nature structural & molecular biology

Article

https://doi.org/10.1038/s41594-022-00853-0

The transcriptional terminator XRN2 and the RNA-binding protein Sam68 link alternative polyadenylation to cell cycle progression in prostate cancer

In the format provided by the authors and unedited

The transcriptional terminator XRN2 and the RNA binding protein Sam68 link alternative polyadenylation to cell cycle progression in prostate cancer

Marco Pieraccioli[1,2], Cinzia Caggiano[1,2], Luca Mignini[3], Chuwei Zhong[4], Gabriele Babini[2], Rossano Lattanzio[5,6], Savino Di Stasi[7], Bin Tian[4], Claudio Sette[1,8,§] and Pamela Bielli[3,8,§]

[1]Department of Neuroscience, Section of Human Anatomy, Catholic University of the Sacred Hearth, Rome, Italy; [2]Fondazione Policlinico Agostino Gemelli IRCCS, Rome, Italy; [3]Department of Biomedicine and Prevention, University of Rome Tor Vergata, Rome, Italy; [4]Gene Expression and Regulation Program, The Wistar Institute, Philadelphia, USA, [5]Department of Innovative Technologies in Medicine & Dentistry, G. d'Annunzio University, Chieti, Italy; [6]Center for Advanced Studies and Technology (CAST), G. d'Annunzio University, Chieti, Italy; [7]Department of Experimental Medicine and Surgery, University of Rome Tor Vergata, Rome, Italy; [8]Laboratory of Neuroembryology, IRCCS Fondazione Santa Lucia, Rome, Italy.

Running title: XRN2 and Sam68 regulate alternative polyadenylation in PC
Keywords: XRN2, Sam68, alternative polyadenylation, cell cycle, prostate cancer

[§]Corresponding Authors: Pamela Bielli
Department of Biomedicine and Prevention, University of Rome Tor Vergata, Rome, Italy; Email: pamela.bielli@uniroma2.it

Claudio Sette
Department of Neuroscience, Section of Human Anatomy, Università Cattolica del Sacro Cuore, Largo Francesco Vito 1, 00168 Rome, Italy; Email: claudio.sette@unicatt.it

Supplemental Methods

Bioinformatic analysis

XRN2 promoter characterization was performed utilizing bioinformatics tools present in the UCSC Genome Browser (https://genome.ucsc.edu; GRCh37/h19) and ChIP-seq data for H3K4Me3, H3K27Ac and RNAPII from ENCODE project (https://www.encodeproject.org). POLR2A and MYC/MAX ChIP-seq Cluster analyses was performed using Transcription Factor ChIP-seq Clusters (338 factors, 130 cell types) from ENCODE3.

3'READS analysis of LNCaP silenced for Sam68 and XRN2 was performed as described[1]. Reads were mapped to the human (hg19) genome. PCA and sample distance analysis on 3'READS samples were generated using custom R scripts.

Metaplot of average binding enrichment of CSTF64, CPSF30 and Sam68 within a ±100nt genomic region surrounding upregulated, downregulated or unregulated pA were computed using PAR-CLIP CSTF64 and CPSF30 bigWig files (GSE37401)[2]. Sam68 uvCLAP CLIP-seq (GSE85164)[3] bam files were converted into bigWig files using deepTools bamCoverage function. Meta-region plots were calculated and plotted using genomation and ggplot2 R packages, respectively.

Plasmid constructs

Human wild-type XRN2 promoter were amplified from genomic DNA of LNCaP using primer reported in Supplementary Table S1 in presence of Phusion Hot Start High-Fidelity DNA polymerase (Thermo Fisher Scientific). Inserts were cloned into pGL3-basic vector (Promega) using MluI/XhoI (New England Biolabs) and plasmids sequenced by Cycle Sequencing (Eurofins Genomics). pGL3-basic and pCDNA3 plasmids carrying intergenic control region and MYC gene have been previously described[4]. The *FLNB* minigene was generated by amplifying the second to last exon, the last intron, the last exon and 200nt of the genomic region downstream from genomic DNA of LNCaP using FLNB/EcoRI/FW and FLNB/SalI/Rev primers and cloned into the EcoRI/SalI restriction sites of pCI vector (Promega) lacking SV40 polyA site (SmaI/NaeI). The *FLNB* mutant minigene was generated by megaprimer strategy[5] using FLNB FLNB/EcoRI/FW, FLNB/SalI/Rev and FLNB/mut primers. In *FLNB* minigene scheme (Figure 6H) boxes and black line represent exons and intron, respectively. The red box and line indicate the alternative 3'UTR and the genomic region downstream of the d-pA, respectively. XRN2 was amplified from LNCaP cells using XRN2/HindIII/Fw and XRN2/BamHI/Rv primers and cloned in the HindIII/BamHI restriction sites of p3XFlag-CMV-7 expression vector (Sigma) and the D235A mutation was introduced by megaprimer strategy[5] using XRN2/D235A/Rv primer. sh-RNA targeting the 3'UTR region of human

XRN2 was generated using shXRN2-UTR/Fw and shXRN2-UTR/Rv primers and cloned in the AgeI/EcoRI restriction sites of pLKO.1-puro plasmid (Sigma). Short and long 3'UTR isoforms of MCM10 were amplified using MCM10-UTR-Fw-XhoI and MCM10-UTR-Rv-NotI primers and cloned in psiCHECK2-2 vector (Promega). All primers are listed in Supplementary Table 2.

Luciferase-based report assay

Luciferase-based report assay was performed in 293T and LNCaP cells[4] and carried out using the Dual Luciferase Reporter Assay (Promega). The luciferase assay of long and short MCM10 3′UTR was performed in LNCaP cells after 8 hrs from transfection and activity was reported as Renilla/Firefly ratio.

GST pull-down

For GST pull down assays[6], Sam68-GST fusion proteins[6] were purified from bacterial lysates on glutathione-agarose (Merck, Sigma-Aldrich) and analysed by SDS-PAGE and Coomassie blue staining to test purity and integrity. GST fusion proteins were equilibrated for 1 h in RSB100 Buffer [10 mM Tris/HCl 7.4, 2.5 mM $MgCl_2$, 100mM NaCl, 0.5% Triton X-100] supplemented with 20 mM β-glycerophosphate, 0.5 mM Na_3VO_4, 1 mM dithiothreitol (DTT) and protease inhibitor cocktail (Merck, Sigma-Aldrich) and then 500 µg of LNCaP nuclear extract was added to the beads. After 2h incubation at 4°C under constant shaking, beads were washed 3 times with RSB100 buffer and boiled in Laemmli Sample buffer.

Isolation of 4-thiouridine-labeled RNA

For isolation of 4-thiouridine (4sU)-labeled RNA[7], growing LNCaP cells stably silenced by lentiviral transduction with control and XRN2 sh-RNAs were treated with 400uM 4sU and vehicle, as negative control. After one-hour cells were collected and RNA was isolated by standard procedures. A total of 150 µg of RNA samples were labeled in biotinylation buffer (100mM Tris-HCl pH7.4, 10mM EDTA, 0.2mg/ml biotin-HDPD) and incubated at RT for 1.5 hours under rotation. RNA was then purified by standard procedure and biotinylated RNA was isolated by streptoavidin beads for 15 min at RT. After three extensive washing with washing buffer (100mM Tris-HCl pH 7.4; 10mM EDTA; 1M NaCl; 0,1% Tween20) at 65°C, and three more washing at RT, biotinylated RNA was eluted with 100mM DTT, purified according to standard procedure, retrotranscribed with random primers and used for qPCR analysis. Bound RNA is represented as percentage (%) of input.

Fluorescence-activated cell sorting (FACS) analysis

For cell cycle and cell proliferation analyses[4], LNCaP cells were silenced and pulse-labelled with 30 μM BrdU (Sigma-Aldrich) for 1h. Cells were collected, fixed overnight with a solution of 70% ethanol in PBS (vol/vol) (4°C) and, after ice-cold PBS washes, stained with 50 μl of anti-BrdU antibody (0.02μg/μl) diluted in PBS supplemented with 1% BSA and 0.5% Tween-20, for 1h at RT. After washes in PBS, cells were incubated with 50 μl of Alexa-488-conjugated secondary antibody (0.02μg/μl), for 45 min at RT and treated with 10 μg/ml RNAse A (Merck, Sigma-Aldrich) in presence of propidium iodide (20 μg/ml) (Merck, Sigma-Aldrich). For double thymidine block, cells were treated with 2 mM thymidine (Merck, Sigma-Aldrich) as described[8] and fixed in ice-cold 70% ethanol overnight at 4°C. A total of 20,000 events were counted with FACS-Calibur flow cytometer (Becton Dickinson) using Cell Quest software (Becton Dickinson) and analysed using FlowJo software.

References

1. Hoque, M. et al. Analysis of alternative cleavage and polyadenylation by 3' region extraction and deep sequencing. *Nat Methods* **10**, 133-9 (2013).
2. Martin, G., Gruber, A.R., Keller, W. & Zavolan, M. Genome-wide analysis of pre-mRNA 3' end processing reveals a decisive role of human cleavage factor I in the regulation of 3' UTR length. *Cell Rep* **1**, 753-63 (2012).
3. Aktas, T. et al. DHX9 suppresses RNA processing defects originating from the Alu invasion of the human genome. *Nature* **544**, 115-119 (2017).
4. Caggiano, C., Pieraccioli, M., Panzeri, V., Sette, C. & Bielli, P. c-MYC empowers transcription and productive splicing of the oncogenic splicing factor Sam68 in cancer. *Nucleic Acids Res* **47**, 6160-6171 (2019).
5. Barik, S. Site-directed mutagenesis in vitro by megaprimer PCR. *Methods Mol Biol* **57**, 203-15 (1996).
6. Bielli, P. et al. The transcription factor FBI-1 inhibits SAM68-mediated BCL-X alternative splicing and apoptosis. *EMBO Rep* **15**, 419-27 (2014).
7. Radle, B. et al. Metabolic labeling of newly transcribed RNA for high resolution gene expression profiling of RNA synthesis, processing and decay in cell culture. *J Vis Exp* (2013).
8. Chen, G. & Deng, X. Cell Synchronization by Double Thymidine Block. *Bio Protoc* **8**(2018).

Supplementary Table 1. Oligonucleotides used for ChIP

Name	Sequence (5' to 3')
human ChIP Sam68 promoter MYCBS F	GCAGACAAAGACATTCCACAGC
human ChIP Sam68 promoter MYCBS R	CCAGAGCTTGATGCGCATG
human ChIP XRN2 promoter MYCBS F	GTCACTTCCGTTCTCCAGCG
human ChIP XRN2 promoter MYCBS R	GGCTGACGAGCGTAACCAAAG
human ChIP 16q22 F	CTACTCACTTATCCATCCAGGCTAC
human ChIP 16q22 R	ATTTCACACACTCAGACATCACAG

Supplementary Table 2. Oligonucleotides used for cloning

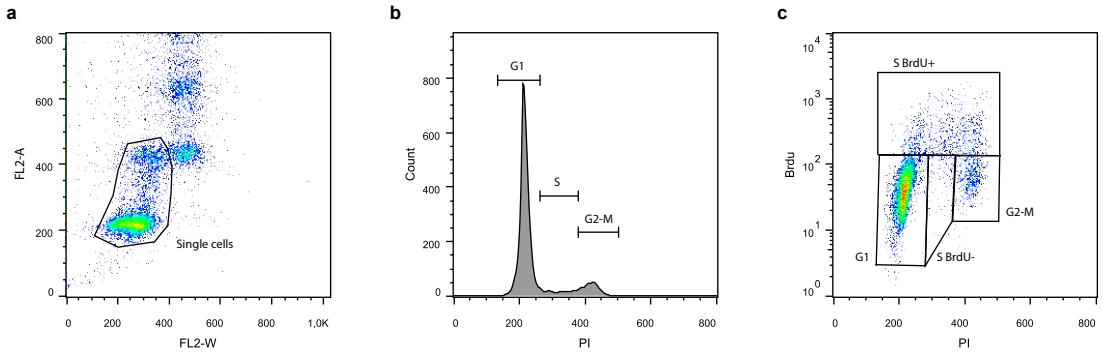
Name	Sequence (5' to 3')
human intergenic F MluI	TCTAACGCGTTTTCGGCATTTTTCACAGAGATGT
human intergenic R XhoI	CAGGCTCGAGCATGCCAGTCTCTTTGTAAGCC
human XRN2 promoter F MluI	TCTAACGCGTCATGCAGAGCGCAGCGATTC
human XRN2 promoter R XhoI	CAGGCTCGAGCCGAGGGTCCTGAATTCCTC
MCM10-UTR-Fw-XhoI	CAACTCGAGCCCGAACTTCAGACATTTTCC
MCM10-UTRL-Rv-NotI	TCTGCGGCCGCTTATGTTTTAAATTTATTTTAAAAGTTAAAATCA AATTGTAAAAAGGC
MCM10-UTRS-Rv-NotI	TCTGCGGCCGCAGATGGGAAAGTTTTATTATTTGTAGAAATC CAGTGCT
FLNB/EcoRI/Fw	AGAGAATTCGCCAGCGTCTAGTTAGCCCTG
FLNB/SalI/Rv	AGAGTGGACACCTTGGCCTCTCAGGCTCAAG
FLNB/mut	CCGACTGTAAGTTGTATCTTGAAGCTTGTTGAATAAAACATTTGC TCCGGAAAGTTGC
XRN2/HindIII/Fw	AGAAGCTTATGGGAGTCCCGGCGTTCTTC
XRN2/BamHI/Rv	GGGATCC TTAATTCCAATTGTATCTTCCTGAGGGT
XRN2 D235A/Rev	GCATAATGAGAgCAGCATCTG
shXRN2-UTR/Fw	CCGGTGTGTATTCTAGATCATCTAAGCTCGAGCTTAGATGATCTA GAATACACTTTTTG
shXRN2-UTR/Rv	AATTCAAAAAGTGTATTCTAGATCATCTAAGCTCGAGCTTAGATG ATCTAGAATACACA

Supplementary Table 3. Oligonucleotides used for RT-PCR and qRT-PCR

Name	Sequence (5' to 3')
human Actin F	GTTGCTATCCAGGCTGTGCTA
human Actin R	AATGTCACGCACGATTTCCCG
human Myc F	AATGAAAAGGCCCCCAAGGTAGTTATCC
human Myc R	GTCGTTTCCGCAACAAGTCCTCTTC
human XRN2 F	CAGTGAACCTGAGCCAGAGGAT
human XRN2 R	GACTGCACAACTTTCCGACGGA
human Sam68 F	TGACGGCAGAAATTGAGAAG
human Sam68 R	CCAAGAATCTTCCCCACAAA
human SCARB2 p-pA F	CACCTCCTTGCAGGCTTGTT
human SCARB2 p-pA R	TGTCAGCCTGCTCTTTAACCT
human SCARB2 d-pA F	ATCCTTGGCCCTCTGGACAT
human SCARB2 d-pA R	CCGCATGAAGTGAACCAACTG
human SCARB2 CLIP p-pA F	GGCTCTGAAATTGTTCTTTTATGTGTC
human SCARB2 CLIP p-pA R	ATTCTGGAGCTACCAGCACCCA
human SCARB2 CLIP d-pA F	CCCTTAATGCAATGAAATGTTTCAT
human SCARB2 CLIP d-pA R	CTCCATAAACAACAGGACATCCA
human FLNB p-pA F	CCAACATGCTGCTGATCGGG
human FLNB p-pA R	CCCCTCTCCTTGACGACGTA
human FLNB d-pA F	AGTATCTGGGGCGAGTGTTG
human FLNB d-pA R	TCCGTGTCCCAGCTTACAAC
human FLNB CLIP p-pA F	CATGTCACAGTGCCTTAAAACAG
human FLNB CLIP p-pA R	CAGGCTGGAGGGCTTTGTAT

human FLNB CLIP d-pA F	TGGAATCAAACGCCGACTGT
human FLNB CLIP d-pA R	CAGAAGCGACATCATCCAAGAAC
human RNF130 p-pA F	GATGTCTCTTATCCCCTTCACTCA
human RNF130 p-pA R	TGGGGAGATTTCAGTGTATTTAGGA
human RNF130 d-pA F	CCTATGTGGGCATTCGGGAC
human RNF130 d-pA R	GCAACACACATCACTTTCTTGGT
human CEP70 p-pA F	ATTCTCTCTTATGTACTAGCACCTTC
human CEP70 p-pA R	CAAGATTTCCAGTCCTACTAGGTAA
human CEP70 d-pA F	GTTGCCCTTTCAAAGCAGGAG
human CEP70 d-pA R	TACATCAAGATTGCAAAAGACACT
human CD164 p-pA F	AATGGGCTTTTGCTTTTCTAGGT
human CD164 p-pA R	AATCAAGTGCGAAACTCAGCC
human CD164 d-pA F	TTGTTGCACAGGCAGATACT
human CD164 d-pA R	ACATGATACAGATTGGTTTTGCAGT
human PDIA6 p-pA F	TAGCCGGTGAATTGAGTCGT
human PDIA6 p-pA R	GCCTCCAAGAACATTCAAGCA
human PDIA6 d-pA F	GTTGTTAATGAACCTCAAAATAGCC
human PDIA6 d-pA R	TACCTGTTCGCATTGGTAACCT
human RCC2 p-pA F	AGTGGTGCTTATGGGGGAGA
human RCC2 p-pA R	CATGGAGACAGTGACAAGTGGTA
human RCC2 d-pA F	AAGAGGCTTAGCGTGAAGGG
human RCC2 d-pA R	ACAGCACTAAACCAGGCACC
human SCAMP2 p-pA F	CCTCAGCCTTTCTCTCGCCTG
human SCAMP2 p-pA R	GAGGAGGAAGAGCCACGGCAAG
human SCAMP2 d-pA F	GCCTTCCCACAAAACTAGCCT
human SCAMP2 d-pA R	GAGTTCATAAAACCCCAGCCG
human SETD6 p-pA F	CGAAACTCAGCTGGAGGGAAC
human SETD6 p-pA R	TTCCTTCAGGGAACAGGGAAAC
human SETD6 d-pA F	TCCAATAGTGAGCCCTTCCCA
human SETD6 d-pA R	GGGGACCTTGAAAATCACTAAGTTA
human CDKN1B p-pA F	TCAGCAGAATGGTGATCACTCCAG
human CDKN1B p-pA R	TAACATTCAAAACTCCCAAGCACC
human CDKN1B d-pA F	AGTGGCATGTTTTGTGCATTTG
human CDKN1B d-pA R	TGCAACCTTTTAAGCATAGCCA
human RIF1 p-pA F	TGCTATGCGTGGTTTTTCAG
human RIF1 p-pA R	ATGCACGGGTTTAAGAAAGCA
human RIF1 d-pA F	TTCGTTTCATGACTCTGTGCC
human RIF1 d-pA R	TACTGTTCGGCTTCCCAAGAA
human MCM10 p-pA F	TTCATTCACTGAAGTTTTTGCCCA
human MCM10 p-pA R	TTGCCTCTGGGTGTTCGTAT
human MCM10 p-Pa polysome R	TTTTTTTTTTTTTTTTTTTTAGATG
human MCM10 d-pA F	AGGTTGGTGTCTGAGCAATCC
human MCM10 d-pA R	TCCAGATACAGCTCACCCTCA
human MCM10 dPa polysome R	TTTTTTTTTTTTTTTTTTTTGACAA
human MCM10 GE F	AGGATGAATTGGAGCCTGCC
human MCM10 GE R	CATCTCAGCCTCGGCCTCTTTC
human ORC2 p-pA F	TGAACAGTGTGCTCAGGGAC
human ORC2 p-pA R	CAATCAGCTCGGGGAGTGTT
human ORC2 p-Pa polysome R	TTTTTTTTTTTTTTTTTTTTTTACG
human ORC2 d-pA F	CCAGCCTAGAAAGCACAGGA
human ORC2 d-pA R	TCTCTAGCCTGGCTTCTCACT
human ORC2 dPa polysome R	TTTTTTTTTTTTTTTTTTTTACAAG
human ORC2 GE F	CTTATTGAGCAAGGTTTCCCCTTC
human ORC2 GE R	GCACAATGTTGAACCCAAGGTGT
human NPAT p-pA F	GCATGACCTGCACTTTCATTGT
human NPAT p-pA R	GCTATACAGTTTTGCAGATTGGCT
human NPAT d-pA F	TGCTGAACCTTACTGCAAAAACT
human NPAT d-pA R	CCTTAGGAATAGACCACTGCCA

human LAMC1 p-pA F	CCACAGAGTGCCTTGACACAA
human LAMC1 p-pA R	AAAAGGACAGTCATGGACAGCA
human LAMC1 d-pA F	GAATGCTGGTTGACACAGGGA
human LAMC1 d-pA R	TTATTAACCACCCCGGAGAGC
human FEM1C p-pA F	GCTAATGCCACAGAAAGGGCT
human FEM1C p-pA R	GCAGGGCAAATATGCCATCATAAA
human FEM1C d-pA F	CTCTCCCTTTTGTCAGTTCCCA
human FEM1C d-pA R	ACCTTACAATGAATGTTGTGGCA
human CEBPG p-pA F	ACTCAAAGACCCATTGGAGGC
human CEBPG p-pA R	GGGGGATTTTGATTCCCACA
human CEBPG d-pA F	AGCTGGGAAGGAAGAAGTCCA
human CEBPG d-pA R	CACAAACTGATGAAAGCCAGGAA
human MALAT1 p-pA F	ACTTGTTCCTGTGGGCTTCA
human MALAT1 p-pA R	CCTTAGTTGGCATCAAGGCAC
human MALAT1 d-pA F	TAATAGCCTGCAGCTGGTGT
human MALAT1 d-pA R	ATTCACAAAAGAAGGACAACAATTT
human FAM57A p-pA F	TCTGGTGCCAACACTCATTCA
human FAM57A p-pA R	TGAGAACACATGGTAGCACAGT
human FAM57A d-pA F	GTTGAGTCTTTAATGGGTGCCA
human FAM57A d-pA R	GGGTGACTTTTAAAGATACACTGAT
human KTM2C p-pA F	TGTGTGCACAATTCAAGTGGAC
human KTM2C p-pA R	ACATCATGCCATATACAAGTTCAGA
human KTM2C d-pA F	ACCAGTATTTGTCCTACCGGG
human KTM2C d-pA R	TGTCCACAAATTCCAAAACCCA
human NUSAP1 p-pA F	AAGCTGGGATAGAAAGGCCA
human NUSAP1 p-pA R	CCAAACTATATGGCCCCGTTTC
human NUSAP1 d-pA F	TAGGCTTTGCTTAGTATCATGTCCA
human NUSAP1 d-pA R	TTGAGCAGTATAAAACTCAGAAGCTC
human CCNB1 p-pA F	GCACTCTACCACAGCTGAAT
human CCNB1 p-pA R	TACAAGTTACACCTTTGCCACA
human CCNB1 d-pA F	GAGTTACTGAAGGTGATGGAGGT
human CCNB1 d-pA R	GCAGATTCTCCATGACTTGGC
human SRSF3 p-pA F	GGATACATGGCTGTTCGTGAC
human SRSF3 p-pA R	GCACAGTAACTCAAAGTTCGGC
human SRSF3 d-pA F	ATTTGAGGTGTTGCATGGGAA
human SRSF3 d-pA R	CACCTGCCATTTTCTAAGAGGA



Supplementary Figure 1. Representative flow cytometry gating strategy.
a, doublet cells were excluded by gating for FL2-Area and FL2-Width signals to obtain single cells population. **b**, cell cycle progression of gated cells was measured in the PI histogram plot using markers set within the analysis program. **c**, BrdU incorporation was measured in the PI-BrdU dot plot by gating for PI and and BrdU signals.