J. Inher. Metab. Dis. 15 (1992) 439–447

Structure an Function of the Mitochondrial Geno e

D. A. CLAYTON
Department of Developmental Biology, Stanford University School of Medicine, Stanford, CA 94305-5427, USA

Summary: It is now clear that molecular defects in human mitochondrial DNA play a significant role in human disease. Mitochondrial DNA mutations range from single base changes in the 16.5 kilobase-pair genome up to large deletions and rearrangements. Independently of the actual cause of a given mutation, it is possible to predict at least some of the consequences of changes in mitochondrial DNA sequence. This paper reviews our overall understanding of the mode and mechanism of mitochondrial DNA replication and transcription and how this relates to mitochondrial gene expression. This provides a background to anticipate the nature and extent of mitochondrial DNA sequence changes that might be of physiological consequence.

It will soon be 25 years since the first structural characterization of human mitochondrial DNA (mtDNA) appeared in the literature (Clayton and Vinograd 1967; Hudson and Vinograd 1967). It is perhaps fitting that those early works showed that mtDNA could exist in different topological forms and sizes. One of these, the unicircular dimer, was associated with human malignancy, in particular leukaemia (Clayton and Vinograd 1969). Although the relevance of mtDNA size changes to any fundamental event in neoplasia is unknown, its correlation is well documented (Clayton and Smith 1975) and stands as the first example of at least suspecting that mtDNA alterations might play a role in altered cell physiology.

There has been substantial progress in learning the basic modes of mtDNA replication (Clayton 1982) and transcription (Clayton 1984), and a good beginning has been made in characterizing the nuclear gene products required for these events (Clayton 1991). The essential features of these processes and what they tell us regarding the types of defects we might expect to cause problems for mitochondrial function are described below.

Replication of mitochondrial DNA

Mammalian mtDNAs have two separate and distinct origins of replication (Table 1). The origin of heavy (H)-strand synthesis (O_H) is located within the displacement-loop (D-loop) region of the genome and the origin of light (L)-strand synthesis (O_L)

Table 1 Definition of terms

Nomenclature	*Regulatory sequence/activity*
H strand	Heavy strand
L strand	Light strand
O_H	Origin of H-strand replication
O_L	Origin of L-strand replication
HSP	H-strand promoter
LSP	L-strand promoter
mtTF1	Mitochondrial transcription factor 1
mtTERM	Mitochondrial termination protein factor
RNase MRP	Site-specific endoribonuclease

is nested within a cluster of five tRNA genes well away from the D loop (Figure 1). A commitment to mtDNA replication begins by an initiation of H-strand synthesis that results in strand elongation for the entire length of the genome. Initiation of L-strand synthesis only occurs after O_L is exposed as a single-stranded template by displacement, and genomic replication has never been found to begin at this origin. Thus in this system, the H-strand (leading-strand) origin is the dominant element for

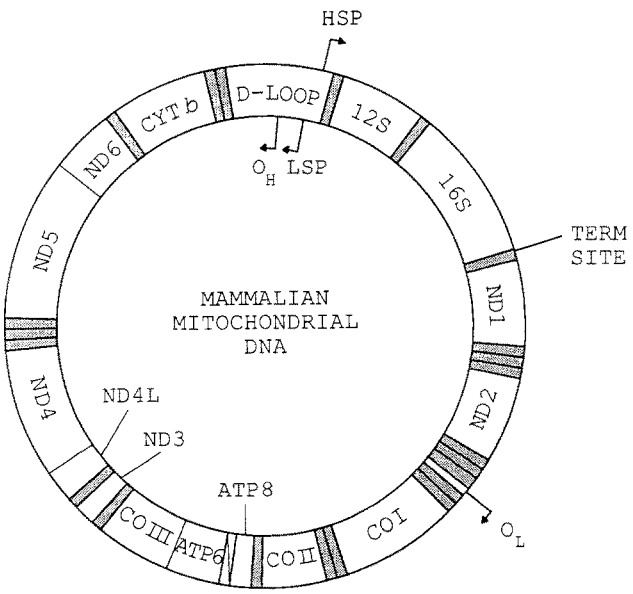

Figure 1 Representation of mammalian mtDNA. The shaded areas represent the 22 tRNA genes. Open regions denote control regions, rRNA genes, and protein-coding genes. These include the displacement-loop (D-loop) region, the 12S and 16S rRNA genes, and the genes for cytochrome *c* oxidase subunits I, II, and III (COI, COII, COIII), ATPase subunits 6 and 8 (ATP6, ATP8), cytochrome *b* (CYT*b*), and subunits 1,2,3,4,4L,5, and 6 of NADH dehydrogenase (ND1, ND2, ND3, ND4, ND4L, ND5, ND6). O_H and O_L demarcate the respective origins of heavy (H)- and light (L)-strand mtDNA synthesis. HSP and LSP mark the respective promoters for transcription from the H- and L-template strands. Transcripts terminate downstream of the 16S rRNA gene at a site in the gene for $tRNA^{Leu(UUR)}$, marked TERM

initiation of replication, while the L-strand (lagging-strand) origin plays an essential but secondary role.

Transcription of mitochondrial DNA

The D-loop region of vertebrate mtDNA has evolved as the control site for both transcription and replication (Figure 2). Biochemical analyses have been aimed at understanding the nature and complexity of protein–DNA interactions within this unique regulatory domain and have resulted in the characterization of a highly faithful *in vitro* system for transcriptional initiation at both promoters in the D loop. Each strand of the mammalian circular mtDNA genome is transcribed from a single major promoter. The two promoters, defined by mutational analyses, are situated approximately 150 base pairs apart, do not overlap, and function as complete and independent entities in *in vitro* assays.

Transcription of human mtDNA involves copying almost the entirety of the H-strand template sequence (producing RNA information of L-strand sequence) as well as an extensive portion of the L strand. Another feature of H-strand transcription is that it provides the necessary ribosomal RNA (rRNA) species as well as most of the messenger RNAs (mRNAs) for protein synthesis. Production of rRNAs appears to involve the only documented site of transcription termination (Christianson and Clayton 1986, 1988; Kruse et al 1989; Hess et al 1991) in this system, and it is located at the end of the 16S rRNA gene. With regard to L-strand transcription, an important

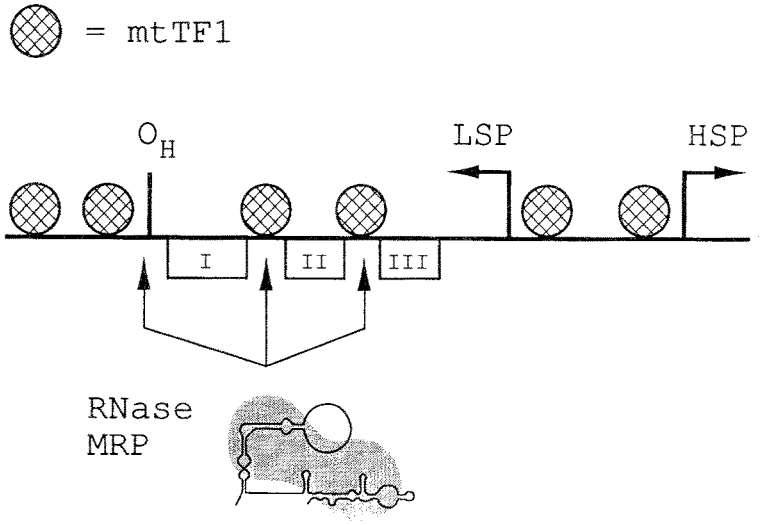

Figure 2 Schematic of the D loop. Boxes marked I, II, and III represent conserved sequence blocks I, II, and III in the D-loop region. Cross-hatched circles denote sites of transcription termination, and potential mtTF1 protein binding, in the D-loop region downstream of the LSP (Hess et al 1991) in addition to mtTF1 binding at the LSP and HSP. RNase MRP is shown diagrammatically based on its proposed structure (Topper and Clayton 1990). The vertical arrows pointing upward indicate sites of cleavage by RNase MRP (Karwan et al 1991)

point to note is that priming of leading-strand mtDNA replication also depends on initiation at the L-strand promoter.

After primary transcripts are produced, some enzymatic activity or activities must process these long RNAs into mature tRNAs and mRNAs. The most likely candidate for such an activity is one that would share some features with RNase P; this is because mitochondrial genes are interspersed between tRNA sequences, and RNase P is an activity known to be able to recognize and process tRNA sequences in a precise manner.

MITOCHONDRIAL GENE EXPRESSION

The overall flow of genetic information in the mitochondrial system involves the cooperative action of mtDNA and proteins and factors that are nuclear gene products and are present within the mitochondrial matrix. Figure 3 shows a few examples of such proteins involved. The important point is that all critical enzymatic activities and accessory proteins are nuclear gene products; none is known to be encoded by mtDNA. Gene expression begins by transcriptional activation at one or both mtDNA promoters. The transcriptional accessory protein (mtTF1) is involved in regulating the amount of transcripts produced from each promoter. It is possible that this protein is able to regulate transcription in a promoter-specific manner given that its relative affinities for the HSP and the LSP are different. That is, differing amounts of mtTF1 might result in different levels of H- and L-strand transcripts being produced.

The next likely level of control is the rate at which primary transcripts are processed. Given the very low abundance of primary transcripts in all cases studied

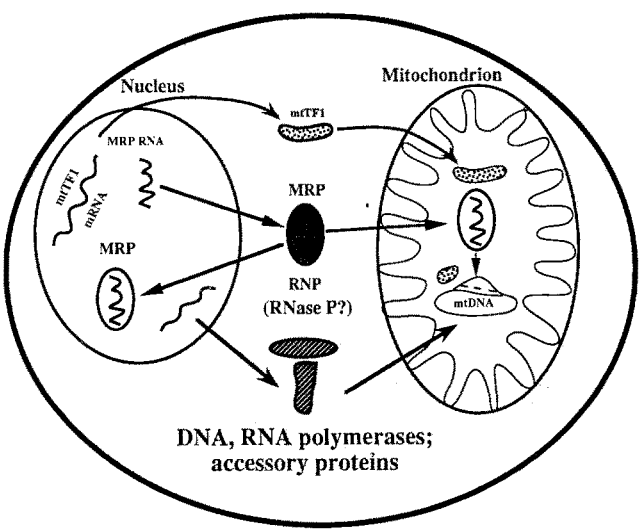

Figure 3 Schematic of transport pathways required for mtDNA replication and transcription. All known or suspected proteins required for these events are nuclear gene products that must be imported. At the current time, RNase MRP (a ribonucleoprotein) is the only activity likely to be located in both the nucleus and the mitochondrion

to date, it is probable that RNA processing, in general, is not rate-limiting. Of course, this need not be true in diseased tissues, and it is conceivable that defects in the RNA processing machinery might result in lack of proper mitochondrial gene expression. In the case of deletion mutations of mtDNA, it is possible that newly created RNA molecules (that encompass deletion end points) might possess structure that could tie up the RNA processing or translation machinery. If so, the overall phenotype could be lower levels of normal mtDNA expression.

Assuming a normal level of production of mature RNA species, the next regulation point is at the level of translation of mRNA into protein. This is an area of mitochondrial molecular biology that has received very little attention. This is particularly true in the vertebrate systems, owing in large part to the unexpected complexity of mitochondrial ribosomes and their natural low abundance relative to cellular ribosomal species. The only item of documented physiological consequence is the fact that chloramphenicol resistance in mammalian cells is known to be due to point mutations at the 3′ end of the large mitochondrial rRNA species (Blanc et al 1981). In fact, this mutation was the first identified that related to the functioning of animal mtDNA.

WHAT TO EXPECT IN HUMAN DISEASE

It is perhaps obvious to indicate that one would anticipate that deletions of mtDNA that removed critical regulatory regions for replication and expression would be lethal events for the molecule (Figure 4). This would include loss of the origins of

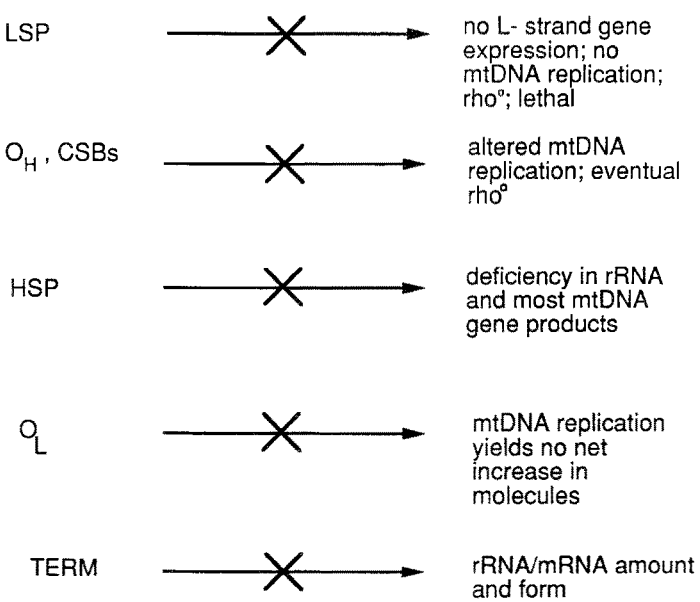

Figure 4 Predictions for loss of function of known regulatory sequences in mtDNA. The sequences are those defined in Table 1. The likely consequence of functional mutation in these sequences is shown on the right

replication and transcriptional promoters. In particular, loss of LSP function would be expected to result in the complete loss of a mtDNA molecule that carried such a mutation. The mutation could, in principle, occur in any critical region of the promoter. But based on our knowledge of human mtDNA promoter function, the most likely site of lethal mutation would be at the transcriptional start point itself (Hixson and Clayton 1985; Topper and Clayton 1989).

Loss of HSP function would not be expected to have any consequence on the ability of mtDNA to replicate. It would, however, create a defunct population of molecules from which both rRNA and most mitochondrial mRNAs would not be produced. Thus the phenotype would be one of lack of mitochondrial gene expression (28 out of the 37 genes would be silent) in a situation where a normal amount of mtDNA might well be present.

In cases of normal LSP function, but defects in the origin of leading-strand mtDNA replication (O_H), one would imagine a selective loss of such molecules. In fact, one could suppose that any sort of deleterious origin mutations would be simply being selected against in a multicopy genomic system.

Loss of the origin of L-strand replication (O_L) would result in a situation where such molecules, when replicated, would yield only one progeny circle; that is, there could never be a net increase in such molecules. However, in principle, a segregated progeny single-stranded circle (see Clayton 1982), if stable, could be primed for duplex synthesis by any available complementary RNA species.

Finally, as mentioned earlier, there is a termination site for transcription at the rRNA region. It has been shown that a very short sequence, a tridecamer, is a target for a protein that is able to sponsor transcription termination in the presence of a mitochondrial polymerase activity (Christianson and Clayton 1988). This sequence is of particular interest, since it has been independently implicated to be a site of mutation in the MELAS disorder (Goto et al 1991; Kobayashi et al 1990; Tanaka et al 1991). It has recently been shown that the standard MELAS mutation renders this sequence ineffective in supporting transcription termination in an *in vitro* assay system (Hess et al 1991). This sequence is also part of a tRNA gene, and it will be interesting to learn to what degree the mutation *in vivo* is affecting termination as opposed to being one of several candidates for tRNA dysfunction.

CONSEQUENCES OF SOME NUCLEAR MUTATIONS

Of course, not all problems of mitochondrial genetic expression need be related to changes in mtDNA itself. For example, loss of mtDNA polymerase or mitochondrial RNA polymerase might be expected to result in complete loss mtDNA. This should be a lethal condition (Figure 5). Mutations in the gene for the currently described mitochondrial transcriptional activator mtTF1 would be expected to affect, but not necessarily abolish, transcription and mtDNA replication. It is possible that other proteins can substitute for mtTF1, or that mitochondrial RNA polymerase itself (or with another accessory protein) can initiate transcription at a low level. This might be sufficient for maintaining mitochondrial gene expression in a fully differentiated cell and might provide sufficient primers for mtDNA replication under these

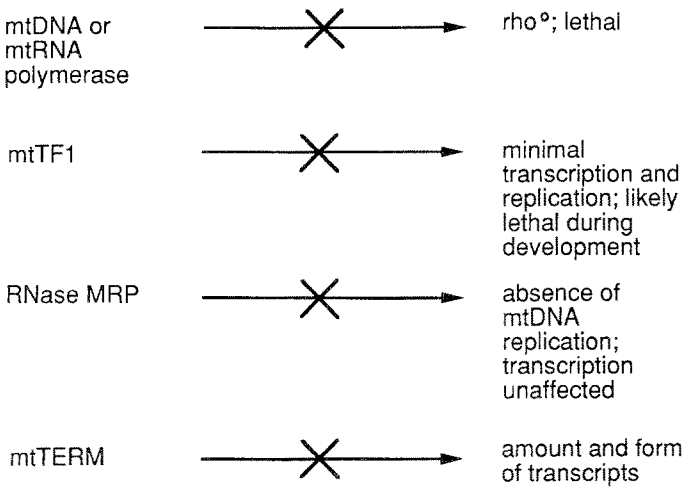

Figure 5 Consequences of loss of key nucleus-encoded proteins. Currently identified nuclear gene products involved in mtDNA replication and transcription are shown on the left. The likely consequence of dysfunction of each is shown on the right. This is only a small portion of the activities encoded in the nucleus that are required for mitochondrial nucleic acid synthesis

circumstances. More serious consequences would be expected during development, when demands on organelle biogenesis are the greatest.

Another activity thought to be important in mtDNA replication is site-specific endonuclease RNase mitochondrial RNA processing (MRP). This activity is involved in primer RNA metabolism for leading-strand DNA synthesis and its absence would be expected to affect the ability of mtDNA to replicate. At present, there is no role for this activity in mitochondrial transcription.

If the protein involved in transcription termination (mtTERM) were to be rendered ineffective, one would expect that the amount and form of H-strand transcripts would be altered. It is not clear whether there are alternative mechanisms or sufficient RNA processing capacity to produce the required amounts of rRNA under these circumstances. That is, in the absence of termination, it might be possible to provide the normal amount of ribosomal capacity by an enhanced level of RNA processing and turnover of excess mRNAs.

PROSPECTS FOR THE FUTURE

The clear examples from many laboratories that implicate mtDNA mutations with human disease have provided an important and exciting stimulus to work in this field. Given the power of today's molecular technology to delineate many of these mutations, one can predict that most of the spectrum of human disease relative to mitochondrial disorders will soon be made available. Along with this goes the hope that our understanding of how mtDNA is replicated and expressed will be similarly enhanced. This is likely to be the case given the inroads that have been made in the

detailed characterization of activities that involve early commitment to the processes of replication and transcription.

Further progress seems assured by more recent studies that suggest a greater commonality between these events in human cells and that of the classical and very well studied yeast mitochondrial system (Parisi and Clayton 1991; Diffley and Stillman 1991). Here we are learning that similar macromolecules are playing parallel roles in the two systems, which argues that genetic manipulations will possibly permit the introduction of human activities into a system that can then be approached both biochemically and genetically. Furthermore, such commonality argues that fundamental observations that may be made in the yeast system itself may be more applicable to the human case than previously imagined.

In any event, the correlations between mitochondrial sequence changes and altered physiology of mammalian cells provides a rich opportunity to identify the most critical regions of mtDNA and should permit investigators to focus sharply on those elements most relevant to organelle biogenesis and cellular vitality. In addition, the phenotypic study of known mtDNA mutations should be enhanced by the ability to introduce such mutations into human tissue culture lines for subsequent biochemical analyses (Chomyn et al 1991).

ACKNOWLEDGEMENTS

This work has been supported by grants from the National Institute of General Medical Sciences (GM-33088) and the American Cancer Society, Inc. (NP-9).

REFERENCES

Blanc H, Wright CT, Bibb MJ, Wallace DC, Clayton DA (1981) Mitochondrial DNA of chloramphenicol-resistant mouse cells contains a single nucleotide change in the region encoding the 3′ end of the large ribosomal RNA. *Proc Natl Acad Sci USA* **78**: 3789–3793.

Chomyn A, Meola G, Bresolin N, Lai ST, Scarlato G, Attardi G (1991) In vitro genetic transfer of protein synthesis and respiration defects to mitochondrial DNA-less cells with myopathy-patient mitochondria. *Mol Cell Biol* **11**: 2236–2244.

Christianson TW, Clayton DA (1986) *In vitro* transcription of human mitochondrial DNA: Accurate termination requires a DNA sequence that can function bidirectionally. *Proc Natl Acad Sci USA* **83**:6277–6281.

Christianson TW, Clayton DA (1988) A tridecamer DNA sequence supports human mitochondrial RNA 3′-end formation in vitro. *Mol Cell Biol* **8**: 4502–4509.

Clayton DA (1982) Replication of animal mitochondrial DNA. *Cell* **28**: 693–705.

Clayton DA (1984) Transcription of the mammalian mitochondrial genome. *Annu Rev Biochem* **53**: 573–594.

Clayton DA (1991) Nuclear gadgets in mitochondrial DNA replication and transcription. *TIBS* **16**: 107–111.

Clayton DA and Smith CA (1975) Complex mitochondrial DNA. *Int Rev Exp Pathol* **14**: 1–67.

Clayton DA, Vinograd J (1967) Circular dimer and catenate forms of mitochondrial DNA in human leukaemic leucocytes. *Nature* **216**: 652–657.

Clayton DA, Vinograd J (1969) Complex mitochondrial DNA in leukemic and normal human myeloid cells. *Proc Natl Acad Sci USA* **62**: 1077–1084.

Diffley JFX, Stillman B (1991) A close relative of the nuclear, chromosomal high-mobility group protein HMG1 in yeast mitochondria. *Proc Natl Acad Sci USA* **88**: 7864–7868.

Goto Y-i, Nonaka I, Horai S (1991) A mutation in the $tRNA^{Leu(UUR)}$ gene associated with the MELAS subgroup of mitochondrial encephalomyopathies. *Nature* **348**: 651–653.

Hess JF, Parisi MA, Bennett JL, Clayton DA (1991) Impairment of mitochondrial transcription termination by a point mutation associated with the MELAS subgroup of mitochondrial encephalomyopathies. *Nature* **351**: 236–239.

Hixson JE, Clayton DA (1985) Initiation of transcription from each of the two human mitochondrial promoters requires unique nucleotides at the transcriptional start sites. *Proc Natl Acad Sci USA* **82**: 2660–2664.

Hudson B, Vinograd J (1967) Catenated circular DNA molecules in HeLa cell mitochondria. *Nature* **216**: 647–652

Karwan R, Bennett JL, Clayton DA (1991) Nuclear RNase MRP processes RNA at multiple discrete sites: Interaction with an upstream G box is required for subsequent downstream cleavages. *Genes Dev* **5**: 1264–1276.

Kobayashi Y, Momoi MY, Tominaga K et al (1990) A point mutation in the mitochondrial $tRNA^{Leu}$(UUR) gene in MELAS (mitochondrial myopathy, encephalopathy, lactic acidosis and strokelike episodes). *Biochem Biophys Res Commun* **173**: 816–822.

Kruse B, Narasimhan N, Attardi G (1989) Termination of transcription in human mitochondria: Identification and purification of a DNA binding protein factor that promotes termination. *Cell* **58**: 391–397.

Parisi MA, Clayton DA (1991) Similarity of human mitochondrial transcription factor 1 to high mobility group proteins. *Science* **252**: 965–969.

Tanaka M, Ino H, Ohno K et al (1991) Mitochondrial DNA mutations in mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episodes (MELAS). *Biochem Biophys Res Commun* **174**: 861–868.

Topper JN, Clayton DA (1989) Identification of transcriptional regulatory elements in human mitochondrial DNA by linker substitution analysis. *Mol Cell Biol* **9**: 1200–1211.

Topper JN, Clayton DA (1990) Secondary structure of the RNA component of a nuclear/mitochondrial ribonucleoprotein. *J Biol Chem* **265**: 13254–13262.