RESEARCH ARTICLE

Reading canonical and modified nucleobases in 16S ribosomal RNA using nanopore native RNA sequencing

Andrew M. Smith[1], Miten Jain[1], Logan Mulroney[1], Daniel R. Garalde[2], Mark Akeson[1] *

1 UC Santa Cruz Genomics Institute, University of California, Santa Cruz, California, United States of America, **2** Oxford Nanopore Technologies, Oxford, United Kingdom

* makeson@soe.ucsc.edu

Abstract

The ribosome small subunit is expressed in all living cells. It performs numerous essential functions during translation, including formation of the initiation complex and proofreading of base-pairs between mRNA codons and tRNA anticodons. The core constituent of the small ribosomal subunit is a ~1.5 kb RNA strand in prokaryotes (16S rRNA) and a homologous ~1.8 kb RNA strand in eukaryotes (18S rRNA). Traditional sequencing-by-synthesis (SBS) of rRNA genes or rRNA cDNA copies has achieved wide use as a 'molecular chronometer' for phylogenetic studies, and as a tool for identifying infectious organisms in the clinic. However, epigenetic modifications on rRNA are erased by SBS methods. Here we describe direct MinION nanopore sequencing of individual, full-length 16S rRNA absent reverse transcription or amplification. As little as 5 picograms (~10 attomole) of purified E. coli 16S rRNA was detected in 4.5 micrograms of total human RNA. Nanopore ionic current traces that deviated from canonical patterns revealed conserved *E. coli* 16S rRNA 7-methylguanosine and pseudouridine modifications, and a 7-methylguanosine modification that confers aminoglycoside resistance to some pathological *E. coli* strains.

Introduction

Nanopore-based native RNA strand sequencing is conceptually similar to nanopore DNA sequencing [1]. An applied voltage across a single protein pore in an impermeable membrane results in an ionic current through the pore [2]. This current varies when a DNA or RNA strand is captured by the electric field and then moved through the pore in single nucleotide steps regulated by a processive enzyme [3–5]. The output is a time series of discrete ionic current segments that correspond to the sequence of bases (kmers) that occupy the pore at any given time [6,7]. Thus, each base within a strand is determined computationally from the series of ionic current segments that it influences [6].

Other PCR-free RNA sequencing technologies (often referred to as direct RNA sequencing because the RNA is present during sequencing) have been implemented using SBS combined with optical readout of fluorophore-labelled DNA nucleotides [8,9]. They share some of the

benefits of nanopore native RNA sequencing (e.g. absence of PCR biases), however their reported read lengths are short (typically <25 nt [8] and <34 nt [9] respectively).

Nanopore RNA sequencing was first implemented by Oxford Nanopore Technologies (ONT) for mRNA using adapters designed to capture polyadenylated RNA strands[1]. We reasoned that this technique could be modified to sequence 16S rRNA. 16S rRNA is a logical substrate for nanopore sequencing because of its abundance and broad use for identifying bacteria and archaea [10,11]. In addition, numerous antibiotics target prokaryotic ribosomes [12], which can acquire resistance via nucleotide substitutions or by gain or loss of base modifications [13]. These base modifications are difficult to detect using SBS methods. A significant advantage of nanopore sequencing is that modifications can be resolved because each nucleoside touches the nanoscale sensor as the strand translocates through the pore.

Results and discussion

Fig 1A illustrates the strategy we used to prepare 16S rRNA for MinION sequencing. Briefly, 16S rRNA was ligated to an adapter bearing a 20-nt overhang complementary to the 3′-end of the 16S rRNA (**Fig 1A** and **S1A Fig**). This overhang included the Shine-Dalgarno sequence [14], which targets the conserved anti-Shine-Dalgarno sequence in prokaryotic 16S rRNA [15]. Next, a modular Oxford Nanopore Technologies (ONT) adapter bearing a proprietary RNA motor protein was hybridized and ligated to the adapted RNA strands thus facilitating capture and sequencing on the MinION. While this manuscript was in preparation a similar targeted approach has been described for nanopore native RNA sequencing of viral RNA [16].

Fig 1B shows a representative ionic current trace caused by translocation of a purified *E. coli* 16S rRNA strand through a nanopore in the MinION array. The read begins with an ionic current pattern characteristic of the ONT RNA sequencing adapter strand followed by the 16S rRNA adapter strand. The 16S rRNA is then processed through the nanopore one base at a time in the 3′ to 5′ direction. The ionic current features are typical of long nucleic acid polymers processed through a nanopore [17,18].

Sequencing of purified 16S rRNA from *E. coli* strain MRE600 using a single flow cell produced 219,917 reads over 24 hours that aligned to the reference sequence (16S rRNA *rrnD* gene) (**Fig 1C**). This represents 94.6% of the total MinION read output for that experiment. Median read length was 1349 bases. We identified 142,295 reads that had sequence coverage within twenty-five nucleotides of the 16S rRNA 5′-end and within fifty nucleotides of the 3′-end.

We calculated percent read identities for nanopore experiments using 16S rRNA and Enolase 2 RNA (a calibration standard supplied by ONT) (**S2 Fig**). The median read identity for 16S rRNA was 81.6% compared to 87.1% for Enolase 2 (**S1 Table**). Close examination of 16S rRNA reads revealed frequent deletion errors in G-rich regions, which are abundant in noncoding structural RNAs (**S2 Table** and **S3 Fig**). This is observed as drops in coverage when unsmoothed read coverage is plotted across the *E. coli* 16S rRNA reference (**S2 Fig**). Other sequencing errors may represent true single nucleotide variants (SNVs) from the 16S rRNA reference sequence used for alignment. *E. coli* strains typically have seven 16S rRNA gene copies, with some of the gene copies differing by as much as 1.1%. Modified nucleotides could also alter ionic current from canonical nucleotides [19,20]. *E. coli* 16S rRNA contains 11 known nucleotide modifications [21].

As a diagnostic for detecting SNVs and nucleotide modifications, we scanned for bases that were reproducibly mis-called when individual 16S rRNA MinION reads were aligned to the *E. coli* MRE600 16S rRNA reference. Using marginCaller at a posterior probability threshold of 0.3 [17], we detected 24 such reproducible miscalls in the nanopore 16S rRNA reads (**S3**
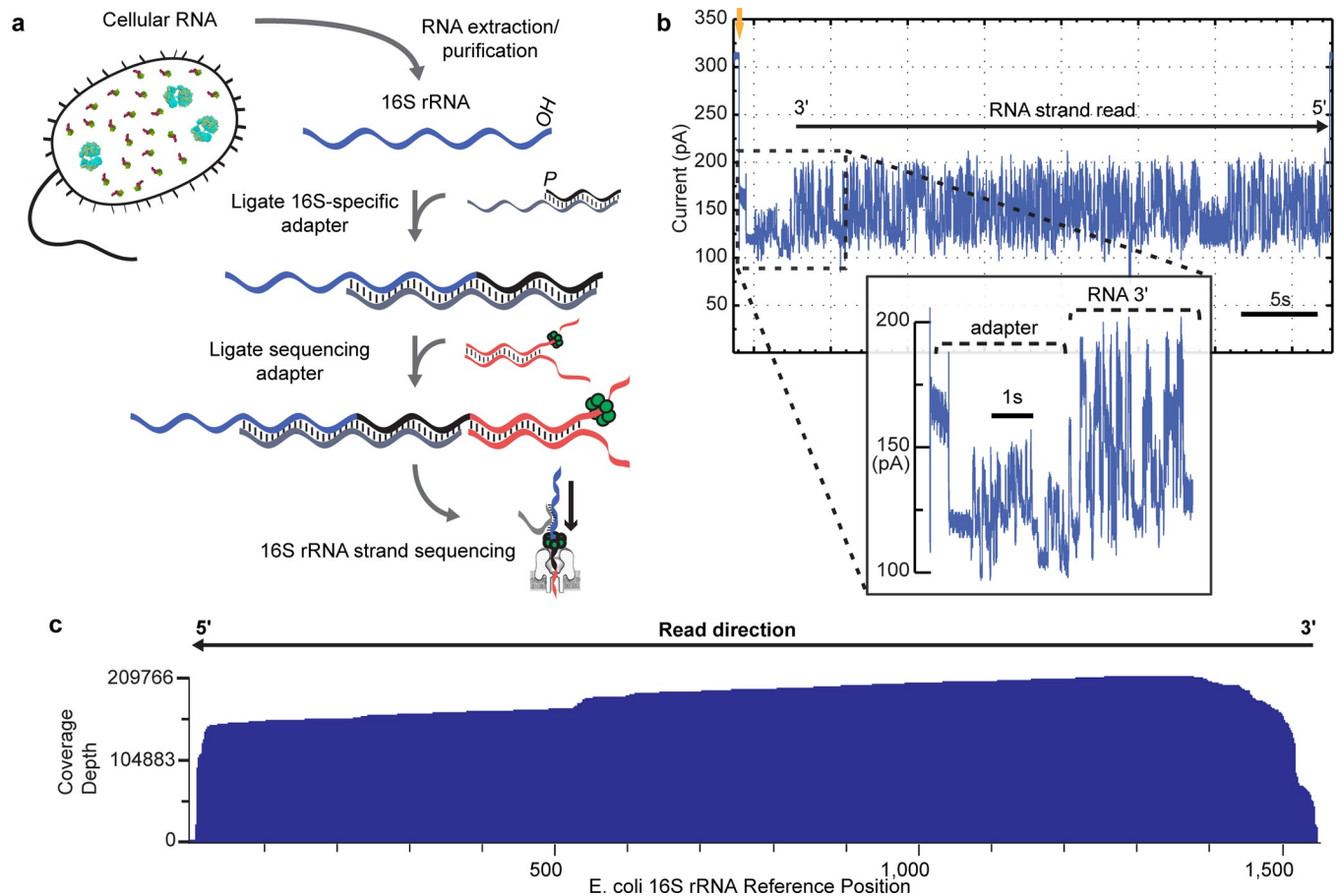
Fig 1. Nanopore sequencing of individual *E. coli* 16S ribosomal RNA strands. (a) Library preparation for MinION sequencing. Following RNA extraction, a 16S rRNA-specific adapter is hybridized and ligated to the 16S rRNA 3′ end. Next, a sequencing adapter bearing a RNA motor protein is hybridized and ligated to the 3′ overhang of the 16S rRNA adapter. The sample is then loaded into the MinION flowcell for sequencing. **(b)** Representative ionic current trace during translocation of a 16S rRNA strand from *E. coli* str. MRE600 through a nanopore. Upon capture of the 3′ end of an adapted 16S rRNA, the ionic current transitions from open channel (310 pA; gold arrow) to a series of discrete segments characteristic of the adapters (inset). This is followed by ionic current segments corresponding to base-by-base translocation of the 16S rRNA. The trace is representative of thousands of reads collected for individual 16S rRNA strands from *E. coli*. **(c)** Alignment of 200,000+ 16S rRNA reads to *E. coli* str MRE600 16S rRNA *rrnD* gene reference sequence. Reads are aligned in 5′ to 3′ orientation, after being reversed by the base-calling software. Numbering is according to canonical *E. coli* 16S sequence. Coverage across reference is plotted as a smoothed curve. In this experiment, 94.6% of reads that passed quality filters aligned to the reference sequence. Data presented here are from a single flow cell.

Table). Five of these correlated with variants in the reference sequence relative to the other 16S rRNA gene copies in *E. coli* MRE600. For example, at position 79 the reference is adenine (A79), whereas the other six *E. coli* MRE600 16S rRNA gene copies are guanosine, in agreement with the majority of nanopore reads.

One of the highest probability deviations from reference was at G527 (predicted C, posterior probability 0.7785), which was systematically mis-called by the nanopore as a C (Fig 2A and Fig 2B). This residue is located in a conserved region of the 16S rRNA 530 loop, near the A-site in the ribosome [22]. The guanosine at this position is known to be methylated at N7 (m7G527) [23], which creates a delocalized positive charge on the base. We hypothesized that this modification could significantly alter the ionic current segments that contain m7G527, thus resulting in the systematic base-call error.

As a test, we compared *E. coli* str. MRE600 (wild type) 16S rRNA nanopore reads with reads for an *E. coli* strain (see methods) that lacks the enzyme RsmG, which responsible for N7
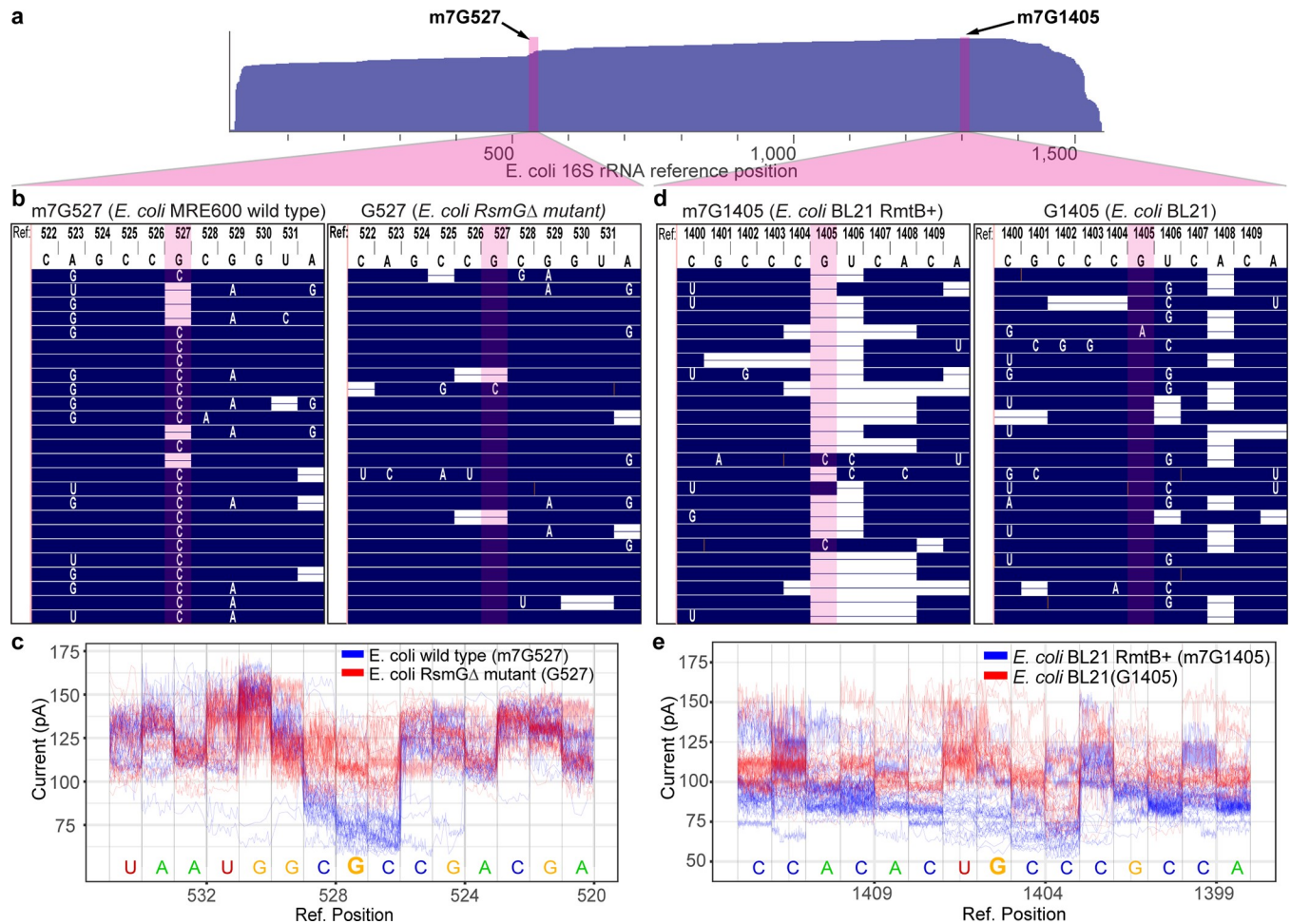
Fig 2. Detection of 7mG modifications in *E. coli* 16S rRNA. (a) Diagram showing the positions along *E. coli* 16S rRNA that correspond to the expanded sequence alignments in panels b-e. Arrows indicate the positions of G527 and G1405 in the *E. coli* reference. **(b)** Alignment of nanopore RNA sequence reads proximal to position 527 of *E. coli* 16S rRNA. Numbered letters at the top represent DNA bases in the reference 16S rRNA gene. Blue regions in the body of the panel denote agreement between reference DNA bases and nanopore RNA strand base-calls. White letters denote base call differences between the reference and the nanopore reads, and horizontal white bars represent base deletions in the nanopore RNA reads. Columns highlighted in red correspond to position 527. The left inset is *E. coli* str. MRE600 (wild type) 16S rRNA (m7G527), and the right inset is *RsmG* mutant strain 16S rRNA (canonical G527). **(c)** Nanopore ionic current traces proximal to position 527 of the *E. coli* 16S rRNA reference. Blue traces are for wild type *E. coli* 16S rRNA translocation events bearing m7G at position 527. Red traces are for mutant strain 16S rRNA translocation events bearing a canonical G at position 527. **(d)** Alignment of nanopore RNA sequence reads proximal to position 1405 of *E. coli* 16S rRNA. Use of colors, shapes, and letters are as described for panel (b). The left inset is engineered mutant *E. coli* str. BL21 (*RmtB+*) 16S rRNA (m7G1405); the right inset is *E. coli* str. BL21 16S rRNA (G1405). **(e)** Nanopore ionic current traces proximal to position 1405 of the *E. coli* 16S rRNA reference. Blue traces are for mutant strain 16S rRNA translocation events bearing m7G at position 1405. Red traces are for wild type 16S rRNA translocation events bearing a canonical G at position 1405.

https://doi.org/10.1371/journal.pone.0216709.g002

methylation at G527 [24]. We validated the absence of methylation at G527 in the *RsmG* deficient strain by chemical cleavage (**S4A Fig**). As predicted, a canonical guanosine base at position 527 in the mutant strain eliminated the reproducible base-call error seen in the wild type *E. coli* strain at that position (**Fig 2B** and **S4** Table).

As a further direct test, we compared ionic current signatures caused by G527 versus m7G527 by aligning traces at kmers bearing those bases. We found a reproducible decrease in ionic current for m7G527 compared to canonical G527, as anticipated (**Fig 2C**).

Typically, *E. coli* 16S rRNA contains only one m7G at position 527. However, some pathogenic strains that are resistant to aminoglycosides contain an additional m7G at position 1405

[25]. The enzymes responsible for G1405 methylation, such as RmtB [26], are thought to have originated from microbes that produce aminoglycosides and are shuttled on multidrug-resistance plasmids [27]. Given the pronounced signal difference for m7G at position 527, we thought it should also be possible to detect m7G in this context.

To this end, we engineered an *E. coli* strain that carried *RmtB* on an inducible plasmid (pLM1-RmtB, see Methods). We confirmed that this *RmtB+* strain was aminoglycoside resistant, (**S4B Fig**) consistent with N7 methylation of G1405. We then compared 16S rRNA sequence reads for this strain (*RmtB+*) with reads from the parent *E. coli* strain (BL21) without the plasmid (**Fig 2A** and **2D**). We observed an increase in deletions and base mis-calls in 16S rRNA reads for the *RmtB+* strain at position G1405 and the adjacent U1406. These mis-calls were absent in the 16S rRNA reads for the parent BL21 strain, which bears a canonical guanosine at G1405. Examination of ionic current segments containing G1405 and m7G1405 in RNA strands for the respective strains confirmed that m7G alters ionic current relative to canonical G (**Fig 2E**), as was observed at position 527. In this region, N4,2′-O-dimethylcytosine (position 1402) and 5-methylcytosine (position 1407) may also contribute to an aberrant ionic current, which could account for the sequence mis-calls proximal to those bases in the parent strain (**Fig 2D,** right panel).

Nanopore detection of epigenetic RNA modifications is not limited to m7G. While examining base mis-calls proximal to G527, we also noted a systematic miscall at U516 (predicted C, posterior probability 0.9086) (**S5 Fig**). This mis-called position had the highest probability variant in our marginCaller analysis (**S3 Table**). We hypothesized that this was due to pseudouridylation at U516 which is typical in *E. coli* 16S rRNA [28]. As a test, we compared nanopore reads for the wild type strain with reads for a mutant strain (*RsuAΔ*) bearing a canonical uridine at position 516 (see Methods). We found that mis-calls and ionic current deviations present at U516 in the wild type were absent in the mutant strain (**S5 Fig** and **S5 Table**) consistent with the hypothesis.

Another important feature of native nanopore 16S rRNA reads is that they are predominantly full-length. It has been established that more complete 16S rRNA sequences allow for improved taxonomic classification [29]. To test if full-length MinION 16S rRNA reads gave better classification than short reads, we sequenced purified 16S rRNA from three additional microbes (*Methanococcus maripaludis* str. S2, *Vibrio cholerae* str. A1552, and *Salmonella enterica* str. LT2). These were chosen to give a range of 16S rRNA sequence similarities to *E. coli* (68.1%, 90.4%, and 97.0% identity respectively). The 16S rRNA adapter sequence was altered slightly for each microbe to accommodate minor sequence differences at the 3′ end (see Methods). We binned reads by length, sampled 10,000 reads per bin for each microorganism, mixed them *in silico*, and aligned them to 16S rRNA sequences for all four microbes. A read was counted as correctly classified if it aligned to a 16S rRNA reference sequence for the source microorganism. As predicted, the classification accuracy increased with read length from 67.9% for short reads (200–600 bases) to 96.9% for long reads (>1000 bases) (**Fig 3A**).

We next assessed the utility of MinION 16S rRNA reads for microbial classification by aligning 1000+ base 16S rRNA MinION reads against two ribosomal RNA reference sequence databases, SILVA (release 128) and GRD [30]. SILVA is a broad database that consists of ~1.9 million 16S rRNA reference sequences. The GRD database consists of ~9,000 full-length curated 16S rRNA sequences. We then calculated the specificity and sensitivity of these reads from the alignments (**Table 1**, **S6 Table** **and S7 Table**). When using the SILVA database, both the specificity and sensitivity for 16S rRNA reads were 97% at the family level, 95% at the genus level, and 83% at the species level. We observed better performance for the curated GRD database, with both specificity and sensitivity of 99% at the family level, 94% at the genus level, and 93% at the species level.
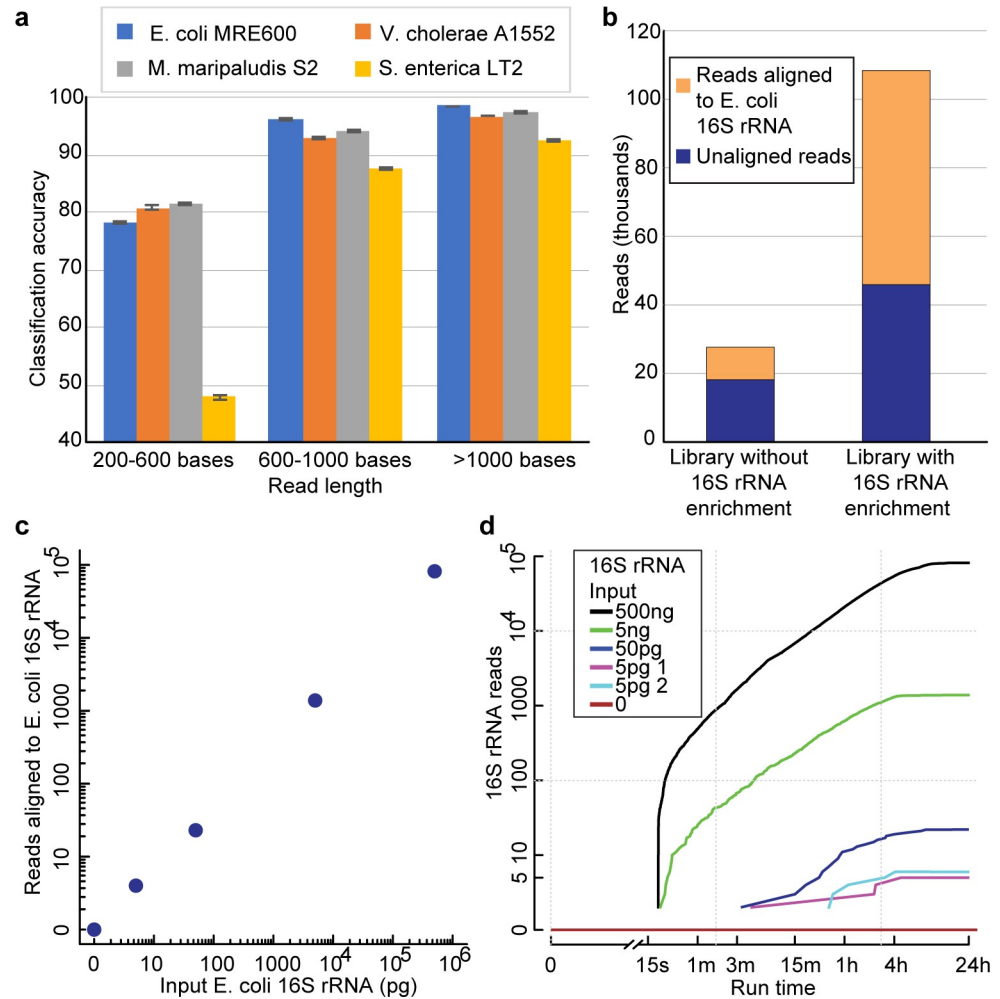
Fig 3. Nanopore 16S rRNA sequencing discriminates among microbes and detects *E. coli* 16S rRNA at low concentration. (**a**) Classification accuracy from an *in silico* mixture of 16S rRNA reads from four microbes. Reads were binned based on length and 10 iterations of classification using 10,000 randomly sampled reads from at least 15,000 reads per microbe were performed. A read was called as correctly classified if it aligned to one of the 16S rRNA reference sequences for that microbe. The error bars indicate one standard deviation for the 10 iterations. (**b**) 16S rRNA sequencing yield for libraries prepared from *E. coli* str. K12 total RNA with and without enrichment. Sequencing libraries were prepared from 5 μg total RNA. The enrichment library used a desthiobiotinylated version of the 16S rRNA-specific adapter, which was hybridized and selected for using magnetic streptavidin beads (see Methods). The two 16S rRNA sequencing libraries were then prepared essentially the same way. (**c**) 16S rRNA reads from sequencing libraries prepared with *E. coli* str. MRE600 16S rRNA titered into 4.5 μg total RNA from HEK 293T cells. (**d**) 16S read accumulation over time in titration sequencing runs. The lines correspond to libraries shown in c.

https://doi.org/10.1371/journal.pone.0216709.g003

Our early sequencing experiments required purifying 16S rRNA, which would be prohibitively slow for most practical applications. Therefore, we devised an enrichment strategy that

Table 1. Computed statistics for long 16S MinION rRNA reads classified against 16S rRNA sequence databases.

Database	Specificity (%)		Sensitivity (%)	
	SILVA	GRD	SILVA	GRD
Family	97.38	98.72	97.37	98.72
Genus	95.20	94.18	95.17	94.23
Species	82.48	92.60	82.45	92.66

https://doi.org/10.1371/journal.pone.0216709.t001

permits selective preparation of 16S rRNA from total bacterial RNA. This involved adding a desthiobiotin to the 16S rRNA adapter. The adapter was hybridized to 16S rRNA in a mixture, and then bound to streptavidin-conjugated magnetic beads. This allowed washing and removal of non-specific RNA in approximately one hour. 16S rRNA with associated adapter could be displaced from the beads with excess biotin, and the library preparation was then performed as usual (see methods). To test the enrichment method, we prepared 16S rRNA sequencing libraries from the same *E. coli* total RNA preparation with and without enrichment. We observed that enrichment increased the number of reads that aligned to 16S *E. coli* rRNA sequence >5-fold relative to the library without enrichment (**Fig 3B**).

This suggested that 16S rRNA could be selectively sequenced from a human total RNA background. To test this, we titered 5 pg to 500 ng of *E. coli* 16S rRNA into 4.5 μg total RNA from human embryonic kidney cells (HEK 293T) and prepared sequencing libraries (**Fig 3C**). The lowest mass (5 pg) approximates the amount of 16S rRNA from 300 *E. coli* cells [31]. 4.5 μg of total human RNA approximates the total RNA typically extracted from 1 ml of blood.

We observed a linear correlation between *E. coli* 16S rRNA reads and *E. coli* 16S rRNA concentrations over a 100,000-fold sample range (**Fig 3C**). In replicate 5 pg experiments, we observed only 4–5 16S rRNA reads, which nonetheless could be distinguished from the total human RNA negative control (0 16S rRNA reads in 24 hours). Because nanopore data are collected in real-time, we examined how rapidly *E. coli* 16S rRNA was detected in these MinION runs. We extracted acquisition times for all reads that aligned to *E. coli* 16S rRNA (**Fig 3D**). At concentrations ≥5ng, we found that the first 16S rRNA read occurred within ~20 seconds of the start of sequencing. This means that some 16S rRNA strands were immediately captured and processed by the MinION upon initiation of the sequencing run. At lower input amounts (<5 ng), we detected *E. coli* 16S rRNA strands in less than one hour.

Conclusions

In summary, we have shown that full length 16S rRNA can be sequenced in its native form using nanopore technology. Additionally, we have demonstrated that nanopore sequencing is sensitive to 7-methyl-guanosine and pseudouridine modifications in *E. coli* 16S rRNA. Further development of the sequencing adapter design will be necessary to capture the variability in 16S rRNA 3′ ends. This is a prerequisite for applying our selective 16S rRNA sequencing method to a metagenomic sample. A targeted approach similar to the one described here could also be applied to eurkaryotic ribosomal 18S rRNA and other rRNAs. Some nanopore RNA sequencing applications (e.g. strain-level taxonomic identification or detection of splice sites in transcript isoforms) will require better base-call accuracy than achieved in this study. These improvements seem likely based on prior evidence for MinION DNA sequencing where base call accuracies increased from 66% in 2014 [17] to 92% in 2015 [32]. Future efforts will be needed to demonstrate sensitivity to other modified nucleotides that are found in RNA, such as 1-methyladenosine and 5-methylcytosine [33–36]. With additional work on bioinformatic tools, modified nucleotides could be called directly from native nanopore RNA sequencing data. We anticipate that native 16S rRNA sequencing will ultimately find use for microbial identification in clinical and environmental settings.

Materials and methods

Cell culture and total RNA Isolation for 16S rRNA sequencing

E. coli strains BW25113 JW3718Δ and BW25113 JW2171Δ (strains hereafter referred to by gene deletion names *RsmGΔ* and *RsuAΔ*, respectively), deficient for 16S rRNA modifying enzymes RsmG and RsuA respectively, were purchased from the Keio Knockout collection

[37] (GE Dharmacon). *E. coli* strains K12 MG1655, *RsmG*Δ, *RsuA*Δ and *S. enterica* strain LT2 were grown in LB media (supplemented with 50 μg/ml kanamycin for *RsmG*Δ and *RsuA*Δ) at 37°C to an A_{600} = 0.8–1.0. Cells were harvested by centrifugation and total RNA was extracted with Trizol (Thermo Fisher) following the manufacturer's recommended protocol. All total RNA samples were treated with DNase I (NEB) (2U/10 ug RNA) in the manufacturer's recommended buffer at 37°C for 15 minutes. Following the DNase I reaction, RNA was extracted by acid phenol/chloroform extraction (pH 4.4, Fisher Scientific) and two rounds of chloroform extraction. RNA was precipitated with sodium acetate (pH 5.2) and ethanol. RNA was resuspended in nuclease-free water and stored at -80°C. For experiments where human RNA was used as a background, total RNA was extracted from 10^7 HEK 293T cells following the same steps.

16S rRNA purification

E. coli strain MRE600 16S rRNA was isolated from sucrose-gradient purified 30S subunits. *Vibrio cholerae* strain A1552 and *Methanococcus maripaludis* strain S2 16S rRNAs were isolated by gel purification from total RNA. 50–100 μg total RNA (DNase I treated) was heated to 95°C for 3 minutes in 7M urea/1xTE loading buffer and run on a 4% acrylamide/7M urea/TBE gel for 2.5 hours at 28W. Gel bands corresponding to 16S rRNA were cut from the gel. 16S rRNA was electroeluted into Maxi-size D-tube dialyzers (3.5 kDa MWCO, EMD Millipore) in 1X TBE for 2 hours at 100V. RNA was precipitated with sodium acetate and ethanol overnight at -20°C. RNA was pelleted washed once with 80% ethanol. Recovered RNA was resuspended in nuclease free water and quantitated using a Nanodrop spectrophotometer.

Oligonucleotides and 16S rRNA adapters

The 16S rRNA adapter was designed as a double-stranded DNA oligo. The bottom 40-nt strand has one 20-nucleotide region complementary to the 3' end of the 16S RNA, and a second 20-nt region complementary to the top strand (**S1A Fig**), with the sequence 5′-CCTAAG AGCAAGAAGAAGCCTAAGGAGGTGATCCAACCGC-3′. The top strand, which is directly ligated to the 16S rRNA, used the sequence 5′-pGGCTTCTTCTTGCTCTTAGGTAGTAGGT TC-3′ (p, 5′ phosphate). For *V. cholerae* and *M. maripaludis*, the 3′ terminal 20-nt of the bottom strand were slightly changed to yield adapters with perfectly complementary to their respective 16S rRNA 3′ ends. This resulted in the strands 5′-CCTAAGAGCAAGAAGAAGCCT AAGGAGGTGATCCAGCGCC-3′ and 5′-CCTAAGAGCAAGAAGAAGCCAGGAGGTGATCCAG CCGCAG-3′, respectively. To make a 16S rRNA adapter, top and the bottom strands were hybridized at 10 μM each in a buffer containing 10 mM Tris-HCl (pH 8.0), 1 mM EDTA, and 50 mM NaCl. The mixtures were heated to 75°C for 1 minute before being slowly cooled to room temperature in a thermocycler. We confirmed the adapter hybridizes and ligates to *E. coli* str. MRE600 16S rRNA 3′ end by a gel electrophoresis-based assay with a 6-FAM-labeled version of the top strand (**S1B Fig**). For experiments where 16S rRNA was enriched from a total RNA background, a desthiobiotin was added to the 5′ terminus of the bottom strand. All adapter oligonucleotides were synthesized by IDT.

Purified 16S rRNA Sequencing Library Preparation

Sequencing libraries of purified 16S rRNA for *E. coli* str. MRE600, *V. cholerae* str. A1552, and *M. maripaludis* str. S2 were prepared as follows: 2 pmol 16S rRNA adapter and 1.5 μg purified 16S rRNA (approximately 3 pmol) were added to a 15 μL reaction in 1x Quick Ligase buffer with 3000U T4 DNA ligase (New England Biolabs). The reaction was incubated at room temperature for 10 minutes. These reactions were cleaned up using 1.8x volume of RNAclean XP

beads (Beckman Coulter), washed once with 80% ethanol and resuspended in 20 µl nuclease-free water. The RNA sequencing adapter (Oxford Nanopore Technologies) was ligated to the RNA library following manufacturer recommended protocol.

Preparation of RNA Sequencing libraries enriched for 16S rRNA

Enrichment-based 16S sequencing libraries were prepared for *E. coli* strains K-12 MG1655, BL21 DE3 pLys, BL21 DE3 pLys pLM1-*RmtB*+, BL21 DE3 pLys pLM1-*RmtB*Δ, *RsmG*Δ, *RsuA*Δ, and *S. enterica* strain LT2. 16S rRNA-enriched sequencing libraries were essentially prepared as described for purified 16S rRNA with the following exceptions: 15 pmol of 5′ desthiobiotinylated 16S rRNA adapter was added to 4.5–5 µg total RNA in 10 µL buffer containing 10 mM Tris-HCl pH 8, 1 mM EDTA and 50 mM NaCl. The mixture was heated to 50˚C for 1 minute and slowly cooled to room temperature in a thermocycler (~10 minutes). The mixture was then incubated at room temperature for 20 minutes with 100 µL MyOne C1 magnetic streptavidin beads (Thermo Fisher) in 10 mM Tris-HCl (pH 8), 1 mM EDTA, 500 mM NaCl, and 0.025% NP-40 (Buffer A). The beads were washed once with an equal volume of Buffer A and once with an equal volume of buffer containing 10 mM Tris-HCl (pH 8), 1 mM EDTA, 150 mM NaCl (Buffer B). To elute 16S rRNA-enriched RNA, 20 µl Buffer B amended with 50 µM biotin was incubated with the beads at 37˚C for 30 minutes. The hybridized 16S rRNA adapter was then ligated by bringing the mixture to 40 µL 1x Quick Ligase buffer (New England Biolabs) and adding 3000U of T4 DNA ligase (New England Biolabs). The rest of the library preparation was performed the same as described for purified 16S rRNA sequencing libraries.

In vivo methylation of 16S rRNA G1405

The *RmtB* gene was purchased as a synthetic gBlock from IDT with the sequence from Gen-Bank accession EU213261.1. pET-32a+ (EMD Millipore) and *RmtB* gBlock were digested with XhoI and NdeI. Digested plasmid and gBlock were ligated with T4 DNA ligase (NEB) to create plasmid pLM1-*RmtB*+. To create *RmtB* null plasmid, pLM1-*RmtB*Δ, XhoI and NdeI digested pET-32a+ was end repaired and ligated. Plasmids were transformed into *E. coli* DH5a cells (NEB) and confirmed by Sanger sequencing. Confirmed clones for pLM1-*RmtB*+ and pLM1-*RmtB*Δ were transformed into *E. coli* BL21 DE3 pLysS cells to create expression strains. To methylate G1405 in 16S rRNA, *E. coli* BL21 DE3 pLys pLM1-*RmtB*+ cells were cultured in 150 ml LB at 37˚C with Ampicillin (100 ug/ml) until OD_{600} ~ 0.4. Cultures were diluted into 1 L in pre-warmed LB media with Ampicillin (100 ug/ml), and plasmid expression was induced with 1 mM IPTG. Cultures were grown at 37˚C to an OD_{600} ~ 0.4. Cells were then pelleted and resuspended in 30 ml of 25 mM Tris-HCl (pH 7.5), 100 mM NH_4Cl, 15 mM $MgCl_2$, 5 mM β-mercaptoethanol. Cells were harvested for RNA purification or flash frozen in liquid nitrogen and stored at -80˚C.

Chemical probing for m7G

Chemical probing for 7-methylguanosine in *E. coli* 16S rRNA was carried out essentially as described previously (Recht et al. 1996). Approximately 10 pmol 16S rRNA or RNA extracted from 70S ribosomes was resuspended in 20 µl 0.5 M Tris-HCl (pH 8.2). Selective reduction of m7G was performed by adding 5 µl freshly made 0.5 M sodium borohydride solution. The reaction was incubated on ice in the dark for 30 minutes. The reaction was ended by the addition of 10 µl 3 M sodium acetate (pH 5.2) and precipitated with ethanol. Pellets were washed once with 80% ethanol. RNA was pelleted by centrifugation and resuspended in 20 µl 1 M aniline/glacial acetic acid solution (1:1.5) (pH 4.5). RNA cleavage proceeded by incubating the

reaction at 60°C for 10 minutes in the dark. The reaction was ended by the addition of 20 μl 0.5 M Tris-HCl (pH 8.2), and the RNA was isolated by extracting with phenol/chloroform/iso-amyl alcohol. RNA was precipitated from the aqueous phase, pelleted and washed with 80% ethanol. RNA pellets were resuspended in 2.5 μl nuclease free water. Primer extension to determine the site of m7G-specific cleavage was carried out as described (Merryman and Noller 1998). To detect G527 methylation, the primer 5′–CGTGCGCTTTACGCCCA–3′ was used.

MinION sequencing of 16S rRNA

MinION sequencing of 16S rRNA libraries was performed using MinKNOW version 1.1.30. The flow cells used were FLO-MIN106 SpotON version. ONT's Metrichor base-calling software (1D RNA Basecalling for FLO-MIN106 v1.134 workflow) takes this raw signal and produces base-called FASTQ sequence in the 5′ to 3′ order after reads are reversed. During the course of these experiments, ONT made a new local base-caller available, named Albacore. We performed base-calling for the sequencing runs using Albacore v1.0.1, and performed all alignment-based analyses with the newer sequence data.

Data analysis

FastQ sequences were extracted using poretools v0.6 [38] and then sequence alignment was performed using marginAlign v0.1 [17] (using BWA-MEM version 0.7.12–41044; parameter "-x ont2d" [39]). The statistics were calculated using marginStats v0.1 [17]. We then created assembly hubs to visualize these alignment on the UCSC genome browser using createAssemblyHub utility in marginAlign suite [17]. We calculated read identity as matches / (matches + mismatches + insertions + deletions). We used marginAlign expectation maximization (EM) to estimate the error model from the sequence data. Using these high-quality alignments, we estimated substitution rates for the RNA nucleotides in MinION data. Using these high-quality alignments, we then performed variant calling using marginCaller v0.1 [17] to predict variants and associate systematic sequence mis-calls with putative base modifications. To test for systematic k-mer biases in MinION RNA data, we compared 5-mers in reads and the known 16S rRNA reference.

Nanopore ionic current visualization

We used nanoraw v0.4.2 [40] to visualize ionic current traces for 16S rRNA reads from different *E. coli* strains that were sequenced on the MinION. We used the software with its default settings. We chose graphmap [41] as the aligner in nanoraw, and the argument 'ont' (now 'pA' in nanoraw v0.4.2) as the option for normalizing raw ionic currents. The ionic current plots were created using the plot_genome_location function. For all of the ionic current analysis, we inverted the reference sequence since the present MinION native RNA sequencing chemistry sequences native RNA molecules in the 3′-5′ direction.

Assessing read-length effect on microbial classification

Binning reads by length (200–600, 600–1000, >1000 bases), we randomly sampled 10,000 reads from at least 15,000 reads per bin for each microbe. These reads were then mixed *in silico* and aligned using marginAlign v0.1 [17]. A read was called as correctly classified if its best alignment was to one of the 16S rRNA reference sequences for that microbe (in a dataset containing 16S rRNA sequences for all 4 microbes). Ten classification iterations were performed for each of the bins.

Microbial classification from long nanopore 16S rRNA reads

Long RNA reads (>1000 bases) collected independently for each of the four microbes were searched against two large microbial databases, SILVA[42] and GRD[43]. We aligned 10,000 randomly sampled long MinION reads from each microbial 16S rRNA sequencing data to the broad databases and computed the following summary statistics:

Sensitivity—Ratio of true positives over the sum of true positives and false negatives. For example, for E. coli this would be the ratio of the number of reads correctly classified as E. coli over the total number of reads from E. coli data that were classified (as anything).

Specificity—Ratio of true negatives over the sum of true negatives and false positives. For example, for E. coli this would be the ratio of the total number of reads correctly classified as non-E. coli over the total number of reads from non-E. coli data that were classified.

Supporting information

S1 Fig. Design of an oligonucleotide adapter to prepare 16S rRNA for nanopore direct RNA strand sequencing. (a) Schematic of oligonucleotide adapter hybridized to the 3′ end of 16S rRNA. The adapter bottom (splint) strand can hybridize 20 nt to the conserved Shine-Dalgarno sequence on the 16S rRNA 3′ end. The adapter top strand can hybridize 20 nt to the 5′ end of the adapter bottom strand as shown. The 3′ terminal hydroxyl of the 16S rRNA and the 5′ terminal phosphate of the adapter top strand can be covalently joined by T4 DNA ligase. The broken line indicates an unshown ssDNA 3′ overhang, which is necessary for ligating the ONT sequencing adapter (see Methods). **(b)** Denaturing acrylamide gel analysis of a ligation reaction demonstrating the 16S adapter hybridizes and ligates to *E. coli* 16S rRNA 3′ ends. The left panel shows the unstained gel image. The lower band is a fluorescent, 3′-6-FAM-labeled version of adapter top strand. Lanes 1–3 show pre-ligation reaction samples for: Lane 1) RNA adapter ligation positive control with a polyA-specific adapter containing a 3′ terminal oligo dT_{10} overhang (replaces 16S-specific overhang) and the 6-FAM labeled top strand. A synthetic 288mer polyA RNA is used as the control substrate. Lane 2) 6-FAM-labeled 16S rRNA-specific adapter and purified 16S rRNA from *E. coli*. Lane 3) ligation negative control with adapter present, E. coli 16S rRNA, but without T4 DNA ligase. Lanes 4–6 show post-ligation reaction samples for: Lane 4) RNA adapter ligation positive control with polyA RNA 288mer. Lane 5) 16S rRNA reaction with 6-FAM labeled 16S rRNA-specific adapter. Lane 6) ligation negative control. The size-shifted fluorescent top strand indicates ligation to the 16S rRNA 3′ end (Open arrow). The right image is the same gel stained with SybrGold. Position of the 16S rRNA is indicated by open arrows.
(PNG)

S2 Fig. Alignment metrics for Enolase 2 polyA calibration strand and E. coli 16S rRNA. Alignments were performed using marginAlign (guide alignments from BWA MEM "-x ont2d" followed by chaining). **(a)** Identity vs. read length for Enolase 2. The intensity of blue shading represents the density of the data distribution. The black dots at the edge of the distributions represent regions where there are only a few data points. **(b)** Identity vs. read length for 16S *E. coli* rRNA. The blue shading and black dots are as in panel (a). **(c)** Coverage across reference for Enolase 2 calibration strand. **(d)** Coverage across reference for 16S *E. coli* rRNA. **(e)** Alignment of 100,000+ Enolase 2 reads to the reference sequence.
(PDF)

S3 Fig. The matrix for substitution emissions for Enolase 2 calibration strand and E. coli 16S rRNA. This matrix was determined using marginAlign EM. The matrix shows low rates of C-to-G and G-to-C substitutions, relative to the other substitutions. The color scheme is

fitted on a log scale, and the substitution values are on an absolute scale.
(PDF)

S4 Fig. Confirmation of guanosine N7-methylation (m7G) at positions 527 and 1405 in E. coli 16S rRNA. (a) Canonical m7G527 is present in wild-type *E. coli* and absent in the RsmG deficient *E. coli* strain. Sodium borohydride/aniline cleavage was used to assess the presence or absence of m7G527 in 16S rRNA from wild-type *E. coli* str. MRE600 or RsmG deficient (mutant) *E. coli* str. BW25113 JW3718Δ. Sequencing lanes 1–4 are labeled for A, C, G, and U of the RNA sequence. The respective lanes are reactions containing the complementary ddNTP to the RNA sequence. Wild-type 16S rRNA from *E. coli* str. MRE600 is used as the template. Lanes 6 and 8: sodium borohydride/aniline treatment (labeled +) of 16S rRNA from wild type *E. coli* and 16S rRNA from RsmG mutant *E. coli*, respectively. Strand cleavage, which is dependent on the presence of m7G, should result in an primer extension stop at C528, 1-nt preceding G527 (position 527 is marked by an asterisk). Lane 5 and 7: untreated 16S rRNA for wild type and mutant 16S rRNA. Primer extension products were run on denaturing 6% acrylamide gel, and imaged using a phosphorimager. A spontaneous RT stop appears at m7G527 in all lanes where wild-type 16S rRNA was used as the template, which has been observed previously. **(b)** RmtB confers a kanamycin resistance phenotype consistent with G1405 N7-methylation in 16S rRNA from an engineered *E. coli* strain. Serial dilutions from 10^{-2} to 10^{-6} (Left to Right) of *E. coli* BL21 DE3 pLysS strains transformed with pLM1-RmtB and negative control pLM1-RmtBΔ were spotted on LB agar plates. The pLM1 plasmids use pET32a as the backbone, which contains an ampicillin resistance gene. The RmtB gene is under the control of a lactose inducible T7 promoter. Plates are supplemented with: 100 μg/ml Ampicillin (top), 100 μg/ml Ampicillin + 200 μg/ml Kanamycin + 1% glucose (middle), 100 μg/ml Ampicillin + 200 μg/ml Kanamycin + 1 mM IPTG (bottom). Apparent leaky expression of the pLM-RmtB vector leads to some cell survival, even under non-inducing conditions and in the presence of glucose. RmtB is known to confer high-level resistance to kanamycin.
(PNG)

S5 Fig. Inference of pseudouridine in E. coli 16S rRNA direct sequencing reads. (a) Comparison of aligned reads from strands containing putative pseudouridine versus strands bearing canonical uridine at position 516. Reads are aligned to the *E. coli* MRE600 *rrnD* 16S rRNA reference sequence. Shown are twenty-five 16S rRNA reads from separate sequencing runs for *E. coli* str. MRE600 (wild type), which bears a pseudouridine at U516 (Ψ516) and an *RsuA* deficient strain (*RsuA*Δ mutant), which has a canonical U at position 516. Green shading indicates the position of U516 (shown as a T in the reference gene sequence). **(b)** Aligned ionic current traces from approximately thirty 16S rRNA reads covering position U516 from wild-type *E. coli* and *RsuA*Δ mutan strain. Pseudouridylation site, U516, is shown in large font. The sequence is shown 3′-to-5′ because ionic current signal is 3′-to-5′. Numbering uses standard *E. coli* 16S rRNA numbering.
(PNG)

S1 Table. Error rate profile for Enolase 2 transcript and 16S E. coli rRNA. Error models were estimated using marginAlign (guide alignments from BWA MEM "-x ont2d" followed by chaining). Statistics were generated using marginStats.
(DOCX)

S2 Table. Over and Under represented 5-mers comparison for Enolase 2 (left) and *E. coli* 16S rRNA (right). 5-mers were counted and compared for RNA read data and their respective reference sequences. LogFC represents log fold-change.
(DOCX)

S3 Table. Sequence variants in E. coli MRE600 16S rRNA MinION reads based on alignments to E. coli str. MRE600 reference sequence for rrnD 16S rRNA.
(DOCX)

S4 Table. Sequence variants in E. coli RsmG 16S rRNA MinION reads based on alignments to E. coli str. MRE600 reference sequence for rrnD 16S rRNA.
(DOCX)

S5 Table. Sequence variants in E. coli RsuAΔ 16S rRNA MinION reads based on alignments to E. coli str. MRE600 reference sequence for rrnD 16S rRNA.
(DOCX)

S6 Table. Performance measurements for long 16S rRNA reads (>1000 bases) when classified against the SILVA database.
(DOCX)

S7 Table. Performance measurements for long 16S rRNA reads (>1000 bases) when classified against the GRD database.
(DOCX)

Acknowledgments

Author Contributions

Conceptualization: Andrew M. Smith, Miten Jain, Daniel R. Garalde, Mark Akeson.

Data curation: Miten Jain.

Formal analysis: Andrew M. Smith, Miten Jain, Mark Akeson.

Funding acquisition: Mark Akeson.

Investigation: Andrew M. Smith, Miten Jain.

Methodology: Andrew M. Smith, Miten Jain, Logan Mulroney.

Project administration: Mark Akeson.

Software: Miten Jain.

Supervision: Mark Akeson.

Validation: Andrew M. Smith, Miten Jain, Mark Akeson.

Writing – original draft: Andrew M. Smith, Miten Jain.

Writing – review & editing: Daniel R. Garalde, Mark Akeson.

References

1. Garalde DR, Snell EA, Jachimowicz D, Sipos B, Lloyd JH, Bruce M, et al. Highly parallel direct RNA sequencing on an array of nanopores. Nat Methods. 2018; 15: 201–206. https://doi.org/10.1038/nmeth.4577 PMID: 29334379

2. Kasianowicz JJ, Brandin E, Branton D, Deamer DW. Characterization of individual polynucleotide molecules using a membrane channel. Proc Natl Acad Sci U S A. 1996; 93: 13770–13773. https://doi.org/10.1073/pnas.93.24.13770 PMID: 8943010

3. Lieberman KR, Cherf GM, Doody MJ, Olasagasti F, Kolodji Y, Akeson M. Processive replication of single DNA molecules in a nanopore catalyzed by phi29 DNA polymerase. J Am Chem Soc. 2010; 132: 17961–17972. https://doi.org/10.1021/ja1087612 PMID: 21121604

4. Cherf GM, Lieberman KR, Rashid H, Lam CE, Karplus K, Akeson M. Automated forward and reverse ratcheting of DNA in a nanopore at 5-Å precision. Nat Biotechnol. 2012; 30: 344–348. https://doi.org/10.1038/nbt.2147 PMID: 22334048

5. Manrao EA, Derrington IM, Laszlo AH, Langford KW, Hopper MK, Gillgren N, et al. Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase. Nat Biotechnol. 2012; 30: 349–353. https://doi.org/10.1038/nbt.2171 PMID: 22446694

6. Jain M, Olsen HE, Paten B, Akeson M. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome Biol. 2016; 17: 239. https://doi.org/10.1186/s13059-016-1103-0 PMID: 27887629

7. Smith AM, Abu-Shumays R, Akeson M, Bernick DL. Capture, Unfolding, and Detection of Individual tRNA Molecules Using a Nanopore Device. Front Bioeng Biotechnol. 2015; 3: 91. https://doi.org/10.3389/fbioe.2015.00091 PMID: 26157798

8. Vilfan ID, Tsai Y-C, Clark TA, Wegener J, Dai Q, Yi C, et al. Analysis of RNA base modification and structural rearrangement by single-molecule real-time detection of reverse transcription. J Nanobiotechnology. 2013; 11: 8. https://doi.org/10.1186/1477-3155-11-8 PMID: 23552456

9. Ozsolak F, Milos PM. Single-molecule direct RNA sequencing without cDNA synthesis. Wiley Interdiscip Rev RNA. 2011; 2: 565–570. https://doi.org/10.1002/wrna.84 PMID: 21957044

10. Woese CR, Fox GE. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc Natl Acad Sci U S A. 1977; 74: 5088–5090. https://doi.org/10.1073/pnas.74.11.5088 PMID: 270744

11. Clarridge JE 3rd. Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. Clin Microbiol Rev. 2004; 17: 840–62, table of contents. https://doi.org/10.1128/CMR.17.4.840-862.2004 PMID: 15489351

12. Moazed D, Noller HF. Interaction of antibiotics with functional sites in 16S ribosomal RNA. Nature. 1987; 327: 389–394. https://doi.org/10.1038/327389a0 PMID: 2953976

13. Sergeeva OV, Bogdanov AA, Sergiev PV. What do we know about ribosomal RNA methylation in Escherichia coli? Biochimie. 2015; 117: 110–118. https://doi.org/10.1016/j.biochi.2014.11.019 PMID: 25511423

14. Shine J, Dalgarno L. The 3′-Terminal Sequence of Escherichia coli 16S Ribosomal RNA: Complementarity to Nonsense Triplets and Ribosome Binding Sites. Proceedings of the National Academy of Sciences. 1974; 71: 1342–1346.

15. Nakagawa S, Niimura Y, Miura K-I, Gojobori T. Dynamic evolution of translation initiation mechanisms in prokaryotes. Proc Natl Acad Sci U S A. 2010; 107: 6382–6387. https://doi.org/10.1073/pnas.1002036107 PMID: 20308567

16. Keller MW, Rambo-Martin BL, Wilson MM, Ridenour CA, Shepard SS, Stark TJ, et al. Author Correction: Direct RNA Sequencing of the Coding Complete Influenza A Virus Genome. Sci Rep. 2018; 8: 15746. https://doi.org/10.1038/s41598-018-34067-6 PMID: 30341398

17. Jain M, Fiddes IT, Miga KH, Olsen HE, Paten B, Akeson M. Improved data analysis for the MinION nanopore sequencer. Nat Methods. 2015; 12: 351–356. https://doi.org/10.1038/nmeth.3290 PMID: 25686389

18. Laszlo AH, Derrington IM, Ross BC, Brinkerhoff H, Adey A, Nova IC, et al. Decoding long nanopore sequencing reads of natural DNA. Nat Biotechnol. 2014; 32: 829–833. https://doi.org/10.1038/nbt.2950 PMID: 24964173

19. Rand AC, Jain M, Eizenga JM, Musselman-Brown A, Olsen HE, Akeson M, et al. Mapping DNA methylation with high-throughput nanopore sequencing. Nat Methods. 2017; https://doi.org/10.1038/nmeth.4189 PMID: 28218897

20. Simpson JT, Workman RE, Zuzarte PC, David M, Dursi LJ, Timp W. Detecting DNA cytosine methylation using nanopore sequencing. Nat Methods. 2017; https://doi.org/10.1038/nmeth.4184 PMID: 28218898

21. Modomics—A Database of RNA Modifications [Internet]. [cited 13 Apr 2017]. Available: http://modomics.genesilico.pl

22. Moazed D, Noller HF. Transfer RNA shields specific nucleotides in 16S ribosomal RNA from attack by chemical probes. Cell. 1986; 47: 985–994. PMID: 2430725

23. Zueva VS, Mankin AS, Bogdanov AA, Baratova LA. Specific fragmentation of tRNA and rRNA at a 7-methylguanine residue in the presence of methylated carrier RNA. Eur J Biochem. Wiley Online Library; 1985; 146: 679–687. PMID: 2578958

24. Okamoto S, Tamaru A, Nakajima C, Nishimura K, Tanaka Y, Tokuyama S, et al. Loss of a conserved 7-methylguanosine modification in 16S rRNA confers low-level streptomycin resistance in bacteria. Mol Microbiol. 2007; 63: 1096–1106. https://doi.org/10.1111/j.1365-2958.2006.05585.x PMID: 17238915

25. Beauclerk AA, Cundliffe E. Sites of action of two ribosomal RNA methylases responsible for resistance to aminoglycosides. J Mol Biol. 1987; 193: 661–671. PMID: 2441068

26. Doi Y, Yokoyama K, Yamane K, Wachino J-I, Shibata N, Yagi T, et al. Plasmid-mediated 16S rRNA methylase in Serratia marcescens conferring high-level resistance to aminoglycosides. Antimicrob Agents Chemother. 2004; 48: 491–496. https://doi.org/10.1128/AAC.48.2.491-496.2004 PMID: 14742200

27. Wachino J-I, Arakawa Y. Exogenously acquired 16S rRNA methyltransferases found in aminoglycoside-resistant pathogenic Gram-negative bacteria: an update. Drug Resist Updat. 2012; 15: 133–148. https://doi.org/10.1016/j.drup.2012.05.001 PMID: 22673098

28. Bakin A, Kowalak JA, McCloskey JA, Ofengand J. The single pseudouridine residue in Escherichia coli 16S RNA is located at position 516. Nucleic Acids Res. 1994; 22: 3681–3684. https://doi.org/10.1093/nar/22.18.3681 PMID: 7524026

29. Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Schleifer K-H, et al. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nat Rev Microbiol. 2014; 12: 635–645. https://doi.org/10.1038/nrmicro3330 PMID: 25118885

30. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2013; 41: D590–6. https://doi.org/10.1093/nar/gks1219 PMID: 23193283

31. Milo R, Philips R. » How many ribosomes are in a cell? [Internet]. [cited 12 Apr 2017]. Available: http://book.bionumbers.org/how-many-ribosomes-are-in-a-cell/

32. Ip CLC, Loose M, Tyson JR, de Cesare M, Brown BL, Jain M, et al. MinION Analysis and Reference Consortium: Phase 1 data release and analysis. F1000Res. 2015; 4: 1075. https://doi.org/10.12688/f1000research.7201.1 PMID: 26834992

33. Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Salmon-Divon M, Ungar L, Osenberg S, et al. Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature. 2012; 485: 201–206. https://doi.org/10.1038/nature11112 PMID: 22575960

34. Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE, Jaffrey SR. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. Cell. 2012; 149: 1635–1646. https://doi.org/10.1016/j.cell.2012.05.003 PMID: 22608085

35. Safra M, Sas-Chen A, Nir R, Winkler R, Nachshon A, Bar-Yaacov D, et al. The m1A landscape on cytosolic and mitochondrial mRNA at single-base resolution. Nature. 2017; 551: 251–255. https://doi.org/10.1038/nature24456 PMID: 29072297

36. Li X, Xiong X, Zhang M, Wang K, Chen Y, Zhou J, et al. Base-Resolution Mapping Reveals Distinct m1A Methylome in Nuclear- and Mitochondrial-Encoded Transcripts. Mol Cell. 2017; 68: 993–1005.e9. https://doi.org/10.1016/j.molcel.2017.10.019 PMID: 29107537

37. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, et al. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol. 2006; 2: 2006.0008.

38. Loman NJ, Quinlan AR. Poretools: a toolkit for analyzing nanopore sequence data. Bioinformatics. 2014; 30: 3399–3401. https://doi.org/10.1093/bioinformatics/btu555 PMID: 25143291

39. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM [Internet]. arXiv [q-bio.GN]. 2013. Available: http://arxiv.org/abs/1303.3997

40. Stoiber MH, Quick J, Egan R, Lee JE, Celniker SE, Neely R, et al. De novo Identification of DNA Modifications Enabled by Genome-Guided Nanopore Signal Processing [Internet]. bioRxiv. 2016. p. 094672. https://doi.org/10.1101/094672

41. Sović I, Šikić M, Wilm A, Fenlon SN, Chen S, Nagarajan N. Fast and sensitive mapping of nanopore sequencing reads with GraphMap. Nat Commun. 2016; 7: 11307. https://doi.org/10.1038/ncomms11307 PMID: 27079541

42. Silva [Internet]. [cited 28 Aug 2017]. Available: https://www.arb-silva.de/

43. RDP Release 11—Sequence Analysis Tools [Internet]. [cited 28 Aug 2017]. Available: https://rdp.cme.msu.edu/