Evolution of cell-to-cell variability in stochastic, controlled, heteroplasmic mtDNA populations

Iain G. Johnston¹, Nick S. Jones²
¹ School of Biosciences, University of Birmingham, B15 2TT, UK
² Department of Mathematics, Imperial College London, SW7 2BB, UK

Abstract
Populations of physiologically vital mitochondrial DNA (mtDNA) molecules evolve in cells under control from the nucleus. The evolution of populations of mixed mtDNA types is complicated and poorly understood, and variability of these controlled admixtures plays a central role in the inheritance and onset of genetic disease. Here, we develop a mathematical theory describing the evolution and variability in these stochastic populations for any type of cellular control, showing that cell-to-cell variability in mtDNA, and mutant load, inevitably increases with time, according to rates which we derive and which are notably independent of the mechanistic details of feedback signalling. We show with a set of experimental case studies that this theory explains disparate quantitative results from classical and modern experimental and computational studies on heteroplasmy variance in different species. We demonstrate that our general model provides a host of specific insights, including a modification of the often-used but hard-to-interpret Wright formula to correspond directly to biological observables, the ability to quantify selective and mutational pressure in mtDNA populations, and the pronounced variability inevitably arising from the action of possible mtDNA quality-control mechanisms. Our general theoretical framework, supported by existing experimental results, thus helps understand and predict the evolution of stochastic mtDNA populations in cell biology.

Introduction
Molecules of mitochondrial DNA (mtDNA) form dynamic evolutionary populations within cells, replicating and degrading according to cellular control signals [1, 2]. MtDNA can vary due to mutation or artificial manipulation [3]; the proportion of mutant mtDNA in a cell is referred to as heteroplasmy. MtDNA encodes vital aspects of the bioenergetic machinery of eukaryotic cells; mtDNA variability can thus have dramatic cellular consequences, including devastating genetic diseases and numerous other conditions [3], making a theoretical understanding of this complex evolutionary system important. Understanding the natural feedback control acting on mtDNA populations is also a vital step in the development of artificial approaches to control mitochondrial behaviour with genetic tools [4, 5]. The population variances of mtDNA types, and heteroplasmy variance, are of particular importance, owing to their implications for maternal transmission of dangerous mutations [6] and the manifestation of pathologies dependent on the range of heteroplasmy in present in a tissue [7], including the demonstration that a very small proportion of cells exceeding a heteroplasmy threshold can lead to pathologies [8].

Stochasticity underlies cell biology; cellular processes including gene expression [9, 10, 11], DNA replication [12], and mitochondrial and mtDNA dynamics [13, 14, 15, 16] are subject to fundamentally stochastic influences. Variability in mitochondria can be a leading contributor to cell physiological behaviour, making mitochondria an important target for explanatory stochastic models [15]. Existing studies have included stochastic modelling and numerical treatments of mitochondrial [17] and mtDNA populations with the assumption of specific control mechanisms [18, 19, 1, 14, 20]. Other theoretical studies have drawn on classical statistical genetics, notably including the well-known Wright formula [21, 22], to produce a description of partitioning at cell divisions, but the role of stochastic mtDNA dynamics between cell divisions is largely omitted. Although recent experimental studies are starting to shed light on cellular control of mtDNA [14, 16], a general theoretical framework is currently absent. Here we address this open question by constructing a general, bottom-up stochastic description of mtDNA populations subject to arbitrary cellular control mechanisms, providing analytic results for the predicted behaviour associated with any mtDNA control mechanism, and adapting the classic Wright formula to account for and interpret stochastic mtDNA dynamics. Notably, our approach and results hold independently of the details of specific regulatory mechanisms underlying mtDNA feedback signalling, providing a general theoretical framework for control of stochastic mtDNA populations across different species and environments.

As we develop the theoretical framework to address mtDNA dynamics below, we will consider a set of appli-
I. When mean copy numbers are constant, at least one variance (wildtype or mutant) increases linearly with time (Eqns. 8-10).

II. When mean copy numbers are constant, heteroplasmic variance increases linearly with time with a rate that depends only on mtDNA copy numbers and the rate of mtDNA turnover, and not on any details of mechanism or signaling (Eqn. 12); many copy-number control mechanisms are therefore indistinguishable even from dynamical measurements of heteroplasmy (Fig. 2).

III. Control applied through (a) biogenesis rates and (b) degradation rates induce comparable behaviour in the cellular mtDNA population (Fig. 2; SI Section 2).

IV. A modified Wright formula (Eqn. 14; Fig. 4) gives a general method to establish a quantitative link between observed normalised heteroplasmy variance $h^2$ and observable (as opposed to \langle \rangle) and effective quantities (Fig. 5).

V. Quality control does not guarantee the clearing or stabilisation of mutant load, and can induce substantial variance in wildtype mtDNA (potentially challenging cells with diminished populations; Fig. 6 & Fig. 7).

Figure 1: Key biological findings from our mathematical model. These results arise when the assumptions of our modelling approach hold — that cells are heteroplasmic, mtDNA replication and degradation are Poisson processes with the same rates for mutant and wildtype mtDNA, and feedback mechanisms depend only on the current state of the system. We justify these assumptions with reference to experimental data in the text and in Appendix A1.

Key biological findings from our mathematical model. These results arise when the assumptions of our modelling approach hold — that cells are heteroplasmic, mtDNA replication and degradation are Poisson processes with the same rates for mutant and wildtype mtDNA, and feedback mechanisms depend only on the current state of the system. We justify these assumptions with reference to experimental data in the text and in Appendix A1.

These results arise when the assumptions of our modelling approach hold — that cells are heteroplasmic, mtDNA replication and degradation are Poisson processes with the same rates for mutant and wildtype mtDNA, and feedback mechanisms depend only on the current state of the system. We justify these assumptions with reference to experimental data in the text and in Appendix A1.

Results

Copy number variance with stable population means

Any control mechanism of the form Eqns. 1-4 (including manifestations of feedback control) can be represented to linear order by a Taylor expansion of its rates about \{w, m\} (the steady state exists by construction with our previous assumption):

\[
\lambda(w, m) \simeq \beta_0 + \beta_w(w - \bar{w}) + \beta_m(m - \bar{m}), \tag{5}
\]

\[
\nu(w, m) \simeq \delta_0 + \delta_w(w - \bar{w}) + \delta_m(m - \bar{m}). \tag{6}
\]

It will readily be seen that to support a stable population mean at \{\bar{w}, \bar{m}\}, $\delta_0 = \beta_0$. Assuming that $w$ and $m$ can be written as the sum of a deterministic and a fluctuating component, we use Van Kampen’s system size expansion to find a Fokker-Planck equation describing the behaviour of $w$ and $m$ governed by Eqns. 5-6 [24, 25]. From this equation we extract expressions for the time behaviour of the mean and variance of $w$ and $m$ (see SI Sections 1-2).

We show in SI Section 4 that attempting to identify a stable state for population variances and covariance yields the condition:

\[
\frac{2\bar{m}(\bar{m} + \bar{w})\beta_0}{\bar{w}} = 0; \tag{7}
\]
hence, any population mean state in which mutant content $\bar{m}$ is nonzero does not admit a stationary solution for variances, unless $\beta_0 = 0$. If $\beta_0 = 0$ then there is no further change to the system once steady state has been reached (no stochastic turnover occurs) and the system remains frozen thereafter. In other words, for a nonzero mutant population and nonzero mtDNA turnover, the variance of at least one mtDNA population will change with time.

The Fokker-Planck equation can be used to compute the expected behaviours of $\langle w^2 \rangle$ (wildtype variance), $\langle m^2 \rangle$ (mutant variance), and $\langle w m \rangle$ (wildtype-mutant covariance) for a given control mechanism. The variance and covariance solutions display some transient behaviour, involving terms on the timescale $t'$ of $\exp((\beta_m - \delta_m)\bar{m} + (\beta_w - \delta_w)\bar{w})t$. As, for stability, $\beta_i$ are nonpositive and $\delta_i$ are nonnegative, $t'$ is either a constant or an exponentially decaying function of time $t$. The expressions thus subsequently converge to linear trends for large $t$:

$$\langle w^2 \rangle = F_1^{\text{decay}}(t') + \theta_1 t + \phi_1 \tag{8}$$
$$\langle w m \rangle = F_2^{\text{decay}}(t') + \theta_2 t + \phi_2 \tag{9}$$
$$\langle m^2 \rangle = F_3^{\text{decay}}(t') + \theta_3 t + \phi_3 \tag{10}$$

The forms of the transient functions $F_i^{\text{decay}}$, and the constants $\theta_i$ and $\phi_i$ are given in SI Section 2 and are functions only of the difference between replication and degradation rates ($\beta_i - \delta_i$), steady-state copy numbers $\bar{m}$ and $\bar{w}$, and mitophagy rate $\beta_0$. Furthermore, the structure of these expressions is such that for $\beta_0 \neq 0$ and nonzero $w$ and $m$, at most one of the $\theta_i$ can be zero, $\theta_1 \geq 0$, and $\theta_3 \geq 0$. Thus, around the mean $(\bar{w}, \bar{m})$, either wildtype variance or mutant variance, or both, increase linearly with time (Fig. 1 I). As time continues, the increasing variance means that extinction of one mtDNA becomes increasingly likely: implications of this behaviour are explored below.

The mathematical structure of the solutions only ever involves the difference between replication and degradation rates ($\beta_i - \delta_i$), showing that control of (a) biogenesis rates and (b) degradation rates induce comparable behaviour in the cellular mtDNA population (Fig. 1 III).

**Heteroplasmy statistics**

As shown in SI Section 5, a first order Taylor expansion of a function of random variables gives an approximation for the variance of $h = m/(w + m)$:

$$\langle h^2 \rangle = \frac{\langle (w^2) \rangle \langle m^2 \rangle + \langle m^2 \rangle \langle w^2 \rangle - 2\langle w m \rangle \langle w \rangle \langle m \rangle}{((\langle m \rangle + \langle w \rangle)^4} \tag{11}$$

![Figure 2: Behaviour of different specific control mechanisms. All control mechanisms have at least one population with time-increasing variance and so have increasing heteroplasmy variance; mean heteroplasmy ($h$) remains constant through these simulations. (i) Replication $\lambda(w, m)$ and degradation $\nu(w, m)$ rates for model control mechanisms explored in the text, from existing studies and newly proposed here. (ii)-(iii) Copy number and heteroplasmy moments with time for these control mechanisms, (ii) in the absence of cell divisions and (iii) with binomial cell divisions every 10 time units. Analytic results (lines; full expressions in SI Section 3) match stochastic simulation (crosses) throughout. Parameters used are chosen to support the same steady state ($\bar{w} = 900, \bar{m} = 100$) with a turnover timescale of $\tau = 5$ days. Other parameters: $\gamma = 0, \alpha_m = 0.001, \alpha = 2$ (except for models B, E, F, for which $\alpha = 0.002, 0.002, 10^2$ respectively); $10^5$ stochastic simulations used.](Image 315x283 to 555x663)
and we can then use previously obtained expressions for \(\langle w \rangle, \langle m \rangle, \langle w^2 \rangle, \langle m^2 \rangle, \langle w m \rangle\) to compute this approximate heteroplasmy variance. Neglecting transient terms and using \(\langle w \rangle = \bar{w}\) and \(\langle m \rangle = \bar{m}\) in Eqn. 11, gives, after some algebra,

\[
\langle h^2 \rangle = \frac{2\beta_0}{n} h(1-h)t; \quad \frac{\langle h^2 \rangle}{\langle h \rangle(1-\langle h \rangle)} = \frac{\langle h^2 \rangle}{\langle h \rangle} = \frac{2\beta_0 t}{n},
\]

where \(n = \bar{w} + \bar{m}\), with \(\bar{w} = (1-h)n\) and \(\bar{m} = hn\); thus \(h\) is (mean) heteroplasmy and \(n\) is (mean) total copy number (recall that \(\beta_0 = \delta_0\) in steady state). In Eqn. 12 we have used normalised heteroplasmy variance \(\langle h^2 \rangle\), accounting for the dependence of \(\langle h^2 \rangle\) on the magnitude of \(h\); \(\langle h^2 \rangle\) is the quantity most often reported in experimental studies.

In other words, when the system size expansion is valid (see below), for any control mechanism, heteroplasmy variance in the copy number steady state increases linearly with time with a rate that depends only on the copy numbers of the system and the timescale of random turnover (Fig. 1 II) (Eqn. 12 is independent of the \(\beta\) and \(\delta\) terms in Eqs. 5-6). As we discuss later, this observation implies that many possible mechanisms could be responsible for the same observed trend in heteroplasmy variance, meaning that measurements of heteroplasmy variance alone, even if repeated many times, can be used to discriminate between different control mechanisms.

**Transient behaviour and cell divisions, and validity of the expansion**

To obtain analytic insight, we have thus far focussed on modelling mtDNA behaviour using the system size expansion when a steady-state assumption had already been applied. Transient behaviour can also be explored by employing the system size expansion directly on the appropriate master equation, using the full expressions for \(\lambda(w, m)\) and \(\nu(w, m)\) (see SI Sections 2 & 6). Relaxing the steady state assumption means that the ODEs describing variance behaviour are analytically intractable for many forms of \(\lambda(w, m), \nu(w, m)\). However, they can simply be solved numerically and, as shown in subsequent sections, well match stochastic simulation (which of course is numerically far more intensive). This ODE approach fully accounts for non-equilibrium behaviour – including transient relaxation, cycling, and so on – while the system size expansion remains appropriate (see below).

This analysis can readily be used to characterise the effect of partitioning mtDNAs at cell divisions. To compute the time behaviour of variance where cell divisions occur at arbitrary times, we invoke a linear noise assumption [24, 25], first using the ODEs above to compute the variance behaviour within one cell cycle. Partitioning rules for copy number statistics are then applied, and the resulting post-partition statistics are used as the initial condition for a next phase of ODE solution. We here illustrate this process for binomial partitioning of mtDNAs to connect with recent studies in mice [14] and HeLa [15], although with an appropriate choice of partitioning rules, this approach can be used to address any partitioning regime (for example, the sub-binomial case recently reported in fission yeast [16]). In the case of binomial partitioning, the appropriate partitioning rules are \(\langle w \rangle \to \langle w \rangle/2, \langle m \rangle \to \langle m \rangle/2, \langle w^2 \rangle \to \langle w^2 \rangle/4 + \langle w \rangle/4, \langle w m \rangle \to \langle w m \rangle/4, \langle m^2 \rangle \to \langle m^2 \rangle/4 + \langle m \rangle/4\), following straightforwardly from the variance of a binomial distribution with \(p = 1/2\) as \(n/4\). We will see below that this picture well describes the behaviour of stochastic populations in dividing cells: hence, the total variance contributions of turnover between divisions and partitioning at divisions can be modelled as a linear sum, and the behaviour of mechanisms across cell cycles is comparable to that within a cell cycle.

The results above hold for a nonzero mutant population. As copy number variance increases, we expect extinction of one mtDNA type to become increasingly likely. To address this behaviour, we must consider when the system size expansion itself, which is reliant on the validity of the linear noise approximation, holds. An important threat to this validity is a non-negligible extinction probability for one mtDNA type, whereupon the mean and variance of mtDNA populations in dividing cells: hence, the total variance contributions of turnover between divisions and partitioning at divisions can be modelled as a linear sum, and the behaviour of mechanisms across cell cycles is comparable to that within a cell cycle.

The results above hold for a nonzero mutant population. As copy number variance increases, we expect extinction of one mtDNA type to become increasingly likely. To address this behaviour, we must consider when the system size expansion itself, which is reliant on the validity of the linear noise approximation, holds. An important threat to this validity is a non-negligible extinction probability for one mtDNA type, whereupon the mean and variance of mtDNA populations in dividing cells: hence, the total variance contributions of turnover between divisions and partitioning at divisions can be modelled as a linear sum, and the behaviour of mechanisms across cell cycles is comparable to that within a cell cycle.

The results above hold for a nonzero mutant population. As copy number variance increases, we expect extinction of one mtDNA type to become increasingly likely. To address this behaviour, we must consider when the system size expansion itself, which is reliant on the validity of the linear noise approximation, holds. An important threat to this validity is a non-negligible extinction probability for one mtDNA type, whereupon the mean and variance of mtDNA populations in dividing cells: hence, the total variance contributions of turnover between divisions and partitioning at divisions can be modelled as a linear sum, and the behaviour of mechanisms across cell cycles is comparable to that within a cell cycle.

...
linear); (iii) when extinction of one type is almost certain, the system tends towards its behaviour if only one type was present (ultimately stalling variance increase, as if \( \hat{m} = 0 \) in Eqn. 7). The results we focus on in this main text can be viewed as describing the ‘quasi-stationary state’ where extinction is negligible; further quantitative details can be derived using, for example, adaptations of the system size expansion that address extinction [26].

**Specific control mechanisms and comparison with simulation**

The previous results make no assumptions about the specific form of control applied to the mtDNA population, other than it depends only on current state and is manifest through the rates of Poissonian replication and degradation which are equal for both mtDNA species (and can be described with the system size expansion, as discussed above). We can exploit the generality of the preceding formalism to obtain results for any given (feedback) control mechanism, defined by a specific form of \( \lambda(w, m) \) and \( \nu(w, m) \) in Eqns. 1–4.

We first consider the well-known ‘relaxed replication’ model [19, 1], which involves stochastic mtDNA degradation, coupled with mtDNA replication which is physically modelled as a deterministic process. We propose that, if degradation is regarded as a stochastic process (due to its microscopic reliance on complicated processes and localisations in the cell), picturing replication (which also relies on complicated interactions on the microscopic scale) as a stochastic process leads to a consistent stochastic generalisation (see Appendix A1). The corresponding model has exactly the same expressions for rates as in Ref. [1] (A in Fig. 2 i), but replication rate is now interpreted as the rate of a stochastic, rather than a deterministic, process.

We also introduce several other models for mtDNA control, to consider a range of potential functional forms, including differential and ratiometric control based on a target wildtype copy number, an absence of any feedback control, and others (Fig. 2 i B-G). The presentation or absence of \( w \) and \( m \) in these expressions reflects what quantity is being sensed by the cell (wildtype mtDNA alone, mutant mtDNA alone, or a combination of the two). We further note that this general formalism can also incorporate physical constraints on the mtDNA population. For example, the hypothesis that mitochondrial concentration is controlled between cell divisions [15] (recently confirmed in fission yeast [16]) could correspond to \( w_{opt} \), the ‘target’ mtDNA number, being a linear function of cell volume in the models above; or could arise through passive birth-death dynamics (model D) with control implemented at the cell division stage (see previous section). The interpretations of these control mechanisms in terms of cellular sensing and the language of stochastic population processes are given in SI Section 3.

Fig. 2 ii illustrates the application of our analysis to these example control mechanisms in the absence of cell divisions. The close agreement between stochastic simulation and analytic results in steady state shows the generality of our theory. The long-term linear increases in one or both mtDNA variances are clear (Fig. 1 I), and trajectories of \( \langle h^2 \rangle \) with the same steady state and turnover timescale are identical (Fig. 1 II). Fig. 2 iii, including cell divisions, demonstrates close agreement between ODE solutions and stochastic simulation, further showing that the linear noise treatment successfully captures stochastic behaviour over cell divisions. The close similarity of \( \langle h^2 \rangle \) trajectories across divisions is a consequence of their aforementioned identity in steady state conditions with no cell divisions (Fig. 2 ii); differences are due to the difference in mechanism behaviour away from the steady state.

**Applications I: Heteroplasmy variance increases at constant mean copy number**

The increase of heteroplasmy variance \( \langle h^2 \rangle \) with time is of profound importance in determining the inheritance and onset of mtDNA diseases. Because disease symptoms often manifest only when heteroplasmy exceeds a certain threshold [7], increasing heteroplasmy variance with time can lead to pathologies even if mean heteroplasmy does not change (because a higher cell-to-cell variance implies a greater probability of a given cell exceeding a threshold) [14].

We sought experimental evidence to support the linear increase of \( \langle h^2 \rangle \) predicted by our theory. Time course measurements of single-cell heteroplasmy values remain limited; we identified results from the *Drosophila* germline [27] and in the mouse germline for the NZB/BALB model [28, 29] and the HB model [14]. For these data, we compared the ability to fit the data of a null model (\( \mathcal{H}_0 : \langle h^2 \rangle' = \alpha + \epsilon \), where \( \alpha \) is a constant), and an alternative model (\( \mathcal{H}_1 : \langle h^2 \rangle = \alpha + \beta t + \epsilon \), where \( \langle h^2 \rangle' \) changes linearly with time as our theory predicts (Eqn. 12; Fig. 1 II). Using the Akaike information criterion (AIC) and assuming normally-distributed noise on mean \( \langle h^2 \rangle' \) (\( \epsilon \sim N(0, \sigma^2) \), an assumption consistent with our linear approximation, but which can be further refined as in Ref. [29]), we found that the alternative, time-varying model was favoured in all cases, providing support for our theory (Fig. 3A-C).

These results are qualitatively consistent with a previous study on the dynamics of heteroplasmy variance during the mtDNA bottleneck in mice [14] where a mechanism involving random mtDNA turnover and random mtDNA partitioning at cell divisions was found to best explain experimental observations. MtDNA in mice and rats often has a half-life of 10–100 days [23]. This corresponds to \( \beta_0 = \delta_0 = \log 2/t_{1/2} = 0.03 - 0.003 \text{ days}^{-1} \) and \( \tau = t_{1/2}/\log 2 = 5 - 50 \text{ days} \). In Fig. 2
we show the patterns of copy number and heteroplasmy means and variances for $\tau = 5$ days and $10^3$ mtDNA molecules per cell under different specific control strategies. The increase of $\langle h^2 \rangle$ from 0 to $10^{-3}$, corresponding for $h = 0.1$ to an increase in $\langle h^2 \rangle'$ from 0 to 0.011, matches the scale of change observed in the mtDNA bottleneck (though the bottleneck is complicated by changing population size $n$ and compensatory changing turnover $\beta_0$) [14]. Previous work has shown that the case with no feedback (D in Fig. 2) describes well the behaviour of mtDNA with cell divisions in mouse development [14]. In Appendix A1 we discuss further connections with previous theoretical studies; experimental cell-to-cell measurements in more quiescent cell types, while currently lacking, will provide valuable further tests of our theory.

MtDNA turnover in the Wright formula

Powerful existing analyses of mtDNA population variance [22, 29] with widespread influence [3] have drawn upon a classical theory by Wright (and Kimura) [21, 30], describing stochastic sampling of a population of elements between generations. The resulting expression for expected heteroplasmy variance is the well-known equation sometimes referred to as the ‘Wright formula’ (though other equations also bear this name) [22, 3]:

\[
\langle h^2 \rangle' = 1 - (1 - (2n_e)^{-1})^g,
\]

where $n_e$ is an effective population size and $g$ is a number of generations. The mapping of this effective theory to the complicated mtDNA system is valuable to develop intuition but cannot capture the detailed dynamics of individual mtDNA molecules, due to assumptions (see Appendix A2) that mean the effective parameters of the theory ($n_e$ and $g$) cannot generally be interpreted as biological observables [22, 31], preventing quantitative analyses of mechanisms and dynamics [3]. In particular, $n_e$ does not generally correspond to a minimum mtDNA copy number (see Appendix A2), and, being a genetic rather than a physical parameter, ‘is unlikely ever to correspond closely to the number of anything’ [31].

The Wright formula, however, does accurately describe the heteroplasmy variance due to binomial sampling of $2n_e$ real elements at cell divisions for an observable population size, and, as seen in the previous section, the additional effect of mtDNA turnover based on observable values can be included as an extra linear contribution. In general, this term will depend on the dynamics controlling the mtDNA population and can easily be calculated using the ODE approach above (Fig. 2).

In the case where no systematic change in mtDNA population size occurs with time, we can use the observation that $\langle h^2 \rangle$ trajectories are often comparable across a variety of different possible cellular control mechanisms (and identical in the steady state; Fig. 2 and Eqn. 12) to produce a simple approximate description linking $\langle h^2 \rangle'$ to observables. The simple steady state behaviour is given by Eqn. 12. To construct an approximation we use a simple estimate of mean population size over a cell cycle, writing $n' = \frac{3}{4} n$, where $n$ is the mtDNA population size immediately after division, and $n'$ thus gives a population size ‘average’ over the changes within a cell cycle. Using the above analysis with $w_0 = (1 - \langle h \rangle)n'$, $m_0 = \langle h \rangle n'$ (representing the ‘average’ populations of wildtype and mutant mtDNA), and $\beta_0 = 1/\tau$ (so that $\tau$ is the timescale of mtDNA degradation), the corresponding expression in terms of $h$ and $n$ is then given by a ‘turnover-adjusted’ Wright formula (see Appendix.
Applications II: Linking physical and genetic rates with the modified Wright formula

The Wright formula is traditionally used to compare a heuristic, effective ‘bottleneck size’ across experimental systems (for example, in studies of different organisms [22] and of human disease [35]). In its uncorrected form this ‘bottleneck size’ can only be semi-quantitatively treated – bottleneck sizes can be ranked, but absolute values and differences cannot be straightforwardly interpreted. Our adaptation allows us to use this formula to connect the rates of physical subcellular processes with the resulting rates of genetic change.

To illustrate this connection, we focus on a particular period during mouse development. Between 8.5 and 13.5 days post conception (dpc) in the developing mouse germ line, cell divisions occur with a period of about 16 hours [36], giving $g = 7$ or 8 cell divisions in this period (of length $t = 5$ days). Copy number measurements during this period show that the mean total number of mtDNA molecules per cell remains of the order of $n = 2000$ (Fig. 5(i)) [28, 33, 34]. During this period, heteroplasmy variance $\langle h^2 \rangle'$ increases on average (but with substantial variability) from around 0.01 to 0.02 (Fig. 5(ii)) [32, 28]. Fig. 5(ii) shows a best-fit line to $\langle h^2 \rangle'$ data, with slope $1.52 \times 10^{-3}$ day$^{-1}$ (95%-confidence intervals $(1.11 - 1.92) \times 10^{-3}$ day$^{-1}$).

We can use these measurements in conjunction with the turnover-adjusted Wright formula (Eqn. 14) to obtain estimates for the rate of mtDNA turnover during this period. Using Eqn. 14 with the best-fit $\langle h^2 \rangle'$, $t = 5$ days gives the resulting estimate $\tau \approx 0.57$ days (5-95% confidence intervals 0.43-0.88 days, using the same values for $n, g, t$) for the characteristic timescale of mtDNA degradation. This increase in mtDNA turnover (relative to the $\tau \approx 5-50$ day timescale in differentiated tissues [23]) in the germ line during this developmental period matches quantitative results from a more detailed study of the bottleneck reporting $\tau$ within the range 0.38-2.1 days (based on posteriors for $\nu = 1/\tau$ between 0.02-0.11 hr$^{-1}$) [14], and illustrates how a suitable mathematical model can be used to estimate biological quantities that are challenging to directly address with experiment [37].

Influence of mutations, replication errors, and selective differences: quality control of replication errors

Our approach is easily generalised to include other processes than those described by Eqn. 1-4: in SI Section 6 we demonstrate that adding and changing appropriate processes allows us to analyse the effects on mtDNA mean and variance due to de novo mutations, replication errors, and multiple selective pressures. Our approach can be thus used to characterise variability arising from selection and mutation under any control mechanisms.
Relaxed replication (model A), which combines a target copy number with a strong coupling between mutant and wildtype mtDNA, displays an intermediate increase on wildtype variance as mutant is cleared.

Theoretical approaches which only consider the mean behaviour of mtDNA populations (Fig. 6 (i)) cannot account for this cellular heterogeneity, and the important fact that quality control acting to remove mutant mtDNA can also induce variability in wildtype mtDNA (Fig. 6 (ii)). The action of quality control may therefore yield a subset of cells with wildtype mtDNA substantially lower than the mean value across cells – potentially placing a physiological challenge on those cells where wildtype mtDNA is decreased. In addition to the important point that the simple presence of quality control does not guarantee the clearing or stabilisation of mutant load, we thus find that quality control may have substantial effects on wildtype as well as mutant mtDNA if cellular control couples the two species (Fig. 1 V).

Applications III: Variance induced through mutant clearing

The A > G mutation at position 3243 in human mtDNA is the most common heteroplasmic pathological mtDNA mutation, giving rise to MELAS (mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes), a multi-system disease. The dynamics of 3243A > G heteroplasmy are complex and tissue-dependent; its behaviour in blood has been characterised in particular detail using a fluorescent PCR assay for heteroplasmy [38] in a way that allows us to explore our theoretical predictions about mtDNA statistics with mutant clearing.

The authors of Ref. [38] took two blood samples, several years apart, from human patients, and quantified 3243A > G heteroplasmy (h) and mean mtDNA copy number per cell for both samples. While single-cell data is not presented in the publication, progress can be made with the averaged quantities. A strong decrease in (h) with time is observed for all patients, confirming that mutant mtDNA is being cleared (Fig. 7(i)). The behaviour of total mtDNA molecules per cell (w + m) is less consistent, with a range of large increases and moderate decreases in total number. As shown in Fig. 7(ii)-(iii), patient-to-patient variance in both w and m increases with time in conjunction with h decreases.

Although the measurements in Ref. [38] are averages over groups of cells, an approximate quantitative comparison of these data with the predictions of our theory can be made. In the spirit of ‘back-of-the-envelope’ calculations [39], we estimate that each sample of cells giving rise to a measurement corresponds to approximately $10^9$ cells (see SI Section 8). Then, using an estimate of $\tau = 5$ days (by comparison with other mammalian species, as above), we find that a selective pressure of $\epsilon_s \simeq 1.2 \times 10^{-4}$ day$^{-1}$ matches the observed decrease in

Figure 6: Variability in quality control clearing of mutants from replication errors. Co-evolution of (i) means and (ii) variances of mutant m and wildtype w mtDNA in cells where replication errors occur. Solid lines (downward trajectories; labels [A-F]0) give the case where quality control is sufficiently strong to clear the remaining mutant molecules (hence decreasing (m) in (i)); dashed lines (upward trajectories; labels [A-F]1) give the case where quality control cannot clear mutants (hence increasing (m)). Colours and labels correspond to different control models (see text and Fig. 2). Points show stochastic simulation. Importantly, regardless of the success of quality control in clearing mutants, wildtype mtDNA variability ($\langle w^2 \rangle$) can increase substantially during the action of quality control (rightwards movement in (ii)), and some strategies lead to order-of-magnitude differences in this increase (starred arrow).
heteroplasmy, corresponding, for example, to a decrease from 0.15 to 0.14 over 8 years, as in Fig. 7(i). Solving the ODEs resulting from this system, using control model D as the simplest case, predicts increases in cell-to-cell copy number variance of approximate magnitude \( \langle w^2 \rangle \sim 10^3 \) and \( \langle m^2 \rangle \sim 2 \times 10^3 \) over 8 years. We can translate these cell-to-cell values into the variance expected across samples of cells by dividing by the number of samples (taken as \( 10^3 \) as above). The resulting variance corresponds, for example, to expected standard deviations in wildtype and mutant copy number after 8 years of 10.0 and 4.0 respectively for a sample with \( \langle w \rangle = 85 \) and \( \langle m \rangle = 15 \) (consistent with the increasing spread of values in Fig. 7(ii)-(iii)). We used the Kolmogorov-Smirnov test respectively to test the alternative hypotheses that the experimentally-observed \( \langle w \rangle \) and \( \langle m \rangle \) at the later time point differed from those predicted by our model; no test yielded \( p < 0.05 \) (see SI Section 8). Of course, an absence of support for an alternative hypothesis cannot be taken as support for a null hypothesis, but shows that the existing experimental data is not incompatible with our model.

The observations in Fig. 7 support the predictions made in Fig. 6, where mutant load is decreased but variance in wildtype copy number (and mutant copy number) increases. The bottom-left quadrant of Fig. 6 (ii) shows that different control mechanisms display similar initial behaviour; follow-up studies on these patients could be used to distinguish possible mechanisms for mtDNA control. For example, if the rate of wildtype variance increase decreases over time, models B, C, and E are more likely; if wildtype variance continues to increase, models A, D, F, and G and more likely. More detailed model discrimination based on the time behaviour of \( w \) and \( m \) variances are possible (Fig. 2), and can be performed using statistical methods accounting for mean and variance behaviour [29, 40].

Discussion

A general, bottom-up theory has been produced to describe the time behaviour of cell-to-cell variance in mtDNA populations subject to controlled biogenesis and/or degradation, mutation, selection, and cell divisions. This theory is based around the microscopic behaviour of mtDNA molecules, allowing a hitherto absent connection between widely-used ‘effective’ statistical genetics approaches (Eqn. 13) and measurable biological quantities, and motivating experiments to further elucidate the mechanisms acting to control mtDNA (described in SI Section 7). We have shown that the predictions of this theory agree with experimental observations of mixed mtDNA populations, and that the application of appropriately validated mathematical theory allows us to make estimates of important biological quantities that remain challenging to directly address with experiments. Our theory describes the cell-to-cell variability in mtDNA populations and thus provides a framework with which to understand the inheritance and onset of mtDNA diseases [6, 14].

Our theoretical platform unifies several existing modelling approaches that have driven advances in the study of mtDNA populations. We have specifically demonstrated that the ‘relaxed replication’ model [19, 1, 33] (our model A), simple birth-death models [14, 20] (our model D), and cellular controls based on homeostatic principles [15, 16] (our models B, C, G) can naturally be represented within our framework. As a result, analytic expressions for the expected behaviour of heteroplasmy variance and other population statistics can readily be extracted for these and other mtDNA models (see SI Section 3), allowing the detailed characterisation of mtDNA dynamics, including the probability of crossing disease thresholds [7], which can be computed from heteroplasmy statistics [14]. We have also used the theoretical ideas developed herein to refine a widely-used model for mtDNA populations of changing size (the Wright formula), explicitly connecting it with cellular processes and allowing a link between physical and genetic quantities (Eqn. 13, Fig. 5).

Further, our theory also describes the dynamics of heteroplasmy change with time in the presence of selective pressure for one mtDNA type. Given an initial heteroplasmy \( h_0 \) and a selective pressure \( \beta \) (positive \( \beta \) corresponding to positive selection for the mutant mtDNA type), we find (see SI Section 5) that heteroplasmy evolves according to:

\[
h = \frac{1}{1 + \frac{1}{h_0} e^{-\beta t}}.
\]

This behaviour immediately motivates a transformation allowing the evolution of heteroplasmy to be compared across different starting values \( h_0 \):
allowing, as in our previous work [23], heteroplasmy results from different biological samples to be compared together, accounting for different initial heteroplasmies (in other words, the same selective pressure will produce the same $\Delta h'$ regardless of $h_0$).

It is likely that control mechanisms found in biology have nonlinear forms (for example, sigmoidal response curves are common in cellular signalling). We have shown that a linearisation satisfactorily describes some non-equilibrium behaviour (for example, in the case of our cell division model) but further investigation of more general nonlinear behaviour, and modulation of wider cell behaviour by mtDNA populations (for example, by influencing cell cycle progression [15]), are important future developments. In SI Section 1 we discuss a linear stability analysis of our expressions for mean mtDNA behaviour, which highlights a link between the ‘sensing’ of an mtDNA species (in the sense that the presence of that species modulates replication or degradation rates) and the ability to control the mean level of that species.

The control of stochastic systems is a well established field within control theory [41]. Optimal control mechanisms addressing the mean and variance of stochastic processes have been derived in a variety of contexts (see, for example, [42] and citations therein), particularly in financial applications [43], and often find tradeoffs between controlling the mean and variance of a process. We observe a comparable tradeoff, that tight control on moments of one species leads to loose control on another. We have focussed on providing a general theoretical formalism with which to treat any given control mechanism; it is anticipated that the above treatment may also be of value in describing heterogeneity in other systems where replication and/or death rates of individuals depend on feedback from current numbers of individuals (for example, through terms describing competition for resources in ecology). Within the context of mtDNA populations, we anticipate that this theoretical framework will assist in understanding natural processes of mtDNA inheritance and evolution within an organismal lifetime (including segregation and increasing variance with age) [14, 23], and informing applied approaches to control mitochondrial behaviour with genetic tools [4, 5].

Appendix

A1. Interpretation of relaxed replication.

The relaxed replication model [19, 1] describes a cellular population of mtDNA molecules according to the following algorithm. MtDNAs randomly degrade as a Poisson process with rate $1/\tau$. Every timestep $\Delta t$, the value of $C(w, m)$, a deterministic function of $w$ and $m$, is computed, then $\Delta t C(w, m)$ mtDNAs are added to the population. The genetic properties of these added mtDNAs are random – each is assigned a genetic type based on a random sampling of the existing populations – but their physical properties (i.e. the total copy number added at each step) is deterministic. We argue that, as both replication and degradation of mtDNAs depend on complicated behaviour and thermal, microscopic interactions, it makes more sense to model both processes as stochastic. Thus, $C(w, m)$ is interpreted as the rate of a Poisson process describing replication, just as $1/\tau$ is the rate of a Poisson process describing degradation. This interpretation reconciles the nature of the two processes.

Although this feature is less interesting than the underlying scientific behaviour, the original algorithm also raises a (not insurmountable) technical problem with implementation. If a timestep $\Delta t < 1/C(w, m)$ is chosen, the algorithm will never add any mtDNA to the system. But, in order to suitably characterise the stochastic degradation without using the Gillespie algorithm [44], it is desirable to choose a timestep $\Delta t$ as low as possible. There is therefore a risk that one or other of the deterministic replication and stochastic degradation processes is inadequately captured in a given simulation protocol.

We have illustrated the excellent agreement between our theoretical approaches and stochastic simulation (for example, Fig. 2). To quantitatively connect with previous analyses of specific control strategies, we confirm that the behaviour for model A (relaxed replication) matches that observed in previous simulation studies [19] with a back-of-the-envelope calculation [39].

The rate of variance increase with time with $\tau = 5$ (comparable in magnitude to the $(1 − 10) \times \ln 2$ days used in Ref. [19]) and $w_{opt} = 1000$ from Fig. 2 is roughly $40 \text{ day}^{-1}$. Considering 50 years of evolution of this system, we expect a standard deviation of roughly $\sqrt{40 \times 50 \times 365} \approx 850$ in mutant copy number. This value is consistent with the simulations in Ref. [19].

We connect to an additional numerical result in Ref. [1]. In the absence of a mutant population, the variance of the wildtype population was reported to be stable at $w_{opt}/(2\alpha)$, with the original model interpretation of mtDNA replication as deterministic. Under the interpretation of stochastic replication, an absent mutant population ($m_{ss} = 0$) permits stability in the wildtype population variance, which after a little algebra is calculated to be $w_{opt}/\alpha$. Intuitively, modelling both replication and degradation as stochastic does not affect mean copy number but does increase variance.

A2. Interpretation of Wright formula for mtDNA.

The Wright formula

$$\langle h^2 \rangle = \langle h \rangle (1 − \langle h \rangle) \left( 1 − \left( 1 − \frac{1}{n_e} \right)^g \right).$$

(17)
has been proposed as a model for the time evolution of heteroplasmy variance ($\langle h^2 \rangle$) in a population with effective size $n_e$ subject to random partitioning at each of $g$ generations. This picture has been successfully employed to investigate heteroplasmy distributions in real systems [22], with $n_e$ and $g$ interpreted as parameters of the theory without immediate biological interpretation.

The mapping of the original genetic system considered by Wright [21] to cellular populations of mtDNA requires some discussion. If ‘generations’ are interpreted as cell divisions, the mechanism by which mtDNA copy number is redoubled between divisions is assumed by the Wright model to be deterministic. Cell divisions will result in a halving of the mtDNA population. Application of the Wright model assumes that the original population is thenceforth recovered with no increased variance in the population. In other words, the mtDNA population is assumed to exactly double between divisions with no stochasticity in the process. As we underline in the Main Text, the effects of (inevitable) stochasticity due to mtDNA turnover are not explicitly captured by the Wright formula. Other complications exist, as described in [22], but play less important roles here. As a result, the ‘bottleneck size’ $n_e$ cannot immediately be interpreted as an observable minimum cellular copy number of mtDNA molecules (a quantity that is reported by, for example, a qPCR experiment measuring cellular mtDNA content), but rather the size of an effective ‘founder’ population.

References


Supplementary Information

S1. Expansion of the control law.

If we write $\epsilon_w = w - w_{ss}, \epsilon_m = m - m_{ss}$, the expansion of replication and degradation rates $\lambda(w, m)$ and $\nu(w, m)$ about the steady state $\{w_{ss}, m_{ss}\}$ gives

$$
\lambda(w, m) \simeq \lambda(w_{ss}, m_{ss}) + \frac{\partial \lambda(w, m)}{\partial w} \epsilon_w + \frac{\partial \lambda(w, m)}{\partial m} \epsilon_m + O(\epsilon^2) \tag{18}
$$

$$
\nu(w, m) \simeq \nu(w_{ss}, m_{ss}) + \frac{\partial \nu(w, m)}{\partial w} \epsilon_w + \frac{\partial \nu(w, m)}{\partial m} \epsilon_m + O(\epsilon^2). \tag{19}
$$

This expansion represents a model of a given control strategy $\lambda(w, m), \nu(w, m)$, which, if the original function is well behaved, we expect to reasonably reflect behaviour of the system close to $\{w_{ss}, m_{ss}\}$. Simulation results show that this expectation is fulfilled for a wide variety of cases (see figures in the Main Text).

To find $w_{ss}$ and $m_{ss}$ we solve the equations describing the deterministic behaviour of the system:

$$
\frac{dw}{dt}\bigg|_{(w_{ss}, m_{ss})} = w_{ss} \lambda(w_{ss}, m_{ss}) - w_{ss} \nu(w_{ss}, m_{ss}) = 0 \tag{20}
$$

$$
\frac{dm}{dt}\bigg|_{(w_{ss}, m_{ss})} = m_{ss} \lambda(w_{ss}, m_{ss}) - m_{ss} \nu(w_{ss}, m_{ss}) = 0 \tag{21}
$$

It will be observed that for this steady state to exist, the terms $\lambda(w_{ss}, m_{ss})$ and $\nu(w_{ss}, m_{ss})$ in Eqns. 18-19 must be equal. We can write the general expansion form of $\lambda(w, m)$ and $\nu(w, m)$, truncated to first order, as

$$
\lambda(w, m) \simeq \beta_0 + \beta_w (w - w_{ss}) + \beta_m (m - m_{ss}), \tag{22}
$$

$$
\nu(w, m) \simeq \delta_0 + \delta_w (w - w_{ss}) + \delta_m (m - m_{ss}), \tag{23}
$$

with $\beta_w = \partial \lambda / \partial w|_{w_{ss}, m_{ss}}, \beta_m = \partial \lambda / \partial m|_{w_{ss}, m_{ss}}, \delta_w = \partial \nu / \partial w|_{w_{ss}, m_{ss}}, \delta_m = \partial \nu / \partial m|_{w_{ss}, m_{ss}}$. Clearly, to support convergence to a steady state, $\beta_w$ and $\beta_m$ must be negative and $\delta_w$ and $\delta_m$ must be positive. Given this model for control dynamics, we next characterise the variance of the system. We can thus describe the system with a set of $R = 4$ processes with rates

$$
f_1 = w (\beta_0 + \beta_w (w - w_{ss}) + \beta_m (m - m_{ss})) \tag{24}
$$

$$
f_2 = m (\beta_0 + \beta_w (w - w_{ss}) + \beta_m (m - m_{ss})) \tag{25}
$$

$$
f_3 = w (\delta_0 + \delta_w (w - w_{ss}) + \delta_m (m - m_{ss})) \tag{26}
$$

$$
f_4 = m (\delta_0 + \delta_w (w - w_{ss}) + \delta_m (m - m_{ss})) \tag{27}
$$

and stoichiometry matrix describing the effects of these reactions on the $N = 2$ species we consider as

$$
S = ((1, 0), (0, 1), (-1, 0), (0, -1))^T. \tag{28}
$$

Using index $i = 1$ to correspond to species $w$ and $i = 2$ to correspond to species $m$, the master equation for the system, describing the time evolution of $P_{w,m}$ (the probability of observing $w$ wildtype and $m$ mutant mtDNAs) can then be written

$$
\frac{\partial P_{w,m}}{\partial t} = \sum_{j=1}^{R} \left( \prod_{i=1}^{N} \mathbb{E}^{-S_{ij}} - 1 \right) f_j(w, m) P_{w,m} \tag{29}
$$

where $\mathbb{E}^{-S_{ij}}$ takes its normal meaning as a raising and lowering operator [24], adding $-S_{ij}$ to each occurrence of index $i$ that follows it on the right (so e.g. as $S_{11} = 1$ and $w$ corresponds to index 1, $\mathbb{E}^{-S_{11}} f_j(w, m) P_{w,m} \rightarrow f_j(w-1,m) P_{w-1,m}$).

The potential nonlinearities and coupling between species in this equation prevents a full general solution. To make progress, we employ Van Kampen’s system size expansion [24, 25] and write $w = \phi_w \Omega + \xi_w \Omega^{1/2}$, $m = \phi_m \Omega + \xi_m \Omega^{1/2}$,
representing copy numbers as the sum of deterministic components \( \phi_i \) and fluctuation components \( \xi_i \) scaled by powers of system size \( \Omega \). Following the standard expansion procedure, by writing \( E, P_{w,m} \) and \( f_i \) in terms of \( \Omega \) and collecting powers of \( \Omega \) in Eqn. 29, first gives equations for the deterministic components of the system (corresponding straightforwardly to the macroscopic rate equations):

\[
\frac{\partial \phi_i}{\partial t} = \sum_{j=1}^{R} S_{ij} f_j,
\]

then gives a Fokker-Planck equation for the time behaviour of the fluctuation components in terms of the bivariate probability distribution \( \Pi(\xi, t) \) of \( \xi = (\xi_w, \xi_m) \) at time \( t \):

\[
\frac{\partial \Pi(\xi, t)}{\partial t} = \sum_{i,j=1}^{N} A_{ij} \frac{\partial(\xi \Pi)}{\partial \xi_i} + \frac{1}{2} \sum_{i,j=1}^{N} B_{ij} \frac{\partial^2 \Pi}{\partial \xi_i \partial \xi_j},
\]

where

\[
A_{ij} = \sum_{k=1}^{R} S_{ik} \frac{\partial f_k}{\partial \xi_i},
\]

\[
B_{ij} = \sum_{k=1}^{R} S_{ik} S_{jk} f_k.
\]

The form of Eqns. 24-27 and Eqn. 28 gives, for steady state copy numbers and \( \delta_0 = \beta_0, A_{11} = \kappa_w w_{ss}, A_{12} = \kappa_m A_{21} = \kappa_w m_{ss}, A_{22} = \kappa_m m_{ss}, B_{11} = 2\beta_0 w_{ss}, B_{22} = 2\beta_m m_{ss}, B_{12} = B_{21} = 0, \) where \( \kappa_w = (\beta_w - \delta_w), \kappa_m = (\beta_m - \delta_m). \) From this Fokker-Planck equation expressions for the moments of \( \xi_i \) can be extracted [24], leading to the expressions:

\[
\frac{d\langle \xi^2 \rangle_w}{dt} = 2 A_{11} \langle \xi^2 \rangle_w + 2 A_{12} \langle \xi_w \xi_m \rangle + B_{11}
\]

\[
\frac{d\langle \xi_w \xi_m \rangle}{dt} = (A_{11} + A_{22}) \langle \xi_w \xi_m \rangle + A_{12} \langle \xi^2_m \rangle + A_{21} \langle \xi^2_w \rangle + B_{12}
\]

\[
\frac{d\langle \xi^2 \rangle_m}{dt} = 2 A_{22} \langle \xi^2_m \rangle + 2 A_{21} \langle \xi_w \xi_m \rangle + B_{22},
\]

A linear stability analysis of the deterministic ODEs describing mean behaviour is straightforward to perform. Linearising Eqns. 20-21 about \( (w_{ss}, m_{ss}) \) gives

\[
\frac{dw}{dt} \simeq (\beta_w - \delta_w) w_{ss}(w - w_{ss}) + O(w^2) + O(wm)
\]

\[
\frac{dm}{dt} \simeq (\beta_m - \delta_m) m_{ss}(m - m_{ss}) + O(m^2) + O(wm),
\]

from which it is straightforward to see that if \( \kappa_w < 0 \) and \( \kappa_m < 0 \), the mean dynamics of \( w \) and \( m \) respectively are linearly stable. This condition is met for \( w \) and \( m \) in control laws A and E, for \( w \) in B, C, and F, and for neither in D. These specific examples illustrate the principle that if a species is explicitly ‘sensed’ – in the sense that it modulates replication or degradation rate – its mean dynamics can be controlled to be linearly stable. If a species is not explicitly sensed (replication and degradation are not functions of its copy number) then its mean dynamics are not explicitly linearly stable, but may be ‘balanced’, as the corresponding \( \kappa \) term is zero. Perturbations in these unsensed ‘balanced’ variables are neither damped away by control nor guaranteed to explode with time; hence the variables are unconstrained but not explicitly unstable.

**S2. Full solutions for steady state ODEs.**

Eqns. 34-36 can be solved exactly for the \( A_{ij}, B_{ij} \) corresponding to the steady state condition above. The complete solutions for arbitrary \( \langle \xi^2_w \rangle, \langle \xi_w \xi_m \rangle, \langle \xi^2_m \rangle \) at \( t = 0 \) are lengthy and do not allow much intuitive interpretation. For \( \langle \xi^2_w \rangle = \langle \xi_0 \xi_m \rangle = \langle \xi^2_m \rangle = 0 \) at \( t = 0 \) (noiseless initial conditions), we can solve and separate the long-term
behaviour from the transient behaviour. The transients, given by Eqns. 42-44 below, involve terms in \( t' = \exp((\kappa \tau \text{behaviour from the transient behaviour. The transients, given by Eqns. 42-44 below, involve terms in } t' = \exp((\kappa \tau (\kappa mss + \kappa w wss) t)). As the \( \kappa \) are nonpositive (for stability, \( \beta \) are nonpositive and \( \delta \) are nonnegative), \( t' \) is either a constant or an exponentially decaying function of time \( t \); in the cases we consider, it either decays with time (all models except D) or the associated term is always zero (model D).

\[
\begin{align*}
\langle \xi^2 \rangle_w & = F_1^{\text{decay}}(t') + \frac{2\beta_0 m_{ss} w_{ss} (m_{ss} + w_{ss}) \kappa_m^2}{(m_{ss} \kappa_m + w_{ss} \kappa_w)^2} t \\
+ \left( -\beta_0 w_{ss}^2 (w_{ss} \kappa_m^2 + m_{ss} \kappa_m (4 \kappa_w - 3 \kappa_m)) \right)/ (m_{ss} \kappa_m + w_{ss} \kappa_w)^3
\end{align*}
\]

\[
\langle \xi_w \xi_m \rangle = F_2^{\text{decay}}(t') + \frac{(m_{ss} \kappa_m + w_{ss} \kappa_w)^2}{(m_{ss} \kappa_m + w_{ss} \kappa_w)^3}
\]

\[
\langle \xi^2 \rangle_m = F_3^{\text{decay}}(t') + \frac{2\beta_0 m_{ss} w_{ss} (m_{ss} + w_{ss}) \kappa_w^2}{(m_{ss} \kappa_m + w_{ss} \kappa_w)^2} t \\
+ \left( -\beta_0 m_{ss}^2 (m_{ss} \kappa_m^2 + w_{ss} \kappa_w (4 \kappa_m - 3 \kappa_w)) \right)/ (m_{ss} \kappa_m + w_{ss} \kappa_w)^3
\]

where the \( F_1^{\text{decay}} \) functions characterising the transient behaviour are

\[
F_1^{\text{decay}}(t') = \frac{\beta_0 w_{ss}^2 t' (\kappa_m^2 w_{ss} m_{ss} t' + \kappa_m m_{ss} (\kappa_m (t' - 4) + 4 \kappa_w))}{(m_{ss} \kappa_m + w_{ss} \kappa_w)^3}
\]

\[
F_2^{\text{decay}}(t') = \frac{\beta_0 w_{ss} m_{ss} t' (2m_{ss} + w_{ss}) \kappa_m \kappa_w + \kappa_w w_{ss} (t' - 2))}{(m_{ss} \kappa_m + w_{ss} \kappa_w)^3}
\]

\[
F_3^{\text{decay}}(t') = \frac{\beta_0 m_{ss}^2 t' (\kappa_m^2 m_{ss} t' + \kappa_w w_{ss} (\kappa_w (t' - 4) + 4 \kappa_m))}{(m_{ss} \kappa_m + w_{ss} \kappa_w)^3}
\]

Eqns. 39-41 with Eqns. 42-44 then give the full transient behaviour displayed in Fig. 2, decaying to the aforementioned linear behaviour with characteristic timescale \( (\kappa_m m_{ss} + \kappa_w w_{ss}) \).

S3. Interpretation and behaviour of specific control strategies.

The mechanisms described in the Main Text have simple interpretations in the language of stochastic processes. Mechanism A, as discussed, corresponds to the ‘relaxed replication’ picture studied previously (Appendix A1). Mechanism D corresponds to independent birth-death processes acting on both species (analysed in [14]). Mechanism F corresponds to an immigration-death process acting on the wildtype (the dependence of replication rate on \( 1/w \) means that overall production is constant with time), and model C can be regarded as a birth-immigration-death process on the wildtype (analysed in [20]). Both these mechanisms are thus expected to tightly control wildtype behaviour (including controlling variance: immigration-death processes yield a constant steady-state variance), but do not sense (and, therefore, do not apply feedback to) mutant load. We also note that mechanisms F and G are in a sense ‘dual’, in that they apply similar control manifest through replication with rate \( \sim 1/w \) and degradation with rate \( w \) respectively. The previous result that control applied to (a) biogenesis rates and (b) degradation rates yields similar population behaviour is visible in the behaviour of F and G in Fig. 2.

Table 1 gives, for each example control strategy in the Main Text, the corresponding steady state \( \{w_{ss}, m_{ss}\} \) and the expansion terms for the strategy \( \beta, \delta \). The corresponding behaviours of variances, from Eqns. 39-41, are given in Table 2.

In the Main Text we discuss the implications of an increasing extinction probability of one mtDNA type. A non-negligible extinction probability challenges the validity of the linear noise approximation and leads to departure from results derived using the system size expansion. To illustrate this behaviour, we reduce the characteristic timescale of the parameterisations used to explore models A-G in the Main Text, setting \( \tau = 1 \) rather than \( \tau = 5 \), and simulate for a longer time window (see Fig. 8). It will be observed that as extinction probability increases (as \( \sqrt{\langle m_s \rangle} \rightarrow \langle m \rangle \)), the numerical behaviour departs from that predicted analytically; in particular, the increase of \( \langle h^2 \rangle \) slows from a linear to sublinear regime as discussed in the Main Text.
Figure 8: Influence of fixation on expansion analysis. The models from the Main Text, simulated for a longer time window and for a shorter characteristic timescale \( \tau \), illustrating the behaviour of the systems when extinction becomes possible. \( P_{m0} \) gives the numerically computed probability that \( m = 0 \); it can be seen that an increase in this quantity corresponds to a moderate increase of \( \langle m \rangle \) and \( \langle m^2 \rangle \) relative to their predicted values, and a decrease of \( \langle h^2 \rangle \) relative to its predicted value (shifting towards a sublinear increase as discussed in the Main Text).

We note that, in some physiological circumstances, the representation of one mtDNA type in a cellular population may be low – for example, the appearance of one mutant mtDNA through \textit{de novo} mutation or replication error, or the presence of a small percentage of a foreign mtDNA haplotype due to carryover in gene therapies [6]. In these cases, a non-negligible extinction probability may occur quickly and the transition of \( \langle h^2 \rangle \) to a sublinear, or flat, regime will be an important aspect of the long-term dynamics. In the case of mtDNA disease inheritance, however, situations with a macroscopic fraction of mutant mtDNA are often the most important, due to the presence of a ‘heteroplasmy threshold’ [7] beyond which disease symptoms manifest. With two mtDNA haplotypes represented in comparable proportions in the cell, our linear analysis holds and can be used to describe heteroplasmy variance in somatic and germline cells.

S4. Steady state solution.

Eqns. 34-36 give, for steady state,
### Table 2: Post-transient time behaviour of copy number and heteroplasmy variances in control strategies A-G.

<table>
<thead>
<tr>
<th>Control</th>
<th>Time coefficient of ((\xi^*_A))</th>
<th>Time coefficient of ((\xi^*_B))</th>
<th>Time coefficient of ((\xi^*_C))</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>(\frac{1}{\tau} )</td>
<td>(\frac{w_{opt}}{\tau} )</td>
<td>(\frac{w_{opt}}{\tau} )</td>
</tr>
<tr>
<td>B</td>
<td>(\frac{w_{opt}}{\tau} )</td>
<td>(\frac{w_{opt}}{\tau} )</td>
<td>(\frac{w_{opt}}{\tau} )</td>
</tr>
<tr>
<td>C</td>
<td>(\frac{w_{opt}}{\tau} )</td>
<td>(\frac{w_{opt}}{\tau} )</td>
<td>(\frac{w_{opt}}{\tau} )</td>
</tr>
<tr>
<td>D</td>
<td>(\frac{w_{opt}}{\tau} )</td>
<td>(\frac{w_{opt}}{\tau} )</td>
<td>(\frac{w_{opt}}{\tau} )</td>
</tr>
<tr>
<td>E</td>
<td>(\frac{w_{opt}}{\tau} )</td>
<td>(\frac{w_{opt}}{\tau} )</td>
<td>(\frac{w_{opt}}{\tau} )</td>
</tr>
<tr>
<td>F</td>
<td>(\frac{w_{opt}}{\tau} )</td>
<td>(\frac{w_{opt}}{\tau} )</td>
<td>(\frac{w_{opt}}{\tau} )</td>
</tr>
<tr>
<td>G</td>
<td>(\frac{w_{opt}}{\tau} )</td>
<td>(\frac{w_{opt}}{\tau} )</td>
<td>(\frac{w_{opt}}{\tau} )</td>
</tr>
</tbody>
</table>

**Notes:**
- Table entries are calculated based on the provided equations in the document.
- The equations are complex and involve time coefficients, variable parameters, and functions specific to each control strategy.

The table outlines how different control strategies impact the copy number and heteroplasmy variances over time, which is crucial for understanding the dynamics of genetic inheritance and stability in cellular processes.
processes above, and the selective pressures (\(\epsilon\) mutant biogenesis, increasing wildtype degradation, and increasing mutant degradation. We thus have two ways of provoking selective advantages in each case: increasing wildtype biogenesis, increasing differences manifest either through replicative or degradation advantages (or both) are incorporated with even-indexed \(\epsilon_i\) (providing multiplicative changes) and odd-indexed \(\epsilon_i\) (providing additive changes). We thus have two ways of provoking selective advantages in each case: increasing wildtype biogenesis, increasing mutant biogenesis, increasing mutant degradation, and increasing mutant degradation.

Fig. 9 shows example trajectories arising from each of our control mechanisms in the presence of the mutation processes above, and the selective pressures (\(\epsilon_3, \epsilon_4, \epsilon_5, \epsilon_6\) that favour mutant mtDNA. An excellent agreement between ODE theory and stochastic simulation is again illustrated, and there is substantial similarity between the behaviours caused by selection (favouring mutant mtDNA) and mutation (producing mutant mtDNAs). In several cases (mechanisms B, C, F, G), \(\langle m \rangle\) and \(\langle m^2 \rangle\) simply increase exponentially with time under favourable selective or mutational pressures; this situation straightforwardly gives rise to a sigmoidal change in heteroplasmy (\(h\) \(\sim 1/(1 + e^{-\Delta f(1 - h_0)/h_0})\), with \(\Delta f\) an effective selective difference, as used in previous work [45, 23]. In mechanisms coupling wildtype and mutant content (A and E), mutant increase is slower and accompanied by a decrease in wildtype, attempting to keep total copy number constant. In these circumstances, variance behaviour can be more complex: for example, under relaxed replication with pressure favouring mutant mtDNA, \(\langle w^2 \rangle\) initially increases then subsequently decreases as \(\langle w \rangle\) decreases in magnitude. Mechanism D, where control does not couple mutant and wildtype, has correspondingly perpendicular trajectories in (\(\langle w\rangle, \langle m \rangle\)) space under different selective pressures, but the coupling action of the mutation operations lead to curved trajectories under mutational influence.

\[
2w_{ss}(\beta_0 + \kappa_m \langle \xi w \xi_m \rangle + \kappa_w \langle \xi_w^2 \rangle) = 0 \tag{46}
\]

\[
w_{ss}\kappa_m \langle \xi_m^2 \rangle + m_{ss}\kappa_w \langle \xi_w^2 \rangle + (m_{ss}\kappa_m + w_{ss}\kappa_w) \langle \xi_w \xi_m \rangle = 0 \tag{47}
\]

\[
2m_{ss}(\beta_0 + \kappa_m \langle \xi_m^2 \rangle + \kappa_w \langle \xi_w \xi_m \rangle) = 0 \tag{48}
\]

Attempting to solve these equations for \(\langle \xi_w^2 \rangle\), then \(\langle \xi_w \xi_m \rangle\), then \(\langle \xi_w \rangle\) first gives \(\langle \xi_w^2 \rangle\) = \((-\beta_0 - \kappa_m \langle \xi_w \xi_m \rangle) / \kappa_w\), then \(\langle \xi_w \xi_m \rangle\) = \((\beta_0 m_{ss} - \kappa_m w_{ss} \langle \xi_m^2 \rangle) / (w_{ss} \kappa_w)\), leaving Eqn. 47 reduced to

\[
\frac{2\beta_0 m_{ss}(m_{ss} + w_{ss})}{w_{ss}} = 0, \tag{49}
\]

a condition only fulfilled (due to the non-negativity of \(m_{ss}\) and \(w_{ss}\)) if \(m_{ss} = 0\). If one proceeds through the analysis by first solving for \(\langle \xi_m^2 \rangle\), then \(\langle \xi_w \xi_m \rangle\), then \(\langle \xi_w \rangle\), a symmetric expression is obtained

\[
\frac{2\beta_0 w_{ss}(m_{ss} + w_{ss})}{m_{ss}} = 0. \tag{50}
\]

Eqns. 49-50 illustrate the symmetry in the system: if the copy number of either species is zero then a situation where variance does not increase is supported (but not inevitable: compare the behaviour of relaxed replication model (A, fixed variance) and the birth-death model (D, increasing variance) in the case of zero mutant population).

The effect of selection, mutations, and replicative errors on mtDNA variances can straightforwardly be included in this analysis. In this general case, we replace Eqns. 1-4 in the Main Text with:

\[
\begin{align*}
\{w, m\} & \xrightarrow{\epsilon_1 + (1+\epsilon_2)w\lambda(w,m)} \{w + 1, m\} \tag{51} \\
\{w, m\} & \xrightarrow{\epsilon_3 + (1+\epsilon_4)m\lambda(w,m)} \{w, m + 1\} \tag{52} \\
\{w, m\} & \xrightarrow{\epsilon_5 + (1+\epsilon_6)w\eta(w,m)} \{w - 1, m\} \tag{53} \\
\{w, m\} & \xrightarrow{\epsilon_7 + (1+\epsilon_8)m\eta(w,m)} \{w, m - 1\} \tag{54} \\
\{w\} & \xrightarrow{w\mu_1} \{w - 1, m + 1\} \tag{55} \\
\{w, m\} & \xrightarrow{w\mu_2} \{w, m + 1\} \tag{56} \\
\{w\} & \xrightarrow{w\mu_3} \{w - 1, m + 2\}. \tag{57}
\end{align*}
\]

Here, we have added processes corresponding to spontaneous mutation of a given wildtype mtDNA (\(\mu_1\)), and two types of replicative error affecting wildtype mtDNA, giving rise to one (\(\mu_2\): original molecule remains intact, new molecule is mutated) and two (\(\mu_3\): both original and new molecules are mutated) mutant mtDNAs respectively. Differences manifest either through replicative or degradation advantages (or both) are incorporated with even-indexed \(\epsilon_i\) (providing multiplicative changes to the bare rates) and odd-indexed \(\epsilon_i\) (providing additive changes). We thus have two ways of provoking selective advantages in each case: increasing wildtype biogenesis, increasing mutant biogenesis, increasing wildtype degradation, and increasing mutant degradation.
Figure 9: MtDNA copy number and variability under mutational and selective advantages for mutant mtDNA. Mean, variance, and CV trajectories with (i) selection pressures, and (ii) mutation rates favouring mutant mtDNA, under models A-G from the text. Some control strategies (A, E) keep mutant relatively bound but sacrifice wildtype and provoke large increases in variability; others (B, C, F, G) focus on wildtype stability, allowing mutant to grow unbound. Labels give the control model (letter) and the parameter varied ($\epsilon_3 = 20; \epsilon_4 = 1; \epsilon_5 = 20; \epsilon_6 = 1$ for (i); $\mu_1 = 0.1; \mu_2 = 0.1; \mu_3 = 0.1$ for (ii)); all other $\epsilon, \mu$ parameters are set to zero. Results are shown for theory (lines) and stochastic simulation (points), progressing from an initial condition with $w_0 = 900, m_0 = 100$ with the parameterisations in Fig. 2.
S5. Heteroplasmy.

For a general function \( h = h(x, y) \),

\[
\langle h^2 \rangle = \langle (h - \langle h \rangle)^2 \rangle.
\]  (58)

We will consider an expansion about \((x_0, y_0)\), a state such that \( h(x_0, y_0) = \langle h \rangle \). Using the first-order Taylor expansion of \( h(x, y) \) around \((x_0, y_0)\):

\[
\langle h^2 \rangle = \langle (h - \langle h \rangle)^2 \rangle
\]

\[
\simeq \left( \langle h(x_0, y_0) + (x - x_0) \frac{\partial h}{\partial x} \bigg|_{(x_0,y_0)} + (y - y_0) \frac{\partial h}{\partial y} \bigg|_{(x_0,y_0)} - h(x_0, y_0) \right)^2.
\]  (59)

\[
= \langle (x - x_0)^2 \left( \frac{\partial h}{\partial x} \bigg|_{(x_0,y_0)} \right)^2 + (y - y_0)^2 \left( \frac{\partial h}{\partial y} \bigg|_{(x_0,y_0)} \right)^2 + 2(x - x_0)(y - y_0) \left( \frac{\partial h}{\partial x} \bigg|_{(x_0,y_0)} \frac{\partial h}{\partial y} \bigg|_{(x_0,y_0)} \right) \rangle
\]

\[
= \langle x^2 \rangle \left( \frac{\partial h}{\partial x} \bigg|_{(x_0,y_0)} \right)^2 + \langle y^2 \rangle \left( \frac{\partial h}{\partial y} \bigg|_{(x_0,y_0)} \right)^2 + 2\langle xy \rangle \left( \frac{\partial h}{\partial x} \bigg|_{(x_0,y_0)} \frac{\partial h}{\partial y} \bigg|_{(x_0,y_0)} \right).
\]  (61)

We now consider \( h(x, y) = x/y \), so that

\[
\langle h^2 \rangle \simeq \left( \langle x^2 \rangle \frac{1}{y^2} + \langle y^2 \rangle \frac{x^2}{y^4} - 2\langle xy \rangle \frac{x}{y^3} \right)
\]

\[
= \left( \frac{x^2}{y^2} \left( \langle x^2 \rangle \frac{x}{y^2} + \frac{\langle y^2 \rangle}{y^2} - \frac{2\langle xy \rangle}{xy} \right) \right)
\]  (63)

Finally, given that \( x_0 = \langle x \rangle \) and \( y_0 = \langle y \rangle \), and setting \( x \equiv m \) and \( y \equiv w + m \), we obtain

\[
\langle h^2 \rangle \simeq \frac{\langle m \rangle^2}{\langle w + m \rangle^2} \left( \frac{\langle m \rangle^2}{\langle w + m \rangle^2} + \frac{\langle (w + m)^2 \rangle}{\langle w + m \rangle^2} - 2\frac{\langle m(w + m) \rangle}{\langle m \rangle \langle w + m \rangle} \right)
\]  (65)

To see that exponential growth or decay in one mtDNA type while the other remains constant gives rise to sigmoidal heteroplasmy dynamics, consider (without loss of generality) \( m = h_0n_0e^{\beta t}, w = (1 - h_0)n_0 \), where \( n_0 \) is an initial population size which will cancel. Then, as \( m \) (and hence \( n = m + w \)) increases with time,

\[
h = \frac{m}{m + w} = \frac{h_0n_0e^{\beta t}}{n_0(h_0e^{\beta t} + (1 - h_0))} = \frac{1}{1 + \frac{h_0}{n_0}e^{-\beta t}},
\]  (66)

as used in Refs. [6] and [14], with \( \beta \) corresponding to a selective pressure (in this derivation, positive \( \beta \) favours mutant mtDNA).

S6. Fokker-Planck terms for nonequilibrium regimes.

The system size expansion approach above can be applied to the general system without employing an expansion of the control strategy about a steady state, by considering the processes

\[
f_1 = w\lambda_n(w, m)
\]  (67)

\[
f_2 = m\lambda_n(w, m)
\]  (68)

\[
f_3 = w\nu_n(w, m)
\]  (69)

\[
f_4 = m\nu_n(w, m)
\]  (70)

If the expansion about steady state is not used, the corresponding terms are
\[ A_{11} = \lambda_w(\phi_w, \phi_m) - \nu_w(\phi_w, \phi_m) + \phi_w(\partial_w \lambda_w(\phi_w, \phi_m) - \partial_w \nu_w(\phi_w, \phi_m)) \]  
\[ A_{12} = \phi_w(\partial_m \lambda_w(\phi_w, \phi_m) - \partial_m \nu_w(\phi_w, \phi_m)) \]  
\[ A_{21} = \phi_m(\partial_w \lambda_m(\phi_w, \phi_m) - \partial_w \nu_m(\phi_w, \phi_m)) \]  
\[ A_{22} = \lambda_m(\phi_w, \phi_m) - \nu_m(\phi_w, \phi_m) + \phi_m(\partial_m \lambda_m(\phi_w, \phi_m) - \partial_m \nu_m(\phi_w, \phi_m)) \]  
\[ B_{11} = \phi_w(\lambda_w(\phi_w, \phi_m) + \nu_w(\phi_w, \phi_m)) \]  
\[ B_{22} = \phi_m(\lambda_w(\phi_w, \phi_m) + \nu_w(\phi_w, \phi_m)) \]  
\[ B_{12} = B_{21} = 0, \]

where \( \partial_x f(\phi_1, \phi_2) \) means \( \frac{\partial f}{\partial x} |_{\phi_1, \phi_2} \). We include the mutational processes in the text by adding \( f_5 = \mu_1 w, f_6 = \mu_2 w, f_7 = \mu_3 w \) and setting the corresponding stoichiometry matrix to

\[ S = ((1, 0), (0, 1), (-1, 0), (0, -1), (-1, 1), (0, 1), (-1, 2))^T. \]

If \( \lambda_w = \lambda_m = \lambda \) and \( \nu_w = \nu_m = \nu \) (no selective differences between mtDNA types), the Fokker-Planck terms become

\[ A_{11} = -\mu_1 - \mu_3 + \lambda(\phi_w, \phi_m) - \nu(\phi_w, \phi_m) + \phi_w(\partial_w \lambda(\phi_w, \phi_m) - \partial_w \nu(\phi_w, \phi_m)) \]  
\[ A_{12} = \phi_w(\partial_m \lambda(\phi_w, \phi_m) - \partial_m \nu(\phi_w, \phi_m)) \]  
\[ A_{21} = \mu_1 + \mu_2 + 2\mu_3 + \lambda(\phi_w, \phi_m) - \partial_w \nu(\phi_w, \phi_m)) \]  
\[ A_{22} = \lambda(\phi_w, \phi_m) - \nu(\phi_w, \phi_m) + \phi_m(\partial_m \lambda(\phi_w, \phi_m) - \partial_m \nu(\phi_w, \phi_m)) \]  
\[ B_{11} = \phi_w(\mu_1 + \mu_3 + \lambda(\phi_w, \phi_m) + \nu(\phi_w, \phi_m)) \]  
\[ B_{22} = (\mu_1 + \mu_2 + 4\mu_3) + \mu(\lambda(\phi_w, \phi_m) + \nu(\phi_w, \phi_m)) \]  
\[ B_{12} = B_{21} = -(\mu_1 + 2\mu_3) \phi_w. \]

Including selection terms (without mutation) requires no change to the original structure of reactions and stoichiometries and immediately gives

\[ A_{11} = (1 + \epsilon_2)\lambda(\phi_w, \phi_m) - (1 + \epsilon_6)\nu(\phi_w, \phi_m) + \phi_w((1 + \epsilon_2)\partial_w \lambda(\phi_w, \phi_m) - (1 + \epsilon_6)\partial_w \nu(\phi_w, \phi_m)) \]  
\[ A_{12} = \phi_w((1 + \epsilon_2)\partial_m \lambda(\phi_w, \phi_m) - (1 + \epsilon_6)\partial_m \nu(\phi_w, \phi_m)) \]  
\[ A_{21} = \phi_m((1 + \epsilon_4)\partial_w \lambda(\phi_w, \phi_m) - (1 + \epsilon_8)\partial_w \nu(\phi_w, \phi_m)) \]  
\[ A_{22} = (1 + \epsilon_4)\lambda(\phi_w, \phi_m) - (1 + \epsilon_8)\nu(\phi_w, \phi_m) + \phi_m((1 + \epsilon_4)\partial_m \lambda(\phi_w, \phi_m) - (1 + \epsilon_8)\partial_m \nu(\phi_w, \phi_m)) \]  
\[ B_{11} = \epsilon_1 + \epsilon_5 + \phi_w((1 + \epsilon_2)\lambda(\phi_w, \phi_m) + (1 + \epsilon_6)\nu(\phi_w, \phi_m)) \]  
\[ B_{22} = \epsilon_3 + \epsilon_7 + \phi_m((1 + \epsilon_4)\lambda(\phi_w, \phi_m) + (1 + \epsilon_8)\nu(\phi_w, \phi_m)) \]  
\[ B_{12} = B_{21} = 0. \]

The same approach as above can be used to obtain Eqs. 34-36 for the time evolution of fluctuation moments, this time valid for a full temporal trajectory of the system.

**S7. Experimental observations to distinguish mechanisms.**

Our theoretical results suggest measurements to further elucidate the control mechanisms underlying mtDNA evolution within cells, without using heteroplasmy variance \( \langle h^2 \rangle \) (the shortcomings of which are manifest because seven different feedback controls all yield the same dynamics in \( \langle h^2 \rangle \) – Fig. 2 ), and in conjunction with further molecular elucidation of processes governing mtDNA [46, 2] which providing bounds on the types and rates of molecular processes involved (for example, disallowing unphysically high rates of mtDNA replication).

If \( \langle w^2 \rangle \) increases with time, mechanisms with weaker constraints on wildtype copy number are more likely (including relaxed replication (A), mechanisms sensing a combination of mutant and wildtype copy number (E), and the case with no feedback (D)). If \( \langle w^2 \rangle \) is low and constant, mechanisms involving differential (B) or ratiometric (C) control are likely. If \( \langle w^2 \rangle \) is high and constant (of the order of \( \langle w \rangle \)), mechanisms resembling immigration-death processes (with propagation scaled by inverse copy number, F and G) are more likely. The behaviour of \( \langle wm \rangle \) can
be used to further distinguish mechanisms which strongly couple wildtype and mutant (including relaxed replication and total copy number control) from those with less coupling.

In all these cases, the likelihood functions associated with specific biological observations will be complicated. Model selection and inference in this case could be performed through comparison to simulation, or using likelihood-free inference [47] for the mean and variance of mtDNA populations [40].


Average cellular mtDNA copy number measurements in Ref. [38] are made by normalising the signal from the mtDNA-encoded \(ND1\) gene by that from the nuclear-encoded \(GADPH\) genes using real-time PCR using iQ Sybr Green on the BioRad iCycler. The published protocol [48] for this technique suggests using 50ng-5pg of genomic DNA. Diploid human cells contain \(\sim 6pg\) of genomic DNA; the mass of several hundred (much smaller) mtDNA genomes is negligible by comparison. The protocol thus implies the presence of 1-10000 cells’ genomic DNA content; we assume 1000 as an estimate consistent with qPCR standards (Joerg Burgstaller, personal communication).

In our analysis of the data from Ref. [38] we use \(\tau = 5\) days and the processes:

\[
\begin{align*}
\{w, m\} & \xrightarrow{\lambda w} \{w + 1, m\} \\
\{w, m\} & \xrightarrow{\lambda m} \{w, m + 1\} \\
\{w, m\} & \xrightarrow{\nu w} \{w - 1, m\} \\
\{w, m\} & \xrightarrow{(1 + \epsilon_8)m\nu} \{w, m - 1\}
\end{align*}
\]  

(93) \hspace{2cm} (94) \hspace{2cm} (95) \hspace{2cm} (96)

with \(\lambda = \nu = 1/\tau\), and \(\epsilon_8\) a selective difference acting to increase degradation of the mutant mtDNA species. We first estimate a value for \(\epsilon_8\) consistent with the heteroplasmy changes involved. Using the transformation

\[
\beta t = \log \left( \frac{h(h_0 - 1)}{h_0(h - 1)} \right),
\]

(97)

from Eqn. 66 above, where \(h_0\) is initial heteroplasmy and \(h\) is heteroplasmy at time \(t\), we obtain an estimate \(\beta = -1.2 \times 10^{-4}\) day\(^{-1}\). We thus set \(\epsilon_8 = 1.2 \times 10^{-4}\) day\(^{-1}\), to produce the required selective difference manifest through mutant degradation.

Solving the ODEs arising from our theoretical approach (Eqns. 34-36) then give values for \(\langle w^2 \rangle\) and \(\langle m^2 \rangle\) over time for a given initial condition. Assuming that each datapoint consists of a sample of \(10^3\) cells, we divide these values by \(10^3\) to obtain an estimated distribution for each later \(w, m\) pair, given the paired initial \(w, m\) state. We combine these distributions to build an overall distribution over later results, and use the Kolmogorov-Smirnov test to test the alternative hypothesis that the later results were incompatible with draws from this distribution. The results were \(p = 0.054\) for wildtype mtDNA copy number and \(p = 0.861\) for mutant mtDNA copy number. As highlighted in the text, the absence of a \(p < 0.05\) result cannot be interpreted as support for the null hypothesis, but this analysis suggests that the available data is not incompatible with the predictions of our model.