





## DATA NOTE

# Ribo-Seq and RNA-Seq of *TMA46 (DFRP1)* and *GIR2 (DFRP2)* knockout yeast strains [version 1; peer review: 3 approved]

Artyom A. Egorov <sup>1-3</sup>, Desislava S. Makeeva<sup>2,3</sup>, Nadezhda E. Makarova<sup>2-4</sup>, Dmitri A. Bykov<sup>2,5</sup>, Yanislav S. Hrytseniuk<sup>2,6</sup>, Olga V. Mitkevich<sup>7</sup>, Valery N. Urakov<sup>7</sup>, Alexander I. Alexandrov<sup>7</sup>, Ivan V. Kulakovskiy<sup>2,3,8</sup>, Sergey E. Dmitriev <sup>2-5</sup>

<sup>1</sup>Phystech School of Biological and Medical Physics, Moscow Institute of Physics and Technology (State University), Dolgoprudny, 141700, Russian Federation

<sup>2</sup>Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Moscow, 119234, Russian Federation

<sup>3</sup>Sirius University of Science and Technology, Sochi, 354340, Russian Federation

<sup>4</sup>Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University, Moscow, 119234, Russian Federation

<sup>5</sup>Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991, Russian Federation

<sup>6</sup>Department of Chemistry, Lomonosov Moscow State University, Moscow, 119234, Russian Federation

<sup>7</sup>Bach Institute of Biochemistry, FRC of Biotechnology of the Russian Academy of Sciences, Moscow, 119071, Russian Federation

<sup>8</sup>Institute of Protein Research, Russian Academy of Sciences, Pushchino, 142290, Russian Federation

**V1** First published: 16 Nov 2021, 10:1162  
<https://doi.org/10.12688/f1000research.74727.1>

Latest published: 16 Nov 2021, 10:1162  
<https://doi.org/10.12688/f1000research.74727.1>

## Abstract

In eukaryotes, stalled and collided ribosomes are recognized by several conserved multicomponent systems, which either block protein synthesis *in situ* and resolve the collision locally, or trigger a general stress response. Yeast ribosome-binding GTPases RBG1 (DRG1 in mammals) and RBG2 (DRG2) form two distinct heterodimers with TMA46 (DFRP1) and GIR2 (DFRP2), respectively, both involved in mRNA translation. Accumulated evidence suggests that the dimers play partially redundant roles in elongation processivity and resolution of ribosome stalling and collision events, as well as in the regulation of GCN1-mediated signaling involved in ribosome-associated quality control (RQC). They also genetically interact with SLH1 (ASCC3) helicase, a key component of RQC trigger (RQT) complex disassembling collided ribosomes. Here, we present RNA-Seq and ribosome profiling (Ribo-Seq) data from *S. cerevisiae* strains with individual deletions of the *TMA46* and *GIR2* genes. Raw RNA-Seq and Ribo-Seq data as well as gene-level read counts are available in NCBI Gene Expression Omnibus (GEO) repository under GEO accession [GSE185458](#) and [GSE185286](#).

## Keywords

Transcriptome, translome, ribosome profiling, ribosome stalling, ribosome collision, *Saccharomyces cerevisiae*, TMA46, GIR2, GCN1/GCN20, eIF2A, YGR054W, STM1, PUB1

## Open Peer Review

Reviewer Status 

Invited Reviewers

	1	2	3
<b>version 1</b>			
16 Nov 2021	report	report	report

1. **Anton A. Komar** , Cleveland State University, Cleveland, USA
2. **Olivier Namy**, Université Paris-Saclay, Gif-sur-Yvette, France
3. **Nikolay E. Shirokikh** , The Australian National University, Canberra, Australia  
**Katrina Woodward**, The Australian National University, Canberra, Australia

Any reports and responses or comments on the article can be found at the end of the article.

**Corresponding authors:** Ivan V. Kulakovskiy ([ivan.kulakovskiy@gmail.com](mailto:ivan.kulakovskiy@gmail.com)), Sergey E. Dmitriev ([sergey.dmitriev@belozersky.msu.ru](mailto:sergey.dmitriev@belozersky.msu.ru))

**Author roles:** **Egorov AA:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Makeeva DS:** Investigation, Methodology, Validation; **Makarova NE:** Investigation, Methodology; **Bykov DA:** Investigation; **Hrytseniuk YS:** Investigation; **Mitkevich OV:** Investigation, Methodology; **Urakov VN:** Investigation; **Alexandrov AI:** Investigation, Methodology, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing; **Kulakovskiy IV:** Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Software, Supervision, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; **Dmitriev SE:** Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Validation, Writing – Original Draft Preparation, Writing – Review & Editing

**Competing interests:** No competing interests were disclosed.

**Grant information:** Analyses of the tma46Δ, gir2Δ, stm1Δ, and ygr054wΔ strains were supported by the Russian Foundation for Basic Research (grant 19-34-51047 to S.E.D.) and the Ministry of Science and Higher education of the Russian Federation (to A.I.A, V.N.U, O.V.M). The pub1Δ strain analyses were supported by the Russian Science Foundation (grant 18-14-00291 to S.E.D.).  
*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

**Copyright:** © 2021 Egorov AA *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**How to cite this article:** Egorov AA, Makeeva DS, Makarova NE *et al.* **Ribo-Seq and RNA-Seq of TMA46 ( DFRP1) and GIR2 ( DFRP2) knockout yeast strains [version 1; peer review: 3 approved]** F1000Research 2021, 10:1162  
<https://doi.org/10.12688/f1000research.74727.1>

**First published:** 16 Nov 2021, 10:1162 <https://doi.org/10.12688/f1000research.74727.1>

## Introduction

Here, we present Ribo-Seq and RNA-Seq data for *S. cerevisiae* strains lacking translation-associated proteins Tma46 and Gir2, as well as for the wild type BY4742 parent strain. Tma46 and Gir2 are yeast orthologs of two mammalian DRG family regulatory proteins: DFRP1 and DFRP2, respectively.<sup>1–3</sup> Mammalian DFRP1 and DFRP2 are binding partners of two closely related proteins, developmentally regulated GTPases DRG1 and DRG2,<sup>1,4</sup> while yeast Tma46 and Gir2 dimerize with their orthologs, the ribosome-binding GTPases Rbg1 and Rbg2, respectively.<sup>2,3,5</sup> Thus, in both yeast and mammals, two distinct heterodimers exist, RBG1•TMA46 (DRG1•DFRP1) and RBG2•GIR2 (DRG2•DFRP1), although under some conditions RBG1 may interact with GIR2 as well.<sup>5</sup> The RBG1 (DRG1) containing complexes associate with mono- and polysomes.<sup>2,3,5–7</sup> Using 5P-Seq, it was recently shown that RBG1•TMA46 promotes efficient translation in yeast, alleviating ribosome pausing at Arg/Lys-rich regions.<sup>7</sup> In contrast, the RBG2 (DRG2) containing dimers are not bound to ribosomes under normal conditions.<sup>3,6</sup> However, they are also clearly related to translation, as GIR2 interacts with the ribosome-bound GCN1, and RBG2•GIR2 is responsible for efficient cell growth under amino acid starvation.<sup>5,8</sup> GCN1 is a large protein necessary for activation of GCN2, the evolutionary conserved eIF2 kinase.<sup>9</sup> Recently, the RBG2•GIR2 complex was detected on the leading stalled ribosome on the Cryo-EM reconstruction of a GCN1-disome complex.<sup>10</sup> These results suggest that GIR2 is a physical linker between RBG2 and GCN1 and that this interaction could prevent excessive activation of the GCN2 pathway upon incidental ribosome stalling.

Interestingly, neither the yeast *rbg1Δ* or *rbg2Δ* knockout strains nor the double *rbg1Δrbg2Δ* mutants display any defects in translation or cell growth.<sup>3</sup> However, a genetic screen for triple synthetic interaction demonstrates that RBGs have redundant function with SLH1,<sup>3</sup> an RNA helicase involved in ribosome-associated quality control (RQC). SLH1 is an ortholog of mammalian ATCC3, a component of the ASC-1 complex that disassembles collided ribosomes (see 11 and references therein).

Taken together, the above data suggest that the RBG1•TMA46 (DRG1•DFRP1) and RBG2•GIR2 (DRG2•DFRP1) complexes play a role in elongation processivity and resolution of ribosome stalling and collision events, as well as in control of GCN1-mediated signaling accompanying these processes. However, many questions remain unanswered. In particular, the individual roles of the two distinct complexes are still unclear. To improve our understanding of their functions, we systematically characterized translational defects in *S. cerevisiae* strains with individual deletions of the *TMA46* or *GIR2* genes using ribosome profiling.<sup>12</sup> We present RNA-Seq and Ribo-Seq data for the yeast *tma46Δ* and *gir2Δ* knockout strains. For comparison, we also provide corresponding data for three strains bearing deletions of other translation-related genes: *STM1*, *PUB1* and *YGR054W* (encoding translation factor eIF2A), as well as for the wild type BY4742 parent strain. Raw sequencing data are available online in the NCBI Gene Expression Omnibus (GEO accession: GSE185458 and GSE185286).

## Materials and methods

### Yeast strains, cell maintenance, RNA-Seq and Ribo-Seq library preparation

RNA-Seq and Ribo-Seq cDNA libraries were prepared from total RNA samples or ribosome-bound RNA samples, respectively, for the wild-type BY4742 (*MATa his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0*) yeast strain and five knockout strains. The data were obtained in two independent series. The first one included the wild-type BY4742 yeast strain (*wt*) and strains with individually deleted *YGR054W* (eIF2A), *STM1*, and *TMA46* genes, which were obtained from the Yeast Gene Deletion Collection.<sup>13</sup> The second one included the wild-type BY4742 yeast strain and strains lacking *GIR2* and *PUB1*, which were created during this study. KanMX disruption cassettes were obtained via polymerase chain reaction (PCR) using the primers (*GIR2* - 5'-CAATTGTAAAATCAGCAGGCA-3'; 5'-AAACTTGTCTATTTCTCTCTC-3'; *PUB1* - 5'-TTGTCCTTCATTTTCTCTCGT-3'; 5'-AGGCCCTTTTATTTTCGAGC-3') and genomic DNA of the corresponding deletion collection strains as a template. Phusion HF polymerase (TFS F-530L) was used as recommended by the supplier. PCR protocol: 95°C – 5 min, [95°C – 30 sec, 64°C – 30 sec, 72°C – 60 sec] x 28 cycles, 72°C – 90 sec. The resulting cassettes were used to transform the BY4742 strain. The correctness of the integration was checked using pairs of verification primers (*GIR2*-5'-GAAAAAGAAAGAAGAAAAATTTGGG-3'; *PUB1*-5'-ACGACCACAAAGGATCCAGGGCTT-3'; Universal primer inside the KanMX cassette - 5'-CTGCAGCGAGGAGCCGTAAT-3').

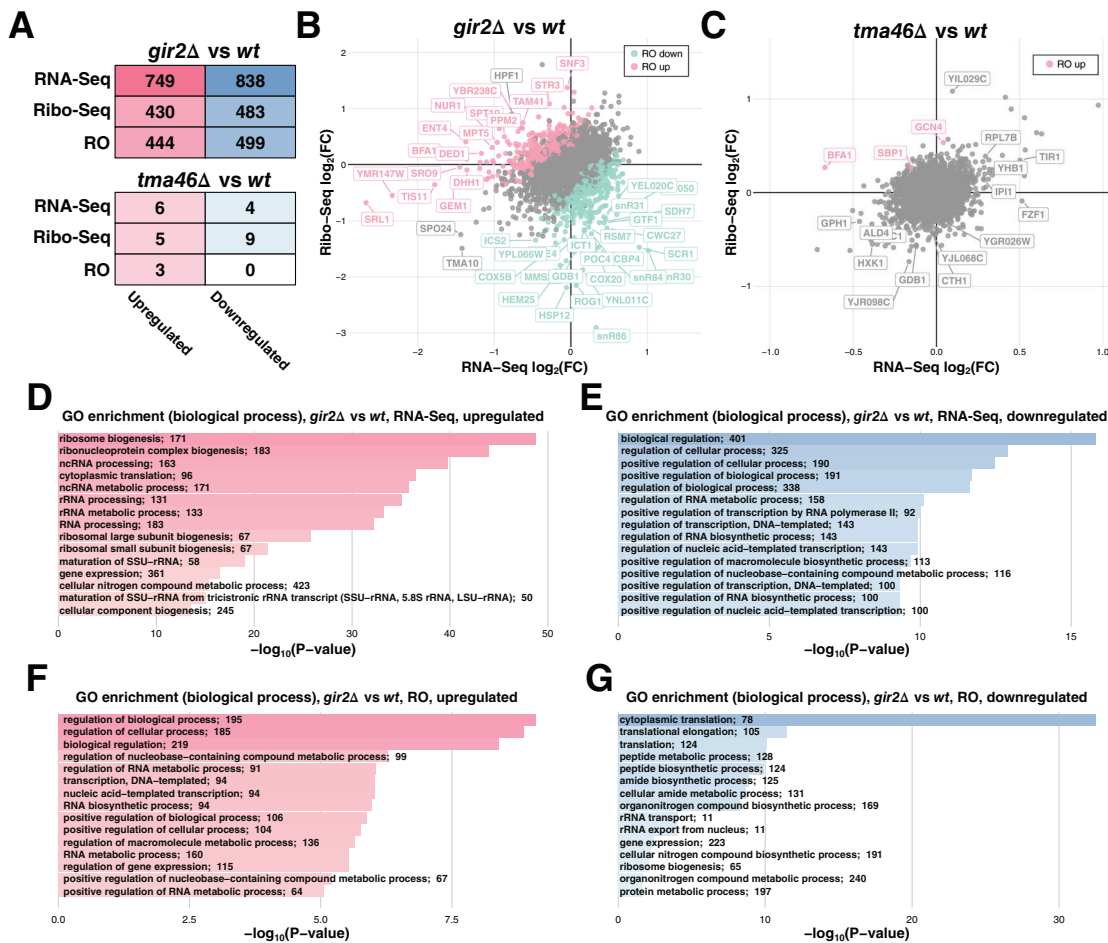
Here we focus on *wt*, *tma46Δ*, and *gir2Δ* strains. The data from the other strains were used to correct for batch effects within each series. The libraries were sequenced, resulting in 31 RNA-Seq and 28 Ribo-Seq data sets, including 18 RNA-Seq and 16 Ribo-Seq data sets for *wt*, *gir2Δ*, and *tma46Δ* strains. Supplementary Table 1 in the *Extended data*<sup>14</sup> summarizes information about the sequencing experiments.

The experimental procedure followed the ribosome profiling protocol described in.<sup>15</sup> Briefly, yeast cells were grown to exponential phase (OD = 0.5-0.6) in yeast extract peptone dextrose (YPD) media (1% yeast extract, 2% peptone, 2% glucose). Cells were harvested by filtration, scraped into liquid nitrogen, and ground using a liquid nitrogen-cooled

mortar and pestle with drop-by-drop addition of polysome lysis buffer (20 mM Tris-HCl pH 8.0, 140 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.1 mg/ml cycloheximide, 1% Triton). Cell lysates were clarified by two sequential centrifugation steps - 3000g, 5 minutes, 4°C, and 20000g, 10 minutes, 4°C. The cell lysate was partially used for mRNA isolation using oligo (dT) beads. Another portion was treated with ribonuclease I for polysome disassembly and applied to a linear 10-50% sucrose gradient in fractionation buffer (20 mM Tris pH 8.0, 140 mM KCl, 15 mM MgCl<sub>2</sub>, 1 mM DTT, 0.1 mg/ml cycloheximide, 1% Triton) and separated on a SW-41 rotor (Beckman) at 35000 rpm, 3 hours, 4°C. Subsequently, ribosome-bound RNA fragments were collected from the monosome fraction. Ribosome-bound RNA was isolated using acidic-phenol extraction. Further Ribo-Seq and RNA-Seq library preparations were performed as described previously.<sup>12</sup>

**Sequencing data processing and analysis**

Reads were trimmed using cutadapt v. 2.10<sup>16</sup> with the following parameters for RNA-Seq (-a AGATCGGAAGAGCA CACGTCTGAACTCCAGTCAC -minimum-length 20 -q 20) and Ribo-Seq samples (-a CTGTAGGCACCATCAA TAGATCGGAAGAGCACACGTCTGAACTCCAGTCAC -trimmed-only -q 20). Additionally, for Ribo-Seq, the reads were deduplicated with seqkit mdup v. 0.10.1,<sup>17</sup> and unique barcodes were then removed with cutadapt v. 2.10 (-q 20 -minimum-length 20 -u -4). Afterwards, reads were aligned against eukaryotic rRNA sequence set obtained from silva-euk<sup>18</sup> and rfam<sup>19</sup> databases using bowtie2 v. 1.2.3.<sup>20</sup> Only unmapped non-rRNA reads were used in the further analysis. Read mapping and counting against the *Saccharomyces cerevisiae*.R64-1-1.95 (Ensembl)<sup>21</sup> genome assembly was performed with STAR v. 2.7.9a.<sup>22</sup> We estimated the position of the P-site for each dataset from the 5' end of the reads



**Figure 1. Differential gene expression upon deletions of *GIR2* and *TMA46*.** (A) Total number of differentially expressed genes (passing FDR < 0.05) in each test. (B) and (C) Scatter plots illustrating expression changes at the level of transcription (X-axis, RNA-Seq) and translation (Y-Axis, Ribo-Seq) in the *gir2Δ* and *tma46Δ* strains, respectively. Translationally upregulated or downregulated (FDR < 0.05) genes marked in pink/blue, respectively. (D-G) Results of Gene Ontology (GO) enrichment analysis of RNA-Seq (D-E) and ribosome occupancy (RO) (F-G) considering upregulated (D, F) and downregulated (E, G) genes in *gir2Δ vs wt* comparison. The numbers of differentially expressed genes with particular GO terms are shown in labels. X-axes show the enrichment P-value in log-scale.

on the basis of the length of each footprint using plastid v0.5.1.<sup>23</sup> Fraction of reads in each phase and read length distribution were also obtained with plastid, see Figure S1 in the *Extended data*.<sup>24</sup> The results show that almost 90% 28nt reads are in 0 phase through the annotated coding sequences (CDSs). Then we produced BedGraph profiles from SAM data with *samtools* v. 1.10<sup>25</sup> and *bedtools* v2.27.1.<sup>26</sup> Coverage profiles were normalized using normalization factor and library size estimates from differential expression analysis (see below) separately for each bedGraph profile. Finally, we visualized coverage tracks in the modified genomic loci using *svist4get*.<sup>27</sup> Figure S2 in the *Extended data*<sup>24</sup> shows that the read counts originating from the mRNA encoded by the knockout gene in the corresponding strain are negligible. The coverage of the neighboring genes remains unaltered, i.e. there are no indications of the so-called neighboring gene effect (NGE).<sup>28</sup>

### Differential expression and Gene Ontology (GO) enrichment analysis

Statistical analyses were performed in R v. 4.1.2 using edgeR Bioconductor package.<sup>29</sup> As mentioned above, the data were produced in two independent series which were analyzed separately. Genes not reaching 10 read count per million (CPM) in at least 4 RNA-Seq and 4 Ribo-Seq libraries were excluded from the analysis. Then, we performed the batch correction using ComBat-seq R package.<sup>30</sup> Principal component analysis (PCA) plots of the raw and batch corrected expression profiles are shown in Figure S3 in the *Extended data*.<sup>24</sup> A generalized linear model (*glmQLFit*, *glmQLFTest* of the edgeR package) was used to detect differentially expressed genes (for RNA-Seq, Ribo-Seq, and ribosome occupancy (RO) defined as the Ribo-Seq coverage of a CDS normalized to its RNA-Seq coverage) with the strain as a categorical variable. The false discovery rate (FDR < 0.05) was used for identification of differential expressed genes. We also performed Gene Ontology (GO) enrichment analysis for upregulated and downregulated genes with *yeastmine*.<sup>31</sup> The results are shown in Figure 1.

### Data availability

#### Underlying data

NCBI Gene Expression Omnibus: RNA Sequencing and Ribosome profiling of *TMA46*, *STM1* and *YGR054W* knockout yeast strains. Accession number [GSE185458](https://identifiers.org/geo:GSE185458); <https://identifiers.org/geo:GSE185458>.

NCBI Gene Expression Omnibus: RNA Sequencing and Ribosome profiling of *GIR2* and *PUB1* knockout yeast strains. Accession number [GSE185286](https://identifiers.org/geo:GSE185286); <https://identifiers.org/geo:GSE185286>.

#### Extended data

Figshare: Supplementary Table 1.csv <https://doi.org/10.6084/m9.figshare.16818505>.<sup>14</sup>

This project contains the following extended data:

- Supplementary Table 1.csv (Table with information about sequencing experiments)

Figshare: Supplementary Figures. <https://doi.org/10.6084/m9.figshare.16818610.v1>.<sup>24</sup>

Data are available under the terms of the [Creative Commons Zero “No rights reserved” data waiver](https://creativecommons.org/licenses/by/4.0/) (CC0 1.0 Public domain dedication).

### Acknowledgments

A.A.E., N.E.M., D.A.B., and S.E.D. are members of the Interdisciplinary Scientific and Educational School of Moscow University “Molecular Technologies of the Living Systems and Synthetic Biology”.

### References

1. Ishikawa K, Azuma S, Ikawa S, *et al.*: **Identification of DRG family regulatory proteins (DFRPs): specific regulation of DRG1 and DRG2.** *Genes Cells.* 2005; **10**: 139–150. [PubMed Abstract](#) | [Publisher Full Text](#)
2. Fleischer TC, Weaver CM, McAfee KJ, *et al.*: **Systematic identification and functional screens of uncharacterized proteins associated with eukaryotic ribosomal complexes.** *Genes Dev.* 2006; **20**: 1294–1307. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
3. Daugeron MC, Prouteau M, Lacroute F, *et al.*: **The highly conserved eukaryotic DRG factors are required for efficient translation in a manner redundant with the putative RNA helicase Slh1.** *Nucleic Acids Res.* 2011; **39**: 2221–2233. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

4. Li B, Trueb B: **DRG represents a family of two closely related GTP-binding proteins.** *Biochim. Biophys. Acta.* 2000; **1491**: 196–204.  
[PubMed Abstract](#) | [Publisher Full Text](#)
5. Wout PK, Sattlegger E, Sullivan SM, et al.: **Saccharomyces cerevisiae Rbg1 protein and its binding partner Gir2 interact on Polyribosomes with Gcn1.** *Eukaryot. Cell.* 2009; **8**: 1061–1071.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
6. Ishikawa K, Akiyama T, Ito K, et al.: **Independent stabilizations of polysomal Drg1/Dfrp1 complex and non-polysomal Drg2/Dfrp2 complex in mammalian cells.** *Biochem. Biophys. Res. Commun.* 2009; **390**: 552–556.  
[PubMed Abstract](#) | [Publisher Full Text](#)
7. Zeng F, Li X, Pires-Alves M, et al.: **Conserved heterodimeric GTPase Rbg1/Tma46 promotes efficient translation in eukaryotic cells.** *bioRxiv.* 2021.2020.2007.2006.190082.
8. Ishikawa K, Ito K, Inoue J, et al.: **Cell growth control by stable Rbg2/Gir2 complex formation under amino acid starvation.** *Genes Cells.* 2013; **18**: 859–872.  
[PubMed Abstract](#) | [Publisher Full Text](#)
9. Castilho BA, Shanmugam R, Silva RC, et al.: **Keeping the eIF2 alpha kinase Gcn2 in check.** *Biochim. Biophys. Acta.* 2014; **1843**: 1948–1968.  
[PubMed Abstract](#) | [Publisher Full Text](#)
10. Pochopien AA, Beckert B, Kasvandik S, et al.: **Structure of Gcn1 bound to stalled and colliding 80S ribosomes.** *Proc. Natl. Acad. Sci. U. S. A.* 2021; **118**: e2022756118.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
11. Juszkiwicz S, Speldewinde SH, Wan L, et al.: **The ASC-1 Complex Disassembles Collided Ribosomes.** *Mol. Cell.* 2020; **79**: 603–614 e608.  
[Publisher Full Text](#)
12. Ingolia NT, Ghaemmaghami S, Newman JR, et al.: **Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling.** *Science.* 2009; **324**: 218–223.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
13. Giaever G, Nislow C: **The yeast deletion collection: a decade of functional genomics.** *Genetics.* 2014; **197**: 451–465.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
14. Egorov A: **Supplementary Table 1.csv.** *figshare. Dataset.* 2021.  
[Publisher Full Text](#)
15. Kasari V, Pochopien AA, Margus T, et al.: **A role for the Saccharomyces cerevisiae ABCF protein New1 in translation termination/recycling.** *Nucleic Acids Res.* 2019; **47**: 8807–8820.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
16. Martin M: **Cutadapt removes adapter sequences from high-throughput sequencing reads.** 2011; **17**: 3.2011.
17. Shen W, Le S, Li Y, et al.: **SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation.** *PLoS One.* 2016; **11**: e0163962.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
18. Quast C, Pruesse E, Yilmaz P, et al.: **The SILVA ribosomal RNA gene database project: improved data processing and web-based tools.** *Nucleic Acids Res.* 2013; **41**: D590–D596.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
19. Kalvari I, Nawrocki EP, Ontiveros-Palacios N, et al.: **Rfam 14: expanded coverage of metagenomic, viral and microRNA families.** *Nucleic Acids Res.* 2021; **49**: D192–D200.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
20. Langmead B, Salzberg SL: **Fast gapped-read alignment with Bowtie 2.** *Nat. Methods.* 2012; **9**: 357–359.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
21. Zerbino DR, Achuthan P, Akanni W, et al.: **Ensembl 2018.** *Nucleic Acids Res.* 2018; **46**: D754–D761.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
22. Dobin A, Davis CA, Schlesinger F, et al.: **STAR: ultrafast universal RNA-seq aligner.** *Bioinformatics.* 2013; **29**: 15–21.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
23. Dunn JG, Weissman JS: **Plastid: nucleotide-resolution analysis of next-generation sequencing and genomics data.** *BMC Genom.* 2016; **17**: 958.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
24. Egorov A: **Supplementary Figures.** *figshare. Figure.* 2021.  
[Publisher Full Text](#)
25. Li H, Handsaker B, Wysoker A, et al.: **The Sequence Alignment/Map format and SAMtools.** *Bioinformatics.* 2009; **25**: 2078–2079.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
26. Quinlan AR, Hall IM: **BEDTools: a flexible suite of utilities for comparing genomic features.** *Bioinformatics.* 2010; **26**: 841–842.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
27. Egorov AA, Sakharova EA, Anisimova AS, et al.: **svist4get: a simple visualization tool for genomic tracks from sequencing experiments.** *BMC Bioinform.* 2019; **20**: 113.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
28. Egorov AA, Alexandrov AI, Urakov VN, et al.: **A standard knockout procedure alters expression of adjacent loci at the translational level.** *Nucleic Acids Res.* 2021; **49**: 11134–11144.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
29. Robinson MD, McCarthy DJ, Smyth GK: **edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.** *Bioinformatics.* 2010; **26**: 139–140.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
30. Zhang Y, Parmigiani G, Johnson WE: **ComBat-seq: batch effect adjustment for RNA-seq count data.** *NAR Genom. Bioinform.* 2020; **2**: lqaa078.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
31. Balakrishnan R, Park J, Karra K, et al.: **YeastMine—an integrated data warehouse for Saccharomyces cerevisiae data as a multipurpose tool-kit.** *Database (Oxford).* 2012; **2012**: bar062.

# Open Peer Review

Current Peer Review Status:   

---


Version 1

Reviewer Report 07 December 2021

<https://doi.org/10.5256/f1000research.78509.r100404>

© 2021 Shirokikh N et al. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Nikolay E. Shirokikh** 

Division of Genome Sciences and Cancer, The John Curtin School of Medical Research, The Australian National University, Canberra, Australia

**Katrina Woodward**

Division of Genome Sciences and Cancer, The John Curtin School of Medical Research, The Australian National University, Canberra, Australia

The work by Artyom Egorov and others touches upon proteins implicated in the processivity of translation of mRNA into polypeptides by eukaryotic ribosomes, specifically in resolving the stalled (or 'collided') translational intermediates. Research in this area is accelerating with many new studies uncovering unforeseen depths of the cells' mechanisms of translation fidelity control. Not too many datasets exist in this space to date and few are multi-target studies that provide a package of data obtained in a highly similar fashion for several translation-related genes, which is the rationale that justifies well the generated datasets.

In a series of yeast knock-out strains, the authors accurately employ common approaches of RNA sequencing and RNA-sequencing-based ribosome (translation) footprint profiling, to provide high-throughput short-read sequencing data resulting from the individual gene knock-out effects. The data are provided for the wild-type yeast of BY4742 strain used as a control, the *GIR2* and *TMA46* knock-out mutant strains which are in the focus of the dataset, and additionally the *YGR054W* (*eIF2A*), *STM1*, and *PUB1* knock-out mutant strains, either commercially purchased or generated in the work.

The methods are written succinctly but capture most of the relevant detail where it matters. Perhaps, expanding on the particulars of the method used to isolate mRNA with oligo(dT) beads could add to the better understanding of the RNA-seq data (including any information about the other treatments and purifications used, such as RNA fragmentation and size cut-offs). Regarding the ribosomal material isolation, although not strictly necessary, it would add to the value of the datasets by outlining their applicability if the authors could provide more detail on the intensity of the digestion with RNase I to generate the footprints, the homogeneity of the resultant sedimentation profiles, and the boundaries used to define the monosomal fraction. For the lysate

clarification, the second spin uses very high acceleration; comments about whether the spin does or does not result in any loss of the heavy polysomes (e.g., if a suitably long particle sedimentation path was employed) could be beneficial. Although the authors are referring to the prior works, to bolster the applicability of the datasets and make their uptake convenient, any information about the additional size selection of the ribosome-protected fragments used (biochemical and/or bioinformatic) would be very helpful.

The data appears as accurate. The authors provide some differential abundance and GO term enrichment information for the more and less abundant gene categories within the pairwise comparisons, to characterise the datasets. This is very useful information; more information regarding the general properties of the datasets (such as total read/mappable/unique read number tables or histograms, gene coverage saturation plots etc.) could be further attractive in the main figure. The replicated datasets are validated through principal component analysis plots in Supplementary Figure 3, which show generally good groupings between the replicated sample types. The validity of the ribo-seq data is confirmed by Supplementary Figure 1 through demonstrating triplet periodicity of the dominant-length footprint fraction usually attributed to the cycloheximide-stabilised ribosomal protection extent (28 nt; although this is shown only for the wild-type, *GIR2* and *TMA46* knock-out strains). The datasets are most definitely usable, presented accessibly including uploads to the NCBI Gene Expression Omnibus, and are of a high utility.

Altogether, the authors provide an excellent resource to the field, which perfectly fits the scope of F1000Research Data Note publication, and will likely be cited.

**Is the rationale for creating the dataset(s) clearly described?**

Yes

**Are the protocols appropriate and is the work technically sound?**

Yes

**Are sufficient details of methods and materials provided to allow replication by others?**

Yes

**Are the datasets clearly presented in a useable and accessible format?**

Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Molecular biology, biochemistry, cell biology, genetics and gene expression, RNA, protein biosynthesis, ribosome profiling

**We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

Reviewer Report 01 December 2021

<https://doi.org/10.5256/f1000research.78509.r100401>



© 2021 Namy O. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



### Olivier Namy

CNRS, I2BC, Université Paris-Saclay, Gif-sur-Yvette, France

In this manuscript Egorov A.A and colleagues aim to understand the translational consequences of the inactivation of two genes TMA46 and GIR2 in yeast *S. cerevisiae*. The experiments are clearly detailed and basic quality controls have been done for the RiboSeq analysis. It is not clear for me why the authors provided data for STM1, PUB1, and YGR054W without showing or analyzing them. I have several comments:

- The authors should justify why they have used a generalized linear model to calculate differential expression instead of a robust and well established negative binomial distribution (DESeq2). Why the authors used two different databases (silva-euk and rfam) to align reads to the rRNA? Did they use sequences specific to *S. cerevisiae*?
- Legend of Supplementary Fig1 needs to be improved. In the actual format, it is not obvious how the authors calculate the phasing of their RiboSeq data. Moreover, it is a pity that the analyses were not pursued a little further. For example, triplet periodicity could be shown, ribosome collision, and/or ribosome occupancy calculated. These additional qualitative analyses would strengthen the manuscript.

#### Is the rationale for creating the dataset(s) clearly described?

Yes

#### Are the protocols appropriate and is the work technically sound?

Yes

#### Are sufficient details of methods and materials provided to allow replication by others?

Partly

#### Are the datasets clearly presented in a useable and accessible format?

Yes

**Competing Interests:** No competing interests were disclosed.

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

Reviewer Report 30 November 2021

<https://doi.org/10.5256/f1000research.78509.r100403>

© 2021 Komar A. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Anton A. Komar** 

Center for Gene Regulation in Health and Disease, Cleveland State University, Cleveland, OH, 44115, USA

In this manuscript, Egorov and co-authors presented transcriptome and ribosome profiling data for *Saccharomyces cerevisiae* strains with individual deletions of the *TMA46* and *GIR2* genes. The interest in the protein products encoded by these genes has increased substantially in recent years, as these protein products have been shown to be involved in translation elongation and ribosome stalling. However, to date, there were limited “omics” data available from the respective knockout strains that could help shed additional light on the function of these genes. Particularly, ribosomal profiling of the knockout strains has not been performed. The present study fills in this gap and presents comparative ribosome profiling data obtained using the wild-type and knockout yeast strains. The ribosome profiling was followed by the standard quality checks and some basic downstream analysis, i.e. analysis of the differentially expressed genes and Gene Ontology (GO) enrichment analysis. Overall, these data will be useful for further research on *TMA46* and *GIR2* genes’ functions.

The sample preparation was performed using the standard protocols. The raw and processed data were deposited in the relevant open access repositories such as the Gene Expression Omnibus (GEO). Data processing has been described with reasonable detail (including the description of the command-line tools options and adapter sequences). The reliability of Ribo-Seq data has been demonstrated by read length distribution and the corresponding fraction of reads in each frame (Supplementary figure 1). All in all, the data seem to be technically sound and accurately described, but I would suggest adding illustrative plots showing triplet periodicity of the ribosome footprints.

**Is the rationale for creating the dataset(s) clearly described?**

Yes

**Are the protocols appropriate and is the work technically sound?**

Yes

**Are sufficient details of methods and materials provided to allow replication by others?**

Yes

**Are the datasets clearly presented in a useable and accessible format?**

Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Translational control of gene expression

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

---

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact [research@f1000.com](mailto:research@f1000.com)

**F1000Research**