Metabolic implications for the mechanism of mitochondrial endosymbiosis and human hereditary disorders

Benjamin Lovegren de Bivort[a],*, Chun-Chung Chen[b], Fabrizio Perretti[b], Giacomo Negro[b], Thomas M. Philip[b], Yaneer Bar-Yam[a,b]

[a]*Department of Molecular and Cellular Biology, Sherman Fairchild Room 337, 7 Divinity Avenue, Harvard University, Cambridge, MA 02138, USA*
[b]*New England Complex Systems Institute, 24 Mt. Auburn Street, Cambridge, MA 02138, USA*

Abstract

The endosymbiosis of proto-mitochondrial prokaryotes (PMP) into proto-eukaryotic host-cells was a major advance in eukaryotic evolution. The nature of the initial relationship remains the subject of controversy. Various conceptual models have been proposed, but none has definitive support. We construct a model of inter-species interactions based upon well-established respiratory pathways, describing the respective energy gain of host-cell and PMP resulting from varying levels of cooperation. The model demonstrates conflicting evolutionary strategies ("Prisoner's Dilemmas") in the interspecies molecular transfers. Nevertheless, we show that coercion and iterated, multilevel selection on both species encourage endosymbiosis. Mutualism is favored if host-cells are significantly more effective than PMPs at gathering food. Otherwise, an unambiguous asymmetry between host-cell and PMP benefits implies that the initial relationship consisted of the host-cell deriving a reproductive advantage at the PMPs' expense—a cellular version of farming. Other initial relationships such as oxygen-detoxification mutualism and parasitism are not strongly supported by the model. We compare the model behavior with experiments on mutant human mitochondria and find the model predicts proliferation rates consistent with that data. We derive from the evolutionary dynamics counter-intuitive therapeutic targets for two human hereditary mitochondrial disorders that reflect the ongoing effect of short-term selection at the mitochondrial level.

1. Introduction

Mitochondria evolved from free-living α-proteobacteria-like prokaryotes (Gray et al., 1999; Margulis, 1970), ca. 1.8 billion years ago (Doolittle et al., 1996; Knoll, 1992; Sicheritz-Ponten et al., 1998). The resulting symbiotic relationship of cells with organelles constituted the defining moment of origin of the extant mitochondrial eukaryotes, if not all eukaryotes (Embley and Martin, 2006). Moreover, the proliferative success of eukaryotes and ultimately their multicellular body plans are due to mitochondrial incorporation (Pfeiffer et al., 2001). Despite the importance of this event in eukaryote evolution—and as a definitive example of coevolution (Rand et al., 2004)—there is no consensus model of the evolutionary mechanisms that drove the establishment of stable symbiotic populations. It was initially thought (Margulis, 1970) that the endosymbiosis was driven by energy-based mutualism because mitochondria are the means by which eukaryotes generate ATP via oxidative respiration. Alternatively, it has been suggested that the host-cell exploited the proto-mitochondrial prokaryotes (PMP) as a food source (Maynard-Smith and Szathmáry, 1998). Recently, it has been argued that the PMP was a parasite unlikely to have yielded any ATP to its host-cell (Andersson and Kurland, 1999). They conclude that the selective advantage of PMP incorporation was to purge the host-cell of poisonous molecular oxygen, which came into atmospheric prevalence around

*Corresponding author.

E-mail addresses: bivort@fas.harvard.edu (B.L. de Bivort), cjj@u.washington.edu (C.-C. Chen), fabrizio_perretti@harvard.edu (F. Perretti), giacomo.negro@unibocconi.it (G. Negro), tmphilip@uclink2.berkeley.edu (T.M. Philip), yaneer@necsi.org (Y. Bar-Yam).

the time of endosymbiosis, as opposed to supplying it with energy. It has alternatively been hypothesized that a hydrogen dependent autotrophic archaebacterium fused with a respiration-capable heterotrophic eubacterium that produced hydrogen as a metabolic byproduct (Martin and Muller, 1998). This provides an immediate selective advantage for symbiosis given these environmental roles. The PMP may also have provided an advantage to the host-cell by metabolizing its excess NADH, thereby liberating the glycolytic pathway to produce ATP and pyruvate (Gest, 1980). Lastly, it may have served to help cycle sulfur metabolites with a sulfur producing archae-bacterium host (Searcy, 1992). An analysis of these possibilities is given by (Embley and Martin, 2006).

The dynamics of mitochondrial evolution has a critical role in several human mitochondrial disorders localized to cells with high energy demands, e.g. neurons and muscles, in which mitochondria vastly over proliferate relative to healthy cells—an intracellular analogy of cancer. In Mitochondrial Myopathy with Encephalopathy Lactaci-dosis and Stroke (MELAS), and Myoclonic Epilepsy with Ragged Red Fibers (MERRF) mitochondrial replication can even cause cell lysis (Lombes et al., 1989). These syndromes appear to be caused by point mutations in transfer RNAs encoded in the mitochondrial genome (MELAS: A3243G tRNA^{Leu}, MERRF: A8344G tRNA^{Lys}) (Larsson et al., 1995; Tanaka et al., 1991). These mutations are heteroplasmic; the severity of the disease phenotype grows with the number of copies of the mutated genomes present in the mitochondrial population (Chinnery et al., 1998; Szuhai et al., 2001). Bentlage and Attardi (1996) observed that mitochondria with mutated genomes have a reproductive advantage over healthy mitochondria in these diseases; hence ongoing mitochondrial selection within the host-cell directly contributes to the disease state.

To investigate the transition of the PMP-host symbiont into a full-blown eukaryote, we modeled the interactions between PMPs and their host-cells, which consist of the exchange of energetic molecules. By following individual classes of molecules through the process of respiration, we calculate the number of high-energy ATP product molecules produced and consumed by both host-cell and PMP (Fig. 1). Since these pathways have been thoroughly investigated biochemically, the values of several model parameters can be inferred from well-established experimental results. These experimental results describe the unique mechanisms and sites of glycolysis, the tricarboxylic acid cycle, and anaerobic and aerobic respiration. While it is unknown whether these precise metabolic conditions existed at the initiation of endosymbiosis, details such as whether glucose was used as a carbon source, or the specific number of ATP molecules generated per carbon source molecule, do not qualitatively affect our model. Therefore, we focus our analysis on those aspects that determine the persistence of symbiosis given a host-cell and PMP that derive energy from the same metabolic precursor and are capable of manipulating the extent to which the precursor
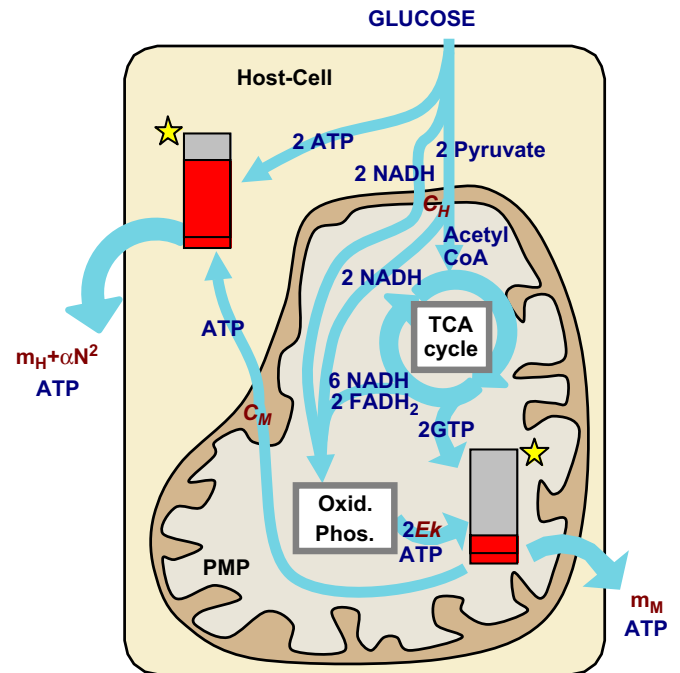

Fig. 1. *Schematic of the metabolic model.* Glucose entering a cell is metabolized via glycolysis into ATP, 2 NADH, and 2 pyruvate molecules. A fraction (C_H) of this pyruvate and NADH enters the PMP where it is converted by oxidative decarboxylation, the TCA cycle and oxidative phosphorylation (Oxid. Phos.) into up to $2E+2$ ATP molecules per pyruvate. A number of PMPs (N) are needed to metabolize all metabolic products in each iteration of the simulation. A portion of the PMP's ATP (C_M) is then transported back to the host-cell cytoplasm. Both cells have "pools" (shown in red) in which ATP is cached, and from which maintenance costs of $\alpha N^2 + m_H$ and m_M are respectively debited. When the total non-maintenance energy costs of the cell-cycle (yellow star) are paid, the cell divides; if the pool empties, the cell dies (is removed from the simulation). The ATP generated by the TCA cycle and oxidative phosphorylation are separated only for the ease of understanding the sources of the additive terms of the equations. They play no distinct role in the subsequent energy pools.

and its energetic products are shared. We first consider metabolic processes assuming equal rates of gathering of energetic molecules. We also consider the possibility of unequal rates of food gathering. Given the metabolic constraints, we ultimately identify two plausible symbiotic regimes: (1) mutualism with asymmetric specialization to food gathering and food metabolizing or (2) coercive farming by host-cells of PMPs.

Once glucose (or equivalent metabolic precursor) is imported into a solitary host-cell, PMP or host containing PMP, glycolysis generates 2 ATP, 2 NADH and 2 pyruvate molecules. In the models we describe in detail below, only PMP were able to convert this pyruvate to acetyl CoA (also generating 2 NADH) and then metabolize acetyl CoA using the tricarboxylic acid (TCA) cycle to generate 2 GTP, 6 NADH, and 2 FADH$_2$, since this process now occurs only in mitochondria (for our results we treated GTP and ATP as equivalent because their energetic yield is similar).

For completeness, we also considered the possibility that before symbiosis, host-cells could utilize the TCA cycle.

This is because the TCA cycle does not strictly depend on the double membrane created by the endosymbiosis of a single-membrane PMP into the host-cell cytoplasm. The enzyme succinate dehydrogenase is partially embedded in the inner mitochondrial membrane and mediates the conversion of succinate to fumarate in the TCA cycle, generating $FADH_2$ which is utilized within the protein to further the electron transport chain in the inner membrane. The dehydrogenase activity of this enzyme is mediated by a domain of the protein that is entirely within the matrix of the mitochondria, and it is plausible that this function could be performed by a protein unbound to the membrane, as long as $FADH_2$ was then trafficked to the membrane. According to this scenario, the TCA cycle might proceed without immediate association to a membrane. Therefore, we also performed simulations in which we made the assumption that both cell types could utilize the TCA cycle (results not shown). This assumption did not qualitatively change any of our results, or the conclusions.

In the case of host-cells containing PMPs, some of the pyruvate and NADH generated by glycolysis is transferred to the encapsulated PMPs. The proportion of host-generated pyruvate and NADH allocated to the mitochondria is C_H, the first parameter whose value is unknown and relevant to endosymbiotic dynamics. Using their allocated portion of pyruvate, the PMPs use oxidative decarboxylation to generate acetyl CoA and two more NADH, and the TCA cycle to generate an additional 6 NADH and 2 $FADH_2$. The pool of NADH and $FADH_2$ is then used to drive oxidative phosphorylation, generating a number of ATP molecules, E, from the cumulative pyruvate and NADH. The amount of ATP generated by oxidative phosphorylation per glucose molecule, i.e. for every 2 NADH and 2 pyruvate transported into the PMP is 2E. While E can vary, a reference value of 16 is appropriate to modern eukaryotes (Gabaldon and Huynen, 2003; Hinkle et al., 1991). E is an important parameter in considering the evolutionary advantages of symbiosis and we will discuss the effect of variations of its value due to evolutionary changes in the efficiency of aerobic respiration. The effects of varying E also provide insights into the relative strengths of various models of the mechanism for mitochondrial endosymbiosis, particularly as regards the sensitivity of our results to model parameters.

The rate at which a single mitochondrion can perform oxidative phosphorylation is limited. We use a parameter N to indicate how many PMPs are needed on average to metabolize the glucose at the rate of ingestion by the host. For convenience, we write the equations in terms of $k = 1/N$.

A fraction C_M of the total accumulated ATP (including those molecules generated by the TCA cycle) is then yielded up to the host-cells. The value of C_M at the time of the endosymbiosis is unknown, and it constitutes the second evolutionarily critical parameter. The proportions C_M and C_H represent aggregate values determined by many genes encoded in both genomes; they represent the levels of cooperation between the symbiotic species.

The ATP in PMPs and host-cells is considered to be in intracellular pools from which "maintenance costs" are debited (m_M and m_H for the PMPs and host-cells, respectively). These costs correspond to the energy needed to maintain the house-keeping functions of the cells. The host-cell pays an additional cost of αN^2 ATP molecules, where N is the number of PMPs it contains. This penalty reflects an indirect cost of harboring PMPs; to the extent that the host-cell must synthesize new proteins or devote new regulatory systems to accommodate the PMPs, these processes will ultimately be paid for with ATP. In particular, this penalty grows non-linearly with respect to the number of PMPs because the effect of additional organelles would be detrimental as their increasing numbers began to spatially interfere with other cellular compartments (the qualitative results of the model are insensitive to this non-linearity exponent over the range of 1.5–5—data not shown). When the total ATP available in one of the intracellular pools reaches thresholds g_M or g_H, the respective PMP or host-cell divides, and thus these values represent the total energy the cells need to implement all phases of the cell-cycle.

2. Results

Using the model in Fig. 1, we performed conventional steady state and game theoretic analysis, and agent-based simulations. Our primary conclusions will be based upon the agent-based simulations, which do not rely upon simplifying assumptions present in the analytical approaches.

More specifically, we (1) derive analytic equations for the growth rates of host-cell and PMP and solve them for steady state conditions; (2) treat these rate equations as payoff values in a 2-player economic game, an analysis that suggests that energetic symbiosis would not have been favorable to both species; and (3) use agent-based simulation that models reproductive dynamics and multi-level selection. All of these analyses attempt to identify conditions that favor symbiosis, i.e. conditions in which host-cells containing PMPs will have a selective advantage over PMPs outside of host-cells and host-cells containing no PMPs. Selection for the symbiont depends on the persistence of PMPs inside host-cells even when independent reproduction of either cell-type is possible. Our analysis reveals that the three analytic techniques yield differing predictions about the stability of symbiosis due to their varying assumptions about the multi-level aspects of symbiotic selection: The game theoretic analysis enables us to conclude that an equilibrium two-player game where each player's choices are independent does not favor the transition to symbiosis. The agent-based simulation reveals biochemically feasible conditions that promote stable symbiosis through an asymmetric relationship analogous to farming.

2.1. Continuous analytic equations

Analytic equations for population growth rates have been successfully applied to several problems in evolutionary biology; the most well known application characterized the dynamics governing predator–prey interactions (Takeuchi, 1996). This approach assumes that population sizes are continuously valued, and that the collective behavior of many individuals can be treated as a single population-level behavior. While biologically we can think about the continuous variable as consistent with the dynamic growth, budding and fusion of mitochondria observed in modern cells such as yeast (Shaw and Nunnari, 2002), the approach is a general strategy for converting discrete time, space and number quantities into continuous equations that can be treated analytically. Under these assumptions, the growth rates of each cell type (r_H and r_M), in terms of number of genome copies gained per glucose imported into the host-cell per existing cell are (A_{in} = ATP produced, A_{out} = ATP consumed, A_{Glyc} = ATP generated by glycolysis, A_{TCA} = ATP generated by the TCA cycle, A_{Oxid} = ATP generated by oxidative phosphorylation, A_{maint} = maintenance cost in ATP, A_{cost} = ATP spent by the host-cell to accommodate the PMPs):

$$r_H = 1 + \frac{1}{g_H}(A_{in} - A_{out}),$$

$$r_H = 1 + \frac{1}{g_H}((A_{Glyc} + A_{TCA} + A_{Oxid}) - (A_{maint} + A_{cost})),$$

$$r_H = 1 + \frac{1}{g_H}(2 + kNC_MC_H(2 + 2E) - m_H - \alpha N^2), \quad (1)$$

$$r_M = 1 + \frac{1}{g_M}(A_{in} - A_{out}),$$

$$r_M = 1 + \frac{1}{g_M}((A_{TCA} + A_{Oxid}) - A_{maint}),$$

$$r_M = 1 + \frac{1}{g_M}(kNC_H(1 - C_M)(2 + 2E) - Nm_M). \quad (2)$$

For a higher-level symbiont to remain stable across generations, the growth rate of the PMPs must match that of the host-cells, i.e. the parameters must have values that satisfy the self-consistent relationship $r_H = r_M$. Subject to this constraint, the evolutionarily favored values of the parameters are those that maximize the growth rates. The only real-valued solution to the constraining steady state condition is given by

$$C_M = \frac{-2g_M + g_Mm_H + 2C_Hg_HkN + 2C_HEg_HkN - g_Hm_HN + \alpha g_MN^2}{2C_H(1 + E)(g_Hg_M)kN}$$
$$(3)$$

If the values of most of these parameters are fixed, one can analytically solve for those values of the remaining parameters that maximize the growth rate. We illustrate how this can be done using a reasonable set of parameters, whose precise values are not essential to the results. The parameters m_H, m_M, g_H, g_M determine the fraction of a

cell-cycle completed per glucose molecule, and thus control the value of the growth rates, but not their qualitative behavior. We assumed that $m_H = m_M$ and $g_H = g_M$ and set them equal to 0.25 and 50 respectively. The equilibrium number of PMPs per host-cell is determined by k: $N = 1/k$. Values of N do not change the qualitative behavior of the model as long as $\alpha N^2 + m_H$ is smaller than the ATP acquired by the host-cell. Unless otherwise specified, we used $k = 0.5$ and $N = 2$. The scaling parameter α does not greatly change equilibrium properties (such as those values of C_H and C_M) that maximize growth rates over the range of 0–2.5, and we set $\alpha = 0.25$. Finally, E was set to 16 because this reflects the current number of ATP molecules that can be generated from a single molecule of pyruvate (Hinkle et al., 1991). While we use $E = 16$ for most of the analyses, since this is a critical parameter and it might be argued that at the time of endosymbiosis the respiratory machinery might not have been as efficient despite prior evolution of the PMP, we vary E below to investigate its influence on the model behavior. The parameter values of the model are summarized in Table 1.

Given these parameters, we can solve for the values of C_H and C_M that maximize the symbiont growth rate: at $C_H = 1$ and $C_M = 0.482$, $r_H = r_M = 1.34$. Biologically, $C_H = 1$ is not possible, as some glycolysis-derived pyruvate is dedicated to non-energetic biosynthesis, but this indicates that the mutual growth rate is maximized in part by maximizing C_H subject to metabolic constraints of biosynthesis. Such a symbiont would not have a selective advantage compared to non-symbiotic competitors, since the growth rate of an isolated PMP is higher, $r_M = 1 + (1/g_M) (2 + k(2 + 2E) - m_M) = 1.375$. On the other hand, the host-cell growth rate without PMPs is lower, $r_H = (1/g_H)(2 - m_H) = 1.035$, indicating that it would only be to the host-cell's advantage to acquire a population of PMPs. However, the growth rate of the isolated PMP is not strikingly greater than that of the PMP in symbiosis, and if E were lower at the time of mitochondrial incorporation, symbiosis might have been favored by both cell-types

Table 1
Fixed and varying model parameters

Fixed parameters	Value
m_H	0.25
m_M	0.25
g_H	50
g_M	50
K	0.5
A	0.25
N (analytic model, game theory)	2

Varied parameters	
E	0–20
C_H	0–1
C_M	0–1
N (agent model)	Varied by simulation

(for example, if $E = 13$, $r_M = 1.26$). These steady-state analytic results are in reasonable agreement with the agent-based simulation described below; however it does little to reveal the mechanisms and dynamics that governed the first endosymbiotic incorporation. A first step in the direction of exploring such dynamics is provided by a game-theoretic treatment.

2.2. Two player game theoretic analysis

As a second approach, we considered a game-theoretic analysis of metabolic benefits for each of the populations. This approach does not require a continuous variable approximation, or that the populations are in a fixed steady-state relationship. The analysis is specifically designed to separate the individual benefits to each of the partners (PMP and host-cell) in the relationship. In the game-theoretic analysis the degrees to which the PMP and host-cell share nutrients (C_M and C_H) can be interpreted as their willingness to cooperate with the other cell (Turner and Chao, 1999). Rationality within the cells is not implied. Mutation will move the cells through the phenotypic space of cooperation levels, and selection will cause populations to "choose" a particular level of cooperation if it is advantageous. However, the *cooperativities* that the cells can adopt reflect strategies in a non-zero-sum game, and allow a game theoretic interpretation (Axelrod, 1984; Maynard-Smith, 1982; Von Neumann and Morgenstern, 1944). If all other parameters are fixed at the same values given above, then associated with every set of cooperativities will be specific net reproductive rates for the host and the PMP (a payoff for each cell type), derivable from Eq. (1) and 2. The payoff matrix per glucose cycle is shown in Table 2a.

The initial effect of the cooperativity choices is very similar to the well-known Prisoner's Dilemma (Fudenberg and Tirole, 1991; Nowak and Sigmund, 1992) for relevant values of C_M and C_H (Tables 2b and c). When both cells cooperate and share molecules, then both can derive an

energetic benefit, but if either player defects by claiming "more than its share" of nutrients, it will gain a reproductive advantage at the expense of the other (Turner and Chao, 1999). As in the Prisoner's Dilemma, mutual defection (non-cooperativity) is a Nash Equilibrium (Nash, 1950) of the game—a combination of strategies at which neither player can improve their payoff with a unilateral move. However, this equilibrium is unstable to drift in PMP cooperativity. In the presence of this drift, there will always be selective pressure on one species to change their level of cooperativity, resulting in a phenotypic trajectory that oscillates between favoring one cell-type and the other. Secondary adaptations that would solidify a symbiotic relationship are unlikely to go to fixation under such circumstances.

Under the assumptions of two-player game-theoretic equilibrium analysis it appears that stable symbiosis could not have arisen because of the selective pressure on PMP and host-cell to share nutrients strictly according to their *individual* short-term reproductive interests. We therefore must release some of these assumptions. For example, the two-player game does not include the impact of host-cell survival on encapsulated PMP survival and assumes that each player is free to invoke its own rational strategy independently of the other player (Fudenberg and Tirole, 1991). However, host death impacts PMP survival, and both the PMP and host-cell could possess the biological means to change the cooperativity of the other cell. The following are actual examples of such coercion: the establishment of the hydrogen ion gradient in the inter-membrane space enables coupled transport of negatively charged pyruvate into the mitochondria (increasing C_H) (Halestrap, 1978), phagolysosomes lyse mitochondria and potentially release their ATP content into the host-cell cytosol (Campbell and Thorsness, 1998) and the ATP exporting translocases are encoded in the host-cell genome (increasing C_M) (Karlberg et al., 2000). For these reasons a more comprehensive model is needed.

Table 2

(a) Energetic payoff as a function of cellular strategies for varying C_H and C_M (given parameter values $m_H = m_M = 0.25$, $E = 16$, k $= 0.5$, $a = 0.25$, $N = 2$, and $g_M = g_H = 50$). The first value of each pair is the net ATP gained by the PMP per glucose; the second is the net ATP gained by the host-cell. For values of C_H and C_M other than 1 or 0, the cell-type-specific ATP yield depends on the parameters C_M and C_H as given by the equations. (b) The payoff matrix for the classical Prisoner's Dilemma with entries equal to proportional benefits. (c) A similar relationship in the host-cell and PMP relationship for $C_H = (0.1, 1)$ and $C_M = (0, 0.42)$

(a)	Host-cell		
PMP	$C_H = 0$	$0 < C_H < 1$	$C_H = 1$
$C_M = 0$	$(-0.5, 0.75)$	$(34C_H - 0.5, 0.75)$	$(33.5, 0.75)$
$0 < C_M < 1$	$(-0.5, 0.75)$	$(34C_H(1-C_M) - 0.5, 34C_HC_M - 0.75)$	$(34(1-C_M) - 0.5, 34C_M - 0.75)$
$C_M = 1$	$(-0.5, 0.75)$	$(-0.5, 34C_H - 0.75)$	$(-0.5, 33.25)$

(b)	Player 1		(c)	C_H	
Player 2	Cooperate	Defect	C_M	1	0
Cooperate	(3,3)	(0,5)	0.43	(15.03, 19.22)	(2.18, 1.42)
Defect	(5,0)	(1,1)	0	(33.5, 0.75)	(2.9, 0.75)

2.3. Agent-based model simulations

To take these effects into account, we implemented an agent model in which multiple cells and their progeny were individually tracked. This third approach to analyzing the energetic model reflects the state of the host/PMP relationship at the time of the first endosymbiosis when PMP were either absent or discretely present, in contrast with the continuous analytical model. It also represents the coupling of host and PMP survival, a key feature of multilevel selection. Starting with a single host-cell containing a single PMP, individual glucose molecules were imported into the host-cell and their metabolites distributed among the host-cell and PMP according to Fig. 1, allowing them to divide when they accumulated at least g_H or g_M ATP (Fig. 2). The PMPs contained in any dividing host-cell were allocated into one of the daughter cells with a probability of $p = 0.5$ for each PMP. A variant of this model, in which PMPs were divided as evenly as possible between daughter cells (and hence deterministically) was also simulated with similar results.

If any host-cell or PMP attained a negative ATP balance in the agent model, it died by removal from the population (Fig. 2b). For all simulations 60 glucose iterations were implemented, allowing the original to proliferate into ~50–25,000 progeny depending on which other parameters were used. After these 60 iterations the net growth rate of the PMP and host-cell populations could be calculated by raising the total number of PMP or host-cells present to the 1/60th power.

Unlike the game theoretic analysis, in which the scale of selection is fixed to that of the low-level components (PMP and host-cell as players), the agent-based model integrates the behaviors of both cell-types composing the symbiont, thereby by reflecting the group-level behavior over time. The effects of multi-level selection coupling host survival to PMP survival can be clearly seen in the growth rates of host-cell and PMP obtained for a range of values of C_H and C_M (Fig. 3). For example, for almost all values of E, and low levels of C_M (and compounded by high C_H), the growth rates imply that both host-cell and PMP go extinct. This is a result of the PMPs hoarding all the ATP they produce, spurring their reproduction to a level that kills off the host-cell, thus killing all symbiotic PMPs.

While the agent model includes no explicit provisions to guarantee that the PMP and host-cell reproductive rates remained equal through time, this behavior is nevertheless observed. In both populations in Fig. 2, the ratio of PMPs to host-cells at intermediate times ($t = 17$) and the end of the simulation are quite close (in a, the PMP to host ratios at these times are 3:3, 13:13 respectively; in b the ratios are 7:3, 24:11 respectively). This parity is not explicitly regulated, but is determined by the group-selective dynamics; those lineages deficient in PMPs reproduce more slowly and are under-represented in subsequent generations, while those lineages with too many PMPs die out.
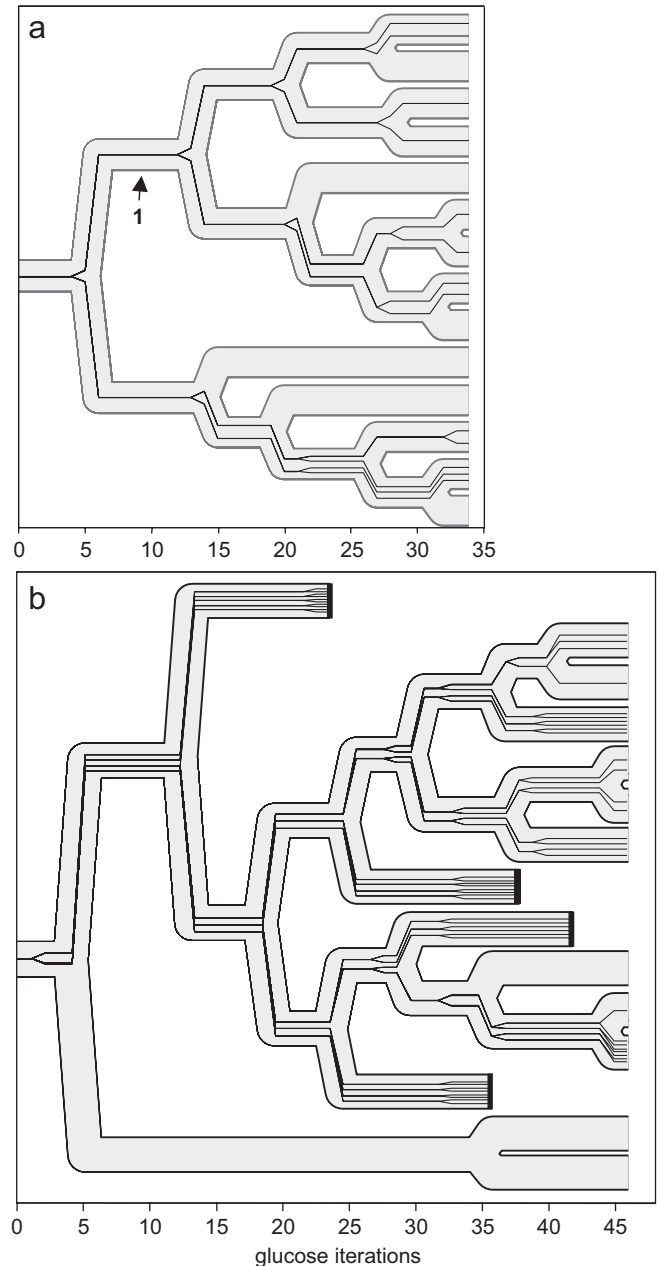

Fig. 2. *Sample lineages from the agent-based model.* Host-cell family trees are shown in gray, PMP family trees are superimposed in black. (a) For the parameters given in Table 2c, with $C_H = 1$ and $C_M = 0.42$, after 34 glucose iterations 14 host-cells and 13 PMP have emerged. It is striking that the lineage (1), which coincidentally had the most even (uniform) PMP allotment, proliferated more than its sister lineage. This indicates that immediately after endosymbiosis there would be strong selective pressure on the cells to evolve mechanisms to equalize the number of PMPs placed in to each daughter cell. (b) With $\alpha = 0.5$ (twice the host-cell cost per PMP) and $g_M = 25$ (twice the PMP division rate), there are many times more PMP per host-cell, which contributes to the death of several host-cells (indicated by a black bar and truncated lineage).

This dynamic is analogous to the stochastic corrector model of Szathmary and Demeter (1987). The remaining lineages, with an intermediate number of PMPs, constitute the viable portion of the population, dynamically bringing about the steady-state condition described in the first
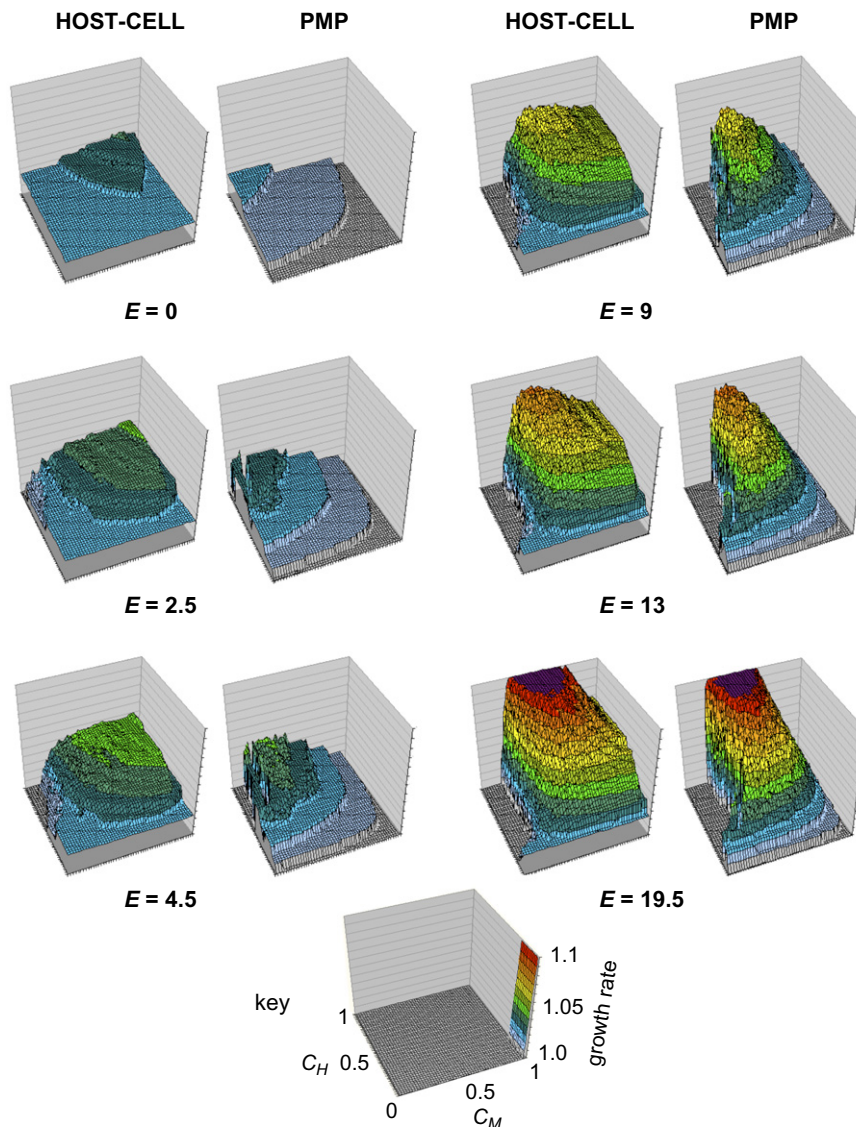
Fig. 3. *Strategic landscapes of host-cell and PMP*. The net growth rates of PMP and host-cell derived over 60 iterations of the agent model (in terms of fractional increase in the number of genomes per iteration) are plotted over all possible values of C_M and C_H, and representative values of E (landscape for $E = 19.5$ is shown as a hypothetical). These surfaces represent the agent model payoff functions for PMP and host-cell, and are comparable to the analytic payoff functions in Table 2. The growth rates of PMPs and host-cells not in symbiotic relationships are calculated as 1.375 and 1.035, respectively. The axis values of each plot are shown in the key at bottom.

analytical model. These forces would furthermore have conferred a strong selective advantage on any symbiont lineage that could actively control the allotment of PMP into daughter cells, thereby minimizing group-selective losses.

One paradoxical aspect of this multiscale selection is that while PMPs with intermediate growth rates foster the stability of symbionts, and ultimately their own population within the symbionts, they are at a short-term selective disadvantage compared to competitor PMPs with lower cooperativity and higher growth rates. Thus, both symbionts are also under selective pressure to evolve mechanisms of controlling the total PMP reproductive rate. The migration of mitochondrial genes into the nucleus (Karlberg et al., 2000; Kurland and Andersson, 2000) is

consistent with the transfer of mitochondrial reproductive control from the PMP to the host-cell.

These dynamics further emphasize the reproductive advantage that the host-cell can gain from PMPs under the assumptions we have used above. However, this advantage does not imply that a relationship of mutualism was favored since the simulation does not allow the alternative of independent PMPs. In order to draw more specific conclusions about the PMP/host-cell relationship at its very beginning, we consider the competition of symbiotic host-PMP communities with independent PMPs. This competition is sensitive to assumptions about the energy efficiency of oxidative phosphorylation. However, at all respiratory efficiencies, the PMPs have lower rates of reproduction in the symbiotic condition than they would as

free-living cells (Fig. 4a). Indeed, since it is likely that the PMPs were able to perform at least a limited form of aerobic respiration and oxidative phosphorylation outside of host-cells (Andersson and Kurland, 1999), the PMPs likely had rapid reproductive rates outside of any host-cells—suggesting that there was no selective advantage of incorporation. We note that if we make the restrictive assumption that PMP were incapable of aerobic respiration before the symbiosis, the symbiotic PMP growth does surpass the non-respiring free-living rate at E greater than 12 (data not shown), i.e. in the regime of current values.

On the other hand, the host-cells have *increased* reproductive rates as long as the total respiratory efficiency slightly surpasses the total energy costs to host and PMP ($E \approx 2$), and as long as the PMP is moderately cooperative. While it is appropriate to assume the free-living PMP could undergo aerobic respiration, the same does not apply to the host-cell, as the most critical functional proteins that bring about aerobic respiration are of strictly mitochondrial origin (Kurland and Andersson, 2000).

3. Discussion

These relationships, along with the fact that 90% of mitochondrial metabolite transport molecules are of eukaryotic origin (Campbell and Thorsness, 1998), provides evidence in favor of the suggestion of (Maynard-Smith and Szathmáry, 1998), that during the transition from PMP to mitochondrion, the host-cell exploited the PMP, manipulating it for its own energetic benefit. This is consistent with a prolonged co-evolutionary struggle culminating in an eventual détente (Smith and Douglas, 1987) brought on by the horizontal transfer of the majority of mitochondrial genes into the host-cell nucleus (Kurland and Andersson, 2000). Indeed, except in the case of parasitism, the PMP can never increase its own growth rate at the expense of the host-cell, because any growth rate greater than that of the host would ultimately interfere with the symbionts' reproductive rate. The growth rate of the PMPs at most values of E slightly surpasses that of the host-cells because of the initial burst of growth within a single host-cell by the single founder PMP until it reaches a stable population of N PMPs per host-cell (Fig. 2b).

We note that fusion of phagolysosomes to yeast mitochondria has been demonstrated as a mechanism by which mitochondrial DNA can be liberated (Campbell and Thorsness, 1998). This mechanism could additionally provide the host-cell a means to release mitochondrially-generated ATP, before the evolution of ATP transport proteins. This possibility is made even more plausible by the following observations: (1) lysosomal fusion proteins were present in early eukaryotes (Pevsner et al., 1996); (2) relatively low levels of mitochondrial ATP yielding (C_M) are sufficient (Fig. 3, 4c) to give the host-cell a reproductive advantage given high C_H; (3) it is to the advantage of host-cells to lyse freshly endocytosed PMPs in both predator/prey and parasite/host relationships, and delaying this
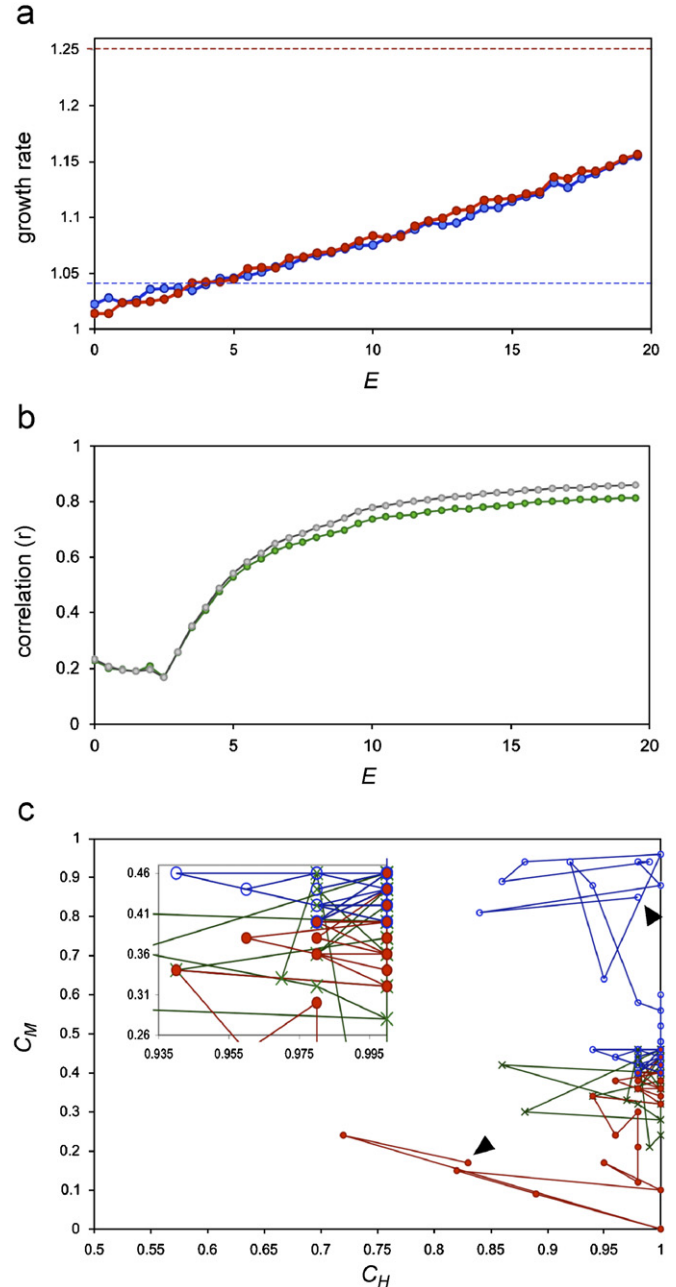

Fig. 4. *Properties of the endosymbiotic relationship for increasing E.* (a) Dashed lines are the growth rates for each cell type (blue for host-cells, red for PMPs) before endosymbiosis. For fixed $C_H = 1$ and $C_M = 0.42$ (values that approximately maximize the growth rates derived by both PMP and host-cell), because the PMP can aerobically respire outside of symbiosis, the host-cell has the greatest potential energetic gain from symbiosis. (b) Plotted are the correlation coefficients between the payoff surfaces shown in Fig. 2, in terms of E. Gray indicates optimal determinist allotment of PMP into daughter cells, green stochastic allotment. As respiratory efficiency (E) increases past 4.5, the benefits derived by PMP and host-cell become increasing correlated. (c) The maxima of the fitness landscapes of PMP and host-cell are plotted in red and blue, respectively, through C_H, C_M parameter space as E increases. Plotted values are the running averages of 4 consecutive points with E ranging from 1.5 (arrowheads) to 18.75 (E > 16 is shown as a hypothetical to illustrate the trends associated with varying E) Green indicates the location of the maxima of the geometric mean of the PMP and host-cell landscapes (a rough indication of where group-level selection would optimize the growth of both populations). The inset is a close-up view of the region of convergence.

fusion is a simple mechanism to allow the PMPs to accumulate ATP aerobically before lysis.

Our model does support energetic mutualism as the driving mechanism that favored endosymbiosis rather than coercion if we allow for different rates of nutrient collection by host-cells and independent PMP. So far, we have assumed that PMPs living autonomously and host-cells are equally capable of retrieving nutrients from the environment. If the host-cell is more efficient at collecting glucose, even when they contain PMPs, the benefit to the individual PMPs of receiving molecules as a captive of the host increases in direct proportion. An increase in nutrient collection by their hosts would increase their rate of growth compared to autonomous ones. In particular, if host-cells can deliver glucose metabolites to captive PMP at a rate 10% greater than their ability to gather nutrients independently, then energetic mutualism is favored (for $E = 16$). Whether or not this is feasible depends both on the rate of host-cell nutrient collection and on the degree to which the host-cell must use its pyruvate for non-respiratory purposes (i.e. amino acid biosynthesis). For example, if $C_H = 1$ were biosynthetically possible, as little as a 10% greater food acquisition by the host-cell would support mutualism. However, if host-cells are biosynthetically constrained to only giving $C_H = 0.8$ of their pyruvate to PMPs, then they would need a glucose acquisition rate 1.38-fold greater than that of independent PMPs in order to deliver the equivalent 10% nutrient enhancement to captive PMPs. This form of mutualism would reflect an inherently asymmetric capability of larger surface area hosts to efficiently gather nutrients, and the smaller captive PMPs to efficiently metabolize them.

As previously mentioned, it is possible that during the period of endosymbiosis, individual PMPs within the host-cell may have exhibited varying levels of cooperative nutrient sharing, resulting in inter-PMP competition. This possibility is implicitly addressed by our agent models. Because the host-cell (and therefore the symbiont) viability is determined by the size of the ATP pool accumulated from all PMPs, over a single generation, a mixed population of PMPs is equivalent to a uniform population with averaged cooperativity. Over multiple generations (and reasonable mutation rates) the small numbers of PMPs in a host-cell would result in drift causing homogeneous PMP populations in a single cell. We will address, however, the short-term effects of such heteroplasmic populations due to mutations resulting in disease states in current cell populations below.

Whether the initial relationship was coercive or mutualistic, the ultimate stability of the symbiosis was reinforced by a convergence in the "strategic interests" of both cell types; the growth-rates derived by the two cell types for all possible strategies (Fig. 3) is correlated except for the lowest values of E (Fig. 4b). To the extent that both cell types derive maximal growth from the same set of strategies, evolution will drive them to these solutions. In particular, the location of the optima of both PMP and host-cell's fitness landscapes converge in C_H, C_M parameter space (Fig. 4c). Indeed, for values of E greater than 6, the global maxima of the strategic interest landscapes of both host-cell and PMP almost coincide. This makes energy sharing, and symbiosis, a Nash Equilibrium of the agent model, provided that the PMP is bound to stay within a symbiotic relationship.

3.1. Implications for mitochondrial syndromes

The set of living mitochondria within a cell forms a population under strenuous selective pressure to conform to group level behaviors that are grossly violated in the MELAS and MERRF disorders. The A3243G tRNA[Leu] and A8344G tRNA[Lys] point mutations present in these syndromes modify the translation of *mitochondrially* encoded genes (Yasukawa et al., 2001), and these mutations result in the paradoxical observation that afflicted cells are filled with mitochondria, but unable to meet their cellular ATP demands (James et al., 1999).

Both syndrome phenotypes—ATP deficiency and mitochondrial proliferation—can be explained parsimoniously by a decrease in C_M. Mutated mitochondria with reduced C_M would horde the ATP they produce, depriving the cell, and speeding their own reproduction. Indeed, the alternative explanation of decreasing E alone, cannot explain these behaviors, since lowering E only lowers mitochondrial growth rate (Fig. 4a). Any population of mitochondria with decreased reproductive rates will be outcompeted by wild-type mitochondria when its phenotype is heteroplasmic; an *increased* reproductive rate is required for mitochondria to contribute to proliferative syndromes.

It has been suggested (Bentlage and Attardi, 1996) that host-cells with mutated mitochondria populations detect their low ATP yield, and feedback mechanisms increase the reproduction of all their mitochondria (compounding the ATP deficiency by sinking more resources into deficient mitochondria). In terms of our model, this is equivalent to a greater investment of mitochondria-produced ATP in mitochondrial reproduction—a decrease in the mitochondrial cooperativity C_M indirectly induced by reduced E.

The details of such a feedback system are unknown. However, our model makes predictions consistent with some types of simple feedback. For example, if a cell were to up-regulate mitochondrial synthesis proportionally to its fraction of mutant alleles—so that if a fraction x of its mitochondria were defective, it would invest in additional mitochondrial growth by the proportion λx, ($C'_M = \lambda x C_M + (1-x)C_M$)—then its effective population-level C_M would be a function of its heteroplasmy. Presuming wild-type mitochondria have $C_M = 0.42$, since this optimizes growth rates of host-cell and PMP, we would then predict that heteroplasmic cells exhibit growth rates intermediate to those derived by mitochondria with $C_M = 0$ and 0.42 depending on λ. Using λ as the only fitting parameter, the rate-adjusted results of the agent model, under this feedback assumption, are quite similar to
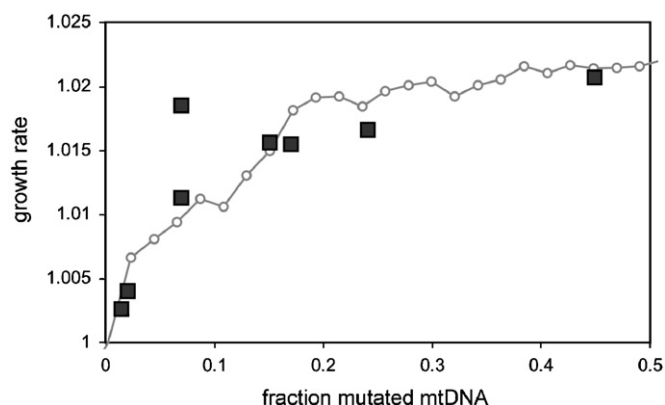
Fig. 5. *Model predictions match in vivo observations.* Theoretical predictions (solid line) of mitochondrial growth rate with increasing mutated mitochondrial DNA (mtDNA) are consistent ($r^2 = 0.94$) with published experimental observations (Bentlage and Attardi, 1996) (shaded squares). Patient isolated mitochondria with varying levels of heteroplasmy were introduced into mtDNA-less cells. By isolate, the percent wild-type mtDNA was derived from Fig. 6a of [12], and the relative mitochondrial growth rate was determined by exponential fit of the curves in Fig. 5 of [12]. The theoretical results are growth rates derived from the agent model (scaled to per-day time units, with $C_H = 1$) versus composite C_M—the weighted average of $C_M = 0.42$ in w.t., and $C_M = 0.29$ in mutant mitochondria. This change in C_M indicates a 45% increase in mitochondrial growth investment for each defective mitochondrion. Along with the adjustment of time scale, this was one of only two variables used to fit the model to the data.

those obtained experimentally by Bentlage and Attardi (1996) when they introduced MELAS patient mitochondrial populations into mitochondrialess culture cells (Fig. 5).

From the host-cell perspective, the feedback induced increase in mitochondria is unlikely to succeed in improving energy supply because the cell would need to specifically reproduce the functional mitochondrial subpopulation. Ironically, it is the *eukaryotic* encoding of mitochondrial import proteins (immune to the mitochondrial translational defects) that will guarantee both mutant and wild-type mitochondria receive their share of pyruvate; E cannot be restored for any number of new mitochondria. This conclusion points to a therapeutic target that might potentially reduce the lysis of muscle cells and neurons due to mitochondrial proliferation, although it would not remedy the deficiency of ATP produced in cells containing mutated mitochondrial genomes. Blocking mitochondrial protein import molecules, such as the Inner and Outer Membrane Transport Complexes, would arrest the reproduction of all mitochondria and potentially reduce the ragged red fiber phenotype. Inhibiting the feedback mechanism itself would have the additional benefit of reducing the futile energy investment in protein production for mitochondrial reproduction. While energy deficiency would remain to the extent cells are heteroplasmic for the mutant allele, and a poorly targeted inhibitor molecule might slow mitochondrial production in otherwise healthy cells, in the most strongly affected neurons and muscle cells, all ATP which would otherwise be invested in

a ten-fold increase in the number of mitochondria could be used for normal cellular functions. Without experimental studies, we cannot be confident that this calculated energetic advantage would provide a net clinical benefit, but established MELAS and MERFF animal models (Clark et al., 1998) provide an ideal environment for testing these hypotheses.

Acknowledgments

References

Andersson, S.G., Kurland, C.G., 1999. Origins of mitochondria and hydrogenosomes. Curr. Opin. Microbiol. 2 (5), 535–541.

Axelrod, R., 1984. The Evolution of Cooperation. Basic Books, New York, NY, USA.

Bentlage, H.A., Attardi, G., 1996. Relationship of genotype to phenotype in fibroblast-derived transmitochondrial cell lines carrying the 3243 mutation associated with the MELAS encephalomyopathy: shift towards mutant genotype and role of mtDNA copy number. Hum. Mol. Genet. 5 (2), 197–205.

Campbell, C.L., Thorsness, P.E., 1998. Escape of mitochondrial DNA to the nucleus in yme1 yeast is mediated by vacuolar-dependent turnover of abnormal mitochondrial compartments. J. Cell Sci. 111 (Part 16), 2455–2464.

Chinnery, P.F., Howell, N., Lightowlers, R.N., Turnbull, D.M., 1998. MELAS and MERRF. The relationship between maternal mutation load and the frequency of clinically affected offspring. Brain 121 (Pt 10), 1889–1894.

Clark, K.M., Watt, D.J., Lightowlers, R.N., Johnson, M.A., Relvas, J.B., Taanman, J.W., Turnbull, D.M., 1998. SCID mice containing muscle with human mitochondrial DNA mutations: an animal model for mitochondrial DNA defects. J. Clin. Invest. 102 (12), 2090–2095.

Doolittle, R.F., Feng, D.F., Tsang, S., Cho, G., Little, E., 1996. Determining divergence times of the major kingdoms of living organisms with a protein clock. Science 271 (5248), 470–477.

Embley, T.M., Martin, W., 2006. Eukaryotic evolution, changes and challenges. Nature 440 (7084), 623–630.

Fudenberg, D., Tirole, J., 1991. Game Theory. MIT Press, Cambridge, MA, USA.

Gabaldon, T., Huynen, M.A., 2003. Reconstruction of the proto-mitochondrial metabolism. Science 301 (5633), 609.

Gest, H., 1980. The evolution of biological energy-transducing systems. FEMS Microbiol. Lett. 7, 73–77.

Gray, M.W., Burger, G., Lang, B.F., 1999. Mitochondrial evolution. Science 283 (5407), 1476–1481.

Halestrap, A.P., 1978. Pyruvate and ketone-body transport across the mitochondrial membrane: exchange properties, pH-dependence and mechanism of the carrier. Biochem. J. 172 (3), 377–387.

Hinkle, P.C., Kumar, M.A., Resetar, A., Harris, D.L., 1991. Mechanistic stoichiometry of mitochondrial oxidative phosphorylation. Biochemistry 30 (14), 3576–3582.

James, A.M., Sheard, P.W., Wei, Y.H., Murphy, M.P., 1999. Decreased ATP synthesis is phenotypically expressed during increased energy demand in fibroblasts containing mitochondrial tRNA mutations. Eur. J. Biochem. 259, 462–469.

Karlberg, O., Canback, B., Kurland, C.G., Andersson, S.G., 2000. The dual origin of the yeast mitochondrial proteome. Yeast 17 (3), 170–187.

Knoll, A.H., 1992. The early evolution of eukaryotes: a geological perspective. Science 256 (5057), 622–627.

Kurland, C.G., Andersson, S.G., 2000. Origin and evolution of the mitochondrial proteome. Microbiol. Mol. Biol. Rev. 64 (4), 786–820.

Larsson, N.G., Tulinius, M.H., Holme, E., Oldfors, A., 1995. Pathogenetic aspects of the A8344G mutation of mitochondrial DNA associated with MERRF syndrome and multiple symmetric lipomas. Muscle Nerve 3, S102–S106.

Lombes, A., Bonilla, E., Dimauro, S., 1989. Mitochondrial encephalomyopathies. Rev. Neurol. (Paris) 145 (10), 671–689.

Margulis, L., 1970. Origin of Eukaryotic Cells. Yale University Press, New Haven, CT.

Martin, W., Muller, M., 1998. The hydrogen hypothesis for the first eukaryote. Nature 392 (6671), 37–41.

Maynard-Smith, J., 1982. Evolution and the Theory of Games. Cambridge University Press, Cambridge, UK.

Maynard-Smith, J., Szathmáry, E., 1998. The Major Transitions in Evolution. Oxford University Press, Oxford, UK.

Nash, J.F., 1950. Equilibrium points in *n*-person games. Proc. Natl Acad. Sci. 36, 48–49.

Nowak, M.A., Sigmund, K., 1992. Tit for tat in heterogenous populations. Nature 355, 250–253.

Pevsner, J., Hsu, S.C., Hyde, P.S., Scheller, R.H., 1996. Mammalian homologues of yeast vacuolar protein sorting (vps) genes implicated in Golgi-to-lysosome trafficking. Gene 183 (1-2), 7–14.

Pfeiffer, T., Schuster, S., Bonhoeffer, S., 2001. Cooperation and competition in the evolution of ATP-producing pathways. Science 292 (5516), 504–507.

Rand, D.M., Haney, R.A., Fry, A.J., 2004. Cytonuclear coevolution: the genomics of cooperation. Trends Ecol. Evol. 19 (12), 645–653.

Searcy, D.G., 1992. In: Matsuno, H.H., Matsuno, K. (Eds.), The Origin and Evolution of the Cell. World Scientific, Singapore, pp. 47–78.

Shaw, J.M., Nunnari, J., 2002. Mitochondrial dynamics and division in budding yeast. Trends Cell Biol. 12 (4), 178–184.

Sicheritz-Ponten, T., Kurland, C.G., Andersson, S.G., 1998. A phylogenetic analysis of the cytochrome *b* and cytochrome *c* oxidase I genes supports an origin of mitochondria from within the Rickettsiaceae. Biochim. Biophys. Acta 1365 (3), 545–551.

Smith, D.C., Douglas, A.E., 1987. The Biology of Symbiosis. Edward Arnold Ltd., London, UK.

Szathmary, E., Demeter, L., 1987. Group selection of early replicators and the origin of life. J. Theor. Biol. 128 (4), 463–486.

Szuhai, K., Ouweland, J., Dirks, R., Lemaitre, M., Truffert, J., Janssen, G., Tanke, H., Holme, E., Maassen, J., Raap, A., 2001. Simultaneous A8344G heteroplasmy and mitochondrial DNA copy number quantification in myoclonus epilepsy and ragged-red fibers (MERRF) syndrome by a multiplex molecular beacon based real-time fluorescence PCR. Nucleic Acids Res. 29 (3), E13.

Takeuchi, Y., 1996. Global Dynamical Properties of Lotka–Volterra Systems. World Scientific Publishing Company, Singapore.

Tanaka, M., Ino, H., Ohno, K., Ohbayashi, T., Ikebe, S., Sano, T., Ichiki, T., Kobayashi, M., Wada, Y., Ozawa, T., 1991. Mitochondrial DNA mutations in mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episodes (MELAS). Biochem. Biophys. Res. Commun. 174 (2), 861–868.

Turner, P.E., Chao, L., 1999. Prisoner's dilemma in an RNA virus. Nature 398 (6726), 441–443.

Von Neumann, J., Morgenstern, O., 1944. Theory of Games and Economic Behavior. Princeton University Press, Princeton, NJ USA.

Yasukawa, T., Suzuki, T., Ishii, N., Ohta, S., Watanabe, K., 2001. Wobble modification defect in tRNA disturbs codon–anticodon interaction in a mitochondrial disease. EMBO J. 20 (17), 4794–4802.