Original Contributions

Incorporating Mouse Genome
Mammalian Genome
From Genotype to Phenotype

Characterization of the mouse genes for mitochondrial transcription factors B1 and B2

Anja Rantanen, Martina Gaspari, Maria Falkenberg, Claes M. Gustafsson, Nils-Göran Larsson

Department of Medical Nutrition and Department of Biosciences, Karolinska Institutet, NOVUM, Huddinge Hospital, S-14186 Huddinge, Sweden

Received: 10 July 2002/Accepted: 18 September 2002

Abstract

We have recently fully reconstituted the basal human mitochondrial transcription machinery in a pure *in vitro* system. Surprisingly, we found two different transcription factors (TFB1M and TFB2M) that each interact with mitochondrial RNA polymerase in human mitochondria, whereas there is only one such factor in budding yeast mitochondria. This unexpected finding raised important questions concerning the regulation of mitochondrial transcription in mammals in general and in other metazoans. We have now further analyzed putative homologs to TFB1M and TFB2M in different species. We mapped the mouse homologs, *Tfb1m* and *Tfb2m*, by linkage analysis to mouse Chr 17 and Chr 1, respectively. These regions display conserved linkage synteny with human Chr 6 and Chr 1, where *TFB1M* and *TFB2M* map. The intron-exon arrangements of *Tfb1m* and *TFB1M* and of *Tfb2m* and *TFB2M* were identical, and the promoter regions had similar predicted recognition elements for transcriptional factors NRF2 and Sp1. Northern blot analyses showed that *Tfb1m* and *Tfb2m* were ubiquitously expressed and had expression patterns that were very similar to the previously reported expression patterns for *TFB1M* and *TFB2M*. These findings show that *Tfb1m* and *Tfb2m* indeed are orthologs to *TFB1M* and *TFB2M*. Bioinformatic analyses indicated that most metazoans have two *TFBM* genes, since putative homologs to both *TFB1M* and *TFB2M* were found in *D. melanogaster*. Our data thus suggest that a duplication event of the *TFBM* gene in early metazoan evolution has permitted a more flexible regulation of mtDNA transcription, possibly in response to the complex physiological demands of multicellular organisms.

Introduction

Respiratory chain dysfunction is an important cause of inherited disease (Larsson and Clayton 1995) and has also been implicated in common age-associated human diseases and the ageing process (Larsson and Luft 1999; Wallace 1999). The mitochondrial DNA (mtDNA) is present in 10^3–10^4 copies per cell and encodes 13 proteins, which are all respiratory chain subunits (Anderson et al. 1981; Bibb et al. 1981). Nuclear genes encode the remaining ~80–100 respiratory chain subunits, all proteins necessary for replication and transcription of mtDNA and all other mitochondrial proteins. The biogenesis of the respiratory chain is thus dependent on the concerted expression of both nuclear and mtDNA genes. Dissection of the molecular mechanisms regulating this intergenomic cross-talk is necessary for understanding the regulation of oxidative phosphorylation capacity and may also provide novel insights into the pathogenesis of respiratory chain disorders (Lightowlers et al. 1997).

We have recently succeeded in defining the basal human mtDNA transcription machinery and fully reconstituted the system *in vitro* with pure proteins (Falkenberg et al. 2002). Surprisingly, there are two different factors, mitochondrial transcription factors B1 and B2 (TFB1M and TFB2M), that each interact with mitochondrial RNA polymerase, whereas there is only one such factor in budding yeast mitochondria. Faithful transcription initiation was obtained from templates containing the heavy and light strand promoters (HSP and LSP) of human mtDNA by adding purified recombinant mitochondrial RNA

Correspondence to: C.M. Gustafsson or Nils-Göran Larsson; E-mail: claes.gustafsson@mednut.ki.se or nils-goran.larsson@mednut.ki.se

polymerase (POLRMT) (Tiranti et al. 1997), mitochondrial transcription factor A (TFAM) (Parisi and Clayton 1991), and TFB1M or TFB2M (Falkenberg et al. 2002). Both TFB1M and TFB2M function as transcription initiation factors and form heterodimeric complexes with POLRMT. TFB2M was found to be at least 10 times more active than TFB1M in promoting transcription initiation. The *Saccharomyces cerevisiae* basal mitochondrial transcription machinery consists of only two factors, mitochondrial RNA polymerase (RPO41) (Masters et al. 1987) and its specificity factor MTF1 (Schinkel et al. 1987). The yeast TFAM homolog, ABF2, is not necessary for transcription, but rather functions as an mtDNA packaging factor (Parisi et al. 1993).

The human *TFB1M* and *TFB2M* genes were first identified because of their peptide sequence homology to the putative *Schizosaccharomyces pombe* MTF1 homolog but surprisingly they also showed strong homology to bacterial rRNA adenosine dimethyl transferases (Falkenberg et al. 2002). The TFB1M protein has been shown to bind the methyl group donor S-adenosylmethionine (McCulloch et al. 2002), but *in vitro* methylation activity has not been documented. The finding of two different factors that interact with POLRMT in mammalian mitochondria was unexpected and raised important questions concerning the regulation of mitochondrial transcription in mammals in general and in other metazoans. We have previously identified putative mouse orthologs to human *TFB1M* and *TFB2M*, denoted *Tfb1m* and *Tfb2m* (Falkenberg et al. 2002). In this paper, we have further characterized the *Tfb1m* and *Tfb2m* genes and their expression patterns. We have also used bioinformatics tools to identify possible homologs in other species. Our data show that most metazoans have two *TFBM* genes, thus suggesting an important role for both of these factors in regulating mtDNA gene expression.

Materials and methods

Gene mapping. The genes for *Tfb1m* and *Tfb2m* were mapped by using The Jackson Laboratory, BSB [(C57BL/6J × *Mus spretus*) × C57BL/6J] backcross panel services (http://www.jax.org/resources/documents/cmdata/). PCR assays were designed to distinguish between the C57BL/6J and *M. spretus Tfb1m* and *Tfb2m* alleles. For *Tfb2m*, the PCR products from *M. spretus* but not from the C57BL/6J template contained two *Hae*III restriction sites. For *Tfb2m*, the PCR products from C57BL/6J but not from the *M. spretus* template contained an *Ava*I restriction site. PCR products from DNA from the two parental mouse strains and the 94 backcross animals were digested with the appropriate enzymes (*Hae*III or *Ava*I), and the restriction fragments were analyzed by gel electrophoresis. The map loci for *Tfb1m and Tfb2m* were computed based on the previously mapped loci and markers (n > 920) on the BSB panel map.

Northern blot analyses. The *Tfb1m* and *Tfb2m* expression patterns were analyzed in different mouse tissues by probing a Clontech® MTN membrane with the corresponding cDNAs. The sizes of the transcripts were determined by comparison with RNA markers.

Phylogenetic studies. TFB1M and TFB2M homologs were aligned with Basic GeneBee ClustalW 1.75 sequence alignment tool (http://www.genebee.msu.su/clustal/). A phylogenetic tree was generated from ClustalW alignment data by using TreeView 1.6.6 (http://taxonomy.zoology.gla.ac.uk/rod/treeview.html).

Results

Tfb1m and Tfb2m map loci and gene structures are conserved in evolution. The mouse *Tfb1m* and *Tfb2m* genes were mapped by using The Jackson Laboratory BSB [(C57BL/6J × *Mus spretus*) × C57BL/6J] backcross panel (Fig. 1). The mapping results are based on linkage analysis of 94 BSB backcross genotypes. This analysis placed the *Tfb1m* gene locus to the most proximal (centromeric) part of Chr 17 with a distance of 2.13 cM to the closest distal markers *D17Mit113*, *D17Mit156*, and *D17Mit246*. The *Tfb1m* locus is in a region of conserved linkage synteny with human Chr 6, where *TFB1M* maps. The *Tfb2m* locus co-segregated with markers *D1Mit315*, *D1Mit273*, and *D1Mit509* located to the distal end of Chr 1. The distance to the closest proximal (*D1Mit404*) or distal (*Tlr5*, *D1Mit37*, *D1Mit360*, and *D1Mit458*) markers was 1.06 cM. The *Tfb2m* locus is in a region of conserved linkage synteny with human Chr 1 where *TFB2M* maps.

We obtained genomic sequences for *Tfb1m* and *Tfb2m* genes by BLAST searches against the Celera mouse genome database. Sequence homology to the *Tfb1m* cDNA nucleotide sequence was found on Chr 17, consistent with the linkage analysis data. Alignment of the genomic sequence with the cDNA sequence revealed that the *Tfb1m* gene spans 38.5 kb and consists of seven exons. BLAST search with the *Tfb2m* cDNA nucleotide sequence identified two loci on different mouse chromosomes. The *Tfb2m* gene locus on Chr 1 spans 18.2 kb and is divided into eight exons, while the locus on Chr 12 spans 1.9 kb, contains no introns, and encodes an open reading

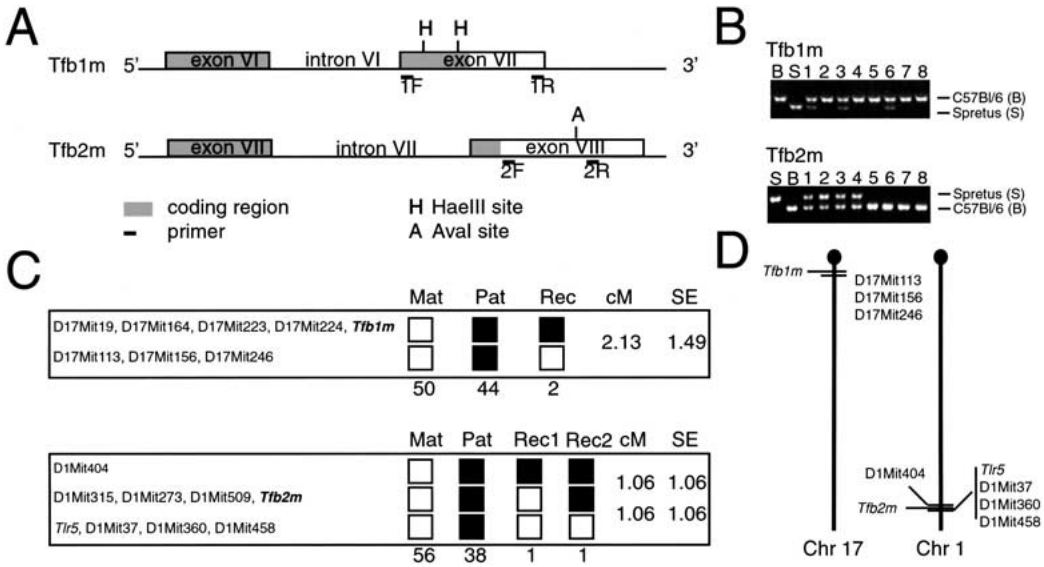

Fig. 1. Mapping of the mouse *Tfb1m* and *Tfb2m* genes. Genes encoding mouse TFB1M and TFB2M proteins were mapped by using The Jackson Laboratory BSB [(C57BL/6J × *Mus spretus*) × C57BL/6J] back cross panel services. **A**. Two PCR assays were designed to distinguish between *M. spretus* (S) and C57BL/6J (B) alleles. In the case of *Tfb1m*, only PCR products from *M. spretus* template contained the two *Hae*III sites shown. For *Tfb2m* only PCR products from C57BL/6J template contained the *Ava*I site shown on the figure. **B**. The PCR products from the two parental (B and S) and 94 backcross DNA templates were digested with the appropriate enzymes (*Hae*III or *Ava*I), and the restriction fragments were separated in gel electrophoresis. Results are shown for parental alleles (B and S) and 8 backcross samples. **C**. The loci for *Tfb1m* and *Tfb2m* were computed based on the previously mapped loci (n > 920) on the BSB panel map. *Tfb1m* co-segregated with markers (*D17Mit19*, *D17Mit164*, *D17Mit223*, and *D17Mit224*) located to the most proximal part of Chr 17. The distance to the closest more distal markers (*D17Mit113*, *D17Mit156*, and *D17Mit246*) was 2.13 cM. *Tfb2m* co-segregated with markers (*D1Mit315*, *D1Mit273*, and *D1Mit509*) located to the distal end of Chr 1. The distance to the closest proximal (*D1Mit404*) or distal (*Tlr5*, *D1Mit37*, *D1Mit360*, and *D1Mit458*) markers was 1.06 cM. **D**. A schematic view of the loci for *Tfb1m* and *Tfb2m* and the closest markers on the Jackson BSB backcross panel Chrs 17 and 1.

frame (ORF) of 347 aa. The ORF on Chr 12 lacks coding sequence corresponding to the 49 most amino terminal amino acids of the ORF encoded by the *Tfb2m* cDNA. The absence of introns and the lack of a methionine codon at the beginning of the truncated ORF suggest that the locus on Chr 12 is a processed pseudogene lacking potential to encode a functional TFB2M protein. Similarly, a human processed pseudogene for *TFB2M* exists on human Chr 6.

Tfb1m and Tfb2m are ubiquitously expressed. We probed multi-tissue Northern blots with *Tfb1m* and *Tfb2m* cDNAs (Fig. 2) and found ubiquitous expression of both genes. The *Tfb1m* transcripts had the highest relative levels in heart, liver, and testis, while *Tfb2m* transcript had the highest relative levels in heart, liver, kidney, skeletal muscle, and brain. We only observed single messages of ~1.4 kb for *Tfb1m* and ~2.4 kb for *Tfb2m*. The transcript sizes corresponded well to the length of known cDNAs for *Tfb1m* and *Tfb2m*. The absence of any additional transcripts detectable by *Tfb2m* probe makes it likely that the locus on Chr 12 contains a non-transcribed pseudogene. Next, we analyzed the proximal promoter regions for *Tfb1m*, *Tfb2m*, *TFB1M*, and *TFB2M* with the MatInspector program. Both the mouse and the human promoters contain multiple conserved sequence elements with predicted binding sites for the transcription factors Sp1

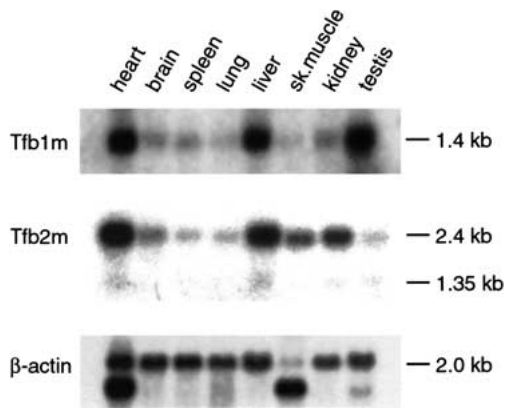

Fig. 2. The expression pattern of *Tfb1m* and *Tfb2m* mRNAs. Multi-tissue Northern blot shows the ubiquitous expression of *Tfb1m* and *Tfb2m* genes. The *Tfb1m* cDNA probe detected an mRNA of ~1.4 kb, and the *Tfb2m* cDNA probe detected a single mRNA of ~2.4 kb. β-actin is a loading control.

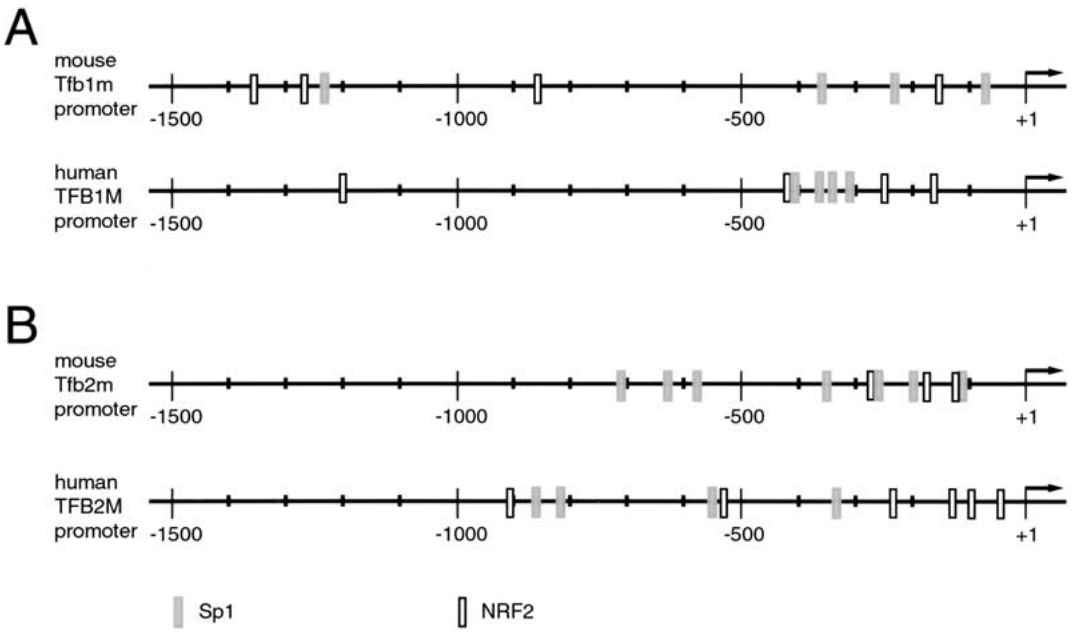

Fig. 3. A schematic view over the proximal promoters. Distribution of putative binding sites for the transcription factors NRF2 (white boxes) or Sp1 (gray boxes) in the promoter regions of *Tfb1m*/TFB1M (**A**) and *Tfb2m*/TFB2M (**B**).

and NRF2 (Fig. 3). No good consensus binding sites for NRF1 were found in any of the promoters.

TFB1M and TFB2M homologs exist in most metazoans. We performed BLAST searches with peptide sequences to find TFB1M and TFB2M homologs in yeast and animals. We found homologs to both proteins in *Drosophila melanogaster*, whereas there was only a single TFBM homolog in *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and *Caenorhabditis elegans*. We studied the relationship between these TFBM homologs by generating a phylogenetic tree based on a multi sequence alignment with the ClustalW program. These different factors form three clusters in an unrooted tree (Fig. 4), where one branch includes the TFB1M peptides from human, mouse and *D. melanogaster* and the single *C. elegans* TFBM. The second branch contains the human, mouse, and *D. melanogaster* TFB2M peptide sequences, and the third branch contains the single TFBM homologs of *S. cerevisiae* and *S. pombe*.

Discussion

We have recently identified two novel human mitochondrial transcription factors, TFB1M and TFB2M, that interact directly with mitochondrial RNA polymerase and activate mitochondrial transcription *in vitro* (Falkenberg et al. 2002). In this paper, we have further characterized *Tfb1m* and *Tfb2m*, encoding putative mouse homologs of *TFB1M* and *TFB2M*. We used linkage analysis and mapped *Tfb1m* and *Tfb2m* to mouse Chr 17 and Chr 1, respectively, in regions of conserved linkage synteny with human Chr 6 and Chr 1, where *TFB1M* and *TFB2M* map. The intron-exon arrangement between *Tfb1m and TFB1M* and between *Tfb2m* and *TFB2M* was identical, and the promoter regions had similar predicted binding sites for the transcription factors NRF2 and Sp1. Furthermore, Northern blot analyses showed that *Tfb1m* and *Tfb2m* were ubiquitously expressed and had expression patterns that were very similar to the previously reported expression patterns for *TFB1M* and *TFB2M* (Falkenberg et al. 2002). The amino acid sequence identity is greater for *Tfb1m* and *TFB1M* (86% identity in the first 322 aa) or *Tfb2m* and *TFB2M* (53% identity) than for *Tfb1m* and *Tfb2m* (21% identity) or *TFB1M* and *TFB2M* (25% identity) (Falkenberg et al. 2002). Taken together, these data strongly suggest that *Tfb1m* and *Tfb2m* are indeed orthologous genes to *TFB1M and TFB2M*.

The presence of two different factors that interact with POLRMT in mammals is unexpected and unexplained. The regulation of mitochondrial gene expression in response to the metabolic requirements in mammals is not well understood. TFAM is likely to have a key role in this process, because it can differentially regulate the relative levels of mtDNA transcription from LSP and HSP (Falkenberg et al. 2002). Mitochondrial transcription termination events may also be of importance in regulating the

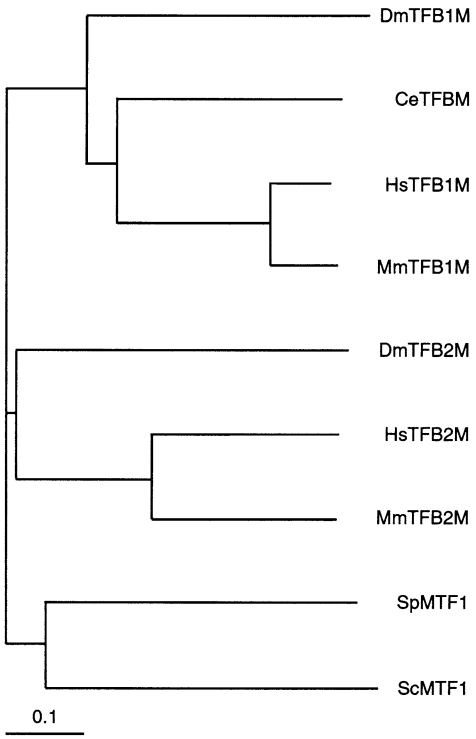

Fig. 4. TFB1M and TFB2M homologs are present in most metazoans. A phylogenetic tree based on a ClustalW alignment and generated with the program Tree View. The protein accession numbers associated with each factor are HsTFB1M, NP_057104; HsTFB2M, NP_071761; MmTFB1M, AF508971; MmTFB2M, NP_032275; CeTFBM, T29195; DmTFB1M, AAF50022; DmTFB2M, AAF54482. The scale below the tree indicates an amino acid replacement distance of 0.1.

transcription of protein encoding genes on the heavy strand of mtDNA (Kruse et al. 1989; Shang and Clayton 1994). TFB1M and TFB2M could constitute yet another level of regulation of mtDNA gene expression in mammalian cells (Falkenberg et al. 2002). *In vitro* transcription experiments have demonstrated that TFB2M is at least ten times more effective than TFB1M in supporting transcription initiation from mitochondrial promoters. The preferential use of TFB1M or TFB2M can, therefore, directly influence overall mitochondrial transcription initiation.

There was an interesting discrepancy between the expression of *Tfb1m* and *Tfb2m* in mouse testis; the levels of *Tfb1m* transcripts were high, whereas the levels of *Tfb2m* transcripts were low. Consistent with this observation, we have previously reported much higher levels of *TFB1M* than *TFB2M* transcripts in human testis (Falkenberg et al. 2002). The mtDNA copy number is downregulated during mammalian spermatogenesis, while the mitochondrial transcript levels remain unchanged (Rantanen and Larsson 2000). This process, dissociating mtDNA replication from mtDNA transcription in testis, may be of importance to avoid paternal transmission of mtDNA (Rantanen and Larsson 2000). We have previously identified specific *TFAM/Tfam* transcript isoforms in human, mouse, and rat testis. These findings suggest that there may be specific mechanisms that regulate *TFAM/Tfam*, *TFB1M/Tfb1m* and *TFB2M/Tfb2m* expression, which in turn regulate mtDNA replication and transcription, in the mammalian testis (Larsson et al. 1996, 1997; Rantanen et al. 2001; Rantanen and Larsson 2000).

Homology searches of protein databases allowed us to identify putative homologs for both *TFB1M* and *TFB2M* in *D. melanogaster*. However, we found only a single homolog in two different yeast species and in the primitive metazoan *C. elegans*. These findings show that duplicated *TFBM* genes are present from early metazoan evolution. Our data thus suggest that most animals will have genes orthologous to both *TFB1M* and *TFB2M*. Computer-assisted protein function predictions indicate that the TFB1M proteins are RNA diadenosine methyltransferases (Falkenberg et al. 2002). The TFB2M proteins are not predicted to be RNA methyltransferases, and their homology to known bacterial transferases is significantly lower (Falkenberg et al. 2002). It is thus possible that an enzyme originally devoted to RNA modification has been recruited during evolution to serve as a mitochondrial transcription factor and that a subsequent gene duplication event in early metazoans generated the *TFB1M* and *TFB2M* genes. TFB2M may have lost its ability to methylate RNA during metazoan evolution and instead became a more efficient transcription factor.

Acknowledgments

M. Falkenberg is supported by a postdoctoral fellowship from the Karolinska Institutet. N.-G. Larsson is supported by grants from the Swedish Research Council, Funds of Karolinska Institutet, Torsten and Ragnar Söderbergs stiftelse, the Swedish Heart and Lung Foundation, and the Swedish Foundation for Strategic Research (Functional Genomics and INGVAR). C.M. Gustafsson is supported by grants from the Swedish Cancer Society, the Swedish Research Council, Human Frontiers Science Program, the Swedish Foundation for Strategic Research (INGVAR), the Swedish Society for Medical Research, the Åke Wiberg Foundation, and the Magn. Bergwall Foundation.

References

1. Anderson S, Bankier AT, Barell BG, deBruijn MHL, Coulson AR et al. (1981) Sequence and organization of the human mitochondrial genome. Nature 290, 457–465
2. Bibb MJ, VanEtten RA, Wright CT, Walberg MW, Clayton DA (1981) Sequence and organization of mouse mitochondrial DNA. Cell 26, 167–180
3. Falkenberg M, Gaspari M, Rantanen A, Trifunvic A, Larsson NG, et al. (2002) Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA. Nat Genet 31, 289–294
4. 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
5. Larsson NG, Clayton D (1995) Molecular genetic aspects of human mitochondrial disorders. Annu Rev Genet 29, 151–178
6. Larsson NG, Luft R (1999) Revolution in mitochondrial medicine. FEBS Lett 455, 199–202
7. Larsson NG, Garman JD, Oldfors A, Barsh GS, Clayton DA (1996) A single mouse gene encodes the mitochondrial transcription factor A and a testis-specific nuclear HMG-box protein. Nat Genet 13, 296–302
8. Larsson NG, Oldfors A, Garman JD, Barsh GS, Clayton DA (1997) Down-regulation of mitochondrial transcription factor A during spermatogenesis in humans. Hum Mol Genet 6, 185–191
9. Lightowlers RN, Chinnery PF, Turnbull DM, Howell N (1997) Mammalian mitochondrial genetics: heredity, heteroplasmy and disease. Trends Genet 13, 450–455
10. Masters BS, Stohl LL, Clayton DA (1987) Yeast mitochondrial RNA polymerase is homologous to those encoded by bacteriophages T3 and T7. Cell 51, 89–99
11. McCulloch V, Seidel-Rogol BL, Shadel GS (2002) A human mitochondrial transcription factor is related to RNA adenine methyltransferases and binds S-adenosylmethionine. Mol Cell Biol 22, 1116–1125
12. Parisi MA, Clayton DA (1991) Similarity of human mitochondrial transcription factor 1 to high mobility group proteins. Science 252, 965–969
13. Parisi MA, Xu B, Clayton DA (1993) A human mitochondrial transcriptional activator can functionally replace a yeast mitochondrial HMG-box protein both in vivo and in vitro. Mol Cell Biol 13, 1951–1961
14. Rantanen A, Larsson NG (2000) Regulation of mitochondrial DNA copy number during spermatogenesis. Hum Reprod 15 (Suppl. 2), 86–91
15. Rantanen A, Jansson M, Oldfors A, Larsson NG (2001) Downregulation of Tfam and mtDNA copy number during mammalian spermatogenesis. Mamm Genome 12, 787–792
16. Schinkel AH, Koerkamp MJAG, Touw EPW, Tabak HF (1987) Specificity factor of yeast mitochondrial RNA polymerase. Purification and interaction with core RNA polymerase. J Biol Chem 262, 12785–12791
17. Shang J, Clayton DA (1994) Human mitochondrial transcription termination exhibits RNA polymerase independence and biased bipolarity in vitro. J Biol Chem 269, 29112–29120
18. Tiranti V, Savoia A, Forti F, D'Apolito MF, Centra M et al. (1997) Identification of the gene encoding the human mitochondrial RNA polymerase (h-mtRPOL) by cyberscreening of the Expressed Sequence Tags database. Hum Mol Genet 6, 615–625
19. Wallace DC (1999) Mitochondrial diseases in man and mouse. Science 283, 1482–1488