OXFORD FEMS

FEMS Yeast Research, 2022, **22**, 1–9
DOI: 10.1093/femsyr/foac024
Advance access publication date: 6 May 2022
Protocol

Development of a ribosome profiling protocol to study translation in *Kluyveromyces marxianus*

Darren A. Fenton[1,2], Stephen J. Kiniry[1], Martina M. Yordanova[1], Pavel V. Baranov[1], John P. Morrissey[2,*]

[1]School of Biochemistry and Cell Biology, University College Cork, Cork, T12 XF62, Ireland
[2]School of Microbiology, Environmental Research Institute, APC Microbiome Institute, SUSFERM Fermentation Science Centre, University College Cork, Cork T12 K8AF, Ireland
*Corresponding author: School of Microbiology, University College Cork, Cork T12 K8AF, Ireland. Tel: +353-87-219-6075; E-mail: j.morrissey@ucc.ie
One sentence summary: Development of a ribosome profiling protocol to study gene expression in the thermotolerant yeast *Kluyveromyces marxianus*.
Editor: Carol Munro

Abstract

Kluyveromyces marxianus is an interesting and important yeast because of particular traits such as thermotolerance and rapid growth, and for applications in food and industrial biotechnology. For both understanding its biology and developing bioprocesses, it is important to understand how *K. marxianus* responds and adapts to changing environments. For this, a full suite of omics tools to measure and compare global patterns of gene expression and protein synthesis is needed. We report here the development of a ribosome profiling method for *K. marxianus*, which allows codon resolution of translation on a genome-wide scale by deep sequencing of ribosome locations on mRNAs. To aid in the analysis and sharing of ribosome profiling data, we added the *K. marxianus* genome as well as transcriptome and ribosome profiling data to the publicly accessible GWIPS-viz and Trips-Viz browsers. Users are able to upload custom ribosome profiling and RNA-Seq data to both browsers, therefore allowing easy analysis and sharing of data. We also provide a set of step-by-step protocols for the experimental and bioinformatic methods that we developed.

Keywords: nonconventional yeast (NCY), Ribo-Seq, RNA-Seq, transcriptome visualization, global analysis, industrial biotechnology

Introduction

As with other microbes, yeasts have evolved elaborate mechanisms to sense and respond to changing extracellular and intracellular environments. External influences include phenomena such as altered nutrient availability, toxic molecules, temperature fluctuations, low pH and high osmotic pressure, while internally, cells can experience changes such as reduced intracellular pH, ion fluxes, energy depletion or nutrient starvation (Martínez-Montañés *et al.* 2010, Broach 2012, Ljungdahl and Daignan-Fornier 2012, Morano *et al.* 2012, de la Torre-Ruiz *et al.* 2015, Sui *et al.* 2015, Taymaz-Nikerel *et al.* 2016). The best-studied response mechanisms in yeast involve sensor systems, signal transduction pathways and changes in gene expression (de Nadal and Posas 2010). Ultimately, this gives rise to a new set of proteins that enable the cell to adapt, if necessary, and to survive as well as proliferate in this new environment. Dissecting these response mechanisms is central to understanding the fundamental biology of a species, but it is also important for the development of yeast for biotechnological applications (Liu and Nielsen 2019). Yeasts are used for diverse applications in the food, biopharma and industrial biotechnology sectors (Arevalo-Villena *et al.* 2017, Nandy and Srivastava 2018, Parapouli *et al.* 2020) and, very often, they need to perform under suboptimal conditions or deal with a fluctuating environment. This is a particular problem when scaling engineered yeast cell factories in industrial biotechnology (Takors 2012, Delvigne *et al.* 2014, Wehrs *et al.* 2019). Developing a comprehensive understanding of adaptive responses is a key requirement for the construction of yeast cell factories that are both robust and resilient, and capable of optimal performance in an industrial bioprocess.

Most adaptive responses involve increased or reduced activity of specific proteins, which can be achieved at the level of synthesis, stability or activity. While some specific responses can be at the protein level, for example mediated by allosteric regulation, adaptation usually requires changes in the expression of many genes, and is considered to be a 'global' response. Changes in global gene expression can occur at various levels, most notably via transcription, translation or mRNA stability. Transcriptional changes are due to chromatic restructuring or changes in the activity of particular transcriptional regulators leading to increased or decreased levels of mRNA (Hahn and Young 2011). This is by far the best-understood adaptive process, deployed in response to heat shock (Masser *et al.* 2020), osmotic stress (de Nadal and Posas 2010), oxidative stress (Morano *et al.* 2012) and cell wall challenges (Sanz *et al.* 2017, Jiménez-Gutiérrez *et al.* 2020). Transcriptional responses can be studied at the level of individual genes using northern blots and reverse transcription quantitative real-time PCR (RT-qPCR), or globally by DNA microarrays or massively parallel sequencing (RNA-Seq), the latter of which has become the method of choice to study changes in gene expression (Schena *et al.* 1995, Gibson *et al.* 1996, Wang *et al.* 2009). Translation results in protein synthesis and, as such, is a better indicator of protein levels than transcription, though in many cases, higher abundance of mRNA due to increased transcription leads to a corresponding increase in the amount of translation. This is not always the

Received: February 11, 2022. **Revised:** April 17, 2022. **Accepted:** May 4, 2022
© The Author(s) 2022. Published by Oxford University Press on behalf of FEMS. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

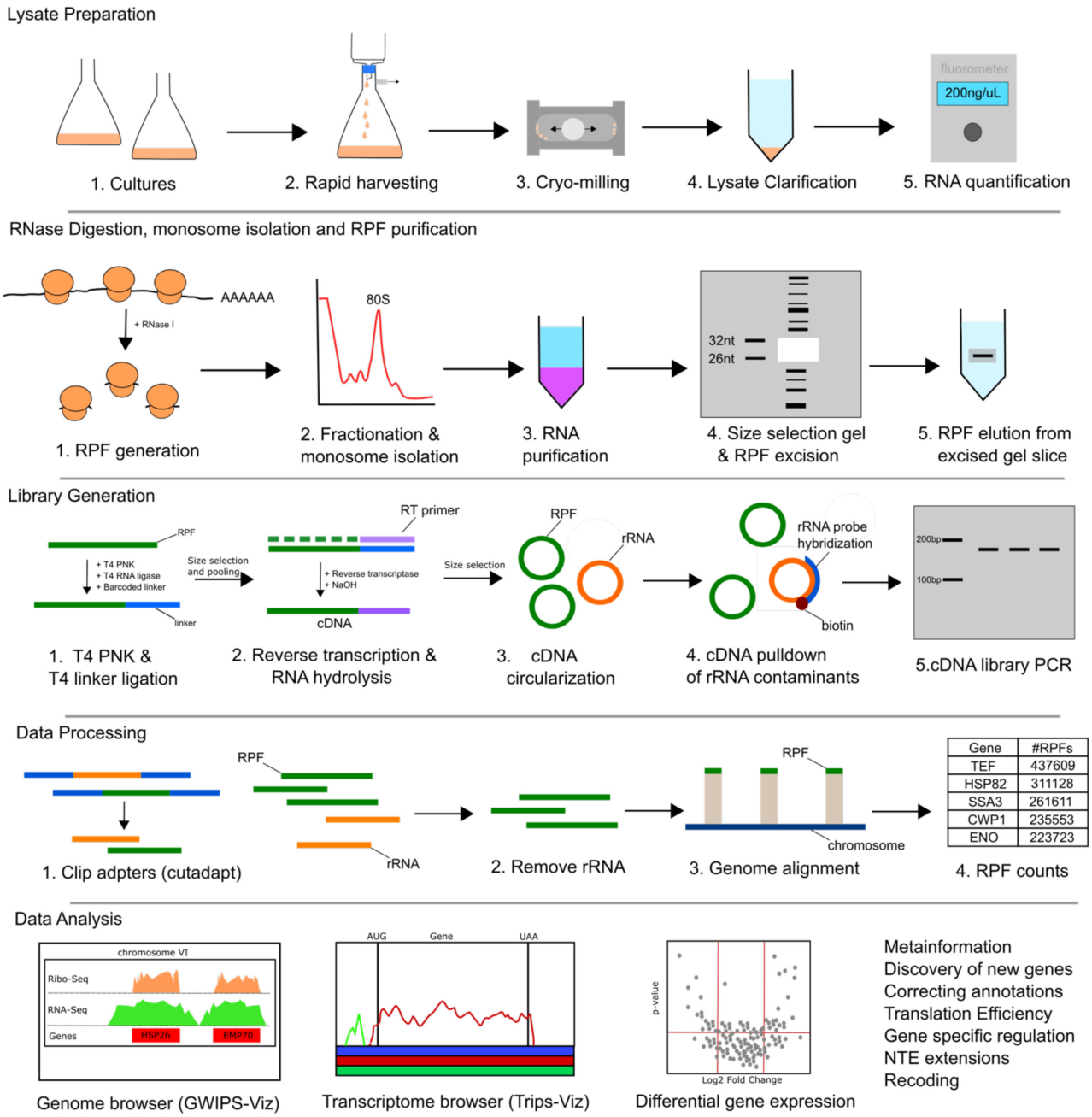

Figure 1. Summary of the ribosome profiling workflow. This summary is broken into five parts: lysate preparation, RPF generation and purification, library generation, data processing and data analysis. Lysate preparation includes culturing, lysis and the quantification of total RNA in a lysate. RNase digestion, monosome isolation and RPF purification represent the generation of RPFs. Library generation involves the conversion of small RNAs (RPFs) to a cDNA library, ready to be sequenced on an Illumina sequencing platform. Data processing involves removing of the sequencing adapters to leave only RPF sequences that are aligned to the genome. Data analysis typically involves visualization of data via genome and/or transcriptome browser, differential gene expression and a range of others as listed in the figure.

case, however, and there are also instances where translation is regulated without any changes in the mRNA abundance. A well-documented example of this in yeast is regulation of the translation of the transcriptional activator encoded by *GCN4*. In this case, short open reading frames (ORFs) upstream of the main *GCN4* coding sequence regulate the rate of *GCN4* translation in response to intracellular amino acid levels (Hinnebusch 2005). The target of rapamycin (TOR) growth control system also mediates some of its effects by regulating translation via controlling access of the small ribosomal subunit to the cap structure at the 5′ end of the mRNA (Merrick 2015). Indeed, as will be mentioned later, there is an increased awareness that translational regulation is a central part of the yeast system for controlling gene expression.

Ribosome profiling, sometimes termed Ribo-Seq, is a method that allows for the visualization and quantification of translation at a global level. First developed in *Saccharomyces cerevisiae* (Ingolia *et al.* 2009), it has since been widely used in bacteria, yeast and mammalian systems for genome-wide studies of translation (Andreev *et al.* 2017, Ingolia *et al.* 2019, Mohammad *et al.* 2019). During mRNA translation, a ribosome translocates an mRNA one codon at a time and protects a fragment of mRNA within its mRNA tunnel (Steitz 1969). Ribosome profiling is a method to

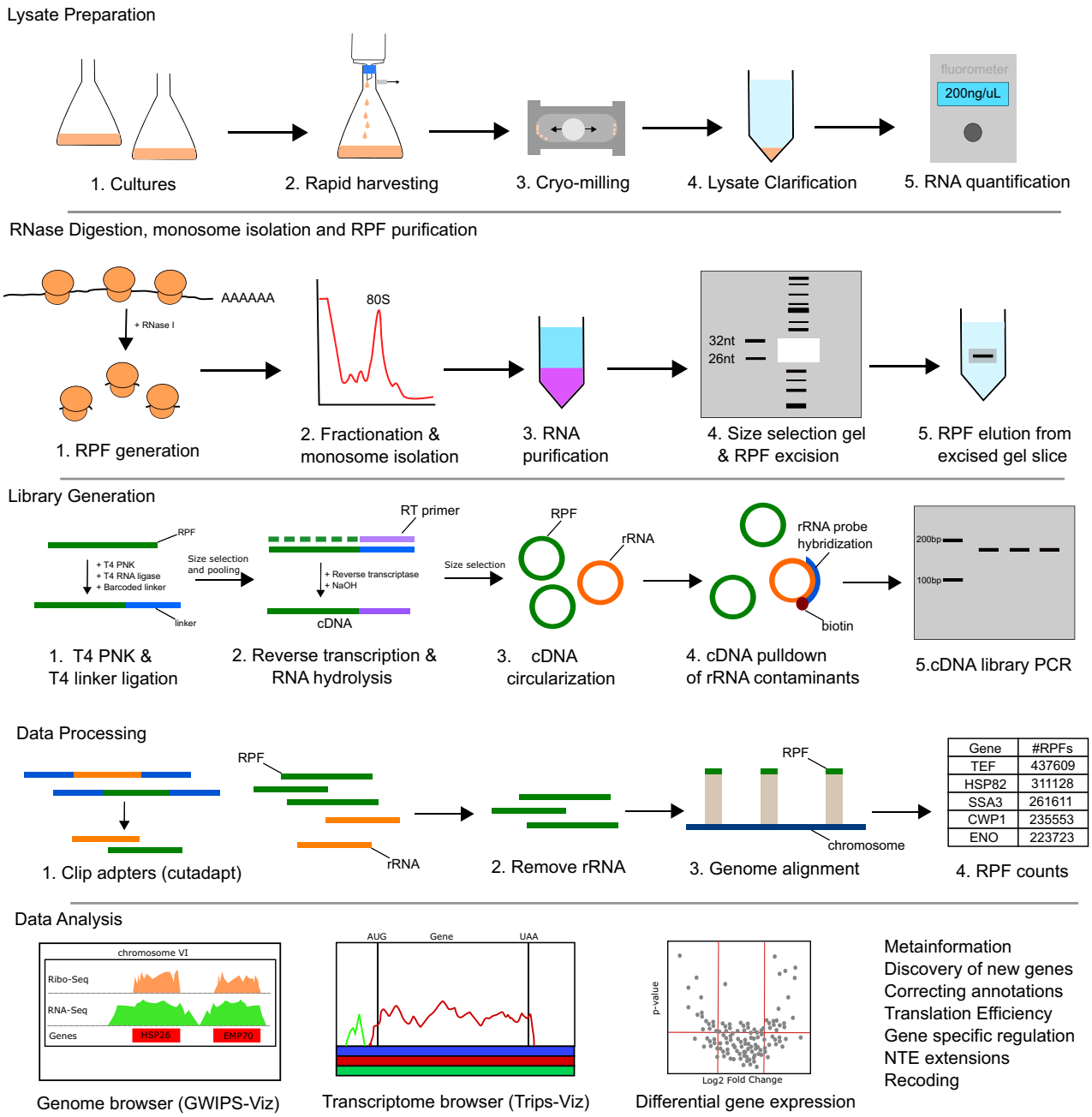

Figure 2. Ribosome profiling data from *K. marxianus*. **(A)** Pearson's correlation of biological replicates for each experimental condition. Axis values represent $\log_2$ read counts. **(B)** Composition of ribosome profiling library with rRNA depletion. **(C)** Triplet periodicity of aligned RPFs for each read length. Panels **(D)** and **(E)** display a metagene profile of aligned RPFs near the start codon and stop codon, respectively.

Table 1. Biotinylated oligos for rRNA depletion. These oligos contain a 5′ biotin modification to allow pulldown of specific rRNA contaminants using magnetic streptavidin beads.

Name	rRNA target	Sequence
rRNA#1	26S	5′AAGGGTGCATCATCGACCGATCCTG
rRNA#2	26S	5′GTTTCTTTACTTATTCAATTAAGCGGA
rRNA#3	Mitochondrial	5′TAAAGAATGGTACAGCTATAAATATT
rRNA#4	18S	5′GCTCGAATATATTAGCATGGAATAATGGA
rRNA#5	26S	5′TATAGAAGGATACGAATAAGGCGTC
rRNA#6	26S	5′TTTCCACGTTCTAGCATTCAAAGTCCT

identify these ribosome-protected fragments (RPFs), thereby reporting which mRNAs are being translated at a given point in time. When applying the method, translation is arrested, usually by the addition of translation inhibitors, ribosomes are isolated and the RPFs are identified by deep sequencing. RNA-Seq is usually carried out in parallel to ribosome profiling, allowing estimation of changes in mRNA translation efficiency (Ingolia *et al.* 2009). In *S. cerevisiae*, ribosome profiling has uncovered widespread translation of upstream ORFs (Ingolia *et al.* 2009), non-AUG initiation at canonical genes (Monteuuis *et al.* 2019, Eisenberg *et al.* 2020) and small translated ORFs throughout the genome (Smith *et al.* 2014). Ribosome profiling combined with RNA-Seq has been useful in deciphering both transcriptional and translation regulation in the

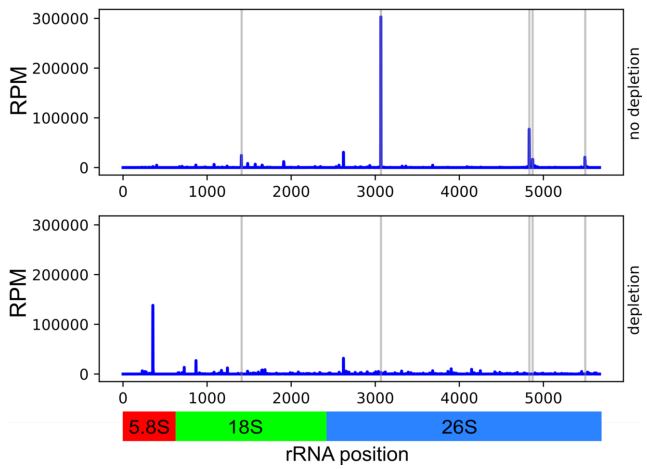

Figure 3. Targeted removal of nuclear-encoded rRNA contaminants. Abundance is represented in reads per million (RPM) and position is relative to the generated rRNA index presented in the bottom track. Top panel represents rRNA composition and abundance with no targeted rRNA depletion employed. Bottom panel represents rRNA composition and abundance with targeted rRNA depletion protocol. Targets for rRNA depletion are highlighted as dark grey areas. Abundance is represented in RPM and position is relative to the generated rRNA index presented in the bottom track. The grey vertical lines highlight the rRNA contaminants that are targeted in the oligo depletion step.

yeast meiotic programme (Brar and Weissman 2015) and in the response to oxidative stress (Blevins *et al.* 2019). Ribosome profiling has also been carried out on a range of other yeast species, including *Saccharomyces paradoxus* (McManus *et al.* 2014), *Schizosaccharomyces pombe* (Duncan and Mata 2014), *Saccharomyces uvarum* (Spealman *et al.* 2018), *Komagataella phaffii* (Alva *et al.* 2021) and *Candida albicans* (Sharma *et al.* 2021).

We are especially interested in another budding yeast, *Kluyveromyces marxianus*, which originally attracted interest because of its role in food fermentation (Coloretti *et al.* 2017) but is now increasingly being considered as a platform of industrial biotechnology (Fonseca *et al.* 2008, Lane and Morrissey 2010, Karim *et al.* 2020). *Kluyveromyces marxianus* has some intrinsic traits such as thermotolerance, a broad substrate range and rapid growth that are useful for biotechnology (Groeneveld *et al.* 2009), and molecular and genomic tools to aid its development as an industrial platform (Cernak *et al.* 2018, Rajkumar *et al.* 2019, Rajkumar and Morrissey 2020). To date, all studies that addressed gene expression in this yeast focused on transcriptional effects via RNA-Seq experiments. Aspects that have been studied include growth and ethanol production on alternative sugar substrates such as xylose (Schabort *et al.* 2016, Kwon *et al.* 2019) and inulin (Gao *et al.* 2015), ethanol tolerance during adaptive laboratory evolution (Mo *et al.* 2019), response to growth inhibitors derived from lignocellulosic substrates (Wang *et al.* 2018) and the ability to grow at high temperatures (Fu *et al.* 2019). Recently, mainly using transcriptome analysis, we determined that young genes specific to *K. marxianus* are enriched in the response to stresses such as high temperature, low pH and high osmolarity (Doughty *et al.* 2020).

To complement the molecular toolbox, and as a resource to study the biology of this yeast, here we report the development of a protocol to carry out ribosome profiling in *K. marxianus*. For this, we adapted and applied the methods previously used for *S. cerevisiae* (Ingolia *et al.* 2009). We also developed a suite of bioinformatics tools to visualize and analyse *K. marxianus* RNA-Seq and ribosome profiling results. This involved addition of the *K. marxianus* data to publicly available genome (GWIPS-viz) and transcriptome (Trips-Viz) browsers, which, in turn, can be uploaded with user-generated expression data and used in private or public configurations. To facilitate the use of ribosome profiling as a very valuable tool to explore gene expression, we also include a detailed step-by-step protocol for users.

Materials and methods

Strains and growth conditions

Kluyveromyces marxianus strain CBS 6556 (CBS-KNAW culture collection, Westerdijk Institute) was used in these studies following standard growth and handling procedures. This particular strain is also available from other collections under the strain names ATCC 26548, NRRL Y-7571, KCTC 17555 and NCYC 2597 and has been quite widely used as a representative *K. marxianus* strain. For ribosome profiling experiments, standard growth conditions used synthetic minimal medium (Verduyn *et al.* 1992) and an incubation temperature of 30°C with shaking. The mineral medium consisted of the following per litre amounts: $(NH_4)_2SO_4$, 5.0 g; KH_2PO_4, 3.0 g; $MgSO_4{\cdot}7H_2O$, 0.5 g; trace elements (EDTA, 15 mg; $ZnSO_4{\cdot}7H_2O$, 4.5 mg; $MnCl_2{\cdot}2H_2O$, 0.84 mg; $CoCl_2{\cdot}6H_2O$, 0.3 mg; $CuSO_4{\cdot}5H_2O$, 0.3 mg; $Na_2MoO_4{\cdot}2H_2O$, 0.4 mg; $CaCl_2{\cdot}2H_2O$, 4.5 mg; $FeSO_4{\cdot}7H_2O$, 3.0 mg; H_3BO_3, 1.0 mg; KI, 0.1 mg); and silicone antifoam, 0.05 mL. It was adjusted to pH 6.0 with KOH before autoclaving (121°C, 20 min). The medium was cooled to room temperature and a filter-sterilized solution of vitamins prepared in demineralized water was added to a final concentration, per litre, of D-biotin, 0.05 mg; calcium pantothenate, 1.0 mg; nicotinic acid, 1.0 mg; *myo*-inositol, 25 mg; thiamine HCl, 1.0 mg; pyridoxine HCl, 1.0 mg; and *para*-aminobenzoic acid, 0.20 mg. Glucose was sterilized separately and added to a final concentration of 10 g/L. One hundred fifty microlitres of cultures in 500-mL conical flasks were grown to early log phase at A_{600} ~0.8 and either harvested or transferred to a shaking water bath at 40°C, with cells harvested at 5, 15, 30 and 60 min. All experiments were carried out with two biological replicates.

Ribosome profiling

For cell harvesting, cultures were quickly poured into a Durapore glass filter assembly (Merck, New Jersey, USA) using 0.45-μm pore nitrocellulose filter membrane (#7184-009, General Electric, New York, USA). A vacuum pump was immediately turned on and once liquid medium was removed, cells were quickly scraped into a 50-mL Falcon tube filled with liquid nitrogen. Once a 150 mL culture is added to the filtration assembly, it takes ~9–12 seconds until the medium is removed and the scraped cell pellet is collected and submerged in liquid nitrogen. After harvesting, 1.5 mL of polysome lysis buffer (5 mM $MgCl_2$, 150 mM KCl, 20 mM Tris–HCl, 100 μg/mL cycloheximide (Sigma-Aldrich, St. Louis, USA), 1 mM DTT and 1% Triton X-100) was slowly added dropwise to the liquid nitrogen and cells to create a frozen mixture of buffer and cells. The 50-mL Falcon tube (with pierced cap from screwdriver) was placed in −80°C to allow boiling off of the liquid nitrogen. Frozen cells/buffer was disrupted using cryogenic grinding using a Retsch Mixer Mill 400 (Haan, Germany) and 10-mL steel grinding jars and balls. Samples were ground for six cycles of 3 min each at 20 Hz; the steel jars were submerged in liquid nitrogen to cool samples between each cycle. After lysis, lysates were gently thawed on ice and quantified with Qubit 4.0 fluorometer and BR assay kit (#Q10211, Waltham, Massachusetts, USA). Thirty micrograms of lysate was diluted to 200 μL in polysome

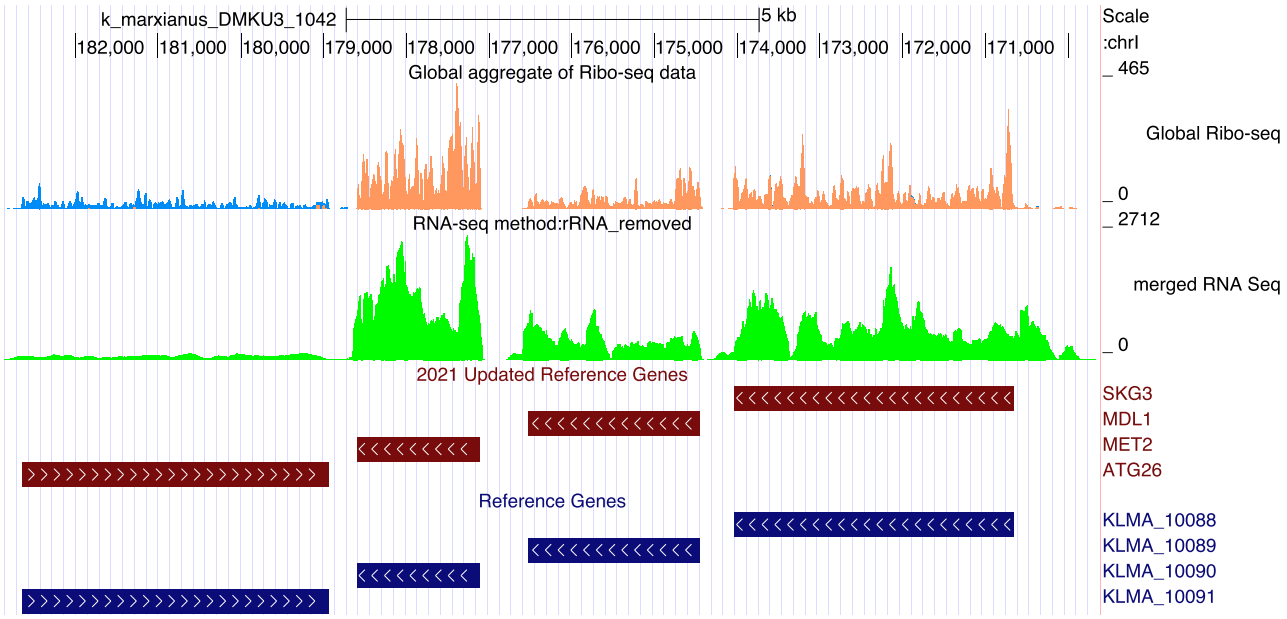

Figure 4. GWIPS-viz browser screenshot surrounding the *SKG3*, *MDL1*, *MET2* and *ATG26* locus of chromosome 1. Arrows on reference gene bars represent strand orientation. For ribosome profiling, orange reads represent positive strand RPFs, while blue reads represent negative strand RPFs.

buffer (lysis buffer without Triton X-100) and 1.5 μL RNase I was added (#N6901K, Epicentre/Lucigen, Middleton, Wisconsin, USA). RNase digestion was carried out at 200 rpm at 37°C for 45 min. To halt digestion, SUPERase•In (Invitrogen) was added, and samples were placed on ice before loading onto cold 10–50% sucrose gradients, which were prepared using a Biocomp gradient master. Gradients were spun for 3 h at 4°C and 36 000 rpm (221 632 × *g*) on SW 41 Ti rotor (Beckman Coulter). Monosome fractions were isolated from each sucrose gradient with Brandel density gradient fractionator using 1.5 mL/min flow speed and 60% CsCl, aliquoting fractions every 12 s on a UV–visible 96-well plate. Reading the 96-well plate at 260 nm determined which well(s) contained the monosome fractions. RNA from monosome fractions were isolated using TRIzol (Invitrogen) (Chomczynski and Sacchi 2006). Ribosome footprints were size selected with a 15% PAGE–urea (polyacrylamide gel electrophoresis) gel (70 min in 1× TBE and 300 V constant) using a 26 and 34 nt RNA marker (IDT) as a guide for excision. Using a scalpel, a slice representing the RPFs was cut from gel and placed into a 1.5-mL RNase-free Eppendorf tube and 500 μL of RNA elution buffer (300 mM NaOAc, pH 5.5, 1 mM EDTA and 0.25% v/v sodium dodecyl sulfate (SDS)) was added. Following overnight shaking at room temperature to elute the RPFs, RPFs were precipitated using standard alcohol precipitation using ice-cold isopropanol, 80% ethanol and 1.5 μL GlycoBlue coprecipitant (#AM9515, Ambion, Austin, Texas, USA).

Library construction

cDNA library construction with ribosome footprints is based on McGlincy and Ingolia (2017) protocol with minor modifications. In brief, size selected ribosome footprints were treated with T4 Polynucleotide Kinase (#M0201L, New England Biolabs (NEB), Ipswich, Massachusetts, USA)) followed by ligation to a DNA linker using T4 RNA ligase 2, truncated K227Q (#M0351L, NEB). The footprints were reverse transcribed using ProtoScript II (#M0368L, NEB). cDNA products were circularized using CircLigase II (#CL9025K, Epicentre/Lucigen). The major rRNA contaminants were removed using subtractive hybridization with custom biotinylated oligos (Sigma-Aldrich) and streptavidin beads (#65001, Invitrogen) as described in Ingolia *et al.* (2012). The remaining circularized products were amplified by PCR using Phusion polymerase (#M0530L, NEB). In a pilot experiments, libraries were sequenced on MiSeq platform at the Teagasc Next Generation DNA Sequencing Facility, Moorepark, Moorepark West, Fermoy, Co. Cork, Ireland. Prepared libraries using the protocol described within were sequenced on Illumina HiSeq 4000 using SE-75 sequencing at the Genomics & Cell Characterization Core Facility (GC3F), University of Oregon, Eugene, Oregon, USA.

Ribosome profiling data analysis preprocessing and genome annotation update

Adapter sequences were removed from reads using Cutadapt (Martin 2011). For genomic alignments, rRNA contaminants were removed and remaining reads were aligned to the *K. marxianus* DMKU3-1042 reference genome (Lertwattanasakul *et al.* 2015) with Bowtie (Langmead *et al.* 2009) using parameters -n 2 -m 1. Transcriptome alignments were made with Trips-Viz (Kiniry *et al.* 2019). For all genomic and transcriptome analysis, Ribosome profiling and RNA-Seq reads were aligned to the *K. marxianus* DMKU3-1042 reference strain (Lertwattanasakul *et al.* 2015).

RNA-Seq

RNA was isolated from clarified lysates using TRIzol (Chomczynski and Sacchi 2006) (Invitrogen #15596026) and quantified with a Qubit 4.0 fluorometer (Invitrogen). 1 μg of total RNA from each sample was analysed on an agarose bleach gel to determine RNA quality. Samples were sent to BGI Hong Kong for yeast rRNA removal (Ribo-Zero Gold rRNA Removal kit by Illumina, San Diego, California, USA (now discontinued)), library generation and sequencing with paired-end chemistry. Alternatively, RNA-Seq was carried out using polyA selection using Poly(A)Purist Mag Kit (#AM1922, Ambion) as per the manufacturer's instructions. PolyA selected cDNA library was generated and sequenced in the same method as ribosome profiling.

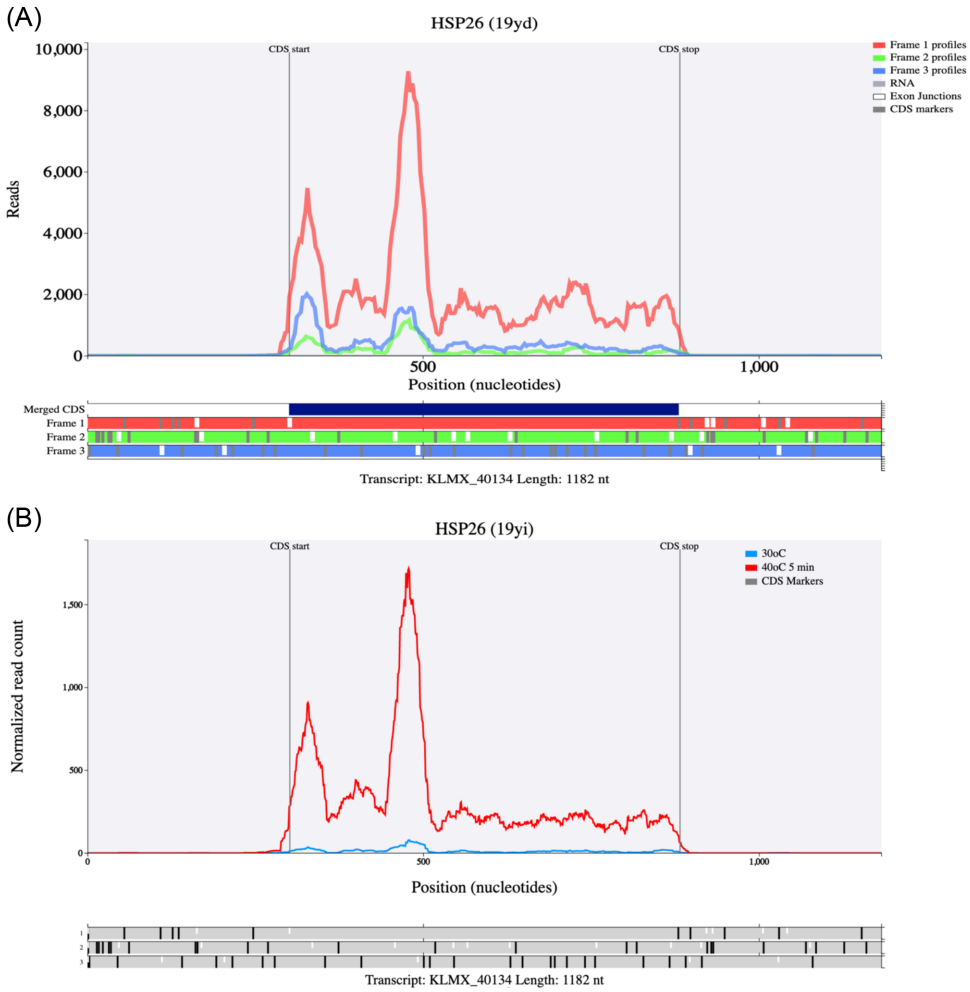

Figure 5. Trips-Viz transcript plots of *HSP26*. **(A)** Ribosome profiling coverage for aggregated data is displayed. Note that the dominant red line corresponds to the *HSP26* CDS region. **(B)** Normalized transcript comparison plot of *HSP26* showing increased mRNA translation at 30 and 40°C temperatures. Unique Trips-Viz plot identifiers are presented as '19yd' and '19yi' for panels (A) and (B), respectively. These identifiers can be added at the end of the following link to regenerate these specific plots (https://trips.ucc.ie/short/) on a web browser (example for panel (A): https://trips.ucc.ie/short/19yd).

Results and discussion

Development of a ribosome profiling protocol to study translation in *K. marxianus*

The previously described *S. cerevisiae* protocol (McGlincy and Ingolia 2017) was used as the basis for development of a ribosome profiling procedure for *K. marxianus*. An overview of the pipeline from culturing to downstream bioinformatic analyses is shown in Fig. 1. Summarizing the first part of the protocol, cultures are rapidly harvested and flash frozen to preserve the translational state of the cell. Frozen cells are then lysed cryogenically in the presence of cycloheximide, which ensures that ribosomes remain stalled even if the cells thaw, and the clarified lysate containing polysomes is treated with RNase I to digest unprotected mRNA surrounding the ribosomes, retaining the RPF. Monosomes are isolated from a sucrose gradient and loaded on a polyacrylamide gel to allow size selection of RPFs of ∼28nt. A cDNA library of these fragments is created and sequenced to identify the RPFs.

A limited-scale pilot experiment was first carried out to validate the methods and to identify the most abundant rRNA contaminants in the library. These arise because of RNase I digestion of rRNA and subsequent copurification of fragments of the same size as the RPFs. Due to natural polymorphisms in the rRNA encoding genes between species, the sequence of the major contaminated rRNA fragments needs to be determined empirically for each yeast. Knowing these sequences allows the design of synthetic biotinylated oligos that can be used to reduce rRNA contamination (Ingolia *et al.* 2012). In our pilot library, we identified six highly abundant rRNA fragments, four from the 25S rRNA, and one each from the 18S rRNA and from the mitochondrial 21S rRNA (Table 1). One of these sequences from 25S rRNA (GGGTGCATCATCGACCGATCCT) comprised ∼33% of all rRNA contaminants. By reducing rRNA contamination, the proportion of RPFs in a library is increased and thus more usable data are generated per experiment. The ribosome profiling protocol with rRNA depletion was then tested on a larger scale using 150 mL cultures of *K. marxianus* growing at different temperatures to increase the total number of genes that would be expressed. The focus at this time was on assessing the quality of the data generated and the robustness of the protocol rather than on analysis of changes in gene expression. Ribosome profiling was performed in duplicate on flask cultures at 30°C and at 5, 15, 30 and 60 min after a transfer from 30 to 40°C and, as is standard, RNA-Seq was also performed to measure transcript levels. Several analyses were performed to assess the robustness of the data that were obtained. First, the degree

of correlation of the number of mapped reads per gene between biological replicates for each condition was assessed and found to be high with a Pearson's correlation of >0.96 (Fig. 2A). Second, we checked whether the RPFs actually represented known protein coding genes (Fig. 2B). We found that 14% of total reads aligning to the genome represented uniquely mapping RPFs; only ∼0.5% of reads represented ambiguous RPFs, aligning to more than one location on the genome/transcriptome; and ∼85% of the reads mapped to rRNA encoding genes. Third, we examined whether our data showed the distinctive triplet periodicity (or subcodon phasing) of the aligned reads reflecting the 'codon-wise' movement of elongating ribosomes that is seen in ribosome profiling data. In the dataset, footprints of length 28 nt (approximately half of total footprints) displayed a remarkable strong periodicity signal with ∼95% of RPFs in phase with one of the three subcodon positions (Fig. 2C). Finally, RPFs are expected to be massively enriched in coding sequence (CDS) regions of genes. Using metagene profiles, we found that RPFs are largely present with CDS regions (Fig. 2D and E). In combination, these data demonstrate that the protocol generates robust ribosome profiling data.

Despite the oligo rRNA depletion, in our dataset from the large-scale experiment, ∼85% of the total reads were rRNA fragments. To determine the efficiency of the targeted rRNA contamination depletion, specific rRNA contaminant sequences were analysed before and after depletion. After depletion, we see almost 100% efficiency in removal of targeted rRNA contaminants, but there is not complete removal of all reads mapping to rRNA genes (Fig. 3). In part, this is because the rRNA abundance and composition may vary between samples and experiments due to slicing of RPFs from size selection gels by free hand. If desired, more oligonucleotides could be designed to further reduce rRNA contamination, thus increasing the proportion of RPFs in the sequencing pool. It was interesting to note that while ambiguously mapped reads can represent >10% of all reads in many studies from *S. cerevisiae* (seen looking at data in the Trips-Viz browser; https://trips.ucc.ie/) these comprised <1% of all reads in *K. marxianus*. Ambiguous mapping, whereby an RPF maps to two or more loci in the genome or transcriptome, arises because of the very short reads generated by ribosome profiling. As a result, it is not possible to determine the origin of the reads and these are generally discarded/ignored. This difference is most likely due to the large number of paralogous genes in *S. cerevisiae*, which arose through the proposed whole-genome duplication/hybridization (WGD) event in the evolutionary history of this species (Wolfe and Shields 1997, Marcet-Houben and Gabaldón 2015). As *K. marxianus* is a pre-WGD yeast, the same issue does not apply.

Visualization of *K. marxianus* ribosome profiling data on public browsers

Visualization of ribosome profiling data is important to examine translation/transcription of particular loci of interest. We previously developed two tools to allow visualization of these data at a genome level (GWIPS-viz) and at the level of individual RNAs (Trips-Viz). These tools are freely accessible via RiboSeqOrg portal at https://riboseq.org. GWIPS-viz is a genome browser that displays RPFs mapped to each chromosome of a reference genome (Michel *et al.* 2014). The GWIPS-viz database already contained reference genomes for ∼24 animal, plant, protozoal, fungal and viral genomes, and we added *K. marxianus* using the genome sequence and annotation from *K. marxianus* DMKU3-1042 strain as this was the most complete genome sequence available (Lertwattanasakul *et al.* 2015). It is possible to search GWIPS-viz by gene name or gene ID, and to zoom in/out of loci and as an extra feature that is new to GWIPS-viz, we included strand orientation of our ribosome profiling data to allow users to determine the strand to which an RPF is mapped (orange for +/forward strand, blue for −/negative strand) (Fig. 4). The browser is free to use and any user who generates their own ribosome profiling data or RNA-Seq tracks (bigWigs) can upload those data as custom tracks that can be viewed privately or made public. Once uploaded, a user is able to visualize and analyse their data using all the functionality of GWIPS-viz.

GWIPS-viz is mainly designed for analysis at a global level, allowing users to visualize any part of a genome, regardless of whether or not it is included in the annotations. In contrast, the second tool Trips-Viz is a transcriptome-level browser that focuses on individual mRNAs and allows a deep analysis of translation of each mRNA (Kiniry *et al.* 2019, 2021). This transcriptome browser allows users to generate single transcript plots displaying the ORF that is being translated. It also allows users to visualize the distribution of RPFs along an individual mRNA while also utilizing the triplet periodicity signal and differential colouring to identify potential translation in each ORF. As Trips-Viz did not include a reference transcriptome for *K. marxianus*, we created this reference transcriptome using our data. The application of Trips-Viz to study an individual mRNA is illustrated with an analysis of *HSP26*, using the (ribosome profiling) data for translation at 30 and 40°C (Fig. 5). The top panel uses aggregate data and shows the distribution of RPFs between each reading frame and across the transcript. It is clear that reads from the first ORF (red) dominate, which match the position and frame of the annotated CDS. The increase in the number of reads (RPFs) at certain positions indicates ribosome stalling during translation, for example at difficult-to-translate codons. The bottom panel compares the normalized read count of the correct ORF between the samples coming from cells grown at 30 and 40°C. The huge increase in translation at 40°C is evident. This was to be expected as *HSP26* encodes a heat shock protein and is strongly transcriptionally induced by temperature shift. Thus, the increase in translation in this case is due to an increase in mRNA abundance. Although not shown in this simple example, in addition to single transcript plots, the Trips-Viz browser contains a large amount of metadata analyses such as triplet periodicity, read breakdowns, metaplot, protein count tables and differential expression analysis that is useful for detailed studies of translation and its regulation (Kiniry *et al.* 2019, 2021).

Integrated omics studies with *K. marxianus*

Ultimately, a full suite of omics technologies, ranging from genomics to proteomics, is a requisite for comprehensive studies of any microbe. Analysis of genome sequences and transcriptomes is now relatively straightforward for diverse yeasts, but the development of other tools still lags. Now, with the laboratory and *in silico* methods that we developed, it is possible to perform ribosome profiling with a nonmodel yeast, *K. marxianus*. By including both transcriptome (RNA-Seq) and translatome (ribosome profiling) analyses in future studies, it will be possible to generate a comprehensive view of gene expression at a point of time and in response to a perturbation. This can be very useful to understand biological processes and for the development of strains for biotechnology. The strategy taken, and the pipeline used, can also serve as a prototype for the development of ribosome profiling methods for other yeasts of biological and biotechnological interest. To facilitate the application of the tools by as many users as possible, a comprehensive step-by-step protocol is provided in the Supporting Information.

Acknowledgements

We greatly appreciate the help of Dr Gary Loughran for advice on developing ribosome profiling and Luke Power (RiboMaps) for his critical reading of the manuscript. We thank Dr Simon Lawrence, School of Chemistry, University College Cork, for providing access to the Retsch Mixer Mill and Matthew Leibovitch (McGill University) for his help in installing the TracerDAQ hardware.

Supplementary data

Supplementary data are available at FEMSYR online.

Funding

This project received funding from the European Union's Horizon 2020 Framework Programme for Research and Innovation Grant Agreement No. 720824. This work was also supported by the SFI(Science Foundation Ireland)–HRB(Health Research Board)–Wellcome Trust Biomedical Research Partnership (210692/Z/18/ to PVB).

Data availability

The data underlying this article are available in European Nucleotide Archive at www.ebi.ac.uk, and can be accessed with the project assession number PRJEB45612. The data have also been deposited to GWIPS-viz (https://gwips.ucc.ie/) and Trips-Viz (https://trips.ucc.ie/).

Conflict of interest statement. PVB is a cofounder of RiboMaps Ltd, a company that provides ribosome profiling as a service.

References

Alva TR, Riera M, Chartron JW. Translational landscape and protein biogenesis demands of the early secretory pathway in *Komagataella phaffii*. *Microb Cell Fact* 2021;**20**:19.

Andreev DE, O'Connor PBF, Loughran G *et al*. Insights into the mechanisms of eukaryotic translation gained with ribosome profiling. *Nucleic Acids Res* 2017;**45**:513–26.

Arevalo-Villena M, Briones-Perez A, Corbo MR *et al*. Biotechnological application of yeasts in food science: starter cultures, probiotics and enzyme production. *J Appl Microbiol* 2017;**123**:1360–72.

Blevins WR, Tavella T, Moro SG *et al*. Extensive post-transcriptional buffering of gene expression in the response to severe oxidative stress in baker's yeast. *Sci Rep* 2019;**9**:11005.

Brar GA, Weissman JS. Ribosome profiling reveals the what, when, where and how of protein synthesis. *Nat Rev Mol Cell Biol* 2015;**16**:651–64.

Broach JR. Nutritional control of growth and development in yeast. *Genetics* 2012;**192**:73–105.

Cernak P, Estrela R, Poddar S *et al*. Engineering *Kluyveromyces marxianus* as a robust synthetic biology platform host. *mBio* 2018;**9**:e01410–8.

Chomczynski P, Sacchi N. The single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction: twenty-something years on. *Nat Protoc* 2006;**1**:581–5.

Coloretti F, Chiavari C, Luise D *et al*. Detection and identification of yeasts in natural whey starter for Parmigiano Reggiano cheese-making. *Int Dairy J* 2017;**66**:13–7.

de la Torre-Ruiz MA, Pujol N, Sundaran V. Coping with oxidative stress. The yeast model. *Curr Drug Targets* 2015;**16**:2–12.

de Nadal E, Posas F. Multilayered control of gene expression by stress-activated protein kinases. *EMBO J* 2010;**29**:4–13.

Delvigne F, Zune Q, Lara AR *et al*. Metabolic variability in bioprocessing: implications of microbial phenotypic heterogeneity. *Trends Biotechnol* 2014;**32**:608–16.

Doughty TW, Domenzain I, Millan-Oropeza A *et al*. Stress-induced expression is enriched for evolutionarily young genes in diverse budding yeasts. *Nat Commun* 2020;**11**:2144.

Duncan CDS, Mata J. The translational landscape of fission-yeast meiosis and sporulation. *Nat Struct Mol Biol* 2014;**21**:641–7.

Eisenberg AR, Higdon AL, Hollerer I *et al*. Translation initiation site profiling reveals widespread synthesis of non-AUG-initiated protein isoforms in yeast. *Cell Syst* 2020;**11**:145–60.

Fonseca GG, Heinzle E, Wittmann C *et al*. The yeast *Kluyveromyces marxianus* and its biotechnological potential. *Appl Microbiol Biotechnol* 2008;**79**:339–54.

Fu X, Li P, Zhang L *et al*. Understanding the stress responses of *Kluyveromyces marxianus* after an arrest during high-temperature ethanol fermentation based on integration of RNA-Seq and metabolite data. *Appl Microbiol Biotechnol* 2019;**103**:2715–29.

Gao J, Yuan W, Li Y *et al*. Transcriptional analysis of *Kluyveromyces marxianus* for ethanol production from inulin using consolidated bioprocessing technology. *Biotechnol Biofuels* 2015;**8**:115.

Gibson UE, Heid CA, Williams PM. A novel method for real time quantitative RT-PCR. *Genome Res* 1996;**6**:995–1001.

Groeneveld P, Stouthamer AH, Westerhoff HV. Super life: how and why 'cell selection' leads to the fastest-growing eukaryote. *FEBS J* 2009;**276**:254–70.

Hahn S, Young ET. Transcriptional regulation in *Saccharomyces cerevisiae*: transcription factor regulation and function, mechanisms of initiation, and roles of activators and coactivators. *Genetics* 2011;**189**:705–36.

Hinnebusch AG. Translational regulation of GCN4 and the general amino acid control of yeast. *Annu Rev Microbiol* 2005;**59**:407–50.

Ingolia NT, Brar GA, Rouskin S *et al*. The ribosome profiling strategy for monitoring translation *in vivo* by deep sequencing of ribosome-protected mRNA fragments. *Nat Protoc* 2012;**7**:1534–50.

Ingolia NT, Ghaemmaghami S, Newman JRS *et al*. Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. *Science* 2009;**324**:218–23.

Ingolia NT, Hussmann JA, Weissman JS. Ribosome profiling: global views of translation. *Cold Spring Harb Perspect Biol* 2019;**11**:a032698.

Jiménez-Gutiérrez E, Alegría-Carrasco E, Sellers-Moya Á *et al*. Not just the wall: the other ways to turn the yeast CWI pathway on. *Int Microbiol* 2020;**23**:107–19.

Karim A, Gerliani N, Aïder M. *Kluyveromyces marxianus*: an emerging yeast cell factory for applications in food and biotechnology. *Int J Food Microbiol* 2020;**333**:108818.

Kiniry SJ, Judge CE, Michel AM *et al*. Trips-Viz: an environment for the analysis of public and user-generated ribosome profiling data. *Nucleic Acids Res* 2021;**49**:W662–70.

Kiniry SJ, O'Connor PBF, Michel AM *et al*. Trips-Viz: a transcriptome browser for exploring Ribo-Seq data. *Nucleic Acids Res* 2019;**47**:D847–52.

Kwon D-H, Park J-B, Hong E *et al*. Ethanol production from xylose is highly increased by the *Kluyveromyces marxianus* mutant 17694-DH1. *Bioprocess Biosyst Eng* 2019;**42**:63–70.

Lane MM, Morrissey JP. *Kluyveromyces marxianus*: a yeast emerging from its sister's shadow. *Fungal Biol Rev* 2010;**24**:17–26.

Langmead B, Trapnell C, Pop M *et al*. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 2009;**10**:R25.

Lertwattanasakul N, Kosaka T, Hosoyama A *et al.* Genetic basis of the highly efficient yeast *Kluyveromyces marxianus*: complete genome sequence and transcriptome analyses. *Biotechnol Biofuels* 2015;**8**:47.

Liu Y, Nielsen J. Recent trends in metabolic engineering of microbial chemical factories. *Curr Opin Biotechnol* 2019;**60**:188–97.

Ljungdahl PO, Daignan-Fornier B. Regulation of amino acid, nucleotide, and phosphate metabolism in *Saccharomyces cerevisiae*. *Genetics* 2012;**190**:885–929.

Marcet-Houben M, Gabaldón T. Beyond the whole-genome duplication: phylogenetic evidence for an ancient interspecies hybridization in the baker's yeast lineage. *PLoS Biol* 2015;**13**:e1002220.

Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* 2011;**17**:10–2.

Martínez-Montañés F, Pascual-Ahuir A, Proft M. Toward a genomic view of the gene expression program regulated by osmostress in yeast. *OMICS* 2010;**14**:619–27.

Masser AE, Ciccarelli M, Andréasson C. Hsf1 on a leash: controlling the heat shock response by chaperone titration. *Exp Cell Res* 2020;**396**:112246.

McGlincy NJ, Ingolia NT. Transcriptome-wide measurement of translation by ribosome profiling. *Methods* 2017;**126**:112–29.

McManus CJ, May GE, Spealman P *et al.* Ribosome profiling reveals post-transcriptional buffering of divergent gene expression in yeast. *Genome Res* 2014;**24**:422–30.

Merrick WC. eIF4F: a retrospective. *J Biol Chem* 2015;**290**:24091–9.

Michel AM, Fox G, Kiran AM *et al.* GWIPS-viz: development of a ribo-seq genome browser. *Nucleic Acids Res* 2014;**42**:D859–64.

Mo W, Wang M, Zhan R *et al.* *Kluyveromyces marxianus* developing ethanol tolerance during adaptive evolution with significant improvements of multiple pathways. *Biotechnol Biofuels* 2019;**12**:63.

Mohammad F, Green R, Buskirk AR. A systematically-revised ribosome profiling method for bacteria reveals pauses at single-codon resolution. *eLife* 2019;**8**:e42591.

Monteuuis G, Miścicka A, Świrski M *et al.* Non-canonical translation initiation in yeast generates a cryptic pool of mitochondrial proteins. *Nucleic Acids Res* 2019;**47**:5777–91.

Morano KA, Grant CM, Moye-Rowley WS. The response to heat shock and oxidative stress in *Saccharomyces cerevisiae*. *Genetics* 2012;**190**:1157–95.

Nandy SK, Srivastava RK. A review on sustainable yeast biotechnological processes and applications. *Microbiol Res* 2018;**207**:83–90.

Parapouli M, Vasileiadis A, Afendra A-S *et al.* *Saccharomyces cerevisiae* and its industrial applications. *AIMS Microbiol* 2020;**6**:1–31.

Rajkumar AS, Morrissey JP. Rational engineering of *Kluyveromyces marxianus* to create a chassis for the production of aromatic products. *Microb Cell Fact* 2020;**19**:207.

Rajkumar AS, Varela JA, Juergens H *et al.* Biological parts for *Kluyveromyces marxianus* synthetic biology. *Front Bioeng Biotechnol* 2019;**7**:97.

Sanz AB, García R, Rodríguez-Peña JM *et al.* The CWI pathway: regulation of the transcriptional adaptive response to cell wall stress in yeast. *J Fungi (Basel)* 2017;**4**:1.

Schabort DTWP, Letebele PK, Steyn L *et al.* Differential RNA-seq, multi-network analysis and metabolic regulation analysis of *Kluyveromyces marxianus* reveals a compartmentalised response to xylose. *PLoS One* 2016;**11**:e0156242.

Schena M, Shalon D, Davis RW *et al.* Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 1995;**270**:467–70.

Sharma P, Wu J, Nilges BS *et al.* Humans and other commonly used model organisms are resistant to cycloheximide-mediated biases in ribosome profiling experiments. *Nat Commun* 2021;**12**:5094.

Smith JE, Alvarez-Dominguez JR, Kline N *et al.* Translation of small open reading frames within unannotated RNA transcripts in *Saccharomyces cerevisiae*. *Cell Rep* 2014;**7**:1858–66.

Spealman P, Naik AW, May GE *et al.* Conserved non-AUG uORFs revealed by a novel regression analysis of ribosome profiling data. *Genome Res* 2018;**28**:214–22.

Steitz JA. Polypeptide chain initiation: nucleotide sequences of the three ribosomal binding sites in bacteriophage R17 RNA. *Nature* 1969;**224**:957–64.

Sui Y, Wisniewski M, Droby S *et al.* Responses of yeast biocontrol agents to environmental stress. *Appl Environ Microbiol* 2015;**81**:2968–75.

Takors R. Scale-up of microbial processes: impacts, tools and open questions. *J Biotechnol* 2012;**160**:3–9.

Taymaz-Nikerel H, Cankorur-Cetinkaya A, Kirdar B. Genome-wide transcriptional response of *Saccharomyces cerevisiae* to stress-induced perturbations. *Front Bioeng Biotechnol* 2016;**4**:17.

Verduyn C, Postma E, Scheffers WA *et al.* Effect of benzoic acid on metabolic fluxes in yeasts: a continuous-culture study on the regulation of respiration and alcoholic fermentation. *Yeast* 1992;**8**:501–17.

Wang D, Wu D, Yang X *et al.* Transcriptomic analysis of thermotolerant yeast *Kluyveromyces marxianus* in multiple inhibitors tolerance. *RSC Adv* 2018;**8**:14177–92.

Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 2009;**10**:57–63.

Wehrs M, Tanjore D, Eng T *et al.* Engineering robust production microbes for large-scale cultivation. *Trends Microbiol* 2019;**27**:524–37.

Wolfe KH, Shields DC. Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* 1997;**387**:708–13.