Translatome and transcriptome analysis of TMA20 (MCT-1) and TMA64 (eIF2D) knockout yeast strains

Desislava S. Makeeva, Andrey S. Lando, Aleksandra Anisimova, Artyom A. Egorov, Maria D. Logacheva, Alexey A. Penin, Dmitry E. Andreev, Pavel G. Sinitcyn, Ilya M. Terenin, Ivan N. Shatsky, Ivan V. Kulakovskiy, Sergey E. Dmitriev

A B S T R A C T

TMA20 (MCT-1), TMA22 (DENR) and TMA64 (eIF2D) are eukaryotic translation factors involved in ribosome recycling and re-initiation. They operate with P-site bound tRNA in post-termination or (re-)initiation translation complexes, thus participating in the removal of 40S ribosomal subunit from mRNA stop codons after termination and controlling translation re-initiation on mRNAs with upstream open reading frames (uORFs), as well as de novo initiation on some specific mRNAs. Here we report ribosomal profiling data of S. cerevisiae strains.
with individual deletions of TMA20, TMA64 or both TMA20 and TMA64 genes. We provide RNA-Seq and Ribo-Seq data from yeast strains grown in the rich YPD or minimal SD medium. We illustrate our data by plotting differential distribution of ribosomal-bound mRNA fragments throughout uORFs in 5'-untranslated region (5' UTR) of GCN4 mRNA and on mRNA transcripts encoded in MAT locus in the mutant and wild-type strains, thus providing a basis for investigation of the role of these factors in the stress response, mating and sporulation. We also document a shift of transcription start site of the APC4 gene which occurs when the neighboring TMA64 gene is replaced by the standard G418-resistance cassette used for the creation of the Yeast Deletion Library. This shift results in dramatic deregulation of the APC4 gene expression, as revealed by our Ribo-Seq data, which can be probably used to explain strong genetic interactions of TMA64 with genes involved in the cell cycle and mitotic checkpoints. Raw RNA-Seq and Ribo-Seq data as well as all gene counts are available in NCBI Gene Expression Omnibus (GEO) repository under GEO accession GSE122039 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE122039).

&© 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

**Specifications table**

| Subject area          | Biology |
|-----------------------|---------|
| More specific subject area | Protein Synthesis, Translational Control, Ribosome, Bioinformatics, Transcriptomics, Translatomics |
| Type of data          | Table and figures |
| How data was acquired | Ribosome profiling and RNA-Seq of wild-type or knockout yeast strains were performed. Sequences were obtained using Illumina HiSeq 2000. |
| Data format           | Raw and analyzed |
| Experimental factors  | Saccharomyces cerevisiae BY4741 wild-type strain and BY4741-based strains with TMA20, TMA64 or both TMA20 and TMA64 knockouts were maintained in rich (YPD) or minimal (SD) media. |
| Experimental features | In the mid-log exponential phase, yeast cells were pretreated with cycloheximide and collected. cDNA libraries of ribosome-bound mRNA and total mRNA from wild-type and knockout strains were performed as described previously [1]. Sequenced reads were trimmed, read mapping and counting was performed. |
| Data source location  | Moscow State University (Moscow, Russia) |
| Data accessibility    | Analyzed data is presented in the article. Raw RNA-Seq and Ribo-Seq data as well as all gene counts are available in NCBI Gene Expression Omnibus (GEO) repository under GEO accession GSE122039 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE122039). |
| Related research article | N/A |

**Value of the data**

- The data provides a gene expression landscape of yeast strains lacking TMA20 and/or TMA64 proteins, which are orthologous to mammalian translation factors MCT-1 and eIF2D, thus...
expanding our knowledge about individual functional roles of these two translation factors in a living cell.

- An abnormal translation of the MATα2 mRNA derived from the MAT locus of MATα yeast strain is detected, which can be used for explanation of sporulation defects previously detected in the TMA64 deletion strain.
- Quantitative Ribo-Seq data provides essential information of translational changes in the knockout strains including altered uORFs translation in 5’ UTR of mRNA encoding important transcription regulator GCN4, thus providing a basis for investigating the role of these proteins in the stress response.
- The RNA-Seq data highlights transcriptional abnormalities within the APC4 gene locus, caused by replacement of the adjacent TMA64 gene by the standard G418 or HYG resistance cassettes commonly used for generating gene deletions, which can be probably used to explain previously observed strong genetic interactions of TMA64 with genes involved in the cell cycle and mitotic checkpoints.
- The deep sequenced Ribo-Seq and RNA-Seq are applicable for detailed bioinformatics analysis of translation events, such as prediction of alternative open reading frames.

### 1. Data

In this study we present ribosome profiling data generated from the wild-type BY4741 *S. cerevisiae* strain and strains lacking translation factors TMA20 (MCT-1), TMA64 (eIF2D) or both of them at the same time. Information on all performed experiments is shown in Table 1. Raw Ribo-Seq and RNA-Seq data are available online in the NCBI Gene Expression Omnibus repository (GEO accession: GSE122039, [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE122039](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE122039)). Supplementary Table S1 and Supplementary Figure S1 contain analyzed NGS data, as described below. Examples of differentially translated and transcribed genes in wild-type and knockout strains are presented in Figs. 1–3.

In contrast to a dataset previously obtained in a study by Young et al. [2], here we present ribosome profiling data not only for double knockout yeast strains, but also for strains with individual deletions of TMA20 and TMA64. This allows studying transcriptional and translational changes caused

| Sample | Sample name | Yeast strain name | Yeast strain genotype | Growth media | Sample type |
|--------|-------------|------------------|-----------------------|--------------|-------------|
| 1      | wt1_ribo    | wt MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 | YPD         | Ribo-Seq    |
| 2      | wt1_rna     | wt              | YPD                   | RNA-Seq     |
| 3      | wt2_ribo    | Δtma20 MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 tma20Δ::KanMX4 | YPD         | RiboSeq     |
| 4      | wt2_rna     | Δtma64 MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 tma64Δ::KanMX4 | YPD         | RNA-Seq     |
| 5      | wt_sd_ribo  | wt          | SD                     | Ribo-Seq    |
| 6      | wt_sd_rna   | Δtma64     | SD                     | RNA-Seq     |
| 7      | tma20_ribo  | Δtma20 MATα his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 tma20Δ::KanMX4 | YPD         | RiboSeq     |
| 8      | tma20_rna   | Δtma64     | YPD                   | RNA-Seq     |
| 9      | tma20_sd_ribo | Δtma20Δ::KanMX4 | SD                     | Ribo-Seq    |
| 10     | tma20_sd_rna | Δtma64     | SD                     | RNA-Seq     |
| 11     | tma64_ribo  | Δtma64     | YPD                   | RNA-Seq     |
| 12     | tma64_rna   | Δtma20Δ::KanMX4 | YPD         | RNA-Seq     |
| 13     | tma64_sd_ribo | Δtma20Δ::KanMX4 | YPD         | RNA-Seq     |
| 14     | tma64_sd_rna | Δtma64     | SD                     | RNA-Seq     |
| 15     | tma20tma64_ribo | Δtma20Δ::KanMX4 | YPD         | Ribo-Seq    |
| 16     | tma20tma64_rna | Δtma20Δ::KanMX4 | YPD         | RNA-Seq     |
| 17     | tma20tma64_sd_ribo | Δtma20Δ::KanMX4 | YPD         | RNA-Seq     |
| 18     | tma20tma64_sd_rna | Δtma20Δ::KanMX4 | YPD         | RNA-Seq     |
specifically by the absence of individual translation factors TMA20 (MCT-1) or TMA64 (eIF2D). Due to cycloheximide addition to yeast culture medium before harvesting the cells, distribution of mapped reads from Ribo-Seq sets may slightly vary from data described by Young et al. According to a previous study [3], use of the inhibitor in such a fashion is likely to cause blurring of local density effects, but also strengthens Ribo-Seq signals at translation initiation sites, facilitating analysis of ribosome distribution over uORFs.

2. Experimental design, materials, and methods

2.1. Cell maintenance and cDNA libraries preparation

RNA-Seq and Ribo-Seq cDNA libraries were prepared from total RNA samples or ribosome-bound RNA samples, respectively, for both the wild-type BY4741 yeast strain and three knockout strains.

![Fig. 1. Ribo-Seq and RNA-Seq coverage of MATa locus in the studied yeast strains. According to Ribo-Seq signals, MATa2-2 ORF is probably translated in mutant strains. The Y axis tracks show total read coverage (positive and negative values correspond to the coverage of the direct and reverse complementary strands respectively).](image)
(with individually deleted TMA20 or TMA64 genes, or with a double deletion of TMA20 and TMA64, hereafter referred as wt, Δtma20, Δtma64 and ΔΔtma20tma64, respectively). The libraries were sequenced, resulting in 9 RNA-Seq and 9 Ribo-Seq data sets. Table 1 summarizes information about all of the sequencing experiments.

The experimental procedure in general followed the ribosome profiling protocol described in [1]. Briefly, yeast cells were grown to an exponential phase in either rich YPD (1% yeast extract, 2% peptone, 2% glucose) or minimal SD (0.67% YNB w/o amino acids with ammonium sulfate, 2% glucose, complete amino acid supplementation) media. Cycloheximide was added to yeast media to a final concentration of 100 μg/ml and growth was continued for 3 more minutes; then cells were harvested by filtration, resuspended in polysome lysis buffer (20 mM Tris pH 8.0, 140 mM KCl, 1.5 mM MgCl2, 100 g/ml cycloheximide, 1% Triton), flash frozen in liquid nitrogen and homogenized by grinding. Then a portion of each cell lysate was used for total RNA isolation, while another part was treated with RNase I for polysome disassembly, applied to a sucrose gradient for fractionation, followed by isolation of a monosome fraction and extraction of ribosome-protected mRNA fragments for ribosome profiling. mRNA was isolated using Oligo(dT) beads and ribosome-bound RNA was isolated from sucrose fractions using acidic-phenol extraction. Further ribosome profiling and RNA-Seq library preparations were performed as described previously [1]. Two biological replicates indicated as WT1 and WT2 were performed for wild-type strain maintained in YPD.

Fig. 2. Ribo-Seq and RNA-Seq coverage of the 5’ UTR of the yeast GCN4 gene. Ribo-Seq profiles show ribosome occupancy at different uORFs that regulate translation of the main coding region. The Y axis tracks show total read coverage (positive and negative values correspond to the coverage of the forward and reverse complementary strands respectively).
2.2. Sequence data processing and analysis

Reads were trimmed with cutadapt v1.18 [4] and mapped to the Saccharomyces_cerevisiae.R64-1-1.90 (Ensembl) genome assembly using the respective genome annotation. The read mapping and counting were performed with STAR v2.5.3a [5]. The genomic signal plots (Figs. 1–3) were generated by the svist4get software [6] using bedGraph files constructed from bam alignments by samtools and bedtools [7,8]. Transcription start site coordinates for MAT locus [9] and GCN4, TMA64 and APC4 genes [10] were used for 5’ UTR mapping in Figs. 2 and 3.

3. Data analysis

3.1. Ribosome profiling of yeast strains lacking TMA20 and/or TMA64 genes

Translation factor TMA64 and homologs of its N- and C-terminal regions, TMA20 and TMA22 respectively (eIF2D, MCT-1, and DENR in mammals) are proteins involved in translation termination, re-initiation, and ribosome recycling. Initially, eIF2D and heterodimer MCT-1●DENR were assumed to provide a non-canonical translation initiation pathway as they facilitate GTP-independent delivery of Met-tRNA^{Met} and some elongator tRNAs to the 40S ribosomal P-site [11,12]. In addition, in vitro and in vivo studies demonstrated that TMA64/eIF2D, TMA20/MCT-1, and TMA22/DENR are able to promote the post-terminational tRNA and mRNA release from the 40S ribosomal subunit both in yeast and mammals [2,13]. The absence of these factors, together with the 40S recycling failure, led to deregulated translation re-initiation downstream of both short and full-size translated open reading frames in different organisms [2,14–16].

The C-terminal regions of TMA64/eIF2D and TMA22/DENR contain the SUI1 domain, which is also present in the translation factor SUI1/eIF1. Structural data indicate that the SUI1 domains of all three factors have similar positions in the P-site of the 40S ribosomal subunit with a conserved β-loop protruding toward a codon-anticodon duplex formed by mRNA and a P-site tRNA [17,18]. In accordance with biochemical data, this suggests that during recycling, TMA64/eIF2D and the heterodimer
TMA20-TMA22 (MCT-1-DENR) may operate in a manner similar to SUJ1/eIF1 in translation initiation, or control initiator tRNA access to re-initiating ribosomal complexes after uORF translation.

Raw and analyzed Ribo-Seq and RNA-Seq data sets for wild-type, individual Δtma20 and Δtma64, as well as double ΔΔtma20tma64 knockout yeast strains were obtained and uploaded into the NCBI Gene Expression Omnibus repository (GEO accession: GSE122039, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE122039). The initial analysis revealed that they included deep RNA-Seq and Ribo-Seq data with more than 10 million uniquely mapped and counted reads within gene CDS. In general, in a half of the samples, ~50% of the total library length was composed of uniquely mapped reads. Across all samples, 60 to 90% of them were located within annotated CDS and included in the gene counts. Metagene start- and stop-centric profiles for Ribo-Seq data exhibited clear triplet periodicity (Supplementary Fig. S1). Supplementary Table S1 provides an overview of mapped reads as well as gene-level read counts for all experiments and descriptive statistics on the generated sequencing data.

3.2. Examples of data illustrating differential transcription and translation in wt and knockout strains

Two different gene cassettes, MATα and MATa, either of which can be present in the MAT locus of S.cerevisiae genome, define the mating type of yeast. Each mating-specific cassette encodes two transcripts directed from the opposite DNA strands by a shared bidirectional promoter: either MATa1 and MATa2 in a-type strains, or MATα1 and MATα2 in α-type strains (Fig. 1) [9]. All of the transcripts except MATa2 encode functional proteins – transcription factors that determine mating type or diploid phenotype (reviewed in [19,20]). While MATa2 is considered to be non-functional [21,22], it nevertheless contains two ORFs (Fig. 1), presumably originating from an original coding region with similarity to MATα2 via an internal frameshift [9,23]. The second ORF could still encode a remarkably conserved amino acid sequence with a high similarity to a portion of MATα2 [9,24] that represents its DNA-binding domain [25,26]. The corresponding protein (MATa2-2) could compete with MATα2 for DNA binding or even have its own transcription factor activity [27]. However, its synthesis should be inhibited by presence of the first ORF in the MATa2 mRNA, which can be regarded as an uORF for the MATa2-2 coding region. Since TMA20 and TMA64 knockout strains have an upregulated translation re-initiation and/or readthrough activities [2,16,28], it was interesting to illustrate our ribosome profiling data with a footprint coverage of MATa locus present in BY4741 strain derivatives. As the sequenced S288C strain is MATα, Ribo-Seq and RNA-Seq reads were re-mapped to MATa locus sequence taken from GenBank (accession number V01313.1)[9]. Fig. 1 provides data on Ribo-Seq and RNA-Seq read coverage of MATa locus of the studied yeast strains. This data can be used to explain the role played by TMA20 and TMA64 translation factors in mating and sporulation programs [29–32].

Another example involves GCN4, the global transcriptional regulator, which is activated during amino acid starvation. Expression of the GCN4 mRNA is controlled by a peculiar mechanism based on differential translation re-initiation on four short uORFs in its 5’ UTR (reviewed in [33]). Fig. 2 shows the 5’ proximal region of the GCN4 transcript, with differential ribosome footprint coverage of uORFs in different strains. Our data can be used for further investigation of TMA20 and TMA64 roles in uORF-mediated translational control of stress response.

APC4, the gene encoding a subunit of anaphase-promoting complex, is located in the same genetic locus as TMA64 and shares a 238-bp promoter region with it. The corresponding mRNAs are synthesized from opposite DNA strands. In the Δtma64 and ΔΔtma20tma64 strains the TMA64 coding sequence was replaced with G-418 or hygromycin resistance gene cassettes (KanMX or HygMX), respectively. Fig. 3 shows differential RNA-Seq coverage of the 238 bp region, flanked by segments of the APC4 and TMA64 coding regions or KanMX/HygMX cassettes, in different yeast strains. This data may likely account for the observed strong genetic interactions of TMA64 with genes involved in the cell cycle and mitotic checkpoints [34] and cell cycle abnormalities of TMA64 knockout strains [35].
Acknowledgments

We are grateful to Alan Hinnebusch (NIH) and Nicholas Guydosh (NIH) for providing the knockout yeast strains and helpful discussion. Ribosome profiling and library sequencing were supported by the Russian Science Foundation (grant RSF 18-14-00291 to S.E.D.). Adaptation of the visualization software was supported by the Russian Foundation for Basic Research (grant 18-34-20024 to I.V.K.). Bioinformatics pipeline was supported by the Program of Fundamental Research for State Academies for 2013–2020 years (No. 01201363825).

Transparency document. Supporting information

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.dib.2019.103701.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.dib.2019.103701.

References

[1] N.T. Ingolia, S. Ghaemmaghami, J.R. Newman, J.S. Weissman, Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling, Science 324 (2009) 218–223.
[2] D.J. Young, D.S. Makeeva, F. Zhang, A.S. Anisimova, E.A. Stolboushkin, F. Ghobakhlou, I.N. Shatsky, S.E. Dmitriev, A.G. Hinnebusch, N.R. Guydosh, Tma64/eIF2D, Tma20/MCT-1, and Tma22/DENR recycle post-termination 40S subunits in vivo, Mol. Cell 71 (2018) 761–774 (e765).
[3] M.V. Gerashchenko, V.N. Gladyshev, Translation inhibitors cause abnormalities in ribosome profiling experiments, Nucleic Acids Res 42 (2014) e134.
[4] M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads, EMBO Journal 2011 (17) (2011) 3.
[5] A. Dobin, C.A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T.R. Gingeras, STAR: ultrafast universal RNA-seq aligner, Bioinformatics 29 (2013) 15–21.
[6] A.A. Egorov, svist4get: a simple visualization tool for genomic tracks from sequencing experiments, (2019): https://bitbucket.org/artegorov/svist4get/.
[7] H. Li, B. Handsaker, A. Wyssoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, The Sequence Alignment/Map format and SAMtools, Bioinformatics 25 (2009) 2078–2079.
[8] A.R. Quinlan, I.M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features, Bioinformatics 26 (2010) 841–842.
[9] C.R. Astell, L. Ahlstrom-Jonasson, M. Smith, K. Tatchell, K.A. Nasmyth, B.D. Hall, The sequence of the DNAs coding for the mating-type loci of Saccharomyces cerevisiae, Cell 27 (1981) 15–23.
[10] F. Miura, N. Kawaguchi, J. Sese, A. Toyoda, M. Hattori, S. Morishita, T. Ito, A large-scale full-length cDNA analysis to explore the budding yeast transcriptome, Proc. Natl. Acad. Sci. USA 103 (2006) 17846–17851.
[11] S.E. Dmitriev, I.M. Terenin, D.E. Andreev, P.A. Ivanov, J.E. Dunaevsky, W.C. Merrick, I.N. Shatsky, GTP-independent tRNA delivery to the ribosomal P-site by a novel eukaryotic translation factor, J Biol. Chem. 285 (2010) 26779–26787.
[12] M.A. Skabkin, O.V. Skabkina, V. Dhote, A.A. Komar, C.U. Hellen, T.V. Pestova, Activities of Ligatin and MCT-1/DENR in eukaryotic translation initiation and ribosomal recycling, Genes Dev. 24 (2010) 1787–1801.
[13] M.A. Skabkin, O.V. Skabkina, C.U. Hellen, T.V. Pestova, Reinitiation and other unconventional post-termination events during eukaryotic translation, Mol. Cell 51 (2013) 249–264.
[14] S. Schleich, K. Strasserburg, P.C. Janiesch, T. Koledachkina, K.K. Miller, K. Hanke, Y.S. Cheng, K. Kuchler, G. Stoecklin, K.E. Duncan, A.A. Teleman, DENR-MCT-1 promotes translation re-initiation downstream of uORFs to control tissue growth, Nature 512 (2014) 208–212.
[15] S. Schleich, J.M. Acevedo, K. Clemm von Hohenberg, A.A. Teleman, Identification of transcripts with short stuORFs as targets for DENR’MCTS1-dependent translation in human cells, Sci. Rep. 7 (2017) 3722.
[16] V. Torrance, D. Lydall, Overlapping open reading frames strongly reduce human and yeast STN1 gene expression and affect telomere function, PLoS Genet. 14 (2018) e1007523.
[17] M. Weisser, T. Schafer, M. Leibundgut, D. Bohringer, C.H.S. Aylett, N. Ban, Structural and functional insights into human re-initiation complexes, Mol. Cell 67 (2017) 447–456 (e447).
[18] J.B. Lomakin, E.A. Stolboushkin, A.T. Vaidya, C. Zhao, M.B. Garber, S.E. Dmitriev, T.A. Steitz, Crystal structure of the human ribosome in complex with DENR-MCT-1, Cell Rep. 20 (2017) 521–528.
[19] J.E. Haber, Mating-type gene switching in Saccharomyces cerevisiae, Annu. Rev. Genet 32 (1998) 561–599.
