The function of genomes in bioenergetic organelles.

Allen, John

Published in:
Philosophical Transactions of the Royal Society B: Biological Sciences

DOI:
10.1098/rstb.2002.1191

2003

THE ROYAL SOCIETY

Published online 19 December 2002

The function of genomes in bioenergetic organelles

John F. Allen

Plant Biochemistry, Center for Chemistry and Chemical Engineering, Lund University, Box 124, SE-221 00 Lund, Sweden (*john.allen@plantbio.lu.se*)

Mitochondria and chloroplasts are energy-transducing organelles of the cytoplasm of eukaryotic cells. They originated as bacterial symbionts whose host cells acquired respiration from the precursor of the mitochondrion, and oxygenic photosynthesis from the precursor of the chloroplast. The host cells also acquired genetic information from their symbionts, eventually incorporating much of it into their own genomes. Genes of the eukaryotic cell nucleus now encode most mitochondrial and chloroplast proteins. Genes are copied and moved between cellular compartments with relative ease, and there is no obvious obstacle to successful import of any protein precursor from the cytosol. So why are any genes at all retained in cytoplasmic organelles? One proposal is that these small but functional genomes provide a location for genes that is close to, and in the same compartment as, their gene products. This co-location facilitates rapid and direct regulatory coupling. Redox control of synthesis *de novo* is put forward as the common property of those proteins that must be encoded and synthesized within mitochondria and chloroplasts. This testable hypothesis is termed CORR, for *co*-location for *r*edox *r*egulation. Principles, predictions and consequences of CORR are examined in the context of competing hypotheses and current evidence.

Keywords: chloroplasts; mitochondria; photosynthesis; respiration; gene expression; redox regulation

1. INTRODUCTION

Mitochondria and chloroplasts are cytoplasmic organelles of eukaryotic cells. Their primary function is energy transduction. Mitochondria carry out respiration, where energy is provided by differences in chemical redox potential, whereas chloroplasts carry out photosynthesis, where the source of energy is light. Both processes couple vectorial electron transfer (redox chemistry) to ATP synthesis, by means of a transmembrane proton motive force. The primary steps in energy transduction are intrinsic to internal membranes—mitochondrial inner membranes and chloroplast thylakoids—that are homologous with energy-transducing, cytoplasmic membranes of bacteria. Mitochondria and chloroplasts also contain, respectively, associated metabolic cycles of carbon dioxide production (the oxidative tricarboxylic acid pathway) and assimilation (the reductive pentose phosphate pathway). These cycles are coupled tightly to energy transduction but involve soluble intermediates and are catalysed by soluble enzymes. They take place in internal aqueous phases—the mitochondrial matrix and chloroplast stroma—that are homologous with the bacterial cytoplasm.

A prominent feature of mitochondria and chloroplasts is, at first sight, apparently unrelated to their primary function: their internal, aqueous phases also contain discrete, quasi-autonomous, genetic systems. Each organelle's genetic system contains DNA, RNA and all the components necessary both for DNA replication and for synthesis of the proteins encoded therein. Nuclearly encoded components are always required for the operation of an organellar genetic system, but mitochondria and chloroplasts may nevertheless be described as 'quasi-autonomous': they arise only from pre-existing mitochondria and chloroplasts, and they are capable of performing synthesis of some of their own components, even, transiently, after isolation *in vitro*. The proteins synthesized that are not part of the genetic system are a subset of proteins with related functions in electron transport and closely related events in photosynthesis and respiration. A central core of this subset of proteins contains proteins that are universally organelle-encoded. These include membrane-spanning subunits that are central to protein complexes of electron transfer and proton translocation.

Figure 1 is a schematic diagram of generalized animal mitochondrial inner membrane. Figure 2 is the corresponding diagram of part of a generalized chloroplast from a green plant. In both diagrams, protein subunits are presented as cartoon shapes that are coloured in such as way as to indicate the location of the genes encoding them. Inventories of mitochondrial (Lang *et al.* 1999) and chloroplast (Stoebe *et al.* 1998; Martin *et al.* 2002) genes are available and the list of organisms from which they come is growing rapidly (Korab-Laskowska *et al.* 1998). Table 1 presents a summary of the functions of chloroplast genes retained from cyanobacteria.

The retention of protein-coding genes in organelle DNA must be the result of an evolutionary process. The evolutionary mechanisms at work, and the selection pressures for loss and retention of genes, are open and controversial questions. A broad spectrum of contrasting views can be found in the literature and some of these are considered in § 4.

One contribution of 21 to a Discussion Meeting Issue 'Chloroplasts and mitochondria: functional genomics and evolution'.

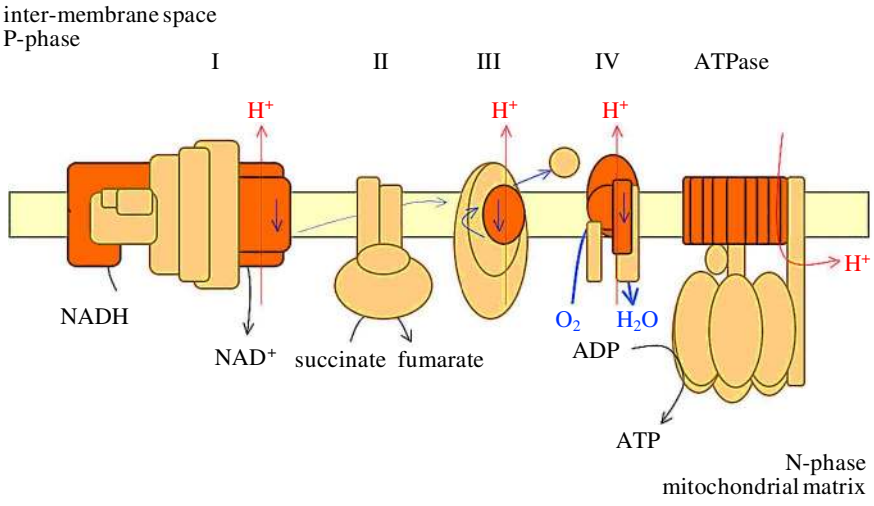

Figure 1. Elements of energy transduction in respiration and oxidative phosphorylation in mitochondria. The mitochondrial inner membrane is shown in yellow. The principal complexes involved in energy transduction are complex I (NADH dehydrogenase), complex II (succinate dehydrogenase), complex III (the cytochrome b–c_1 complex), complex IV (cytochrome c oxidase), and the coupling ATPase. Vectorial electron transfer is depicted as thin, dark-blue arrows. Proton (hydrogen ion; H^+) translocation is depicted as thin, red arrows. Other chemical conversions are given as black arrows. The major, variable environmental input is oxygen (O_2), shown in blue. Subunits of protein complexes are coloured according to the location of the genes encoding them. Mitochondria are usually pink or reddish-brown, the colour of cytochromes and iron–sulphur proteins, so reddish-brown subunits have genes in the mitochondrion and are synthesized in the mitochondrial matrix; light brown subunits have genes in the nucleus, and are imported from the cytosol as precursors. The depiction of sites of synthesis is schematic only and corresponds roughly to the arrangement in vertebrates. Adapted from Ozawa (1995).

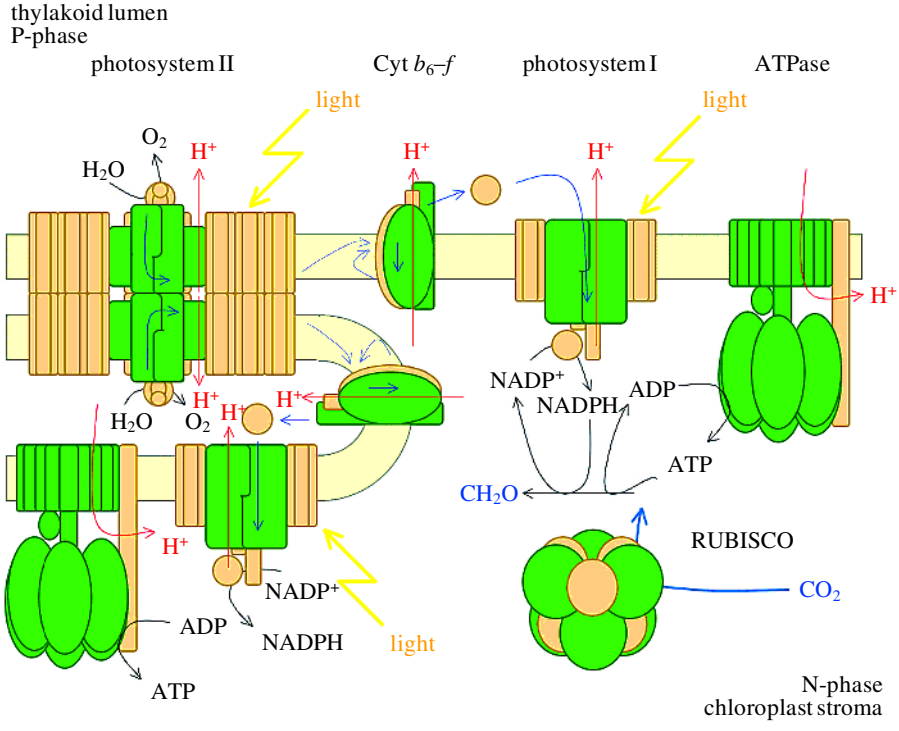

Figure 2. Elements of energy transduction in photosynthesis in chloroplasts. The chloroplast thylakoid membrane is shown in yellow. The principal complexes involved in energy transduction are photosystem II, the cytochrome b_6–f complex (Cyt b_6–f), photosystem I, the coupling ATPase, and RUBISCO. Vectorial electron transfer is depicted as thin, dark-blue arrows. Proton (hydrogen ion; H^+) translocation is depicted as thin, red arrows. Other chemical conversions are given as black arrows. The major, variable environmental inputs are light (orange arrows) and carbon dioxide (blue). Subunits of protein complexes are coloured according to the location of the genes encoding them. Chloroplasts are usually green, so green subunits have genes in the chloroplast and are synthesized in the chloroplast stroma; light brown subunits have genes in the nucleus, and are imported from the cytosol as precursors. Biochemically, green suggests chlorophyll, and the colouring is then counterintuitive: only the reaction centre subunits of photosystems I and II are both chlorophyll-binding and synthesized in the chloroplast. The depiction of sites of synthesis is schematic only and corresponds roughly to the arrangement in land plants. Adapted from Martin & Herrmann (1998) and Race *et al.* (1999).

Table 1. Retention of cyanobacterial genes in chloroplast DNA.
(Protein-coding genes are sorted into the functional categories assigned by Kaneko *et al.* (1996*a*,*b*). The genes are named, and their distribution recorded in 16 completely sequenced chloroplast genomes to date, in Martin *et al.* (2002), from which the data are taken. Electronic Appendix A (available on The Royal Society's Publications Web site) lists the genes, sorted into the functional categories of their gene products. GPC, genes per functional category. The number of genes present, in each category, in the cyanobacterium *Synechocystis* PCC 6803 (Syn) and in any chloroplast genome completely sequenced (cp). Cat Ret, category retention. The number of times that a *Synechocystis* functional category is represented by a gene in cpDNA, divided by the number of genes in that category in *Synechocystis* itself. Cat Ret = cp/Syn. For example, *Synechocystis* has 84 genes for amino-acid biosynthesis, 10 of which are still present in at least one of 16 sequenced plastid genomes. Hence, Cat Ret = 10/84 = 11.9% for amino-acid biosynthesis. Deg Ret, degree of retention. The number of genes present in 16 chloroplast genomes, divided by 16, and multiplied by the number of times that the functional category is represented in chloroplast DNA. For example, in the 16 genomes analysed, there are 10 different amino-acid biosynthetic genes that occur in at least one cpDNA. But of the 160 elements of the resulting matrix that could be positive, only 27 are. Hence, Deg Ret = $27/(16 \times 10) = 16.9\%$ for amino-acid biosynthesis. F(ret), fraction of retention. The number of genes present in 16 chloroplast genomes, divided by 16, and multiplied by the number of times that the functional category is represented in *Synechocystis*. For example, in the 16 genomes sequenced, 27 different amino-acid biosynthetic genes occur in cpDNA. *Synechocystis* has 84 genes for amino-acid biosynthesis that can be assumed to have been present in the common ancestor of these 16 plastid genomes. Hence, for amino-acid biosynthesis, $F(ret) = 27/(16 \times 84) = 2.0\%$. The genes analysed here correspond to those tabulated by Martin *et al.* (2002). Only genes with known homologues in cyanobacterial genomes were included in the analysis. Values were calculated from the supplementary material of Martin *et al.* (2002), which was downloaded from http://www.pnas.org/cgi/content/full/182432999/DCI at the PNAS Web site.)

	no. of GPC				
	Syn	cp	Cat Ret (%)	Deg Ret (%)	F(ret) (%)
amino-acid biosynthesis	84	10	11.9	16.9	2.0
cofactor/prosthetic group biosynthesis	108	21	19.4	20.2	3.9
cell envelope	64	7	10.9	25.0	2.7
cellular processes	68	19	27.9	21.0	5.9
central metabolism	31	1	3.2	6.3	0.2
energy metabolism	86	3	3.5	12.5	0.4
fatty-acid biosynthesis	35	8	22.9	17.2	3.9
photosynthesis and respiration	139	74	53.2	60.3	32.1
photosystem I	12	11	91.7	62.5	57.3
photosystem II	25	20	80.0	80.6	64.5
purine and pyrimidine biosynthesis	39	1	2.6	6.3	0.2
regulatory functions	136	7	5.1	17.8	0.9
DNA replication	49	2	4.1	15.6	0.6
transcription	23	4	17.4	100.0	13.0
translation	144	56	38.9	53.1	20.7
transport	158	2	1.3	6.3	0.1
other plus hypothetical	587	43	7.3	20.2	1.5

(a) *The bacterial nature of mitochondria and chloroplasts*

It is generally accepted that organellar genetic systems are evidence for the evolutionary origin of mitochondria and chloroplasts as free-living eubacteria, or α-proteobacteria, which entered into close inter-dependency with a host cell (Douglas 1998; Delwiche 1999; Lang *et al.* 1999; McFadden 1999, 2001; Roger 1999; Hackstein *et al.* 2001; Stoebe & Maier 2002). The genetic systems retained by organelles have many distinctively bacterial features, and heterologous systems for gene expression can be reconstituted *in vitro* from organellar and bacterial components. By sharp contrast, the genetic systems of the eukaryotic cell nucleus and cytosol are only distantly related to those of bacteria and to those of their own cytoplasmic organelles. Instead, nucleo-cytoplasmic genetic systems more closely resemble the genetic systems of archaea (also known as archaebacteria). It is has been suggested that eukaryotic cells are an archaeal system of information storage, replication, transcription and translation that is powered by a eubacterial system of energy transduction (Martin & Russell 2003). The genetic systems of mitochondria and chloroplasts can then be viewed as eukaryotic cells' residual, bacterial systems of information storage and transmission. Eukaryotic cells' residual, archaeal systems of energy transduction might then include direct vectorial ion pumps such as P- and V-ATPases, and seven-helix transporters now associated primarily with information and signalling. However, the central processes of energy transduction, respiration and photosynthesis, are clearly bacterial in origin.

(b) *Currencies of symbiosis*

One hypothesis for the origin of mitochondria states that the nucleo-cytoplasmic system of eukaryotes is derived from that of an archaeal, methanogenic chemoautotroph that came to rely on the precursor of mitochondria for its supply of molecular hydrogen (Martin & Muller 1998). This 'hydrogen hypothesis for the first eukaryote' resolves a long-held but paradoxical feature of the endosymbiont hypothesis: how could a heterotrophic host cell have been able to supply abundant organic sub-

strates that the bacterial precursor of mitochondria could not already obtain for itself? Furthermore, a second paradox is resolved, or at a least avoided: export of ATP from the bacterial partner was not an immediate requirement for the symbiosis to become established and could have evolved subsequently. Thus, it is not necessary to postulate a suicidal bacterium that excretes ATP, nor a deferred selective advantage to the partnership, implying foresight, or an assurance of eventual contract fulfilment.

The hydrogen hypothesis (Martin & Muller 1998) seeks to provide an explanation of the origin of mitochondria that accounts for the immediate benefit of the primordial symbiosis to each partner, but the implications of the hydrogen hypothesis for the origin of chloroplasts seem not to have been considered in detail. 'What was in it for the proto-chloroplast?' may be an even more acute question than 'what was in it for the proto-mitochondrion?', as, speaking metabolically and energetically, cyanobacteria can do almost anything.

Eukaryotes, whether autotrophic or heterotrophic, have now mostly adopted the aerobic part of the energy metabolism of their original, bacterial constituents, though interesting exceptions are known. Respiration and photosynthesis are antiparallel hydrogen transfer (redox) reactions.

$$CH_2O + A_2 + H_2O \underset{\text{light, photosynthesis}}{\overset{\text{respiration}}{\rightleftharpoons}} 2H_2A + CO_2$$

As in the Van Niel equation for photosynthesis, A may be an organic moiety, for example giving H_2A as succinate; an inorganic element such as sulphur; or even 'nothing', so that H_2A becomes H_2. When A is O (oxygen), it becomes obvious that O_2 in photosynthesis comes from H_2O, not CO_2, and the equation simplifies to:

$$CH_2O + O_2 \underset{\text{light, photosynthesis}}{\overset{\text{respiration}}{\rightleftharpoons}} H_2O + CO_2$$

Phosphate group transfer and electron-transfer-linked ATP synthesis are intermediates in all cases. In respiration, CO_2 is produced from stepwise oxidative decarboxylations of pyruvate dehydrogenase and reactions of the tricarboxylic acid cycle. In most forms of photosynthesis CO_2 is assimilated to form, initially, sugar phosphates.

The mitochondrion exports ATP and CO_2, and imports ADP, inorganic phosphate, an electron acceptor, and pyruvate. Energetically, the preferred electron acceptor is molecular oxygen. Oxygen comes from the environment: cells and cell membranes are freely permeable to oxygen. The primary metabolic division of labour between the chloroplast and the plant cytosol is that the chloroplast exports oxygen and reduced sugar phosphates (triose phosphate), and imports CO_2, less reduced sugars (3-phosphoglycerate), and inorganic phosphate. The light comes directly from the external environment, and, in this respect, the chloroplast thylakoid is the external boundary of the plant cell, where light, both as energy and information, is first received. The CO_2 is also an environmental input, though its supply may be regulated locally by CO_2- or HCO_3^--concentrating mechanisms.

At the cellular level, the mitochondrial inner membrane is the first place where altered O_2 supply exerts its profound metabolic effects. Even granted specialized photoreceptors, the chloroplast thylakoid membrane is first place where changes in light quality and intensity affect a plant's primary input of energy. Temperature changes selectively affect diffusion-limited metabolism involving soluble intermediates. In both locations, abiotic environmental changes therefore have immediate and very far-reaching effects on electron transfer and redox chemistry. The internal membranes of mitochondria and chloroplasts stand at the energetic interface of the cell with the physical environment, and are likely to be first to know about, and respond to, its changes.

2. THE PROBLEM OF CYTOPLASMIC GENETIC SYSTEMS

Given that the genomes of mitochondria and chloroplasts are derived from those of once free-living bacteria, it is clear that most of the bacterial genes that have survived in any form have now been removed to the cell nucleus. If most genes can be moved to the nucleus, then why not any gene? And if any gene can be moved, why not all? Why have mitochondria and chloroplasts retained functional genomes at all? What do the genes of mitochondria and chloroplasts have in common that has resulted in their retention?

Alberts *et al.* (1994, p. 400) succinctly describe the problem posed by the persistence of cytoplasmic genetic systems, as follows:

> Why do mitochondria and chloroplasts require their own separate genetic systems when other organelles that share the same cytoplasm, such as peroxisomes and lysosomes, do not?... The reason for such a costly arrangement is not clear, and the hope that the nucleotide sequences of mitochondrial and chloroplast genomes would provide the answer has proved unfounded. We cannot think of compelling reasons why the proteins made in mitochondria and chloroplasts should be made there rather than in the cytosol.

3. CORR: *CO*-LOCATION OF GENES AND GENE PRODUCTS FOR *R*EDOX *R*EGULATION OF GENE EXPRESSION

This hypothesis states that mitochondria and chloroplasts contain genes whose expression is required to be under the direct regulatory control of the redox state of their gene products, or of electron carriers with which their gene products interact. These genes comprise a core, or primary subset, of organellar genes. The requirement for redox control of each gene in the primary subset then confers a selective advantage upon location of that gene within the organelle instead of in the cell nucleus. Chloroplast and mitochondrial genomes also contain genes for

components of the chloroplast and mitochondrial genetic systems themselves. These genes comprise a secondary subset of organellar genes; genetic system genes. There is generally no requirement for redox control of expression of genes in this secondary subset, though retention of redox control of genetic system genes may, in some cases, allow amplification of redox signals acting upon genes in the primary subset. Retention of genes of the secondary subset (genetic system genes) is necessary for the operation of redox control of expression of genes in the primary subset (bioenergetic genes). Without genes in the primary subset, the function of genes in the secondary subset is eventually lost, and organelles lose their genomes.

This hypothesis of *co*-location for *r*edox *r*egulation of gene expression, CORR, was first outlined in a review (Allen 1992). The hypothesis was then systematically put forward in two articles (Allen 1993*a*,*b*), developed (Allen & Raven 1996), and has been independently reviewed (Race *et al.* 1999). Some of its predictions have been tested experimentally, as described later.

Figure 3 illustrates the general idea of redox regulatory control operating on mitochondrial genes to give the distribution of sites of protein synthesis seen in figure 1. Figure 4 shows the corresponding, predicted pattern of chloroplast redox regulatory control giving rise to the distribution of sites of protein synthesis seen in figure 2. Table 1 shows a functional classification of cyanobacterial genes retained in 16 completely sequenced chloroplast genomes. This analysis (table 1) reveals a strong selective retention of genes for subunits of complexes involved in respiration and photosynthesis, and especially in chloroplast photosystems I and II. The pattern (table 1) seems to be broadly consistent with CORR (figure 4) and inconsistent with alternative hypotheses (see § 4). It is reasonable to view photosynthesis as a process where precise, rapid and coordinated redox regulatory control of gene expression repays the cost of maintaining a chloroplast genetic system.

The CORR hypothesis is based on several propositions, assumptions or principles, 10 of which are stated explicitly and listed below. Some of the principles are already generally accepted; some are shared with competing hypotheses; and some make direct experimental predictions. However, other underlying principles of CORR are themselves general principles, or subsidiary hypotheses, for which no direct proof is available, but for which predictions are made and potentially falsifying observations or results can be imagined.

The 10 principles that form the basis of the CORR hypothesis follow, and each will be critically discussed in the light of available evidence.

(a) *The principle of endosymbiotic origin*

As now generally agreed, bioenergetic organelles have evolved from free-living bacteria.

Although once controversial (Mahler & Raff 1975), this principle will be taken as an axiom and not considered further here. The reader is referred to Gray & Doolittle (1982); Gray (1992); Martin (1999); Martin *et al.* (2001), to § 1, and to accompanying articles in this journal issue.

(b) *The principle of unselective gene transfer*

Gene transfer between the symbiont or organelle and the cell nucleus may occur in either direction and is not selective for particular genes.

This principle takes the form of a testable assertion that seems to be consistent with current knowledge. Nothing certain seems to be known about the mechanism of intracellular gene relocation (Palmer 1997; Martin & Herrmann 1998; Martin *et al.* 1998; Adams *et al.* 2000; Henze & Martin 2001; Millen *et al.* 2001; Rujan & Martin 2001; Daley *et al.* 2002). Plausible scenarios can be imagined on the basis of the classical genetical phenomena of transformation, transduction and transfection. However, as there is no agreed mechanism, there is no basis for asserting a mechanistic argument for selectivity. Gene transfer between the organelle and nucleus seems to be a special case of lateral, or horizontal, gene transfer (Doolittle 1999), and no mechanistic basis seems to be currently available for selectivity in what may be transferred. It should also be pointed out that lateral gene transfer between the organelle and nucleus does not have to occur within a single cell, nor within a multicellular individual (although here any vertically transmitted gene must be incorporated into the germ line), for the outcome of copying genes to novel compartments to be secured. Thus, any plausible mechanism for heterologous recombination may apply equally to gene transfer between compartments. Mitochondrial departures from the 'universal' genetic code and biased codon usage in chloroplasts (see later) may be of selective value in that a barrier is thereby put in place to expression of genes in the 'wrong' compartment. This barrier is unlikely to be insuperable, especially if the 'wrong' compartment turns out actually to be a better location. Impeded relocation might confer stability without being a primary reason for the existence of cytoplasmic genomes.

(c) *The principle of unselective protein import*

There is no barrier to the successful import of any precursor protein, nor to its processing and assembly into a functional, mature form.

As with § 3b, this principle takes the form of a testable assertion that seems to be consistent with current knowledge. The principle is discussed in more detail in § 4, because its negation (some proteins cannot be imported) has previously been put forward, in itself, as a reason for the retention of cytoplasmic genomes. Good hypotheses are easily falsified, and they therefore forbid certain observations. The principle of unselective protein import forbids the existence of a protein, even a synthetic one, that cannot be imported by some means, as a precursor.

(d) *The principle of continuity of redox control*

Direct redox control of expression of certain genes was present in the bacterial progenitors of chloroplasts and mitochondria, and was vital for selectively advantageous cell function before, during and after the transition from bacterium to organelle. The mechanisms of this control have been conserved.

This principle has been tested experimentally in the past few years and seems to be corroborated by an emerging picture of redox signalling in bacteria and organelles. Experiments carried out *in vitro* on the products of protein synthesis (Allen *et al.* 1995; Galvis *et al.* 1998) or RNA synthesis (Pearson *et al.* 1993; Konstantinov *et al.* 1995; Wilson *et al.* 1996*b*) are consistent with a general phenomenon of redox dependency of gene expression in bioenergetic organelles. Redox control of chloroplast transcription

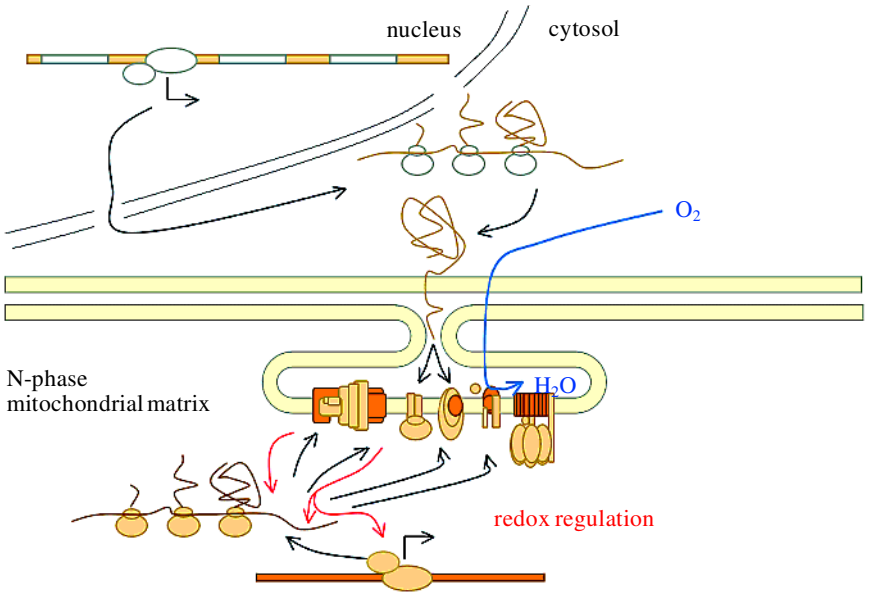

Figure 3. Gene expression and principal pathways of biosynthesis of subunits of protein complexes involved in respiration and oxidative phosphorylation in animal mitochondria. As in figure 1, reddish-brown DNA, RNA and protein subunits are located and synthesized in the mitochondrial matrix; light brown protein subunits have genes (also light brown) in the nucleus, and are imported from the cytosol as precursors. White genes and ribosomal and protein subunits are nuclear–cytoplasmic and of archaebacterial origin. Reddish-brown and light-brown genes and ribosomal and protein subunits are of bacterial origin. The major, variable environmental input is oxygen (blue). The essence of the CORR hypothesis is that it is beyond the ability of the nuclear–cytoplasmic system to respond rapidly and directly to changes in oxygen concentration or partial pressure, and so redox regulation of gene expression (red arrows), has been retained from the ancestral endosymbiont. This redox regulation requires co-location of certain genes, with their gene products, within the mitochondrion.

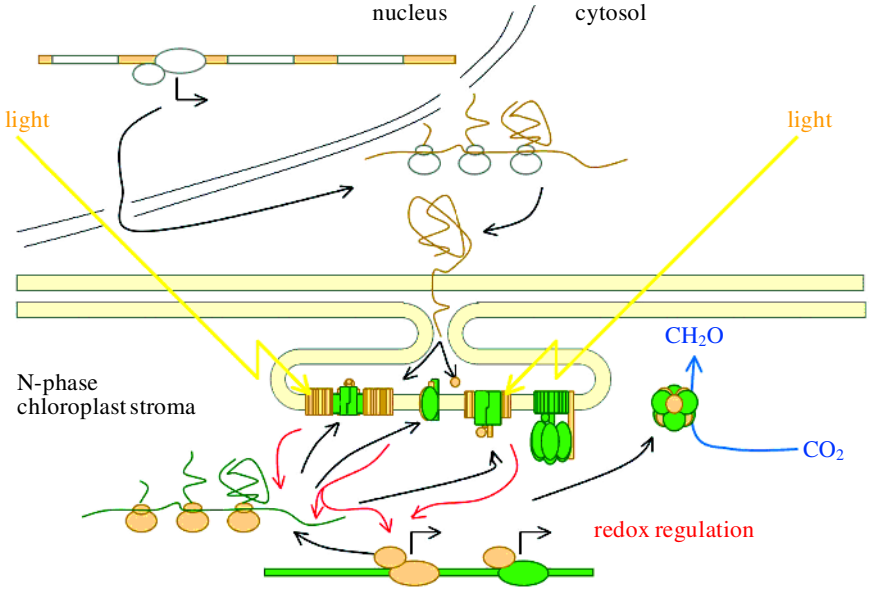

Figure 4. Gene expression and principal pathways of biosynthesis of subunits of protein complexes involved in photosynthesis in chloroplasts. As in figure 2, green DNA, RNA and protein subunits are located and synthesized in the chloroplast; light brown protein subunits have genes (also light brown) in the nucleus, and are imported from the cytosol as precursors. White genes and ribosomal and protein subunits are nuclear–cytoplasmic and of archaebacterial origin. Green and light-brown genes and ribosomal and protein subunits are of bacterial origin. The major environmental inputs are light (orange arrows) and carbon dioxide (blue). The essence of the CORR hypothesis is that it is beyond the ability of the nuclear–cytoplasmic system to respond rapidly and directly to changes in light and carbon dioxide, and so redox regulation of gene expression (red arrows), has been retained from the ancestral endosymbiont. This redox regulation requires co-location of certain genes, with their gene products, within the chloroplast.

has been demonstrated (Pfannschmidt *et al.* 1999*a*,*b*; Allen & Pfannschmidt 2000; Tullberg *et al.* 2000) and was prompted by this principle.

The level of redox control of chloroplast transcription established unequivocally so far is that of plastoquinone, an electron and hydrogen atom carrier located strategically between photosystems I and II of photosynthesis. The directions in which the effects are exerted are functionally intelligible: reduction of plastoquinone increases photosystem I transcription and decreases photosystem II transcription; oxidation of plastoquinone decreases photosystem I transcription and increases photosystem II transcription. Plastoquinone, in normal photosynthesis, is reduced by the action of photosystem II and oxidized by that of photosystem I. Therefore redox control of chloroplast gene expression provides a compensatory response (Pfannschmidt *et al.* 1999*a*), and adds a previously missing part of an explanation for the ability of plants and algae to adjust the ratio of the two photosystems in response to changes in light quality (Chow *et al.* 1990). By the same argument, a functionally equivalent post-translational mechanism of redox control of light-harvesting function by photosystems I and II may have been conserved during the transition from prokaryotes to eukaryotes (Allen 1992, 1995; Allen & Nilsson 1997; Allen & Race 2002).

The precedent of chloroplast redox control of transcription (Pfannschmidt *et al.* 1999*a*,*b*), taken together with the principle in § 3d, predicts plastoquinone redox control of cyanobacterial reaction centre gene transcription, and this is borne out by experiment (El Bissati & Kirilovsky 2001). This complementary, self-adjusting transcriptional response (Fujita 1997; Murakami *et al.* 1997; El Bissati & Kirilovsky 2001), indicates strongly that redox control of photosynthetic reaction centre gene transcription was present in the proto-chloroplast. Thus, the mechanism of regulation is likely to have been conserved throughout the evolutionary transition from bacterial symbiont to cytoplasmic organelle. A conserved mechanism of redox regulation of transcription in photosynthetic prokaryotes (Unden & Bongaerts 1997; Bauer *et al.* 1999) and chloroplasts (Allen 1993*c*; Pfannschmidt *et al.* 1999*b*), is predicted if evolutionary continuity of the process has been maintained.

Thus, apart from being consistent with experimental data, the principle of continuity of redox control has been of heuristic value in pointing to the underlying bacterial character, previously unsuspected, in the control of gene expression, especially of transcription, in chloroplasts of higher plants. Other possible precedents for redox signalling components of cytoplasmic organelles include the prr (Eraso & Kaplan 2000, 2002; Oh & Kaplan 2000) and reg (Bauer *et al.* 1999) redox regulatory systems of purple non-sulphur bacteria. The TspO 'oxygen sensor' of *Rhodobacter sphaeroides* is present in mammalian mitochondria as a protein known there as a benzodiazepine receptor, and each protein will complement loss of the other (Yeliseev & Kaplan 2000). As outlined in § 3j, genes for such components are to be expected primarily in nuclear genomes. This independent evolutionary hypothesis (§ 3d) would be strengthened still further by a demonstration of one or more decisive redox regulatory links shared by photosynthetic prokaryotes and all eukaryotes with mitochondria and, where present, with chloroplasts (Bauer *et al.* 2003).

(e) *The principle of the selective value of redox control*

For each gene under redox control (§ 3d), it is selectively advantageous for that gene to be retained and expressed only within the organelle.

The twin selective advantages of redox control are likely to be energetic efficiency and the suppression of the harmful side effects of electron transport operating on the wrong substrates. No direct experimental support for this principle is available, and, equally, there is no evidence against it. Allotopic mutants, in which genes have been experimentally relocated from organelles to the nucleus, have been produced in yeast, higher plants (Kanevski & Maliga 1994) and human cells (Manfredi *et al.* 2002) and are predicted by CORR to be poor adapters to environmental change. Some authors believe there is potential therapeutic value in allotopic transformation, moving mitochondrial genomes to the nucleus (Kanevski & Maliga 1994; de Grey 2000), and predict the possibility of ρ^- human cell lines in the foreseeable future (de Grey 2000). This principle also has implications for ageing and the evolution of separate sexes (Allen 1996) and possible consequences for development (Blackstone 1999, 2000, 2001; Harvey *et al.* 2002) and somatic cloning in animals (Allen & Allen 1999).

(f) *The principle of the selective value of nuclear location for genes not under redox control*

For each bacterial gene that survives and is not under redox control, it is selectively advantageous for that gene to be located in the nucleus and expressed only in the nucleus and cytosol. If the mature gene product functions in chloroplasts or mitochondria, the gene is first expressed in the form of a precursor for import.

There are several possible reasons for the selective advantage of nuclear location of genes for organellar proteins. One is the decreased probability of mutation arising from free radical by-products of 'incorrect' electron transfers (Raven *et al.* 1994*a*,*b*; Allen & Raven 1996). There would seem to be obvious economies in resource use, and advantages in coordination of gene expression, if genes are kept in one compartment. One might say that the bipartite or tripartite genetic system of eukaryotic cells is untidy and inefficient.

Another suggestion is that organellar genes do not undergo recombination, and are present in relatively small copy numbers, so that disadvantageous mutations will spread relatively quickly through a clonal population of organelles. Incorporation of functional gene copies, instead, into the nuclear genome may then be an irreversible step analogous to the click of a ratchet (Saccone *et al.* 2000). There are examples of genes shared between all three compartments (Stern & Lonsdale 1982; Lonsdale *et al.* 1983).

The cost of retaining an organellar genetic system is not increased appreciably when additional genes are retained in organelles: if just one gene is retained, then the cost of a cytoplasmic genetic system is incurred, and there may be relatively little extra cost for each additional gene retained (Allen & Raven 1996). For a functional explanation of the

existence of chloroplast and mitochondrial genomes, the focus of attention should therefore be on the relatively constant, primary subset of bioenergetic genes (§ 3e).

(g) *The principle of contemporary operation of selection on gene location*

For any species, the distribution of genes between organelle (by the principle in § 3e) and nucleus (by the principle in § 3f) is the result of selective forces that continue to operate.

This principle is essentially a statement contradicting both the 'frozen accident' hypothesis and the hypothesis that gene relocation is progressive, predictable, and still incomplete (§ 4). In support of this principle, it is becoming increasingly clear that mitochondria that lose their function in aerobic respiration also lose their genomes, as seen in the relict mitochondria of microsporidia (Embley & Martin 1998; Van der Giezen *et al.* 2002) and the 'mitosome' of *Entamoeba histolytica* (Tovar *et al.* 1999). In chloroplasts, loss of photosynthesis results in loss of photosynthetic genes. The examples of *Epifagus* (a parasitic higher plant) (Ems *et al.* 1995; Wolfe *et al.* 1992) and the residual apicomplexan plastids of *Plasmodium* (Wilson *et al.* 1996*a*) and *Toxoplasma*, where some plastidic genetic system is retained, mean that some of their gene products must be subject to redox regulation of gene expression if this principle is correct. Experimental test of this principle is possible, using allotopic mutants, as for the predictions of the principle in § 3f. The evidence that the apicoplast (Wilson *et al.* 1996*a*) is a plastid is largely genetic: the residual genome resembles that of photosynthetic plastids (chloroplasts) in gene organization rather than in content. A plastid that has completely lost its genome, by virtue of having lost photosynthesis and all other processes requiring redox regulation, might be difficult to distinguish from an organelle derived in a similar way from a mitochondrion. The persistence of targeted, imported chaperones Hsp70 (Tovar *et al.* 1999; Williams *et al.* 2002) and cp60 (Roger *et al.* 1998) in residual organelles might not distinguish a plastid ancestry from a mitochondrial one.

(h) *The principle of primary involvement in energy transduction*

Those genes for which redox control is always vital to cell function have gene products involved in, or closely connected with, primary electron transfer. These genes are always contained within the organelle.

This principle would be violated, and the CORR hypothesis refuted, if any photosynthetic reaction centre core subunit or respiratory chain subunit currently organelle-encoded is found that functions in primary, vectorial electron transport and is nevertheless encoded in the nucleus and synthesized cytosolically as a precursor for import, all without any selective cost to the individual. From the chloroplast transcriptional experiments of Pfannschmidt *et al.* (1999*a*,*b*), the chloroplast genes whose transcription is regulated by the plastoquinone redox state include *psaAB* (for photosystem I reaction centre core polypeptides); *psbA* (for the photosystem II polypeptide, D1); and *rbcL* (for the large subunit of RUBISCO, interestingly following the same pattern of redox control—upregulated by oxidized plastoquinone—as *psbA*). Transcription of the chloroplast genes *petA* (cytochrome *f*); *trnG* (glycine tRNA) and *rpoB* (RNA polymerase) did not respond to plastoquinone redox state. *petA* for cytochrome *f* is predicted by this principle to be regulated at a different level of redox control, or, if at plastoquinone, at a different stage in gene expression. This prediction might also apply to *trnG* and *rpoB*, though these qualify as members of the secondary subset, that of genetic system genes, to which the principle in § 3i applies.

(i) *The principle of secondary involvement in energy transduction*

Genes whose products contribute to the organelle genetic system itself, or whose products are associated with secondary events in energy transduction, may be contained in the organelle in one group of organisms, but not in another, depending on the physiology and biochemistry of photosynthesis and respiration in the species concerned.

Even phylogenetically moderately related species (within kingdoms) show differences in the location of some genes, for example the ATP synthase subunits of *Neurospora* and *Saccharomyces* (Attardi & Schatz 1988). The CORR hypothesis also does not explicitly account for the wide disparities in the distribution of genetic system genes, for example the almost complete import of tRNAs in *Chlamydomonas* (GoldschmidtClermont 1998). Transport of nucleic acids across membranes is uncommon, and why some RNAs should be organelle-encoded in some species but nuclearly encoded in others is not currently understood. The only suggestion that can be made along the lines of CORR is that in some genetic systems genes may need to be under the control of a regulatory system that responds to a signal arising within the organelle itself. This signal could be the primary signal of altered redox state, or some secondary signal. Among the latter might be found signals that themselves report on the current state of gene expression, another piece of information that the nucleus might not have at its immediate disposal. Such secondary signals could result, at least in part, from CORR, though other developmental and environmental inputs may be imagined and would not detract from CORR unless it could be shown that they provided a sufficient reason for the retention of the genetic system as a whole.

(j) *The principle of the nuclear encoding of redox signalling components*

Components of the redox-signalling pathways upon which CORR *depends are themselves not involved in primary electron transfer, and so their genes have been relocated to the nucleus, in accordance with § 3f.*

This principle states that the redox signalling pathways that are required for CORR should not themselves be expected to utilize components whose biosynthesis must be under direct redox control. Such components should therefore fall into the major category of organellar proteins, and be imported as precursors from the cytosol, in line with the principle in § 3f. However, the principle above might be need to be modified in cases where amplification of the redox signal is useful to the cell. In the bacterial *ntr* system, which was a precedent for CORR (Harrison *et al.* 1990; Allen 1992), the histidine kinase *ntrB* and aspartate response regulator *ntrC* form part of the *gln* operon whose transcription they control. *Odontella*

and *Porphyra* have plastids that encode more proteins than those of most chlorophytes (Stoebe *et al.* 1998; Turmel *et al.* 1999), and these proteins include two-component signal transduction proteins prr and ycf27 (Stoebe *et al.* 1998). Extreme, or extremely rapid, environmental fluctuations might be catered for by amplification of signals that are transmitted by signalling pathways whose components are under their own control. It is not known whether ycf27 and prr are in operons whose promoters their gene products control. It is also uncertain whether ycf27 is a redox signalling component, and its distribution, at least in its plastidic location, seems to be correlated with the presence of a phycobilin-based light-harvesting system, as found in red algae and cyanobacteria (Ashby & Mullineaux 1999*a*,*b*; Ashby *et al.* 2002). The best guess of the original statement of the CORR hypothesis (Allen 1992, 1993*a*,*c*) was that bacterial two-component systems (Iuchi & Lin 1992; Allen 1993*c*; Grebe & Stock 1999) provide a plausible means of exerting redox regulatory control over mitochondrial and chloroplast gene expression, and subsequent research confirms the presence of two-component genes in genomes of *Arabidopsis* and yeast. Several *Arabidopsis* two-component genes are predicted to be targeted to mitochondria and chloroplasts (Forsberg *et al.* 2001). These must be considered as prime candidates for components involved in redox regulation of chloroplast transcription seen in the experiments of Pfannschmidt *et al.* (1999*a*, 2001*a*). The presumably homologous regulatory process serving to control photosystem stoichiometry in cyanobacteria, the prr system (Li & Sherman 2000), is now characterized as a two-component system involving a redox sensor and redox response regulator as originally defined (Allen 1992, 1993*a*,*c*).

4. CONTRASTING HYPOTHESES CONCERNING THE RETENTION OF CYTOPLASMIC GENOMES

CORR is a recent addition to explanations that have been put forward to account for the persistence of genomes in organelles. These models and their underlying mechanisms are now discussed and compared.

(a) *There is no reason: that is just how it is*

One reaction to the problem of cytoplasmic genetic systems (§ 2) is that there is no problem, and no particular reason for their retention. However, there is a very wide phylogenetic distribution of a common core, or primary subset, of organellar-encoded genes, and one which has withstood considerable genomic reorganization over evolutionary time. This constancy is seen, for example, in the same genes being present in the chloroplasts of land plants, with multiple copies of a single, circular chloroplast chromosome, and in dinoflagellate chloroplasts, with a collection of different DNA minicircles, each with essentially one gene (Zhang *et al.* 1999, 2001; Barbrook *et al.* 2001). Another suggested explanation of the retention of mitochondrial and chloroplast genomes is simply to restate the endosymbiont hypothesis, and to point out that chloroplast and mitochondrial genes 'came in' with the bacterial endosymbiont. This suggestion, however, does not address the question of why some genes were retained whereas others were not. It is also now clear that most genes that came in and still survive have nuclear and cytosolic gene products, whereas many chloroplast and mitochondria proteins had a eukaryotic origin (Salzberg *et al.* 2001; Martin & Russell 2003). The odds against a conserved pattern in gene location (e.g. table 1) arising by chance must be astronomical. Thus there is a reason for the retention of organellar genomes, and a reason for the retention of the same primary subset of genes. Because evolution proceeds by natural selection of chance variation, 'reason' in this context requires identification of the function of the location of a gene, that is, the selective value of the phenotype produced by that location. Evolution sometimes retains apparently functionless relics of previously useful features, but not for long, especially if there is a cost.

(b) *The lock-in hypothesis*

The 'lock-in' hypothesis of Bogorad (1975) was proposed independently of the endosymbiont hypothesis, and does not depend upon it. According to the lock-in hypothesis, core components of multi-subunit complexes must be synthesized, *de novo*, in the correct cellular compartment, otherwise photosystems or respiratory chain complexes, for example, might assemble in the plasma membrane, vacuolar membrane or endoplasmic reticulum. The lock-in hypothesis made no clear prediction about the continuity of cytoplasmic genetic systems, even though cytoplasmic, uniparental, non-Mendelian inheritance for both mitochondria (Roodyn & Wilkie 1968) and plastids (Kirk & Tilney-Bassett 1978) made it likely that organellar DNA, the chondriome and the plastome, replicated independently of nuclear DNA (Whitehouse 1969).

Mechanisms of protein import and targeting (Schatz 1998; Koehler *et al.* 1999; Leuenberger *et al.* 1999; Gabriel *et al.* 2001; Jarvis & Soll 2001; Pfanner & Geissler 2001; Duby & Boutry 2002; Pfanner & Wiedemann 2002) now effectively answer the problem of specifying the correct locations for membrane protein complexes for which the lock-in hypothesis was proposed as a solution. However, regulation of synthesis of an entire, functional, multisubunit complex by control of synthesis of individual subunits is still an important principle. This seems to be a feature of the 'CES' model (*c*ontrol by *e*pistasy of *s*ynthesis) of Wollman (Wollman *et al.* 1999; Eberhard *et al.* 2002), which envisages nuclear control as all-embracing, with organellar genetic systems doing, as it were, what they are told. In the hypothesis of CORR propounded here, CES is equally important, but the essence of CORR is that the organellar genetic system alone knows what are the immediate and most important signals of environmental change. The organelle may keep the nucleus informed, but this is a lower priority than initiating an appropriate response. CES and CORR agree that assembly of multisubunit complexes is a complex business, and requires coordinated control, in which no step is rate-limiting but many may be obligatory.

(c) *Going, going, gone: the evolutionary process of transfer of genes from organelles to the nucleus is still underway, but incomplete*

This hypothesis envisages that, for most eukaryotes, we are now witnessing the last stages in a slow evolutionary transfer of genes to the nucleus, and this transfer may be expected to proceed to completion in what is, in evol-

utionary terms, the near future (Palmer 1997; Gray 1999, 2000; Gray *et al.* 1999; Adams *et al.* 2000; Daley *et al.* 2002). A clear and compelling statement of this widely held position is that of Herrmann & Westhoff (2001, p. 10):

> …the transfer of genes…tells that the partite plant genome is not in a phylogenetic equilibrium. All available data suggest that the ultimate aim of genome restructuring in the plant cell, as in the eukaryotic cell in general, is the elimination of genome compartmentation while retaining physiological compartmentation.

One serious problem with this hypothesis is the distinctly non-random sample of genes that remain to be transferred. A relatively constant subset of genes is found in bioenergetically functional mitochondria from fungi to mammals and angiosperms, as in photosynthetic plastids (chloroplasts) from unicellular algae to angiosperms. This hypothesis therefore requires an additional, *ad hoc* assumption that some gene transfers are inherently more probable than others (this too is widely supposed, see later), and thus that gene transfer events proceed in the same sequence in quite independent lineages. There seems to be no evidence for this, and the order of gene loss from mitochondria and chloroplasts has been used to infer phylogeny (Martin *et al.* 1998; Adams *et al.* 2000; Millen *et al.* 2001). 'Going, going, gone' (Palmer 1997) is difficult to refute categorically, but it does not seem to be consistent with recent evidence. For example, chromosome two of *Arabidopsis thaliana* (Lin *et al.* 1999) has a single insertion encompassing, with about 99% sequence similarity, the complete *Arabidopsis* mitochondrial genome (Unseld *et al.* 1997; Marienfeld *et al.* 1999; Stupar *et al.* 2001). Unless this single transfer is a freak occurrence, this finding strongly indicates that gene transfer occurs with relative ease (Adams *et al.* 2000; Henze & Martin 2001; Rujan & Martin 2001; Salzberg *et al.* 2001). In addition, there is experimental evidence that randomly generated polypeptides are mitochondrially imported at high frequency (Lemire *et al.* 1989). The hypothesis that we now see the same few stragglers in diverse organellar genomes is not impossible, but should be regarded as a last resort, perhaps applicable only in extreme cases such as the greatly reduced apicomplexan plastomes and chondriomes. Even there, the hypothesis carries with it an obligation to explain why organisms with few organellar genes have 'evolved' faster than those with many.

(d) *The frozen accident*

What may be described as 'the frozen accident hypothesis' states that the evolutionary process of gene transfer from organelle to nucleus was underway when something happened that stopped it. One example of the frozen accident is a scenario (Von Heijne 1986) in which import into mitochondria of the precursor proteins arising from newly translocated genes proceeded for long enough, after the primordial endosymbiosis, for most of the symbiont-derived genes to be moved to the cell nucleus. The 'accident' that halted the process was the evolutionary origin of exocytosis and protein secretion, after which it became impossible for precursors properly destined for mitochondrial import to be distinguished from those destined for export from the cell. As with the 'lock-in' hypothesis, the 'frozen accident' seems incompatible with the remarkable precision and specificity of protein targeting (Schatz 1998; Koehler *et al.* 1999; Leuenberger *et al.* 1999; Gabriel *et al.* 2001; Jarvis & Soll 2001; Pfanner & Geissler 2001; Duby & Boutry 2002; Pfanner & Wiedemann 2002). Protein sorting mechanisms do not seem to forbid controlled import into any cellular compartment. Another 'accident' that may have impeded subsequent gene transfer in the case of mitochondria is their departure from what is otherwise a universal genetic code. Genetic code differences are discussed later in the context of whether some genes cannot be moved.

(e) *Some proteins (with cofactors) cannot be imported*

We may describe as the 'unimportability hypothesis' the proposal that some proteins are inherently unimportable and must therefore be synthesized *de novo* in the correct compartment within the cell. The unimportability hypothesis is most often proposed as its special case, the hydrophobicity hypothesis, with hydrophobicity being the reason for unimportability.

(i) *Hydrophobicity*

A commonly entertained notion is that proteins that are encoded and synthesized within organelles comprise a primary subset characterized by shared hydrophobicity: all are intrinsic membrane proteins (Von Heijne 1986; Popot & de Vitry 1990; Claros *et al.* 1995). As with the CORR hypothesis favoured here, the primary subset then requires a secondary subset of genetic system components, the latter not themselves being hydrophobic, but being required for synthesis of the others. Certainly the mitochondrial inner membrane contains some intrinsic protein complexes whose subunits broadly satisfy the criterion of hydrophobicity. However, the core subunits that are mitochondrially encoded are also explained by the CORR hypothesis. This would make hydrophobicity a requirement for their function in vectorial electron and proton transfer, and not a reason, in itself, for retention of their genes in the cytoplasm.

The hydrophobicity hypothesis has some very clear counter-examples in chloroplasts, and cannot therefore stand if one wishes for a single principle that embraces both types of bioenergetic organelle. One clear counter-example is the large (L) subunit of the enzyme RUBISCO, an abundant but entirely membrane-extrinsic protein of *ca*. 55 kDa molecular mass. The RUBISCO large subunit is generally chloroplast-encoded, with the exception of the nuclearly encoded, bacterial-type L subunit of L_2 RUBISCO in peridinin-containing dinoflagellates (Morse *et al.* 1995). In most photosynthetic eukaryotes RUBISCO has a chloroplast encoded large subunit and a nuclearly encoded small (S) subunit (Ellis 1984; Gatenby & Ellis 1990), giving an L_8S_8 holoenzyme, and this RUBISCO is widely taken as a paradigm for the need for coordination of nuclear and chloroplast gene expression (Rodermel 2001).

By complete contrast to the RUBISCO large subunit, which is a chloroplast-encoded and water-soluble protein, a second major violation of the hydrophobicity hypothesis is the existence of the universally nuclearly encoded but always hydrophobic subunits of the chloroplast LHC II and LHC I (Chitnis & Thornber 1988). These apopro-

teins of *ca.* 27–30 kDa are quite water-insoluble, and are usually fully extracted only in chloroform–methanol or in the presence of detergents. The Lhc gene family is diverse (Green & Kuhlbrandt 1995), but most gene products or Lhcp (*l*ight-*h*arvesting *c*hlorophyll *p*rotein) polypeptides have three membrane-spanning, hydrophobic α-helices (Kühlbrandt 1994). It is interesting that three-helix Lhc proteins seem to be absent from extant cyanobacteria and chloroxybacteria (Green & Kuhlbrandt 1995).

(ii) *The five-helix rule*

Perhaps with the three-helix model of LHC II in mind, a variation of the hydrophobicity hypothesis is the empirical 'five-helix rule' (Popot & de Vitry 1990; Claros *et al.* 1995; Wollman *et al.* 1999), which states that no intrinsic membrane protein with five or more membrane-spanning helices can be encoded in the nucleus. Five is a good guess for a threshold number of helices demanding organellar-encoding in the context of photosynthetic reaction centres (Heathcote *et al.* 2002), where the earliest known structure, that of type II centres (Deisenhofer *et al.* 1985), as in photosystem II (Zouni *et al.* 2001), shows exactly five helices, with type I centres containing 11 (Jordan *et al.* 2001). There are, however, many nuclearly encoded subunits in both mitochondria and chloroplasts that have fewer than five membrane helices, so the five-helix rule can only describe a sufficient condition, and not a necessary condition, for nuclear encoding.

It should also be noted that ADP–ATP carriers (AACs) are encoded in the nucleus. AACs of the mitochondrial inner membrane have six transmembrane helices (van der Giezen *et al.* 2002) while the those of the chloroplast inner envelope (Winkler & Neuhaus 1999) have 12 (Tjaden *et al.* 1998).

With the counter-examples of nuclearly encoded AACs with six or 12 membrane-spanning helices, many intrinsic membrane proteins with fewer than five helices, some nuclearly encoded and some not, and with the RUBISCO large subunit as a major product of chloroplast gene expression with no membrane domain at all, the five-helix rule begins to seem rather like a case of special pleading. Even if there are indeed special mitigating circumstances for AACs (this argument is outlined later), the five-helix rule, as with the hydrophobicity hypothesis of which it is a special case, will disappear if it can be shown that possession of five or more membrane helices is required for some other function, such as that of CORR, that determines the organellar location of the gene concerned. Hydrophobicity is, in some cases, a feature of proteins that are encoded in organelles, but it is not, in itself, the issue.

For chloroplasts, the hydrophobicity hypothesis predicts retention of cyanobacterial genes in functional categories associated with the chloroplast envelope and with transport, just as much as with photosynthesis. Table 1 shows that this is not the case.

A mechanism for the unimportability of hydrophobic proteins is rarely stated explicitly, and it is rather left to the imagination whether hydrophobic proteins might, as it were, get stuck halfway through the chloroplast envelope or mitochondrial outer membrane; whether transport cannot get started on a precursor protein that has already spontaneously inserted into the membrane; or whether hydrophobic proteins will be unable to make the journey through the cytosol without precipitating there. In the cytosol, co-translational transport or membrane insertion would, in principle, address the latter point, but it seems to be generally true that chloroplast and mitochondrial proteins are post-translationally inserted into the organellar outer membrane. So the obvious ancillary question about unimportability is 'why then cannot hydrophobic proteins be co-translationally imported into organelles?' There is evidence in yeast for co-translational transport into mitochondria of proteins thought to be derived from the original endosymbiont (Marc *et al.* 2002). If there is any truth in hydrophobicity as a barrier to protein import, the CORR principle of unselective import (§ 3c) should be easily refuted, for example by a demonstration that no known targeting presequence is capable of importing polyphenylalanine encoded by polyU for *in vitro* import studies.

An additional proposal intended to account for the unimportability of some hydrophobic proteins, and for importability of others, is suggested by A. de Grey (personal communication), who argues that bacterially derived protein precursors imported into mitochondria cross the mitochondrial inner membrane completely before insertion into that membrane from the mitochondrial matrix (the N-phase). This is held to be an obstacle to successful import of membrane-intrinsic proteins that might, *en route*, discharge the proton motive force. By contrast, eukaryotically derived proteins can be inserted into the mitochondrial inner membrane from the inter-membrane space (the P-phase), and then there is, for some reason, no similar uncoupling effect. Thus adenylate transporters or AACs can be nuclearly encoded whereas cytochrome oxidase subunits, for example, cannot. It is not obvious why bacterially derived subunits, such as those of cytochrome oxidase, did not acquire eukaryotic insertion mechanisms. Perhaps this acquisition is forbidden by the need to conserve pre-existing routes of assembly into multi-subunit complexes, or perhaps the need for insertion from the inside reflects a need to incorporate a cofactor, for example haem or iron–sulphur clusters, along the way.

(iii) *Unimportability of three dimensional structures and cofactors*

Accumulating evidence on protein transport mechanisms seems to make a specific obstacle to successful transport of certain sorts of protein less and less plausible, especially given the observation that precursors with N-terminal targeting sequences, not mature, functional polypeptides themselves, traverse membranes during biosynthesis.

Protein folding inside the living cell is invariably assisted by molecular chaperones (Gatenby & Ellis 1990; Azem *et al.* 1997; Komiya *et al.* 1997, 1990; Jackson-Constan *et al.* 2001), proteins that impede the formation of 'improper' structural interactions that will otherwise produce aggregation or other non-functional interactions. Protein import mechanisms appear mostly to rely on specific molecular chaperones for guided unfolding, before membrane insertion, and refolding after translocation of the polypeptide into its destined compartment (Gatenby & Ellis 1990; Azem *et al.* 1997; Komiya *et al.* 1997; Jackson-Constan *et al.* 2001). Because polypeptides are mostly 'threaded'

through membranes, no special constraint on transport should be expected to arise from 3D structure or surface properties of a protein.

An exception to this unfolding rule may be provided by holoproteins containing cofactors that are transported along with the apoprotein, presumably with a largely intact 3D structure. The 'TAT' (twin-arginine translocase) system (Bogsch *et al.* 1998; Dalbey & Robinson 1999; Berks *et al.* 2000*a*,*b*; Robinson 2000; Robinson & Bolhuis 2001; Sargent *et al.* 2002) works broadly along these lines, though its mechanism is unresolved. TAT is relevant here because it adds transport of essentially complete holoproteins to the cell's repertoire of transport mechanisms, and even proteins that may not be allowed to unfold, perhaps because of the toxicity of the free cofactor, are not thereby made candidates for an unimportability rule. One variant of the unimportability hypothesis is that transport can work only by unfolding, and dissociated cofactors may be toxic to the cell. The bacterially derived TAT system (Bogsch *et al.* 1998; Dalbey & Robinson 1999; Berks *et al.* 2000*a*,*b*; Sargent *et al.* 2002) overcomes this argument, as would examples of cofactor-free polypeptides, such as (C)F_o subunit *c*, which are nevertheless generally encoded in organelles.

The presence of a vesicular transport system in chloroplasts (Westphal *et al.* 2001) counts against the unimportability hypothesis. Vesicular transport between the chloroplast envelope and the thylakoid membrane should, in principle, allow any protein to be ferried across the chloroplast stroma. Vesicles might allow transport, in particular, of membrane proteins that would discharge the proton motive force if allowed to connect, transiently, the N and P aqueous phases on each side of a coupling membrane.

Another observation that seems to count against unimportability is the increased number of membrane barriers in plastids derived from secondary, tertiary or higher-order endosymbioses. These plastids seem to have retained, if anything, fewer genes rather than more, as unimportability predicts. The nucleomorph, the residual nucleus of the primary host, retains many genes for molecular chaperones (Archibald *et al.* 2001; Cavalier-Smith 2002), in agreement with the conclusion that protein import proceeds by conserved and specialized mechanisms. If each new import added to gain a higher-order endosymbiosis requires additional, targeting presequences, then perhaps there is an advantage in components of an import machinery not being imported themselves.

It seems that unimportability is less easy an assumption to make today than in previous years, and the onus may now be on its proponents to cite specific examples of proteins that can never be imported into chloroplasts or mitochondria.

(f) *Some genes cannot be moved*

(i) *Differences in the mitochondrial genetic code*

The genetic code is the specific relation between ribonucleotide triplets of codons and tRNAs that recognize them, and which thus determine which individual amino acids, if any, will be incorporated into the growing polypeptide chain. Mitochondria show small departures from the otherwise universal genetic code (Attardi & Schatz 1988; Alberts *et al.* 1994). UGA is a 'stop' codon in the universal code and in plant mitochondria, but is the codon for tryptophan in mitochondria of animals and fungi; CUA means leucine everywhere except fungal mitochondria, where it means threonine; and AGA and AGU code for arginine in the 'universal' code, but serine in insect mitochondria and 'stop' in mammalian ones. The significance of this heterogeneity in coding is most unlikely to rest in the early origin of mitochondria, as the 'universal' code is employed by bacteria and protists. Furthermore, the diversity of departures from the 'universal' code among mitochondria of different eukaryotic lineages indicates they had independent origins. A general rule for mitochondrial genetic codes is that reliance on the third, 'wobble' position is decreased, which allows all 20 protein amino acids to be specified by fewer tRNAs (22 in animal mitochondria instead of 31 in the cytosol). This in turn means that mtDNA can specify more proteins for a given length of DNA, or, conversely, that the mitochondrial genome can be condensed or 'streamlined'. Although this arrangement does not explain the retention of mitochondrial genomes, and may be a secondary event in mitochondrial evolution, it almost certainly places an extra barrier in place of gene relocation. As discussed previously, however, there seems to be no single barrier that can explain the retention of a relatively constant subset of mitochondrial genes.

There has clearly been a contraction of mitochondrial DNA, especially in animals. Animal mitochondrial DNA has few, if any, non-coding sequences, and animal mitochondrial genetic codes require fewer tRNAs than the single code operating in the cytosol (Gray 1999; Gray *et al.* 1999; Selosse *et al.* 2001). These may be devices to decrease the total frequency of mutations per mitochondrial genome, assuming a constant frequency per unit length of mitochondrial DNA. It has been proposed that a gene's removal to the nucleus decreases its susceptibility to free radical mutagenesis (Raven *et al.* 1994*a*,*b*; Allen & Raven 1996). If this is the case, then the non-standard code may be a further indication of an overriding need to maintain certain genes within the mitochondrion. Whatever the reason for the retention of genes in organelles it had better be a good one: bioenergetic organelles are clearly the wrong place to keep genetic information, and go to extraordinary lengths to offset the consequences of it being there.

(ii) *Molecular refugees*

A phenomenon related to differences in the genetic code is seen in chloroplasts. This is a bias in codon usage. The 'molecular refugees' hypothesis (Howe *et al.* 2000) suggests that nuclear copies of what were once chloroplast genes escape constraints on codon usage, and thus that what distinguishes chloroplast-encoded proteins is that their amino-acid composition is unusual enough for the bias in codon usage not to apply. It is difficult to see what physico-chemical properties might be uniquely shared by the amino acids of chloroplast-encoded proteins, and, in addition, there seem to be no 'forbidden' amino-acid residue in chloroplasts. As with the departure from the universal genetic code in mitochondria, is seems likely that a bias in codon usage reflects some secondary property of chloroplast genomes and is not the primary reason for their retention. In any case, any restriction on gene move-

ment will act in the wrong direction to explain why chloroplasts have retained genomes at all, provided there is no resulting obstacle to a chloroplast protein being successfully encoded, as a precursor, in the nucleus.

(iii) *Genes are copied and replicated, not relocated in one step*

If complete relocation is envisaged as proceeding by the simultaneous appearance of a gene in one compartment and its deletion from another, then it is possible to imagine that, whatever the vector is, there may be selectivity in gene transfer (Henze & Martin 2001). For example, if viral incorporation, replication and recombination in both organelle and nucleus were the mechanism of relocation, then successful gene transfer might depend on the ability of the virus to complete the process in one step. This in turn might depend on some property of the gene, or of the protein it encodes if expression occurs at any stage in the process of relocation. An assumption of the CORR hypothesis favoured here is that genes are not moved physically, but merely copied, after which either the copy in the original location or the copy in the new location may be lost as a result of selection. A useful analogy may be the action of copying files between computer storage volumes. In this analogy the copy is identical to the original, and either may subsequently be deleted without loss of information.

Because any gene that is not usefully expressed will undergo no selection to maintain a functional allele, the prevailing location will be the one in which the functional copy of the gene resides. The problem of which gene survives is then subsumed into the problem of which is the selectively advantageous location for it. It seems reasonable to assume that there was—and is—an interim period in which homologous genes in different compartments are essentially in competition to see which functions best. As outlined above, there is no obvious obstacle for copying DNA sequences within the cell. The physico-chemical properties of DNA are fairly uniform and not directly dependent on nucleotide sequence, though organelles, like intra-cellular parasites, seem to prefer A and T to G and C. The answer to the question 'which gene location functions best?' may have nothing to do with the gene itself, nor, directly, with the structure and function of its gene product. Instead, the answer may have everything to do with the required regulatory interaction between the gene product and expression of the gene that produces it.

(g) *CORR*

This hypothesis has been propounded earlier (§ 3) and will be summarized as follows. Vectorial electron and proton transfer exerts regulatory control over expression of genes encoding proteins directly involved in, or affecting, redox poise. This regulatory coupling is indispensable for adaptation to environmental changes in chemical activities of oxygen, light and CO_2, and requires co-location of these genes with their gene products. Organelles 'make their own decisions' on the basis of environmental changes affecting redox state. The hypothesis is abbreviated CORR for *co*-location for *r*edox *r*egulation or, more completely, '*co*-location (of gene and gene product) for (evolutionary) continuity of *r*edox *r*egulation of gene expression'.

5. SUMMARY OF THE CURRENT STATUS OF THE HYPOTHESIS OF CORR

Q. Why do mitochondria and chloroplasts have their own genetic systems?

A. CORR: co-location of genes and gene products permits direct and autonomous redox regulation of gene expression.

(a) *Objections to CORR*

(i) *There is no regulation of gene expression in chloroplasts and mitochondria*

This is still commonly stated but incorrect. See the principle in § 3d.

(ii) *There is no regulation of gene expression in chloroplasts and mitochondria at the level of transcription*

For chloroplasts there is now clear evidence for a functionally intelligible redox control of transcription (Pfannschmidt *et al.* 1999*a*). This invalidates the previous, widely held assumption that all control of chloroplast gene expression is post-transcriptional (Gruissem & Tonkyn 1993; GoldschmidtClermont 1998). The evidence for mitochondria is less clear, but redox control of transcription certainly cannot be ruled out, and there are indications that it may occur (Pearson *et al.* 1993; Konstantinov *et al.* 1995; Wilson *et al.* 1996*b*). It must also be restated that transcriptional control is not necessary for CORR, but serves as a compelling example of a conserved, typically bacterial, mode of response that operates still in eukaryotic, cytoplasmic organelles, thus illustrating the principle of continuity of redox control (§ 3d). Post-transcriptional redox regulation (Danon & Mayfield 1994; Kim & Mayfield 1997; Brown *et al.* 2001; Trebitsh & Danon 2001) is consistent with CORR and there is nothing to forbid it being a primary mechanism that satisfies most, or indeed all, of the formal requirements of the hypothesis in particular cases.

(iii) *There are no redox signalling components encoded in chloroplast or mitochondrial DNA*

This is what CORR predicts. See the principle in § 3j.

(iv) *There is redox control of nuclear gene expression*

This is correct and many examples are now known, especially for chloroplasts (Escoubas *et al.* 1995; Kovacs *et al.* 2000; Pfannschmidt *et al.* 2001*a*,*b*; Surpin *et al.* 2002). Redox signals from photosynthetic or respiratory electron transport chains are not at all prohibited by CORR. The essence of CORR is that there is a subset of organellar proteins for which there is an overriding requirement for a redox regulatory control that is initiated within the organelle and whose effects are exerted by the simplest and most direct route. This redox signalling is envisaged to take place without a requirement for prior consultation and coordination with what may be regarded as a nucleocytoplasmic bureaucracy.

(v) *Nuclear genes can be expressed as rapidly as chloroplast and mitochondrial genes*

Zerges (2002), for example, cites the equal times taken for complete expression of nuclear and chloroplast genes in *Chlamydomonas* as evidence against the CORR hypoth-

esis. However, the question is not one of the time taken to express a gene, but the time taken to alter the rate of its expression in response to a (redox) signal originating within the organelle. An equally important, additional factor may be the autonomy of the signalling pathway, that is, its independence from other inputs as described in the previous paragraph. One might say that mitochondria and chloroplasts require a free hand.

(vi) *There is no evidence to support this hypothesis*

The hypothesis is rich in predictions (§§ 3 and 6), and those that have been tested stringently are in agreement with experimental results. None of the predictions of the hypothesis has been found to be incorrect, and none of the underlying principles (§ 3) has been disproved. Further predictions will be the subject of future experiments.

(b) *Evidence for CORR*

(i) *Chloroplasts have selectively retained cyanobacterial genes for primary events in photosynthesis*

See table 1.

(ii) *There is regulation of gene expression in chloroplasts*

There is transcriptional regulation of gene expression in chloroplasts (§ 3d). The case for mitochondria is not yet proven.

(iii) *There is* ***redox*** *regulation of gene expression in chloroplasts*

There is transcriptional redox regulation of gene expression in chloroplasts (§ 3d). The case for mitochondria is not yet proven, though see (Allen *et al.* 1995; Konstantinov *et al.* 1995; Galvis *et al.* 1998).

(iv) *There are no redox signalling components encoded in chloroplast or mitochondrial DNA*

Correct, but with possible exceptions discussed in § 3j.

(v) *There are 'bacterial' redox signalling components in eukaryotes*

See § 3j.

(vi) *Further predictions: DNA-array results and phenotypes of allotopic mutants*

A systematic test of CORR might be to make an oxidized-minus-reduced gene expression profile of the 'transcriptome' of a cyanobacterium to see if the set of genes whose transcription is under most rapid and direct redox control corresponds to the genes of the chloroplast genome. Laboratories with this experimental capability do not seem to have addressed this possibility (Hihara *et al.* 2002; Ikeuchi 2002). Likewise, if a putative modern representative of the proto-mitochondrion can be found and agreed, e.g. *Rhodospirillum rubrum* or *Paracoccus denitrificans*, then a similar experiment should reveal a subset of genes corresponding to mitochondrial genomes. DNA-microarrays also present the possibility of testing the relative rate of redox response, and dependency on other signalling inputs, of transcription of eukaryotic genes. To date organellar genes have not been included in such investigations (Ferea *et al.* 1999). A further experimental possibility is to examine the redox regulatory phenotypes of existing allotopic mutants, as outlined in § 3e.

6. CONCLUSION AND FUTURE PROSPECTS

The CORR hypothesis states that the function of mitochondrial and chloroplast genomes is to provide co-location (of gene and gene product) for (evolutionary) continuity of redox regulation of gene expression. This hypothesis is neither self-contradictory nor devoid of testable predictions. It is also consistent with available evidence, and, in principle, makes sense of the otherwise puzzling distribution of genes between compartments in eukaryotic cells. The explanation that CORR offers applies equally to mitochondria and chloroplasts, and identifies selective forces that still operate to determine the location of genes. No selective barriers need be assumed to gene transfer, protein import, nor membrane insertion. Nor do we need to invoke a 'frozen accident': that something happened to stop gene transfer, for some reason, at some time in the remote and inaccessible past. CORR also explains the absence of genomes from organelles that have lost their primary bioenergetic functions. CORR may act as a stimulus to a search for residual bioenergetic function in those organelles that have lost aerobic respiration and photosynthesis but where residual genomes remain.

CORR also predicts regulatory properties of known components of chloroplasts and mitochondria that in turn indicate important but as yet uncharacterized flexibility in energy transduction. These components include the RUBISCO large subunit, and the CF_o and F_o subunits of coupling ATPase in both chloroplasts and mitochondria. Redox regulation of the relative stoichiometry of components, as seen in regulation of chloroplast photosystem stoichiometry (Pfannschmidt *et al.* 1999*a*), may extend to $(C)F_o$-$(C)F_1$-ATPase, indicating that the function of retention of genes for one or more of these components is flexibility in the stoichiometries of H^+ to e^- and H^+ to ATP in chemiosmotic coupling (Schemidt *et al.* 1998; Stock *et al.* 2000; Allen 2002).

So far, explicit tests of CORR have largely been confined to chloroplasts, and to some extent mitochondria of higher plants, but this merely reflects the expertise and interests of the investigators. Further attempted disproof of CORR is welcomed and encouraged. Novel experimental tests can be devised, as well as attempted refutation through analysis of existing data. It is hoped that explicit attention will be given to the CORR hypothesis in laboratories already suitably equipped, and that resources can be allocated for this purpose in laboratories where there is the interest but not, currently, the technology or expertise.

The heuristic value of the hypothesis, already seen in the role of redox-regulated transcription of photosynthetic reaction centre genes in chloroplasts, may be its most useful property. The central proposal is that mitochondria and chloroplasts contain genes for proteins whose function in electron transfer demands rapid, direct, and unconditional redox regulatory control of their biosynthesis. The full implications of this hypothesis are very broad and only just beginning to be explored.

I thank Carol Allen, Aubrey de Grey, William Martin, John A. Raven and Francis-André Wollman for discussion and correspondence during the preparation of this paper, and William Martin for suggestions, including table 1. Research grants from

The Swedish Natural Sciences Research Council (NFR) and the Crafoord Foundation are gratefully acknowledged.

REFERENCES

Adams, K. L., Daley, D. O., Qiu, Y. L., Whelan, J. & Palmer, J. D. 2000 Repeated, recent and diverse transfers of a mitochondrial gene to the nucleus in flowering plants. *Nature* **408**, 354–357.

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. & Watson, J. D. 1994 *The molecular biology of the cell*. New York and London: Garland Publishing.

Allen, C. A., Hakansson, G. & Allen, J. F. 1995 Redox conditions specify the proteins synthesized by isolated chloroplasts and mitochondria. *Redox Rep.* **1**, 119–123.

Allen, J. F. 1992 Protein phosphorylation in regulation of photosynthesis. *Biochim. Biophys. Acta* **1098**, 275–335.

Allen, J. F. 1993*a* Control of gene expression by redox potential and the requirement for chloroplast and mitochondrial genomes. *J. Theor. Biol.* **165**, 609–631.

Allen, J. F. 1993*b* Redox control of gene expression and the function of chloroplast genomes: a hypothesis. *Photosynth. Res.* **36**, 95–102.

Allen, J. F. 1993*c* Redox control of transcription: sensors, response regulators, activators and repressors. *FEBS Lett.* **332**, 203–207.

Allen, J. F. 1995 Thylakoid protein–phosphorylation, state-1–state-2 transitions, and photosystem stoichiometry adjustment: redox control at multiple levels of gene expression. *Physiol. Plant* **93**, 196–205.

Allen, J. F. 1996 Separate sexes and the mitochondrial theory of ageing. *J. Theor. Biol.* **180**, 135–140.

Allen, J. F. 2002 Photosynthesis of ATP-electrons, proton pumps, rotors, and poise. *Cell* **110**, 273–276.

Allen, J. F. & Allen, C. A. 1999 A mitochondrial model for premature ageing of somatically cloned mammals. *IUBMB Life* **48**, 369–372.

Allen, J. F. & Nilsson, A. 1997 Redox signalling and the structural basis of regulation of photosynthesis by protein phosphorylation. *Physiol. Plant* **100**, 863–868.

Allen, J. F. & Pfannschmidt, T. 2000 Balancing the two photosystems: photosynthetic electron transfer governs transcription of reaction centre genes in chloroplasts. *Phil. Trans. R. Soc. Lond.* B **355**, 1351–1357. (DOI 10.1098/rstb.2000.0697.)

Allen, J. F. & Race, H. L. 2002 Will the real LHC II kinase please step forward? *Science's STKE* http://www.stke.org/cgi/content/full/sigtrans;2002/155/pe43

Allen, J. F. & Raven, J. A. 1996 Free-radical-induced mutation vs redox regulation: costs and benefits of genes in organelles. *J. Mol. Evol.* **42**, 482–492.

Archibald, J. M., Cavalier-Smith, T., Maier, U. & Douglas, S. 2001 Molecular chaperones encoded by a reduced nucleus: the cryptomonad nucleomorph. *J. Mol. Evol.* **52**, 490–501.

Ashby, M. K. & Mullineaux, C. W. 1999*a* Cyanobacterial ycf27 gene products regulate energy transfer from phycobilisomes to photosystems I and II. *FEMS Microbiol. Lett.* **181**, 253–260.

Ashby, M. K. & Mullineaux, C. W. 1999*b* The role of ApcD and ApcF in energy transfer from phycobilisomes to PSI and PSII in a cyanobacterium. *Photosynth. Res.* **61**, 169–179.

Ashby, M. K., Houmard, J. & Mullineaux, C. W. 2002 The ycf27 genes from cyanobacteria and eukaryotic algae: distribution and implications for chloroplast evolution. *FEMS Microbiol. Lett.* **214**, 25–30.

Attardi, G. & Schatz, G. 1988 Biogenesis of mitochondria. *A. Rev. Cell Biol.* **4**, 289–333.

Azem, A., Oppliger, W., Lustig, A., Jeno, P., Feifel, B., Schatz, G. & Horst, M. 1997 The mitochondrial hsp70 chaperone system: effect of adenine nucleotides, peptide substrate, and mGrpE on the oligomeric state of mhsp70. *J. Biol. Chem.* **272**, 20 901–20 906.

Barbrook, A. C., Symington, H., Nisbet, R. E. R., Larkum, A. & Howe, C. J. 2001 Organisation and expression of the plastid genome of the dinoflagellate *Amphidinium operculatum*. *Mol. Genet. Genomics* **266**, 632–638.

Bauer, C. E., Elsen, S. & Bird, T. H. 1999 Mechanisms for redox control of gene expression. *A. Rev. Microbiol.* **53**, 495–523.

Bauer, C. E., Elsen, S., Swem, L. R., Swem, D. L. & Masuda, S. 2003 Redox and light regulation of gene expression in photosynthetic prokaryotes. *Phil. Trans. R. Soc. Lond.* B **358**, 147–154. (DOI 10.1098/rstb.2002.1189.)

Berks, B. C., Sargent, F., De Leeuw, E., Hinsley, A. P., Stanley, N. R., Jack, R. L., Buchanan, G. & Palmer, T. 2000*a* A novel protein transport system involved in the biogenesis of bacterial electron transfer chains. *Biochim. Biophys. Acta Bioenerg.* **1459**, 325–330.

Berks, B. C., Sargent, F. & Palmer, T. 2000*b* The tat protein export pathway. *Mol. Microbiol.* **35**, 260–274.

Blackstone, N. W. 1999 Redox control in development and evolution: evidence from colonial hydroids. *J. Exp. Biol.* **202**, 3541–3553.

Blackstone, N. W. 2000 Redox control and the evolution of multicellularity. *Bioessays* **22**, 947–953.

Blackstone, N. W. 2001 Redox state, reactive oxygen species and adaptive growth in colonial hydroids. *J. Exp. Biol.* **204**, 1845–1853.

Bogorad, L. 1975 Evolution of organelles and eukaryotic genomes. *Science* **188**, 891–898.

Bogsch, E. G., Sargent, F., Stanley, N. R., Berks, B. C., Robinson, C. & Palmer, T. 1998 An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria. *J. Biol. Chem.* **273**, 18 003–18 006.

Brown, E. C., Somanchi, A. & Mayfield, S. P. 2001 Interorganellar crosstalk: new perspectives on signaling from the chloroplast to the nucleus. *Genome Biol.* **2**, 1021.1–1021.4.

Cavalier-Smith, T. 2002 Chloroplast evolution: secondary symbiogenesis and multiple losses. *Curr. Biol.* **12**, R62–R64.

Chitnis, P. R. & Thornber, J. P. 1988 The major light-harvesting complex of photosystem 2. Aspects of its molecular and cell biology. *Photosynth. Res.* **16**, 41–63.

Chow, W. S., Melis, A. & Anderson, J. M. 1990 Adjustments of photosystem stoichiometry in chloroplasts improve the quantum efficiency of photosynthesis. *Proc. Natl Acad. Sci. USA* **87**, 7502–7506.

Claros, M. G., Perea, J., Shu, Y., Samatey, F. A., Popot, J. L. & Jacq, C. 1995 Limitations to *in vivo* import of hydrophobic proteins into yeast mitochondria. The case of a cytoplasmically synthesized apocytochrome *b*. *Eur. J. Biochem.* **228**, 762–771.

Dalbey, R. E. & Robinson, C. 1999 Protein translocation into and across the bacterial plasma membrane and the plant thylakoid membrane. *Trends Biochem. Sci.* **24**, 17–22.

Daley, D. O., Adams, K. L., Clifton, R., Qualmann, S., Millar, A. H., Palmer, J. D., Pratje, E. & Whelan, J. 2002 Gene transfer from mitochondrion to nucleus: novel mechanisms for gene activation from Cox2. *Plant J.* **30**, 11–21.

Danon, A. & Mayfield, S. P. 1994 Light-regulated translation of chloroplast messenger RNAs through redox potential. *Science* **266**, 1717–1719.

de Grey, A. 2000 Mitochondrial gene therapy: an arena for the biomedical use of inteins. *Trends Biotechnol.* **18**, 394–399.

Deisenhofer, J., Epp, O., Miki, K., Huber, R. & Michel, H. 1985 Structure of the protein subunits in the photosynthetic reaction center of *Rhodopseudomonas viridis* at 3Å resolution. *Nature* **318**, 618–624.

Delwiche, C. F. 1999 Tracing the thread of plastid diversity through the tapestry of life. *Am. Nat.* **154**, S164–S177.

Doolittle, W. F. 1999 Phylogenetic classification and the universal tree. *Science* **284**, 2124–2129.

Douglas, S. E. 1998 Plastid evolution: origins, diversity, trends. *Curr. Opin. Genet. Dev.* **8**, 655–661.

Duby, G. & Boutry, M. 2002 Mitochondrial protein import machinery and targeting information. *Plant Sci.* **162**, 477–490.

Eberhard, S., Drapier, D. & Wollman, F. A. 2002 Searching limiting steps in the expression of chloroplast encoded proteins: relations between gene copy number, transcription, transcript abundance and translation rate in the chloroplast of *Chlamydomonas reinhardtii*. *Plant J.* **31**, 149–160.

El Bissati, K. & Kirilovsky, D. 2001 Regulation of psbA and psaE expression by light quality in *Synechocystis* species PCC 6803. A redox control mechanism. *Plant Physiol.* **125**, 1988–2000.

Ellis, R. J. 1984 The nuclear domination of chloroplast development. *Science Progress* **69**, 129–142.

Embley, T. M. & Martin, W. 1998 Molecular evolution: a hydrogen-producing mitochondrion. *Nature* **396**, 517–519.

Ems, S. C., Morden, C. W., Dixon, C. K., Wolfe, K. H., dePamphilis, C. W. & Palmer, J. D. 1995 Transcription, splicing and editing of plastid RNAs in the non-photosynthetic plant *Epifagus virginiana*. *Plant Mol. Biol.* **29**, 721–733.

Eraso, J. M. & Kaplan, S. 2000 From redox flow to gene regulation: role of the PrrC protein of *Rhodobacter sphaeroides* 2.4.1. *Biochemistry* **39**, 2052–2062.

Eraso, J. M. & Kaplan, S. 2002 Redox flow as an instrument of gene regulation. *Meth. Enzymol.* **348**, 216–229.

Escoubas, J. M., Lomas, M., Laroche, J. & Falkowski, P. G. 1995 Light-intensity regulation of cab gene-transcription is signaled by the redox state of the plastoquinone pool. *Proc. Natl Acad. Sci. USA* **92**, 10 237–10 241.

Ferea, T. L., Botstein, D., Brown, P. O. & Rosenzweig, R. F. 1999 Systematic changes in gene expression patterns following adaptive evolution in yeast. *Proc. Natl Acad. Sci. USA* **96**, 9721–9726.

Forsberg, J., Rosenquist, M., Fraysse, L. & Allen, J. F. 2001 Redox signalling in chloroplasts and mitochondria: genomic and biochemical evidence for two-component regulatory systems in bioenergetic organelles. *Biochem. Soc. Trans.* **29**, 403–407.

Fujita, Y. 1997 A study on the dynamic features of photosystem stoichiometry: accomplishments and problems for future studies. *Photosynth. Res.* **53**, 83–93.

Gabriel, K., Buchanan, S. K. & Lithgow, T. 2001 The alpha and the beta: protein translocation across mitochondrial and plastid outer membranes. *Trends Biochem. Sci.* **26**, 36–40.

Galvis, M. L. E., Allen, J. F. & Hakansson, G. 1998 Protein synthesis by isolated pea mitochondria is dependent on the activity of respiratory complex II. *Curr. Genet.* **33**, 320–329.

Gatenby, A. A. & Ellis, R. J. 1990 Chaperone function: the assembly of ribulose bisphosphate carboxylase-oxygenase. *A. Rev. Cell Biol.* **6**, 125–149.

GoldschmidtClermont, M. 1998 Coordination of nuclear and chloroplast gene expression in plant cells. *Int. Rev. Cytol. Surv. Cell Biol.* **177**, 115–180.

Gray, M. W. 1992 The endosymbiont hypothesis revisited. *Int. Rev. Cytol.* **141**, 233–357.

Gray, M. W. 1999 Evolution of organellar genomes. *Curr. Opin. Genet. Dev.* **9**, 678–687.

Gray, M. W. 2000 Evolutionary biology: mitochondrial genes on the move. *Nature* **408**, 302–305.

Gray, M. W. & Doolittle, W. F. 1982 Has the endosymbiont hypothesis been proven? *Microbiol. Rev.* **46**, 1–42.

Gray, M. W., Burger, G. & Lang, B. F. 1999 Mitochondrial evolution. *Science* **283**, 1476–1481.

Grebe, T. W. & Stock, J. B. 1999 The histidine protein kinase superfamily. *Adv. Microb. Physiol.* **41**, 139–227.

Green, B. R. & Kuhlbrandt, W. 1995 Sequence conservation of light-harvesting and stress-response proteins in relation to the 3-dimensional molecular structure of Lhcii. *Photosynth. Res.* **44**, 139–148.

Gruissem, W. & Tonkyn, J. C. 1993 Control mechanisms of plastid gene expression. *Crit. Rev. Plant Sci.* **12**, 19–55.

Hackstein, J. H. P. (and 11 others) 2001 Hydrogenosomes: convergent adaptations of mitochondria to anaerobic environments. *Zool. Anal. Complex Syst.* **104**, 290–302.

Harrison, M. A., Keen, J. N., Findlay, J. B. C. & Allen, J. F. 1990 Modification of a Glnb-like gene-product by photosynthetic electron-transport in the cyanobacterium *Synechococcus* 6301. *FEBS Lett.* **264**, 25–28.

Harvey, A. J., Kind, K. L. & Thompson, J. G. 2002 Redox regulation of early embryo development. *Reproduction* **123**, 479–486.

Heathcote, P., Fyfe, P. K. & Jones, M. R. 2002 Reaction centres: the structure and evolution of biological solar power. *Trends Biochem. Sci.* **27**, 79–87.

Henze, K. & Martin, W. 2001 How do mitochondrial genes get into the nucleus? *Trends Genet.* **17**, 383–387.

Herrmann, R. G. & Westhoff, P. 2001 Thylakoid biogenesis and dynamics: the result of a complex phylogenetic puzzle. In *Regulation of photosynthesis*, vol. 11 (ed. E.-M. Aro & B. Andersson), pp. 1–28. Dordrecht, The Netherlands: Kluwer.

Hihara, Y., Sonoike, K., Kanehisa, M. & Ikeuchi, M. 2002 Search for redox responsive genes in *Synechocystis* sp. PCC 6803 by using DNA microarrays. *Plant Cell Physiol.* **43**, S178.

Howe, C. J., Barbrook, A. C. & Lockhart, P. J. 2000 Organelle genes: do they jump or are they pushed? *Trends Genet.* **16**, 65–66.

Ikeuchi, M. 2002 Gene expression and responses to light stress in cyanobacteria. *Plant Cell Physiol.* **43**, S21.

Iuchi, S. & Lin, E. C. C. 1992 Mutational analysis of signal transduction by Arcb, a membrane sensor protein responsible for anaerobic repression of operons involved in the central aerobic pathways in *Escherichia coli*. *J. Bacteriol.* **174**, 3972–3980.

Jackson-Constan, D., Akita, M. & Keegstra, K. 2001 Molecular chaperones involved in chloroplast protein import. *Biochim. Biophys. Acta Mol. Cell Res.* **1541**, 102–113.

Jarvis, P. & Soll, M. 2001 Toc, Tic, and chloroplast protein import. *Biochim. Biophys. Acta Mol. Cell Res.* **1541**, 64–79.

Jordan, P., Fromme, P., Witt, H. T., Klukas, O., Saenger, W. & Krauss, N. 2001 Three-dimensional structure of cyanobacterial photosystem I at 2.5 ångström resolution. *Nature* **411**, 909–917.

Kaneko, T. (and 23 others) 1996*a* Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.* **3**, 109–136.

Kaneko, T. (and 23 others) 1996*b* Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions [Supplement]. *DNA Res.* **3**, 185–209.

Kanevski, I. & Maliga, P. 1994 Relocation of the plastid rbcL gene to the nucleus yields functional ribulose-1,5-bisphosphate carboxylase in tobacco chloroplasts. *Proc. Natl Acad. Sci. USA* **91**, 1969–1973.

Kim, J. & Mayfield, S. P. 1997 Protein disulfide isomerase as a regulator of chloroplast translational activation. *Science* **278**, 1954–1957.

Kirk, J. T. O. & Tilney-Bassett, R. A. E. 1978 *The plastids. Their chemistry, structure, growth and inheritance*. Amsterdam: Elsevier.

Koehler, C. M., Merchant, S. & Schatz, G. 1999 How membrane proteins travel across the mitochondrial intermembrane space. *Trends Biochem. Sci.* **24**, 428–432.

Komiya, T., Rospert, S., Schatz, G. & Mihara, K. 1997 Binding of mitochondrial precursor proteins to the cytoplasmic domains of the import receptors Tom70 and Tom20 is determined by cytoplasmic chaperones. *FASEB J.* **11**, 1196.

Konstantinov, Y. M., Lutsenko, G. N. & Podsosonny, V. A. 1995 Genetic functions of isolated maize mitochondria under model changes of redox conditions. *Biochem. Mol. Biol. Int.* **36**, 319–326.

Korab-Laskowska, M., Rioux, P., Brossard, N., Littlejohn, T. G., Gray, M. W., Lang, B. F. & Burger, G. 1998 The organelle genome database project (GOBASE). *Nucleic Acids Res.* **26**, 138–144.

Kovacs, L., Wiessner, W., Kis, M., Nagy, F., Mende, D. & Demeter, S. 2000 Short- and long-term redox regulation of photosynthetic light energy distribution and photosystem stoichiometry by acetate metabolism in the green alga, *Chlamydobotrys stellata*. *Photosynth. Res.* **65**, 231–247.

Kühlbrandt, W. 1994 Structure and function of the plant light-harvesting complex, LHC-II. *Curr. Opin. Struct. Biol.* **4**, 519–528.

Lang, B. F., Gray, M. W. & Burger, G. 1999 Mitochondrial genome evolution and the origin of eukaryotes. *A. Rev. Genet.* **33**, 351–397.

Lemire, B. D., Fankhauser, C., Baker, A. & Schatz, G. 1989 The mitochondrial targeting function of randomly generated peptide sequences correlates with predicted helical amphiphilicity. *J. Biol. Chem.* **264**, 20 206–20 215.

Leuenberger, D., Bally, N. A., Schatz, G. & Koehler, C. M. 1999 Different import pathways through the mitochondrial intermembrane space for inner membrane proteins. *EMBO J.* **18**, 4816–4822.

Li, H. & Sherman, L. A. 2000 A redox-responsive regulator of photosynthesis gene expression in the cyanobacterium *Synechocystis* sp. strain PCC 6803. *J. Bacteriol.* **182**, 4268–4277.

Lin, X. Y. (and 36 others) 1999 Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. *Nature* **402**, 761–768.

Lonsdale, D. M., Hodge, T. P., Howe, C. J. & Stern, D. B. 1983 Maize mitochondrial-DNA contains a sequence homologous to the ribulose-1,5-bisphosphate carboxylase large subunit gene of chloroplast DNA. *Cell* **34**, 1007–1014.

McFadden, G. I. 1999 Endosymbiosis and evolution of the plant cell. *Curr. Opin. Plant Biol.* **2**, 513–519.

McFadden, G. I. 2001 Chloroplast origin and integration. *Plant Physiol.* **125**, 50–53.

Mahler, H. R. & Raff, R. A. 1975 The evolutionary origin of the mitochondrion: a nonsymbiotic model. *Int. Rev. Cytol.* **43**, 1–124.

Manfredi, G., Fu, J., Ojaimi, J., Sadlock, J. E., Kwong, J. Q., Guy, J. & Schon, E. A. 2002 Rescue of a deficiency in ATP synthesis by transfer of MTATP6, a mitochondrial DNA-encoded gene, to the nucleus. *Nature Genet.* **30**, 394–399.

Marc, P., Margeot, A., Devaux, F., Blugeon, C., Corral-Debrinski, M. & Jacq, C. 2002 Genome-wide analysis of mRNAs targeted to yeast mitochondria. *EMBO Rep.* **3**, 159–164.

Marienfeld, J., Unseld, M. & Brennicke, A. 1999 The mitochondrial genome of *Arabidopsis* is composed of both native and immigrant information. *Trends Plant Sci.* **4**, 495–502.

Martin, W. 1999 A briefly argued case that mitochondria and plastids are descendants of endosymbionts, but that the nuclear compartment is not. *Proc. R. Soc. Lond.* B **266**, 1387–1395. (DOI 10.1098/rspb.1999.0792.)

Martin, W. & Herrmann, R. G. 1998 Gene transfer from organelles to the nucleus: how much, what happens, and why? *Plant Physiol.* **118**, 9–17.

Martin, W. & Muller, M. 1998 The hydrogen hypothesis for the first eukaryote. *Nature* **392**, 37–41.

Martin, W. & Russell, M. J. 2003 On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells. *Phil. Trans. R. Soc. Lond.* B **358**, 59–85. (DOI 10.1098/rstb.2002.1183.)

Martin, W., Stoebe, B., Goremykin, V., Hansmann, S., Hasegawa, M. & Kowallik, K. V. 1998 Gene transfer to the nucleus and the evolution of chloroplasts. *Nature* **393**, 162–165.

Martin, W., Hoffmeister, M., Rotte, C. & Henze, K. 2001 An overview of endosymbiotic models for the origins of eukaryotes, their ATP-producing organelles (mitochondria and hydrogenosomes), and their heterotrophic lifestyle. *Biol. Chem.* **382**, 1521–1539.

Martin, W., Rujan, T., Richly, E., Hansen, A., Cornelsen, S., Lins, T., Leister, D., Stoebe, B., Hasegawa, M. & Penny, D. 2002 Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc. Natl Acad. Sci. USA* **99**, 12 246–12 251.

Millen, R. S. (and 12 others) 2001 Many parallel losses of infA from chloroplast DNA during angiosperm evolution with multiple independent transfers to the nucleus. *Plant Cell* **13**, 645–658.

Morse, D., Salois, P., Markovic, P. & Hastings, J. W. 1995 A nuclear-encoded form-II RUBISCO in dinoflagellates. *Science* **268**, 1622–1624.

Murakami, A., Kim, S. J. & Fujita, Y. 1997 Changes in photosystem stoichiometry in response to environmental conditions for cell growth observed with the cyanophyte *Synechocystis* PCC 6714. *Plant Cell Physiol.* **38**, 392–397.

Oh, J. I. & Kaplan, S. 2000 Redox signaling: globalization of gene expression. *EMBO J.* **19**, 4237–4247.

Ozawa, T. 1995 Mitochondrial-DNA mutations associated with aging and degenerative diseases. *Exp. Gerontol.* **30**, 269–290.

Palmer, J. D. 1997 Organelle genomes: going, going, gone! *Science* **275**, 790–791.

Pearson, C. K., Wilson, S. B., Schaffer, R. & Ross, A. W. 1993 NAD turnover and utilization of metabolites for RNA synthesis in a reaction sensing the redox state of the cytochrome-$b_{(6)}f$ complex in isolated chloroplasts. *Eur. J. Biochem.* **218**, 397–404.

Pfanner, N. & Geissler, A. 2001 Versatility of the mitochondrial protein import machinery. *Natl Rev. Mol. Cell Biol.* **2**, 339–349.

Pfanner, N. & Wiedemann, N. 2002 Mitochondrial protein import: two membranes, three translocases. *Curr. Opin. Cell Biol.* **14**, 400–411.

Pfannschmidt, T., Nilsson, A. & Allen, J. F. 1999*a* Photosynthetic control of chloroplast gene expression. *Nature* **397**, 625–628.

Pfannschmidt, T., Nilsson, A., Tullberg, A., Link, G. & Allen, J. F. 1999*b* Direct transcriptional control of the chloroplast genes psbA and psaAB adjusts photosynthesis to light energy distribution in plants. *IUBMB Life* **48**, 271–276.

Pfannschmidt, T., Allen, J. F. & Oelmuller, R. 2001*a* Principles of redox control in photosynthesis gene expression. *Physiol. Plant.* **112**, 1–9.

Pfannschmidt, T., Schutze, K., Brost, M. & Oelmuller, R. 2001*b* A novel mechanism of nuclear photosynthesis gene regulation by redox signals from the chloroplast during pho-

tosystem stoichiometry adjustment. *J. Biol. Chem.* **276**, 36 125–36 130.

Popot, J. L. & de Vitry, C. 1990 On the microassembly of integral membrane proteins. *A. Rev. Biophys. Biophys. Chem.* **19**, 369–403.

Race, H. L., Herrmann, R. G. & Martin, W. 1999 Why have organelles retained genomes? *Trends Genet.* **15**, 364–370.

Raven, J. A., Johnston, A. M., Parsons, R. & Kubler, J. 1994*a* The influence of natural and experimental high O_2 concentrations on O_2 evolving phototrophs. *Biol. Rev. Cambridge Phil. Soc.* **69**, 61–94.

Raven, J. A., Johnston, A. M., Parsons, R. & Kubler, J. 1994*b* The occurrence, and influence on photolithotrophs, of high oxygen concentrations. *Proc. R. Soc. Edinb.* B **102**, 193–201.

Robinson, C. 2000 The twin-arginine translocation system: a novel means of transporting folded proteins in chloroplasts and bacteria. *Biol. Chem.* **381**, 89–93.

Robinson, C. & Bolhuis, A. 2001 Protein targeting by the twin-arginine translocation pathway. *Natl Rev. Mol. Cell Biol.* **2**, 350–356.

Rodermel, S. 2001 Pathways of plastid-to-nucleus signaling. *Trends Plant Sci.* **6**, 471–478.

Roger, A. J. 1999 Reconstructing early events in eukaryotic evolution. *Am. Nat.* **154**, S146–S163.

Roger, A. J., Svard, S. G., Tovar, J., Clark, C. G., Smith, M. W., Gillin, F. D. & Sogin, M. L. 1998 A mitochondrial-like chaperonin 60 gene in *Giardia lamblia*: evidence that diplomonads once harbored an endosymbiont related to the progenitor of mitochondria. *Proc. Natl Acad. Sci. USA* **95**, 229–234.

Roodyn, D. B. & Wilkie, D. 1968 *The biogenesis of mitochondria. Methuen's monographs on biological subjects*. London: Methuen.

Rujan, T. & Martin, W. 2001 How many genes in *Arabidopsis* come from cyanobacteria? An estimate from 386 protein phylogenies *Trends Genet.* **17**, 113–120.

Saccone, C., Gissi, C., Lanave, C., Larizza, A., Pesole, G. & Reyes, A. 2000 Evolution of the mitochondrial genetic system: an overview. *Gene* **261**, 153–159.

Salzberg, S. L., White, O., Peterson, J. & Eisen, J. A. 2001 Microbial genes in the human genome: lateral transfer or gene loss? *Science* **292**, 1903–1906.

Sargent, F., Berks, B. C. & Palmer, T. 2002 Assembly of membrane-bound respiratory complexes by the Tat protein-transport system. *Arch. Microbiol.* **178**, 77–84.

Schatz, G. 1998 Protein transport: the doors to organelles. *Nature* **395**, 439–440.

Schemidt, R. A., Qu, J., Williams, J. R. & Brusilow, W. S. 1998 Effects of carbon source on expression of F_o genes and on the stoichiometry of the c subunit in the F_1F_o ATPase of *Escherichia coli*. *J. Bacteriol.* **180**, 3205–3208.

Selosse, M. A., Albert, B. R. & Godelle, B. 2001 Reducing the genome size of organelles favours gene transfer to the nucleus. *Trends Ecol. Evol.* **16**, 135–141.

Stern, D. B. & Lonsdale, D. M. 1982 Mitochondrial and chloroplast genomes of maize have a 12- kilobase DNA-sequence in common. *Nature* **299**, 698–702.

Stock, D., Gibbons, C., Arechaga, I., Leslie, A. G. W. & Walker, J. E. 2000 The rotary mechanism of ATP synthase. *Curr. Opin. Struct. Biol.* **10**, 672–679.

Stoebe, B. & Maier, U. G. 2002 One, two, three: nature's tool box for building plastids. *Protoplasma* **219**, 123–130.

Stoebe, B., Martin, W. & Kowallik, K. V. 1998 Distribution and nomenclature of protein-coding genes in 12 sequenced chloroplast genomes. *Plant Mol. Biol. Rep.* **16**, 243–255.

Stupar, R. M., Lilly, J. W., Town, C. D., Cheng, Z., Kaul, S., Buell, C. R. & Jiang, J. M. 2001 Complex mtDNA constitutes an approximate 620 kb insertion on *Arabidopsis thaliana* chromosome 2: implication of potential sequencing errors caused by large-unit repeats. *Proc. Natl Acad. Sci. USA* **98**, 5099–5103.

Surpin, M., Larkin, R. M. & Chory, J. 2002 Signal transduction between the chloroplast and the nucleus. *Plant Cell* **14**, S327–S338.

Tjaden, J., Schwoppe, C., Mohlmann, T., Quick, P. W. & Neuhaus, H. E. 1998 Expression of a plastidic ATP/ADP transporter gene in *Escherichia coli* leads to a functional adenine nucleotide transport system in the bacterial cytoplasmic membrane. *J. Biol. Chem.* **273**, 9630–9636.

Tovar, J., Fischer, A. & Clark, C. G. 1999 The mitosome, a novel organelle related to mitochondria in the amitochondrial parasite *Entamoeba histolytica*. *Mol. Microbiol.* **32**, 1013–1021.

Trebitsh, T. & Danon, A. 2001 Translation of chloroplast psbA mRNA is regulated by signals initiated by both photosystems II and I. *Proc. Natl Acad. Sci. USA* **98**, 12 289–12 294.

Tullberg, A., Alexciev, K., Pfannschmidt, T. & Allen, J. F. 2000 Photosynthetic electron flow regulates transcription of the psaB gene in pea (*Pisum sativum* L.) chloroplasts through the redox state of the plastoquinone pool. *Plant Cell Physiol.* **41**, 1045–1054.

Turmel, M., Lemieux, C., Burger, G., Lang, B. F., Otis, C., Plante, I. & Gray, M. W. 1999 The complete mitochondrial DNA sequences of *Nephroselmis olivacea* and *Pedinomonas minor*: two radically different evolutionary patterns within green algae. *Plant Cell* **11**, 1717–1729.

Unden, G. & Bongaerts, J. 1997 Alternative respiratory pathways of *Escherichia coli*: energetics and transcriptional regulation in response to electron acceptors. *Biochim. Biophys. Acta Bioenerg.* **1320**, 217–234.

Unseld, M., Marienfeld, J. R., Brandt, P. & Brennicke, A. 1997 The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366 924 nucleotides. *Nature Genet.* **15**, 57–61.

Van der Giezen, M., Slotboom, D. J., Horner, D. S., Dyal, P. L., Harding, M., Xue, G. P., Embley, T. M. & Kunji, E. R. S. 2002 Conserved properties of hydrogenosomal and mitochondrial ADP/ATP carriers: a common origin for both organelles. *EMBO J.* **21**, 572–579.

Von Heijne, G. 1986 Why mitochondria need a genome. *FEBS Lett.* **198**, 1–4.

Westphal, S., Soll, J. & Vothknecht, U. C. 2001 A vesicle transport system inside chloroplasts. *FEBS Lett.* **506**, 257–261.

Whitehouse, H. L. K. 1969 *Towards an understanding of the mechanism of heredity*. London: Edward Arnold Ltd.

Williams, B. A. P., Hirt, R. P., Lucocq, J. M. & Embley, T. M. 2002 A mitochondrial remnant in the microsporidian *Trachipleistophora hominis*. *Nature* **418**, 865–869.

Wilson, R. J. (and 10 others) 1996*a* Complete gene map of the plastid-like DNA of the malaria parasite *Plasmodium falciparum*. *J. Mol. Biol.* **261**, 155–172.

Wilson, S. B., Davidson, G. S., Thomson, L. M. & Pearson, C. K. 1996*b* Redox control of RNA synthesis in potato mitochondria. *Eur. J. Biochem.* **242**, 81–85.

Winkler, H. H. & Neuhaus, H. E. 1999 Non-mitochondrial ATP transport. *Trends Biochem. Sci.* **24**, 64–68.

Wolfe, K. H., Morden, C. W. & Palmer, J. D. 1992 Small single-copy region of plastid DNA in the non-photosynthetic angiosperm *Epifagus virginiana* contains only two genes. Differences among dicots, monocots and bryophytes in gene organization at a non-bioenergetic locus. *J. Mol. Biol.* **223**, 95–104.

Wollman, F. A., Minai, L. & Nechushtai, R. 1999 The biogenesis and assembly of photosynthetic proteins in thylakoid membranes. *Biochim. Biophys. Acta Bioenerg.* **1411**, 21–85.

Yeliseev, A. A. & Kaplan, S. 2000 TspO of *Rhodobacter sphaeroides*: a structural and functional model for the mam-

malian peripheral benzodiazepine receptor. *J. Biol. Chem.* **275**, 5657–5667.
Zerges, W. 2002 Does complexity constrain organelle evolution? *Trends Plant Sci.* **7**, 175–182.
Zhang, Z. D., Green, B. R. & Cavalier-Smith, T. 1999 Single gene circles in dinoflagellate chloroplast genomes. *Nature* **400**, 155–159.
Zhang, Z. D., Cavalier-Smith, T. & Green, B. R. 2001 A family of selfish minicircular chromosomes with jumbled chloroplast gene fragments from a dinoflagellate. *Mol. Biol. Evol.* **18**, 1558–1565.
Zouni, A., Witt, H. T., Kern, J., Fromme, P., Krauss, N., Saenger, W. & Orth, P. 2001 Crystal structure of photosystem II from *Synechococcus elongatus* at 3.8 ångström resolution. *Nature* **409**, 739–743.

Discussion

D. S. Horner (*Dipartimento di Fisiologia e Biochimica Generali, University of Milan, Milan, Italy*). Given that there are components of these membrane-bound energy-transduction pathways that are consistently nuclear encoded, not the key ones you pointed out but others, do you expect perhaps a slower, perhaps a more crude, but yet a redox communication between the organelle and expression of nuclear-encoded genes?

J. F. Allen. There is, and there is no question about it. I think that Bogorad's lock-in hypothesis has something to say here, and there are other people who say it is all to do with assembling multi-subunit complexes. If you have control over the initial construction of the core around which the whole complex is eventually built, then redox control, or some sort of control over synthesis of that core will control construction of the whole thing eventually, and I think that may well be largely correct. I think it is possible in principle for nuclear genes to make precursors according to redox instructions, but it is not really so important that those precursors are here and now present in exactly the right quantities, because the precursors can to a limited extent hang around and wait. Actually there is proteolytic degradation of unattached precursors and so on, but I think the fact that these chloroplast and mitochondrially DNA-encoded subunits are at the core of these energy-transducing complexes is the crucial issue.

T. Cavalier-Smith (*Department of Zoology, University of Oxford, Oxford, UK*). In the case of mitochondria the subset of genes that remains is much more variable than you implied. It runs from three proteins in the minimal cases up to, I think, over 90. That is a big difference, and many of the genes in different lineages of protists that remain have nothing to do with redox control; they are things like ribosomal proteins and so on. So I think that even though what you say about redox control of certain genes may actually be an important thing, it is not the fundamental, overall reason why the genomes are maintained. It may be part of the explanation for why certain genes are always included in that subset, but even in mitochondria I think that cytochrome *b* is the only one that is always present. So I think that both this great diversity of what has been retained in different lineages and the nucleomorph case indicate that evolution was able to do one thing in one lineage 500 million years ago and never got around to doing it for 500 million years in another lineage, so there is a lot more haphazardness, and evolutionary accidents and contingency than a straight functional interpretation; there is a lot of just, sheer 'happened–or–didn't–happen', or carryover.

J. F. Allen. I was not implying anything about the range of mitochondrial genomes. I have a limited amount of time, so I just talked about what I could talk about within that limit. Of course there will be a range and this will depend upon the physiology and bioenergetics of the particular organism. But what I am trying to do is to see what the common properties of the irreducible core are, of components that are always encoded in the organelles, and cytochrome *b* is a perfect example because that really is the thing that binds these two all-important quinones, without which there is no Q cycle, without which there is no proton pumping and bioenergetics, and in *Plasmodium*, where the mitochondrion is reduced to this bare minimum, there is still a mitochondrial gene for cytochrome *b*. Now my argument about the ribosomal proteins, and all the rest of it, is that in order to be able to exert this all-important redox control over gene expression of certain core components, you need an *in situ* genetic system, and this can vary in composition for reasons that I cannot describe. Maybe you need RNA polymerase, or maybe you can import RNA polymerase. In plastids, you have both of these options. You are going to need ribosomes, transfer RNAs and the rest of the apparatus of gene expression, and that needs to be *in situ*, but we are still left with the question, why does there have to be any genetic system at all? The mature tobacco or spinach chloroplast seems to have an entire genetic system just to make one protein, D_1, which is all it ever needs to resynthesize. I think that this is a clue. It is worth having this very uneconomical, crazy system because there is an overriding requirement to connect electron transfer with gene expression directly and unambiguously—without recourse to a bureaucracy, if you like, which is what I think the nucleus is.

K. Willison (*Institute of Cancer Research, Chester Beatty Laboratories, London, UK*). These photosynthetic reaction complexes are a sort of baroque mixture of polypeptide chains and cofactors, and my understanding is that some of them in high photon flux can turn over within, say, 30 minutes, so I am wondering if there is a correlation between gene retention and the turnover rates of some of the components and the complexes; it seems to me that that could really be a limiting factor and the reason that the proteins had to be encoded locally in order to get the complexes to be assembled rapidly.

J. F. Allen. I think that is right. You can only alter the steady-state concentration of a protein complex by affecting the synthesis and breakdown, and therefore turnover. There is an extreme example, the D_1 protein encoded by the *psbA* gene of chloroplasts, which is always in the chloroplast genome, without exception. High light will tend to degrade this protein, and when the high light goes away there is an absolute need very quickly to bring it back up again. I think that is a very good example. Whether this argument applies to genes such as cytochrome oxidase I do not know.

C. J. Leaver (*Department of Plant Sciences, University of Oxford, Oxford, UK*). Your experiments tend to relate mainly to the gene expression side. What about protein, RNA turnover, the whole concept of redox control of that—do we know anything?

J. F. Allen. Well, quite a lot is known about chloroplasts. There is redox control of messenger RNA stability for this same D_1 protein, which has been described by Stephen Mayfield and his co-workers. The evidence is that the site of redox control is ferredoxin via thioredoxin, and the implication is that thioredoxin tells whether non-cyclic electron transport is working at all. When it is working there is a signal that says 'OK, now continue with all the protein synthesis'. It is not particularly specific, but more of a global effect. I think this is what Mayfield argues, but of course this is not inconsistent with my argument. I have concentrated on quinone pools and medium-potential redox regulatory control because that is where I can see signals about the dangerous consequences of getting it wrong, but there is control of other levels of gene expression. In mitochondria, I am not so sure. I would bounce the question back to you on that one.

W. Martin (*Institute of Botany III, Heinrich-Heine Universität, Düsseldorf, Düsseldorf, Germany*). The nucleomorph genome was cited as evidence in conflict with your model, but is that true? The nucleomorph genome is not in a bioenergetic organelle in that sense. It resides in the cytosol, so does the size of the nucleomorph genome work against your model? Is it really evidence against your model?

J. F. Allen. I do not know; why should it be evidence? I cannot explain why the nucleomorph has a genome. That is why I clutched at the straw, that it is storage for nucleic acid precursors.

T. Cavalier-Smith. The reason I introduced the nucleomorph is not to say it is direct evidence against your hypothesis, but to show that there probably was no single functional reason for the retention of every gene in every organelle. It may be that there are specific physiological reasons for the retention of certain genes, and therefore you might be right about a subset of genes, but I very much doubt if there are specific physiological reasons for the retention of every single gene that is retained differentially, because odd things happen in evolution. Some things are easy and sometimes things are difficult. Evolution is history.

J. F. Allen. Of course I agree with that. I do not think I can explain everything, but I do think I can explain something. If I have done that, then I am content.

GLOSSARY

AAC: ADP–ATP carrier
CES: control by epistasy of synthesis
CORR: co-location for redox regulation
LHC: light-harvesting complex
PCC: Pasteur culture collection
ORF: open reading frame
RUBISCO: ribulose-1,5-bisphosphate carboxylase–oxygenase
TAT: twin-arginine translocase

Visit http://www.pubs.royalsoc.ac.uk to see an electronic appendix to this paper.