Human Reproduction, Vol. 15, (Suppl. 2), pp. 11–17, 2000

Transcription and replication of mitochondrial DNA

David A.Clayton[1]

Howard Hughes Medical Institute, Chevy Chase, Maryland, USA

[1]To whom correspondence should be addressed at: Howard Hughes Medical Institute, 4000 Jones Bridge Road, Chevy Chase, MD 20815–6789, USA.
E-mail: clayton@hhmi.org

The physical isolation of mammalian mitochondrial DNA (mtDNA) over 30 years ago marked the beginning of studies of its structure, replication and the expression of its genetic content. Such analyses have revealed a number of surprises: novel DNA structural features of the circular genome such as the displacement loop (D-loop); multiple sized and deleted forms of the circular genome; a minimal set of mitochondrially encoded rRNAs and tRNAs needed for translation; a bacteriophage-like, nuclear-encoded mitochondrial RNA polymerase for transcription; and a direct linkage between transcription and the commitment to replication of the leading mtDNA strand that centres on the nuclear encoded mitochondrial transcription factor A. One of the more recent revelations is the existence, near the D-loop, of an atypical, stable RNA–DNA hybrid (or R-loop) at the origin of mammalian leading-strand DNA replication, composed of the parent DNA strands and an RNA transcript. In mammalian mitochondrial systems, all of the proteins known to be involved in DNA replication are encoded in the nucleus. Thus alterations and deficiencies in mtDNA replication must arise from mutations in mtDNA regulatory sequences and nuclear gene defects. Further studies of the relationships between nuclear-encoded proteins and their mtDNA target sequences could result in strategies to manipulate genotypes within cellular mtDNA populations.

Key words: mitochondrial transcription/mtDNA/mtDNA replication/mtRNA polymerase/mt transcription factor A

Introduction

It was the fact that mammalian mitochondrial DNA (mtDNA) exists as a closed circular DNA species that enabled its initial isolation in dye-caesium chloride buoyant density gradients. This, in turn, permitted studies of its physical size and complexity. Because of a bias in base composition between the two helical strands, it was possible to physically separate and isolate a heavy strand (or H-strand) and a light strand (or L-strand) in high pH density gradients. Thus mtDNA enjoyed an early advantage in being the subject of assessing the differential capacities of complementary DNA strands, e.g. coding or non-coding status.

The hallmark of mammalian mtDNA is the D-loop form (Figure 1) which results from template-directed termination of H-strand

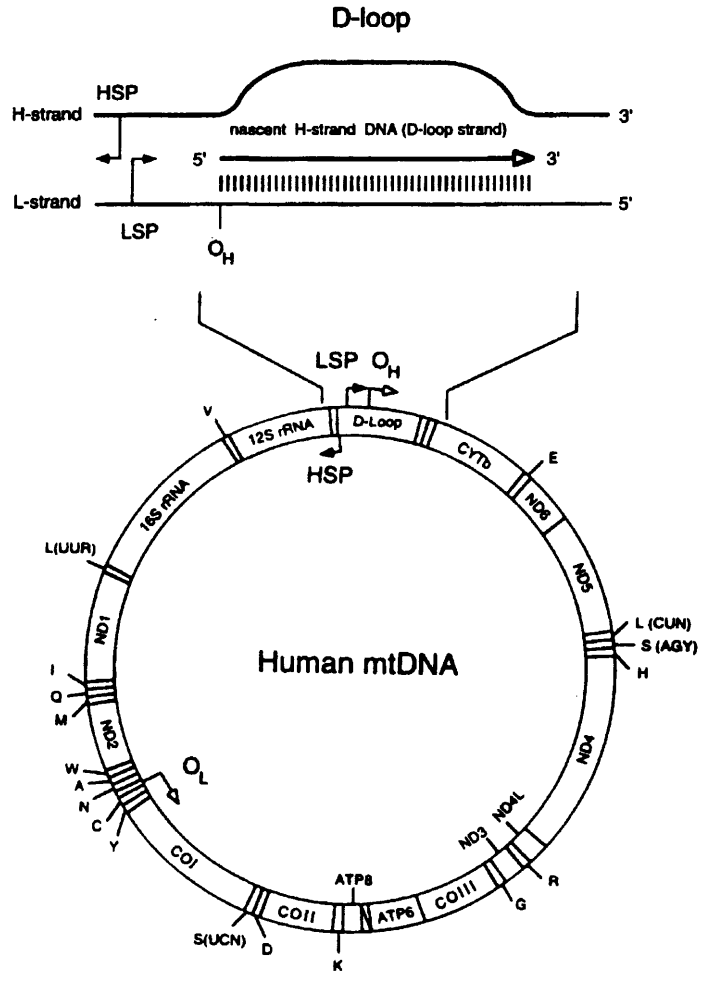

Figure 1. Schematic representation of the human mitochondrial genome, illustrating the position of key transcription promoters and points of origin of replication. ***Transcription:*** Heavy strand (H-strand) transcription starts at the H-strand promoter (HSP) and proceeds counter-clockwise on the diagram, commencing with ribosomal RNAs, and including 12 messenger RNAs that code for oxidative phosphorylation proteins, as well as 14 transfer RNAs. Light strand (L-strand) transcription starts at the L-strand promoter (LSP) and proceeds clockwise on the diagram, commencing with RNA primers for replication and, beyond the D-loop, coding for 8 of 22 transfer RNAs and 1 messenger RNA. Long polycistronic RNA chains are formed that are subsequently cleaved precisely into separate RNA species. ***Replication:*** Mitochondrial DNA replication requires RNA primers synthesized downstream from LSP (see text). The nascent H-strand forms from the L-strand template commencing at O_H and proceeding clockwise on the diagram, first displacing the parental H-strand to form a bubble (the D-loop). When the (single) mitochondrial replisome responsible for replication proceeds clockwise past the D-loop region, two thirds of the nascent H-strand is formed before a point is reached at which nascent L-strand synthesis can start at O_L. As a newly exposed single-stranded template sequence in the H-strand forms a hairpin to constitute O_L, H-strand replication (into a nascent L strand) commences in the opposite direction (counter-clockwise in the diagram), also continuously. Both copy strands are thus replicated as leading strands (5′→3′ directed) rather than lagging strands. The progeny molecules are released as dissimilar free circles. The new double-stranded mtDNA molecule is formed by removal of the RNA primers, gap-filling, introduction of superhelical turns and closure of the circle. ***Translation:*** The 13 protein-coding genes comprise: *complex I, or NADH dehydrogenase*: seven subunits (ND1, ND2, ND3, ND4, ND4L, ND5, ND6); *complex III*: cytochrome b (CYTb); *complex IV, or cytochrome c oxidase*: subunits I, II and III (COI, COII, COIII); and *complex V, or H^+-translocating ATP synthase*: subunits 6 and 8 (ATP6, ATP8).The 22 transfer RNA genes (for 20 amino acids) comprise: *L-strand*: *proline* (P, not labelled), *glutamic acid* (E), *serine* (S-UCN), *tyrosine* (Y), *cysteine* (C), *asparagine* (N), *alanine* (A), *glutamine* (Q); *H-strand*: *phenylalanine* (F, not labelled)), *valine* (V), *leucine* (L-UUR), *isoleucine* (I), *methionine* (M), *tryptophan* (W), *aspartic acid* (D), *lysine* (K), *glycine* (G), *arginine* (R), *histidine* (H), *serine* (S-AGY), *leucine* (L-CUN), *threonine* (T, not labelled). Adapted from Clayton (1991) and Shadel and Clayton (1997), with permission.

Table I. Mitochondrial genomic transcription products from 'clockwise' and 'counter-clockwise' transcription[a]

L-strand (clockwise) transcripts	H-strand (counter-clockwise) transcripts
RNA primers	2 rRNAs (12S, 16S)
8 tRNAs (P,E, S, Y, C, N, A, Q)	14 tRNAs (F, V, L, I, M, W, D, K, G, R, H, S, L, T)
1 mRNA (ND6)	12 mRNAs

[a]For more details see Jansen (2000).

DNA synthesis <1 kb after initiation. Understanding the nature of the D-loop is fundamental to an appreciation of how mtDNA replicates. Intrinsic to the process is the use of employment of DNA-dependent RNA synthesis (i.e. transcription) in the initiation of replication, which commences with the synthesis of nascent H-strand from the parent L-strand in this D-loop region. The first 1000 or so bp of new or nascent H-strand remains associated with the circular parental molecule, forming a triple-stranded structure (Figure 1) until replication is permitted to progress further around the circle. It is not known whether initiation must begin anew, or whether pre-existing D-loop stands are elongated. Once replication is underway, mammalian mtDNA molecules replicate by an asynchronous-displacement mechanism involving two unidirectional origins. The L-strand replicates two thirds of its extent (as a new H-strand) before H-strand replication (as a new L-strand) starts (see Figure 1 for details). Replication is both slow (it takes >1 h to synthesize 16 kb) and relaxed (there are no known cell cycle-dependent restrictions).

Transcription is also unique (Table I). Depending on which mtDNA strand is to be the template, transcription occurs from the light-strand promoter (LSP) or the heavy-strand promoter (HSP), in opposite directions around the entire genomic circle. The nature of promoter selection for a given transcription event is not known, but affinity for mitochondrial transcription factor A (mtTFA) (described below) may play a role. Fisher and Clayton (1988) showed that the affinity of the two major promoters for this activator was different; i.e. the promoter for D-loop synthesis (LSP) responds to lower concentrations of the mtTFA protein, so in principle it would be preferentially activated, at least under limiting concentrations of mtTFA. Transcription from the LSP proceeds from upstream of the D-loop, through it, and then copies the genes on this mtDNA strand. It is unknown whether transcription can occur from both promoters at the same time on the same molecule, but the relatively slow pace of transcription suggests that such an event would not normally be required. At the end of synthesis, polycistronic transcripts from each strand are cleaved precisely to generate the separate mature mRNAs, tRNAs and rRNAs (Clayton, 1984; Attardi and Schatz, 1988).

The first replicative event: transcription initiation at the L-strand promoter

Mammalian mitochondrial promoters have a bipartite structure consisting of one region encompassing the site of transcription initiation, and a target sequence immediately upstream that serves as the binding site for the nuclear-encoded mtTFA (Chang and Clayton, 1984; 1986; Fisher and Clayton, 1988; Fisher *et al.*, 1987, 1989). Currently the only mitochondrial transcription factor known in mammals, mtTFA bends and unwinds DNA upon binding, a characteristic typical of the high-mobility-group (HMG)-box family of proteins that this transcription factor belongs

to. It is most likely the case that the fundamental mtTFA activation step for replication of mtDNA is to bend mtDNA immediately (~10–40 bp) upstream from the transcription initiation site, permitting transcription of RNA to commence from exposed L-strand DNA.

Transcription initiation requires a nuclear encoded mitochondrial RNA polymerase, irrespective of whether the transcription is primarily for RNA synthesis (from either strand's promoter, see Table I) or primarily for initiation of replication (the L-strand promoter). Nucleotide sequencing of the *Saccharomyces cerevisiae* (yeast) mtRNA polymerase gene (the first isolated) revealed that the enzyme is homologous to RNA polymerases from the bacteriophages T7, T3 and SP6 (Masters *et al.*, 1987). Subsequent studies from multicellular animals have shown that their mtRNA polymerases are bacteriophage-like enzymes.

Wang and Shadel (1999) have recently demonstrated that the *N*-terminal domain of mtRNA polymerase is required for the maintenance of wild-type mtDNA in yeast – further evidence that there is a fundamental role for transcription in the process of mtDNA replication.

Formation of a stable RNA–DNA hybrid structure is the second step

Precise mapping of RNA and DNA species complementary to the D-loop region has provided compelling evidence that RNA derived from the LSP can serve as a primer for H-strand DNA replication (Chang and Clayton, 1985; Chang *et al.*, 1985). An early comparison of the nucleotide sequences of human, mouse and rat D-loop mtDNA revealed three *conserved sequence blocks* (termed CSB I, CSB II, and CSB III) (Walberg and Clayton, 1981). The CSBs, particularly CSB I, were thought to be involved in some aspect of mtDNA regulation or replication because of their physical location near the leading-strand origin (O_L). The CSBs could be involved in formation of a properly configured RNA substrate for subsequent processing into the mitochondrial primer RNAs and then elongation by mtDNA polymerase.

A detailed study of the mammalian RNA–DNA hybrid using mouse leading-strand origins has revealed an unusual structure (Lee and Clayton, 1996, 1997). An RNA designed to mimic the mouse LSP transcript can assemble on a supercoiled origin-containing mtDNA template during transcription by SP6 RNA polymerase (similar to mtRNA polymerase) *in vitro*. The resultant molecule is an extremely stable R-loop (a structure consisting of the two parental DNA strands and one RNA strand). Thus, the mtDNA sequence at the origin of replication contains all of the information required for formation of this extremely stable RNA–DNA hybrid. Examination of the RNA structure by enzymic digestion and chemical modification suggests that the RNA–DNA hybrid is of a unique type, involving interactions among all three strands, rather than a conventional R-loop structure with the singular RNA–DNA duplex region described earlier.

Primer RNA processing

For an RNA molecule to serve as a primer for initiation of DNA replication it must have a 3′-hydroxyl group available for extension by DNA polymerase. To provide such a primer, the new *LSP* RNA transcript must either terminate or be processed. A site-specific mtRNA-processing endoribonuclease (RNase MRP) was first identified in mouse and human cells, where it was shown to be able to process origin-containing RNA substrates at sites that match some of the DNA replication priming sites. The enzyme contains, in addition to protein components, an RNA essential for its activity (Chang and Clayton, 1987, 1989; Topper and Clayton,

1990; Schmitt *et al.*, 1993). RNase MRP activity and genes encoding the RNA component of the enzyme have now been characterized from a variety of organisms (Schmitt *et al.*, 1993). The enzymic activity, as well as the structural features of the RNA component, have been well conserved through evolution.

The pattern of RNA cleavage by RNase MRP is consistent with a role for the enzyme in providing primers for mtDNA replication. Curiously, cleavage at the majority of known RNA–DNA transition sites (including major transition sites) does not occur when single-stranded, origin-containing RNAs are used as the substrate. The enzyme acts only on the triple-stranded DNA–RNA hybrid described above. Testing RNase MRP with an appropriate R-loop substrate has given a cleavage pattern matching precisely the majority of RNA to DNA transition sites that have been mapped in mouse mitochondria *in vivo* (Chang and Clayton, 1987). Interestingly, these cleavages were completely dependent on the presence of CSB I. Similar results were obtained using human RNase MRP on a human R-loop substrate. These data argue strongly that RNase MRP is indeed necessary for the processing that produces the RNA primers in vertebrate mitochondria and further implicates the RNA–DNA hybrid as the substrate for the RNA processing that leads to formation of the primers for leading-strand replication.

Analyses now from yeast to humans have provided compelling evidence for transcription-priming mechanisms in the initiation of mtDNA replication. Although these systems' transcription-priming mechanisms share basic features, dependence of initiation on a specific CSB is not strictly conserved; nonetheless CSB I is almost invariably present and is usually associated with the initiation of D-loop DNA synthesis at the leading-strand origin. The variable occurrence of CSB II and CSB III might be a reflection of the fact that the substrate for RNA primer formation is apparently a complex RNA–DNA structure (Chang and Clayton, 1987; Lee and Clayton, 1996), which may be able to be formed despite a variety of different D-loop arrangements. Thus, the high sequence variability in the D-loop region could be a consequence of nucleic acid structural requirements. It has long been known that the D-loop region contains potential cloverleaf (tRNA-like) structures that are relatively conserved in many vertebrates (Brown *et al.*, 1986; Saccone *et al.*, 1991; Quinn and Wilson, 1993). With regard to primer RNA processing, the cloverleaf structure predicted to occur near CSB I (Brown *et al.*, 1986) is relevant because this region marks the boundary of the RNA–DNA hybrid structure in the human and mouse, and flanks the major initiation site for leading-strand DNA synthesis in most vertebrates *in vivo*. In this regard, it is interesting to note the similarities between the RNase MRP and RNase P (RNase P being the common cellular and organellar tRNA processing activity). Although the enzymes are distinct, they are related in structure and they share an ability to recognize tRNA structures.

Mitochondrial DNA polymerase γ

In vertebrates, there is apparently only one DNA polymerase (mitochondrial DNA polymerase γ, or pol γ) devoted to mtDNA synthesis. Pol γ is distinguished from other cellular DNA polymerases by certain chemical criteria, including high activity with synthetic RNA templates *in vitro*, inhibition by both *N*-ethylmaleimide and dideoxynucleoside triphosphates, resistance to aphidicolin, and stimulation by salt. Pol γ from all sources studied so far co-purifies with a 3′-5′ exonuclease domain, probably responsible for the very high fidelity reported for pol γ (Kunkel, 1985; Wernette *et al.*, 1988). (The high mutation rate of mtDNA observed *in vivo* is prob-

ably not due to any inherent inaccuracy in mtDNA polymerase.)

Pol γ consists of two polypeptides: a larger catalytic subunit (~125–140 kDa) that harbours both DNA polymerization and exonuclease activity, and a smaller subunit (35–54 kDa) of unknown function (Shadel and Clayton, 1997). Its heterodimeric composition immediately raised questions about the function of the smaller unit. Isolation of the genes encoding pol γ from various organisms has recently clarified the puzzle, the inference being that the smaller subunit of mtDNA polymerase possesses structural similarities to aminoacyl-tRNA synthetases, suggesting a tRNA-like primer RNA recognition phenomenon (Fan *et al.*, 1999).

Regulating mtDNA copy number

A central question in organelle biogenesis is to determine the mechanism by which the cell controls mtDNA copy number. Based on the mechanism of replication outlined above, it is reasonable to suggest that the rate of mtDNA replication is determined by a balance between transcription initiation from the LSP and subsequent RNA processing steps to provide replication primers. There is no evidence to suggest that mtDNA copy number is regulated by the levels of mtDNA polymerase or other proteins associated with DNA strand elongation (which is not to say that mutations in these proteins would not result in less, or even loss of, mtDNA). This puts the events of transcription-initiation first, which currently is thought to comprise mtRNA polymerase and the mtTFA transcriptional activator. In this regard, the recent report of Wu *et al.* (1999) is relevant. They showed that direct activation of the nuclear transcription factor NRF-1 results in stimulation of mitochondrial biogenesis coincident with increased synthesis of mtTFA by virtue of NRF-1 interaction as an activator of the nuclear gene for mtTFA. This result, along with studies in which loss of mtTFA results in loss of mtDNA and mitochondrial function (Larssen *et al.*, 1998; Wang *et al.*, 1999), raises the probability that mtTFA is a central feature of not only mtDNA transcription and replication, but of organelle biogenesis itself.

References

Attardi, G. and Schatz, G. (1988) Biogenesis of mitochondria. *Ann. Rev. Cell. Biol.*, **4**, 289–333.

Brown, G.G., Gadaleta, G., Pepe, G. *et al.* (1986) Structural conservation and variation in the D-loop-containing region of vertebrate mitochondrial DNA. *J. Mol. Biol.*, **192**, 503–511.

Chang, D.D. and Clayton, D.A. (1984) Precise identification of individual promoters for transcription of each strand of human mitochondrial DNA. *Cell*, **36**, 635–643.

Chang, D.D. and Clayton, D.A. (1985) Priming of human mitochondrial DNA replication occurs at the light-strand promoter. *Proc. Natl Acad. Sci. USA*, **82**, 351–355.

Chang, D.D. and Clayton, D.A. (1986) Precise assignment of the light-strand promoter of mouse mitochondrial DNA: a functional promoter consists of multiple upstream domains. *Mol. Cell. Biol.*, **6**, 3253–3261.

Chang, D.D. and Clayton, D.A. (1987) A mammalian mitochondrial RNA processing activity contains nucleus-encoded RNA. *Science*, **235**, 1178–1184.

Chang, D.D. and Clayton, D.A. (1989) Mouse RNAase MRP RNA is encoded by a nuclear gene and contains a decamer sequence complementary to a conserved region of mitochondrial RNA substrate. *Cell*, **56**, 131–139.

Chang, D.D. and Clayton, DA. (1987) A novel endoribonuclease cleaves at a priming site of mouse mitochondrial DNA replication. *EMBO J.*, **6**, 409–417.

Chang, D.D., Hauswirth, W.W. and Clayton, D.A. (1985) Replication priming and transcription initiate from precisely the same site in mouse mitochondrial DNA. *EMBO J.*, **4**, 1559–1567.

Clayton, D.A. (1984) Transcription of the mammalian mitochondrial genome. *Ann. Rev. Biochem.*, **53**, 573–594.

Clayton, D.A. (1991) Replication and transcription of vertebrate mitochondrial DNA. *Ann. Rev. Cell. Biol.*, **7**, 453–478.

Fan, L., Sanschagrin, P.C., Kaguni, L.S. and Kuhn, L.A. (1999) The accessory subunit of mtDNA polymerase shares structural homology with aminoacyl-tRNA synthetases: implications for a dual role as a primer recognition factor and processivity clamp. *Proc. Natl Acad. Sci. USA*, **96**, 9527–9532.

Fisher, R.P. and Clayton, D.A. (1988) Purification and characterization of human mitochondrial transcription factor 1. *Mol. Cell. Biol.*, **8**, 3496–3509.

Fisher, R.P., Parisi, M.A. and Clayton, D.A. (1989) Flexible recognition of rapidly evolving promoter sequences by mitochondrial transcription factor 1. *Genes Dev.*, **3**, 2202–2217.

Fisher, R.P., Topper, J.N. and Clayton, D.A. (1987) Promoter selection in human mitochondria involves binding of a transcription factor to orientation-independent upstream regulatory elements. *Cell*, **50**, 247–258.

Jansen, R.P.S. (2000) Origin and persistence of the mitochondrial genome. *Hum. Reprod.*, **15** (Suppl. 2). 1–10.

Kunkel, T.A. (1985) The mutational specificity of DNA polymerases-alpha and -gamma during *in vitro* DNA synthesis. *J. Biol. Chem.*, **260**, 12866–12874.

Larsson, N.G., Wong, J., Wilhelmsson, H. *et al.* (1998) Mitochondrial transcription factor A is necessary for mtDNA maintenance and embryogenesis in mice. *Nature Genet.*, **18**, 231–326.

Lee, D.Y. and Clayton, D.A. (1996) Properties of a primer RNA-DNA hybrid at the mouse mitochondrial DNA leading-strand origin of replication. *J. Biol. Chem.*, **271**, 24262–24269.

Lee, D.Y. and Clayton, D.A. (1997) RNase mitochondrial RNA processing correctly cleaves a novel R loop at the mitochondrial DNA leading-strand origin of replication. *Genes Dev.*, **11**, 582–592.

Masters, B.S., Stohl, L.L. and Clayton, D.A. (1987) Yeast mitochondrial RNA polymerase is homologous to those encoded by bacteriophages T3 and T7. *Cell*, **51**, 89–99.

Quinn, T.W. and Wilson, A.C. (1993) Sequence evolution in and around the mitochondrial control region in birds. *J. Mol. Evol.*, **37**, 417–425.

Saccone, C., Pesole, G. and Sbisà, E. (1991) The main regulatory region of mammalian mitochondrial DNA: structure-function model and evolutionary pattern. *J. Mol. Evol.*, **33**, 83–91.

Schmitt, M.E., Bennett, J.L., Dairaghi, D.J. and Clayton, D.A. (1993) Secondary structure of RNase MRP RNA as predicted by phylogenetic comparison. *FASEB J.*, **7**, 208–213.

Shadel, G.S. and Clayton, D.A. (1997) Mitochondrial DNA maintenance in vertebrates. *Ann. Rev. Biochem.*, **66**, 409–435.

Topper, J.N. and Clayton, D.A. (1990) Characterization of human MRP/Th RNA and its nuclear gene: full length MRP/Th RNA is an active endoribonuclease when assembled as an RNP. *Nucleic Acids Res.*, **18**, 793–799.

Walberg, M.W. and Clayton, D.A. (1981) Sequence and properties of the human KB cell and mouse L cell D-loop regions of mitochondrial DNA. Characterization of human MRP/Th RNA and its nuclear gene: full length MRP/Th RNA is an active endoribonuclease when assembled as an RNP. *Nucleic Acids Res.*, **9**, 5411–5421.

Wang, J., Wilhelmsson, H., Graff, C. *et al.* (1999) Dilated cardiomyopathy and atrioventricular conduction blocks induced by heart-specific inactivation of mitochondrial DNA gene expression. *Nature Genet.*, **21**, 133–137.

Wang, Y. and Shadel, G.S. (1999) Stability of the mitochondrial genome requires an amino-terminal domain of yeast mitochondrial RNA polymerase. *Proc. Natl Acad. Sci. USA*, **96**, 8046–8051.

Wernette, C.M., Conway, M.C. and Kaguni, L.S. (1988) Mitochondrial DNA polymerase from *Drosophila melanogaster* embryos: kinetics, processivity, and fidelity of DNA polymerization. *Biochemistry*, **27**, 6046–6054.

Wu, Z., Puigserver, P., Andersson, U. *et al.* (1999) Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1. *Cell*, **98**, 115–124.