A streamlined ribosome profiling protocol for the characterization of microorganisms

Haythem Latif[1], Richard Szubin[1], Justin Tan[1], Elizabeth Brunk[2,3], Anna Lechner[4], Karsten Zengler[1,5], and Bernhard O. Palsson[1,5]
[1]Bioengineering Department, University of California, San Diego, La Jolla, CA, [2]Fuels Synthesis Division, Joint BioEnergy Institute, Emeryville, CA, [3]Department of Chemical & Biomolecular Engineering, Department of Bioengineering, University of California, Berkeley, Berkeley, CA, [4]QB3 Institute, University of California, Emeryville, CA, and [5]Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Lyngby, Denmark

BioTechniques 58:329-332 (June 2015) doi 10.2144/000114302
Keywords: ribosome profiling; next-generation sequencing; bacteria; Illumina

Supplementary material for this article is available at www.BioTechniques.com/article/114302.

Ribosome profiling is a powerful tool for characterizing in vivo protein translation at the genome scale, with multiple applications ranging from detailed molecular mechanisms to systems-level predictive modeling. Though highly effective, this intricate technique has yet to become widely used in the microbial research community. Here we present a streamlined ribosome profiling protocol with reduced barriers to entry for microbial characterization studies. Our approach provides simplified alternatives during harvest, lysis, and recovery of monosomes and also eliminates several time-consuming steps, in particular size-selection steps during library construction. Furthermore, the abundance of rRNAs and tRNAs in the final library is drastically reduced. Our streamlined workflow enables greater throughput, cuts the time from harvest to the final library in half (down to 3–4 days), and generates a high fraction of informative reads, all while retaining the high quality standards of the existing protocol.

Ribosome profiling is a novel, sequencing-based technique that captures ribosomes as they traverse transcripts in vivo, thereby revealing protein synthesis at the genome scale (1,2). This protocol has produced numerous findings on the detailed molecular processes of translation, such as the sequence-based prediction of translational pause sites (3), the characterization of the co-translational chaperone action of trigger factor (4), and the identification of translational regulatory processes (2). In addition to elucidating translation mechanisms, the quantitative nature of ribosome profiling has shown a strong correlation with quantitative proteomics (1,5). More recently, this approach has been utilized in systems-level modeling of translation elongation (6).

The ribosome profiling protocol, which has been previously detailed (7,8), applies endonuclease digestion to cell lysate, thereby generating mRNA fragments protected by actively translating ribosomes that are then recovered and converted into a sequencing library. Although the original protocol is robust, broad utilization of ribosome profiling for microbial applications is limited by technological challenges. The original protocol utilizes equipment (e.g., Retch mill, sucrose gradient fraction collector) for which simpler alternatives can be found. Furthermore, ribosome profiling is a laborious protocol taking 7–8 days to complete (7). The final libraries may also contain a notable fraction of rRNA and tRNA species (7) (http://www.nature.com/nmeth/journal/v10/n5/abs/nmeth.f.363.html), which produce uninformative sequencing reads. Many of these limitations have been addressed for eukaryotic applications in the ARTseq protocol (Epicenter, Madison, WI), but implementation of the ARTseq protocol in bacteria is not currently possible because the ARTseq kits are only compatible with yeast and mammalian samples.

We have updated and streamlined the ribosome profiling protocol with the goal of making it available to the broader microbial research community by (*i*) offering simplified alternatives in the steps leading to footprint recovery, (*ii*) enabling increased throughput and parallelization of samples, (*iii*) reducing the time from harvest to library construction, and (*iv*)eliminating rRNA and tRNA contaminating species in the final sequencing libraries. Our protocol achieves this by first modifying the existing ribosome profiling protocol to

METHOD SUMMARY

Here we report a streamlined method for ribosome profiling in bacteria that retains the high data quality characteristics of the original protocol while incorporating modifications that (*i*) enable increased throughput, (*ii*) reduce the time from harvest to final library production to 3–4 days, and (*iii*) enriche for informative reads by eliminating rRNA and tRNA reads.

A

Day 1	Day 2	Day 3	Day 4
Inoculate Culture	Micrococcal Nuclease Digestion	T4 PNK Treatment	PCR Enrichment
Harvest & Lysis	Ribosome Enrichment (Size Exclusion)	Small RNA Purification Kit	Optional: Size Selection
	Small RNA Isolation/Purification (Kit based)	Small RNA library Prep Kit (NEB) through Reverse Transcription	Library QC
	rRNA Subtraction (Ribo-Zero Kit)	Recommended: tRNA Removal	Sequence

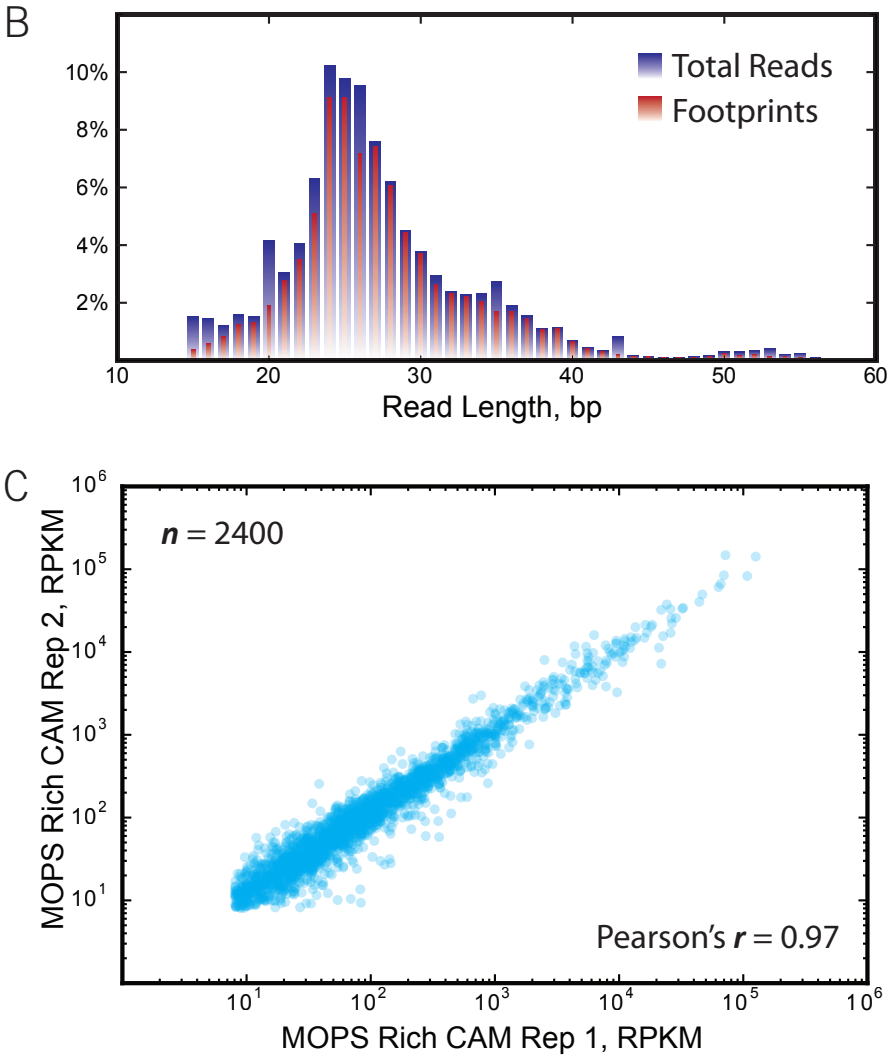

Figure 1. Streamlined protocol for ribosome profiling of microorganisms. (A) Overview of the major steps and associated timing of the streamlined protocol presented here. (B) The read length distribution for all reads (blue) and ribosomal footprints (red) generated using the streamlined protocol. (C) Correlation of reads per kilobase per million reads mapped (RPKM) across all genes with RPKM>8 for biological replicates generated using the streamlined protocol. *E. coli* was grown on MOPS rich media and harvested at the exponential growth phase with chloramphenicol (CAM) addition prior to harvest.

perform harvest and lysis using short centrifugation followed by repeated freeze-thaw lysis. This protocol uses antibiotic treatment prior to harvest as previously described (4,7). Lysate is then treated with micrococcal nuclease in a buffer that maintains ribosome integrity without sacrificing high nuclease activity. Monosomes are then recovered using a size exclusion spin-column analogous to those used for ARTseq. Next, recovered ribosomes are treated with Qiazol (Qiagen, Germantown, MD) to recover RNA footprints. Footprints are subsequently isolated using kit-based purification approaches followed by Ribo-Zero rRNA depletion (Epicenter). Library construction is then performed using a commercially available small RNA purification kit. For this study, the NEBNext Small RNA Library Prep Kit for Illumina was used (New England Biolabs, Ipswich, MA). After 3′ and 5′ adaptor ligation, additional steps are introduced to remove tRNAs by hybridization of custom anti-sense DNA oligos followed by treatment with a thermostable RNase H (Epicenter). This effectively degrades tRNA contaminating species, leaving mRNA footprints unaffected. The anti-tRNA probes carry terminal dideoxynucleotides to prevent their participation during the final library amplification. Unlike the existing ribosome profiling protocols, this procedure for library construction does not require any gel purification steps. Therefore, the time from harvest to finished library is reduced from 7–8 days to 3–4 days (Figure 1A).

We validated our streamlined protocol by generating ribosome profiling data replicating published conditions in *Escherichia coli* K-12 MG1655 wild type cells (5), with the exception that chloramphenicol (CAM) was added to the cultures 2 min prior to harvest as previously described (7). *E. coli* cultures were grown aerobically to exponential phase in MOPS rich medium with 0.2% glucose and a full supplementation of amino acids (Teknova, Hollister, CA). Libraries were sequenced using the Illumina MiSeq platform (GEO accession GSE63858) and processed using the bioinformatics pipeline outlined previously (5,7). Examination of the mapped read length distribution produced by the streamlined protocol yields footprints in a size range comparable to that found using the original protocol (20–42 nucleotides) (5) with 92% of mapped footprints falling within this range (Figure 1B). This plot also shows that 84% of reads mapped to ribosomal footprints (~7 million reads), with the remaining 16% mapping to annotated rRNAs, tRNAs, and other non-coding RNAs (~1.3 million reads). Furthermore, the ribosome density was compared for

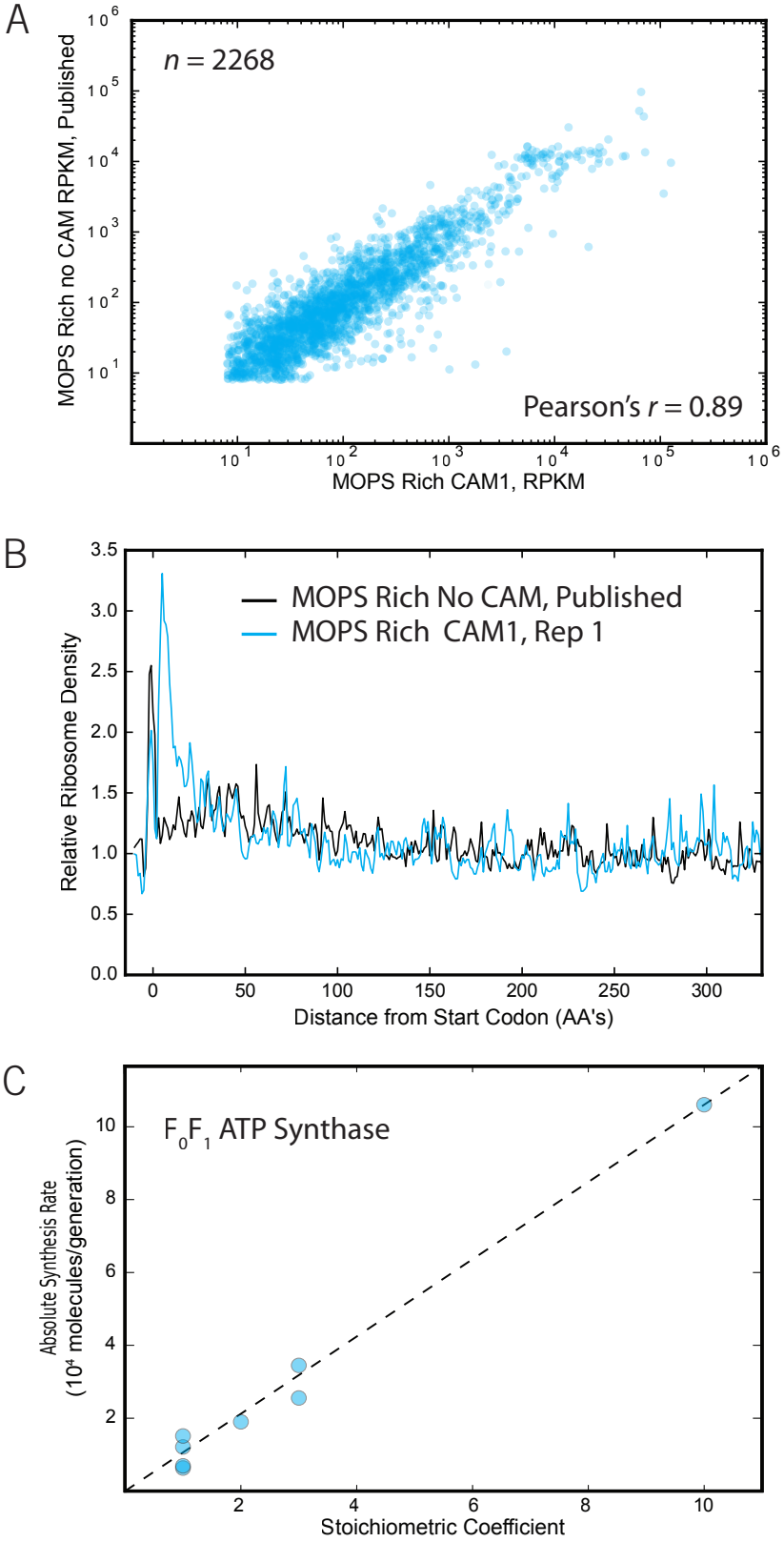

Figure 2. Benchmarking the streamlined protocol against publicly available data. (A) Correlation of reads per kilobase per million reads mapped (RPKM) across all genes with RPKM>8 for the streamlined protocol compared with publicly available data. For the streamlined protocol, *E. coli* was grown on MOPS rich media and harvested at the exponential growth phase with chloramphenicol (CAM) addition prior to harvest. The publicly available data were generated under similar growth conditions but without the use of CAM prior to harvest. (B) Meta-gene analysis of the relative ribosome density profiles observed from the start codon to 333 amino acids downstream. (C) F_0F_1 ATP synthase stoichiometry determined by ribosome profiling.

biological replicates across all protein-coding genes with reads per kilobase per million reads mapped (RPKM) greater than 8 (Figure 1C). These data sets showed a strong linear correlation between biological replicates analogous to those generated by the original protocol (1,3,7). Thus, the streamlined protocol is highly reproducible and yields a large fraction of informative reads in half the time of the original method.

Comparison with publicly available data further affirms that the streamlined protocol provides high-quality ribosome profiling data. Figure 2A compares one of the replicates generated using MOPS rich medium with CAM added prior to harvest with data generated under similar conditions but without CAM addition prior to harvest using the "rapid harvest" methodology (GEO accession GSM1300279). This yields a linear correlation with a Pearson's *r* value of 0.89. Interestingly, a comparison of harvest approaches with and without CAM using *E. coli* MC4100 Δ*tig*::Kan + pTrc-*tig*-TEV-Avi cells showed a similar correlation with $r = 0.90$ (7). Similar differences between CAM-treated and rapid harvest samples are observed in the meta-gene profiles from the start codon to 300 amino acids downstream (Figure 2B). CAM-treated samples produce a pronounced peak near the sixth codon that does not occur when using the rapid harvest approach. This difference in profiles was also observed in *E. coli* MC4100 Δ*tig*::Kan + pTrc-*tig*-TEV-Avi and is thought to occur as a result of poor inhibition of translation initiation by CAM (7). This effect has also been seen with cyclohexamide-treated ribosome profiling data sets in eukaryotes (8). As such, the user should account for this during quantitative analysis. Lastly, the power of ribosome profiling was recently displayed in its ability to capture the stoichiometry of heteroprotein complexes using the calculated absolute synthesis rate (5). This use of ribosomal profiling was perhaps best illustrated in the ability to predict the stoichiometry of the eight proteins making up the F_0F_1 ATP Synthase complex. Using the ribosome profiling data generated here, we were able to accurately predict the stoichiometry of this complex (Figure 2C).

Here we present modification of an increasingly important molecular biology approach to enable broader accessibility for the microbial research community. This method allows for data to be generated in half the time needed for the original protocol, allows for multiple samples to be processed in parallel, and yields more usable data by eliminating undesirable library contaminants. When combined with parallel RNA-seq sample preparation, valuable quantitative measures can be ascertained (e.g., translation efficiency) (4), and potential biases introduced during library preparation can be identified and corrected (9). This protocol simplifies the ribosome profiling procedure while retaining high quality data, a hallmark of the original ribosome profiling protocol.

Author contributions

H.L. and R.S. conceived, developed, and executed the experiments. R.S. performed all experiments. H.L. processed and analyzed all data sets. H.L., J.T., E.B., and A.L. developed the bioinformatics analytical pipeline. H.L., R.S., B.O.P., and K.Z. wrote and edited the manuscript.

Acknowledgments

We would like to thank Gene-Wei Li and Johnathan Weissman (University of California, San Francisco) for their input and guidance. We also would like to thank Epicentre. The Novo Nordisk Foundation and NIH U01 grant GM102098-01 provided financial support for this work. H.L. was supported through the National Science Foundation Graduate Research Fellowship under grant DGE1144086. This paper is subject to the NIH Public Access Policy.

Competing interests

The authors declare no competing interest.

References

1. **Ingolia, N.T., S. Ghaemmaghami, J.R. Newman, and J.S. Weissman.** 2009. Genome-wide analysis *in vivo* of translation with nucleotide resolution using ribosome profiling. Science *324*:218-223.
2. **Ingolia, N.T.** 2014. Ribosome profiling: new views of translation, from single codons to genome scale. Nat. Rev. Genet. *15*:205-213.
3. **Li, G.W., E. Oh, and J.S. Weissman.** 2012. The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. Nature *484*:538-541.
4. **Oh, E., A.H. Becker, A. Sandikci, D. Huber, R. Chaba, F. Gloge, R.J. Nichols, A. Typas, et al.** 2011. Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor *in vivo.* Cell *147*:1295-1308.
5. **Li, G.W., D. Burkhardt, C. Gross, and J.S. Weissman.** 2014. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. Cell *157*:624-635.
6. **Subramaniam, A.R., B.M. Zid, and E.K. O'Shea.** 2014. An integrated approach reveals regulatory controls on bacterial translation elongation. Cell *159*:1200-1211.
7. **Becker, A.H., E. Oh, J.S. Weissman, G. Kramer, and B. Bukau.** 2013. Selective ribosome profiling as a tool for studying the interaction of chaperones and targeting factors with nascent polypeptide chains and ribosomes. Nat. Protoc. *8*:2212-2239.
8. **Ingolia, N.T., G.A. Brar, S. Rouskin, A.M. McGeachy, and J.S. Weissman.** 2012. The ribosome profiling strategy for monitoring translation *in vivo* by deep sequencing of ribosome-protected mRNA fragments. Nat. Protoc. *7*:1534-1550.
9. **Artieri, C.G. and H.B. Fraser.** 2014. Accounting for biases in riboprofiling data indicates a major role for proline in stalling translation. Genome Res. *24*:2011-2021.

Received 09 December 2014; accepted 06 March 2015.

Address correspondence to Karsten Zengler, Bioengineering Department, University of California, San Diego, La Jolla, CA. E-mail: kzengler@ucsd.edu