to specific ethnic groups and have been used by population geneticists to define the migration and colonization of the planet, supporting the 'out of Africa' hypothesis, which proposes that the human mtDNA had its origins in Africa (FIG. 5).

Further large-scale analysis of the human population structure using mtDNA is now being undertaken[95], and increasingly sophisticated models have been developed that, for example, incorporate purifying selection[96]. Criticism of the quality of some of the available mtDNA sequence submissions[97,98] has led to a higher degree of care being taken to ensure high-quality data are obtained, which is crucially important because sequencing artefacts can falsely suggest that expected biological phenomena, such as selection, intermolecular recombination or a biased mutation rate in different regions of the genome, have occurred. With this added care, these large-scale studies have confirmed the traditional view of the relative spread of human mitochondria throughout the world from an African origin ~100,000 years ago. These studies have also provided information on the basic transmission of mtDNA, such as the lack of the occurrence of recombination, a high rate of substitution in humans (leading to many recent mutations specific to small families) and a large number of back-mutations (homoplasies) throughout the human population[95].

Homoplasies
Mutations of the same nucleotide at the same genomic position arising by two (or more) independent mutations.

Selection in mtDNA evolution. Some common polymorphic mtDNA variants are predicted to have functional consequences, raising the possibility that selection might have been important during human mtDNA evolution. Evidence in support of this possibility comes from animal studies, such as those in the brook char, in which specific mtDNA variants seem to have evolved in response to local climatic change and have thereby also shaped the associated nuclear genome[99]. Proving that such an effect occurred in the human population has been extremely difficult. As a haploid genome, mtDNA associations are particularly susceptible to population stratification effects, leading to false-positive results from association studies. This problem has complicated the search for evidence of selection for or against specific mtDNA genotypes in response to climate change in the human population[100,101], and it also provides an explanation for the many reported genetic associations between mtDNA polymorphisms and common late-onset human diseases that have failed to be replicated in subsequent studies[102].

By contrast, several studies have shown that different common polymorphic variants of mtDNA can affect mitochondrial function *in vitro*. This work has largely been based on cybrid cell lines, in which mtDNA has been exchanged on a common nuclear genetic background. In this context, mtDNA and mtRNA levels, mitochondrial protein synthesis, cytochrome oxidase levels and activity, normalized oxygen consumption, mitochondrial inner-membrane potential and growth capacity are different in cybrids from the haplogroup H (found in ~40% of Europeans) when compared with those from the haplogroup Uk (found in ~25% of Europeans)[14]. Similarly, in mice, mtDNA molecules from different laboratory strains are associated with the generation of different levels of reactive oxygen species and a compensatory increase in mtDNA levels to normalize levels of ATP synthesis[103].

Several studies carried out over the past 10 years have clarified the association between polymorphic variants of mtDNA and common human diseases, and work on human sepsis provides one example[104]. Despite the current availability of antibiotics and intensive care therapy, humans who develop multi-organ failure as a consequence of systemic infection have a 50% mortality risk. Several studies have now shown that common mtDNA variants influence the likelihood of survival following acute sepsis[104,105], which raises the possibility that exposure to infectious disease in the past led to the selection of particular mtDNA variants that conferred a survival advantage but conceivably predispose to late-onset degenerative diseases (antagonistic pleiotropy)[106].

Large-scale genetic association studies, including consistent replication studies, have reported an association between Parkinson disease and common mtDNA variants[106,107], and an analysis of several different common late-onset human diseases identified mtDNA variants predisposing to several related disorders as well as those conferring protection against others[12]. Alleles that increased disease risk were more common than those that decreased disease risk, suggesting that deleterious,

novel sub-haplogroup variants have not yet been removed from the population through natural selection. These high-risk alleles are particularly concentrated in the non-coding mtDNA D loop, which is implicated in mtDNA replication and gene expression (BOX 1), suggesting that subtle differences in mtDNA replication or transcription contribute to the pathogenesis of common age-related diseases.

Conclusions and perspectives

The emergence of large-scale mtDNA datasets at both the single-cell and the human-population level provides insight into the origin, fixation and consequences of mtDNA mutations. At one extreme, very low-level mtDNA heteroplasmy seems to be an almost universal finding in healthy humans, and the transmission of such heteroplasmic variants down the maternal line provides a source of potentially deleterious mutations that could affect individual cells, the organism or the population. The mtDNA genetic bottleneck leads to the rapid segregation of heteroplasmic variants to homoplasmy within the developing germ line. For highly deleterious mutations, this rapid segregation will purge the germ line of mutations through selection at the organellar, cellular or organismal level. However, mildly deleterious genetic variants might not be selected against at the cellular level or during germline transmission to the next generation, allowing these variants to accumulate within the population over time through random genetic drift in the absence of any selection pressures. Although mtDNA mutations may confer some protection against acute diseases in the short term, in the long term an increased burden of these mtDNA mutations have been associated with cellular dysfunction and senescence. Given the shift in the demographics of the human population over the past 20 years towards longer life expectancies, the contribution of mtDNA to the risk of developing late-onset multifactorial degenerative diseases is likely to increase.

Although these mtDNA variants could contribute to disease susceptibility of the human host, as an organelle the mitochondrion will continue to evolve, supported by a parallel-evolving nuclear genome. Embedded within a complex multicellular organism that can adapt to complex climates and dietary restrictions, the mitochondrion is somewhat protected from our dramatically changing ecosystems. We may reconsider the notion of the nuclear–mitochondrial symbiosis, and ask the

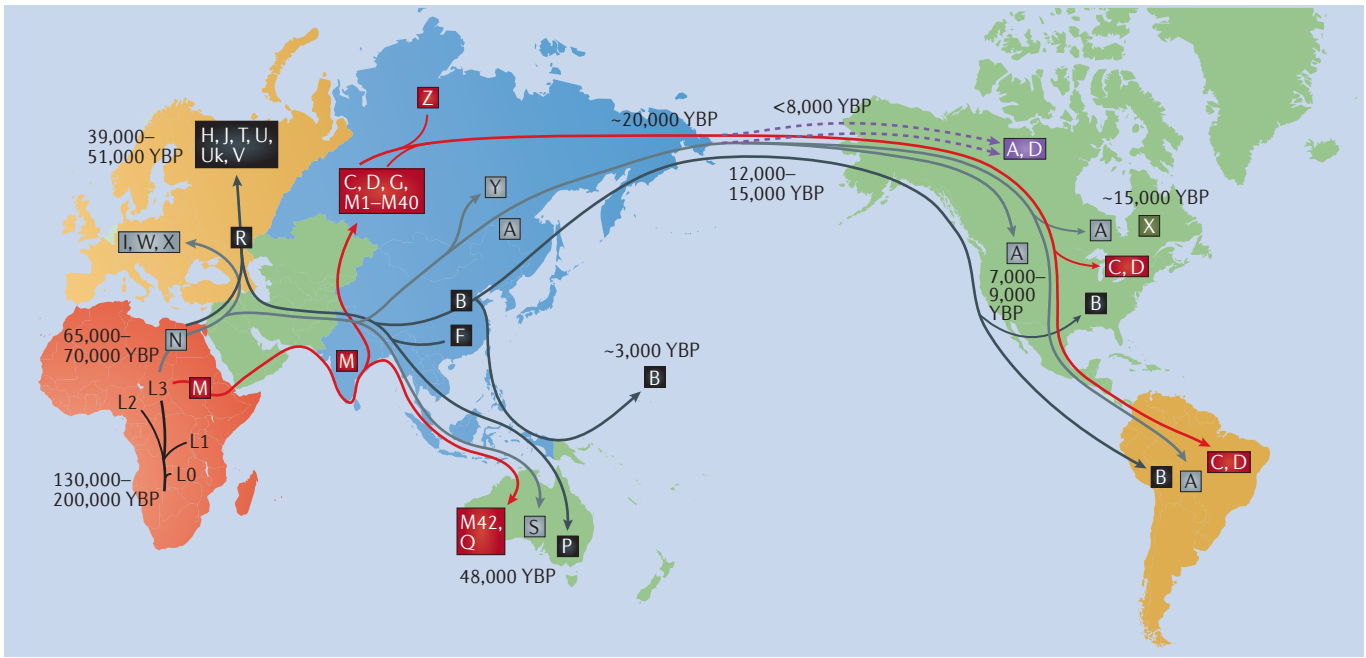


Figure 5 | **Human population migrations and major mtDNA haplogroups.** As mitochondrial DNA (mtDNA) is uniparentally inherited, it undergoes negligible recombination at the population level, and mutations acquired over time have subdivided the human population into several discrete haplogroups. The major haplogroups arose 40,000–150,000 years before present (YBP) and have defined different human populations as they migrated out of Africa and populated the globe. The African root was the source of four lineages specific for sub-Saharan Africa: L0, L1, L2 and L3 (130,000–200,000 YBP). Two more haplogroups, M and N, arose from the African haplogroup L3 65,000–70,000 YBP to populate the rest of the world[125]. As humans migrated, haplogroup N was directed to Eurasia and haplogroup M lineages moved to Asia, giving rise to the haplogroups A, B, C, D, G and F. In Europe, haplogroup N led to haplogroup R, which is the root of the European haplogroups H, J, T, U and V, which emerged 39,000–51,000 YBP[126]. Haplogroups S, P, and Q are found in Australasia and were formed ~48,000 YBP, and haplogroups A, B, C and D arose <20,000 YBP and populated East Asia and the Americas. These low-resolution single-nucleotide variant studies have been superseded by massive whole-mtDNA-genome-sequencing studies, which have identified many different sub-haplogroups that define the contemporary mtDNA phylogenetic tree. For example, haplogroup H, the most common in Europe, is comprised of almost 90 different sub-haplogroups[127]. Adapted with permission from MITOMAP (original authors Lott, M. T. *et al.*; http://creativecommons.org/licenses/by/3.0/).

question: is our two-genome system working in concert for mutual benefit or is the mtDNA driving nuclear genome evolution? This is important because it may be maladaptive for humans that now live into their eighth or ninth decade, so understanding how we might influence this may be one way of enhancing the length and quality of our lifespan. Mammalian and invertebrate model systems are now available to tackle these issues at the cellular, and soon the sub-cellular, level. High-resolution live-cell imaging studies will be useful in clarifying the mechanics of mtDNA compartmentalization and segregation at critical time points, and combining ultra-deep mtDNA sequencing and transcriptomics will be necessary to understand the effect of low-frequency variants on the specific cell types. Together, these approaches are important for characterizing the mechanisms of selection, which may or may not be directly related to oxidative phosphorylation. Much of this work is likely to be done in animal models, but proving that the same mechanisms occur in humans will be crucial if there is to be translation of these findings into the clinic. Given recent concerns about the theoretical long-term consequences of the mitochondrial–nuclear mismatch caused by mitochondrial transfer in human embryos[108,109], this work will have an effect on the new treatments aimed at preventing monogenic mtDNA diseases that were recently approved by the UK Parliament[110].

1. Wallace, D. C. *et al.* Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy. *Science* **242**, 1427–1430 (1988).
Describes the first-identified human pathogenic mtDNA mutation, which was a homoplasmic point mutation in *MTND4*. It transpires that this mutation (m.11778A > G) is actually the most common pathogenic mtDNA mutation encountered in clinical practice.
2. Holt, I., Harding, A. E. & Morgan-Hughes, J. A. Deletion of muscle mitochondrial DNA in patients with mitochondrial myopathies. *Nature* **331**, 717–719 (1988).
3. Gorman, G. S. *et al.* Prevalence of nuclear and mitochondrial DNA mutations related to adult mitochondrial disease. *Ann. Neurol.* **77**, 753–759 (2015).
4. Chinnery, P. F. *et al.* Epidemiology of pathogenic mitochondrial DNA mutations. *Ann. Neurol.* **48**, 188–193 (2000).
The first epidemiological study of pathogenic mtDNA mutations, which shows that mtDNA diseases are not exceptionally rare, as previously thought.
5. Di Mauro, S., Schon, E. A., Carelli, V. & Hirano, M. The clinical maze of mitochondrial neurology. *Nat. Rev. Neurol.* **9**, 429–444 (2013).
6. Vafai, S. B. & Mootha, V. K. Mitochondrial disorders as windows into an ancient organelle. *Nature* **491**, 374–383 (2012).
7. Wallace, D. C. Colloquium paper: bioenergetics, the origins of complexity, and the ascent of man. *Proc. Natl Acad. Sci. USA* **107** (Suppl. 2), 8947–8953 (2011).
8. Pinto, M. & Moraes, C. T. Mechanisms linking mtDNA damage and aging. *Free Radic. Biol. Med.* **85**, 250–258 (2015).
9. Pinto, M. & Moraes, C. T. Mitochondrial genome changes and neurodegenerative diseases. *Biochim. Biophys. Acta* **1842**, 1198–1207 (2014).
10. Ross, J. M. *et al.* Germline mitochondrial DNA mutations aggravate ageing and can impair brain development. *Nature* **501**, 412–415 (2013).
11. Keogh, M. & Chinnery, P. F. Hereditary mtDNA heteroplasmy: a baseline for aging? *Cell Metab.* **18**, 463–464 (2013).
12. Hudson, G., Gomez-Duran, A., Wilson, I. J. & Chinnery, P. F. Recent mitochondrial DNA mutations increase the risk of developing common late-onset human diseases. *PLoS Genet.* **10**, e1004369 (2014).
13. Wallace, D. C. & Chalkia, D. Mitochondrial DNA genetics and the heteroplasmy conundrum in evolution and disease. *Cold Spring Harb. Perspect. Med.* **3**, a021220 (2013).
14. Gomez-Duran, A. *et al.* Unmasking the causes of multifactorial disorders: OXPHOS differences between mitochondrial haplogroups. *Hum. Mol. Genet.* **19**, 3343–3353 (2010).
15. Mereschkowsky, K. Theorie der zwei Plasmaarten als Grundlage der Symbiogenesis, einer neuen Lehre von der Ent-stehung der Organismen. *Biol. Centralbl.* **30**, 353–367 (1910) (in German).
16. Pagliarini, D. J. *et al.* A mitochondrial protein compendium elucidates complex I disease biology. *Cell* **134**, 112–123 (2008).
17. Vafai, S. B. & Mootha, V. K. Medicine. A common pathway for a rare disease? *Science* **342**, 1453–1454 (2013).
18. Anderson, S. *et al.* Sequence and organization of the human mitochondrial genome. *Nature* **290**, 457–465 (1981).
Provides the first human mitochondrial DNA sequence, subsequently referred to as the Cambridge reference sequence (CRS). The published sequence was derived predominantly from a human placenta from a person belonging to haplogroup H, but difficult regions were derived from other sources, including the bovine sequence.
19. Andrews, R. M. *et al.* Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet.* **23**, 147 (1999).
Sequence errors in the original CRS (reference 18) were corrected in this publication, resulting in a new sequence referred to as the revised CRS (rCRS).
20. Calvo, S. E. & Mootha, V. K. The mitochondrial proteome and human disease. *Annu. Rev. Genom. Hum. Genet.* **11**, 25–44 (2010).
21. Kukat, C. *et al.* Super-resolution microscopy reveals that mammalian mitochondrial nucleoids have a uniform size and frequently contain a single copy of mtDNA. *Proc. Natl Acad. Sci. USA* **108**, 13534–13539 (2011).
22. Elson, J. L. *et al.* Analysis of European mtDNAs for recombination. *Am. J. Hum. Genet.* **68**, 145–153 (2001).
23. Schon, E. A., DiMauro, S. & Hirano, M. Human mitochondrial DNA: roles of inherited and somatic mutations. *Nat. Rev. Genet.* **13**, 878–890 (2012).
24. Chinnery, P. F. *et al.* Risk of developing a mitochondrial DNA deletion disorder. *Lancet* **364**, 592–596 (2004).
25. McFarland, R. *et al.* Multiple neonatal deaths due to a homoplasmic mitochondrial DNA mutation. *Nat. Genet.* **30**, 145–146 (2002).
26. Tiranti, V. *et al.* A novel frameshift mutation of the mtDNA *COIII* gene leads to impaired assembly of cytochrome *c* oxidase in a patient affected by Leigh-like syndrome. *Hum. Mol. Genet.* **9**, 2733–2742 (2000).
27. Limongelli, A. *et al.* Variable penetrance of a familial progressive necrotising encephalopathy due to a novel $tRNA^{Ile}$ homoplasmic mutation in the mitochondrial genome. *J. Med. Genet.* **41**, 342–349 (2004).
28. King, M. P. & Attardi, G. Human cells lacking mtDNA: repopulation with exogenous mitochondria by complementation. *Science* **246**, 500–503 (1989).
29. Boulet, L., Karpati, G. & Shoubridge, E. A. Distribution and threshold expression of the $tRNA^{Lys}$ mutation in skeletal muscle of patients with myoclonic epilepsy and ragged-red fibers (MERRF). *Am. J. Hum. Genet.* **51**, 1187–1200 (1992).
30. Carelli, V., Giordano, C. & d'Amati, G. Pathogenic expression of homoplasmic mtDNA mutations needs a complex nuclear–mitochondrial interaction. *Trends Genet.* **19**, 257–262 (2003).
31. Elliott, H. R., Samuels, D. C., Eden, J. A., Relton, C. L. & Chinnery, P. F. Pathogenic mitochondrial DNA mutations are common in the general population. *Am. J. Hum. Genet.* **83**, 254–260 (2008).
32. Cortopassi, G. A. & Arnheim, N. Detection of a specific mitochondrial DNA deletion in tissues of older humans. *Nucleic Acids Res.* **18**, 6927–6933 (1990).
33. Brierley, E. J., Johnson, M. A., Lightowlers, R. N., James, O. F. & Turnbull, D. M. Role of mitochondrial DNA mutations in human aging: implications for the central nervous system and muscle. *Ann. Neurol.* **43**, 217–223 (1998).
34. Corral-Debrinski, M., Shoffner, J. M., Lott, M. T. & Wallace, D. C. Association of mitochondrial DNA damage with aging and coronary atherosclerotic heart disease. *Mutat. Res.* **275**, 169–180 (1992).
One of the early papers describing the presence of somatic mtDNA mutations in aged individuals and the association of these mutations with age-related diseases.
35. Corral-Debrinski, M. *et al.* Marked changes in mitochondrial DNA deletion levels in Alzheimer brains. *Genomics* **23**, 471–476 (1994).
36. Bender, A. *et al.* High levels of mitochondrial DNA deletions in substantia nigra neurons in aging and Parkinson disease. *Nat. Genet.* **38**, 515–517 (2006).
37. Kraytsberg, Y. *et al.* Mitochondrial DNA deletions are abundant and cause functional impairment in aged human substantia nigra neurons. *Nat. Genet.* **38**, 518–520 (2006).
38. Trifunovic, A. *et al.* Premature ageing in mice expressing defective mitochondrial DNA polymerase. *Nature* **429**, 417–423 (2004).
39. Kujoth, G. C. *et al.* Mitochondrial DNA mutations, oxidative stress, and apoptosis in mammalian aging. *Science* **309**, 481–484 (2005).
References 38 and 39 provides the first descriptions of a mouse with a defective mtDNA polymerase. They show that mtDNA mutations accumulated in somatic tissues over time and were accompanied by a premature ageing phenotype.
40. Polyak, K. *et al.* Somatic mutations of the mitochondrial genome in human colorectal tumours. *Nat. Genet.* **20**, 291–293 (1998).
A key paper describing clonally expanded mtDNA mutations in human tumours.
41. He, Y. *et al.* Heteroplasmic mitochondrial DNA mutations in normal and tumour cells. *Nature* **464**, 610–614 (2010).
42. Ju, Y. S. *et al.* Origins and functional consequences of somatic mitochondrial DNA mutations in human cancer. *eLife* **3**, e02935 (2014).
43. Payne, B. A. *et al.* Mitochondrial aging is accelerated by anti-retroviral therapy through the clonal expansion of mtDNA mutations. *Nat. Genet.* **43**, 806–810 (2011).
44. Payne, B. A., Gardner, K., Coxhead, J. & Chinnery, P. F. Deep resequencing of mitochondrial DNA. *Methods Mol. Biol.* **1264**, 59–66 (2015).
45. Hazkani-Covo, E., Zeller, R. M. & Martin, W. Molecular poltergeists: mitochondrial DNA copies (numts) in sequenced nuclear genomes. *PLoS Genet.* **6**, e1000834 (2010).
46. Dayama, G., Emery, S. B., Kidd, J. M. & Mills, R. E. The genomic landscape of polymorphic human nuclear mitochondrial insertions. *Nucleic Acids Res.* **42**, 12640–12649 (2014).
47. Li, M. *et al.* Detecting heteroplasmy from high-throughput sequencing of complete human mitochondrial DNA genomes. *Am. J. Hum. Genet.* **87**, 237–249 (2010).
48. Ameur, A. *et al.* Ultra-deep sequencing of mouse mitochondrial DNA: mutational patterns and their origins. *PLoS Genet.* **7**, e1002028 (2011).
49. Payne, B. A. *et al.* Universal heteroplasmy of human mitochondrial DNA. *Hum. Mol. Genet.* **22**, 384–390 (2013).

50. Rajasimha, H. K., Chinnery, P. F. & Samuels, D. C. Selection against pathogenic mtDNA mutations in a stem cell population leads to the loss of the 3243A→G mutation in blood. *Am. J. Hum. Genet.* **82**, 333–343 (2008).
51. Lehtinen, S. K. *et al.* Genotypic stability, segregation and selection in heteroplasmic human cell lines containing np 3243 mutant mtDNA. *Genetics* **154**, 363–380 (2000).
52. Raap, A. K. *et al.* Non-random mtDNA segregation patterns indicate a metastable heteroplasmic segregation unit in m.3243A > G cybrid cells. *PLoS ONE* **7**, e52080 (2012).
53. Chinnery, P. F. & Samuels, D. C. Relaxed replication of mtDNA: a model with implications for the expression of disease. *Am. J. Hum. Genet.* **64**, 1158–1165 (1999).
54. Elson, J. L., Samuels, D. C., Turnbull, D. M. & Chinnery, P. F. Random intracellular drift explains the clonal expansion of mitochondrial DNA mutations with age. *Am. J. Hum. Genet.* **68**, 802–806 (2001).
55. Diaz, F. *et al.* Human mitochondrial DNA with large deletions repopulates organelles faster than full-length genomes under relaxed copy number control. *Nucleic Acids Res.* **30**, 4626–4633 (2002).
56. Clark, K. A. *et al.* Selfish little circles: transmission bias and evolution of large deletion-bearing mitochondrial DNA in *Caenorhabditis briggsae* nematodes. *PLoS ONE* **7**, e41433 (2012).
57. Samuels, D. C. *et al.* Recurrent tissue-specific mtDNA mutations are common in humans. *PLoS Genet.* **9**, e1003929 (2013).
58. Jenuth, J. P., Peterson, A. C. & Shoubridge, E. A. Tissue-specific selection for different mtDNA genotypes in heteroplasmic mice. *Nat. Genet.* **16**, 93–95 (1997).
59. Takeda, K., Takahashi, S., Onishi, A., Hanada, H. & Imai, H. Replicative advantage and tissue-specific segregation of RR mitochondrial DNA between C57BL/6 and RR heteroplasmic mice. *Genetics* **155**, 777–783 (2000).
60. Burgstaller, J. P. *et al.* mtDNA segregation in heteroplasmic tissues is common *in vivo* and modulated by haplotype differences and developmental stage. *Cell Rep.* **7**, 2031–2041 (2014).
61. Wanrooij, S. *et al. In vivo* mutagenesis reveals that OriL is essential for mitochondrial DNA replication. *EMBO Rep.* **13**, 1130–1137 (2012).
62. Trifunovic, A. *et al.* Somatic mtDNA mutations cause aging phenotypes without affecting reactive oxygen species production. *Proc. Natl Acad. Sci. USA* **102**, 17993–17998 (2005).
63. Durham, S. E., Brown, D. T., Turnbull, D. M. & Chinnery, P. F. Progressive depletion of mtDNA in mitochondrial myopathy. *Neurology* **67**, 502–504 (2006).
64. Durham, S. E., Samuels, D. C., Cree, L. M. & Chinnery, P. F. Normal levels of wild-type mitochondrial DNA maintain cytochrome *c* oxidase activity for two pathogenic mitochondrial DNA mutations but not for m.3243A→G. *Am. J. Hum. Genet.* **81**, 189–195 (2007).
65. Chinnery, P. F., Samuels, D. C., Elson, J. & Turnbull, D. M. Accumulation of mitochondrial DNA mutations in ageing, cancer, and mitochondrial disease: is there a common mechanism? *Lancet* **360**, 1323–1325 (2002).
66. Cree, L. M. *et al.* A reduction of mitochondrial DNA molecules during embryogenesis explains the rapid segregation of genotypes. *Nat. Genet.* **40**, 249–254 (2008).
67. Battersby, B. J., Redpath, M. E. & Shoubridge, E. A. Mitochondrial DNA segregation in hematopoietic lineages does not depend on MHC presentation of mitochondrially encoded peptides. *Hum. Mol. Genet.* **14**, 2587–2594 (2005).
68. Jokinen, R. *et al. Gimap3* regulates tissue-specific mitochondrial DNA segregation. *PLoS Genet.* **6**, e1001161 (2010).
69. Upholt, W. B. & Dawid, I. B. Mapping of mitochondrial DNA of individual sheep and goats: rapid evolution in the D loop region. *Cell* **11**, 571–583 (1977).
70. Olivo, P. D., Van de Walle, M. J., Laipis, P. J. & Hauswirth, W. W. Nucleotide sequence evidence for rapid genotypic shifts in the bovine mitochondrial DNA D-loop. *Nature* **306**, 400–402 (1983).
References 69 and 70 describe rapid shifts in mtDNA heteroplasmy in cows. These observations laid the foundations for the mtDNA bottleneck hypothesis, which explains differences in the level of heteroplasmy observed within human pedigrees transmitting pathogenic mtDNA mutations.
71. Solignac, M., Gĕnermont, J., Monnerot, M. & Mounolou, J.-C. Genetics of mitochondria in *Drosophila*: mtDNA inheritance in heteroplasmic strains of *D.mauritania*. *Mol. Genet. Genomics* **197**, 183–188 (1984).
72. Rand, D. M. & Harrison, R. G. Mitochondrial DNA transmission genetics in crickets. *Genetics* **114**, 955–970 (1986).
73. Birky, C. W. Relaxed and stringent genomes: why cytoplasmic genes don't obey Mendel's laws. *J. Hered.* **85**, 355–365 (1994).
74. Wright, S. *Evolution and the Genetics of Populations* (University of Chicago Press, 1969).
75. Howell, N. *et al.* Mitochondrial gene segregation in mammals: is the bottleneck always narrow? *Hum. Genet.* **90**, 117–120 (1992).
76. Chinnery, P. F., Howell, N., Lightowlers, R. N. & Turnbull, D. M. MELAS and MERRF: the relationship between maternal mutation load and the frequency of clinically affected offspring. *Brain* **121**, 1889–1894 (1998).
77. White, S. L. *et al.* Genetic counseling and prenatal diagnosis for the mitochondrial DNA mutations at nucleotide 8993. *Am. J. Hum. Genet.* **65**, 474–482 (1999).
78. Wai, T., Teoli, D. & Shoubridge, E. A. The mitochondrial DNA genetic bottleneck results from replication of a subpopulation of genomes. *Nat. Genet.* **40**, 1484–1488 (2008).
79. Cao, L. *et al.* The mitochondrial bottleneck occurs without reduction of mtDNA content in female mouse germ cells. *Nat. Genet.* **39**, 386–390 (2007).
References 66, 78 and 79 describe a reduction in the amount of mtDNA present within the germ line of mice, thus providing the first direct evidence of the existence of a mammalian mtDNA genetic bottleneck.
80. Freyer, C. *et al.* Variation in germline mtDNA heteroplasmy is determined prenatally but modified during subsequent transmission. *Nat. Genet.* **44**, 1282–1285 (2012).
81. Cree, L. M., Samuels, D. C. & Chinnery, P. F. The inheritance of pathogenic mitochondrial DNA mutations. *Biochim. Biophys. Acta* **1792**, 1097–1102 (2009).
82. Chinnery, P. F. *et al.* The inheritance of mitochondrial DNA heteroplasmy: random drift, selection or both? *Trends Genet.* **16**, 500–505 (2000).
83. Jenuth, J., Peterson, A. C., Fu, K. & Shoubridge, E. A. Random genetic drift in the female germ line explains the rapid segregation of mammalian mitochondrial DNA. *Nat. Genet.* **14**, 146–151 (1996).
84. Brown, D. T., Samuels, D. C., Michael, E. M., Turnbull, D. M. & Chinnery, P. F. Random genetic drift determines the level of mutant mtDNA in human primary oocytes. *Am. J. Hum. Genet.* **68**, 533–536 (2001).
85. Stewart, J. B. *et al.* Strong purifying selection in transmission of mammalian mitochondrial DNA. *PLoS Biol.* **6**, e10 (2008).
86. Fan, W. *et al.* A mouse model of mitochondrial disease reveals germline selection against severe mtDNA mutations. *Science* **319**, 958–962 (2008).
References 85 and 86 provide evidence of selection against mtDNA variants during maternal transmission in different mouse models.
87. Greaves, L. C. *et al.* Comparison of mitochondrial mutation spectra in ageing human colonic epithelium and disease: absence of evidence for purifying selection in somatic mitochondrial DNA point mutations. *PLoS Genet.* **8**, e1003082 (2012).
88. Ma, H., Xu, H. & O'Farrell, P. H. Transmission of mitochondrial mutations and action of purifying selection in *Drosophila melanogaster*. *Nat. Genet.* **46**, 393–397 (2014).
89. Hill, J. H., Chen, Z. & Xu, H. Selective propagation of functional mitochondrial DNA during oogenesis restricts the transmission of a deleterious mitochondrial variant. *Nat. Genet.* **46**, 389–392 (2014).
90. Chacinska, A., Koehler, C. M., Milenkovic, D., Lithgow, T. & Pfanner, N. Importing mitochondrial proteins: machineries and mechanisms. *Cell* **138**, 628–644 (2009).
91. Piko, L. & Taylor, K. D. Amounts of mitochondrial DNA and abundance of some mitochondrial gene transcripts in early mouse embryos. *Dev. Biol.* **123**, 364–374 (1987).
92. Stewart, J. B. & Larsson, N. G. Keeping mtDNA in shape between generations. *PLoS Genet.* **10**, e1004670 (2014).
93. Krakauer, D. C. & Mira, A. Mitochondria and germ-cell death. *Nature* **400**, 125–126 (1999).
94. Kaare, M. *et al.* Do mitochondrial mutations cause recurrent miscarriage? *Mol. Hum. Reprod.* **15**, 295–300 (2009).
95. Behar, D. M. *et al.* The Genographic Project public participation mitochondrial DNA database. *PLoS Genet.* **3**, e104 (2007).
96. Soares, P. *et al.* Correcting for purifying selection: an improved human mitochondrial molecular clock. *Am. J. Hum. Genet.* **84**, 740–759 (2009).
97. Salas, A., Carracedo, A., Macaulay, V., Richards, M. & Bandelt, H. J. A practical guide to mitochondrial DNA error prevention in clinical, forensic, and population genetics. *Biochem. Biophys. Res. Commun.* **335**, 891–899 (2005).
98. Bandelt, H. J., Yao, Y. G., Salas, A., Kivisild, T. & Bravi, C. M. High penetrance of sequencing errors and interpretative shortcomings in mtDNA sequence analysis of LHON patients. *Biochem. Biophys. Res. Commun.* **352**, 283–291 (2007).
99. Blier, P. U., Dufresne, F. & Burton, R. S. Natural selection and the evolution of mtDNA-encoded peptides: evidence for intergenomic co-adaptation. *Trends Genet.* **17**, 400–406 (2001).
100. Ruiz-Pesini, E., Mishmar, D., Brandon, M., Procaccio, V. & Wallace, D. C. Effects of purifying and adaptive selection on regional variation in human mtDNA. *Science* **303**, 223–226 (2004).
101. Mishmar, D. *et al.* Natural selection shaped regional mtDNA variation in humans. *Proc. Natl Acad. Sci. USA* **100**, 171–176 (2003).
102. Samuels, D. C., Carothers, A. D., Horton, R. & Chinnery, P. F. The power to detect disease associations with mitochondrial DNA haplogroups. *Am. J. Hum. Genet.* **78**, 713–720 (2006).
103. Moreno-Loshuertos, R. *et al.* Differences in reactive oxygen species production explain the phenotypes associated with common mouse mitochondrial DNA variants. *Nat. Genet.* **38**, 1261–1268 (2006).
104. Baudouin, S. V. *et al.* Mitochondrial DNA and survival after sepsis: a prospective study. *Lancet* **366**, 2118–2121 (2005).
105. Jimenez-Sousa, M. A. *et al.* Mitochondrial DNA haplogroups are associated with severe sepsis and mortality in patients who underwent major surgery. *J. Infect.* **70**, 20–29 (2015).
106. Hudson, G. *et al.* Two-stage association study and meta-analysis of mitochondrial DNA variants in Parkinson disease. *Neurology* **80**, 2042–2048 (2013).
107. Ghezzi, D. *et al.* Mitochondrial DNA haplogroup K is associated with a lower risk of Parkinson's disease in Italians. *Eur. J. Hum. Genet.* **13**, 748–752 (2005).
108. Reinhardt, K., Dowling, D. K. & Morrow, E. H. Medicine. Mitochondrial replacement, evolution, and the clinic. *Science* **341**, 1345–1346 (2013).
109. Chinnery, P. F. *et al.* The challenges of mitochondrial replacement. *PLoS Genet.* **10**, e1004315 (2014).
110. Kmietowicz, Z. UK becomes first country to allow mitochondrial donation. *BMJ* **350**, h1103 (2015).
111. Shadel, G. S. & Clayton, D. A. Mitochondrial DNA maintenance in vertebrates. *Annu. Rev. Biochem.* **66**, 409–435 (1997).
112. Holt, I. J., Lorimer, H. E. & Jacobs, H. T. Coupled leading- and lagging-strand synthesis of mammalian mitochondrial DNA. *Cell* **100**, 515–524 (2000).
113. Yang, M. Y. *et al.* Biased incorporation of ribonucleotides on the mitochondrial L-strand accounts for apparent strand-asymmetric DNA replication. *Cell* **111**, 495–505 (2002).
114. Yasukawa, T., Yang, M. Y., Jacobs, H. T. & Holt, I. J. A bidirectional origin of replication maps to the major noncoding region of human mitochondrial DNA. *Mol. Cell* **18**, 651–662 (2005).
115. Terzioglu, M. *et al.* MTERF1 binds mtDNA to prevent transcriptional interference at the light-strand promoter but is dispensable for rRNA gene transcription regulation. *Cell Metab.* **17**, 618–626 (2013).
116. Kuhl, I. *et al.* POLRMT does not transcribe nuclear genes. *Nature* **514**, E7–E11 (2014).
117. Kirkman, M. A. *et al.* Gene–environment interactions in Leber hereditary optic neuropathy. *Brain* **132**, 2317–2326 (2009).
118. Hudson, G. *et al.* Clinical expression of Leber hereditary optic neuropathy is affected by the mitochondrial DNA-haplogroup background. *Am. J. Hum. Genet.* **81**, 228–233 (2007).

119. Chinnery, P. F., Andrews, R. M., Turnbull, D. M. & Howell, N. Leber's hereditary optic neuropathy: does heteroplasmy influence the inheritance and expression of the G11778A mitochondrial DNA mutation? *Am. J. Med. Genet.* **98**, 235–243 (2001).
120. Kennedy, S. R., Salk, J. J., Schmitt, M. W. & Loeb, L. A. Ultra-sensitive sequencing reveals an age-related increase in somatic mitochondrial mutations that are inconsistent with oxidative damage. *PLoS Genet.* **9**, e1003794 (2013).
121. Boore, J. L., Macey, J. R. & Medina, M. Sequencing and comparing whole mitochondrial genomes of animals. *Methods Enzymol.* **395**, 311–348 (2005).
122. Lou, D. I. *et al.* High-throughput DNA sequencing errors are reduced by orders of magnitude using circle sequencing. *Proc. Natl Acad. Sci. USA* **110**, 19872–19877 (2013).
123. Cao, L. *et al.* New evidence confirms that the mitochondrial bottleneck is generated without reduction of mitochondrial DNA content in early primordial germ cells of mice. *PLoS Genet.* **5**, e1000756 (2009).
124. Carling, P. J., Cree, L. M. & Chinnery, P. F. The implications of mitochondrial DNA copy number regulation during embryogenesis. *Mitochondrion* **11**, 686–692 (2011).
125. Wallace, D. C., Brown, M. D. & Lott, M. T. Mitochondrial DNA variation in human evolution and disease. *Gene* **238**, 211–230 (1999).
126. Torroni, A. *et al.* Classification of European mtDNAs from an analysis of three European populations. *Genetics* **144**, 1835–1850 (1996).
127. van Oven, M. & Kayser, M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum. Mutat.* **30**, E386–E394 (2009).

Acknowledgements
J.B.S. is supported by the Max Planck Society and a project grant from the United Mitochondrial Disease Foundation. P.F.C. is an honorary consultant neurologist at Newcastle upon Tyne Foundation Hospitals UK National Health Service (NHS) Foundation Trust, a Wellcome Trust Senior Fellow in Clinical Science (101876/Z/13/Z) and a UK National Institute for Health Research (NIHR) senior investigator. P.F.C. receives additional support from the Wellcome Trust Centre for Mitochondrial Research (096919Z/11/Z), the UK Medical Research Council Centre for Translational Muscle Disease research (G0601943), the European Union FP7 (Seventh Framework Programme for Research and Technological Development) TIRCON (Treat Iron-Related Childhood-Onset Neurodegeneration) consortium, and the NIHR Newcastle Biomedical Research Centre based at Newcastle upon Tyne Hospitals NHS Foundation Trust and Newcastle University. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the UK Department of Health.

Competing interests statement
The authors declare no competing interests.

FURTHER INFORMATION
MITOMAP: http://www.mitomap.org

ALL LINKS ARE ACTIVE IN THE ONLINE PDF