PLOS GENETICS

Ribosomal DNA promoter recognition is determined in vivo by cooperation between UBTF1 and SL1 and is compromised in the UBTF-E210K neuroregression syndrome

Michel G. Tremblay[1‡], Dany S. Sibai[1,2‡], Melissa Valère[1,2], Jean-Clément Mars[1,2¤], Frédéric Lessard[1], Roderick T. Hori[3], Mohammad Moshahid Khan[4], Victor Y. Stefanovsky[1], Mark S. LeDoux[5], Tom Moss[1,2]*

1 Laboratory of Growth and Development, St-Patrick Research Group in Basic Oncology, Cancer Division of the Quebec University Hospital Research Centre, Québec, Canada, 2 Department of Molecular Biology, Medical Biochemistry and Pathology, Faculty of Medicine, Laval University, Québec, Canada, 3 Departments of Microbiology, Immunology and Biochemistry, 4 Departments of Neurology and Anatomy & Neurobiology, University of Tennessee Health Science Center, Memphis, Tennessee, United States of America, 5 Department of Psychology, University of Memphis, Memphis TN and Veracity Neuroscience LLC, Memphis, Tennessee, United States of America

¤ Current address: IRIC, Université de Montréal, Montréal, Québec, Canada
‡ These authors share first authorship on this work
* Tom.Moss@crhdq.ulaval.ca

Abstract

Transcription of the ~200 mouse and human ribosomal RNA genes (rDNA) by RNA Polymerase I (RPI/PolR1) accounts for 80% of total cellular RNA, around 35% of all nuclear RNA synthesis, and determines the cytoplasmic ribosome complement. It is therefore a major factor controlling cell growth and its misfunction has been implicated in hypertrophic and developmental disorders. Activation of each rDNA repeat requires nucleosome replacement by the architectural multi-HMGbox factor UBTF to create a 15.7 kbp nucleosome free region (NFR). Formation of this NFR is also essential for recruitment of the TBP-TAF$_I$ factor SL1 and for preinitiation complex (PIC) formation at the gene and enhancer-associated promoters of the rDNA. However, these promoters show little sequence commonality and neither UBTF nor SL1 display significant DNA sequence binding specificity, making what drives PIC formation a mystery. Here we show that cooperation between SL1 and the longer UBTF1 splice variant generates the specificity required for rDNA promoter recognition *in cell*. We find that conditional deletion of the TAF1B subunit of SL1 causes a striking depletion of UBTF at both rDNA promoters but not elsewhere across the rDNA. We also find that while both UBTF1 and -2 variants bind throughout the rDNA NFR, only UBTF1 is present with SL1 at the promoters. The data strongly suggest an induced-fit model of RPI promoter recognition in which UBTF1 plays an architectural role. Interestingly, a recurrent UBTF-E210K mutation and the cause of a pediatric neurodegeneration syndrome provides indirect support for this model. E210K knock-in cells show enhanced levels of the UBTF1 splice variant and a concomitant increase in active rDNA copies. In contrast, they also display reduced rDNA transcription and promoter recruitment of SL1. We suggest the underlying cause of

the UBTF-E210K syndrome is therefore a reduction in cooperative UBTF1-SL1 promoter recruitment that may be partially compensated by enhanced rDNA activation.

Author summary

The 200 Ribosomal RNA gene copies (the rDNA) account for 80% of total cellular RNA and determine the cell's ribosome complement and hence also its capacity to synthesize proteins. As a consequence, these genes directly control cell growth and their misfunction leads both to hypertrophic and developmental disorders. However, it is still unclear what determines the choice of which rDNA copies are activated and how this relates to the formation of the transcription preinitiation complex. Our study, based on the conditional deletion of TAF1B, a subunit of the TBP complex SL1, reveals how the DNA sequence non-specific HMGbox factor UBTF plays a sequence-specific role in SL1 recruitment and preinitiation complex formation. We further show how this role is degraded by a mutation that is the cause of the E210K recurrent pediatric neuroregression syndrome. The data suggest an induced-fit model of preinitiation complex formation in which UBTF creates a DNA folding that is specifically recognized by SL1.

Introduction

The ribosomal RNA (rRNA) genes encode the catalytic and structural RNAs of the ribosome as a single 47S precursor. As such, transcription of these genes is a major determinant of cell growth, cell cycle progression and cell survival and an essential factor in the formation of hypertrophic diseases such as cancer [1]. Dysregulation of rRNA genes is also the cause of a large range of developmental and neurological disorders that are often also associated with cancer [2–6]. To develop treatment strategies for these diseases it is important that we command an understanding of how these genes are transcribed and regulated. Transcription of the several hundred tandemly repeated and essentially identical [7] rRNA genes, the rDNA, is undertaken exclusively by RNA Polymerase I (RPI, Pol1, POLR1) and a set of basal transcription factors dedicated to this task. This strict correspondence of gene and polymerase has resulted in the rapid coevolution of rDNA promoters with basal factors, leading to a high degree of species specificity of the RPI transcription machinery [8–10]. The functional uniqueness of the RPI machinery provides an obvious target for novel therapeutic approaches [11]. However, what directs the RPI transcription machinery exclusively to the rDNA and how it is specifically recruited to both the major 47S pre-rRNA promoter and enhancer element despite these having little or no DNA sequence commonality are still not understood. Here we show that despite having little or no inherent DNA sequence selectivity, the multi-HMGbox Upstream Binding Factor (UBF/UBTF) plays a crucial role in targeting RPI preinitiation complex formation to the rDNA promoters *in vivo*. Further, we show that this UBTF function is compromised by an E210K mutation recently linked to a recurrent human pediatric neuroregression syndrome [6,12–14].

The basal factors of the mammalian RPI transcription machinery include Selectivity Factor 1 (SL1), the multi-HMGbox factor UBTF/UBF and the RPI-associated initiation factor RRN3 [9,15,16] (Fig 1A). SL1 consists of the TATA-box Binding Protein (TBP) and the RPI specific TBP Associated Factors, TAFIA to D. Evolutionary variability of these TAFs determine their species-specific functions and that of RPI transcription, and for example human or mouse SL1
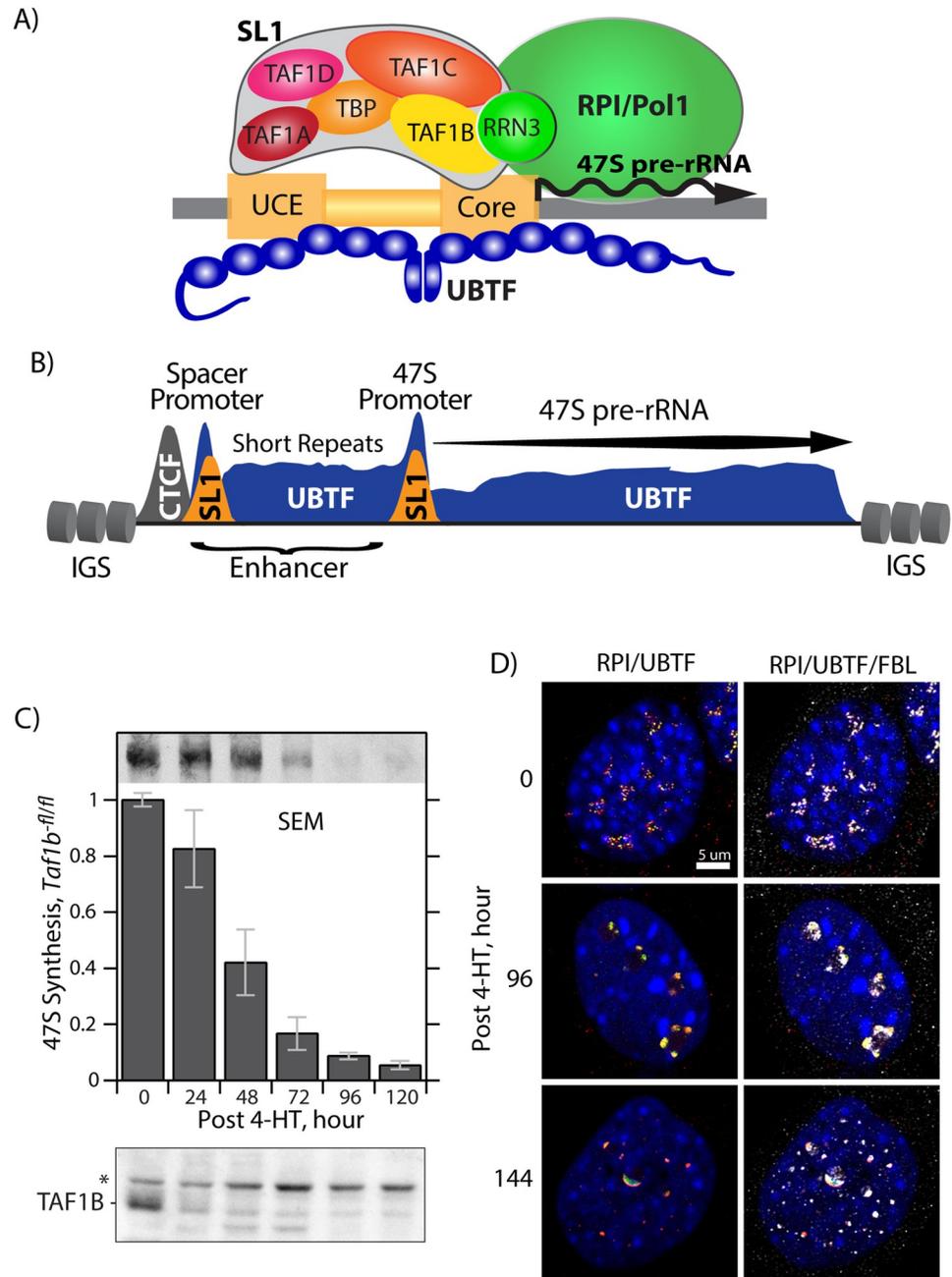
Fig 1. Genetic inactivation of the *Taf1b* gene disrupts rRNA synthesis and nucleolar structure in conditional MEFs. A) Schematic of the basal factors of the RNA Polymerase I (RPI/Pol1) initiation complex and their assembly on the rDNA promoter. The diagram is not intended to indicate precise factor positioning. B) General organisation of the rDNA chromatin structure in mouse. The positions of Spacer and 47S Promoter duplications, the distribution of UBTF, CTCF binding at the upstream boundary and the flanking nucleosomal IGS are indicated. C) Synthesis of 47S pre-rRNA in *Taf1b* conditional MEFs. 47S pre-rRNA synthesis was determined by [3H]-uridine RNA metabolic labelling following homozygous *Taf1b* deletion by 4-HT induction of ER-Cre. The inset in the upper panel shows a typical example of 47S pre-rRNA labelling in *Taf1b$^{fl/fl}$/p53$^{-/-}$/ERcre$^{+/+}$* MEFs at different times after ER-Cre induction. The histogram shows quantification of 47S labelling in *Taf1b$^{fl/fl}$/p53$^{-/-}$/ERcre$^{+/+}$* MEFs normalized to labelling in control *Taf1b$^{wt/wt}$/p53$^{-/-}$/ERcre$^{+/+}$* MEFs. The data were derived from 6 independent analyses of MEFs derived from two floxed and two wild type embryos. Error bars indicate the SEM. The lower panel in C shows a parallel time course of TAF1B depletion, see S2 Fig for more detail. D) RPI, UBTF and fibrillarin (FBL) indirect immunofluorescence labelling in *Taf1b$^{fl/fl}$/p53$^{-/-}$/ERcre$^{+/+}$* MEFs at different time points after 4-HT induction of *Taf1b* deletion, see S3 Fig for a more detail.

https://doi.org/10.1371/journal.pgen.1009644.g001

complexes are not functionally exchangeable [17–19]. In contrast, UBTF is a highly conserved essential factor that is functionally exchangeable between human, mouse and to some extent even Xenopus [19–21]. Our present understanding of how SL1 and UBTF function in rDNA transcription derives predominantly from cell-free studies. These have suggested sequential binding scenarios for the formation of the RPI preinitiation complex, whereby one or two dimers of UBTF bind across the rDNA promoter provide a landing site for SL1, or may stabilize the initial binding of SL1 [22,23]. UBTF was found to bend and loop DNA, suggesting that it could bring together the two distal Upstream Promoter Element (UPE, aka UCE) and Core Elements of the RPI promoter to form such a landing site for SL1 [24]. However, UBTF of itself does not display any significant degree of sequence specific binding, making its role in targeting SL1 difficult to understand [10]. Indeed, UBTF binds continuously throughout the transcribed regions of the active rDNA genes, where it creates a 15.7 kbp long DNase I hypersensitive Nucleosome-Free Region (NFR) that is bounded upstream by CTCF and the enhancer associated Spacer Promoter, and flanked by the nucleosomal InterGenic Spacer (IGS) [25,26] (Fig 1B). Despite this, targeted gene inactivation has unequivocally shown UBTF to be also essential for recruitment of SL1 to both the major 47S promoter and the enhancer associated Spacer Promoter in mouse [25,27].

Here we extend recent studies of UBTF and RRN3 [25,27] to a conditional cell culture model for the TAF1B (TAF68) subunit of SL1, to the roles of the UBTF variants and to a UBTF-E210K mutation recently identified as the cause of a recurrent pediatric neurodegeneration syndrome. The resulting data provide the first *in cell* test of the requirement for an RPI-specific TAF and a significant new insight into RPI preinitiation complex formation. The data resolve key questions surrounding rDNA promoter recognition and RPI pre-initiation complex formation by showing that the UBTF1 splice variant displays a striking specificity for the RPI promoter sequences only when in the presence of a functional SL1. They further suggest an "induced-fit" model of promoter recognition in which UBTF plays an architectural role to model rDNA conformation to fit SL1 and hence catalyze its recruitment. Our findings further suggest that the fundamental cause of the UBTF-E210K pediatric neuroregression syndrome is a partial defect in SL1-UBTF cooperation leading to reduced RPI preinitiation complex formation.

Results

The TAF1B subunit of SL1 is essential for mouse development and rDNA transcription

Mouse lines carrying a targeted "Knockout First" insertion in the gene for TAF1B (TAF68), were established and these crossed to generate lines carrying either conditional $Taf1b^{flox}$ or $Taf1b^{\Delta}$ null-alleles (S1 Fig). Mice heterozygous for a $Taf1b^{\Delta}$ allele were found to be both viable and fertile and the null-allele was propagated at near Mendelian frequency (S1 Table). However, no $Taf1b^{\Delta/\Delta}$ homozygous offspring (pups) were identified and genotyping of embryos detected no $Taf1b^{\Delta/\Delta}$ homozygotes at stages 6.5 and later. It was therefore concluded that $Taf1b$ was essential for mouse development beyond blastula, (see S1 Text for more detail).

To determine the cellular effects of TAF1B loss, $Taf1b$-flox mice (S1A and S1B Fig) were also crossed to introduce ER-Cre and $p53$-null alleles and conditional ($Taf1b^{fl/fl}/ER$-$Cre^{+/+}/p53^{-/-}$) and isogenic control ($Taf1b^{wt/wt}/ER$-$Cre^{+/+}/p53^{-/-}$) mouse embryonic fibroblasts (MEFs) were isolated as previously described [25,28]. Addition of 50nM 4-hydroxytamoxifen (4-HT) to the conditional MEFs induced homozygous recombination of $Taf1b^{flox}$ alleles. This treatment strongly depleted the TAF1B protein already 24h post 4-HT and essentially

eliminated it 96h post 4-HT. As expected for an integral subunit of SL1, depletion of TAF1B was paralleled by the strong suppression of 47S pre-rRNA synthesis (Figs 1C and S2).

Taf1b deletion induces nucleolar stress

Deletion of the *Taf1b* gene also led to a disruption of nucleolar structure characteristic of nucleolar stress. Prior to TAF1B depletion, immunofluorescence imaging (IF) of conditional MEFs revealed the typical punctate sub-nucleolar pattern of RPI and UBTF overlapped by fibrillarin (FBL) staining indicative of transcriptionally active rDNA gene units (0h post 4-HT, in Figs 1D and S3). During TAF1B depletion, RPI and UBTF staining collapsed into common intense foci that were often arranged in pairs around more central FBL staining (e.g. 96 and 120h post 4-HT in Figs 1D and S3). At later times, UBTF became highly condensed and partially segregated from RPI while FBL dispersed throughout the nucleus (120h and 144h post 4-HT in Figs 1D and S3). These changes were consistent with the nucleolar changes previously observed on inactivation of RPI transcription either by RRN3 gene deletion or CX-5461 drug inhibition [25,29]. Despite this, cells continued to slowly divide for at least 120h post 4-HT and remained >95% viable as determined by trypan blue exclusion and MTT tetrazolium reduction assays (S4 Fig).

Loss of TAF1B prevents SL1 recruitment but has only a small effect on "active" rDNA chromatin

Chromatin Immunoprecipitation (ChIP-qPCR) revealed that deletion of the TAF1B gene essentially eliminated TAF1B binding at the 47S and Spacer rDNA promoters in both MEFs and ESCs. It also prevented recruitment of the SL1 subunits TBP and TAF1C at both promoters (Figs 2B and S5). Thus, loss of TAF1B functionally inactivated SL1 and prevented preinitiation complex formation, explaining the suppression of rDNA transcription. In contrast, loss of TAF1B had only a limited effect on UBTF binding across the 47S gene body where it predominantly replaces histone-based chromatin [25,26] (see ETS and 28S amplicons in Figs 2B and S5). This was further confirmed by DNase-Seq analysis that showed little reduction in the pattern of rDNA hypersensitivity on TAF1B loss (S6 Fig). Consistent with this, psoralen accessibility crosslinking (PAC) also indicated only a small reduction in the "active" form of rDNA chromatin previously shown to depend on UBTF [25–27] (Fig 2C). Nevertheless, an increase in the mobility of the "active" PAC band on TAF1B loss suggested some degree of rDNA chromatin compaction, most probably related to the concomitant loss of RPI loading. A similar observation was made when RRN3 was inactivated [25,30].

The recruitment of UBTF and SL1 to the rDNA promoters is highly cooperative

Though TAF1B depletion and the loss of SL1 recruitment to the rDNA had only a small effect on UBTF binding across the gene body, inspection of the ChIP-qPCR mapping suggested that selective depletion did occur at both rDNA promoters (compare SpPr and T0/Pr with ETS and 28S amplicons in Figs 2B and S5). This was confirmed using the higher resolution of Deconvolution ChIP-Seq mapping (DChIP-Seq) [10,25]. Before TAF1B depletion, DChIP-Seq revealed overlapping peaks of SL1 (TAF1B) and UBTF at both rDNA promoters in conditional MEFs (Fig 3A). Subsequent TAF1B depletion essentially eliminated it from the promoters but also strongly suppressed the overlapping peak of UBTF, and this same effect was also seen in TAF1B conditional mESCs (S7A Fig). Despite the TAF1B-dependent loss of UBTF from the rDNA promoters, its binding profile elsewhere across the rDNA was unaffected, though a
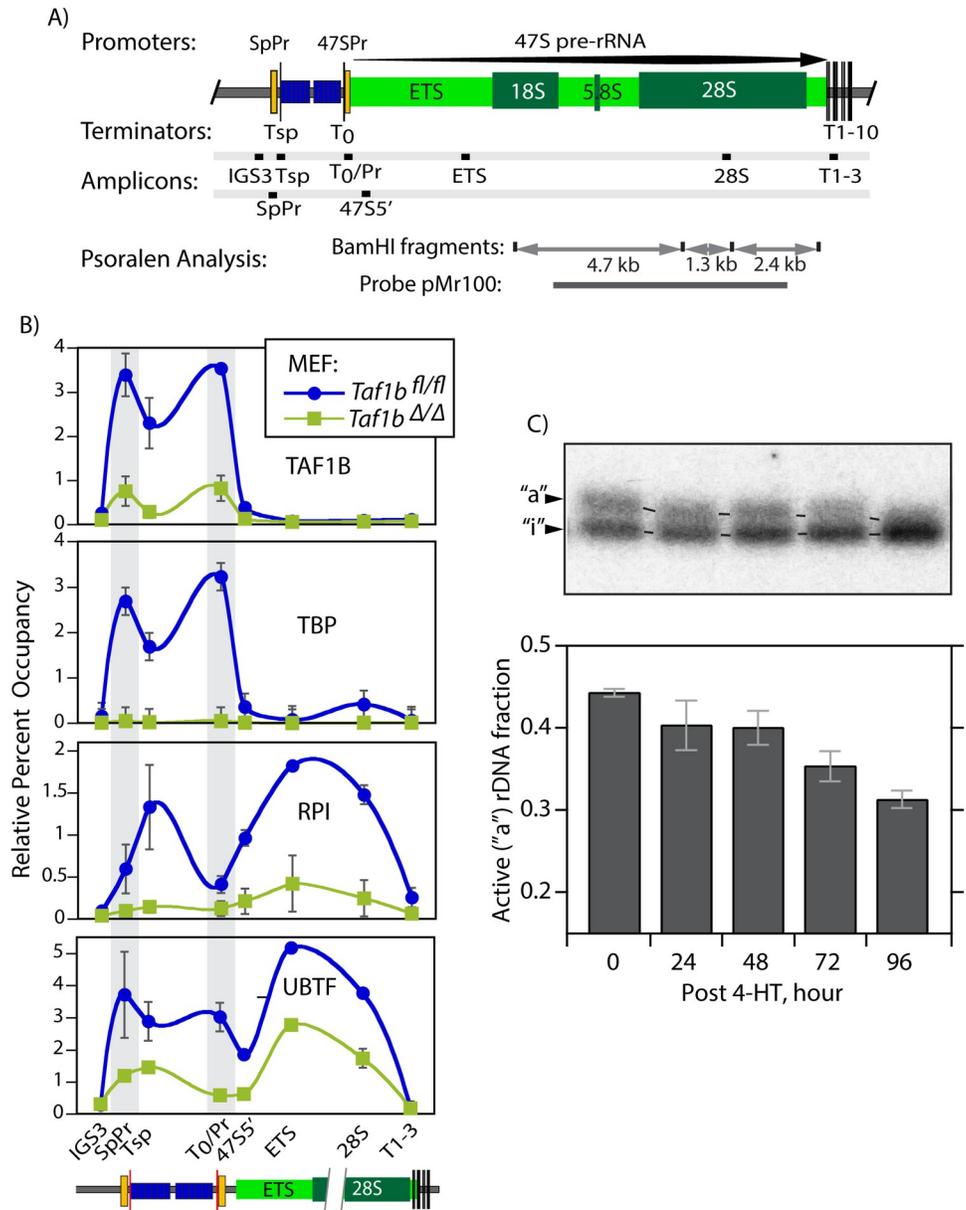
Fig 2. Loss of TAF1B abrogates RPI but not UBTF recruitment to the rDNA. A) Organisation of the mouse rDNA locus indicating the positions of the Spacer (SpPr) and 47S (47SPr) promoter sequences, the RPI termination sites Tsp, T_0 and T1-10, the extent 47S pre-rRNA coding region (light green) and the encoded 18, 5.8 and 28S rRNAs. The positions of the qPCR amplicons used in ChIP analyses are indicated, as are the rDNA fragments and pMr100 probe used in psoralen accessibility cross linking (PAC). B) ChIP-qPCR analysis of TAF1B, TBP, RPI and UBTF occupancy at sites across the rDNA of *Taf1b*$^{fl/fl}$/*p53*$^{-/-}$/*ERcre*$^{+/+}$ MEFs before and 5 days after *Taf1b* inactivation by 4-HT treatment. The data derive from 3 biological ChIP replicas each analyzed by qPCR in triplicate. S5 Fig shows a similar ChIP analysis in *Taf1b*$^{fl/fl}$/*ERcre*$^{+/+}$ mESCs with mapping of TAF1B, TAF1C and TBP subunits of SL1. C) PAC reveals a gradual reduction in the amount and mobility of active rDNA chromatin following *Taf1b* inactivation as in B. Upper panel shows a typical psoralen time course analysis of rDNA chromatin showing the lower mobility of the active "a" and higher mobility of the inactive "i" 1.3kbp BamHI-BamHI fragment from the rDNA 47S coding region. The lower histogram panel shows the mean active rDNA fraction estimated from curve fit analysis of the 1.3, 2.4 and 4.7kbp BamHI-BamHI rDNA fragment profiles in two biological replicas. Error bars indicate the SEM.
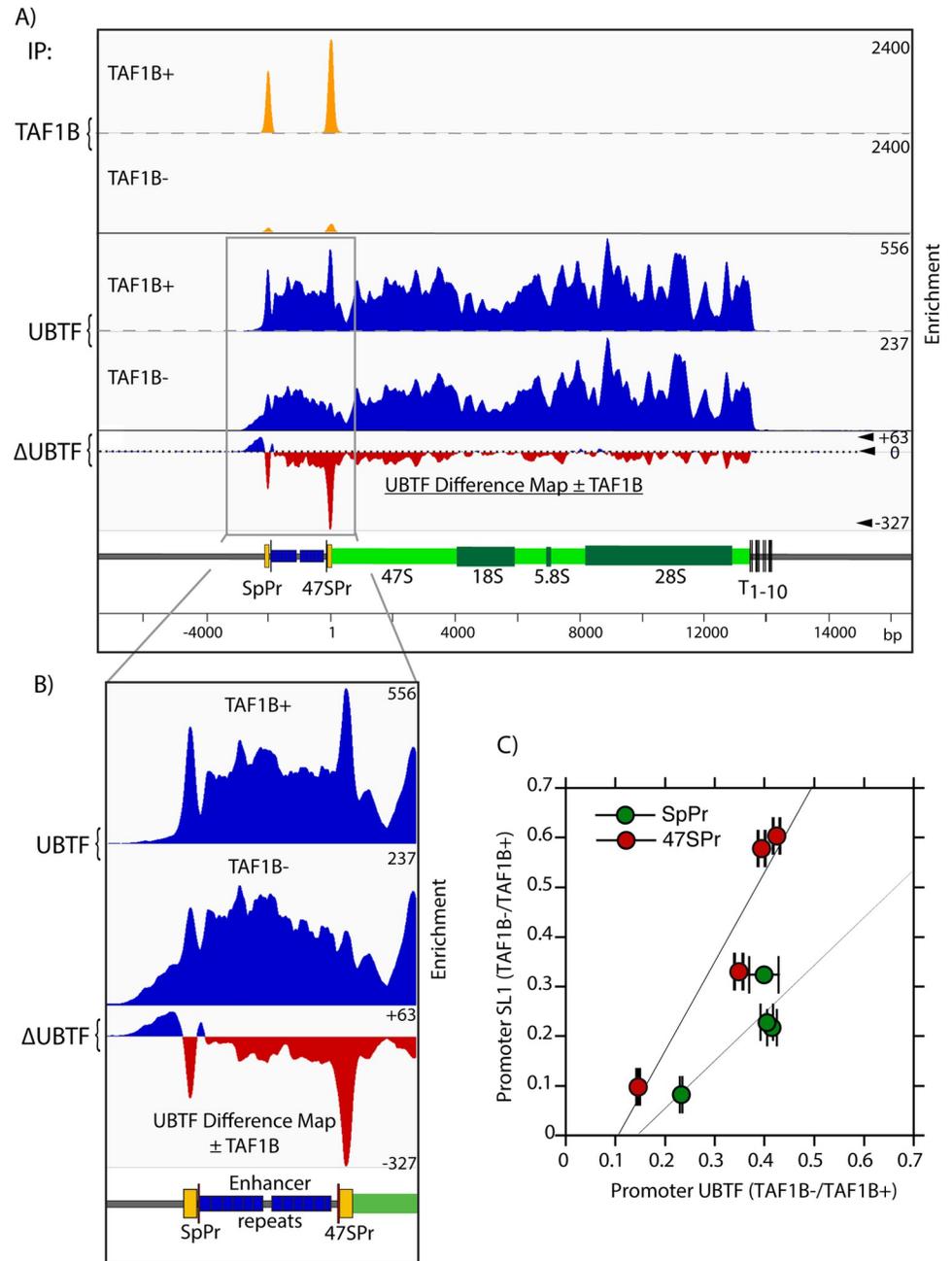
Fig 3. TAF1B loss induces depletion of UBTF from both Spacer and 47S promoters but not from the adjacent enhancer repeats nor from the 47S gene body. A) DChIP-Seq analysis of TAF1B and UBTF occupancy across the rDNA repeat in $Taf1b^{fl/fl}/p53^{-/-}/ERcre^{+/+}$ MEFs before (TAF1B+) and 5 days after $Taf1b$ inactivation (TAF1B-). ΔUBTF indicates the difference map of UBTF occupancy after TAF1B depletion minus the occupancy before TAF1B depletion. B) Magnified view of the DChIP mapping in A showing detail over the promoter and enhancer regions. C) Analysis of Spacer and 47S promoter occupancies reveals a direct proportionality between TAF1B and UBTF. Four independent DChIP UBTF and TAF1B data sets displaying different levels of TAF1B depletion were quantitatively analyzed for UBTF and TAF1B promoter occupancy by peak fit, examples of which are shown in S8 Fig. The fractions of SL1 (TAF1B) and UBTF on each promoter after TAF1B depletion are plotted one against the other and reveal near linear relationships. Error bars in C show the SEM associated with peak fitting.

https://doi.org/10.1371/journal.pgen.1009644.g003

generalized 25 to 50% loss of UBTF respectively in mESCs and MEFs was observed. The dependence of UBTF binding on TAF1B, and hence on functional SL1, was most evident in DChIP-Seq difference maps, which showed strong suppression of UBTF specifically at 47S and Spacer promoters in both MEFs and mESCs types (Figs 3A and 3B and S7A, S7B).

The generation of DChIP-Seq profiles at differing degrees of TAF1B depletion allowed a quantitative estimate of the interdependence of UBTF and SL1 binding at both rDNA promoters (see Materials and Methods and S8 Fig). The data revealed near linear relationships between SL1 and UBTF occupancy and confirmed that their binding was strongly interdependent at either promoter (Fig 3C). This was particularly striking given that these promoters share only 26% base sequence identity, no more than expected for two sequences chosen at random. Interestingly, the data (Fig 3C) potentially also suggested a 2-fold difference in the relative SL1: UBTF stoichiometries between the Spacer and the 47S promoter, possibly a factor in their differential promoter strengths and functionalities.

UBTF1, the longer of the two Ubtf variants is selectively recruited to the rDNA promoters

The observation that UBTF recruitment depended on SL1 specifically at the rDNA promoters but not elsewhere across the rDNA repeat suggested that the UBTF variants might be important in this specificity. Mammals express two splice variants of UBTF, both UBTF1 and UBTF2 encompass six tandem HMGbox DNA binding domain homologies but differ in HMGB-box2, a central segment of which is deleted in UBTF2 (Fig 4A). While MEFs express both forms of UBTF, ESCs naturally express exclusively UBTF1 (S7C Fig). Promoter recruitment of UBTF1 in these cells was found to be strongly suppressed on depletion of functional SL1 (S7A and S7B Fig). However, this left open the question of whether UBTF2 could also cooperate, albeit less efficiently, with SL1 in preinitiation complex formation. To answer this question, we determined the distribution of each UBTF variant across the rDNA in MEFs.

Pools of NIH3T3 MEF clones expressing 3xFLAG-tagged UBTF1 or UBTF2 at sub-endogenous levels were selected and subjected to DChIP-Seq mapping (Figs 4B and S9A and S9B). The profiles of the 3xFLAG-UBTF1 and -UBTF2 binding closely followed that of total endogenous UBTF across most of the rDNA, however, it was significantly different at the Spacer and 47S promoters. Characteristic peaks of UBTF were present at both promoters in the UBTF1 profile but were absent in the UBTF2 profile. This differential promoter binding was most evident in UBTF1-UBTF2 difference maps (Figs 4B and 4C and S9B). Quantitative analysis of the variant UBTF occupancy profiles (Experimental Procedures and S10 Fig) showed greater than 4 times more UBTF1 than UBTF2 at the rDNA promoters (Fig 4D). However, our previous studies of the UBTF-DNA complex showed that a UBTF dimer contacts contiguously 130 to 140 bp of DNA, arguing that each 150-170bp rDNA promoter could interact with at most two dimers of UBTF [10,24]. Hence, the rDNA promoters appear to predominantly, if not exclusively, recruit UBTF1.

The combined data from MEFs and ESCs showed that formation of the RPI preinitiation complex *in cell* involves a cooperation between UBTF1 and SL1. Further, this same cooperation occurs at both Spacer and 47S promoters despite their unrelated base sequences. Since the only difference between UBTF1 and UBTF2 lies in the structure of HMGbox2, this domain must play an important role in UBTF-SL1 cooperation and RPI promoter recognition.

UBTF2 is not detected in the preinitiation complex either in the presence or absence of UBTF1

Given that the UBTF variants can heterodimerize [31], it was surprising that little or no UBTF2 was detected within the preinitiation complexes at either rDNA promoter. This left
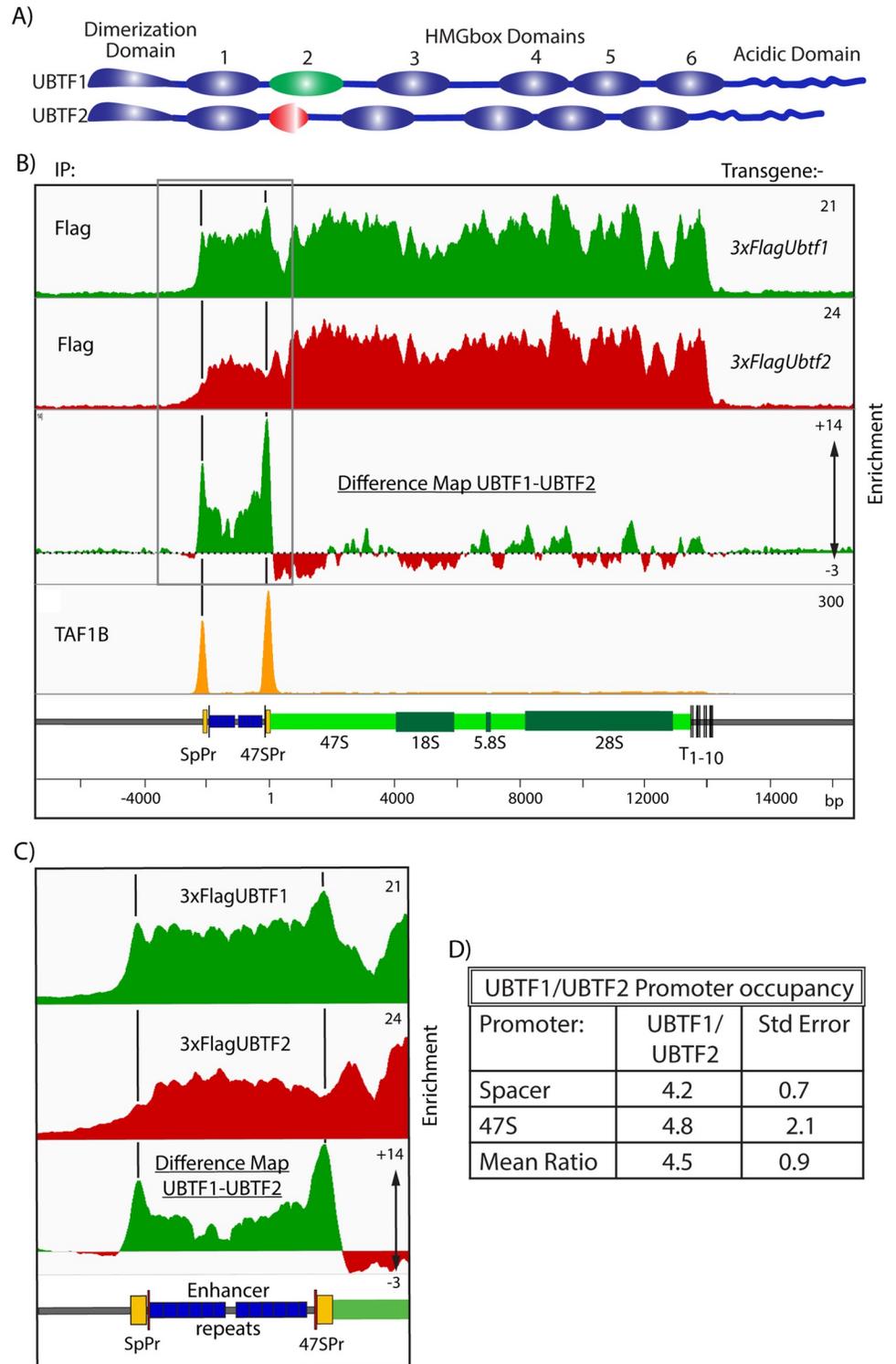
Fig 4. The rDNA promoters selectively recruit the UBTF1 variant. A) Schematic representation of the domain structure of the UBTF splice variants in mouse and human indicating the N-terminal dimerization, 6 HMGbox domains and the C-terminal Acidic Domain. B) DChIP-Seq mapping profiles of exogenously expressed 3xFLAG-UBTF1 or UBTF2, and endogenous TAF1B in NIH3T3 MEFs, see also S9 Fig. The difference map of UBTF1-UBTF2 occupancies reveals a strong selectivity for UBTF1 mapping precisely over the Spacer and 47S promoters. C) Magnified view of the DChIP mapping in B showing detail over the promoter and enhancer regions. D)

Peak fit analysis of UBTF1 and UBTF2 occupancies over the Spacer and 47S promoter revealed that UBTF1 was at least 4 times more prevalent at either promoter. The data derive from two biological replicas and the SEM is shown.

open the question of whether HMGbox2 splicing in UBTF2 eliminates or simply reduces the ability of this variant to cooperate in preinitiation complex formation. To test this, we investigated promoter recruitment of each of the UBTF variants separately. MEFs conditional for both UBTF forms (*Ubtf*^*fl/fl*^/*ER-Cre*^*+/+*^/*p53*^*-/-*^) and expressing 3xFLAG-tagged UBTF1 or UBTF2 at sub-endogenous levels were isolated and subjected to DChIP-Seq mapping before and after deletion of endogenous UBTF (Figs 5A, and S11A and S11B). 3xFlag-UBTF1 was recruited to both Spacer and 47S promoters and this recruitment was enhanced by the loss of endogenous UBTF, as might be expected if this variant were essential. In contrast, 3xFlag-UBTF2 was not detected at the promoters either before or after UBTF deletion, despite 3xFlag-UBTF2 displaying a typical binding profile elsewhere across the rDNA (S11B Fig). Quantitative analysis of the UBTF1 and -2 occupancy profiles in independent experiments suggested that, in the absence of endogenous UBTF, UBTF2 recruitment at the Spacer and 47S promoters was respectively less than 2% and 10% that of UBTF1 (Figs 5B and S12). UBTF1 therefore represents the predominant and very possibly the sole form of UBTF able to cooperate with SL1 in pre-initiation complex formation.

An HMGbox2 mutation associated with neuroregression is unlikely to directly affect DNA interactions

An E>K mutation at residue 210 in HMGbox2 of UBTF was recently shown to be the cause of a recurrent human pediatric neuroregression syndrome [6,12–14]. The key role of HMGbox2 revealed by our study suggested that this mutation might affect the formation of the RPI preinitiation complex *in vivo* and possibly explain the origin of this syndrome. Unfortunately, as yet the structure of HMGbox2 has not been determined experimentally. However, despite a high degree of primary sequence variability, HMGboxes display very similar tertiary structures and DNA contacts, making them accessible to molecular modelling (summarized in S13A Fig). Modelling of UBTF-HMGbox2 revealed a typical HMGB saddle structure with basic residues K198, 200 and 211 lining the DNA binding underside (S13B Fig). Significantly, the sidechain of residue K211, a highly conserved minor groove contact in other HMGboxes, was predicted to be correctly oriented towards the DNA. In contrast, the sidechain of the immediately adjacent E210 residue was predicted to point away from the DNA and lay on the seat of the HMGbox saddle. Furthermore, this predicted sidechain position was unaffected by the E210K mutation (S13C Fig). We concluded that the E210K mutation was extremely unlikely to affect HMGbox2 interactions with the DNA. However, the mutation would create a significant change in the electrostatic surface potential of the seat of HMGbox2 (S13D Fig), suggesting that it could have an effect on interactions with protein factors such as SL1. However, to date we have been unable to detect such effects in UBTF Pull-Down and coIP assays.

The UBTF HMGbox2 E210K mutation suppresses 47S rRNA synthesis in a MEF model

Given that the sequences of human and mouse UBTF are 99% identical, we took advantage of a recently generated *ubtf*^*E210K*^ mouse knock-in model. Mice homozygous for the E210K mutation are viable but exhibit behavioral abnormalities that worsen with increasing postnatal age (details will be described elsewhere). *Ubtf*^*E210K/E210K*^ MEFs were isolated from these mice and found to proliferate somewhat more slowly than MEFs from isogenic wild type littermates,

A)

B)
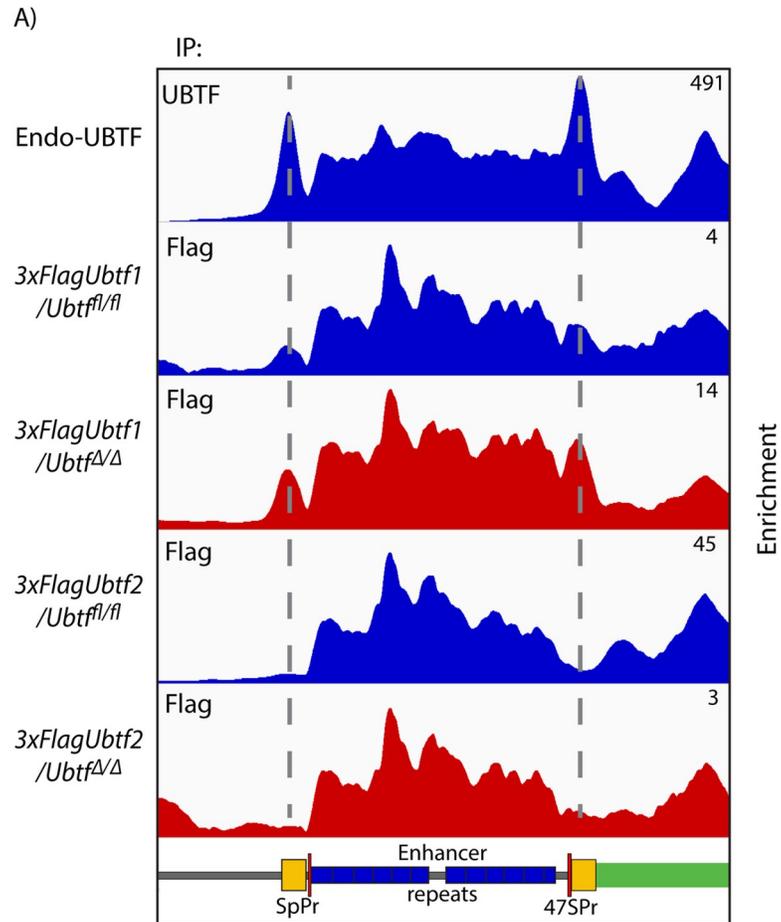
Promoter occupancy, UBTF2/UBTF1 in UBTF-null MEFs ($Ubtf^{\Delta/\Delta}/Er\text{-}cre^{+/+}/p53^{-/-}$)	
Promoter:	Ratio UBTF2/UBTF1 x 100
Spacer	0.5 - 1.5%
47S	6.0 - 8.0%

Fig 5. UBTF1 but not UBTF2 is recruited to the rDNA promoters in the absence of endogenous UBTF. A) DChIP-Seq mapping profiles of exogenous 3xFlag-tagged UBTF1 or UBTF2 in UBTF conditional ($Ubtf^{fl/fl}/ER\text{-}Cre^{+/+}/p53^{-/-}$) MEFs before (blue) and after (red) UBTF deletion. As reference, the upper track shows mapping of the endogenous UBTF shown in Fig 3. B) Quantitation of UBTF1 and 2 at the Spacer and 47S promoters as determined by peak fit analysis, S12 Fig, see Materials and Methods.

doubling times of 35h and 31h respectively (Fig 6A). Metabolic RNA labelling also revealed a >40% lower rate of *de novo* 47S pre-rRNA synthesis in the mutant as compared to the wild type MEFs (Fig 6B), however, no overt rRNA processing defects were detected (S14A Fig). The mutant MEFs also contained 30% less total cellular RNA, (~80% of which is rRNA), than wild type MEFs (Fig 6C). Thus, the E210K mutation in UBTF significantly reduced the capacity of
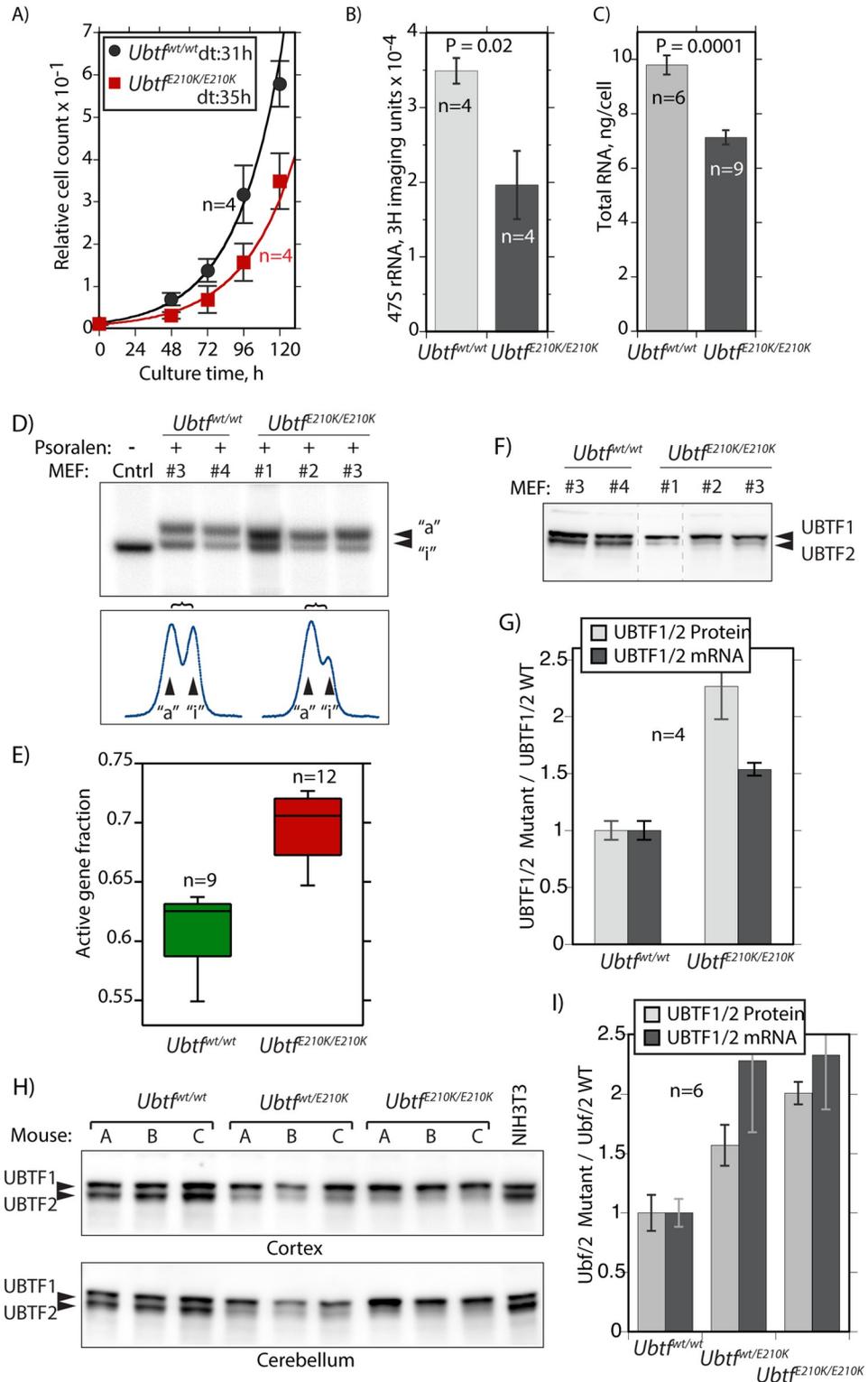
Fig 6. The E210K UBTF mutation suppresses both proliferation and 47S pre-rRNA synthesis but enhances rDNA activation. A) $Ubtf^{E210K/E210K}$ MEFs were found to proliferate significantly more slowly than isogenic wild type ($ubtf^{wt/wt}$) MEFs. Doubling time (dt) was estimated from an exponential curve fit as respectively 35h and 31h for mutant and wild type MEFs. B) 47S pre-rRNA synthesis in $ubtf^{E210K/E210K}$ and wild type MEFs determined by metabolic pulse (30 min) labelling. See also S14A Fig for analysis of processing intermediates at increasing labelling times for individual

MEF isolates. C) Per cell total cellular RNA content of $ubtf^{E210K/E210K}$ and wildtype MEFs. The data in A, B and C derive from two or more biological replicas in each of which a minimum of two independently isolated mutant and wild type MEF cultures were analyzed in parallel. D) PAC analysis of $ubtf^{E210K/E210K}$ and wild type MEFs. Upper panel shows an example of the active rDNA "a" and inactive "i" profiles for the 1.3kbp BamHI-BamHI 47S coding region fragment (see Fig 2A) and the lower panel the corresponding band intensities. E) Active rDNA fractions were estimated from the combined curve fit analysis of 1.3, 2.4 and 4.7kbp BamHI-BamHI rDNA fragment PAC profiles. The data derive from three independent $Ubtf^{E210K/E210K}$ and two wild type MEF isolates in two PAC biological replicas and are plotted to show median, upper and lower data quartiles and outliers. F) and G) Analysis of UBTF1 and 2 levels in $Ubtf^{E210K/E210K}$ and wild type MEF isolates. Panel F shows a typical Western analysis of UBTF variants in these MEFs and panel G quantitative estimates of relative UBTF1/UBTF2 protein and mRNA ratios in these MEFs. H) and I) Show similar estimates of relative UBTF1/UBTF2 protein and mRNA ratios in Cortex and Cerebellum tissue from matched $Ubtf^{wt/wt}$, $Ubtf^{wt/E210K}$ and $Ubtf^{E210K/E210K}$ adult mice. Error bars throughout indicate SEM.

https://doi.org/10.1371/journal.pgen.1009644.g006

MEFs to synthesize rRNA and to assemble ribosomes, explaining their reduced proliferation rate.

The E210K mutation also enhances UBTF1 levels and the fraction of active rDNA repeats

Unexpectedly, the $ubtf^{E210K/E210K}$ MEFs displayed a significant increase in the fraction of activated rDNA copies determined by PAC (Fig 6D and 6E), and this corresponded to an equally significant increase in the expression of the UBTF1 variant both at the protein and mRNA levels (Fig 6F and 6G). A similar bias towards UBTF1 expression was also observed in brain tissue of mutant mice (Fig 6H and 6I). This suggested the interesting possibility that the enhanced levels of UBTF1 in the mutant MEFs revealed an inherent feedback mechanism regulating splicing. In this way the cell might control the fraction of active rDNA copies and hence potentially also rRNA synthesis. However, it will first be necessary to determine whether the E210K mutation directly affected usage of the adjacent splice junctions (see S14B Fig). In either scenario, the increase in active rDNA copies would normally be expected to enhance rRNA synthesis and cell growth in the mutant MEFs. This was clearly not the case since the E210K mutant MEFs displayed both reduced rRNA synthesis and rRNA accumulation and proliferated more slowly than matched wild type MEFs (Fig 6A–6C).

The E210K mutation reduces RPI loading and SL1 and UBTF recruitment to the rDNA promoters

The spatial distribution of UBTF in the nuclei of E210K mutant MEFs closely resembled that in the matched wild type MEFs, UBTF precisely colocalizing with RPI and Fibrillarin (S15 Fig). Hence the mutation had no overt effect on nucleolar topology and did not cause nucleolar stress for example as seen on TAF1B loss (Figs 1D and S3). However ChIP-qPCR analyses revealed that RPI loading across the rDNA was reduce by >40% in the $ubtf^{E210K/E210K}$ mutant MEFs, explaining the observed reduction in pre-rRNA synthesis in these cells (compare RPI loadings in Fig 7A with de novo rRNA synthesis levels in 6B). Recruitment of TAF1B (SL1) and UBTF to both Spacer and 47S rDNA promoters was somewhat reduced in the mutant MEFs, though less than RPI loadings (Fig 7B). Thus, the E210K UBTF mutation most probably reduced pre-initiation complex formation, consistent with it affecting UBTF-SL1 cooperation. The higher resolution of DChIP-Seq further showed that occupancy of UBTF at both 47S and Spacer promoters was selectively reduced by the E210K mutation (Fig 7C), again consistent with a reduced UBTF-SL1 cooperativity. The reduction of UBTF at the rDNA promoters was particularly apparent in difference maps between wild type and E210K mutant MEFs (Figs 7D and S16). The reduction in UBTF was especially strong at the Spacer promoter and corresponded with a similar reduction in TAF1B occupancy and in RPI recruitment (Fig 7D and
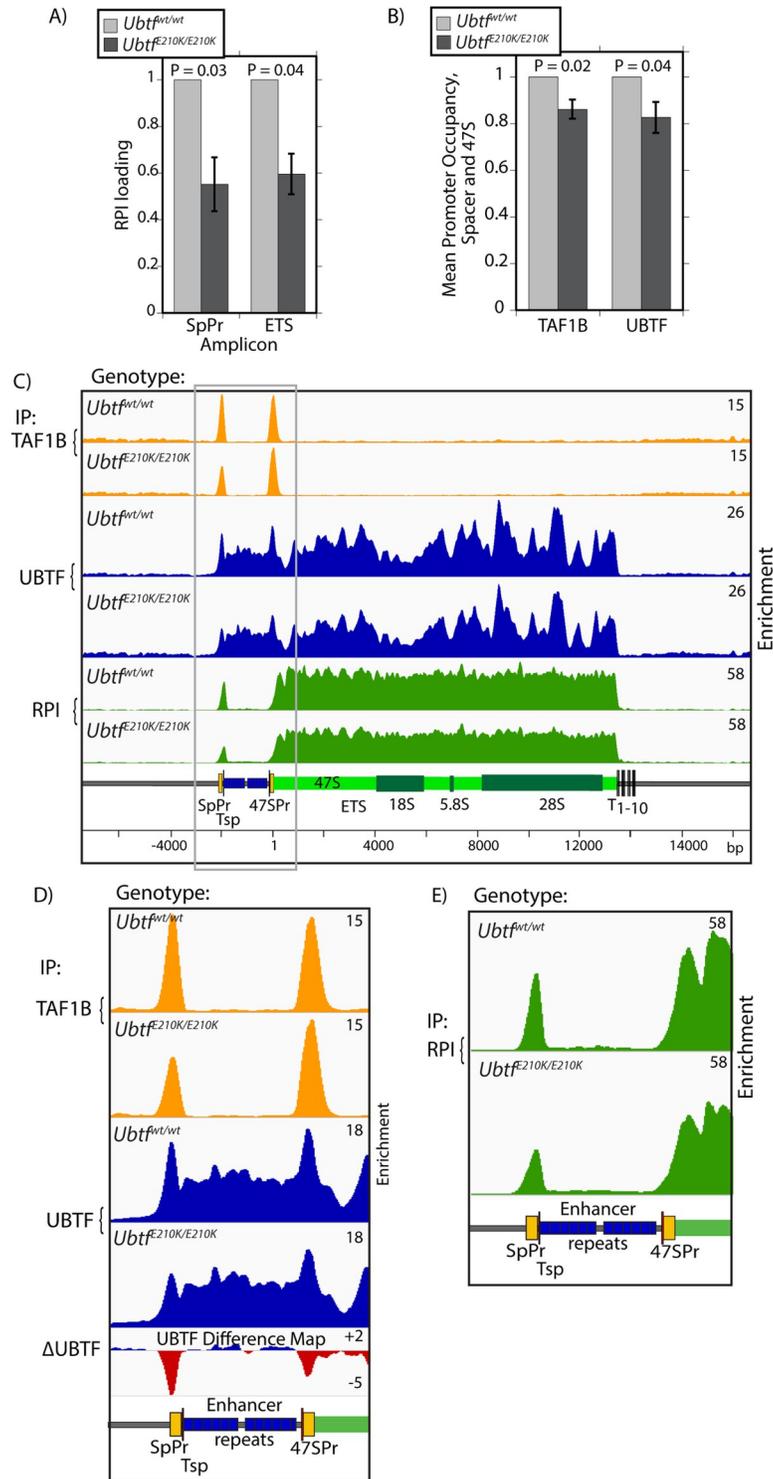
Fig 7. The E210K mutation in UBTF suppresses both the RPI loading across the rDNA and preinitiation complex formation at the Spacer and 47S promoters. A) ChIP-qPCR analysis of RPI occupancy at the Spacer promoter within the ETS region of the 47S coding region in $Ubtf^{E210K/E210K}$ and wild type MEFs. B) ChIP-qPCR analysis of relative preinitiation complex formation in $Ubtf^{E210K/E210K}$ and wild type MEFs. The data show the mean occupancy at amplicons SpPr and T0/Pr by TAF1B or UBTF. The data in A and B derive from 4 independent ChIP-qPCR experiments and error bars indicate the SEM, see Fig 2A and Materials and Methods for amplicon positions. C), D) and E) DChIP-Seq mapping of TAF1B, UBTF and RPI across the rDNA of $Ubtf^{E210K/E210K}$ and wild type MEFs. Panels

D and E show an enlargement of the rDNA promoter and Enhancer region and a difference map of UBTF occupancy (mutant—wild type MEFs). A full-width UBTF difference map is shown in S16 Fig. The data are typical of two biological replicas.

7E). The data strongly suggested that the E210K mutation causes a small but significantly reduced ability of UBTF to cooperate with SL1 in the formation of the RPI preinitiation complex, and together point to a reduction in the efficiency of RPI transcription initiation as the fundamental cause of the UBTF-E210K neuroregression syndrome. Thus, these data indirectly support the central role of the SL1-UBTF1 cooperation in determining RPI preinitiation complex formation and efficient rDNA transcription *in vivo*.

Discussion

Formation of the RPI preinitiation complexes on the rRNA genes (rDNA) in mammals determines as much as 35% of total nuclear RNA synthesis but is still poorly understood. In particular, prior to our study it was unclear how, or indeed if, the multi-HMGbox factor UBTF played a role in targeting the TBP-TAF$_I$ complex SL1 to the rDNA promoters or if it simply acted as a general chromatin replacement protein. UBTF displays little or no DNA sequence binding specificity and binds throughout the 15kbp NFR of the mouse and human rDNA [26]. We previously showed that conditional deletion of the UBTF gene inactivated rDNA transcription and allowed the reformation of nucleosomes across the rDNA [25,27]. Hence, it appeared that UBTF may simply facilitate SL1 recruitment by eliminating the obstacles presented by nucleosomes. Genetic ablation of the SL1 subunit TAF1B has now revealed an unexpected sequence specific role for UBTF and has suggested a novel induced-fit model for RPI preinitiation complex formation.

Though it was assumed from cell-free studies that SL1 would be essential for cell and organism survival due to its role in rDNA transcription, this had not been directly tested. Our data show that homozygous deletion of the gene for the SL1 subunit TAF1B prevented mouse development beyond blastocyst while heterozygous deletants were viable and fertile. Hence, as observed for the other RPI basal factors UBTF and RRN3, TAF1B is an essential factor in mouse. Conditional deletion of *Taf1b* in MEF and mES cell culture was also found to arrest rDNA transcription and to cause severe disruption of nucleolar structure characteristic of nucleolar stress [25,29]. Depletion of TAF1B also prevented promoter recruitment of TAF1C and TBP subunits of SL1 and hence PIC formation at both the 47S pre-rRNA and the Enhancer-associated Spacer rDNA promoters. Quite unexpectedly, this also led to a loss of UBTF at both promoters, though not elsewhere across the rDNA NFR. ChIP-qPCR and high resolution DChIP-Seq showed that the loss of UBTF from the promoters was proportional to the loss of SL1, strongly arguing that binding of these two basal factors was cooperative. Conversely, we had previously shown that *in cell* loss of UBTF eliminated SL1 from the rDNA promoters [25,27], consistent with the cooperative recruitment of these factors. Data from early cell-free studies had suggested two possible scenarios for RPI preinitiation complex formation, either SL1 recruitment depended on pre-binding of UBTF or conversely that UBTF recruitment depended on pre-binding of SL1 [23,32]. Our data resolve this contradiction by showing that in cells UBTF and SL1 binding at the rDNA promoters is strongly interdependent, neither factor being recruited in the absence of the other. The lack of UBTF binding at the promoters in the absence of SL1 was particularly surprising, especially so since UBTF remained bound throughout the rest of the rDNA NFR and even at immediately promoter adjacent sites. Thus, in the absence of SL1 the RPI promoters rather than being preferred sites of UBTF binding, as

usually assumed, are quite on the contrary sites of low affinity within an NFR continuum of higher affinity sites.

Our data further revealed the key importance of the UBTF1 variant in the recruitment of SL1 to the rDNA promoters. Mouse and human cells express varying levels of the UBTF1 and UBTF2 splice variants that differ by a 37a.a. deletion in HMGbox2 of UBTF2 (Fig 4). By mapping these variants across the rDNA we found that UBTF1 was recruited to the rDNA promoters at least four times more often than UBTF2, though the data were also consistent with the exclusive recruitment of UBTF1 at the promoters. In contrast, UBTF1 and UBTF2 bound indistinguishably elsewhere across the rDNA. Since only UBTF1 is present in mESCs, deletion of *Taf1b* in these cells also clearly demonstrated that promoter recruitment of UBTF1 depended on SL1 (S5 and S7 Figs). Thus, formation of the RPI preinitiation complex is driven predominantly if not exclusively by a cooperation between SL1 and UBTF1. This provides the first mechanistic explanation for why UBTF1 is absolutely required for rDNA activity *in vivo* [33].

Recruitment of UBTF1 and SL1 was found to be cooperative not only at the major 47S rDNA promoter but also at the enhancer associated Spacer promoter. Since these promoters display little DNA sequence homology, this raised the question of what in fact defines an RPI promoter and how is it recognized? Our data clearly show that promoter recognition involves the cooperative recruitment of UBTF1 and SL1. Previous data showed that UBTF interacts with SL1 via its highly acidic C-terminal tail, an ~80 a.a. domain containing 65% Asp/Glu residues [34]. However, though this domain is required to enhance cell-free transcription [35] and has been shown to bind TBP [36], it is anyhow present in both UBTF variants. So, while it might play some role in bringing SL1 to the promoters it cannot explain their selective binding of UBTF1. Co-immunoprecipitation also failed to detect a specific interaction between SL1 and the UBTF1 variant. Thus, it seems unlikely that the rDNA promoters are recognized by a pre-formed SL1-UBTF1 pre-initiation complex. Rather we suggest that promoter recognition involves the transient imposition of a specific DNA conformation by UBTF1 that is in turn locked into place by SL1 (Fig 8). There is significant precedent for such a mechanism, since the HMGboxes of UBTF were shown to induce in-phase bending and looping of a DNA substrate. Indeed, it was suggested that such a looping could position UCE and Core promoter elements (Fig 1A) to facilitate their contact by SL1 [24,37,38]. Essentially, UBTF1 might transiently mould the promoter DNA to create an induced-fit for SL1, locking it in place. Since the 37 a.a.
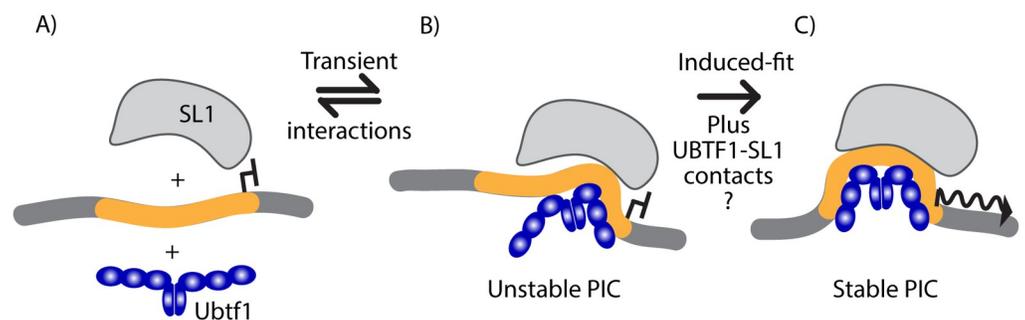
Fig 8. Diagrammatic representation of the induce-fit model for cooperative UBTF1-SL1 recognition and binding at the rDNA promoters. A) Neither UBTF1 nor SL1 alone is able to form a stable interaction with the promoters. B) UBTF1 binding induces a transient reshaping of the promoter that allows SL1 to form weak DNA interactions. C) Tightening of UBTF contacts induces further promoter reshaping inducing a new DNA surface that closely "fits" the DNA interaction surface of SL1. The promoter and flanking DNA sequences are shown respectively in orange and dark grey, the region of UBTF known to bend DNA is shown in blue, SL1 in light grey, and the active transcription initiation site is indicated by an arrow.

deletion in HMGbox2 of UBTF2 prevents this box from bending DNA [38], UBTF2 would mould the promoter DNA differently from UBTF1 and would therefore not induce the appropriate fit for SL1.

Though in the induced-fit model of rDNA promoter recognition direct UBTF1-SL1 contacts would not be essential, they may nonetheless play an important part in stabilizing the pre-initiation complex. Indeed, our study of the UBTF-E210K recurrent pediatric neuroregression syndrome suggested that this was quite probably the case. Molecular modelling showed that while this E210K mutation in HMGbox2 of UBTF was very unlikely to affect interactions with DNA, it might possibly affect interactions with other proteins such as SL1. This said, we did not detect significant effects of the E210K mutation on UBTF1-SL1 interactions in pull-down assays. We found that introduction of the homozygous UBTF-E210K mutation in MEFs significantly reduced rDNA transcription rates, reduce total cellular RNA accumulation and slowed cell proliferation. In apparent contradiction to these effects, the E210K mutation enhanced expression of UBTF1 both in mutant MEFs and mouse tissues, and this led to an increase in the fraction of active rDNA copies, possibly as an attempt to compensate for reduced rDNA transcription. However, ChIP analyses further revealed that the UBTF-E210K mutation reduced the cooperative recruitment of SL1 and UBTF to the rDNA promoters. Thus, it appeared that the primary effect of the E210K-UBTF mutation was to limit PIC formation on the rDNA. This further emphasized the central importance of a functional cooperation between UBTF and SL1 in determining rDNA activity. It further suggested that the UBTF-E210K neurodegeneration syndrome was caused by a subtle defect in PIC formation on the rDNA.

In summary, our study identifies the parameters that determine RNA polymerase I promoter recognition and preinitiation complex formation *in vivo*. We reveal the central importance of a cooperative interaction between the RPI-specific TBP complex SL1 and the UBTF1 splice variant in promoter recognition and propose an induce-fit model for pre-initiation complex formation that explains the functional differences between the UBTF splice variants in terms of their abilities to induce specific conformational changes in the rDNA promoter sequences. Our data further suggest that the UBTF-E210K recurrent neurodegeneration syndrome is caused by a subtle reduction in UBTF-SL1 cooperativity that leads to reduced rDNA transcription.

Materials and methods

Ethics statement

All animal care and animal experiments were conducted in accordance with the guidelines provided by the Canadian Council for Animal Protection, under the surveillance and authority of the institutional animal protection committees of Université Laval and the Centre hospitalier universitaire de Québec (CHU de Québec). The specific studies described were performed under protocols #2014–100 and 2014–101 examined and accepted by the "Comité de protection des animaux du CHU de Québec". This ensured that all aspects of the work were carried out following strict guidelines to ensure careful, consistent, and ethical handling of mice.

Primary antibodies for Immunofluorescence, ChIP and Western blotting

Rabbit polyclonal antibodies against mouse UBTF, RPI large subunit (RPA194/Polr1A), and TAF1B were generated in the laboratory and have been previously described [25], anti-TAF1C was a gift from I. Grummt. All other antibodies were obtained commercially; anti-Fibrillarin (#LS-C155047, LSBio), anti-RPA135 (#SC-293272, Santa Cruz), anti-RPA194 (#SC-48385, Santa Cruz), anti-TBP (#ab818 and #ab51841, Abcam) and anti-FLAG M2 (F3165, Sigma).

Generation of *Taf1b* mutant mice

Taf1b+/- (targeted allele *Taf1btm1a(EUCOMM)Hmgu*) embryonic stem (ES) cells were obtained from EuMMCR and generated using the targeting vector PG00150_Z_6_E02. Two ES clones (HEPD0596_3_G02 and HEPD0596_3_H01) were each used to generate independent mouse lines using the services of the McGill Integrated Core for Animal Modeling (MICAM).

The resulting *Taf1b^fl-neo^* mouse lines carried a "knock-out first" allele in which Lox recombination sites were inserted in intron 3 and intron 5, and a *neo* selective marker gene flanked by FRT sites inserted intron 3 (S1A and S1B Fig). Mouse lines heterozygous for the *Taf1b^fl-neo^* allele derived from the two ES clones were viable, fertile, and appeared phenotypically normal compared to their wild-type littermate but no homozygotes were identified. These mice were then crossed with FLPo (FLP recombinase) and Cre expressing mice (Jackson Laboratory strains FLPo (#012930), Sox2-Cre (#004783)) to generate both *Taf1b^fl^* and *Taf1b^A^* alleles (S1A Fig). Subsequently Cre and FLPo transgenes were removed by backcrossing. *Taf1b^fl/fl^/ER-cre^+/+^*, *Taf1b^fl/fl^/ER-cre^+/+^/p53^-/-^* and corresponding *Taf1b^+/+^* control mouse lines were generated by crossing with ER-Cre expressing and p53-null mice (Jackson Laboratory strains ER-Cre (#004847) and p53 KO (#002101) and used to generate MEF and mES cell lines as previously described [25]. Mice were genotyped by PCR using the primers (S1A Fig): A; 5'-gtcccttcctcactgatcac, B; 5'-tgcagattaggtggcctcag, C; 5'-ccctctcaccttctacccca and D; 5'-ctgggcttggtggctgtaa.

Embryo collection and genotyping

Heterozygous *Taf1b^A/wt^* mice were inter-crossed and embryos isolated, imaged and genotyped from pregnant females at E3.5, 6.5, 7.5, 8.5 and E9.5 as described in [25,27]. DNA from E3.5 embryos was amplified using the REPLI-g Mini kit (QIAGEN). Individual embryos were genotyped by PCR using the same primers as for mouse lines (S1A Fig).

Isolation and culturing of *Taf1b* conditional MEF and mES cells

Conditional TAF1B primary mouse embryonic fibroblasts (MEFs) were generated from E14.5 *Taf1b^fl/fl^/ER-Cre^+/+^/p53^-/-^* and wild type control *ER-Cre^+/+^/p53^-/-^* embryos as previously described [27,39], and were genotyped by PCR as described for mice, see S1A and S1B Fig.

Mouse Embryonic Stem (mES) cells were derived from the inner cell mass of *Taf1b^fl/fl^/ER-Cre^+/+^* and wild type control *ER-Cre^+/+^* blastocysts essentially as published [40]. After establishment of the cells on feeder monolayers, they were adapted to feeder-independence on 2i/LIF N2B27 (ThermoFisher) free serum medium [41] and subsequently maintained in this medium. The *Taf1b^fl/fl^/ER-Cre^+/+^* and control mESCs were genotyped as for MEFs.

Primary MEFs were also generated from E14.5 *Ubtf^E210K/E210K^* and wild type control *Ubtf^wt/wt^* sibling embryos from three independent litters, immortalized by transfection with pBSV0.3T/t [27] and genotyped by base sequencing of PCR products generated using primers 5' CTGGGTGAAGTAGGCCTTGG and 5' CCAGGAGGGTAAGGTGGAGA flanking the mutation site. All MEFs were cultured in Dulbecco's modified Eagle medium (DMEM)-high glucose (Life Technologies), supplemented with 10% fetal bovine serum (Wisent, Life Technologies or other), L-glutamine (Life Technologies) and Antibiotic/Antimycotic (Life Technologies).

Inactivation of *Taf1b* or *Ubtf* in cell culture

Gene inactivation in MEFs and ESCs followed the previously described procedures [25,27]. Briefly, cells were plated in 6 cm petri dishes (0.8×10^6 cells each) and cultured for 18 hours in

DMEM, high glucose, 10% fetal bovine serum or 2i/LIF N2B27 free serum medium as appropriate. For *Taf1b* or *Ubtf* inactivation, 4-hydroxytamoxifen (4-HT) was added in parallel to conditional and wild type control cell cultures to a final concentration of 50nM (the 0h time point for analyses). After 4 hr incubation the medium was replaced with fresh medium without 4-HT. Cell cultures were then maintained for the indicated times and systematically genotyped by PCR on harvesting.

Analysis of TAF1B and UBTF1/2 protein and mRNA levels

TAF1B, and UBTF1/2 protein levels were monitored by Western blotting. At harvesting, cells were quickly rinsed in cold phosphate buffered saline (PBS), recovered by centrifugation (2 min, 2000 r.p.m.) and resuspended directly in SDS–polyacrylamide gel electrophoresis (SDS-PAGE) loading buffer [42]. After fractionation by SDS-PAGE, proteins were analysed by standard Western blotting procedures using an HRP conjugated secondary antibody and Immobilon chemiluminescence substrate (Millipore-Sigma). Membranes were imaged on an Amesham Imager 600 (Cytiva) and UBTF1/2 ratios were determined from lane scans using ImageJ [43] and Gaussian curve fit using MagicPlot Pro (Magicplot Systems). Relative UBTF1/2 mRNA levels were determined by PCR on total cDNA using primers bracketing the spliced sequences (5'TGCCAAGAAGTCGGACATCC and 5'TCCGCACAGTACAGGGAGTA). Products were fractionated by electrophoresis on a 1.5 or 2% agarose EtBr-stained gel, photographed using the G:BOX acquisition system (Syngene) and UBTF1/2 mRNA ratios determined using ImageJ and Gaussian curve fitting as for proteins.

Determination of rRNA synthesis rate

The rate of rRNA synthesis was determined by metabolic labelling immediately before cell harvesting. 10 μCi [^{3}H]-uridine (PerkinElmer) was added per 1ml of medium and cell cultures incubated for a further 30min to 3h as indicated. RNA was recovered with 1 ml Trizol (Invitrogen) according to the manufacturer's protocol and resuspended in Formamide (Invitrogen). One microgram of RNA was loaded onto a 1% formaldehyde/MOPS Buffer gel [44,45] or a 1% formaldehyde/TT Buffer gel [46]. The EtBr-stained gels were photographed using the G:BOX acquisition system (Syngene), irradiated in a UV cross-linker (Hoefer) for 5 min at maximum energy, and transferred to a Biodyne B membrane (Pall). The membrane was UV cross-linked at 70 J/cm^2, washed in water, air dried and exposed to a Phosphor BAS-IP TR 2025 E Tritium Screen (Cytiva). The screen was then analyzed using a Typhoon imager (Cytiva) and quantified using the ImageQuant TL image analysis software.

Psoralen crosslinking accessibility and Southern blotting

The psoralen crosslinking accessibility assay and Southern blotting were performed on cells grown in 60 mm petri dishes and DNA was analyzed as previously described [47,48], using the 6.7kb 47SrRNA gene EcoRI fragment (pMr100) [47]. The ratio of "active" to "inactive" genes was estimated by analyzing the intensity profile of low and high mobility bands revealed by phospho-imaging on an Amersham Typhoon (Cytiva) using a Gaussian peak fit generated with MagicPlotPro (MagicPlot Systems LLC).

Indirect immunofluorescence (IF) microscopy

Cells were plated on poly-lysine treated coverslips and subjected to the standard 4-HT treatment to induce *Taf1b* deletion. At the indicated time points cells were rinsed with PBS, fixed in 4% PFA in PBS for 10 min and permeabilized with 0.5% Triton, PBS for 15 minutes. After a

blocking step in PBS-N (PBS, 0.1% IGEPAL (Sigma)), 5% donkey serum, cover slips were incubated with primary antibodies in PBS-N, 5% donkey serum for ~16h at 4 deg. C. RPI was detected using a combination of mouse anti-A194 and A135 antibodies (#SC-293272, #SC-48385), fibrillarin with goat anti-FBL (#LS-C155047), and UBTF with rabbit anti-UBTF (in-house #8). Cells were incubated for ~ 2h at room temperature with the appropriate AlexaFluor or Dylight 488/568/647 conjugated secondary antibodies (ThermoFisher / Jackson ImmunoResearch) and counterstained with DAPI. After mounting in Prolong Diamond (Thermo-Fisher), epifluorescent 3D image stacks were acquired using a Leica SP5 II scanning confocal microscope and LAS-AF (Leica Microsystems) and Volocity (Quorum Technologies) software.

Chromatin immunoprecipitation (ChIP)

Cells were fixed with 1% formaldehyde for 8 min at room temperature. Formaldehyde was quenched by addition of 125 mM Glycine and cells harvested and washed in PBS. Nuclei were isolated using an ultrasound-based nuclei extraction method (NEXSON: Nuclei Extraction by SONication) [49] with some modifications. Briefly, for all cell types, 33 million cells were resuspended in 1.5 ml of Farnham lab buffer (5 mM PIPES pH 8.0, 85 mM KCl, 0.5% IGEPAL, protease inhibitors). Cell suspensions were sonicated in 15 ml polystyrene tubes (BD #352095) using 3 to 4 cycles of 15 sec on: 30 sec off at low intensity in a Bioruptor (Diagenode). After recovery of the NEXSON-isolated nuclei by centrifugation (1000g, 5 min), nuclei were resuspended in 1.5 ml of shearing buffer (10 mM Tris-HCl pH 8.0, 1 mM NaEDTA, 0.1% SDS, protease inhibitors) and sonicated for 25 min, 30 sec on: 30 sec off, at high intensity. Each immunoprecipitation (IP) was carried out using the equivalent of 16×10^6 cells as previously describe [25]. To map UBTF1 and UBTF2 variants cDNAs encompassing the complete coding regions (M61726 / M61725) were subcloned into pCDNA3Hygro-3xFLAG-C1 (N. Bisson) and verified by Sanger sequencing. The resulting pC3xFLAG-UBF1 and -UBF2 (lab. stocks #2072, #2073) constructs were used to transfect NIH3T3 or conditional *Ubtf* (*Ubtf*$^{fl/fl}$/*ER-Cre*$^{+/+}$/*p53*$^{-/-}$) MEFs [25] and cultures selected with Hygromycin. In the case of NIH3T3, pools of positive clones expressing UBTF1 or UBTF2 at sub-endogenous levels were then subjected to parallel ChIP for FLAG-UBTF1/2 (anti-FLAG) and total UBTF. Essentially the same procedure was used for the MEFs except that FLAG-UBF1 and 2 expressing cell clones were first isolated and these used in ChIP experiments 5 days after *Ubtf* inactivation by 4-HT addition or mock inactivation to eliminate endogenous UBTF.

ChIP-qPCR analysis

All ChIP experiments included a minimum of 2 biological replicates and were analyzed as previously described [25]. For qPCR analysis, reactions (20 μl) were performed in triplicate using 2.5 μl of sample DNA, 20 pmol of each primer, and 10 μl of Quantifast SYBR Green PCR Master Mix (QIAGEN) or PowerUp SYBR Green Master Mix (ThermoFisher). Forty reaction cycles of 10 s at 95˚C and 30 s at 58˚C were carried out on a Multiplex 3005 Plus (Stratagene/Agilent). The amplicon coordinates relative to the 47S rRNA initiation site (BK000964v3) were as follows: IGS3, 42653–42909; SpPr, 43076–43279; Tsp, 43253–43447; T0/Pr (47SPr), 45133–40; 47S-5', 159–320; ETS, 3078–3221; 28S, 10215–10411; T1-3, 13417–13607. Data was analyzed using the MxPro software (Agilent). The relative occupancy of each factor was determined by comparison with a standard curve of amplification efficiency for each amplicon using a range of input DNA amounts generated in parallel with each qPCR run.

ChIP-Seq and data analysis

ChIP DNA samples were quality controlled by qPCR before being sent for library preparation and 50 base single-end sequencing on an Illumina HiSeq 2500 or 4000 (McGill University and Genome Quebec Innovation Centre). Sequence alignment and deconvolution of factor binding profiles to remove sequencing biases (Deconvolution ChIP-Seq, DChIP-Seq) were carried out as previously described [10,25]. The manual for the deconvolution protocol and a corresponding Python script can be found at https://github.com/mariFelix/deconvoNorm. Gaussian curve fitting to transcription factor binding profiles was perform using MagicPlot Pro (Magicplot Systems) on data from the DChIP-Seq BedGraph files. The raw sequence files and the processed deconvolution BedGraphs have been submitted to ArrayExpress under accession E-MTAB-10433 and E-MTAB-11385.

DNase I accessibility mapping (DNase-Seq)

DNase-Seq analysis of MEFs was carried out following the published protocol [50], the only modifications being adjustment of digestion level to obtain the required fragmentation. DNA samples were quality controlled by qPCR before size selection following the published protocol.

Cell Proliferation assay

Taf1b^{fl/fl}/ER-Cre^{+/+}/p53^{-/-} and isogenic control *Taf1b^{wt/wt}*/ER-Cre^{+/+}/p53^{-/-} MEFs were continuously cultured for more than a week prior to assay. Cells were plated at ~500 per well in 96-well plates, treated as indicated above with 50 ng.ml^{-1} 4-HT for just 4h and cultured for a further six days. At each timepoint, duplicate wells were treated with Hoechst 3342 (Invitrogen, Thermo Fisher Scientific) for 45min. Images were acquired using Cytation5 (Cell Imaging Multi-Mode Reader by BioTek) and cell counts for each clone were determined using the Gen5 software. Cells from three *ubf^{E210K/E210K}* MEF clones (#1, #2, #3) and two isogenic *ubf^{wt/wt}* MEF clones (#3, #4) were cultured and cell proliferation analyzed in the same way.

Viability assay

Cell Viability was assayed using the colorimetric MTT kit for cell survival and proliferation (Millipore, sigma) and the Trypan Bleu exclusion test (Gibco, Thermo Fisher Scientific). *Taf1b^{fl/fl}*/ER-Cre^{+/+}/p53^{-/-} and isogenic control *Taf1b^{wt/wt}*/ER-Cre^{+/+}/p53^{-/-} MEFs were cultured in 96-well plates at a density of 1 to 6 x 10^3 cells/well and were treated with 4-HT as described above. The assay was performed at different time points post 4-HT by adding MTT to culture wells following the manufacturer's instructions and incubating for 5h at 37°C. MTT absorbance readings were then normalized to cell numbers. For the Trypan Bleu exclusion test quadruplicate cell cultures cells were stained and manually counted at different time points after 4-HT treatment using a hemacytometer.

Total RNA Extraction and quantification

Cells were trypsinized, counted and total RNA was recovered from 3x10^6 cells using 1 ml of Trizol (Invitrogen, Thermo Fisher Scientific) according to the manufacturer's protocol. RNA yields were determined using Qubit RNA BR (Invitrogen, Thermo Fisher Scientific).

Supporting information

S1 Text. The TAF1B gene is essential for mouse development beyond early blastula.
(DOCX)

S1 Table. Numbers and genotypes of embryos and pups derived from mating of *Taf1b*+/- mice.
(DOCX)

S1 Fig. Construction and phenotypic effects of the mutant *Taf1b* alleles. A) Organisation of the first 8 exons of the mouse *Taf1b* gene (*Taf1bwt)*, and the "flox-neo" insertion, "floxed" and alleles indicating the position of inserted FRT and Lox sites and the inactivated (*Taf1bΔ*) allele after Lox site recombination to delete exons 4 and 5. The positions of genotyping primers A to D are also indicated. B) Examples of mouse PCR genotyping. C) alignment of the N-terminal sequence of wild type TAF1B with the predicted residual TAF1B peptide encoded by the *Taf1bΔ* allele. D) Typical images of mouse embryos at 3.5 dpc derived from *Taf1b*$^{wt/Δ}$ mouse crosses. The corresponding genotypes and numbers of embryos in each class are indicated, see also S1 Table. E) Embryo phenotyping using primers A, B and D shown in panel A.
(TIF)

S2 Fig. Time course of rRNA synthesis after *Taf1b* inactivation in MEFs. Conditional *Taf1b*$^{fl/fl}$/p53$^{-/-}$/ERcre$^{+/+}$ and control *Taf1b*$^{wt/wt}$/p53$^{-/-}$/ERcre$^{+/+}$ MEFs were treated with 50 nM 4-hydroxy-tamoxifen (4-HT) for 4h (4-HT pulse) before removing 4-HT by a change of the culture medium. 47S pre-rRNA synthesis was then determined by [3H]-uridine RNA meta-bolic labelling. The panels show the parallel time course analyses post 4-HT treatment of; A) *Taf1b* inactivation, B) [3H]-rRNA labelling versus steady state 28S rRNA, and C) TAF1B pro-tein depletion. The 47S pre-rRNA and rRNA processing products are indicated in B. The "*" in C indicates a non-specific antibody interaction that serves as a loading control.
(TIF)

S3 Fig. Time course of parallel *in situ* immunofluorescence labelling of RPI, UBTF and Fibrillarin (FBL) in *Taf1b*$^{fl/fl}$/p53$^{-/-}$/ERcre$^{+/+}$ MEFs after 4-HT induction of *Taf1b* inactiva-tion. Single confocal image planes are shown as for each factor and as merged overlays each with DAPI staining of DNA.
(TIF)

S4 Fig. A) Trypan blue exclusion and B) MTT conversion assays for cell viability, and C) cell proliferation at different time points post 4-HT or mock treatment were determined in qua-druplicate for Taf1bfl/fl/p53-/-/ERcre+/+ and Taf1bwt/wt/p53-/-/ERcre+/+MEFs.
(TIF)

S5 Fig. ChIP-qPCR analysis of TAF1B, TAF1C, TBP and UBTF occupancy at sites across the rDNA of *Taf1b*$^{fl/fl}$/ERcre$^{+/+}$ mESCs before and after *Taf1b* inactivation by 4-HT treat-ment. The data for TAF1B and UBTF derive from 3 ChIP biological replicas and for TAF1C and TBP a single ChIP each analyzed by qPCR in triplicate.
(TIF)

S6 Fig. DNase accessibility (DNase-Seq) mapping of the rDNA repeat unit in TAF1B con-ditional MEFs (*Taf1b*$^{fl/fl}$/p53$^{-/-}$/ERcre$^{+/+}$) before and after induction of TAF1B loss. A) Mapping of raw sequencing data obtained from nuclei treated with increasing concentrations of DNase I (0 to 150 units) across the rDNA, see Materials and Methods. B) Same data sets were also mapped across the active euchromatic GAPDH and inactive heterochromatic Tub8a genes. "Q-PCR amplicons" indicate the Gapdh and Tub8a gene regions used to select appro-priate DNaseI digestion to discriminate hyper-accessible gene regions.
(TIF)

S7 Fig. TAF1B depletion in condition *Taf1b^{fl/fl}/ERcre^{+/+}* mESCs also induces depletion of UBTF from both Spacer and 47S promoters but not from the adjacent enhancer repeats or from the 47S gene body. A) DChIP-Seq analysis of TAF1B and UBTF occupancy across the rDNA repeat before (TAF1B+) and 3 days after *Taf1b* inactivation (TAF1B-). ΔUBTF indicates the difference map of UBTF occupancy after TAF1B depletion minus the occupancy before TAF1B depletion. B) Magnified view of the DChIP mapping in A showing detail over the promoter and enhancer regions. C) Comparative Western blots of UBTF from TAF1B conditional MEFs (*Taf1b^{fl/fl}/p53^{-/-}/ERcre^{+/+}*) and mESCs (*Taf1b^{fl/fl}/ERcre^{+/+}*), showing the presence of both UBTF1 and UBTF2 variants are expressed in the MEFs but only the UBTF1 variant is expressed in the mESCs.
(TIF)

S8 Fig. Analysis of DChIP profiles by Gaussian curve fit before and after *Taf1b* inactivation. A) and B) respectively show examples of UBTF and TAF1B enrichment profiles over the 47S promoter region before and after *Taf1b* inactivation are shown, (dark blue line), and the best Gaussian peak fits to these profiles (dashed red line). In the case of TAF1B the profile closely followed a single Gaussian peak from which both the position and relative occupancy were determined. Since UBTF was present not only at the promoter but also over the adjacent regions, curve fits were made using three Gaussians peaks, and the central one used to estimate relative occupancy.
(TIF)

S9 Fig. Mapping of UBTF1 and -2 variants across the rDNA unit. A) Expression of exogenous 3xFlag-UBTF1 and UBTF2 in NIH3T3 MEFs. Total cell protein extracts from Mock, 3xFlag-UBTF1 or 3xFlag-UBTF2 transfected cells were analyzed by Western blot using either anti-Flag (αFlag), left hand panel) or anti-UBTF (UBTF) antibodies to detect total UBTF (right hand panel). B) DChIP-Seq mapping profiles of the exogenously 3xFlag-UBTF1 or 3xFlag-UBTF2 (Flag) expressed in NIH3T3 MEFs (as in Fig 4) and the total endogenous UBTF profiles ChIPped from the same chromatin preparations.
(TIF)

S10 Fig. Analysis of DChIP profiles for Flag-UBTF1 and -UBTF2 by Gaussian curve fit. A) and B) Show examples of Flag-UBTF1 and Flag-UBTF2 enrichment profiles respectively over the 47S and Spacer promoter regions, (dark blue line). The best Gaussian peak fits to these profiles are shown (dashed red line), as are the individual Gaussian peaks used to estimate relative promoter occupancy. Since UBTF was present not only at each promoter but also over adjacent regions, curve fits were made using three, or in the case of the Spacer promoter four, Gaussians peaks.
(TIF)

S11 Fig. A) expression of 3Flag-UBTF1 and 2 and endogenous UBTF in conditional MEFs before and after UBTF deletion as revealed by Western blot using anti-Flag and anti-UBTF antibodies. B) D-ChIP mapping of UBTF forms across the full rDNA transcription unit as in Fig 5. As reference, the upper track shows mapping of the endogenous UBTF shown in Fig 3.
(TIF)

S12 Fig. Analysis of DChIP profiles for Flag-UBTF1 and -UBTF2 in UBTF conditional MEFs by Gaussian curve fit, as in S10 Fig. The area of Flag-UBTF1 or 2 peaks at each promoter was normalized to the summed area adjacent Flag-UBTF peaks to allow comparisons between profiles. Gaussian peak positions were maintained constant between analyses.
(TIF)

S13 Fig. Structure prediction for wild type and E210K mutant HMGbox2 of UBTF1. A) Sequence alignment of the HMGboxes 1 and 2 of human and mouse UBTF1 (NM_014233–2, NP_035681) with HMGbox2 of Xenopus laevis UBTF1a (CAA42523.1) and the HMGbox of SOX2 (P48431). The positions of the predicted DNA intercalating residue, the E210K mutation and the adjacent conserved basic DNA contacting residue are indicated as are the positions of the α-helical segments. B) Comparative molecular modelling of UBTF HMGbox2 using as templates the structures 1k99 (human UBTF HMGbox1) and 6hb4 (human mitochondrial transcription factor A, TFAM). The two predicated structures were generated by SWISS-MODEL [51] and are shown individually and as an aligned overlay generated in ChimeraX-1.1.1 [52]. Comparison of these structures using the Matchmaker routine in ChimeraX-1.1.1 revealed an RMSD of 1.215 Å over 41 of 72 alpha-carbons, including those of helix 1 affected by the E210K mutation. C) The predicted positions and orientations of the E210 and K210 residues within the HMGbox2 of UBTF1 are shown relative to the adjacent conserved basic residue at position 211, which is a lysine in UBTF/UBTF. The likely other DNA minor groove contacting residues K198 and K200 are also shown. D) The predicted surface electrostatic potential of the wild type, left, and the mutant, right, HMGbox2. Blue indicates a positive and red a negative potential. Position of changes in surface potential due to the E210K mutation are enclosed by an ellipse.
(TIF)

S14 Fig. A) RNA metabolic pulse labelling to reveal 47S pre-rRNA synthesis and processing products in $Ubtf^{E210k/E210K}$ knock-in and wild type $Ubtf^{wt/wt}$ MEFs. Gel fractionation of RNA after increasing labelling times is shown for two individual (numbered) MEF isolates. B) DNA base sequence of the differentially spliced region of mouse *Ubtf* gene showing coding exon 6, the differentially spliced coding exon 7 and coding exon 8 in black and the intervening introns in red (taken from GRCm38:11:102303960:102320342). The position of the G>A gene mutation, the cause of the E210K change in the UBTF protein, is indicated as are the potential splice branch sites in the intervening introns that most closely fit the yTnAy consensus [53].
(TIF)

S15 Fig. Indirect immunofluorescence labelling of RPI, UBTF and Fibrillarin (FBL) in mutant $Ubtf^{E210K/E210K}$ knock-in and matched wild type $Ubtf^{wt/wt}$ MEFs reveals no overt changes in nucleolar structure or organisation associated with the UBTF mutation.
(TIF)

S16 Fig. DChIP-Seq difference map of UBTF occupancy ($Ubtf^{E210K/E210K}$ - wild type MEFs) as in Fig 7D but here shown across the full rDNA repeat. TAF1B and UBTF mapping are also shown for reference.
(TIF)

S1 Data. All numerical data underlying graphs and summary statistics in each illustration are presented here in spreadsheet form.
(XLSX)

Acknowledgments

Author Contributions

Conceptualization: Tom Moss.

Data curation: Tom Moss.

Formal analysis: Michel G. Tremblay, Dany S. Sibai, Tom Moss.

Funding acquisition: Roderick T. Hori, Mohammad Moshahid Khan, Mark S. LeDoux, Tom Moss.

Investigation: Michel G. Tremblay, Dany S. Sibai, Melissa Valère, Jean-Clément Mars, Frédéric Lessard, Victor Y. Stefanovsky, Tom Moss.

Methodology: Michel G. Tremblay, Frédéric Lessard, Roderick T. Hori, Mohammad Moshahid Khan, Tom Moss.

Project administration: Tom Moss.

Resources: Michel G. Tremblay, Mohammad Moshahid Khan.

Supervision: Michel G. Tremblay, Frédéric Lessard, Tom Moss.

Validation: Michel G. Tremblay, Tom Moss.

Visualization: Michel G. Tremblay, Tom Moss.

Writing – original draft: Tom Moss.

Writing – review & editing: Michel G. Tremblay, Frédéric Lessard, Mark S. LeDoux, Tom Moss.

References

1. Hannan KM, Sanij E, Rothblum LI, Hannan RD, Pearson RB. Dysregulation of RNA polymerase I transcription during disease. Biochim Biophys Acta. 2013; 1829(3–4):342–60. https://doi.org/10.1016/j.bbagrm.2012.10.014 PMID: 23153826

2. Long PA, Theis JL, Shih YH, Maleszewski JJ, Abell Aleff PC, Evans JM, et al. Recessive TAF1A mutations reveal ribosomopathy in siblings with end-stage pediatric dilated cardiomyopathy. Hum Mol Genet. 2017; 26(15):2874–81. https://doi.org/10.1093/hmg/ddx169 PMID: 28472305

3. Dauwerse JG, Dixon J, Seland S, Ruivenkamp CA, van Haeringen A, Hoefsloot LH, et al. Mutations in genes encoding subunits of RNA polymerases I and III cause Treacher Collins syndrome. Nat Genet. 2011; 43(1):20–2. https://doi.org/10.1038/ng.724 PMID: 21131976

4. Phan T, Khalid F, Iben S. Nucleolar and Ribosomal Dysfunction-A Common Pathomechanism in Childhood Progerias? Cells. 2019; 8(6). https://doi.org/10.3390/cells8060534 PMID: 31167386

5. Kara B, Koroglu C, Peltonen K, Steinberg RC, Maras Genc H, Holtta-Vuori M, et al. Severe neurodegenerative disease in brothers with homozygous mutation in POLR1A. European journal of human genetics: EJHG. 2017; 25(3):315–23. https://doi.org/10.1038/ejhg.2016.183 PMID: 28051070

6. Toro C, Hori RT, Malicdan MCV, Tifft CJ, Goldstein A, Gahl WA, et al. A recurrent de novo missense mutation in UBTF causes developmental neuroregression. Hum Mol Genet. 2018; 27(4):691–705. https://doi.org/10.1093/hmg/ddx435 PMID: 29300972

7. Hori Y, Shimamoto A, Kobayashi T. The human ribosomal DNA array is composed of highly homogenized tandem clusters. Genome Res. 2021. https://doi.org/10.1101/gr.275838.121 PMID: 34407983

8. Moss T, Mitchelson K, de Winter R. The promotion of ribosomal transcription in eukaryotes. Oxf Surv Eukaryot Genes. 1985; 2:207–50. PMID: 3916926

9. Moss T, Langlois F, Gagnon-Kugler T, Stefanovsky V. A housekeeper with power of attorney: the rRNA genes in ribosome biogenesis. Cell Mol Life Sci. 2007; 64(1):29–49. https://doi.org/10.1007/s00018-006-6278-1 PMID: 17171232

10. Mars JC, Sabourin-Felix M, Tremblay MG, Moss T. A Deconvolution Protocol for ChIP-Seq Reveals Analogous Enhancer Structures on the Mouse and Human Ribosomal RNA Genes. G3 (Bethesda). 2018; 8(1):303–14. https://doi.org/10.1534/g3.117.300225 PMID: 29158335

11. Ferreira R, Schneekloth JS Jr, Panov KI, Hannan KM, Hannan RD. Targeting the RNA Polymerase I Transcription for Cancer Therapy Comes of Age. Cells. 2020; 9(2). https://doi.org/10.3390/cells9020266 PMID: 31973211

12. Edvardson S, Nicolae CM, Agrawal PB, Mignot C, Payne K, Prasad AN, et al. Heterozygous De Novo UBTF Gain-of-Function Variant Is Associated with Neurodegeneration in Childhood. Am J Hum Genet. 2017; 101(2):267–73. https://doi.org/10.1016/j.ajhg.2017.07.002 PMID: 28777933

13. Sedlackova L, Lassuthova P, Sterbova K, Haberlova J, Vyhnalkova E, Neupauerova J, et al. UBTF Mutation Causes Complex Phenotype of Neurodegeneration and Severe Epilepsy in Childhood. Neuropediatrics. 2019; 50(1):57–60. https://doi.org/10.1055/s-0038-1676288 PMID: 30517966

14. Bastos F, Quinodoz M, Addor MC, Royer-Bertrand B, Fodstad H, Rivolta C, et al. Childhood neurodegeneration associated with a specific UBTF variant: a new case report and review of the literature. BMC Neurol. 2020; 20(1):17. https://doi.org/10.1186/s12883-019-1586-x PMID: 31931739

15. Grummt I. Life on a planet of its own: regulation of RNA polymerase I transcription in the nucleolus. Genes Dev. 2003; 17(14):1691–702. https://doi.org/10.1101/gad.1098503R PMID: 12865296

16. Russell J, Zomerdijk JC. RNA-polymerase-I-directed rDNA transcription, life and works. Trends Biochem Sci. 2005; 30(2):87–96. https://doi.org/10.1016/j.tibs.2004.12.008 PMID: 15691654

17. Rudloff U, Eberhard D, Tora L, Stunnenberg H, Grummt I. TBP-associated factors interact with DNA and govern species specificity of RNA polymerase I transcription. EMBO J. 1994; 13:2611–6. PMID: 8013460

18. Heix J, Zomerdijk JCBM, Ravanpay A, Tjian R, Grummt I. Cloning of murine RNA polymerase I-specific TAF factors: Conserved interactions between the subunits of the species-specific transcription initiation factor TIF-IB/SL1. ProcNatlAcadSciUSA. 1997; 94:1733–8. https://doi.org/10.1073/pnas.94.5.1733 PMID: 9050847

19. Murano K, Okuwaki M, Momose F, Kumakura M, Ueshima S, Newbold RF, et al. Reconstitution of human rRNA gene transcription in mouse cells by a complete SL1 complex. J Cell Sci. 2014; 127(Pt 15):3309–19. https://doi.org/10.1242/jcs.146787 PMID: 24928901

20. Bell SP, Pikaard CS, Reeder RH, Tjian R. Molecular mechanisms governing species-specific transcription of ribosomal RNA. Cell. 1989; 59:489–97. https://doi.org/10.1016/0092-8674(89)90032-9 PMID: 2805069

21. Cairns C, McStay B. HMG box 4 is the principal determinant of species specificity in the RNA polymerase I transcription factor UBF. Nucleic Acids Res. 1995; 23:4583–90. https://doi.org/10.1093/nar/23.22.4583 PMID: 8524646

22. Bell SP, Jantzen H-M, Tjian R. Assembly of alternative multiprotein complexes directs rRNA promoter selectivity. Genes Dev. 1990; 4:943–54. https://doi.org/10.1101/gad.4.6.943 PMID: 2384215

23. Friedrich JK, Panov KI, Cabart P, Russell J, Zomerdijk JC. TBP-TAF complex SL1 directs RNA polymerase I pre-initiation complex formation and stabilizes upstream binding factor at the rDNA promoter. J Biol Chem. 2005; 280(33):29551–8. https://doi.org/10.1074/jbc.M501595200 PMID: 15970593

24. Stefanovsky VY, Bazett-Jones DP, Pelletier G, Moss T. The DNA supercoiling architecture induced by the transcription factor xUBF requires three of its five HMG-boxes. Nucleic Acids Res. 1996; 24:3208–15. https://doi.org/10.1093/nar/24.16.3208 PMID: 8774902

25. Herdman C, Mars JC, Stefanovsky VY, Tremblay MG, Sabourin-Felix M, Lindsay H, et al. A Unique Enhancer Boundary Complex on the Mouse Ribosomal RNA Genes persists after loss of RRN3 or UBF and the Inactivation of RNA Polymerase I Transcription. PLoS Genetics. 2017; 13(7):e1006899. https://doi.org/10.1371/journal.pgen.1006899 PMID: 28715449

26. Moss T, Mars JC, Tremblay MG, Sabourin-Felix M. The chromatin landscape of the ribosomal RNA genes in mouse and human. Chromosome Res. 2019; 27(1–2):31–40. https://doi.org/10.1007/s10577-018-09603-9 PMID: 30617621

27. Hamdane N, Stefanovsky VY, Tremblay MG, Nemeth A, Paquet E, Lessard F, et al. Conditional inactivation of Upstream Binding Factor reveals its epigenetic functions and the existence of a somatic nucleolar precursor body. PLoS Genetics. 2014; 10(8):e1004505. https://doi.org/10.1371/journal.pgen.1004505 PMID: 25121932

28. Bryja V, Bonilla S, Arenas E. Derivation of mouse embryonic stem cells. Nat Protoc. 2006; 1(4):2082–7. https://doi.org/10.1038/nprot.2006.355 PMID: 17487198

29. Mars JC, Tremblay MG, Valere M, Sibai DS, Sabourin-Felix M, Lessard F, et al. The chemotherapeutic agent CX-5461 irreversibly blocks RNA polymerase I initiation and promoter release to cause nucleolar disruption, DNA damage and cell inviability. NAR Cancer. 2020; 2(4):zcaa032. https://doi.org/10.1093/narcan/zcaa032 PMID: 33196044

30. Wittner M, Hamperl S, Stockl U, Seufert W, Tschochner H, Milkereit P, et al. Establishment and maintenance of alternative chromatin states at a multicopy gene locus. Cell. 2011; 145(4):543–54. https://doi.org/10.1016/j.cell.2011.03.051 PMID: 21565613

31. O'Mahony DJ, Smith SD, Xie W, Rothblum LI. Analysis of the phosphorylation, DNA-binding and dimerization properties of the RNA polymerase I transcription factors UBF1 and UBF2. Nucleic Acids Res. 1992; 20:1301–8. https://doi.org/10.1093/nar/20.6.1301 PMID: 1561086

32. Bell SP, Learned RM, Jantzen HM, Tjian R. Functional cooperativity between transcription factors UBF1 and SL1 mediates human ribosomal RNA synthesis. Science. 1988; 241:1192–7. https://doi.org/10.1126/science.3413483 PMID: 3413483

33. Sanij E, Poortinga G, Sharkey K, Hung S, Holloway TP, Quin J, et al. UBF levels determine the number of active ribosomal RNA genes in mammals. J Cell Biol. 2008; 183(7):1259–74. https://doi.org/10.1083/jcb.200805146 PMID: 19103806

34. Tuan JC, Zhai WG, Comai L. Recruitment of TATA-binding protein-TAF, complex SL1 to the human ribosomal DNA promoter is mediated by the carboxy-terminal activation domain of upstream binding factor (UBF) and is regulated by UBF phosphorylation. MolCell Biol. 1999; 19:2872–9.

35. Jantzen HM, Chow AM, King DS, Tjian R. Multiple domains of the RNA polymerase I activator hUBF interact with the TATA-binding protein complex hSL1 to mediate transcription. Genes Dev. 1992; 6:1950–63. https://doi.org/10.1101/gad.6.10.1950 PMID: 1398072

36. Kwon H, Green MR. The RNA polymerase I transcription factor, upstream binding factor, interacts directly with the TATA box-binding protein. Journal of Biological Chemistry. 1994; 269:30140–6. PMID: 7982918

37. Stefanovsky VY, Pelletier G, Bazett-Jones DP, Crane-Robinson C, Moss T. DNA looping in the RNA polymerase I enhancesome is the result of non-cooperative in-phase bending by two UBF molecules. Nucleic Acids Res. 2001; 29(15):3241–7. https://doi.org/10.1093/nar/29.15.3241 PMID: 11470882

38. Stefanovsky VY, Moss T. The splice variants of UBF differentially regulate RNA polymerase I transcription elongation in response to ERK phosphorylation. Nucleic Acids Res. 2008; 36(15):5093–101. https://doi.org/10.1093/nar/gkn484 PMID: 18676449

39. Giroux S, Tremblay M, Bernard D, Cadrin-Girard JF, Aubry S, Larouche L, et al. Embryonic death of Mek1-deficient mice reveals a role for this kinase in angiogenesis in the labyrinthine region of the placenta. Curr Biol. 1999; 9:369–72. https://doi.org/10.1016/s0960-9822(99)80164-x PMID: 10209122

40. Conner DA. Mouse embryonic stem (ES) cell isolation. Curr Protoc Mol Biol. 2001;Chapter 23:Unit 23.4. https://doi.org/10.1002/0471142727.mb2304s52 PMID: 18265205

41. Ying QL, Wray J, Nichols J, Batlle-Morera L, Doble B, Woodgett J, et al. The ground state of embryonic stem cell self-renewal. Nature. 2008; 453(7194):519–23. https://doi.org/10.1038/nature06968 PMID: 18497825

42. Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. 1970; 227:680–5. https://doi.org/10.1038/227680a0 PMID: 5432063

43. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 2012; 9(7):671–5. https://doi.org/10.1038/nmeth.2089 PMID: 22930834

44. Stefanovsky VY, Moss T. Metabolic Labeling in the Study of Mammalian Ribosomal RNA Synthesis. Methods Mol Biol. 2016; 1455:133–45. https://doi.org/10.1007/978-1-4939-3792-9_11 PMID: 27576716

45. Stefanovsky VY, Pelletier G, Hannan R, Gagnon-Kugler T, Rothblum LI, Moss T. An immediate response of ribosomal transcription to growth factor stimulation in mammals is mediated by ERK phosphorylation of UBF. Mol Cell. 2001; 8(5):1063–73. https://doi.org/10.1016/s1097-2765(01)00384-7 PMID: 11741541

46. Mansour FH, Pestov DG. Separation of long RNA by agarose-formaldehyde gel electrophoresis. Anal Biochem. 2013; 441(1):18–20. https://doi.org/10.1016/j.ab.2013.06.008 PMID: 23800830

47. Conconi A, Widmer RM, Koller T, Sogo JM. Two different chromatin structures coexist in ribosomal RNA genes throughout the cell cycle. Cell. 1989; 57:753–61. https://doi.org/10.1016/0092-8674(89)90790-3 PMID: 2720786

48. Stefanovsky VY, Moss T. Regulation of rRNA synthesis in human and mouse cells is not determined by changes in active gene count. Cell Cycle. 2006; 5(7):735–9. https://doi.org/10.4161/cc.5.7.2633 PMID: 16582637

49. Arrigoni L, Richter AS, Betancourt E, Bruder K, Diehl S, Manke T, et al. Standardizing chromatin research: a simple and universal method for ChIP-seq. Nucleic Acids Res. 2016; 44(7):e67. https://doi. org/10.1093/nar/gkv1495 PMID: 26704968

50. He HH, Meyer CA, Hu SS, Chen MW, Zang C, Liu Y, et al. Refined DNase-seq protocol and data analysis reveals intrinsic bias in transcription factor footprint identification. Nat Methods. 2014; 11(1):73–8. https://doi.org/10.1038/nmeth.2762 PMID: 24317252

51. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 2018; 46(W1):W296–w303. https://doi.org/10.1093/nar/gky427 PMID: 29788355

52. Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci. 2018; 27(1):14–25. https://doi.org/10. 1002/pro.3235 PMID: 28710774

53. Gao K, Masuda A, Matsuura T, Ohno K. Human branch point consensus sequence is yUnAy. Nucleic Acids Res. 2008; 36(7):2257–67. https://doi.org/10.1093/nar/gkn073 PMID: 18285363