

Review of: "Development of a Ribosome Profiling Protocol to Study Translation in the yeast *Kluyveromyces marxianus*"

Pieter Spealman

Potential competing interests: The author(s) declared that no potential competing interests exist.

This work is a valuable addition to the yeast community as it extends the reach of an important molecular method to a up and coming model organism. There are some components which could be improved for clarity, but I feel these are minor revisions at best.

I'll attach my line notes as a supplementary file and summarize my main points here.

Bench methods

These are relatively complete and clear and only need some small typographical edits.

I would like to say however that the authors miss an opportunity to introduce their model organism to a broader audience. Coming from a background with *S. cerevisiae* I am comfortable with some procedures at the bench, things like OD600, doubling rates, cell saturation and induced oxidative shift (and heat stress). But these are species (if not strain dependent). I feel a brief primer in working with *K. marxianus*, laying out simple guidelines like shaker RPMs, expected doubling times, and any other expertise the lab has gained with this organism would make the prospects of working with it more approachable to the average *Saccharomyces* biologist.

Data analysis

There are three issues present with the data analysis as is.

The first issue, if I understood the manuscript correctly, is that the original reads included UMIs but these were not utilized. This is unfortunate, as UMIs can help address PCR amplification artifacts that can distort the underlying signal. And while PCR artifacts may not be a significant issue at 8 cycles, they can be a much greater concern at 14. While not every lab may have access to UMIs, having a method that supports their use and results that set an expectation of UMI duplication rates would be very valuable to the community.

The second issue, although debatable, is that a splicing aware aligner should be used to handle any and all

expression data in organisms (like *K. marxianus*) that use splicing. This becomes more important when other components of data quality control are unreported, like the number of reads per run that failed to align.

This brings up the final issue which is that no quality control data is reported. While FASTQC, cutadapt , etc. are ran, the reader doesn't have an expectation of how many total reads are produced using this method. Basic metrics like how many reads contain the adapter, how many align, how many are duplicated, chimeric, or of low quality, should be reported. Notably, the percent alignment would be beneficial to know since the aligned reference genome comes from a different strain, and potentially one with several thousand SNPs ^[1] difference (and a splicing un-aware aligner that only allows two mismatches). Some of these will results will vary on a run-by-run basis and are dependent on the hands at the bench but establishing a performance baseline would help set the expectations of users following these methods.

Conclusion

I hope the authors address these concerns as they would benefit this paper and help introduce *K. marxianus* to a new audience.

References

1. ^Raúl A. Ortiz-Merino, Javier A. Varela, Aisling Y. Coughlan, Hisashi Hoshida, et al. (2018). Ploidy Variation in *Kluyveromyces marxianus* Separates Dairy and Non-dairy Isolates. *Front. Genet.*, vol. 9 . doi:10.3389/fgene.2018.00094.