frontiers in Genetics

ORIGINAL RESEARCH
published: 03 February 2021
doi: 10.3389/fgene.2020.627007

Human Ubiquitin-Specific Peptidase 18 Is Regulated by microRNAs *via* the 3'Untranslated Region, A Sequence Duplicated in Long Intergenic Non-coding RNA Genes Residing in chr22q11.21

*Erminia Rubino[1,2], Melania Cruciani[1], Nicolas Tchitchek[2,3], Anna Le Tortorec[4], Antoine D. Rolland[4], Önay Veli[1], Leslie Vallet[1], Giulia Gaggi[1], Frédérique Michel[1], Nathalie Dejucq-Rainsford[4] and Sandra Pellegrini[1]**

[1]Unit of Cytokine Signaling, Institut Pasteur, INSERM U1221, Paris, France, [2]École Doctorale Physiologie, Physiopathologie et Thérapeutique, ED394, Sorbonne Université, Paris, France, [3]i3 research unit, Hôpital Pitié-Salpêtrière-Sorbonne Université, Paris, France, [4]UMR_S1085, Institut de recherche en santé, environnement et travail (Irset), EHESP, Inserm, Univ Rennes, Rennes, France

Ubiquitin-specific peptidase 18 (USP18) acts as gatekeeper of type I interferon (IFN) responses by binding to the IFN receptor subunit IFNAR2 and preventing activation of the downstream JAK/STAT pathway. In any given cell type, the level of USP18 is a key determinant of the output of IFN-stimulated transcripts. How the baseline level of USP18 is finely tuned in different cell types remains ill defined. Here, we identified microRNAs (miRNAs) that efficiently target *USP18* through binding to the 3'untranslated region (3'UTR). Among these, three miRNAs are particularly enriched in circulating monocytes which exhibit low baseline *USP18*. Intriguingly, the *USP18* 3'UTR sequence is duplicated in human and chimpanzee genomes. In humans, four *USP18* 3'UTR copies were previously found to be embedded in long intergenic non-coding (linc) RNA genes residing in chr22q11.21 and known as *FAM247A-D*. Here, we further characterized their sequence and measured their expression profile in human tissues. Importantly, we describe an additional lincRNA bearing USP18 3'UTR (here *linc-UR-B1*) that is expressed only in testis. RNA-seq data analyses from testicular cell subsets revealed a positive correlation between *linc-UR-B1* and *USP18* expression in spermatocytes and spermatids. Overall, our findings uncover a set of miRNAs and lincRNAs, which may be part of a network evolved to fine-tune baseline USP18, particularly in cell types where IFN responsiveness needs to be tightly controlled.

Keywords: type I interferon, ubiquitin-specific peptidase, 3'untranslated region, microRNAs, long intergenic non-coding RNA, 22q11.2, Testis

OPEN ACCESS

Edited by:
Qiu-Ning Liu,
Yancheng Teachers University, China

Reviewed by:
Nicholas Delihas,
Stony Brook Medicine, United States
James Dominic Mills,
Amsterdam University Medical Center (UMC), Netherlands
Chunyan Fan,
Xi'an Technological University, China

***Correspondence:**
Sandra Pellegrini
sandra.pellegrini@pasteur.fr

Specialty section:
This article was submitted to RNA, a section of the journal Frontiers in Genetics

Received: *07 November 2020*
Accepted: *30 December 2020*
Published: *03 February 2021*

Citation:
Rubino E, Cruciani M, Tchitchek N, Le Tortorec A, Rolland AD, Veli Ö, Vallet L, Gaggi G, Michel F, Dejucq-Rainsford N and Pellegrini S (2021) Human Ubiquitin-Specific Peptidase 18 Is Regulated by microRNAs via the 3'Untranslated Region, A Sequence Duplicated in Long Intergenic Non-coding RNA Genes Residing in chr22q11.21. Front. Genet. 11:627007. doi: 10.3389/fgene.2020.627007

INTRODUCTION

Ubiquitin-specific peptidase 18 (USP18) is an interferon (IFN)-stimulated gene (ISG) exerting a specific and non-redundant role in the negative regulation of type I IFN (here, IFN) responses (Honke et al., 2016; Basters et al., 2018). USP18 is also an isopeptidase (de-ISGylase) that non-redundantly cleaves the ubiquitin-like ISG15 from protein conjugates (Malakhova et al., 2006; Francois-Newton et al., 2011; Arimoto et al., 2017). USP18 contains a conserved cysteine protease catalytic domain that is flanked by short N- and C-terminal regions providing key binding surfaces. Through the N-ter region (amino acids 36-51) and the C-ter region (amino acids 313-371, encoded by exons 9–11), human USP18 is recruited to the IFNAR2 receptor subunit at the plasma membrane. This interaction is reinforced by adjacent regions binding to STAT2 (Arimoto et al., 2017). The complex interferes with JAK1/TYK2 activation and attenuates STAT-mediated IFN-stimulated genes ISGs induction. The role of the C-ter region is supported by the dysregulated IFN response phenotype of the $Usp18^{Ity9/Ity9}$ mouse strain bearing a single homozygous mutation (Leu361Phe) in this region (Dauphinee et al., 2014). Deficiency of *USP18* causes a severe type I interferonopathy resulting in perinatal death with serious brain malformations due to spontaneous microglia activation, which most likely results from unrestrained response to constitutive IFNβ (Meuwissen et al., 2016). Indeed, high baseline USP18 maintains microglia quiescence and prevents sub-threshold activation (Goldmann et al., 2015; Schwabenland et al., 2019). In mouse models of VSV and LCMV infection, it was shown that the ability of CD169+ spleen macrophages and dendritic cells to present viral antigens and elicit an innate and adaptive immunity relies on the expression of *Usp18* (Honke et al., 2011, 2013). Altogether, these findings point to USP18 as a key determinant of cell responsiveness to IFN, including in the context of constitutive low IFNβ levels. We and others have shown that in humans, but not in mice, free ISG15 sustains the level of USP18, by preventing its proteasomal degradation by the S-phase kinase-associated protein 2 (SKP2; Zhang et al., 2015; Speer et al., 2016; Vuillier et al., 2019). Yet, little is known on how baseline USP18 levels are set in different cell lineages.

MicroRNAs (miRNAs) act as fine-tuners of gene expression and protein output. Each gene target can be regulated by numerous miRNAs, which typically bind to specific sites in 3' untranslated regions (3'UTRs) of messenger RNAs (mRNAs), influencing mRNA stability and translation. Hence, the amount of target varies from cell to cell depending on the identity and the expression level of each miRNA (Bartel, 2004). MiRNAs operate in different physiological and pathological processes (Ebert and Sharp, 2012) and also participate to the complex network that regulates immune cell development, function, and response to stimuli (Lindsay, 2008). As for many immune pathways, miRNAs have been shown to regulate the IFN system both at the level of production and response. Hence, some of the key players of the type I IFN signaling pathway are under the control of miRNAs (Forster et al., 2015), but no miRNA targeting USP18 has been described so far.

Here, we investigated whether miRNAs contribute to tune down baseline USP18 in a cell-context specific manner. We identified four miRNAs that negatively regulate USP18 through binding the 3'UTR. Several copies of the *USP18* 3'UTR are present in the human genome. These duplications reside on chr22q11.21 and are embedded in long intergenic non-coding (linc) RNA (lincRNA) genes that we found expressed in different human tissues. In particular, we describe a lincRNA bearing *USP18* 3'UTR with a remarkable cell type-specific expression in testicular germ cells. The positive correlation of the expression of this lincRNA with *USP18* raises the possibility that the *USP18* 3'UTR not only acts as a cis-regulatory sequence targeted by miRNAs, but also acts in *trans* when embedded in lincRNA molecules expressed in particular cell types.

MATERIALS AND METHODS

Computational Prediction of miRNAs Targeting the *USP18* 3'UTR

Bibiserv[1] and miRWalk 2.0[2] interfaces were used. The human *USP18* 3'UTR sequence [580 nucleotide (nt)] was downloaded from the UCSC genome browser (hg38; https://genome.ucsc.edu/) and used as input. In Bibiserv, RNA hybrid algorithm was used with the following parameters: energy threshold = −20, no G:U in the seed, helix constraint from nt 2 to 8 (7-mer seed), and approximate *p*-value for 3utr_human. In miRWalk, TargetScan, Pita, miRDB, and RNA22 algorithms were used with the following parameters: minimum seed length 7, position 2 (nt 2) as start position of the miRNA seed, and *p* value 0.05.

Monocyte Isolation

Peripheral blood mononuclear cells (PBMCs) were isolated from freshly collected buffy coats obtained from healthy blood donors (Établissement Francais du Sang, Paris, France; CPSL UNT-18/EFS/041) by density gradient centrifugation using lymphocyte separation medium (Eurobio, France, Les Ulis). Monocytes were purified by positive sorting using anti-CD14-conjugated magnetic microbeads (Miltenyi Biotec, Bergisch Gladsbach, Germany). The recovered cells were >98% $CD14^+$ as determined by flow cytometry with anti-CD14 Ab (Miltenyi Biotec). The CD14 negative fraction (here PBL) was also recovered for further analyses.

Testicular Cells Isolation

Normal human adult testes were obtained either after orchidectomy from prostate cancer patients who had not received any hormone therapy or at autopsy. The procedure was approved by Ethics Commitee Ouest V, Rennes, France (authorization DC-2016-2783) and the French National Agency for Biomedical Research (authorization PFS09-015). The presence of full spermatogenesis was assessed by transillumination of freshly dissected seminiferous tubules. Testis fragments were frozen

[1]https://bibiserv.cebitec.uni-bielefeld.de/rnahybrid
[2]http://zmf.umm.uni-heidelberg.de/apps/zmf/mirwalk2

and stored at −80°C until RNA extraction. Pachytene spermatocytes, round spermatids, Leydig cells, and peritubular cells were isolated and cultured according to previously described procedures (Rolland et al., 2019) before freezing at −80°C for RNA extraction. Primary Sertoli cells were purchased from Lonza (Walkersville, MD, United States) and cultured as previously described (Rolland et al., 2019).

RNA Extraction and RT-qPCR Quantification

Total RNA was extracted from Trizol lysates of freshly isolated monocytes or PBL, testis explants, and testicular populations, using miRNeasy mini kit (Qiagen, Germantown, MD, United States) and following the manufacturer's instructions. Total liver RNA was purchased from Thermo Fisher scientific. Quantification and purity were assessed by a Nanodrop spectrophotometer (Nanodrop2000, Thermo Fisher Scientific). For the analysis of *USP18*, *IRF7*, *OAS1*, *IFIT1*, *STAT2*, *FAM247A/C/D*, and *linc-UR-B1* levels, total RNA was reverse-transcribed using the high-capacity complementary DNA (cDNA) reverse Transcription (RT) kit (Applied Biosystems, Thermo Fisher Scientific). To measure the expression of *USP18*, *FAM247A/C/D*, and *linc-UR-B1* in human tissues, the cDNAs from Human Immune System MTC™ Panel, Human MTC™ Panel I, and Human MTC™ Panel II (Takara Bio, Mountain View, CA, United States) were used. Quantitative PCR (qPCR) assays were performed using the FastStart SYBR Green Master Mix (Roche) on a Step One Plus Real-Time PCR system (Applied Biosystems, Thermo Fisher Scientific). Transcript expression was normalized to the *18S* or to *ACTB* levels by using the equation $2^{-\Delta Ct}$. For miRNA expression, total RNA was reverse-transcribed using the TaqMan miRNA RT kit (Applied Biosystems, Thermo Fisher Scientific) and miRNA-specific primers (Applied Biosystems, Thermo Fisher Scientific) for hsa-*miR-191-5p*, hsa-*miR-24-3p*, and hsa-*miR-532-3p*. MiRNA expression levels were then analyzed using the appropriate TaqMan miRNA assay and TaqMan Universal Master Mix II (Applied Biosystems, Thermo Fisher Scientific), according to the manufacturer's instructions on a QuantStudio 3 Real-Time PCR system (Applied Biosystems, Thermo Fisher Scientific). The ubiquitously expressed U6b snRNA (*U6*) was quantified as above and used as endogenous control to normalize miRNA expression using the $2^{-\Delta Ct}$ formula. MiRNA qPCR-array in monocytes, testis fragments, and germ cells was performed using custom 96-well plates using the Taqman Advanced technology (Applied Biosystems, Thermo Fisher Scientific). Here, *miR-425-5p* was used as endogenous control.

Mimic Reverse Transfection

miRIDIAN microRNA mimics for human *miR-191-5p*, *miR-24-3p*, *miR-423-5p*, and *miR-532-3p* were purchased from Horizon Discovery Ltd. A negative control – miRIDIAN miRNA Mimic Negative Control #1 – (Horizon Discovery Ltd) was used to assess the specificity of the effect driven by the specific miRNA sequences. Mimic reverse transfection (50 nM) was performed in HeLa S3 cells cultured in DMEM supplemented with 10% heat inactivated FCS and antibiotics. Transfection was carried out by using Lipofectamine RNAiMAX (Thermo Fisher Scientific) according to the manufacturer's instructions. Briefly, the mimic-RNAiMAX complexes were prepared in Opti-MEM serum-free medium by mixing 50 nM of mimic with a dilution of 1:1,000 of lipofectamine RNAiMAX (calculated on the final volume of the culture). After 15 min at room temperature, the complexes were distributed in a p24- (for RNA isolation and luciferase assays) or a p60- (for protein extraction) well plate and 1×10^5 or 9×10^5 cells, respectively, and were plated on top of the complex. After 48 h of transfection, cells were harvested and lysed for RNA isolation, luciferase assay, or protein extraction.

Identification and Analysis of Duplicated *USP18* Sequences

To identify *USP18* copies, we performed BLAT[3] analysis using human *USP18* gene sequence. The sequence was downloaded from UCSC genome browser (hg38; https://genome.ucsc.edu/). We selected only sequences with more than 95% identity to *USP18* sequence and longer than 100 bp. To find whether the *USP18* copies overlapped genes, we looked at the coordinates of these repeats on the UCSC genome browser (hg38). The genes overlapping *USP18* copies were identified, and their genomic sequences and predicted mRNA sequences were downloaded. To measure the expression of *FAM247A,C,D* and *linc-UR-B1*, we designed specific primers, targeting the unique exon-exon11 junctions to avoid detection of *USP18* mRNA. Next, we performed qPCR analysis as described above. All qPCR products were verified by sequencing.

Analysis of Public Available Datasets

Data on miRNA expression, shown in **Figure 2** and **Supplementary Figure S2** were retrieved from the FANTOM5 dataset[4] and from (Allantaz et al., 2012; Cohort Roche, GEO accession: GSE28492). The FANTOM5 dataset was downloaded and analyzed using Qlucore Omics Explorer software. RNA-seq datasets covering a panel of 32 tissues[5] were analyzed, and the results are shown in **Supplementary Figure S5B**. PolyA+ RNA-seq track from testis provided by the ENCODE project (GEO accession: GSM2453457- GSM2453458) and shown in **Supplementary Figure S6A** were visualized using autoscale mode in IGV browser. The data of the single cell RNA-seq on testis used in **Figure 6** and **Supplementary Figure S7** were retrieved from (Guo et al., 2018). Reads were aligned on the hg38 assembly of the human genome and only unique reads were filtered.

Luciferase Reporter Assay

For cloning the *USP18* 3'UTR in the psiCHECK2 vector (Promega, Madison, WI, United States), the first 559 nt of

[3]https://genome.ucsc.edu/cgi-bin/hgBlat
[4]https://fantom.gsc.riken.jp/5/suppl/De_Rie_et_al_2017/vis_viewer/#/human
[5]https://www.ebi.ac.uk/gxa/experiments/E-MTAB-2836/Results

USP18 3'UTR was amplified from cDNA of HeLa S3 treated with IFN using the primers USP183UTR_XhoI_FW and USP183UTR_NotI_RV. Subsequently, the amplicon and psiCHECK2 vector were digested by Xho1 and Not1 and ligated. The resulting plasmid (psiCHECK2-*USP18* 3'UTR) was mutated (nt 2–6 of the seed-matched sequence) in the 3'UTR binding sites of *miR-24-3p*, *miR-191-5p*, *miR-4 23-5p*, and *miR-532-3p*. Mutations were generated in the psiCHECK2-*USP18* 3'UTR plasmid using QuikChange XL site-directed mutagenesis kit (Aligent Technologies, Santa Clara, CA, United States) and specific primers for each site (191-5p_BSmut_FW/RV, 24-3p_BSmut_FW/RV, 423-5p_BSmut_ FW/RV, and 532-3p_BSmut_FW/RV). All new plasmids were sequenced and named as *miR-191-5p*, *miR-24-3p*, *miR-423-5p*, and *miR-532-3p* BSmut. For the luciferase reporter assay, 1×10^5 HeLa S3 cells/well were reverse-transfected in a 24-well plate (see section on *Mimic reverse transfection*). After 24 h, plasmid transfection was carried out using FuGENE HD (Promega) according to the manufacturer's instructions. Three biological replicates were prepared for each condition. Cells were lysed 24 h after transfection and analyzed with Dual-Luciferase® Reporter Assay System (Promega) according to the manufacturer's instructions. Firefly luciferase was measured as normalizer.

Western Blot Analysis and Antibodies

Cells were lysed in modified RIPA buffer (50 mM Tris-HCl pH 8, 200 mM NaCl, 1% NP40, 0.5% DOC, 0.05% SDS, and 2 mM EDTA) with 100 mM PMSF and phosSTOP and a cocktail of antiproteases (Roche). A total of 40 µg proteins was separated by SDS-PAGE and analyzed by western blot. Membranes were cut horizontally according to molecular size markers, and stripes were incubated with different Abs. Immunoblots were analyzed with the ECL Western blotting Reagent (Pierce) or the more sensitive Western Lightning Chemiluminescence Reagent Plus (PerkinElmer), and bands were quantified with Fuji LAS-4000. For reprobing, blots were stripped in 0.2 M glycine (pH 2.5) for 20 min at rt. The following Abs were used: rabbit anti-USP18 (D4E7, Cell Signaling Technology, Beverly, MA, United States) and mouse anti-actin B (Sigma-Aldrich, St. Louis, MO, United States).

3' of cDNA Ends and PCR

3' of cDNA Ends (3'RACE) analysis was performed on testis and liver RNA using the 3'RACE System kit (Thermo Fisher Scientific) according to manufacturer's instructions. PCR was performed with Platinum SuperFi II DNA Polymerase (Thermo Fisher Scientific).

Primers

For the sequences of all primers used refer to **Supplementary Table S2**.

Sequences

The sequences of human *USP18* exon11, *FAM247A/C/D*, and a partial sequence of *linc-UR-B1* are available as Supplementary sequences.

RESULTS

The 3'UTR of Human USP18 Is Targeted by at Least Four miRNAs

Using the bioinformatic prediction programs miRWalk and RNAhybrid, we identified 27 miRNAs predicted to target with high score the 580 nt-long 3'UTR of the human *USP18* mRNA (**Figure 1A**). *USP18* is basally detected in most human tissues, albeit at variable levels (**Supplementary Figure S1**). We, therefore, analyzed the expression of the 27 miRNA candidates in the 90 cell types of the FANTOM5 dataset (de Rie et al., 2017). Fourteen miRNAs were found to be expressed at >50 counts per million mapped reads (CPM) in at least one cell type (**Figure 1A**). We next tested the functional impact of these 14 miRNAs on USP18. For this, we expressed in HeLa S3 cells, each miRNA as mimic and measured the level of *USP18* mRNA by qPCR. As shown in **Figure 1B**, seven miRNAs led to a reduction of endogenous *USP18*, four of them exhibiting a stronger effect. Consistently, this also resulted in lower USP18 protein levels, at least for those four miRNAs with stronger effects at the RNA level (**Figure 1C**).

In order to validate the binding of each miRNA to the predicted site on *USP18* 3'UTR (**Figures 1D,E**), this latter sequence was subcloned downstream of the *Renilla* luciferase gene in the psiCHECK2 vector and the impact of the four miRNAs on luciferase was measured in transfected HeLa S3 cells. As shown in **Figure 1F**, *miR-191-5p*, *miR-24-3p*, *miR-423-5p*, and *miR-532-3p*, expressed as mimics, significantly decreased luciferase activity with respect to the control mimic. Moreover, the effect of each mimic was abrogated when a 5 nt mutation was introduced in the seed-matched sequence of the predicted binding site on the 3'UTR (**Figure 1F**). Altogether these data demonstrate that the *USP18* 3'UTR sequence can be directly targeted by *miR-191-5p*, *miR-24-3p*, *miR-423-5p*, and *miR-532-3p*, with an effect on USP18 abundance.

To determine in which cell type(s) this post-transcriptional control may operate, we performed a principal component analysis (PCA) based on the expression of the seven active miRNAs identified above in the 90 cell types (**Figure 2A**). The analysis revealed a different distribution of immune vs. non-immune cells, with the first three PC vectors explaining 75% of the total variance. While the distribution of non-immune cells was homogenous, the distribution of immune cells was more scattered and mostly driven by *miR-191-5p*, *miR-24-3p*, *miR-532-3p*, and *miR-361-5p* (**Figure 2A**). These four miRNAs were expressed variably across the 90 cell types, but were enriched in immune cells (**Supplementary Figure S2A**). We therefore performed a second PCA based on expression of these four miRNAs exclusively in circulating PBMCs ($CD14^+$ monocytes, $CD8^+$ T cells, $CD4^+$ T cells, NK cells, and $CD19^+$ B cells). These miRNAs appeared to contribute to the distribution of the cell populations analyzed (**Figure 2B**). In particular, *miR-191-5p*, *miR-24-3p*, and *miR-532-3p* were found to be enriched in monocytes, which clustered apart from all lymphoid cells (**Supplementary Figure S2B**). To reinforce these data, we surveyed an independent dataset (Cohort Roche, GSE28492; Allantaz et al., 2012), which confirmed a higher

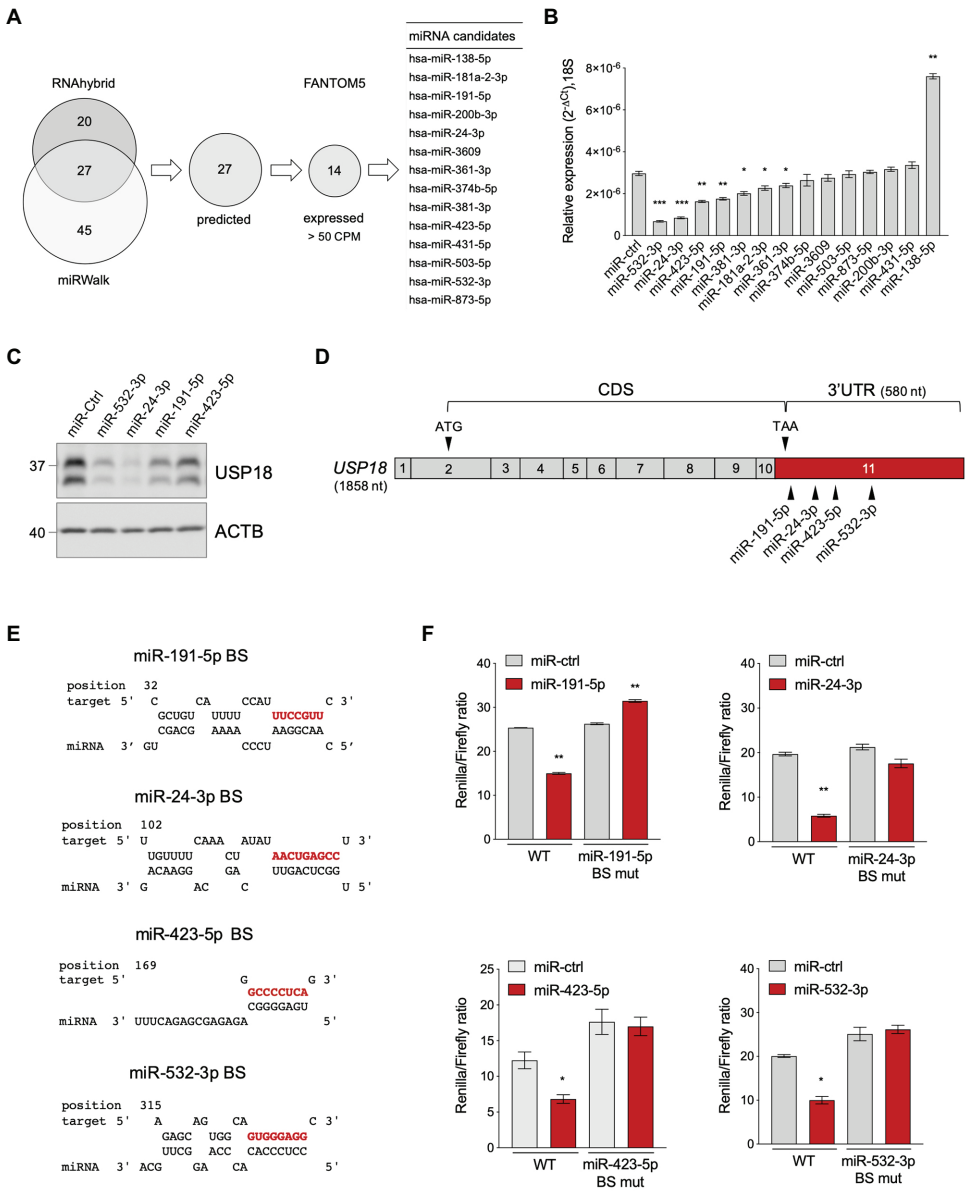

FIGURE 1 | Identification of USP18-targeting miRNAs. **(A)** Prediction of *USP18*-targeting miRNAs using RNA hybrid algorithm and mirWalk software (TargetScan, Pita, miRDB, and RNA22). The common high-scoring miRNAs (27) were further selected for expression (counts per million mapped reads, CPM > 50) in 90 human cell types (FANTOM5 dataset). The selected candidates (14) are listed. **(B)** The 14 selected miRNAs and a miR-control (miR-ctrl) were expressed as mimics in HeLa S3 cells and *USP18* mRNA was measured by qPCR ($n = 4$). Results (± SEM) shown as expression ($2^{-\Delta Ct}$) relative to *18S*, used as housekeeping gene. Raw-matched one-way ANOVA with Geisser-Greenhouse correction was performed (Dunnett's multiple comparison – compared to miR-ctrl), adjusted values shown as stars $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$. **(C)** The miR-ctrl and the indicated miRNAs were expressed as mimics in HeLa S3 cells and 48 h later endogenous USP18 was measured by western blot. Actin B (ACTB) as loading control. The data are representative of three experiments. **(D)** Schematic of the *USP18* mRNA. Numbered boxes correspond to the exons. CDS, coding sequences. Arrowheads, the start (ATG) and stop (TAA) codons and the binding sites for the four indicated miRNAs. Note that exon 11 (in red) contains 43 coding nt, the stop codon, and the 3'untranslated region (UTR). The sequences of *USP18 exon11* are provided as Supplementary Information. **(E)** Binding sites (BS) of *miR-191-5p*, *miR-24-3p*, *miR-423-5p*, and *miR-532-3p* in the *USP18* 3'UTR. Seed-matched sequences are in red. Pairing of miRNA-*USP18* sequences was obtained with the RNAhybrid algorithm. **(F)** HeLa S3 cells were transfected with the indicated miRNAs and 24 h later with the psiCHECK2-*USP18* 3'UTR reporter plasmid, WT, or mutated in the miRNA binding site (BS mut). Each mutant plasmid contains a 5 nt mutation (nt 2–6 of the seed-matched sequence shown in **E**). Luciferase activity was measured 24 h after plasmid transfection ($n = 4$) and expressed as Renilla/Firefly ratio (± SEM).

level of the three miRNAs in monocytes with respect to NK, T, and B cells (**Figure 2C**). Of note, *miR-191-5p* and *miR-24-3p* were more abundant than *miR-532-3p* in monocytes.

The enrichment of *miR-191-5p*, *miR-24-3p*, and *miR-532-3p* in monocytes prompted us to search for a correlation with *USP18*. For this, CD14+ monocytes and monocyte-depleted

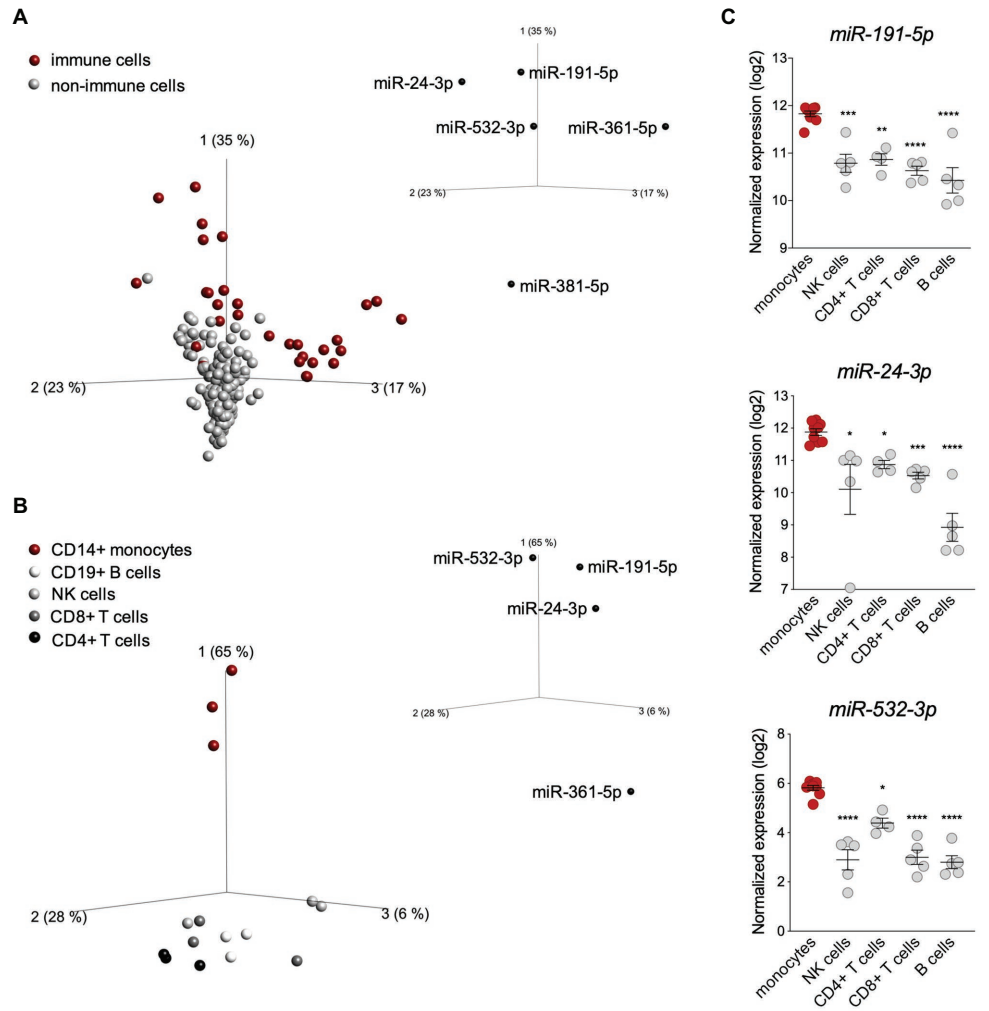

FIGURE 2 | *miR-191-5p*, *miR-24-3p*, and *miR-532-3p* are enriched in monocytes. **(A)** Principal component analysis (PCA) performed on 90 human cell types (FANTOM5 project dataset, three donors per cell population). The distribution is based on expression of the seven microRNAs (miRNAs) significantly targeting ubiquitin-specific peptidase 18 (*USP18*; see **Figure 1B**). Red and gray circles: immune and non-immune cells, respectively. The PCA plot shown captures 75% of the total variance within the selected data set (PCA1 35%, PCA2 23%, and PCA3 17%). value of (ANOVA FDR adjusted *p*-value) $q < 0.05$. The five *USP18*-targeting miRNA candidates significantly contributing to the data distribution are shown on the three PC axes of the plot. **(B)** PCA performed on circulating immune cell populations (FANTOM5 project data set). Analysis was restricted to the four *USP18*-targeting miRNAs enriched in immune cells. Color legend on the left. The PCA plot captures 97% of the total variance within the selected data set (PCA1 65%, PCA2 28%, and PCA3 6%). value of (ANOVA FDR adjusted p-value) $q < 0.05$. The four miRNAs significantly contributing to the data distribution are shown on the three principal component axes of the PCA plot. **(C)** Normalized expression (log2, ±SEM) of *miR-532-3p*, *miR-191-5p*, and *miR-24-3p* in circulating immune cell subsets. Data retrieved from (Allantaz et al., 2012), Cohort Roche, GSE28492. Ordinary one-way ANOVA (Dunnett's multiple comparison – compared to monocytes) was performed, adjusted *p*-values shown as stars $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

PBMCs (here PBL for peripheral blood lymphocytes) were freshly isolated from blood of four donors and levels of the three miRNAs and of *USP18* were measured by RT-qPCR. The three miRNAs were more expressed in monocytes than PBL (**Figure 3A**), while *USP18* exhibited an opposite profile (**Figure 3B**). In addition, a significant negative correlation between each miRNA and *USP18* was observed (**Figure 3C**). Interestingly, the opposite was observed for four other ISGs, i.e., *OAS1*, *IRF7*, *IFIT1*, and *STAT2*, which were found to be more abundant in monocytes than PBL, suggesting a specific profile for USP18 as compared to other ISGs (**Figure 3D**).

Altogether, these results suggest that *miR-191-5p*, *miR-24-3p*, and *miR-532-3p* contribute to restraining baseline *USP18* in human monocytes. Although not enriched in monocytes as compared to PBL, *miR-423-5p* may also contribute as it was relatively highly expressed among the studied miRNAs (**Supplementary Figure S3**).

Copies of USP18 3'UTR Embedded in lincRNA Genes

The results described above suggest that the level of human USP18 can be post-transcriptionally fine-tuned by four miRNAs

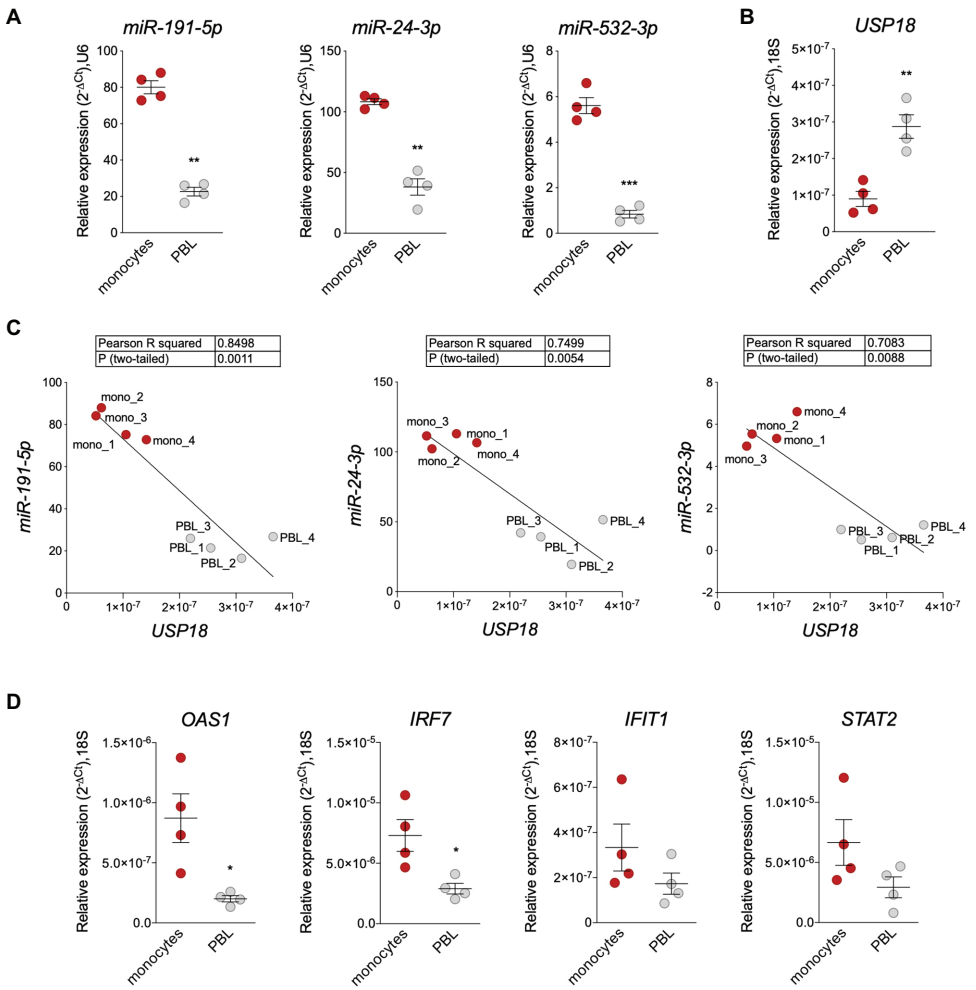

FIGURE 3 | *miR-191-5p*, *miR-24-3p*, and *miR-532-3p* negatively correlate with *USP18* in monocytes vs. PBL. **(A,B)** Expression of *miR-191-5p*, *miR-24-3p*, and *miR-532-3p* **(A)** and of *USP18* **(B)** in *ex vivo* monocytes and monocyte-depleted peripheral blood mononuclear cells (PBMCs; here PBL; four donors). Results (± SEM) shown as expression ($2^{-\Delta Ct}$) relative to *U6* or *18S*, used as housekeeping genes. Paired *t*-test values of *p* shown as stars $^{**}p < 0.01$, $^{***}p < 0.001$. **(C)** Negative correlation between *USP18* and *miR-191-5p*, *miR-24-3p*, and *miR-532-3p* in monocytes and PBL. Expression levels used to calculate the correlation are shown in **(A,B)**. Pearson correlation coefficients and *p*-values (two-tailed) are shown above. **(D)** Expression of *OAS1*, *IRF7*, *IFIT1*, and *STAT2* in monocytes and PBL (four donors). Results (± SEM) shown as expression ($2^{-\Delta Ct}$) relative to *18S*. Paired *t* test *p*-values are shown, when significant, as stars $^{*}p < 0.05$.

binding its 3'UTR. These miRNAs and their genes are conserved down to the mouse (**Supplementary Figure S4A**), but this is not the case for their binding sites on *USP18* 3'UTR. Some of these miRNA-*USP18* interactions (as for *miR-191-5p* and *miR-423-5p*) may be recent, while others (*miR-24-3p* and *miR-532-3p*) may be conserved throughout evolution, as suggested by the conservation of the seed-matched sequences in the *USP18* 3'UTR of other mammals (**Supplementary Figures S4B,C**). In searching for sequence conservation, we found several copies of the *USP18* 3'UTR in the human and chimpanzee genomes (**Table 1**), raising the possibility that transcripts other than the *bona fide* USP18 protein-coding mRNA contain the *USP18* 3'UTR and bind the same miRNAs.

The *USP18* gene and all copies of the 3'UTR sequence map on chr22q11.21, a highly complex region of the human genome that contains four segmental duplications or low-copy repeats (LCR22-A to D; **Figure 4A**; Morrow et al., 2018). The high sequence identity and copy number variation in the LCR22s have indeed hampered accurate sequencing and gene annotation in this region and, even in the most recent version of the human reference genome, LCR22s are rich in gaps and assembly errors (Demaerel et al., 2019). By performing a BLAT analysis of *USP18* gene sequences on the last hg38 assembly, we identified six copies of the 3'UTR in positive or negative strand orientation (**Figure 4A**; **Table 1**). Four copies – named A1–4 here – contain most of intron 10 and the entire exon 11 (i.e., the *USP18* 3'UTR) of the *USP18* gene. The other two copies, named B1 and B2, contain a small segment of intron 10 and the entire exon 11. All six copies reside in LCR-A and LCR-D (**Figures 4A,B**). Importantly, with the exception of B2, the other copies overlap each an annotated lincRNA gene (see list in **Table 2**). The

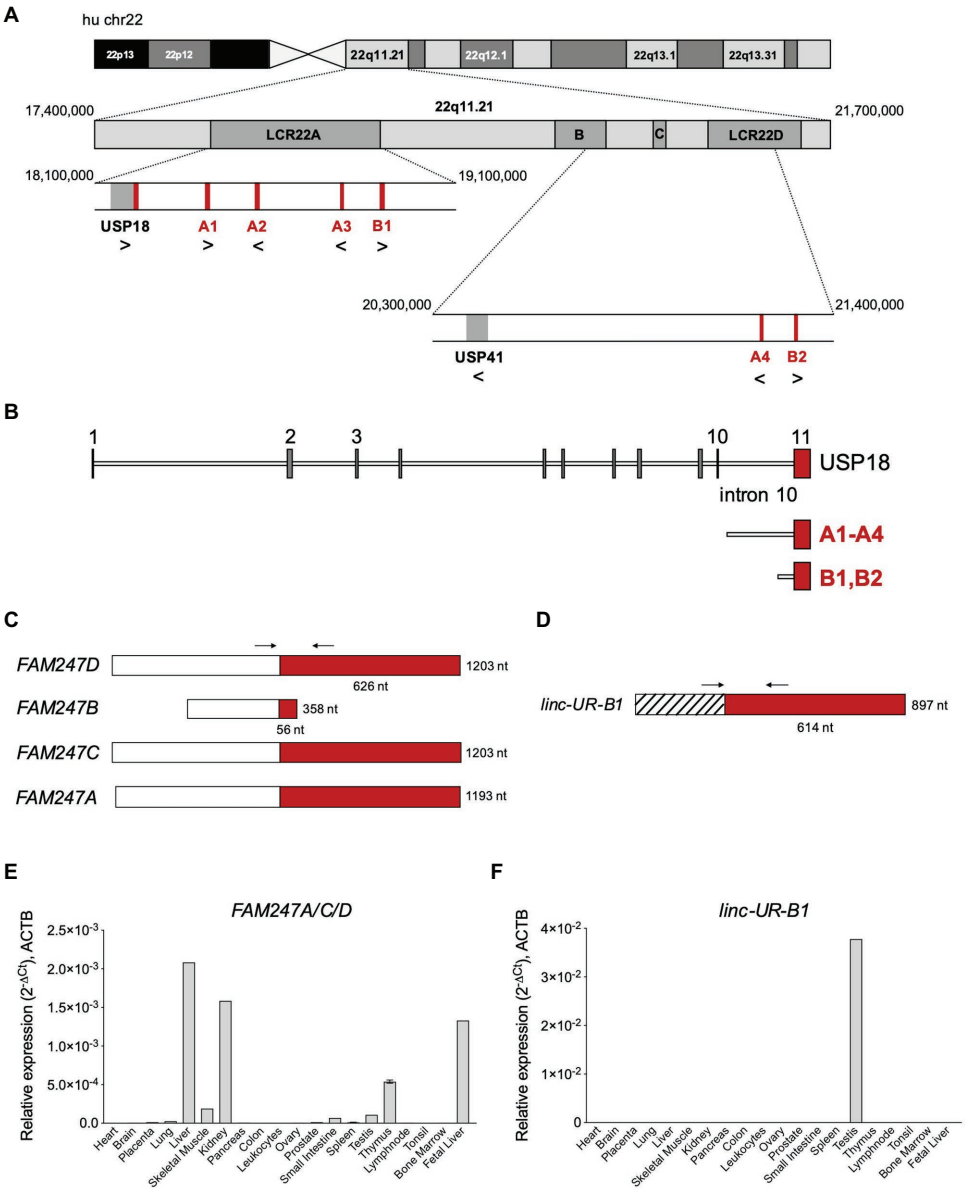

FIGURE 4 | Several *USP18* exon 11 copies are embedded in expressed long intergenic non-coding RNA (lincRNA) genes. **(A)** Schematic of human chr22 (about 50 Mb) and below the chr22q11.21 region with four LCR22s. The *bona fide USP18* gene resides at the boundary of LCR22A. The pseudogene *USP41*, located in LCR22B, contains most *USP18* gene sequences, i.e., from exon 3 to exon 10, but lacks the 5' and 3' UTRs. The six copies of *USP18* exon 11 are indicated in red. Genes and *USP18* copies are shown with their genomic orientation (> for + strand, < for – strand). **(B)** Top, intron-exon organization of the human *USP18* gene. Exons, gray boxes; introns, lines. Exon 11 is highlighted in red. Below are aligned the duplications found. Copies A1–4 contain most of *USP18* intron 10 and the entire exon 11. Copies B1 and B2 contain a smaller sequence of intron 10 and exon 11. **(C)** Map of *FAM247A-D* transcripts containing *USP18* exon 11 in red. The annotated exon upstream of exon 11 is represented as white box. Arrows indicate the primers (A_FW1 and A_RV1) designed across the junction and used to measure expression of *FAM247A/C/D* in **(E)**. **(D)** Map of *linc-UR-B1*. The annotated exon upstream of exon 11 is represented as striped box. Arrows indicate the primers (B_FW1 and B_RV1) designed across the junction and used to measure expression of *linc-UR-B1* in **(F)**. **(E)** Expression of *FAM247A/C/D* in 20 human tissues (pool of donors for each tissue; Human Immune System MTC™ Panel, Human MTC™ Panel I, and Human MTC™ Panel II). Results shown as expression ($2^{-\Delta Ct}$) relative to *ACTB*. SEM is shown for tissues present in two of the above panels. **(F)** Expression of *linc-UR-B1* in the same samples used in **(E)**.

lincRNA genes spanning copies A1, A3, and A4 encode three annotated transcripts of nearly identical sequence, which were previously described by Delihas (2020b) and named *FAM247D*, *FAM247C*, and *FAM247A*, respectively (HUGO nomenclature; **Figure 4C**). The annotated transcript spanning A2 (*FAM247B*) contains only few nucleotides of the 3'UTR and was not further studied. The lincRNA gene spanning B1 (here named *linc-UR-B1*) encodes two transcript isoforms that are annotated as *TCONS_00029753* and *TCONS_00029754* (UCSC genome browser) and differ of only 4 nt at the

TABLE 1 | *USP18* 3'UTR copies in human and non-human primates.

Coordinates[a]	Strand[b]	Identity % (score)[c]
Human (hg38)[d]		
chr22:18176818-18177397[e]	+	100 (580/580)
chr22:21192790-21193369	–	99.9 (578/580)
chr22:18758535-18759114	–	99.7 (576/580)
chr22:18360610-18361189	+	99.7 (576/580)
chr22:18501224-18501803	–	99.5 (574/580)
chr22:21299777-21300356	+	99.5 (574/580)
chr22:18863834-18864413	+	99.4 (572/580)
Chimpanzee (panTro6)		
chr22:2521056-2521635	–	98.8 (566/580)
chr22:134544-135123	–	100 (580/580)
chr22:7470999-7471578	+	99.9 (578/580)
chr22:5364294-5364873	+	99.7 (576/580)
chr22:680203-680782	+	99.5 (574/580)
chr22:5399839-5400413	+	98.3 (570/580)
African Green Monkey (chlSab2)		
chr19:2370494-2371091	+	92.3 (479/580)
Rhesus Macaque (rheMac10)		
chr10:33831575-33832175	–	92.3 (485/580)

[a]*BLAT analysis performed on human and non-human primates genomes using human USP18 3'UTR as input.*
[b]*Refers to the genomic DNA strand.*
[c]*Hits with identity > 90% and length > 100 bp are shown.*
[d]*In brackets, the assembly used as reference for each genome.*
[e]*In bold coordinates referred to the USP18 protein-coding gene.*

exon-exon junction (**Figure 4D**; **Supplementary Figure S5A**). As presently annotated, all these transcripts contain only two exons (**Figures 4C,D**).

Next, we measured the expression of these transcripts by qPCR using a commercial panel of cDNAs from polyA+ RNAs of 20 human tissues. In the NCBI database, *FAM247A*, *FAM247C*, and *FAM247D* appear to be widely expressed in human tissues, as measured by RNA sequencing (BioProjects: PRJEB4337, PRJEB2445, PRJNA280600, and PRJNA270632). However, given the repeated nature of their sequences, a precise quantification of their expression requires either the use of specific qPCR primers or additional filters to extract uniquely mapping reads in the RNA-seq data. Thus, we designed specific primers across the unique exon-exon junction of *FAM247A, C*, and *D* and across the unique exon-exon junction of *linc-UR-B1* (**Figures 4C,D**; primers in **Supplementary Table S2**). *FAM247* transcripts (*FAM247A/C/D*) were lowly expressed in most tissues and showed the highest expression in fetal and adult liver, kidney, and thymus (**Figure 4E**). On the other hand, *linc-UR-B1* was uniquely and abundantly expressed in testis (**Figure 4F**). To confirm the testis-restricted expression of *linc-UR-B1*, we interrogated RNA-seq datasets covering a panel of 32 tissues (E-MTAB-2836). To distinguish *linc-UR-B1* from all other *USP18*-containing transcripts, uniquely mapped reads were selected and aligned again to the human genome (hg38). This analysis revealed that *linc-UR-B1* is expressed only in testis as a single isoform corresponding to the *TCONS_00029754* annotation (**Supplementary Figure S5B**). This was further confirmed by sequencing the only product amplified by qPCR on testis cDNA (**Supplementary Figure S5C**). High inter-individual variability in the level of *linc-UR-B1* was noticed in the RNA-seq dataset (eight donors; **Supplementary Figure S5B**). Our qPCR analyses on testis fragments from six donors displaying normal spermatogenesis and one donor with impaired spermatogenesis (as observed by transillumination of seminiferous tubules) confirmed this variability and showed the lowest *linc-UR-B1* expression in the latter donor (**Supplementary Figure S5D**).

As opposed to protein-coding genes, the annotations of lincRNAs are far from being complete (Lagarde et al., 2016). To validate the existing annotations, we performed 3'RACE on RNAs from liver and testis. We confirmed the 3' end of *FAM247A/C/D* as annotated (**Figure 5A**, right panel) and showed that the 3' end of *linc-UR-B1* extends further than annotated and actually contains the entire *USP18* 3'UTR (580 nt; **Figure 5B**, right panel). Mapping the 5' end of these transcripts proved to be challenging, since the sequences that are annotated as exon 1 (i.e., upstream of the 3'UTR) are themselves duplicated, with several copies residing on chromosome 22 and other chromosomes (**Supplementary Table S1**). Hence, to enrich for transcripts bearing the *USP18* 3'UTR, we performed a RT PCR reaction on RNAs from liver and testis using a specific primer (spRT) complementary to the 3'UTR and using forward primers specific to the annotated 5' ends (**Figures 5A,B**, left panels). These analyses confirmed the annotated sequences for *FAM247A/C/D* and *linc-UR-B1*. Moreover, we could not amplify *FAM247A/C/D* sequence further using primers designed in the upstream genomic region, suggesting that these RNAs do not extend further.

The genomic region upstream of *linc-UR-B1* contains a pseudogene annotated as *NR_135922* and belonging to the large *POM121* family (**Supplementary Figure S6A**). Similar to *linc-UR-B1*, *NR_135922* is expressed only in testis (**Supplementary Figure S6B**). To study whether *linc-UR-B1* transcript extends 5' into *NR_135922* sequences, we designed primers either in the last annotated exon of *NR_135922* or in the intergenic region, and combined them with a reverse primer in the *USP18* 3'UTR. For all primer pairs, the expected products could be amplified from testis RNA (**Figure 5C** left panel and **Figure 5D**), indicating that *NR_135922* and *linc-UR-B1* are part of the same gene. Indeed, 3'RACE on testis RNA yielded the expected *NR_135922* product (537 nt) and a longer isoform (*linc-UR-B1*) that terminates with *USP18* 3'UTR (**Figure 5C**, right panel and **Figure 5D**).

Linc-UR-B1 Is Uniquely Expressed in Male Germ Cells and Positively Correlates With USP18

Given the high and specific expression of *linc-UR-B1* in the testis, we sought to identify in which testicular cell type(s) this RNA is expressed. We first performed qPCR analysis on cDNAs from freshly isolated testicular germ cell populations (spermatocytes and spermatids) and from primary somatic peritubular, Leydig and Sertoli cells and detected *linc-UR-B1* only in spermatocytes and spermatids (**Figure 6A**). To obtain further information on *linc-UR-B1* expression, we surveyed a

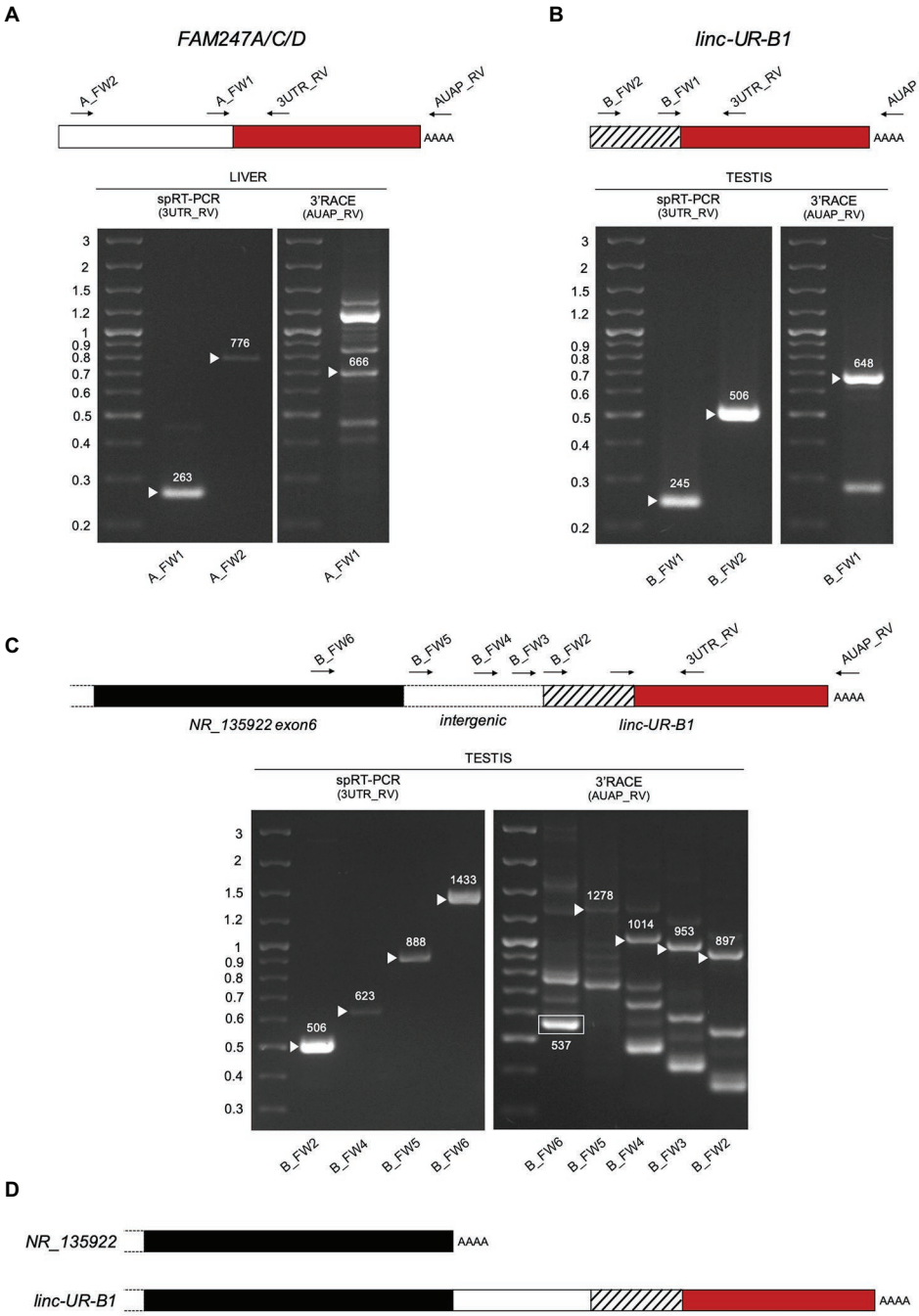

FIGURE 5 | Characterization of *FAM247A/C/D* and *linc-UR-B1* molecules. **(A)** Analysis of *FAM247* transcripts in liver. Right gel, 3' of cDNA Ends (3'RACE). Complementary DNA (cDNA) was obtained by oligo-dT reverse transcription (RT) of polyA+ liver RNA. PCR was performed on this cDNA using the forward (A_FW1) and reverse (AUAP_RV) primers indicated on the schematic above. Several bands were obtained, including the expected 666 bp. Left gel, specific RT-PCR (spRT-PCR). cDNA was obtained from polyA+ liver RNA using a primer designed on the *USP18* 3'UTR (3UTR_GSP1). PCR was then performed using the forward primer (A_FW1 or A_FW2) and a reverse primer designed on the 3'UTR (3UTR_RV) and internal to 3UTR_GSP1. The length in nt is indicated above each amplified product. **(B)** Analysis of *linc-UR-B1* transcripts in testis. The strategies described in **(A)** were used on testis poly A+ RNA. Primers are indicated on the schematic above. The length in nt is indicated above each amplified band. **(C)** Study of *linc-UR-B1* transcript extending on the 5' end. Three forward primers were designed on the annotated intergenic region (B_FW3-4-5). One primer (B_FW6) was designed on the sequence further upstream, corresponding to exon 6 of the annotated *NR_135922* transcript. Right panel, 3'RACE on testis RNA. The 537 nt product obtained using the B_FW6 primer corresponds to the annotated *NR_135922*. Longer products indicated by arrowheads were obtained when using the three primers designed in the intergenic region. Left panel, spRT-PCR was performed on testis RNA using the indicated forward primers and the 3UTR_RV primer. PCR products of the expected lengths were obtained for all primers tested.

(Continued)

FIGURE 5 | Panels **(A–C)**: the PCR products of the size expected from the annotations were sequenced. The additional 3'RACE PCR products likely result from the amplification of other transcripts containing sequences identical to the exon upstream of *USP18* exon11 (**Supplementary Table S1**). **(D)** Schematic of the 3' end of the two *NR_135922* isoforms identified here. Top, the *NR_135922* transcript terminating with exon 6, as in the annotation. Bottom, longer isoform of *NR_135922* terminating with the *USP18* 3'UTR, here called *linc-UR-B1*. The sequences of *FAM247A/C/D* and *linc-UR-B1* are provided as Supplementary Information.

TABLE 2 | Copies of *USP18* 3'UTR embedded in lincRNA genes.

Copies	Coordinates[a]	Strand[b]	Span (bp)[c]	Annotated gene	Annotated transcript	Transcript name[d]
A1	18358013-18361189	+	2889	LOC105372942 (FAM247D)[e]	XR_951230.1	FAM247D
A2	18501226-18504399	–	2884	AC023490.4 (FAM247B)	ENST00000621762.1	FAM247B
A3	18758535-18761708	–	3115	LOC105377182 (FAM247C)	XR_951239.1	FAM247C
A4	21192790-21195963	–	3129	LOC105372935 (FAM247A)	XR_938017.2	FAM247A
B1	18863190-18864411	+	1224	TCONS_00029753/4[f]	TCONS_00029753/4	linc-UR-B1
B2	21299133-21300354	+	1224	no gene	no transcript	

[a]*BLAT analysis performed on human genome using human USP18 intron10-exon11 sequence as input.*
[b]*Refers to the genomic DNA strand.*
[c]*Hits with identity > 90% and length > 100 bp are shown.*
[d]*Transcript name used in this manuscript*
[e]*HUGO nomenclature is indicated in brackets*
[f]*TCONS_000297553 and TCONS_000297554 (here indicated as TCONS000297553/4) represent two transcripts isoforms of the same gene (see **Supplementary Figure S5A**).*
All coordinates, genes and transcripts refers to annotations found in the UCSC genome browser.

single-cell RNA-seq dataset of testicular cells from three individuals (Guo et al., 2018). The expression profile of *linc-UR-B1* was analyzed by tracking the sequence of the unique exon-exon junction between the 3'UTR and the upstream exon. As shown in **Figure 6B**, *linc-UR-B1* was primarily detected in specific germ cell populations (clusters 4–6). Interestingly, the induction of *linc-UR-B1* expression occured in the transition between meiotic early primary and late primary spermatocytes. This burst was followed by a gradual decrease in post-meiotic round and elongated differentiating spermatids. No or low expression was detected in somatic cell populations (clusters 9–13). To identify more precisely the meiotic phase in which *linc-UR-B1* induction occurs, we analyzed the dataset of early and late primary spermatocytes reclustered for the five different stages of meiotic prophase I. *Linc-UR-B1* was first detectable in the late pachytene stage and was maximal during the diplotene stage (**Figure 6C**). Importantly, we observed that the expression of *USP18* paralleled that of *linc-UR-B1* in germ cell populations and during prophase I (**Figures 6D,E**). Since *USP18* is IFN-inducible, its expression may be related to the presence of local IFN. To test this, we analyzed expression of two ISGs, *IFIT1* and *OAS1*, but no or minimal levels of these transcripts were detected in germ cells (**Supplementary Figure S7**). In fact, *IFIT1* and *OAS1* were higher in cells other than germ cells, whereas *USP18* was highest in meiotic and post-meiotic germ cells (**Figure 6D**). The positive correlation between *linc-UR-B1* and *USP18* expression during spermatogenesis (**Figure 6F**) and, in particular, in meiotic cells (**Figure 6G**) suggests that these transcripts – which share the same 3'UTR – may cross-talk. Interestingly, we found that testis, and in particular germ cells, express moderate to high levels of three *USP18*-targeting miRNAs, namely *miR-191-5p*, *miR-24-3p*, and *miR-423-5p* (**Supplementary Figure S8**). These data hint at the possibility of a miRNA-dependent cross-talk between *USP18* and *linc-UR-B1* in germ cells, with the latter possibly acting as a sponge for miRNAs targeting USP18.

DISCUSSION

Ubiquitin-specific peptidase 18 determines the threshold of activation of the type I IFN signaling pathway and ultimately the amount of ISGs in a given cell. To exert its function as a negative regulator, USP18 binds to the IFN receptor/JAK complex and attenuates the magnitude of the response (Francois-Newton et al., 2011; Wilmes et al., 2015; Arimoto et al., 2017). The abundance of other IFN-stimulated components (i.e., STAT2/1, IRF9, and SOCS1) will also impact cell context-specific ISG activation, but in distinct and less specific manners (Kok et al., 2020). Given the critical non-redundant role of USP18, we conceived this study to assess whether USP18 can be post-transcriptionally regulated by miRNAs directed to the 3'UTR. Through the use of target prediction tools, functional validation, public data mining, and correlative analyses, we identified four miRNAs (*miR-191-5p*, *miR-24-3p*, *miR-423-5p*, and *miR-532-3p*) that directly pair to sites within the 580 nt-long 3'UTR and tune down endogenous USP18 at both mRNA and protein levels. These miRNAs are expressed at different levels in numerous immune and non-immune cell types. We do not know whether all four miRNAs collaborate to robustly target USP18. Yet, *miR-191-5p*, *miR-24-3p*, and *miR-532-3p* are enriched in circulating human monocytes, which we found exhibit low baseline USP18 with respect to PBL.

Monocytes are a subset of leukocytes involved in anti-microbial defenses, anti-tumor immune responses, and in tissue homeostasis (Shi and Pamer, 2011). Interestingly, Uccellini et al. using an ISRE-dependent reporter mouse, showed that Ly6C^{hi} inflammatory monocytes – corresponding to the classical human monocytes and representing ≈ 90% of the circulating ones – display a high basal IFN response, further enhanced upon infection (Uccellini and Garcia-Sastre, 2018). Under physiological conditions, circulating monocytes are primed by constitutive IFNβ, in order to be armed with adequate levels of key proteins (transcription factors, i.e., STATs/IRFs, some

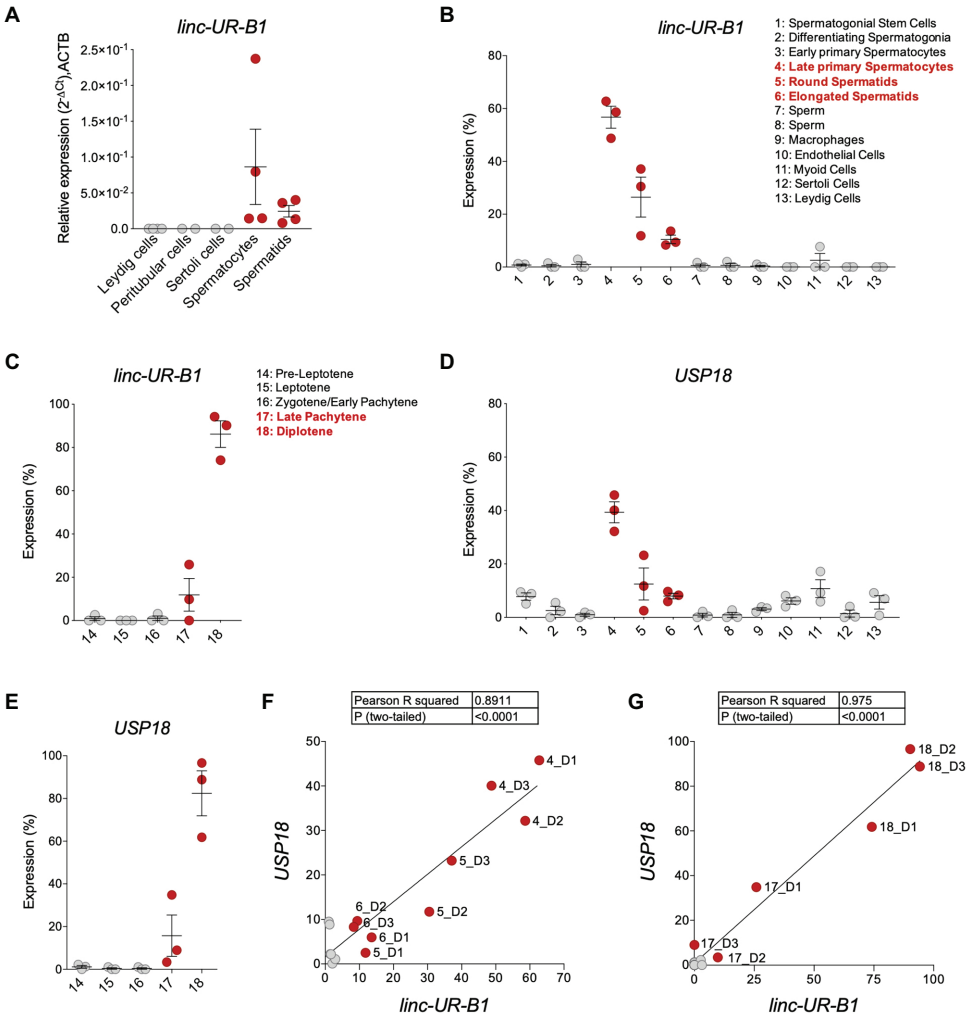

FIGURE 6 | *linc-UR-B1* is expressed in meiotic and post-meiotic germ cells and positively correlates with *USP18*. **(A)** Measure of *linc-UR-B1* levels by quantitative PCR (qPCR) in purified testicular cell populations (spermatocytes, spermatids, peritubular cells, and Leydig cells) and Sertoli cells. Results (± SEM) shown as expression ($2^{-\Delta Ct}$) relative to *ACTB*. **(B,D)** Analysis of *linc-UR-B1* **(B)** and *USP18* **(D)** expression in testicular cell populations (clusters 1–13, three donors). **(C,E)** Analysis of *linc-UR-B1* **(C)** and *USP18* **(E)** expression in clusters of germ cells populations at different stages of prophase I (three donors). Note that clusters 14–18 are subclusters of clusters 3 and 4 studied in **(B)** and **(D)**. **(F,G)** Positive correlation of *linc-UR-B1* and *USP18* expression in germ cell populations (**F**, clusters 1–8) and in germ cells at different stages of prophase I (**G**, clusters 14–18). All data were retrieved from the alignment of single cell RNA-seq reads provided by (Guo et al., 2018), filtered for unique reads, and expressed (± SEM) as percentage expression (normalized expression in one population vs. all populations analyzed). Cell populations expressing *linc-UR-B1* are highlighted in red.

ISGs, and immune response genes products; Gough et al., 2010; Lindqvist et al., 2016). During acute viral infection, in some auto-immune and inflammatory diseases, monocytes encounter higher levels of IFN that in turn promotes their activation and their differentiation into dendritic-like cells with a potent antigen-presenting capacity (Gerlini et al., 2008). Low USP18 may be critical to maintain high responsiveness of these cells to IFN. In line with this, a recent study on *ISG15*-deficient patients showed that, among PBMCs, monocytes displayed the highest IFN signature (Martin-Fernandez et al., 2020). Thus, the targeting of USP18 by *miR-191-5p*, *miR-24-3p*, and *miR-532-3p*, and possibly *miR-423-5p*, may assist the priming of monocytes toward a robust inflammatory and immune response.

The conservation analysis of the *USP18* 3'UTR sequence led us to identify several duplications of the sequences spanning intron 10-exon 11. We found six copies of this sequence mapping in human chr22q11.21. We named these copies as A or B depending on the breakpoint in intron 10 (**Figure 4B**). In chr22q11.2 there are eight LCRs, of which four (LCR22A-D) span most of chr22q11.21, a 3 Mb region that is deleted in 90% of patients affected by the 22q11.2 deletion syndrome. First described by DiGeorge in the 1960s, it represents the most frequent chromosomal microdeletion syndrome, affecting 1 per 3,000–6,000 live births. Patients show heterogeneous clinical features, multi-organ dysfunction, cognitive deficits and neuropsychiatric illness, immunodeficiency, and cardiac and

palatal abnormalities (Mcdonald-Mcginn et al., 2015; Morrow et al., 2018). Such recurrent genomic rearrangements have fueled many studies to resolve the complex architecture of the LCR22s, identify breakpoints, and study the high variability among patients (Demaerel et al., 2019). The LCR22A-D duplications are believed to have occurred recently, in the hominoid-lineage after their divergence from the Old World monkeys (Babcock et al., 2007; Delihas, 2020a). The highly dynamic nature of LCR22s, as of other LCRs, is believed to be a driving force in genome evolution and adaptation, as their rearrangements can give rise to new coding and non-coding genes (Babcock et al., 2003; Delihas, 2018). Interestingly, the *USP18* protein-coding gene resides at the centromeric boundary of LCR22A. Moreover, the gene *USP41* resides in the LCR22B (**Figure 4A**). In databases, *USP41* is annotated as a protein-coding gene. However, this gene contains a sequence nearly identical to *USP18*, but lacks 5' and 3'UTRs. We could not find experimental evidence that *USP41* is transcribed in human tissues. Thus, despite its coding potential, *USP41* is likely an unprocessed pseudogene.

Previous in-depth studies of the LCR22s did report on several copies of *USP18* intron 10-exon 11 and suggested that *Alu* elements present at their 5' breakpoint (**Supplementary Figure S9A**) may have driven duplication events (Babcock et al., 2003; Delihas, 2018). Delihas referred to "*USP18*-linked" sequences as part of a repeat unit present in human and chimpanzee genomes, which includes upstream sequences related to the gamma-glutamyl-transferase (*GGT*) gene family and downstream sequences related to the *FAM230* gene family (Delihas, 2020a). Accordingly, this is the exact genomic context that surrounds the A copies of *USP18* exon 11 (**Supplementary Figure S9B**). Members of the *FAM230* family are also found downstream of the two B copies and of the *USP18* protein-coding gene, while members of the *POM121* family are found upstream (**Supplementary Figure S9B**). Altogether these observations reinforce the hypothesis that the *USP18* intron 10-exon 11 sequence was trapped, together with *FAM230* sequences, in a gene-forming element, and this event may have contributed to the generation of new genes (Delihas, 2020a). Interestingly, four of the six copies of *USP18* exon 11 are embedded in expressed lincRNA genes, three of which were previously described as *FAM247A*, *C*, and *D* and retain an intact *USP18* 3'UTR (Delihas, 2020b). Here, we identified an additional lincRNA bearing *USP18* exon 11, i.e., the entire 3'UTR, to which we refer as *linc-UR-B1* (**Figures 4C,D**). By tracking the unique junction between *USP18* exon 11 and the upstream exon, we showed that *FAM247A/C/D* are more abundant in tissues like liver, kidney, and thymus, while *linc-UR-B1* is uniquely and highly expressed in testis, notably in spermatocytes and spermatids. To our knowledge, this is the first case known of a 3'UTR that is found duplicated independently from the rest of the ancestral gene and embedded in expressed non-coding RNAs. Few non-coding transcripts that bear the 3'UTR sequence of a protein-coding gene have been described. These transcripts are expressed from pseudogenes, notably *PTENP1*, *KRASP1*, and *BRAFP1*. These latter originated by a single duplication of the ancestral protein-coding genes (coding region and UTRs), they are transcribed but not translated due to mutations creating premature stop codons or frameshifts (Poliseno et al., 2010; Karreth et al., 2015). Conversely, the lincRNAs that we have identified bear only the 3'UTR of *USP18*.

In recent years many non-coding RNAs have been discovered, but only few have been assigned functions (Yao et al., 2019). Moreover, a large number of lncRNA genes are transcriptionally active in testis during meiosis and spermatogenesis (Cabili et al., 2011; Soumillon et al., 2013; Rolland et al., 2019). The lincRNAs we describe here may represent junk DNA, proliferating in a selfish manner. Yet, evolution may have repurposed some of them to confer an advantage, be transmitted, and potentially fixed in the population (Dawkins, 1976; Eddy, 2012). It is a fact that the inclusion of exon 11 – possibly favored by a strong constitutive acceptor site (ttctctctagGCAGGAAACT) at the intron 10-exon 11 junction – provides to these transcripts a canonical AATAAA signal for poly-adenylation. The best studied non-coding transcript bearing the 3'UTR of a protein-coding gene is *PTENP1*. This pseudogene undergoes copy number loss in human cancers, this correlating with a decrease in levels of the tumor suppressor *PTEN*. It was proposed that *PTENP1* exerts its function by sponging miRNAs, thus sustaining PTEN levels (Poliseno et al., 2010). Likewise, lincRNAs bearing the *USP18* 3'UTR may act as decoys titrating away *USP18*-targeting miRNAs. Interestingly, our analysis of public RNA-seq data from testicular cell subsets revealed a positive correlation between *linc-UR-B1* and *USP18* in spermatocytes and spermatids. Moreover, we found that three *USP18*-targeting miRNAs are expressed in whole testis and in germ cells. Further work is needed to investigate the possibility that in human germ cells *linc-UR-B1* and *USP18* compete for the binding of these miRNAs.

Of note, the lincRNAs described here contain a small ORF of 43 nt that could encode the last 14 amino acids of USP18 (aa 359–372), eight of which are conserved from zebrafish to primates (Delihas, 2020b). In USP18, these C-ter residues are involved in binding IFNAR2 (Arimoto et al., 2017; Basters et al., 2018). Intriguingly, in *linc-UR-B1* an ATG codon – present at the junction of exon 11 and the upstream exon – may serve as translation initiation codon for the 15 amino acids peptide. The existence of such peptide and its potential impact will have to be investigated in germ cells. Moreover, since *linc-UR-B1* is a hybrid of *NR_135922* (highly similar to *POM121L8P*) and *USP18* sequences, regions other than exon 11 are likely to contribute to its function(s).

Interestingly, studies in the mouse have shown that high/continuous IFNα/β affects spermatogenesis by causing apoptosis of germ cells and induces sterility (Hekman et al., 1988; Iwakura et al., 1988; Satie et al., 2011). It is therefore tempting to speculate that the maintenance of baseline *USP18* represents a recently evolved mechanism to attenuate IFN signaling and protect the precursors of spermatozoa. In this scenario, human spermatocytes and spermatids may be a good shelter for viruses. Indeed, the testis is a reservoir for a number of viruses, such as Zika and Ebola viruses which can be sexually transmitted by infected men after recovery (Matusali et al., 2018; Le Tortorec et al., 2020). We recently revealed the prolonged seminal

excretion of testicular germ cells persistently infected by Zika virus (Mahe et al., 2020). Future studies will be necessary to determine whether *linc-UR-B1* contributes to maintain *USP18* in human male germ cells and indirectly favors viral persistence. For this, we will privilege approaches capable of solving RNA-RNA or RNA-protein interactions of the endogenous *linc-UR-B1* in the unique context of germ cells.

In conclusion, our work reveals the existence of miRNAs regulating USP18 through its 3'UTR, and of several lincRNAs containing the *USP18* 3'UTR. Combined, these molecules may form a non-coding network tuning USP18 levels in cell types where IFN responsiveness needs to be tightly controlled.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics Commitee Ouest V, Rennes, France (authorization DC-2016-2783) and the French National Agency for Biomedical Research (authorization PFS09-015). Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

ER and MC designed and performed experiments and analyzed the data. ÖV, LV, and GG performed experiments under supervision. NT and ER analyzed the RNA-seq and sequence conservation data. AT, AR, and ND-R provided and performed experiments on testis fragments and purified testicular populations. ND-R and FM provided expert advice on experiments. ER, MC, and SP wrote the manuscript. All authors revised the manuscript. SP supervised the work. All authors contributed to the article and approved the submitted version.

FUNDING

Research in the Unit of Cytokine Signaling is funded by the Institut Pasteur, the Fondation pour la Recherche Médicale (Equipe FRM DEQ20170336741) and the Institut National de la Santé et de la Recherche Médicale (Inserm). ER was supported by Sorbonne Université and by FRM. MC was supported by the FRM grant above. GG and ÖV were supported by the Erasmus plus programme of the EU commission. Work at Rennes University is funded by Inserm and Défis Scientifiques Emergents.

ACKNOWLEDGMENTS

We wish to thank P. Miesen for help in initial prediction analyses of miRNAs targeting USP18 3'UTR; J. Guo for assisting us in retrieving scRNA-seq dataset on the Human Testis Cell Atlas; the Flow Cytometry Platform of Institut Pasteur and S. Meunier for technical assistance for cytometric analyses; the members of the Unit of Cytokine Signaling, M. Livingstone, G. Uzé, V. Libri and N. Jouvenet for advice, discussions and support; and M. C. Gauzzi for discusssions and critical reviewing of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.627007/full#supplementary-material

REFERENCES

Allantaz, F., Cheng, D. T., Bergauer, T., Ravindran, P., Rossier, M. F., Ebeling, M., et al. (2012). Expression profiling of human immune cell subsets identifies miRNA-mRNA regulatory relationships correlated with cell type specific expression. *PLoS One* 7:e29979. doi: 10.1371/journal.pone.0029979

Arimoto, K. I., Lochte, S., Stoner, S. A., Burkart, C., Zhang, Y., Miyauchi, S., et al. (2017). STAT2 is an essential adaptor in USP18-mediated suppression of type I interferon signaling. *Nat. Struct. Mol. Biol.* 24, 279–289. doi: 10.1038/nsmb.3378

Babcock, M., Pavlicek, A., Spiteri, E., Kashork, C. D., Ioshikhes, I., Shaffer, L. G., et al. (2003). Shuffling of genes within low-copy repeats on 22q11 (LCR22) by Alu-mediated recombination events during evolution. *Genome Res.* 13, 2519–2532. doi: 10.1101/gr.1549503

Babcock, M., Yatsenko, S., Hopkins, J., Brenton, M., Cao, Q., de Jong, P., et al. (2007). Hominoid lineage specific amplification of low-copy repeats on 22q11.2 (LCR22s) associated with velo-cardio-facial/digeorge syndrome. *Hum. Mol. Genet.* 16, 2560–2571. doi: 10.1093/hmg/ddm197

Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116, 281–297. doi: 10.1016/S0092-8674(04)00045-5

Basters, A., Knobeloch, K. P., and Fritz, G. (2018). USP18- a multifunctional component in the interferon response. *Biosci. Rep.* 38:BSR20180250. doi: 10.1042/BSR20180250

Cabili, M. N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., et al. (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* 25, 1915–1927. doi: 10.1101/gad.17446611

Dauphinee, S. M., Richer, E., Eva, M. M., Mcintosh, F., Paquet, M., Dangoor, D., et al. (2014). Contribution of increased ISG15, ISGylation and deregulated type I IFN signaling in Usp18 mutant mice during the course of bacterial infections. *Genes Immun.* 15, 282–292. doi: 10.1038/gene.2014.17

Dawkins, R. (1976). *The selfish gene*. New York: Oxford University Press.

de Rie, D., Abugessaisa, I., Alam, T., Arner, E., Arner, P., Ashoor, H., et al. (2017). An integrated expression atlas of miRNAs and their promoters in human and mouse. *Nat. Biotechnol.* 35, 872–878. doi: 10.1038/nbt.3947

Delihas, N. (2018). A family of long intergenic non-coding RNA genes in human chromosomal region 22q11.2 carry a DNA translocation breakpoint/AT-rich sequence. *PLoS One* 13:e0195702. doi: 10.1371/journal.pone.0195702

Delihas, N. (2020a). Formation of human long intergenic non-coding RNA genes, pseudogenes, and protein genes: ancestral sequences are key players. *PLoS One* 15:e0230236. doi: 10.1371/journal.pone.0230236

Delihas, N. (2020b). Genesis of non-coding RNA genes in human chromosome 22-a sequence connection with protein genes separated by evolutionary time. *Noncoding RNA* 6:36. doi: 10.3390/ncrna6030036

Demaerel, W., Mostovoy, Y., Yilmaz, F., Vervoort, L., Pastor, S., Hestand, M. S., et al. (2019). The 22q11 low copy repeats are characterized by unprecedented size and structural variability. *Genome Res.* 29, 1389–1401. doi: 10.1101/gr.248682.119

Ebert, M. S., and Sharp, P. A. (2012). Roles for microRNAs in conferring robustness to biological processes. *Cell* 149, 515–524. doi: 10.1016/j.cell.2012.04.005

Eddy, S. R. (2012). The C-value paradox, junk DNA and ENCODE. *Curr. Biol.* 22, R898–R899. doi: 10.1016/j.cub.2012.10.002

Forster, S. C., Tate, M. D., and Hertzog, P. J. (2015). MicroRNA as type I interferon-regulated transcripts and modulators of the innate immune response. *Front. Immunol.* 6:334. doi: 10.3389/fimmu.2015.00334

Francois-Newton, V., Magno de Freitas Almeida, G., Payelle-Brogard, B., Monneron, D., Pichard-Garcia, L., Piehler, J., et al. (2011). USP18-based negative feedback control is induced by type I and type III interferons and specifically inactivates interferon alpha response. *PLoS One* 6:e22200. doi: 10.1371/journal.pone.0022200

Gerlini, G., Mariotti, G., Chiarugi, A., Di Gennaro, P., Caporale, R., Parenti, A., et al. (2008). Induction of $CD83^{+}CD14^{+}$ nondendritic antigen-presenting cells by exposure of monocytes to IFN-alpha. *J. Immunol.* 181, 2999–3008. doi: 10.4049/jimmunol.181.5.2999

Goldmann, T., Zeller, N., Raasch, J., Kierdorf, K., Frenzel, K., Ketscher, L., et al. (2015). USP18 lack in microglia causes destructive interferonopathy of the mouse brain. *EMBO J.* 34, 1612–1629. doi: 10.15252/embj.201490791

Gough, D. J., Messina, N. L., Hii, L., Gould, J. A., Sabapathy, K., Robertson, A. P., et al. (2010). Functional crosstalk between type I and II interferon through the regulated expression of STAT1. *PLoS Biol.* 8:e1000361. doi: 10.1371/journal.pbio.1000361

Guo, J., Grow, E. J., Mlcochova, H., Maher, G. J., Lindskog, C., Nie, X., et al. (2018). The adult human testis transcriptional cell atlas. *Cell Res.* 28, 1141–1157. doi: 10.1038/s41422-018-0099-2

Hekman, A. C., Trapman, J., Mulder, A. H., van Gaalen, J. L., and Zwarthoff, E. C. (1988). Interferon expression in the testes of transgenic mice leads to sterility. *J. Biol. Chem.* 263, 12151–12155. doi: 10.1016/S0021-9258(18)37906-7

Honke, N., Shaabani, N., Cadeddu, G., Sorg, U. R., Zhang, D. E., Trilling, M., et al. (2011). Enforced viral replication activates adaptive immunity and is essential for the control of a cytopathic virus. *Nat. Immunol.* 13, 51–57. doi: 10.1038/ni.2169

Honke, N., Shaabani, N., Zhang, D. E., Hardt, C., and Lang, K. S. (2016). Multiple functions of USP18. *Cell Death Dis.* 7:e2444. doi: 10.1038/cddis.2016.326

Honke, N., Shaabani, N., Zhang, D. E., Iliakis, G., Xu, H. C., Haussinger, D., et al. (2013). Usp18 driven enforced viral replication in dendritic cells contributes to break of immunological tolerance in autoimmune diabetes. *PLoS Pathog.* 9:e1003650. doi: 10.1371/journal.ppat.1003650

Iwakura, Y., Asano, M., Nishimune, Y., and Kawade, Y. (1988). Male sterility of transgenic mice carrying exogenous mouse interferon-beta gene under the control of the metallothionein enhancer-promoter. *EMBO J.* 7, 3757–3762. doi: 10.1002/j.1460-2075.1988.tb03259.x

Karreth, F. A., Reschke, M., Ruocco, A., Ng, C., Chapuy, B., Leopold, V., et al. (2015). The BRAF pseudogene functions as a competitive endogenous RNA and induces lymphoma in vivo. *Cell* 161, 319–332. doi: 10.1016/j.cell.2015.02.043

Kok, F., Rosenblatt, M., Teusel, M., Nizharadze, T., Goncalves Magalhaes, V., Dachert, C., et al. (2020). Disentangling molecular mechanisms regulating sensitization of interferon alpha signal transduction. *Mol. Syst. Biol.* 16:e8955. doi: 10.15252/msb.20198955

Lagarde, J., Uszczynska-Ratajczak, B., Santoyo-Lopez, J., Gonzalez, J. M., Tapanari, E., Mudge, J. M., et al. (2016). Extension of human lncRNA transcripts by RACE coupled with long-read high-throughput sequencing (RACE-Seq). *Nat. Commun.* 7:12339. doi: 10.1038/ncomms12339

Le Tortorec, A., Matusali, G., Mahe, D., Aubry, F., Mazaud-Guittot, S., Houzet, L., et al. (2020). From ancient to emerging infections: the odyssey of viruses in the male genital tract. *Physiol. Rev.* 100, 1349–1414. doi: 10.1152/physrev.00021.2019

Lindqvist, R., Mundt, F., Gilthorpe, J. D., Wolfel, S., Gekara, N. O., Kroger, A., et al. (2016). Fast type I interferon response protects astrocytes from flavivirus infection and virus-induced cytopathic effects. *J. Neuroinflammation* 13:277. doi: 10.1186/s12974-016-0748-7

Lindsay, M. A. (2008). microRNAs and the immune response. *Trends Immunol.* 29, 343–351. doi: 10.1016/j.it.2008.04.004

Mahe, D., Bourgeau, S., Frouard, J., Joguet, G., Pasquier, C., Bujan, L., et al. (2020). Long-term Zika virus infection of non-sperm cells in semen. *Lancet Infect. Dis.* 20:1371. doi: 10.1016/S1473-3099(20)30834-3

Malakhova, O. A., Kim, K. I., Luo, J. K., Zou, W., Kumar, K. G., Fuchs, S. Y., et al. (2006). UBP43 is a novel regulator of interferon signaling independent of its ISG15 isopeptidase activity. *EMBO J.* 25, 2358–2367. doi: 10.1038/sj.emboj.7601149

Martin-Fernandez, M., Bravo Garcia-Morato, M., Gruber, C., Murias Loza, S., Malik, M. N. H., Alsohime, F., et al. (2020). Systemic type I IFN inflammation in human ISG15 deficiency leads to necrotizing skin lesions. *Cell Rep.* 31:107633. doi: 10.1016/j.celrep.2020.107633

Matusali, G., Houzet, L., Satie, A. P., Mahe, D., Aubry, F., Couderc, T., et al. (2018). Zika virus infects human testicular tissue and germ cells. *J. Clin. Invest.* 128, 4697–4710. doi: 10.1172/JCI121735

Mcdonald-Mcginn, D. M., Sullivan, K. E., Marino, B., Philip, N., Swillen, A., Vorstman, J. A., et al. (2015). 22q11.2 deletion syndrome. *Nat. Rev. Dis. Primers.* 1:15071. doi: 10.1038/nrdp.2015.71

Meuwissen, M. E., Schot, R., Buta, S., Oudesluijs, G., Tinschert, S., Speer, S. D., et al. (2016). Human USP18 deficiency underlies type 1 interferonopathy leading to severe pseudo-TORCH syndrome. *J. Exp. Med.* 213, 1163–1174. doi: 10.1084/jem.20151529

Morrow, B. E., Mcdonald-Mcginn, D. M., Emanuel, B. S., Vermeesch, J. R., and Scambler, P. J. (2018). Molecular genetics of 22q11.2 deletion syndrome. *Am. J. Med. Genet. A* 176, 2070–2081. doi: 10.1002/ajmg.a.40504

Poliseno, L., Salmena, L., Zhang, J., Carver, B., Haveman, W. J., and Pandolfi, P. P. (2010). A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* 465, 1033–1038. doi: 10.1038/nature09144

Rolland, A. D., Evrard, B., Darde, T. A., Le Beguec, C., Le Bras, Y., Bensalah, K., et al. (2019). RNA profiling of human testicular cells identifies syntenic lncRNAs associated with spermatogenesis. *Hum. Reprod.* 34, 1278–1290. doi: 10.1093/humrep/dez063

Satie, A. P., Mazaud-Guittot, S., Seif, I., Mahe, D., He, Z., Jouve, G., et al. (2011). Excess type I interferon signaling in the mouse seminiferous tubules leads to germ cell loss and sterility. *J. Biol. Chem.* 286, 23280–23295. doi: 10.1074/jbc.M111.229120

Schwabenland, M., Mossad, O., Peres, A. G., Kessler, F., Maron, F. J. M., Harsan, L. A., et al. (2019). Loss of USP18 in microglia induces white matter pathology. *Acta Neuropathol. Commun.* 7:106. doi: 10.1186/s40478-019-0757-8

Shi, C., and Pamer, E. G. (2011). Monocyte recruitment during infection and inflammation. *Nat. Rev. Immunol.* 11, 762–774. doi: 10.1038/nri3070

Soumillon, M., Necsulea, A., Weier, M., Brawand, D., Zhang, X., Gu, H., et al. (2013). Cellular source and mechanisms of high transcriptome complexity in the mammalian testis. *Cell Rep.* 3, 2179–2190. doi: 10.1016/j.celrep.2013.05.031

Speer, S. D., Li, Z., Buta, S., Payelle-Brogard, B., Qian, L., Vigant, F., et al. (2016). ISG15 deficiency and increased viral resistance in humans but not mice. *Nat. Commun.* 7:11496. doi: 10.1038/ncomms11496

Uccellini, M. B., and Garcia-Sastre, A. (2018). ISRE-reporter mouse reveals high basal and induced type I IFN responses in inflammatory monocytes. *Cell Rep.* 25, 2784.e3–2796.e3. doi: 10.1016/j.celrep.2018.11.030

Vuillier, F., Li, Z., Commere, P. H., Dynesen, L. T., and Pellegrini, S. (2019). USP18 and ISG15 coordinately impact on SKP2 and cell cycle progression. *Sci. Rep.* 9:4066. doi: 10.1038/s41598-019-39343-7

Wilmes, S., Beutel, O., Li, Z., Francois-Newton, V., Richter, C. P., Janning, D., et al. (2015). Receptor dimerization dynamics as a regulatory valve for plasticity of type I interferon signaling. *J. Cell Biol.* 209, 579–593. doi: 10.1083/jcb.201412049

Yao, R. W., Wang, Y., and Chen, L. L. (2019). Cellular functions of long noncoding RNAs. *Nat. Cell Biol.* 21, 542–551. doi: 10.1038/s41556-019-0311-8

Zhang, X., Bogunovic, D., Payelle-Brogard, B., Francois-Newton, V., Speer, S. D., Yuan, C., et al. (2015). Human intracellular ISG15 prevents interferon-alpha/beta over-amplification and auto-inflammation. *Nature* 517, 89–93. doi: 10.1038/nature13801

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.