Transcription profiling suggests that mitochondrial topoisomerase IB acts as a topological barrier and regulator of mitochondrial DNA transcription

Ilaria Dalla Rosa[1#], Hongliang Zhang[1#], Salim Khiati[1], Xiaolin Wu[2], and Yves Pommier[*1]

[1] Laboratory of Molecular Pharmacology, Developmental Therapeutics Branch, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA
[2] Laboratory of Molecular Technology, Frederick National Laboratory for Cancer Research, Frederick, MD 21702, USA
* To whom correspondence should be addressed. Tel.: 240-760-6142; Fax: 240-541-4475; E-mail: pommier@nih.gov.
[#]Contributed equally to this work.

Running title: Transcription profiling reveals TOP1MT function in mtDNA

ABSTRACT

Mitochondrial DNA (mtDNA) is essential for cell viability because it encodes subunits of the respiratory chain complexes. Mitochondrial topoisomerase IB (TOP1MT) facilitates mtDNA replication by removing DNA topological tensions produced during mtDNA transcription, but appears to be dispensable. To test whether cells lacking TOP1MT have aberrant mtDNA transcription, we performed mitochondrial transcriptome profiling. To that end, we designed and implemented a customized tiling array, which enabled genome-wide, strand-specific, and simultaneous detection of all mitochondrial transcripts. Our technique revealed that *TOP1MT* KO mouse cells have a normal processing of the mitochondrial transcripts, but that protein-coding mitochondrial transcripts are elevated. Moreover, we found discrete long noncoding RNAs produced by H-strand transcription and encompassing the noncoding regulatory region of mtDNA in human and murine cells and tissues. Of note, these noncoding RNAs were strongly upregulated in the absence of TOP1MT. In contrast, 7S DNA, produced by mtDNA replication, was reduced in the *TOP1MT* KO cells. We propose that the long noncoding RNA species in the D-loop region are generated by the extension of H-strand transcripts beyond their canonical stop site and that TOP1MT acts as a topological barrier and regulator for mtDNA transcription and D-loop formation.

INTRODUCTION

Mitochondrial DNA (mtDNA) is essential because it encodes subunits of the respiratory chain complexes (1). Each mitochondrion contains multiple copies of a 16 kb circular genome (2) and both mtDNA strands, denoted heavy (H) and light (L), are transcribed. In the mitochondrial genome the only non-coding region (NCR) of a large size is confined to a DNA track of approximately 700 bp in mice and 1100 bp in humans, containing the promoters for mtDNA transcription and the origin of H-strand replication (Fig. 1). In addition to the 13 mRNAs coding for the respiratory chain subunits, mtDNA encodes 2 rRNAs and 22 tRNAs. Ribosomal RNAs associate with nuclear-encoded ribosomal proteins to form the mitochondrial ribosomes and transfer RNAs are used for mitochondrial protein synthesis.

mtDNA transcription starts from 3 different promoters, two on the H-strand (HSP1 and HSP2) and one on the L-strand (LSP) (3) (Fig. 1A). Transcription starting from HSP2 and LSP produces genome-long polycistronic transcripts, whereas transcription from HSP1 stops prematurely, generating high levels of the two ribosomal RNAs (Fig. 1B). The polycistronic transcripts are processed by the excision of tRNAs, which releases individual mRNAs and rRNAs (4) (Fig. 1C). The tRNAs are further processed by the addition of a CCA triplet at their 3'-end (5), whereas rRNAs and mRNAs mature by the addition of 50-60 nt long poly(A) tails (6) (Fig. 1C). Poly(A) tails are known to increase mRNA stability for nuclear eukaryotic genes, whereas the same post transcriptional modification promotes RNA degradation in eubacteria (7). In mitochondria, less is known about the role of polyadenylation (8) and poly(A) tails can have different effects on different transcripts, increasing the stability

of some mtRNAs and promoting degradation of others (9).

Bidirectional transcription and replication of mtDNA generate DNA supercoiling (topological stress) in the circular mitochondrial genome, which is anchored to the mitochondrial inner membrane in nucleoids (10-12). The canonical function of topoisomerases is to release helical tensions and decatenate newly replicated DNA molecules (13-17). Four topoisomerases have been identified in mammalian mitochondria: mitochondrial topoisomerase I (TOP1MT) (18), topoisomerase III-alpha (TOP3A) (19) and more recently topoisomerases II alpha and beta (TOP2A and TOP2B) (20,21). TOP1MT bears a strong mitochondrial targeting sequence (18) and is the only topoisomerase exclusively targeted to mitochondria (17,22). Its catalytic core domain is highly similar to the nuclear topoisomerase IB (TOP1) and it is likely that the two enzymes originated from duplication of a common ancestral gene during vertebrate evolution (22,23).

TOP1MT has recently been shown to facilitate mtDNA replication during liver regeneration and following ethidium bromide exposure (24). Yet, TOP1MT is not essential for mouse development, probably due to the presence of TOP3A, TOP2A and TOP2B in mitochondria, complementing its function (21). TOP1MT associates with transcriptionally active nucleoids and directly interacts with the mitochondrial RNA polymerase (POLRMT) (25). In addition, ChIP-on-Chip experiments showed that TOP1MT preferentially binds to mtDNA promoters, suggesting that the enzyme acts as cofactor for mtDNA transcription. *TOP1MT* KO mouse embryonic fibroblasts (MEFs) show mitochondrial dysfunctions characterized by reduced ATP levels and increased reactive oxygen species (ROS) production (26). Steady-state levels of mitochondrial transcripts (25,26) and negative supercoiling of mtDNA are also increased in *TOP1MT* deficient cells (21), indicating a role of TOP1MT in removing DNA topological tensions produced by mtDNA transcription.

Here we set out to characterize the mtDNA transcription profile in human and mouse cells and tissues in the absence of TOP1MT. Using customized, strand-specific tiling arrays, we provide a method to quantify coding and non-coding mitochondrial transcripts on both mtDNA strands, simultaneously. This approach offers new insights into the role of TOP1MT as a regulator for mtDNA transcription and replication fork.

RESULTS

Tiling array design for the study of mitochondrial transcripts

To perform a comprehensive characterization of the mitochondrial transcriptome in the absence of TOP1MT, a non-essential component of the mitochondrial transcription machinery (25), we designed two customized tiling arrays for the mitochondrial genome, one specific for mouse and another for human. Individual probes on the arrays consisted of 60 bases DNA oligonucleotides matching mtDNA sequences and overlapping adjacent probes by 55 nucleotides (Fig. 2A, sequences are listed in Supplementary Table 1). RNA samples were reverse transcribed and hybridized to the array in order to generate strand specific mtDNA transcription profiles. Because our main interest was to analyze mtDNA transcripts (i.e. mtRNAs processing and stability), RNA samples were reverse transcribed using oligo-dT primers and mature (i.e. polyadenylated) mtRNAs profiles were analyzed.

Figure 2B shows the relative abundance and expression profiles for polyadenylated mtRNAs from WT MEFs. Hybridization of mature RNAs showed a defined pattern characterized by high signals in the coding regions separated by steep drops (valleys) at their junctions corresponding to the tRNAs processing sites (Fig. 2B and C). Such profile is consistent with the mechanism of mtDNA transcripts maturation, with mRNAs processing at tRNAs junctions and successive polyadenylation of their 3'-ends (see Fig. 1). The abundance of mature H-strand transcripts was high, consistent with the fact that 12 of the 13 mitochondrial mRNAs and the two mitochondrial rRNAs are transcribed from the H-strand. By contrast, transcripts generated from the L-strand promoter showed much lower signals, consistent with the fact that L-strand transcription codes only for one gene (*ND6*). Comparison of RNA samples isolated from whole cells (W) or isolated mitochondria (M) showed nearly identical profiles (Fig. 2B);

moreover, the ORFs patterns matched the ones generated by standard RNA-seq (Fig. 3), demonstrating the high specificity of the array for mitochondrial transcripts and its effectiveness using total cellular RNA.

Having established the effectiveness of our tiling array, we explored the mitochondrial transcriptomes of murine and human tissues in parallel. Figure 4 shows that, in all the tissues analyzed, the H-strand transcripts showed consistent and well-defined profiles characterized by high signal blocks within the 12 ORFs separated by deep drops at their junctions corresponding to the tRNA processing sites. The higher levels toward the 3'-end of each transcript is likely caused by the reverse transcription reaction, which produces declining signals as the extension was further away from the poly(A) tails. Mitochondrial rRNAs (12S and 16S) also gave high reads, consistent with the fact that both ribosomal and messenger RNAs are processed by polyadenylation (1,4).

Signals for the L-strand were uniformly weaker than for the H-strand, reflecting that the L-strand only encodes the *ND6* transcript. Unlike the well-defined transcript of the H-stand, the *ND6* transcript was not clearly delineated and appeared to include the *ND5* region, which is consistent with the lack of a canonical polyadenylation tail at the *ND6* 3'-end (28). Detailed examination of the L-strand transcriptome also revealed the existence of mature non-coding RNAs (ncRNAs), which were complementary to the H-strand ORFs across different tissues (see for example CYTB, COX2 and COX3 regions). These results confirm with a different and independent method the presence of antisense non-coding mitochondrial transcripts (29).

Defined signals corresponding to the Non Coding Region (NCR) were also detected on both strands, indicating the presence of non-coding RNAs emanating from both strands of the regulatory NCR. This result reinforces independent observations suggesting the presence of polyadenylation sites within the NCR (28,30). The signals in the NCR were clearly reproducible across tissues, and were detectable both in murine and human samples.

Transcription profile of *TOP1MT* knockout cells reveals increased non-coding RNAs in the mtDNA non-coding regulatory region

Based on the association of TOP1MT with the mitochondrial transcription machinery including POLRMT (25), and the fact that inactivation of nuclear TOP1 selectively impacts the expression of nuclear genes (31-33), we compared the mitochondrial transcriptome profiles of murine embryo fibroblasts (MEFs) from KO and WT mice (21,26). Mitochondrial transcription patterns were globally similar (Fig. 5A). Yet, quantification of the signal intensity within different ORFs showed that mitochondrial transcripts were approximately 1.5-2-fold more abundant in *TOP1MT* KO compared to WT MEFs (Fig. 5B). This quantitative difference is consistent with qRT-PCR results (25,26), while demonstrating that TOP1MT does not affect the processing of mitochondrial transcripts.

In addition, comparison between the transcript profiles of WT and *TOP1MT* KO MEFs showed selective upregulation of non-coding RNAs spanning the region of the 7S DNA and corresponding to the end of the H-strand transcription (Fig. 5C). The levels of these polyadenylated RNAs were about 5-fold higher in the *TOP1MT* KO compared to WT MEFs (Fig. 5D). To validate this finding, we performed RT-PCR experiments and quantified RNA species by semi-quantitative and quantitative RT-PCR (Fig. 5E and F, respectively). Both methods confirmed the increase of the non-coding RNA in *TOP1MT* KO MEFs, and the magnitude of this increase matched the 5-fold changes determined by the tiling array.

To support the causal link between TOP1MT and the presence of non-coding RNA transcripts corresponding to the end of the H-strand transcripts in the regulatory region of the mitochondrial genome, we performed *TOP1MT* knockdown experiments in human colon carcinoma HCT116 cells (Fig. 6A and B). Consistent with the results in the murine *TOP1MT* KO cells, we found that acute *TOP1MT* silencing strongly increased the levels of the H-strand ncRNA spanning the D-loop region.

The increase of the ncRNA species produced by H-strand transcription, occurring in the mtDNA regulatory region in the absence in *TOP1MT*, could be due to enhanced stability or, alternatively, to an increase in transcription rate. In order to understand the basis of this strong upreguation, we treated WT and *TOP1MT* KO

MEFs with ethidium bromide (EtBr). EtBr is a high affinity DNA intercalating dye commonly used to inhibit mtDNA transcription and replication. Due to the relatively long half-life of mtDNA (34), short EtBr treatment effectively blocks mtDNA transcription without affecting mtDNA copy number. EtBr treatment caused a marked decrease of the ncRNA levels of the mtDNA regulatory region (Fig. 6C). This decrease was markedly more pronounced in *TOP1MT* KO MEFs, suggesting a shorter half-life of these ncRNA species. These data suggest that the increase in ncRNA in the mtDNA regulatory region is due to enhanced transcription.

7S DNA alterations in *TOP1MT* KO MEFs

TOP1MT forms high-affinity complexes in the non-coding region of mtDNA (35,36) (Fig.7A) and lack of TOP1MT leads to profound changes in the topological state of the mitochondrial genome, increasing negatively supercoiled mtDNA molecules (19). The non-coding RNA, which we found strongly enhanced in the absence of TOP1MT, is produced by H-strand transcription and we hypothesize that the topological changes occurring in the NCR, upon TOP1MT loss, facilitate the progression of H-strand transcription into this region, generating high levels of ncRNAs. We hypothesized that increased-H strand transcription into the NCR could interfere with the production of 7S DNA, which is synthetized by DNA polymerase gamma in the opposite direction. Indeed, we found that 7S DNA levels were decreased in *TOP1MT* KO MEFs (Fig. 7B). These results suggest that TOP1MT might play a previously unknown role in the mtDNA NCR, regulating the termination of transcription and facilitating replication (21,24,37).

DISCUSSION

To investigate the impact of *TOP1MT* depletion on the mitochondrial transcriptome, we designed customized tiling arrays allowing quantitative and qualitative analysis of all mitochondrial transcripts simultaneously. Although a comprehensive map of the mitochondrial transcriptome has been recently generated using RNA-seq (30), RNA-seq requires specialized bioinformatics support to analyze the results. The advantage of the tiling array technique is its simplicity and strand specificity, even when using total RNAs as starting material (see Fig. 2). Comparison between mitochondrial transcription profiles generated using the tiling array or the RNA-Seq approach showed similar patterns (see Fig. 3). The mtDNA open reading frames were clearly detectable with both techniques and their profiles were comparable, although the RNA-Seq data reflected the combined transcript for both mtDNA strands. The ncRNAs spanning the mtDNA non-coding region were also readily detected in high abundance by RNA-Seq. Moreover, the relative abundance of the different coding and non-coding transcripts was comparable across the two different techniques. Thus, our tiling array method enables the study of mtDNA transcription features in different experimental conditions without purifying mitochondria.

RNA processing and maturation represent key features of gene expression regulation. In mitochondria newly synthetized transcripts, produced as long polycistronic units, are polyadenylated at the 3'-end by the mitochondrial poly(A) polymerase (mtPAP) (38). Contrary to the nuclear compartment where polyadenylation stabilizes mRNAs, the function of poly(A) tails in mitochondria appears different (7), as polyadenylation has opposite effects on the stability of different mitochondrial transcripts (9).

Here we set out to study mtDNA gene expression, profiling all the polyadenylated (i.e. mature) mtRNAs in mouse and human samples under physiological conditions, in the presence of EtBr, and in *TOP1MT* deficient cells. Expression profiles obtained for both mouse and human genomes were similar and consistent with the tRNA punctuation model for mtRNAs processing (4). On the H-strand, regions corresponding to ribosomal and messenger RNAs gave high reads, exactly matching the 5' and 3' ends of each transcript. Products of the L-strand transcription showed much lower signals, which is consistent with the fact that the L-strand contains only one ORF for ND6. Notably, *ND6* transcripts were barely detectable both in human and mouse samples (Fig. 3). This result confirms previous data showing that the *ND6* transcript is polyadenylated ~500 nucleotides downstream of its stop codon (28). Accordingly, we picked up a clear signal in the region complementary to *ND5*, which could represent the 3' end of the processed *ND6* transcript.

Low abundance polyadenylated L-strand transcripts was clearly detectable outside of the ND6 region both in murine and human tissues (Figs. 4-5). These non-coding transcripts (ncRNAs) tended to show the same valleys (tRNA punctuation) as those from the H-strand, and matched the antisense sequence of many H-strand mRNAs (see Figs 3-5, the *CYTB*, *ND5, ND4, COX3, COX2* and *ND1* regions). The existence of three long non-coding mitochondrial RNAs complementary to *ND5*, *ND6* and *CYTB* has been described by Rackham et al. (39). Here we confirm the presence of these three antisense RNAs using an independent method and, additionally, we show that their abundance is variable, generally being fairly high in human solid tissues and lower in MEFs (Figs. 3 and 4). Similarly to nuclear antisense RNAs (40), these mitochondrial RNA species may be involved in the regulation of mtDNA gene expression and their production could depend on different metabolic requirements in different cellular types (29).

Beside the antisense ncRNAs, we report the existence of ncRNAs mapping to the mitochondrial regulatory region (NCR). Non-coding RNAs complementary to 7S DNA have been recently reported in the mitochondrial genome (16). These RNAs form a R-loop on the mtDNA regulatory region and play a role in mtDNA organization and mitochondrial dysfunction.

The functions of TOP1MT have remained relatively unexplored since the discovery of the gene 15 years ago (18,22). In contrast to other nuclear-encoded genes that sustain mtDNA maintenance, *TOP1MT* is not essential (21,24,26,37), and is restricted to vertebrates (22). Two recent studies have shown that TOP1MT is required for mtDNA copy number expansion during liver regeneration (24) and myocardial adaptation to mitochondrial poisoning by doxorubicin (37). TOP1MT also tends to repress mitochondrial transcription (25). Consistently, our genome-wide analyses show increased steady-state levels of most mitochondrial mRNAs in the absence of TOP1MT, a result that confirms previous qRT-PCR data (26). Yet, the overall transcriptome profiles were remarkably similar, indicating that TOP1MT is not critical for mtDNA processing. Interestingly, the most altered transcripts in *TOP1MT*-deficient MEFs were the non-coding RNAs produced by the H-strand transcription and spanning the regulatory region of the D-loop. A non-coding RNA encoded by this region has been reported by other groups using different techniques and model systems (28,30,41-43), but its function and regulation remains unclear. We found that, in cells lacking *TOP1MT*, such non-coding RNA complementary to the 7S DNA was 5-fold enriched, representing a notable perturbation of the entire mtDNA expression profile (Figs. 5-7). Ethidium bromide (EtBr) treatment strongly downregulated this RNA in the *TOP1MT* KO MEFs, whereas it had only a partial effect in the WT MEFs (Fig. 6). This result indicates that the upregulation of this RNA specie in *TOP1MT* KO MEFs is not due to an increase in its stability.

To explain our finding that TOP1MT attenuates the expression of the non-coding RNA mapping the NCR, we propose that TOP1MT, tuning mtDNA topology, limits H-strand transcription into the mtDNA regulatory region (see model in Fig. 7C). Accordingly, TOP1MT cleavage sites accumulate in the mtDNA D-loop region (35) (Fig. 7A). Moreover, the mtDNA of *TOP1MT* KO MEFs is highly negatively supercoiled (i.e. underwound) (21), and this could enable transcription progression behind its usual termination point. In line with this, we found decreased 7S DNA (produced in the opposite direction) in *TOP1MT* KO MEFs. Hence, we surmise that TOP1MT acts as a regulator of mtDNA topology in NCR, facilitating 7S formation and the termination of H-strand transcription at the termination-associated sequence (TAS) (44).

These data suggest a novel regulatory function for TOP1MT as an enzyme able to modulate mtDNA transcription through modification of mtDNA topology.

EXPERIMENTAL PROCEDURES

Cell-culture and *TOP1MT* Knockdown

Human HCT116 cells, WT and *TOP1MT* KO MEFs (26) were cultured in DMEM medium supplemented with 10% FCS. To inhibit mitochondrial transcription, WT and TOP1mt KO MEFs were treated with 50 ng/mL ethidium bromide for 24 hours prior to RNA extraction. For *TOP1MT* knockdown, HCT116 cells were transfected with 5 nM TOP1MT siRNA (ON-TARGET plus SMART pool, Dharmacon) or scrambled sequences using Lipofectamine

RNAiMAX (Invitrogen), according to manufacturer's specifications. Knockdown efficiency was tested by Western blot analysis 72h after transfection.

RNA isolation and mitochondrial DNA tiling array

Total RNA was isolated from cells ($2x10^6$) using RNeasy Mini Kit (QIAGEN), according to the manufacturer's instructions. For RNA isolation from mice tissues, brain and testes were homogenized in 3 ml Trizol using an ultra Turrax homogenizer. After the addition of 600 µl chloroform, samples were thoroughly vortexed and spun at 12000 xg for 15 minutes at 4 degrees. The aqueous phase was then precipitated with equal volumes of RNAse free 70% ethanol and loaded onto the RNeasy Mini Kit columns (QIAGEN). Washing and RNA elution were performed according to the manufacturer's instructions. For all extractions DNase digestion was performed on the columns. Human total RNA was purchased from Agilent Technologies.

Arrays were manufactured by Agilent Technologies (Santa Clara, CA) using Agilent SurePrint 4x44K format. Tiling used 60-mer probes with 5bp stagger for both mtDNA strands (H and L) to detect strand-specific expression within the same array. The arrays also included probes for nuclear genes that served as loading controls. Tiling arrays were customized for the human and murine mitochondrial genomes.

Total RNA isolated from cultured cells or tissues was labeled with the Quick Amp Labeling kit (Agilent, 5190-0442) following the manufacturer recommended protocols. Briefly, cRNA was generated from 500 ng RNA with T7 promoter tagged oligo-dT and MMLV reverse transcriptase. cRNA was labeled with Cy3 and Cy5 dye using T7 RNA polymerase. Three ug of labeled cRNA was fragmented and hybridized to the array at 65°C for 17 hours. Slides were washed in "Agilent GE" wash buffer 1 for 1 min at room temperature and in wash buffer 2 for 1 min at 37°C. The slides were scanned in an Agilent C-Scanner and the data extracted with Agilent Feature extraction 10.7.

RNA-Seq analysis

RNA was iolated from HCT116 cells as described above. Sequencing libraries were generated using the TotalScript™ RNA-Seq Kit (Epicentre) and sequenced using the HiSeq 2000 (Illumina) with paired end 100bp reads. Data was visualized by aligning to the human genome assembly 19.

Semi-quantitative and quantitative PCR

The relative abundance of the non-coding RNA spanning the D-loop region was assessed using semi-quantitative and quantitative PCR. For semi-quantitative measurement, total RNA was retrotranscribed with oligo-dT or random primers and cDNA served as template for the amplification of the D-loop portion spanning nucleotides 15645-15812. Actin B was amplified as reference. To ensure that the abundance of PCR product and template was proportional, we limited the amplification step to 18 cycles. The primers used are the following: D-loop F-GGAAGGGGATAGTCATATGGAAGAGAA G; D-loop R-TGGCCCTGAAGTAAGAACCAGATGTCTG ATAAAG; ACTB F-CTACGAGGGCTATGCTC; ACTB R-CTTTGATGTCACGCACG. For quantitative PCR, PCR reactions were performed in triplicates on 384-well Reaction plates (Applied Biosystems). Each PCR reaction (final volume 10 µl) contained 25 ng cDNA, 5 µl of Power SYBR-Green PCR Master Mix (Applied Biosystems) and 0.5 µM of each forward and reverse primer. The D-loop region and the normalizing control Actin B were amplified using the primers listed above.

Analysis of 7S DNA

Southern Blot analysis of mtDNA was carried out as described (27). Briefly, 2 µg total DNA was digested with the mtDNA single cutter XhoI and run on a 1% agarose gel. DNA was then transferred by capillary transfer onto Hybond-N+ membrane (GE Healthcare) and hybridized with a ^{32}P-labeled probe for mouse-mtDNA (nts 15007-15805). ^{32}P signals were visualized with PhosphorImager system (Molecular Dynamics).

Acknowledgements: Our studies are supported by the Intramural Program of the National Cancer Institute, Center for Cancer Research (Z01 BC006161), National Institutes of Health.

Conflict of interest: The authors declare no conflicts of interest.

Author contributions: IDR, HZ, XW, and YP designed the experiments and analyzed the data. IDR, HZ, SK and XW performed the experiments. IDR and YP wrote the manuscript.

REFERENCES

1. Anderson, S., Bankier, A. T., Barrell, B. G., de Bruijn, M. H., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J., Staden, R., and Young, I. G. (1981) Sequence and organization of the human mitochondrial genome. *Nature* **290**, 457-465

2. D'Erchia, A. M., Atlante, A., Gadaleta, G., Pavesi, G., Chiara, M., De Virgilio, C., Manzari, C., Mastropasqua, F., Prazzoli, G. M., Picardi, E., Gissi, C., Horner, D., Reyes, A., Sbisa, E., Tullo, A., and Pesole, G. (2015) Tissue-specific mtDNA abundance from exome data and its correlation with mitochondrial transcription, mass and respiratory activity. *Mitochondrion* **20**, 13-21

3. Bonawitz, N. D., Clayton, D. A., and Shadel, G. S. (2006) Initiation and beyond: multiple functions of the human mitochondrial transcription machinery. *Mol Cell* **24**, 813-825

4. Ojala, D., Montoya, J., and Attardi, G. (1981) tRNA punctuation model of RNA processing in human mitochondria. *Nature* **290**, 470-474

5. Levinger, L., Morl, M., and Florentz, C. (2004) Mitochondrial tRNA 3' end metabolism and human disease. *Nucleic Acids Res* **32**, 5430-5441

6. Montoya, J., Ojala, D., and Attardi, G. (1981) Distinctive features of the 5'-terminal sequences of the human mitochondrial mRNAs. *Nature* **290**, 465-470

7. Schuster, G., and Stern, D. (2009) RNA polyadenylation and decay in mitochondria and chloroplasts. *Prog Mol Biol Transl Sci* **85**, 393-422

8. Bobrowicz, A. J., Lightowlers, R. N., and Chrzanowska-Lightowlers, Z. (2008) Polyadenylation and degradation of mRNA in mammalian mitochondria: a missing link? *Biochem Soc Trans* **36**, 517-519

9. Wilson, W. C., Hornig-Do, H. T., Bruni, F., Chang, J. H., Jourdain, A. A., Martinou, J. C., Falkenberg, M., Spahr, H., Larsson, N. G., Lewis, R. J., Hewitt, L., Basle, A., Cross, H. E., Tong, L., Lebel, R. R., Crosby, A. H., Chrzanowska-Lightowlers, Z. M., and Lightowlers, R. N. (2014) A human mitochondrial poly(A) polymerase mutation reveals the complexities of post-transcriptional mitochondrial gene expression. *Hum Mol Genet* **23**, 6345-6355

10. Kukat, C., Davies, K. M., Wurm, C. A., Spahr, H., Bonekamp, N. A., Kuhl, I., Joos, F., Polosa, P. L., Park, C. B., Posse, V., Falkenberg, M., Jakobs, S., Kuhlbrandt, W., and Larsson, N. G. (2015) Cross-strand binding of TFAM to a single mtDNA molecule forms the mitochondrial nucleoid. *Proc Natl Acad Sci U S A* **112**, 11288-11293

11. He, J., Mao, C. C., Reyes, A., Sembongi, H., Di Re, M., Granycome, C., Clippingdale, A. B., Fearnley, I. M., Harbour, M., Robinson, A. J., Reichelt, S., Spelbrink, J. N., Walker, J. E., and Holt, I. J. (2007) The AAA+ protein ATAD3 has displacement loop binding properties and is involved in mitochondrial nucleoid organization. *J Cell Biol* **176**, 141-146

12. Bogenhagen, D. F. (2012) Mitochondrial DNA nucleoid structure. *Biochim Biophys Acta* **1819**, 914-920

13. Champoux, J. J. (2001) DNA topoisomerases: structure, function, and mechanism. *Annu Rev Biochem* **70**, 369-413

14. Wang, J. C. (2009) A journey in the world of DNA rings and beyond. *Annu Rev Biochem* **78**, 31-54

15. Pommier, Y., Leo, E., Zhang, H., and Marchand, C. (2010) DNA topoisomerases and their poisoning by anticancer and antibacterial drugs. *Chem Biol* **17**, 421-433

16. Akman, G., Desai, R., Bailey, L. J., Yasukawa, T., Dalla Rosa, I., Durigon, R., Holmes, J. B., Moss, C. F., Mennuni, M., Houlden, H., Crouch, R. J., Hanna, M. G., Pitceathly, R. D., Spinazzola, A., and Holt, I. J. (2016) Pathological ribonuclease H1 causes R-loop depletion and aberrant DNA segregation in mitochondria. *Proc Natl Acad Sci U S A* **113**, E4276-4285

17. Pommier, Y., Sun, Y., Huang, S. N., and Nitiss, J. L. (2016) Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. *Nat Rev Mol Cell Biol* **17**, 703-721

18. Zhang, H., Barcelo, J. M., Lee, B., Kohlhagen, G., Zimonjic, D. B., Popescu, N. C., and Pommier, Y. (2001) Human mitochondrial topoisomerase I. *Proc Natl Acad Sci U S A* **98**, 10608-10613

19. Wang, Y., Lyu, Y. L., and Wang, J. C. (2002) Dual localization of human DNA topoisomerase IIIalpha to mitochondria and nucleus. *Proc Natl Acad Sci U S A* **99**, 12114-12119

20. Low, R. L., Orton, S., and Friedman, D. B. (2003) A truncated form of DNA topoisomerase IIbeta associates with the mtDNA genome in mammalian mitochondria. *Eur J Biochem* **270**, 4173-4186

21. Zhang, H., Zhang, Y. W., Yasukawa, T., Dalla Rosa, I., Khiati, S., and Pommier, Y. (2014) Increased negative supercoiling of mtDNA in TOP1mt knockout mice and presence of topoisomerases IIalpha and IIbeta in vertebrate mitochondria. *Nucleic Acids Res* **42**, 7259-7267

22. Zhang, H., Meng, L. H., Zimonjic, D. B., Popescu, N. C., and Pommier, Y. (2004) Thirteen-exon-motif signature for vertebrate nuclear and mitochondrial type IB topoisomerases. *Nucleic Acids Res* **32**, 2087-2092

23. Dalla Rosa, I., Goffart, S., Wurm, M., Wiek, C., Essmann, F., Sobek, S., Schroeder, P., Zhang, H., Krutmann, J., Hanenberg, H., Schulze-Osthoff, K., Mielke, C., Pommier, Y., Boege, F., and Christensen, M. O. (2009) Adaptation of topoisomerase I paralogs to nuclear and mitochondrial DNA. *Nucleic Acids Res* **37**, 6414-6428

24. Khiati, S., Baechler, S. A., Factor, V. M., Zhang, H., Huang, N., Dalla Rosa, I., Sourbier, C., Neckers, L. M., Thorgeirsson, S. S., and Pommier, Y. (2015) Lack of TOP1mt impairs liver regeneration. *Proc. Nalt. Acad. Sci. U.S.A.* **In press**

25. Sobek, S., Dalla Rosa, I., Pommier, Y., Bornholz, B., Kalfalah, F., Zhang, H., Wiesner, R. J., von Kleist-Retzow, J. C., Hillebrand, F., Schaal, H., Mielke, C., Christensen, M. O., and Boege, F. (2013) Negative regulation of mitochondrial transcription by mitochondrial topoisomerase I. *Nucleic Acids Res* **41**, 9848-9857

26. Douarre, C., Sourbier, C., Dalla Rosa, I., Brata Das, B., Redon, C. E., Zhang, H., Neckers, L., and Pommier, Y. (2012) Mitochondrial topoisomerase I is critical for mitochondrial integrity and cellular energy metabolism. *PLoS One* **7**, e41094

27. Ruhanen, H., Borrie, S., Szabadkai, G., Tyynismaa, H., Jones, A. W., Kang, D., Taanman, J. W., and Yasukawa, T. (2010) Mitochondrial single-stranded DNA binding protein is required for maintenance of mitochondrial DNA and 7S DNA but is not required for mitochondrial nucleoid organisation. *Biochim Biophys Acta* **1803**, 931-939

28. Slomovic, S., Laufer, D., Geiger, D., and Schuster, G. (2005) Polyadenylation and degradation of human mitochondrial RNA: the prokaryotic past leaves its mark. *Mol Cell Biol* **25**, 6427-6435

29. Dietrich, A., Wallet, C., Iqbal, R. K., Gualberto, J. M., and Lotfi, F. (2015) Organellar non-coding RNAs: emerging regulation mechanisms. *Biochimie* **117**, 48-62

30. Mercer, T. R., Neph, S., Dinger, M. E., Crawford, J., Smith, M. A., Shearwood, A. M., Haugen, E., Bracken, C. P., Rackham, O., Stamatoyannopoulos, J. A., Filipovska, A., and Mattick, J. S. (2011) The human mitochondrial transcriptome. *Cell* **146**, 645-658

31. Miao, Z. H., Player, A., Shankavaram, U., Wang, Y. H., Zimonjic, D. B., Lorenzi, P. L., Liao, Z. Y., Liu, H., Shimura, T., Zhang, H. L., Meng, L. H., Zhang, Y. W., Kawasaki, E. S., Popescu, N. C., Aladjem, M. I., Goldstein, D. J., Weinstein, J. N., and Pommier, Y. (2007) Nonclassic Functions of Human Topoisomerase I: Genome-Wide and Pharmacologic Analyses. *Cancer Res.* **67**, 8752-8761

32. Solier, S., Ryan, M. C., Martin, S. E., Varma, S., Kohn, K. W., Liu, H., Zeeberg, B. R., and Pommier, Y. (2013) Transcription Poisoning by Topoisomerase I Is Controlled by Gene Length, Splice Sites, and miR-142-3p. *Cancer Research* **73**, 4830-4839

33. King, I. F., Yandava, C. N., Mabb, A. M., Hsiao, J. S., Huang, H. S., Pearson, B. L., Calabrese, J. M., Starmer, J., Parker, J. S., Magnuson, T., Chamberlain, S. J., Philpot, B. D., and Zylka, M. J. (2013) Topoisomerases facilitate transcription of long genes linked to autism. *Nature* **501**, 58-62

34. Gross, N. J., Getz, G. S., and Rabinowitz, M. (1969) Apparent turnover of mitochondrial deoxyribonucleic acid and mitochondrial phospholipids in the tissues of the rat. *J Biol Chem* **244**, 1552-1562

35. Dalla Rosa, I., Huang, S. Y., Agama, K., Khiati, S., Zhang, H., and Pommier, Y. (2014) Mapping topoisomerase sites in mitochondrial DNA with a poisonous mitochondrial topoisomerase I (Top1mt). *J Biol Chem* **289**, 18595-18602

36. Zhang, H., and Pommier, Y. (2008) Mitochondrial topoisomerase I sites in the regulatory D-loop region of mitochondrial DNA. *Biochemistry* **47**, 11196-11203

37. Khiati, S., Dalla Rosa, I., Sourbier, C., Ma, X., Rao, V. A., Neckers, L. M., Zhang, H., and Pommier, Y. (2014) Mitochondrial topoisomerase I (top1mt) is a novel limiting factor of Doxorubicin cardiotoxicity. *Clin Cancer Res* **20**, 4873-4881

38. Tomecki, R., Dmochowska, A., Gewartowski, K., Dziembowski, A., and Stepien, P. P. (2004) Identification of a novel human nuclear-encoded mitochondrial poly(A) polymerase. *Nucleic Acids Res* **32**, 6001-6014

39. Rackham, O., Shearwood, A. M., Mercer, T. R., Davies, S. M., Mattick, J. S., and Filipovska, A. (2011) Long noncoding RNAs are generated from the mitochondrial genome and regulated by nuclear-encoded proteins. *RNA* **17**, 2085-2093

40. Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C. C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K. C., Hallinan, J., Mattick, J., Hume, D. A., Lipovich, L., Batalov, S., Engstrom, P. G., Mizuno, Y., Faghihi, M. A., Sandelin, A., Chalk, A. M., Mottagui-Tabar, S., Liang, Z., Lenhard, B., Wahlestedt, C., Group, R. G. E. R., Genome

Science, G., and Consortium, F. (2005) Antisense transcription in the mammalian transcriptome. *Science* **309**, 1564-1566

41. Freyer, C., Park, C. B., Ekstrand, M. I., Shi, Y., Khvorostova, J., Wibom, R., Falkenberg, M., Gustafsson, C. M., and Larsson, N. G. (2010) Maintenance of respiratory chain function in mouse hearts with severely impaired mtDNA transcription. *Nucleic Acids Res* **38**, 6577-6588

42. Sbisa, E., Nardelli, M., Tanzariello, F., Tullo, A., and Saccone, C. (1990) The complete and symmetric transcription of the main non coding region of rat mitochondrial genome: in vivo mapping of heavy and light transcripts. *Curr Genet* **17**, 247-253

43. Selwood, S. P., McGregor, A., Lightowlers, R. N., and Chrzanowska-Lightowlers, Z. M. (2001) Inhibition of mitochondrial protein synthesis promotes autonomous regulation of mtDNA expression and generation of a new mitochondrial RNA species. *FEBS Lett* **494**, 186-191

44. Doda, J. N., Wright, C. T., and Clayton, D. A. (1981) Elongation of displacement-loop strands in human and mouse mitochondrial DNA is arrested near specific template sequences. *Proc Natl Acad Sci U S A* **78**, 6116-6120

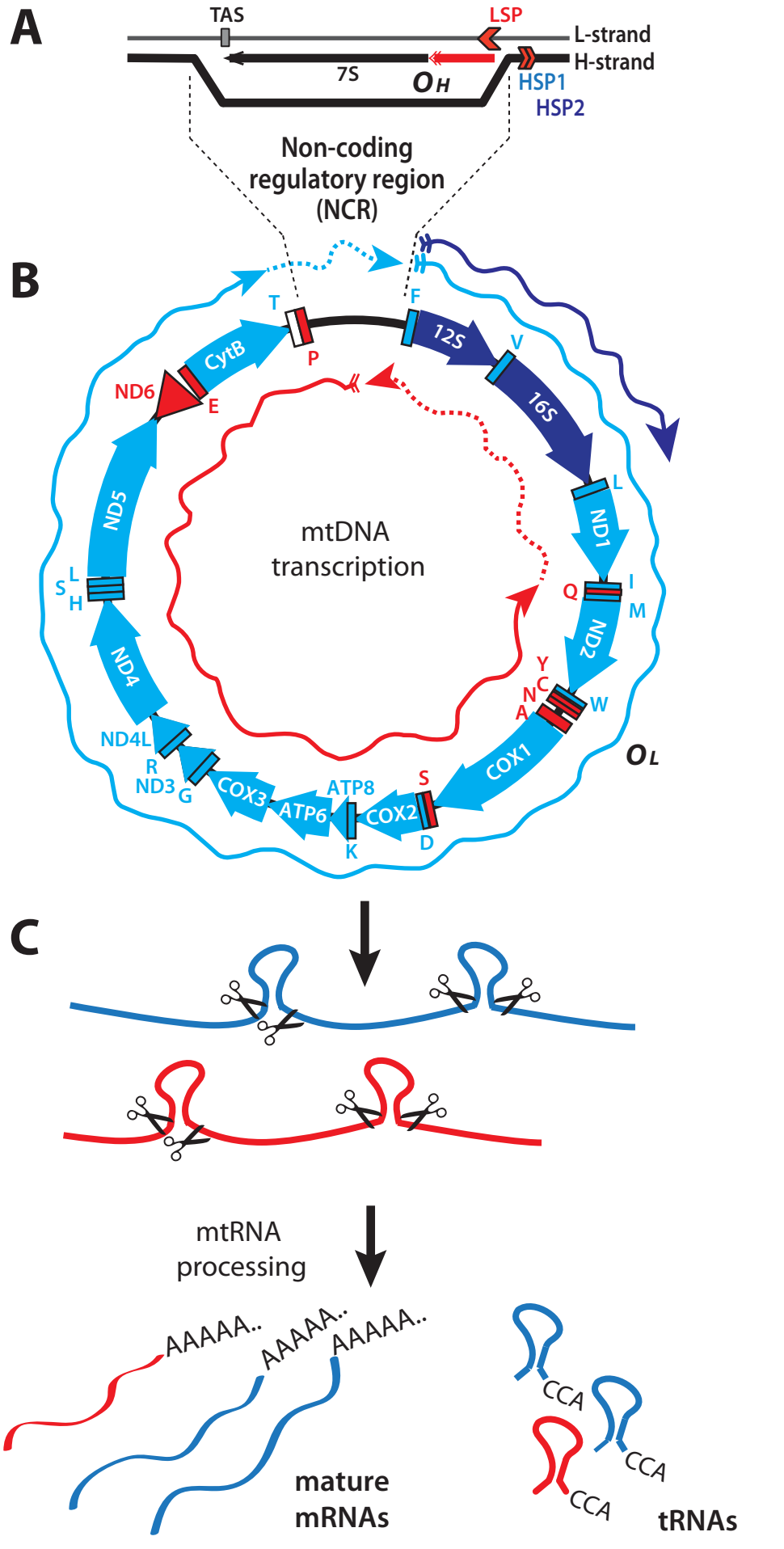

A
TAS
LSP
L-strand
H-strand
7S
O_H
HSP1
HSP2
Non-coding
regulatory region
(NCR)
B
T
F
P
12S
V
16S
L
ND1
I
M
Q
ND2
W
A
N
C
Y
COX1
S
D
COX2
K
ATP8
ATP6
COX3
G
ND3
R
ND4L
ND4
H
S
L
ND5
E
ND6
CytB
mtDNA
transcription
O_L
C
mtRNA
processing
AAAAA..
AAAAA..
AAAAA..
mature
mRNAs
CCA
CCA
CCA
tRNAs

Figure 1

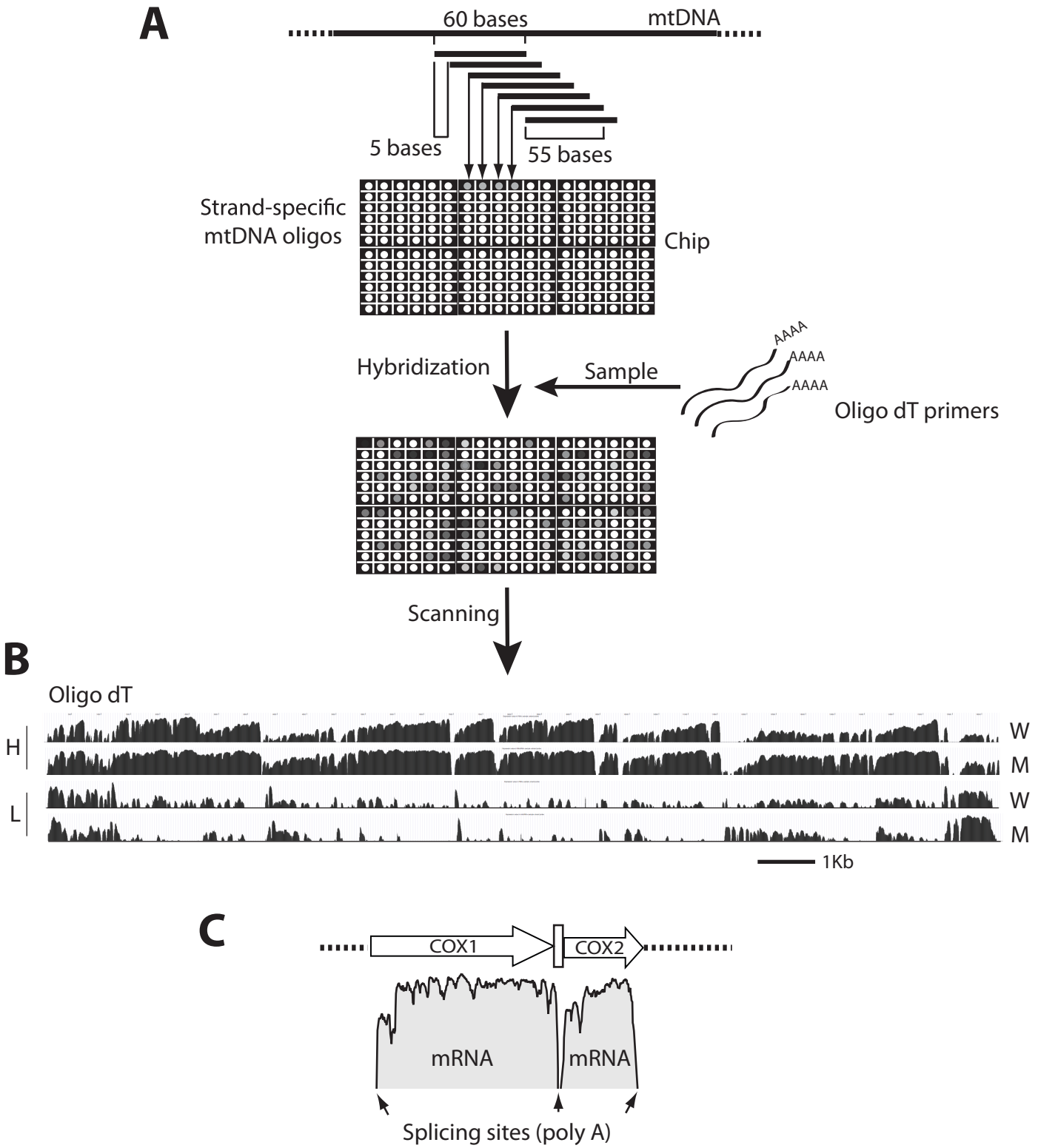

A
60 bases
mtDNA
5 bases
55 bases
Strand-specific
mtDNA oligos
Chip
Hybridization
Sample
AAAA
AAAA
AAAA
Oligo dT primers
Scanning
B
Oligo dT
H
L
W
M
W
M
1Kb
C
COX1
COX2
mRNA
mRNA
Splicing sites (poly A)

Figure 2

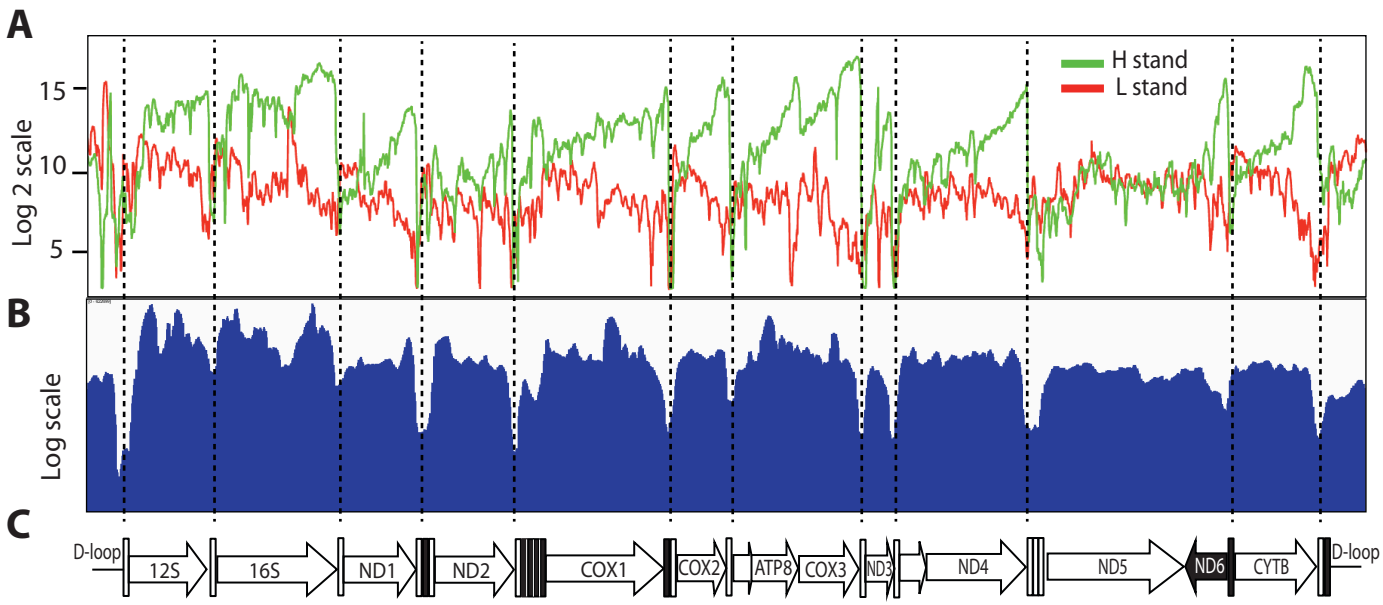

A
Log 2 scale
15
10
5
H stand
L stand
B
Log scale
C
D-loop
12S
16S
ND1
ND2
COX1
COX2
ATP8
COX3
ND3
ND4
ND5
ND6
CYTB
D-loop

Figure 3

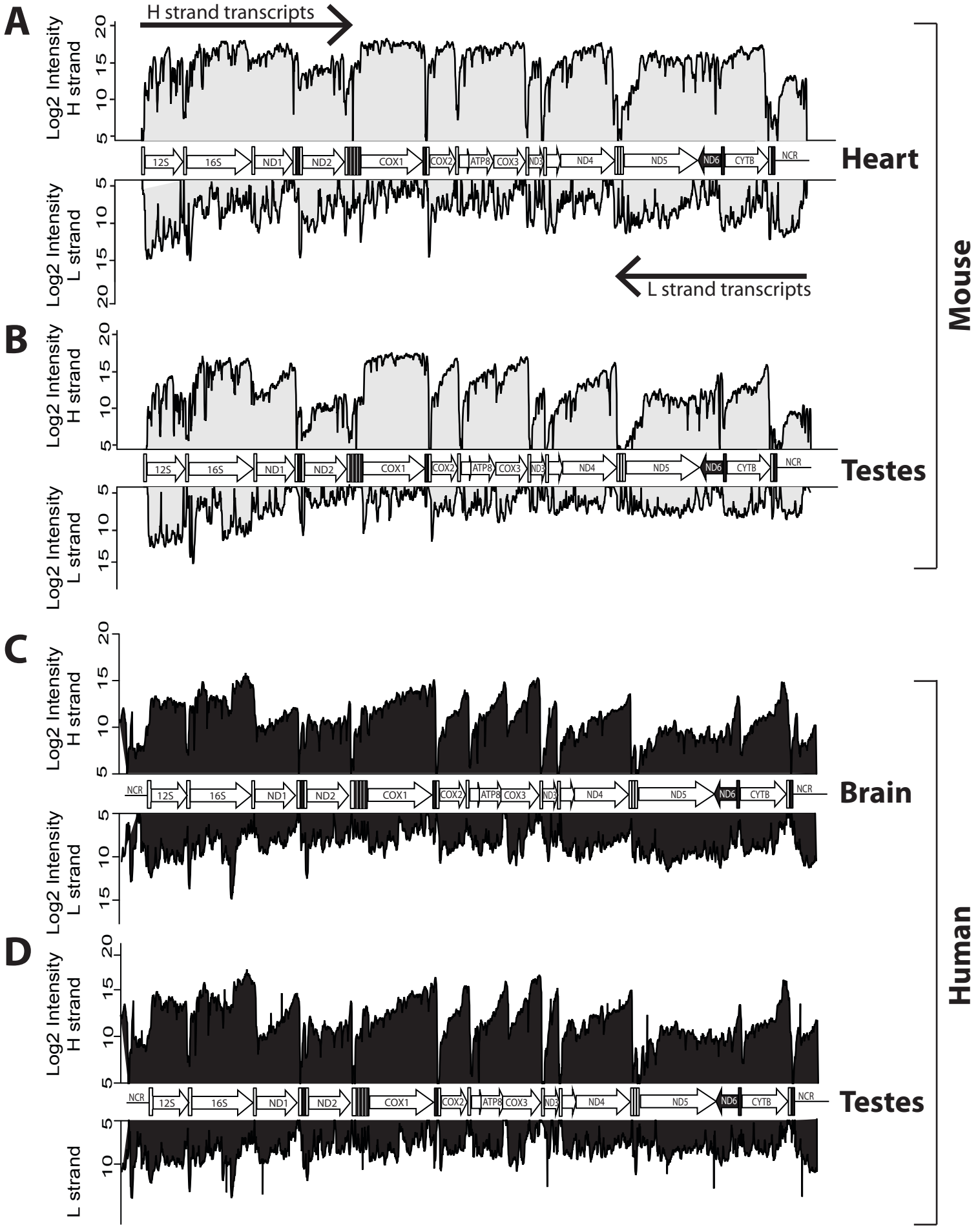

A
H strand transcripts
Log2 Intensity H strand
5 10 15 20
12S 16S ND1 ND2 COX1 COX2 ATP8 COX3 ND3 ND4 ND5 ND6 CYTB NCR
Heart
Log2 Intensity L strand
5 10 15 20
L strand transcripts
B
Log2 Intensity H strand
5 10 15 20
12S 16S ND1 ND2 COX1 COX2 ATP8 COX3 ND3 ND4 ND5 ND6 CYTB NCR
Testes
Log2 Intensity L strand
5 10 15
Mouse
C
Log2 Intensity H strand
5 10 15 20
NCR 12S 16S ND1 ND2 COX1 COX2 ATP8 COX3 ND3 ND4 ND5 ND6 CYTB NCR
Brain
Log2 Intensity L strand
5 10 15
D
Log2 Intensity H strand
5 10 15 20
NCR 12S 16S ND1 ND2 COX1 COX2 ATP8 COX3 ND3 ND4 ND5 ND6 CYTB NCR
Testes
L strand
5 10
Human

Figure 4

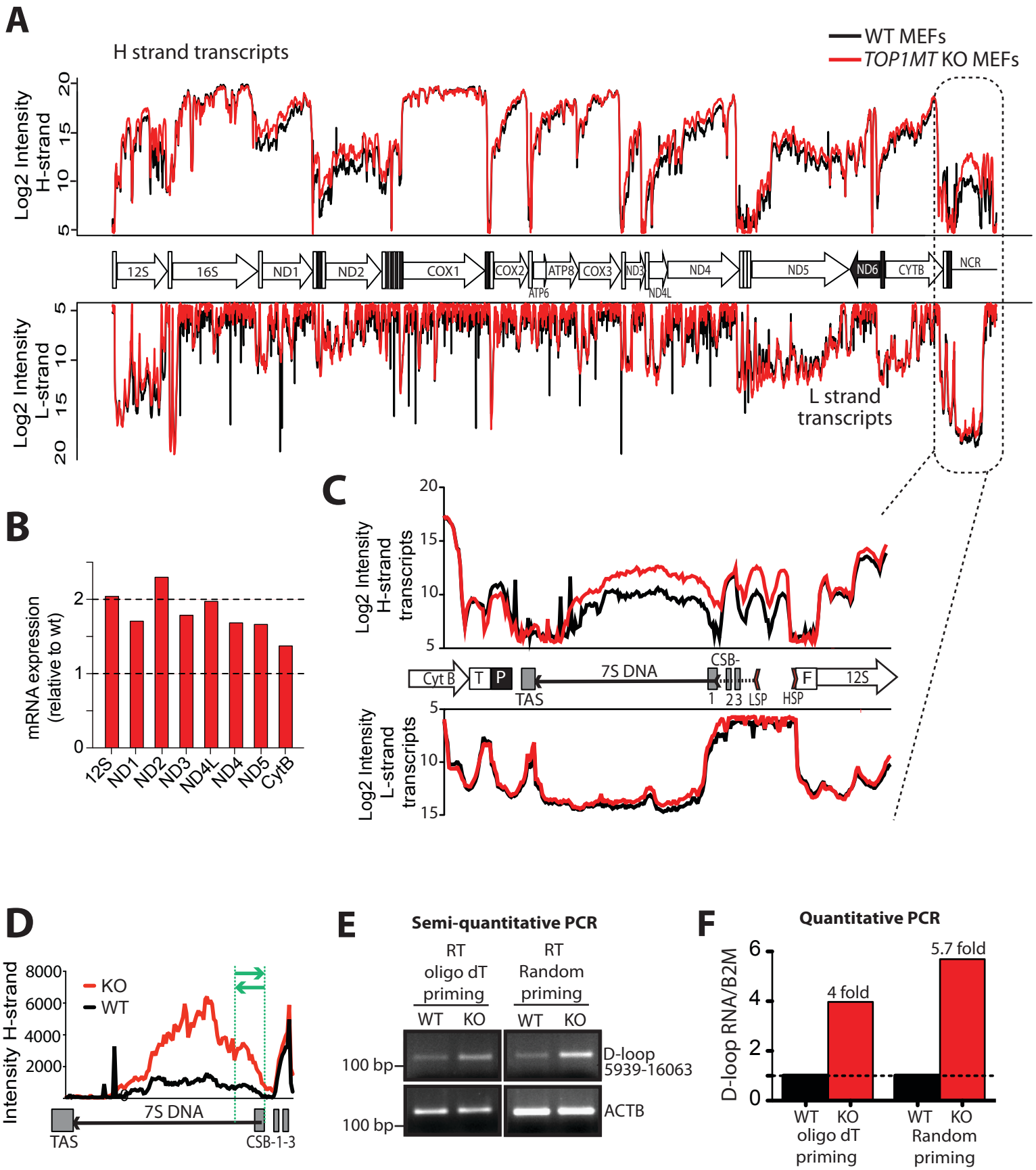

A
H strand transcripts
WT MEFs
TOP1MT KO MEFs
Log2 Intensity H-strand
12S
16S
ND1
ND2
COX1
COX2
ATP8
ATP6
COX3
ND3
ND4L
ND4
ND5
ND6
CYTB
NCR
Log2 Intensity L-strand
L strand transcripts
B
mRNA expression (relative to wt)
12S
ND1
ND2
ND3
ND4L
ND4
ND5
CytB
C
Log2 Intensity H-strand transcripts
Cyt B
T
P
TAS
7S DNA
CSB-1
23
LSP
HSP
F
12S
Log2 Intensity L-strand transcripts
D
Intensity H-strand
KO
WT
7S DNA
TAS
CSB-1-3
E
Semi-quantitative PCR
RT oligo dT priming
RT Random priming
WT
KO
100 bp
D-loop 5939-16063
ACTB
F
Quantitative PCR
D-loop RNA/B2M
4 fold
5.7 fold
WT KO oligo dT priming
WT KO Random priming

Figure 5

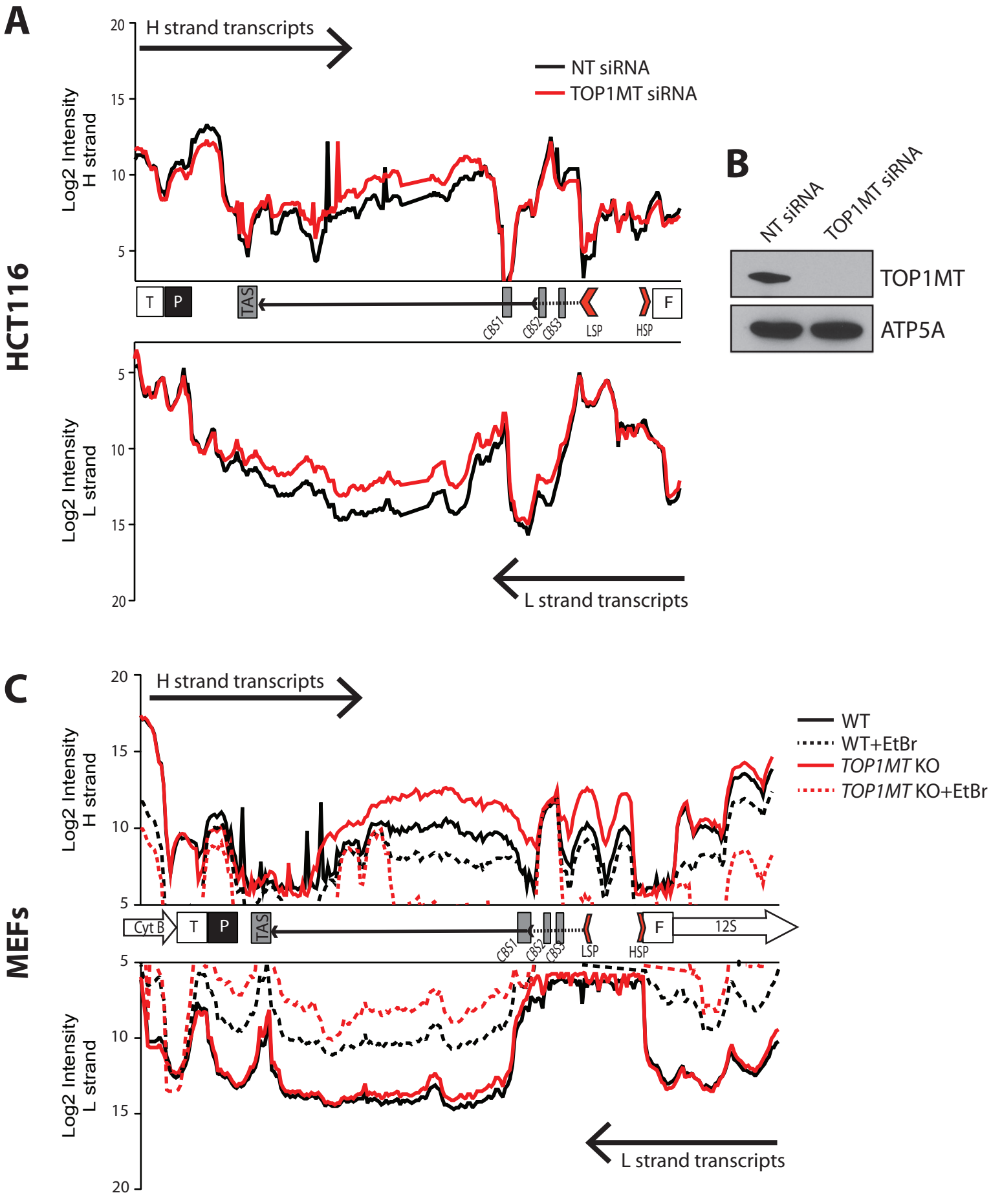
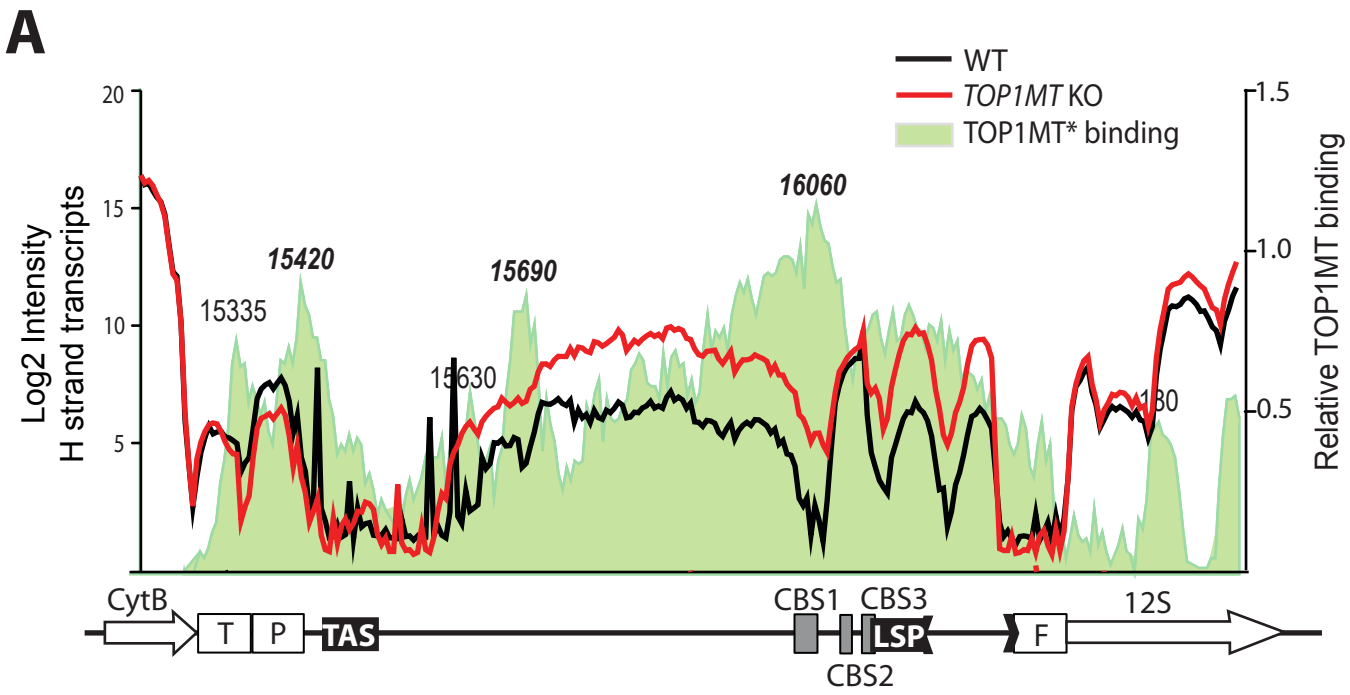
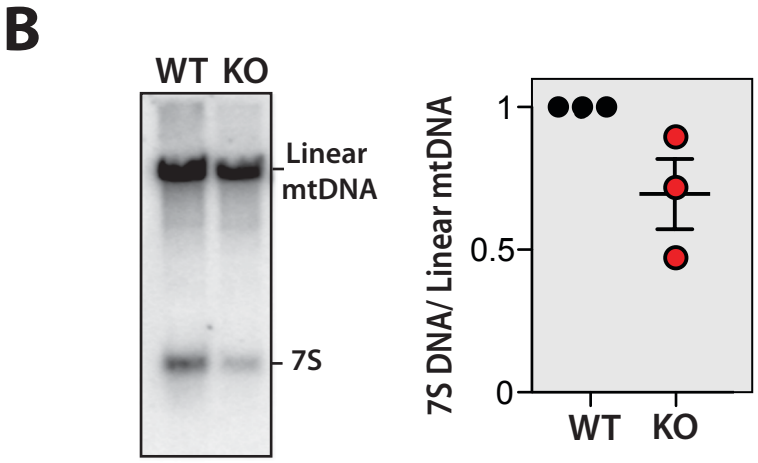

A
HCT116
H strand transcripts
NT siRNA
TOP1MT siRNA
Log2 Intensity H strand
20
15
10
5
T
P
TAS
CBS1
CBS2
CBS3
LSP
HSP
F
Log2 Intensity L strand
L strand transcripts
B
NT siRNA
TOP1MT siRNA
TOP1MT
ATP5A
C
MEFs
H strand transcripts
WT
WT+EtBr
TOP1MT KO
TOP1MT KO+EtBr
Cyt B
12S
L strand transcripts

Figure 6


A
WT
TOP1MT KO
TOP1MT* binding
Log2 Intensity
H strand transcripts
Relative TOP1MT binding
15335
15420
15630
15690
16060
CytB
T
P
TAS
CBS1
CBS3
CBS2
LSP
F
12S


B
WT KO
Linear mtDNA
7S
7S DNA/ Linear mtDNA
WT
KO

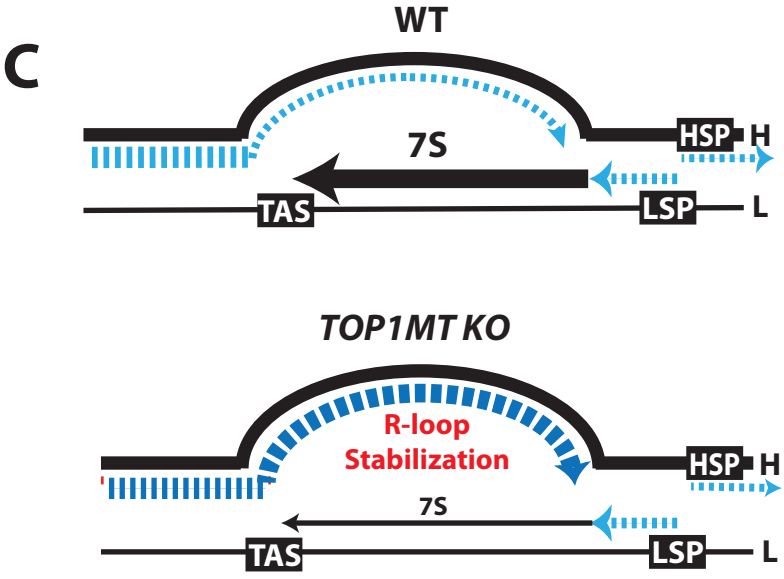

C
WT
7S
HSP
H
TAS
LSP
L
TOP1MT KO
R-loop Stabilization
7S
HSP
H
TAS
LSP
L

Figure 7

Figure legends

Figure 1: Mammalian mitochondrial genome and its transcripts. **A**) Schematic representation of the non-coding regulatory region containing the D-loop mtDNA formed by the extended RNA primer (red) originating from the L-strand promoter (LSP). **B**) The mitochondrial genes are schematized using clockwise and counter-clockwise arrows, representing genes encoded by the Heavy (H) and Light (L) strands, respectively. The H-strand transcripts (rRNAs, mRNAs and tRNAs) are represented in blue (light or dark blue for products generated from HSP2 or HSP1, respectively). tRNAs and the only mRNA (ND6) produced by L-strand transcription are sketched in red. Blue and red arrows represent the polycistronic transcripts produced by transcription initiated at the H and L-strand promoters, respectively. **C**) The polycistronic RNA precursors are processed by excising the tRNAs flanking the open reading frames. Mature messenger and ribosomal RNAs are then maturated by the addition of a 3'-poly(A) tail before being translated or incorporated into mitochondrial ribosomes.

Figure 2: Experimental design. **A**) Array organization and experimental protocol. Each probe of the array is a 60 base strand-specific mtDNA sequence overlapping the two adjacent probes by 55 nucleotides. To detect mitochondrial transcripts, RNAs are retro-transcribed using oligo dT primers. The cDNA fragments obtained are hybridized to the chip and their abundance measured based on the signal intensity for every probe. **B**) Transcription profiles obtained from the hybridization of mature RNAs on the array. W and M indicate profiles obtained using RNA extracted from whole cells or isolated mitochondria, respectively. H and L refer to heavy or light strand profiles. **C**) Data representation: the intensity of each probe (Log scale) is plotted along the mtDNA sequence. Polyadenylated COX1 and COX2 transcripts are shown as examples.

Figure 3: Comparison between mitochondrial transcription profiles of human colon carcinoma HCT116 cells obtained by the tiling array of polyadenylated mtRNAs (A) and RNA-seq (B). Note that the similarity between the ORFs patterns. Sites for endonucleolytic cleavage are indicated as vertical dotted lines.

Figure 4: Polyadenylated mtRNAs profiles in mouse heart (**A**) and testes (**B**), and human brain (**C**) and testes (**D**). The H and L-strand transcripts are plotted mirrored (top and bottom of each panel, respectively). Organization of the mitochondrial genome is sketched in the middle of each panel.

Figure 5: Differential mtRNA profiles in the regulatory region of WT and *TOP1MT* KO MEFs. **A**) Comparison of the mtDNA transcription profiles in WT and KO MEFs (black and red, respectively). **B**) Quantification of mitochondrial transcripts in *TOP1MT* KO MEFs. Average intensities of probes within different ORFs were added together and the values relative to WT for each gene plotted. **C**) Polyadenylated transcripts detected in the mtDNA non-coding region (NCR). The NCR organization is schematized between the H and L-strand transcription profiles. Regulatory sequences within the NCR are abbreviated as HSP and LSP (H- and L- strand promoters, respectively), TAS (termination associated sequence) and CSB-1, 2 and 3 (conserved sequence box 1, 2 and 3). **D**) Quantitation of the abundance of the non-coding RNA transcript produced by H-strand transcription and spanning the 7S DNA. Intensities for each probe of the tiling array are represented as linear scale. **E-F**) Quantitation of the ncRNA by semi-quantitative and quantitative RT-PCR. Total RNA was retro-transcribed with oligo dT or random primers. The cDNA was used as template to amplify mtDNA nucleotides 15939-16063. The region of the non-coding RNA amplified by PCR is showed in green in panel C. For semi-quantitative PCR, PCR products were run on 2% agarose gels and band intensities were visualized after UV transillumination. Beta actin (ACTB) was co-amplified as control. For quantitative PCR, the amounts of PCR products in TOP1mt KO samples were quantified using the ΔΔCt method and are expressed as relative to WT.

Figure 6: ncRNA in the regulatory region of WT and *TOP1MT* KO MEFs. **A)** Polyadenylated transcripts detected in the mtDNA non-coding region of human HCT116 cells after acute *TOP1MT* knockdown. **B)** Western Blot showing the efficiency of *TOP1MT* knockdown 72 hours after siRNA transfection. ATP5A served as control for equal loading. **C)** Stability of the regulatory region

ncRNAs in WT (black) and *TOP1MT* KO MEFs (red). The stability of polyadenylated ncRNAs was assessed by blocking mtDNA transcription with ethidium bromide (EtBr). The abundance of polyadenylated mtRNAs before and after the EtBr treatment is represented by solid and dotted lines, respectively.

Figure 7: Regulatory role of TOP1MT on transcription and replication of the NCR. **A)** Alterations of the NCR transcription patterns in TOP1MT KO cells in relation to TOP1MT binding sites. TOP1MT binding sites are represented by the shaded area colored in light green and determined using the tiling array after isolation of TOP1MT cleavage complexes (see Materials & Methods). **B)** Southern blot analysis of 7S DNA in WT and TOP1MT KO cells. Right panel shows results for 3 independent determinations. **C)** Schematic representation of the NCR transcript and 7S aberrations in *TOP1MT* KO MEFs. DNA and RNA species are represented in black and blue, respectively (see Discussion for details).