High-throughput mutate-map-rescue evaluates SHAPE-directed RNA structure and uncovers excited states

SIQI TIAN,1 PABLO CORDERO,2 WIPAPAT KLADWANG,3 and RHIJU DAS1,2,3
1Department of Biochemistry, Stanford University, Stanford, California 94305, USA
2Biomedical Informatics Program, Stanford University, Stanford, California 94305, USA
3Department of Physics, Stanford University, Stanford, California 94305, USA

ABSTRACT
The three-dimensional conformations of noncoding RNAs underpin their biochemical functions but have largely eluded experimental characterization. Here, we report that integrating a classic mutation/rescue strategy with high-throughput chemical mapping enables rapid RNA structure inference with unusually strong validation. We revisit a 16S rRNA domain for which SHAPE (selective 2′-hydroxyl acylation with primer extension) and limited mutational analysis suggested a conformational change between apo- and holo-ribosome conformations. Computational support estimates, data from alternative chemical probes, and mutate-and-map (M2) experiments highlight issues of prior methodology and instead give a near-crystallographic secondary structure. Systematic interrogation of single base pairs via a high-throughput mutation/rescue approach then permits incisive validation and refinement of the M2-based secondary structure. The data further uncover the functional conformation as an excited state (20 ± 10% population) accessible via a single-nucleotide register shift. These results correct an erroneous SHAPE inference of a ribosomal conformational change, expose critical limitations of conventional structure mapping methods, and illustrate practical steps for more incisively dissecting RNA dynamic structure landscapes.

Keywords: RNA folding; secondary structure; SHAPE; compensatory rescue; mutate-and-map; ribosome

INTRODUCTION
RNA plays critical roles in diverse cellular and viral processes ranging from information transfer to metabolite sensing to translation (Nudler and Mironov 2004; Amaral et al. 2008; Zhang et al. 2009, 2010; Breaker 2012). For the most complex of these processes, RNAs must adopt and interconvert between specific three-dimensional structures (Zhang et al. 2010), but for the vast majority of systems, these conformations remain experimentally uncharacterized. In particular, current prediction methods do not yet offer clear metrics of statistical confidence, probe the possibility of multiple conformations (including weakly populated “excited” states), or provide routes to cross-validation through independent experiments. The situation is particularly problematic since RNAs can form multiple alternative secondary structures whose helices are mutually exclusive (Nudler and Mironov 2004; Henkin 2008; Haller et al. 2011).

One approach for probing large numbers of RNA molecules involves chemically modifying RNA and reading out these events via electrophoresis or deep sequencing (Mitra et al. 2008; Lucks et al. 2011; Yoon et al. 2011). Numerous reagents, including protein nucleases (Walczak et al. 1996; Grover et al. 2011; Siegfried et al. 2011), alkylating chemicals such as dimethyl sulfate (DMS) (Wells et al. 2000; Tijerina et al. 2007; Cordero et al. 2012a), and hydroxyl radicals (Adilakshmi et al. 2006; Das et al. 2008; Ding et al. 2012), have been leveraged to modify or cleave RNA in a structure-dependent manner. Protection of nucleotides from modification, typically signaling the formation of base pairs, can guide manual or automatic secondary structure inference (Mathews et al. 2004; Mitra et al. 2008; Vasa et al. 2008). Strong cases have been made for using reagents that covalently modify 2′-hydroxyls followed by readout via primer extension (SHAPE) (Merino et al. 2005; Mortimer and Weeks 2007; Deigan et al. 2009; McGinnis et al. 2012) and then applying these data as pseudoenergy bonuses in free-energy minimization algorithms, such as RNAstructure (Mathews et al. 2004; Reuter and Mathews 2010; Hajdin et al. 2013).

Assessing the accuracy of structure mapping methods is becoming a major issue as these approaches are being applied to large RNA systems—such as entire cellular transcriptomes—for which crystallographic, spectroscopic, or phylogenetic
methods cannot be brought to bear (Kertesz et al. 2010; Underwood et al. 2010; Kladwang et al. 2011a). In some cases, the resulting models have disagreed with accepted structures (Quarrier et al. 2010; Kladwang et al. 2011c; Hajdin et al. 2013; Sükös et al. 2013; Rice et al. 2014), motivating efforts to estimate uncertainty (Kladwang et al. 2011c; Ramachandran et al. 2013), to incorporate alternative mapping strategies (Cordero et al. 2012a; Kwok et al. 2013), and to integrate mapping with systematic mutagenesis (mutate-and-map [M2] (Kladwang and Das 2010; Kladwang et al. 2011a,b; Cordero et al. 2014). Even these improvements do not provide routes to validation through independent experiments, and structure inferences remain under question (Wenger et al. 2011).

A powerful technique for validating RNA structures is the rescue of disruptive perturbations from single mutations through compensatory mutations that restore Watson-Crick pairs. Mutation/rescue has typically been read out through catalytic reactions such as self-cleavage (Wu and Huang 1992; Macnaughton et al. 1993). For RNAs without well-established functional assays, chemical mapping offers a potentially general readout for mutation/rescue, but has been explored in only a limited fashion (Huang et al. 1992; Macnaughton et al. 1993). For RNAs without well-established functional assays, chemical mapping offers a potentially general readout for mutation/rescue, but has been explored in only a limited fashion (Huang et al. 1992; Macnaughton et al. 1993). For RNAs without well-established functional assays, chemical mapping offers a potentially general readout for mutation/rescue, but has been explored in only a limited fashion (Huang et al. 1992; Macnaughton et al. 1993).

RESULTS

Reproducibility and robustness of SHAPE modeling

Before applying newer approaches, we carried out standard SHAPE chemical mapping on the 126–235 RNA readout by capillary electrophoresis (CE), seeking to reproduce prior studies. Our SHAPE data with the acetylation reagent 1-methyl-7-nitroisatoic anhydride (1M7) (Mortimer and Weeks 2007), averaged across four replicates, agreed with previously published data (Deigan et al. 2009) for the 126–235 RNA in the context of the full ribosome and in different flanking sequences (Fig. 1D). In particular, sequence assignments agreed with prior work and were additionally verified by co-loading chemical mapping samples with dideoxy sequencing ladders and measurements using the Illumina-based MAP-seq (Multiplexed Accessibility Probing readout through sequencing) protocol (Supplemental Fig. S1A,B; Seetin et al. 2014). In ribosome crystals and phylogenetic analyses, the 126–235 RNA domain forms a junction of four helices (Fig. 1B; Zhang et al. 2009); whereas the prior SHAPE-directed model (referred to hereafter as the “1D-data-guided model”) proposed two different helices (Fig. 1A). The SHAPE data were consistent with both the ribosome crystallographic secondary structure as well as with the prior 1D-data-guided secondary structure. In both models, the nucleotides observed to be SHAPE-reactive either do not form Watson-Crick base pairs or are next to such nucleotides.

We then carried out secondary structure predictions guided by these SHAPE data, using the RNAstructure modeling algorithm with default SHAPE pseudo-energy parameters, which were derived from prior 16S rRNA analysis (Deigan et al. 2009). The resulting model
the SHAPE experimental method and modeling procedure (Cordero et al. 2013). Overall, these data confirmed the reproducibility of RNAstructure modeling software versions (Fig. 1D; Supplementary quantitation software (Yoon et al. 2011; Karabiber et al. 2012a).)

We obtained results that were the same, within the robustness of our experimental SHAPE protocol and the correctness of the analysis (Ramachandran et al. 2013), the strengths of these supports did not simply reflect helix length. For example, the short P3 (three base pairs) attained a high support of 83%, whereas the second-longest helix alt-P1d attained a lower support of 43%. Instead, the helices with low supports lay in regions that could form numerous alternative structures calculated to have energies nearly as low as the final 1D-data-guided model, including the crystallographic secondary structure (Fig. 2A). On the other hand, helices that were rearranged in the 1D-data-guided model gave lower bootstrap supports (43%, 78%, 38%, and 13% for alt-P1d, P2*, alt-P4, and alt-P5, respectively). In contrast to prior suggestions (Ramachandran et al. 2013), the strengths of these supports did not simply reflect helix length. For example, the short P3 (three base pairs) attained a high support of 83%, whereas the second-longest helix alt-P1d attained a lower support of 43%. Instead, the helices with low supports lay in regions that could form numerous alternative structures calculated to have energies nearly as low as the final 1D-data-guided model, including the crystallographic secondary structure (Fig. 2B, blue arrows).

Nonparametric bootstrapping gives low confidence estimates

Procedures for estimating uncertainties in SHAPE-directed modeling have not yet become widely accepted. We recently proposed that useful helix-by-helix support values might be calculated by a resampling procedure called nonparametric bootstrapping (Efron et al. 1996; Kladwang et al. 2011c). This conceptually simple procedure has found wide use in complex statistical problems, such as phylogenetic inference, in which parametric models for likelihood or posterior probabilities are unavailable or untrustworthy. Mock data sets are “bootstrapped” from the experimental data set by resampling with replacement from the collection of data-derived energy bonuses. These data sets mimic scenarios in which data at particular residues might be missing or extra data at particular residues are available (e.g., from multiple-probe methods). The data are then input into the same secondary structure prediction algorithm. The frequencies at which helices arise in these bootstrap replicates provide “bootstrap supports.” Low support values indicate that alternative structural models exist and are nearly as consistent with the experimental data as the original model. Although this procedure has been criticized as being overly conservative (Ramachandran et al. 2013), experimental benchmarks on noncoding RNAs of known structure (Kladwang and Das 2010; Kladwang et al. 2011a,b) and simulation-based studies (S Tian and R Das, unpubl.) confirm that bootstrap supports provide numerically accurate indicators of helix confidence.

For the 126–235 RNA SHAPE modeling, bootstrapping calculations gave a wide range of values for the 1D-data-guided model (Fig. 2). On one hand, helices that were shared between the 1D-data-guided model and the crystallographic model gave bootstrap supports >80% (see P1a, P1c, P2b, and P3 in Fig. 2A). On the other hand, helices that were rearranged in the 1D-data-guided model gave lower bootstrap supports (43%, 78%, 38%, and 13% for alt-P1d, P2*, alt-P4, and alt-P5, respectively). In contrast to prior suggestions (Ramachandran et al. 2013), the strengths of these supports did not simply reflect helix length. For example, the short P3 (three base pairs) attained a high support of 83%, whereas the second-longest helix alt-P1d attained a lower support of 43%. Instead, the helices with low supports lay in regions that could form numerous alternative structures calculated to have energies nearly as low as the final 1D-data-guided model, including the crystallographic secondary structure (Fig. 2B, blue arrows).

The low bootstrap supports for the rearranged helices, as well as the difficulty of discriminating between multiple alternative structures with the available data, motivated us to acquire measurements on the 126–235 RNA that might validate or falsify their existence. Independent information from alternative modifiers highlighting bases whose Watson-Crick edges are available for alkylation (A-N1 or C-N3, by DMS) (Tijerina et al. 2007) or for the carbodiimide reaction (G-N1 or U-N3, by CMCT) (Walczak et al. 1996; Cordero et al. 1998; Leonard et al. 2013), the strengths of these supports did not simply reflect helix length. For example, the short P3 (three base pairs) attained a high support of 83%, whereas the second-longest helix alt-P1d attained a lower support of 43%. Instead, the helices with low supports lay in regions that could form numerous alternative structures calculated to have energies nearly as low as the final 1D-data-guided model, including the crystallographic secondary structure (Fig. 2B, blue arrows).

The low bootstrap supports for the rearranged helices, as well as the difficulty of discriminating between multiple alternative structures with the available data, motivated us to acquire measurements on the 126–235 RNA that might validate or falsify their existence. Independent information from alternative modifiers highlighting bases whose Watson-Crick edges are available for alkylation (A-N1 or C-N3, by DMS) (Tijerina et al. 2007) or for the carbodiimide reaction (G-N1 or U-N3, by CMCT) (Walczak et al. 1996; Cordero et al. 1998; Leonard et al. 2013), the strengths of these supports did not simply reflect helix length. For example, the short P3 (three base pairs) attained a high support of 83%, whereas the second-longest helix alt-P1d attained a lower support of 43%. Instead, the helices with low supports lay in regions that could form numerous alternative structures calculated to have energies nearly as low as the final 1D-data-guided model, including the crystallographic secondary structure (Fig. 2B, blue arrows).

The low bootstrap supports for the rearranged helices, as well as the difficulty of discriminating between multiple alternative structures with the available data, motivated us to acquire measurements on the 126–235 RNA that might validate or falsify their existence. Independent information from alternative modifiers highlighting bases whose Watson-Crick edges are available for alkylation (A-N1 or C-N3, by DMS) (Tijerina et al. 2007) or for the carbodiimide reaction (G-N1 or U-N3, by CMCT) (Walczak et al. 1996; Cordero et al. 1998; Leonard et al. 2013), the strengths of these supports did not simply reflect helix length. For example, the short P3 (three base pairs) attained a high support of 83%, whereas the second-longest helix alt-P1d attained a lower support of 43%. Instead, the helices with low supports lay in regions that could form numerous alternative structures calculated to have energies nearly as low as the final 1D-data-guided model, including the crystallographic secondary structure (Fig. 2B, blue arrows).

The low bootstrap supports for the rearranged helices, as well as the difficulty of discriminating between multiple alternative structures with the available data, motivated us to acquire measurements on the 126–235 RNA that might validate or falsify their existence. Independent information from alternative modifiers highlighting bases whose Watson-Crick edges are available for alkylation (A-N1 or C-N3, by DMS) (Tijerina et al. 2007) or for the carbodiimide reaction (G-N1 or U-N3, by CMCT) (Walczak et al. 1996; Cordero et al. 1998; Leonard et al. 2013), the strengths of these supports did not simply reflect helix length. For example, the short P3 (three base pairs) attained a high support of 83%, whereas the second-longest helix alt-P1d attained a lower support of 43%. Instead, the helices with low supports lay in regions that could form numerous alternative structures calculated to have energies nearly as low as the final 1D-data-guided model, including the crystallographic secondary structure (Fig. 2B, blue arrows).
et al. 2012a) were obtained, and secondary structure prediction guided by these data show similarly low bootstrap values (Supplemental Fig. S1D,E). Except for noncanonical pairings in P2*, this analysis provided weak or no support to the 1D-data-guided model above and indicated that confidently disambiguating the RNA’s solution structure would require methods with higher information content.

**Two-dimensional mutate-and-map (M²) resolves ambiguities in modeling**

Compared to the “one-dimensional” (1D) structure mapping approaches above, a recently developed two-dimensional (2D) expansion of chemical mapping offered the prospect of higher confidence modeling of the 126–235 RNA. The mutate-and-map method (M²) involves the parallel synthesis of separate constructs harboring single mutations at each nucleotide of the RNA (Supplemental Table S2), followed by chemical mapping of each construct at nucleotide-resolution (Kladwang and Das 2010; Kladwang et al. 2011a,b). Observation of an initially protected region that becomes reactive upon mutation of a sequence-distant region provides evidence for pairing between the two regions. Mutations that are unique in their effect and that do not release nucleotides other than their partner appear as punctate features in the M² data; such signals provide the strongest evidence for nucleotide–nucleotide interactions (Kladwang et al. 2011b).

M² electropherograms for the 126–235 RNA were acquired as in prior work (Fig. 3A; Kladwang et al. 2011b). Several features provided consistency checks in the M² analysis. Perturbations near the site of the mutations were visible as a diagonal feature from top left to bottom right. In addition, in several expected regions, mutations led to punctate features corresponding to “release” of sequence-distant nucleotides. For example, G138 was exposed by C225G and no other mutation. This data, as well as nearby features (marked I in Fig. 3A) supported P1c. Similarly, punctate features of increased reactivity at G168 by C153G and G184 by C193G defined hairpin P2b and P3 (marked II and III, respectively, on Fig. 3A). These segments, P1c, P2b, and P3, were present in all the models above, including the prediction by RNAstructure (Mathews et al. 2004). This analysis takes into account the single-nucleotide-resolution features as above but also leverages additional, less punctate features that correspond to, e.g., disruption of multiple base pairs upon mutations (marked VIII in Fig. 3A). The weights of these features are calculated as Z-scores (Fig. 3B), which down-weight any regions that are highly variable across constructs (see Methods). Consistent with visual analysis above, the resulting secondary structure (referred to hereafter as the “2D-data-guided model”), recovered helices P1a–c, P2b, P3, and the noncanonical P2* with high confidence (support values of 96% or greater) (Fig. 3C,D). Furthermore, this 2D-data-guided model agreed with the crystallographic secondary structure in helix P4b (support value 99%), which is entirely absent from the 1D-data-guided structure. Minor discrepancies with the crystallographic secondary structure occurred in P2a and in P4a. Some bootstrap replicates (36%) recovered two base pairs in helix P2a, but a larger fraction of replicates (63%) returned this helix with the pairings shifted by one nucleotide, a secondary structure we called shift-P2a (cf. Fig. 1B,C). Nevertheless, the overall support totaled over P2a and shift-P2a was strong (98%). Similarly, bootstrap replicates sampled alternative registers for the P4a helix (shift-P4a and the crystallographic P4a at 85% and 13%, respectively, summing to 98%). In contrast, only 2% of bootstrap replicates gave secondary structures consistent with the alt-P4 rearrangement in the 1D-data-guided model (Figs. 2B, 3D).

As an independent confirmation of the M² analysis, we repeated the M² experiments with the SHAPE modifier 1M7 (instead of NMIA) and again with the DMS modifier. Automated secondary structure prediction guided by these separate M² data sets returned models indistinguishable from the NMIA-based analysis above, with similar residual ambiguities in the registers of P2a and P4b (Supplemental Fig. S3). Overall, the M² analysis recovered the crystallographic secondary structure, up to potential single-nucleotide register shifts in helices P2a and P4b.

**Systematic falsification of 1D-data-guided model through mutation/escape**

To more deeply interrogate and test the 126–235 RNA’s secondary structure, we sought to validate or falsify base pairs in the models above through mutation/escape experiments.
A classic technique for validating RNA models is the rescue of disruptive perturbations from single mutations through compensatory mutations. Rescue of function in double mutants predicted to restore Watson-Crick pairs disrupted by separate single mutations provides strong evidence for base-pairing of those nucleotides. Mutation/rescue approaches are well developed for RNA molecules whose structure is coupled to a functional readout, such as self-cleavage (Wu and Huang 1992; Macnaughton et al. 1993); but such a general and systematic validation method has not been well developed for RNA molecules whose structure is coupled to a functional readout, such as self-cleavage (Wu and Huang 1992; Macnaughton et al. 1993); but such a general and systematic validation method has not been well
explored for cases without well-established functional assays. Here, chemical mapping offers a single-nucleotide-resolution readout of perturbation and restoration of structure (Kladwang et al. 2011b). The crystallographic model, 1D-data-guided and 2D-data-guided models provided specific pairing hypotheses to test. The same synthesis pipeline (Kladwang et al. 2011b) leveraged to synthesize 96 single mutants for M² measurements permitted facile synthesis of these additional RNA variants (Supplemental Table S2). Figure 4 shows the capillary electropherograms of pairings in question as well as quantitated data.

Discriminating the secondary structure models required identification of nucleotides with different pairings across models. As an example, G201 paired with C207 within the alt-P4 helix of the 1D-data-guided model but with C217 in the shift-P4a helix of the 2D-data-guided model. Single mutation of G201C resulted in clear changes in SHAPE data over a ~20-nucleotide region (180–198) (Fig. 4F). The 1D-data-guided model predicted that mutation C207G would restore this pairing. However, the double mutant G201C/C207G retained the disruptions observed in the G201C single mutant as well as additional changes observed in the C207G single mutant (Fig. 4F). An analogous experiment on the same pairing but different mutations (G201U/C207A) gave similar results, with no observed rescue (Fig. 4E).

In contrast to the above experiment, the 2D-data-guided model predicted that a different mutation, C217G, would restore the pairing disrupted by the G201C mutation. Such a rescue was indeed observed in G201C/C217G (Fig. 4S). The effect was even more striking given that the rescuing C217G mutation produced disruptions throughout the entire RNA when implemented as a single mutant (Fig. 4S). The SHAPE reactivities for the entire double mutant RNA were indistinguishable from the wild-type RNA with a minor exception: Weak reactivity at G204 in the wild type was suppressed in the double mutant. Together with similar observations with G201U/C217A (Fig. 4R), these
experiments provided strong evidence for shift-P4a predicted by the 2D-data-guided model.

In addition to the pairing options G201 above, we tested additional base-pairings by mutation/rescue studies on alt-P1d, alt-P4, and alt-P5 (Fig. 4A–O). None of these experiments exhibited disruptions by single mutants that were rescued by double mutants (Supplemental Fig. S4A, E). Three double mutants (A199U/U209A, A143U/U219A, and A143C/U219G) gave reactivity profiles similar to wild type, but the single mutations alone did not introduce significant disruptions either (Fig. 4A, K, L).

For the 2D-data-guided model, mutation/rescue resolved ambiguities in the P2a/shift-P2a and P4a/shift-P4a regions. Mutation/rescue for helix P4b was rendered difficult by its low stability (see Supplemental Fig. S4D). For the pairings of shift-P2a, double mutants did not exhibit rescue of disruptions observed in single mutants (Fig. 4W–Z). Therefore, another set of double mutants was designed based on the register of helix P2a seen in the crystallographic model (Fig. 4AG–AI). In each of these cases, the double mutants did rescue the perturbations induced by single mutants, restoring the SHAPE reactivity seen in the wild-type RNA. (In G145U/G177A, the SHAPE reactivity of G177 was suppressed, consistent with replacement of the original noncanonical G/G pair with a canonical A/U pair.) Taken together, these data strongly favored the formation of the crystallographic P2a.

In the case of shift-P4a, as with G201C/C217G above, all double mutants rescued disruptions observed in single mutants, giving strong evidence for helix shift-P4a (Fig. 4P–V). Furthermore, the quantitated data showed strong agreement with one another (Supplemental Fig. S4B,F).

Quantitative dissection of the P4a helix register shift

The data above strongly favored the crystallographic secondary structure. The only exception was a subtle one: a single-nucleotide register shift in helix P4a, which was experimentally validated through mutation/rescue of the shifted pairings. We sought to further probe this register shift by carrying out mutation/rescue experiments based on crystallographic P4a but not in shift-P4a (Fig. 4AA–AF). Surprisingly, the double mutants SHAPE data showed rescue of disruptions induced by single mutations, and led to SHAPE data indistinguishable from the wild type data, except for stronger reactivity at G204 and the disappearance of weak reactivity at G198 (Supplemental Fig. S4C). These data strongly supported the presence of the crystallographic helix P4a, in apparent contradiction to the results above supporting the 2D-data-guided helix shift-P4a.

The paradox of mutation/rescue data supporting both registers could be resolved if 126–235 RNA interconverts between P4a and shift-P4a registers at equilibrium. The double mutants chosen to test shift-P4a perturbed the equilibrium and favored the shift-P4a structure, and alternative stabilization occurred for the P4a double mutants. This model provided testable predictions: Double mutants should converge to two slightly distinct SHAPE reactivity profiles depending on which register they preferentially stabilized. Furthermore, the wild-type SHAPE profile should be decomposable into a linear combination of the P4a-stabilized and shift-P4a-stabilized profiles.

Indeed, the double mutants observed to stabilize P4a gave distinct SHAPE reactivity data from those that stabilized shift-P4a at nucleotides G204 (increased) and G198 (decreased); see Figure 4AA–AF (blue arrows) and Supplemental Figure S4B,C,F,G. In accord with the model, these are the two nucleotides brought into and out of the helix, respectively, by the register shift; cf. Figure 4S and AD. Finally, the wild-type data could be fitted to an equilibrium mixture of the shift-P4a-stabilized and P4a-stabilized profiles, as determined from the double mutants (Fig. 5A). To estimate the population fraction of shift-P4a and P4a in the wild-type RNA, we carried out $\chi^2$-based fits of the SHAPE reactivity across all nucleotides (126–235), or G204 and A205 (where changes were largest), or G204 only (Supplemental Fig. S4H–I). The equilibrium fractions of shift-P4a and P4a were fitted to be 80 ± 10% and 20 ± 10%, with the error reflecting small differences in fits using the different nucleotides. This ratio of populations corresponds to a free energy difference of 0.3–0.5 kcal/mol. In the context of the full ribosome, this equilibrium may be perturbed to favor the crystallographic/phylogenetic structure by RNA tertiary contacts or S20 protein binding, providing possible checkpoints for 16S rRNA assembly that can be tested in future experiments (Supplemental Fig. S6; below). Additional tests, including SHAPE reactivities from other mutants, suggest that if states other than the P4a and shifted-P4a state are present, either their population fractions are negligible (<5%) (Supplemental Fig. S4K) or their SHAPE reactivities are very similar to the already described two states.

DISCUSSION

Toward validated models of RNA structure from chemical mapping

As a test case for single-nucleotide-resolution chemical mapping, we have carried out an in-depth dissection of the
solution structure of a model system with complex RNA structure, a 110-nucleotide segment of the 16S ribosomal RNA that was previously suggested to undergo a large-scale conformational rearrangement between its intrinsic solution structure and its state within the assembled small ribosomal subunit. Through a simple computational procedure, we first discovered that prior SHAPE-directed modeling (Deigan et al. 2009) as well as DMS and CMCT chemical mapping measurements did not carry sufficient information content to determine the 126–235 RNA’s secondary structure. We then leveraged a more information-rich technology, the mutate-and-map (M2) approach, to infer a more confident model for the 126–235 RNA. An enlarged view of G204 reactivities is on the right. Averaged reactivity profiles of P4a-stabilized (blue) and shift-P4a-stabilized (green) double mutants are plotted. SDs are shown, N = 5. (B) Model of 126–235 RNA secondary structure in solution, including P4a/shift-P4a dynamics.

FIGURE 5. Equilibrium fitting tests a register-shift “excited state” model for the 126–235 RNA. (A) Linear fitting of P4a/shift-P4a equilibrium based on G204. An enlarged view of G204 reactivities is on the right. Averaged reactivity profiles of P4a-stabilized (blue) and shift-P4a-stabilized (green) double mutants are plotted. SDs are shown, N = 5. (B) Model of 126–235 RNA secondary structure in solution, including P4a/shift-P4a dynamics.

Excited states

Nearly all experimental RNA structural methods, ranging from solution chemical mapping to crystallography, focus on inferring the dominant RNA conformation in solution. Obtaining an experimental description of weakly populated states necessarily requires more information and has typically involved mutations to trap such states (Pan and Woodson 1998; Russell et al. 2002; Gilbert et al. 2007). The mutate-map-rescue (M2R) approach herein involved a systematic panel of such mutations and provided a clear view of one such excited state. Several different sets of mutations (Fig. 4P–V,AA–AF) in the region 199–203 converged to a chemical
reactivity profile that was distinct from the wild type. The effects of single mutations and the reactivity perturbations could be explained by formation of the crystallographic P4a conformation, which is shifted in register by a single nucleotide from the dominant solution conformation shift-P4a. This register shift model was additionally confirmed by double-mutant analysis and by demonstrating the recovery of the wild-type reactivity from a linear combination of measurements on mutants stabilizing the conformations. Notably, our approach is analogous to—but significantly faster than—recent mutation-coupled NMR methods for inferring and stabilizing excited states of RNA model systems (Dethoff et al. 2012).

Here, the “excited” state of the 126–235 RNA, present at 20% population, has the register observed in crystallography of the entire 16S rRNA (Fig. 5). Changes in solution condition such as temperature, Mg2+, or monovalent concentration would be expected to shift the populations of the dominant and excited state and any other states, offering the possibility of environmental sensing by the ribosome through this register shift; initial characterization of Mg2+ versus Na+ titrations lend support to the hypothesis that the same states are present, but with different populations, across a wide range of solution conditions (data not shown). Within the full ribosome, this functional helix register may be stabilized by tertiary interactions with the rest of the 16S RNA or with ribosomal proteins (Supplemental Fig. S6B,C). Indeed, the 1D SHAPE profile of full-length 16S RNA from previous study resembles P4a-stabilized mutants at G204 (Fig. 1D). An A-minor interaction between G203-C214 and A465 is observed in crystallographic structure (Supplemental Fig. S6B), suggesting that this tertiary contact may stabilize the P4a conformation in full-length 16S rRNA context. The small ribosomal protein S20 can also bind the 126–235 region, and it plays a crucial role in stabilizing this region of the 5′ domain (Rydén-Aulin et al. 1993; Brodersen et al. 2002). These scenarios suggest a novel check-point for ribosome assembly based on locking the P4a/shift-P4a register shift and should be resolvable through future experiments. More generally, the mutate-map-rescue pipeline holds promise for discovering and validating excited states for other RNAs, especially if the throughput of double mutation can be increased and if inference of a structure ensemble from the available data can be fully automated.

A general mutate-map-rescue (M2R) pipeline

This study has delineated an expansion of conventional chemical mapping that enables systematic inference, testing, and refinement of RNA structure domains, including the correction of prior misleading inferences and possibility of inferring excited states. Although unusually detailed for a chemical mapping study, the entire mutate-map-rescue pipeline described herein was carried out with the same commercially available reagents, equipment, and synthesis strategy as our standard high-throughput chemical mapping protocol (Kladwang et al. 2011b; Lucks et al. 2011). With current technologies, the presented pipeline is generally applicable to noncoding RNA domains up to 300 nucleotides in length. Use of random mutagenesis or modification, and deconvolution through paired-end next-generation sequencing, may allow routine extension to longer transcripts and to RNAs in vivo (Cheng et al. 2014). We therefore propose that this mutate-map-rescue approach can be generally adopted as a “best practice” for finalizing RNA models inferred from chemical mapping.

MATERIALS AND METHODS

RNA synthesis

Double-stranded DNA templates were prepared by PCR assembly of DNA oligomers with maximum length of 60 nt ordered from IDT (Integrated DNA Technologies). DNA templates contain a 20-nt T7 RNA polymerase promoter sequence (TTCTAATACGACTCAG)10 TATA) on the 5′ end, followed by sequence of interest. One hairpin with single-stranded buffering region was added on both ends to flank the region of interest. A 20-nt Tail2 sequence (AAAGAAAC AACAAAACACAC) was put on the 3′ end (Supplemental Table S1). The assembly schemes for all constructs were designed by an automated MATLAB script (design_primers, available at https://github.com/DasLab/NA_thermo) (Kladwang et al. 2011b).

PCR reactions, consisting of 200 pmol of terminal primers and 2 pmol of internal primers, were carried out as previously described (Kladwang et al. 2011b). PCR products were purified using Ampure XP magnetic beads (Agencourt) on a 96-well Greiner microplate format following the manufacturer’s instructions. DNA concentration were measured by UV absorbance on Nanodrop 1000 spectrophotometer (Thermo Scientific). DNA templates were verified by sequencing (PAN core facility, Stanford University). In vitro transcription reactions were described previously (Kladwang et al. 2011b), followed by same purification and quantification steps as DNA. Sequences and purities of RNA samples were confirmed by reverse transcription in presence of each ddNTP.

Chemical modification

One-dimensional chemical mapping, mutate-and-map (M2), and mutation/rescue were carried out in 96-well format as described previously (Kladwang et al. 2011a,b; Cordero et al. 2014). Prior to chemical modification, 1.2 pmol of RNA was heated and cooled to remove secondary structure heterogeneity (90°C for 2 min and cooled on ice for 2 min) and folded for 20 min at 37°C in 15 μL of one of the following buffers as noted in the text: 10 mM MgCl2, 50 mM Na-HEPES, pH 8.0 (our standard) (Kladwang et al. 2011b); 5 mM MgCl2, 200 mM KOAc, 50 mM Na-HEPES, pH 8.0 (Deigan et al. 2009); 10 mM MgCl2, 100 mM NaCl, 50 mM Na-HEPES, pH 8.0 (Wilkinson et al. 2006); or 50 mM Na-HEPES, pH 8.0.

RNA was modified by adding 5 μL of modification reagent (0.5% dimethyl sulfate [DMS] prepared by mixing 1 μL 10.5 M DMS into 9 μL ethanol, and then 190 μL water; 21 mg/mL 1-cyclohexyl-[2-morpholinooethyl] carbodiimide metho-p-toluene sulfonate [CMCT]; 5 mg/mL 1-methyl-7-nitroisatoic anhydride [1M7]; or 12 mg/mL N-methylisatoic anhydride [NMA]) (Merino et al.
Doubley deionized H₂O or anhydrous DMSO was used as background control. Modification reactions were incubated at room temperature for 20 min and then quenched appropriately (5 μL of 0.5 M Na-MES, pH 6.0 for SHAPE and CMCT or 2-mercaptoethanol for DMS). All modifiers were made fresh before use. Quenches also included 1 μL poly(dT) magnetic beads (Ambion) and 0.065 pmols of FAM-labeled Tail2-A20 primer (AGGGTTGGTTGGTTCCCTT) for reverse transcription. Samples were separated and purified using magnetic stands, washed with 100 μL 70% ethanol twice, and air-dried. Beads were resuspended in 5.0 μL reverse transcription mix, then incubated for 30 min at 55°C. RNAs were degraded by adding 5 μL 0.4 M NaOH and incubating for 3 min at 90°C. Solutions were cooled down on ice then neutralized with 5 μL acid quench (1.4 M NaCl, 0.6 M HCl, and 1.3 M NaOAc). Fluorescent-labeled cDNA was recovered by magnetic bead separation, rinsed twice with 40 μL 70% ethanol, and air-dried. The beads were resuspended in 10 μL Hi-Di formamide (Applied Biosystems) with 0.0625 μL ROX-350 ladder (Applied Biosystems) and eluted for 20 min. Supernatants were loaded to capillary electrophoresis sequencer (ABI3100). Sequencing ladders were prepared analogously, without any chemical modification but with inclusion of each 2′-3′-dideoxy-NTP (ddNTP) equimolar to each dNTP during reverse transcription. To verify sequence assignments in downstream analysis, an additional “coloaded” sample composed of each sequencing ladder (5 μL), and cDNA derived from SHAPE-probed RNA (5 μL) was also measured. RNA chemical mapping using an alternative deep-sequencing readout was carried out analogously to the method above but included an additional ligation step to permit Illumina sequencing (Seetin et al. 2014).

**Data processing and structural modeling**

The HiTRACE software package version 2.0 was used to analyze CE data (MATLAB toolbox is available at https://github.com/hitrace) (Yoon et al. 2011), and a web server is also available at http://hitrace.org (Kim et al. 2013). Electrophoretic traces were aligned and baseline subtracted using linear and nonlinear alignment routines as previously described (Kim et al. 2009). Sequence assignment was accomplished manually with verification from sequencing ladders and the coloaded samples. Band intensities were obtained by fitting profiles to Gaussian peaks and integrating.

Rigorous normalization, correction for signal attenuation, and background subtraction were enabled by inclusion of referencing hairpin loop residues (GAGUA) at both 5′ and 3′ ends, 10x dilution replicates, and no-modification controls (get_reactivities in HiTRACE). Briefly, values for saturated peaks were obtained from 10x dilutions. Signal attenuation was corrected from 5′ to 3′ ends based on the relative reactivity between 5′ and 3′ referencing hairpin loop intensities. Reactivities of all the chemical profile were normalized against GAGUA (Kladwang et al. 2014). For most comparisons, the average of reactivities in each of the two GAGUA hairpins were set to two for best comparison with previously reported data (Deigan et al. 2009).

Data-driven secondary structure models were obtained using the Fold program of the RNAstructure package (Mathews et al. 2004; Reuter and Mathews 2010). For secondary structure models guided by 1D chemical mapping, pseudoenergy slope and intercept parameters of 2.6 kcal/mol and −0.8 kcal/mol (RNAstructure version 5.4) (Mathews et al. 2004) or, where stated, 1.8 kcal/mol and −0.6 kcal/mol (version 5.5) (Hajdin et al. 2013) were used. To obtain 2D-data-guided secondary structure models, Z-score matrices for M² data sets were calculated as previously described (Kladwang et al. 2011b). Briefly, Z-scores were calculated for each nucleotide reactivity by subtracting the average reactivity of this nucleotide across all mutants and dividing by standard deviation (output_Zscore_from_rdat in HiTRACE). M² seeks to identify release of putative base-pair partners of a nucleotide upon mutation; therefore, negative Z-scores and positions with high average reactivity (cutoff is 0.8) were excluded. Background subtraction and signal attenuation correction were not applied to M² data, since Z-scores are independent of those steps (which would otherwise introduce noise). Z-score matrices were used as base-pair-wise pseudoenergies with a slope and intercept of 1.0 kcal/mol and 0 kcal/mol (Kladwang et al. 2011b). Helix-wise confidence values were calculated via bootstrapping as described previously: Mock data sets were generated by sampling the mutants with replacement and comparing the helices of the resulting mock-data-driven models with those in the model obtained using the full data (Kladwang et al. 2011b). An independent analysis using the QuSHAPE software was performed following given instructions (Karabiber et al. 2013). In all modeling steps, full-length sequences (including flanking elements) were used for prediction.

**Structural equilibrium fitting**

Equilibrium fractions of each structure were determined by assuming that shift-P4a-stabilized and P4a-stabilized double mutants completely stabilize the register-shifted and crystallographic structure, respectively; their reactivity profiles therefore represent the reactivity profile for each state. The reactivities of shift-P4a-stabilized double mutants and P4a-stabilized double mutants were taken to fit the wild-type reactivity by $\chi^2$ score, which is calculated as follows:

$$\chi^2 = \sum \frac{(d_{\text{WT}} - d_{\text{PRED}})^2}{\sigma_{\text{WT}}^2 + \sigma_{\text{PRED}}^2}$$

$$d_{\text{PRED}} = \alpha \cdot d_{\text{P4a}} + (1 - \alpha) \cdot d_{\text{shift-P4a}}$$

$$\sigma_{\text{PRED}}^2 = \alpha^2 \cdot \sigma_{\text{P4a}}^2 + (1 - \alpha)^2 \cdot \sigma_{\text{shift-P4a}}^2$$

Where $d_{\text{WT}}$, $d_{\text{P4a}}$, and $d_{\text{shift-P4a}}$ are mean SHAPE reactivity profiles of wild-type, P4a-stabilized, and shift-P4a-stabilized double mutants, and $\sigma_{\text{WT}}$, $\sigma_{\text{P4a}}$, and $\sigma_{\text{shift-P4a}}$ are errors (standard deviation across A199U/U218A, A199C/U218G, G201U/U216A, G201C/U216G, G203U/C214A, G203C/C214G for P4a; and A199U/U219A, A199C/U219G, G201U/C217A, G201C/C217G, G203U/C215A, G203C/C215G for shift-P4a) of $d_{\text{WT}}$, $d_{\text{P4a}}$, and $d_{\text{shift-P4a}}$. The parameter $\alpha$ is the fraction of P4a register shift, ranging from 0 to 1. $\chi^2$ is summed over all or a subset of nucleotide positions $i$ as specified in the text. $\chi^2$ scores are plotted against $\alpha$, and the $\alpha$ value with minimum $\chi^2$ is taken as the best fit. Fitting with a third state was carried out analogously, optimizing over the weight $\beta$ on a third state; SHAPE reactivities from mutants C175G, C176G, A199C, C207G, and C215G were tested for the third state.

**Structural visualization**

Secondary structure images were generated by VARNA (Darty et al. 2005; Mortimer and Weeks 2007; Tijerina et al. 2007; Cordero et al. 2012a). The atomic model of crystallographic data from PDB entry
Mutate-map-rescue corrects SHAPE-guided RNA model

31M (Zhang et al. 2009) was visualized in PyMol (The PyMOL Molecular Graphics System, Version 1.5.0.4; Schrödinger, LLC.). Noncanonical base-pairing and long-range interactions were mapped with tool at the RNA 3D Hub (Petrov et al. 2013). Secondary structure diagram in Leontis/Westhoff nomenclature (Leontis and Westhof 2001) was drawn in Illustrator (Adobe).

DATA DEPOSITION
All chemical mapping data sets, including one-dimensional mapping, mutate-and-map, and mutation/rescue, have been deposited at the RNA Mapping Database (http://rmdb.stanford.edu) (Cordero et al. 2012b) under the following accession codes: 16S_STD_0001, 16S_NMIA_0001, 16S_IM7_0001, 16S_DMS_0001, and 16S_RSQ_0001.

SUPPLEMENTAL MATERIAL
Supplemental material is available for this article.

ACKNOWLEDGMENTS
We acknowledge financial support from a Stanford Graduate Fellowship (S.T.), a Conacyt fellowship (P.C.), the Burroughs Wellcome Foundation (CASI to R.D.), and NIH RO1 R01GM102519. We thank members of the Das laboratory for comments on the manuscript.

Received January 13, 2014; accepted June 15, 2014.

REFERENCES


Supporting information for
“High-throughput mutate-map-rescue evaluates SHAPE-directed RNA structure and uncovers excited states”

Siqi Tian,1 Pablo Cordero,2 Wipapat Kladwang,3 and Rhiju Das1,2,3,*

1Department of Biochemistry, 2Biomedical Informatics Program, and 3Department of Physics, Stanford University, Stanford, California 94305, United States

*Corresponding Author

This document contains the following sections:

Supporting Figures S1-S6
Supporting Tables S1-S2
Supporting References
SUPPORTING FIGURES

Figure S1. Confirming sequence assignments and comparisons of SHAPE/DMS/CMCT modeling. (A) Electropherogram of SHAPE and ddNTP ladder co-loading. Sequence assignment is marked on top. Since reverse transcriptase is blocked prior to acylation and ddNTP termination occurs after the incorporation, there is a single register shift for the same nucleotide across profiles. (B) Normalized SHAPE reactivity derived from capillary electrophoresis (CE) and next-generation sequencing (MAP-Seq). Standard deviations (SD) are shown, N = 7 for CE, N = 2 MAP-Seq. (C) Secondary structure prediction of 126-235 RNA without data guidance. (D-E) Secondary structure prediction and bootstrap support matrix using 1-dimensional DMS/CMCT data. For A and C residues, DMS reactivity is taken; and for G and U residues, CMCT reactivity is taken. The model contains a different version of helices compared to 1D-data-guided model, labeled as alt-P4´ and alt-P5´ respectively. Difference from crystallographic model is drawn in yellow/gray lines. Percentage labels give bootstrap support values.
Figure S2 (next page). Consistent but uncertain SHAPE-directed models from different data processing software, secondary structure software and experimental conditions. (A) Similar normalized SHAPE reactivities from different data processing packages. Flanking sequences are grayed and GAGUA pentaloop (Kladwang et al. 2014) highlighted in yellow. (B-K) Secondary structure predictions and bootstrap support matrices using variants of the SHAPE analysis match default analysis (Figure 2A-B). (B-C) Data processed by QuSHAPE give results consistent with default HiTRACE analysis. (D-E) Data normalized so that mean reactivity of flanking GAGUA pentaloops are set to 1.0 instead of 2.0 (default procedure to match prior work; see Methods), gives results consistent with default analysis. (F-G) Data normalized as described in (Deigan et al. 2009) do not give alt-P5a, which is present but with low bootstrap support value (13%) in default analysis. (H-I) Use of RNAstructure version 5.5 (Hajdin et al. 2013) instead of version 5.4 (Deigan et al. 2009) gives a similar secondary structure model but with high confidence in helices alt-P1d and alt-P4, which are falsified by independent data (see main text). (J-K) Inclusion of 200 mM KOAc (with 10 mM MgCl2 and 50 mM Na-HEPES, pH 8.0) during SHAPE probing, to match (Deigan et al. 2009), gives results consistent with default solution conditions without added KOAc; see (M). Difference from crystallographic model is drawn in yellow/gray lines. Percentage labels give bootstrap support values. (L) Data using dGTP vs. dITP in reverse transcription can give different systematic effects at G vs. C positions (Kladwang et al. 2011), but give data in close agreement for the 126-235 RNA. Asterisks mark bands with variable background. (M) Normalized SHAPE reactivity of wild-type 126-235 RNA under different folding conditions.
Figure S3. Mutate-and-map (M²) data of 126-235 RNA using 1M7 and DMS modifiers give consistent results as NMIA M² analysis. (A) Entire mutate-and-map dataset (1M7) across 110 single mutations. (B) Z-score contact-map extracted from (A), used for secondary structure inference. (C-D) Secondary structure prediction and bootstrap support matrix using 2-dimensional M² data (1M7). (E-F) Entire mutate-and-map dataset (DMS) and Z-score contact-map. (G-H) Secondary structure prediction and bootstrap support matrix using 2-dimensional M² data (DMS). The model contains a different version of helix compared to 2D-data-guided model, labeled as shift-P2α'. Difference from crystallographic model is drawn in yellow/gray lines. Percentage labels give bootstrap support values.
Figure S4. Overlays of compensatory mutants test modeled secondary structures and allow tests of proposed P4a helix register shift. (A, E) Compensatory mutants designed from the 1D-data-guided model (alt-P4, alt-P5 and alt-P1d in magenta, brown, and red, respectively) give SHAPE profiles in poor agreement with each other, giving no support for these pairings. (B, F) Compensatory mutants designed from the 2D-data-guided model give SHAPE profiles in agreement with the wild type (WT) for shift-P4a (light green) but not shift-P2a (dark green). (C, G) Compensatory mutants give strong support for both P4a (blue) and P2a (cyan). Blue arrows mark G204 exposure in P4a-stabilized mutants. (D) Quartet of electropherograms testing mutation/rescue for P4b (C206G/G213C) does not give support for this pairing. (H-J) $\chi^2$ score curve of fraction of P4a helix, fitting based on (H) whole RNA (126-235), (I) G204 and A205, and (J) G204 only, using component profiles from P4a-stabilized and shift-P4a-stabilized mutants in (F) and (G). Parameters at which optimal $\chi^2$ scores are obtained are circled; errors are estimated by values at which $\chi^2$ increases by one. (K) $\chi^2$ tests of any additional third state, based on fitting reactivities across whole RNA (126-235) and use of SHAPE profiles for mutants that differed substantially from wild type. All cases show substantial worsening of $\chi^2$ (by more than one) at >2% population fractions, ruling out significant population fractions of such states.
Figure S5. Repeated SHAPE analysis of M1 and M2 mutants from previous study (Deigan et al. 2009). (A) Normalized SHAPE reactivity of WT, M1 and M2 resolved by capillary electrophoresis (CE). Flanking sequences are grayed and GAGUA pentaloop highlighted in yellow. Standard deviations (SD) are shown, N = 3. (B) Normalized SHAPE reactivity resolved by next-generation sequencing (MAP-Seq). Error bars are estimated from MAP-Seq raw counts. (C-D) Quartet electropherogram (wild type, single mutants, and double mutant) of M1 (U208C/U218C) and M2 (G145C/U219C). For M1, no significant change is observed, consistent with these mutants preserving the wild type structure. Several models, including a prior model (Deigan et al. 2009) and the distinct secondary structure presented in main text Figure 5 are consistent with these data. For M2, blue arrow marks exposure of G204, consistent with stabilizing crystallographic P4a (the excited state register shift) in this mutant. See main text Discussion for further description.
Figure S6. Non-canonical base-pairs and tertiary contacts intrinsic to the 126-235 RNA or formed during full ribosome assembly. (A) Base-pairing of G145 and G177 in crystallographic structure may be an intrinsic feature of the 126-235 RNA; the replacement G145C reduces G177 SHAPE reactivity (SI Figure S5D). G177 is in syn-conformation and its Hoogsteen edge is forming base-pair with the Watson/Crick edge of G145. (B) RNA-RNA tertiary interaction of P4a that can form in the context of the full 16S rRNA. A-minor contact between A465 and G203/C214 is shown. (C) Secondary structure diagram of 126-235 RNA highlighting contacts with sequences elsewhere in the full-length 16S rRNA using Leontis/Westhof nomenclature (Leontis and Westhof 2001).
**SUPPORTING TABLES**

**Table S1. Schematic of PCR primer assembly of wild-type sequence.** A total of four primers are used to assemble the double stranded DNA template listed in SI Table S2. The annealing sites are shown by vertical lines, annealing temperature for each site is shown on the left. Nucleotide numbering from the full-length 16S rRNA is shown above or under each primer. The forward strand of assembled DNA template is highlighted: T7 promoter (blue), flanking sequences (buffering region / stem of referencing hairpin / pentaloop of referencing hairpin) (yellow/green/red), sequence of interest (black), and tail2 (magenta).

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>primer-1-F</td>
<td>TTCTAATACGACTCACTATAGGGAACGACTCGAGTAGACTCCAAAAAGGAACTGCTGATGGAGGGGGATAACTACTGGAAACGGTAGCTAATACCGCATAACGTCGCAAGACCAAAGG</td>
</tr>
<tr>
<td>primer-2-R</td>
<td>TGCGACGTTATGCGGTATTAGCTACCGTTTCCAGTAGTTATCCCCCTCCATCAGGTTGTTTCTTTTGGTTGTTGTTGTTGTTG</td>
</tr>
<tr>
<td>primer-3-F</td>
<td>CCGCATAACGTCGCAAGACCAAAGAGGGGGACCTTCGGGCCTCTTGCCATCGGATGTGCC</td>
</tr>
<tr>
<td>primer-4-R</td>
<td>GTTGGTTGTTGTTTCTTTTGGTAGCTACCGACTCTTGGGCACATCCGGATGTC</td>
</tr>
</tbody>
</table>

**Sequence**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Primers</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTCTAATACGACTCACTATAGGGAACGACTCGAGTAGACTCCAAAAAGGAACTGCTGATGGAGGGGGATAACTACTGGAAACGGTAGCTAATACCGCATAACGTCGCAAGACCAAAGG</td>
<td>primer-1-F</td>
</tr>
<tr>
<td>TGCGACGTTATGCGGTATTAGCTACCGTTTCCAGTAGTTATCCCCCTCCATCAGGTTGTTTCTTTTGGTTGTTGTTGTTGTTG</td>
<td>primer-2-R</td>
</tr>
<tr>
<td>CCGCATAACGTCGCAAGACCAAAGAGGGGGACCTTCGGGCCTCTTGCCATCGGATGTGCC</td>
<td>primer-3-F</td>
</tr>
<tr>
<td>GTTGGTTGTTGTTTCTTTTGGTAGCTACCGACTCTTGGGCACATCCGGATGTC</td>
<td>primer-4-R</td>
</tr>
</tbody>
</table>
Table S2. List of 16S sequences. Each sequence composes similar elements (SI Table S1). It starts with a promoter sequence for *in vitro* transcription using T7 RNA polymerase. A 5’ flanking sequence contains three Gs to facilitate efficient transcription elongation, a small hairpin (6 base-pairs in the stem and GAGUA pentaloop) for reactivity referencing, as well as four As in buffering region. Followed is the sequence of interest (126-235). A 3’ flanking sequence similar to 5’ flanking sequence appends, containing three As in buffering region, a small hairpin for reactivity referencing, and AAC buffering the tail2. Tail2 sequence is at the 3’ end of each DNA and RNA molecule, which is used to bind the nucleic acid to beads and reverse transcription.

<table>
<thead>
<tr>
<th>Construct Name</th>
<th>DNA Sequence</th>
</tr>
</thead>
</table>
| 16S-WT         | TTCTAATACGACTGACTCATATTGGAAGACGACTCGAGTAGACTGCAGAAGGGGAACTCTGCTGGAGGGGTAAGTACTATTACGGCATACACGTC<br>GAAGACCAAGGGGACCTTCCGGGCTTCGCACTCGATGCTGCCCAAGAGGAGCTGAGTATGCTCCACAAAGAAAGAACCAACACAC<br>16S-A199T      | TTCTAATACGACTGACTCATATTGGAAGACGACTCGAGTAGACTGCAGAAGGGGAACTCTGCTGGAGGGGTAAGTACTATTACGGCATACACGTC<br>GAAGACCAAGGGGACCTTCCGGGCTTCGCACTCGATGCTGCCCAAGAGGAGCTGAGTATGCTCCACAAAGAAAGAACCAACACAC<br>16S-T219A      | TTCTAATACGACTGACTCATATTGGAAGACGACTCGAGTAGACTGCAGAAGGGGAACTCTGCTGGAGGGGTAAGTACTATTACGGCATACACGTC<br>GAAGACCAAGGGGACCTTCCGGGCTTCGCACTCGATGCTGCCCAAGAGGAGCTGAGTATGCTCCACAAAGAAAGAACCAACACAC<br>16S-T219A      | TTCTAATACGACTGACTCATATTGGAAGACGACTCGAGTAGACTGCAGAAGGGGAACTCTGCTGGAGGGGTAAGTACTATTACGGCATACACGTC<br>GAAGACCAAGGGGACCTTCCGGGCTTCGCACTCGATGCTGCCCAAGAGGAGCTGAGTATGCTCCACAAAGAAAGAACCAACACAC<br>16S-A199T;T219A| TTCTAATACGACTGACTCATATTGGAAGACGACTCGAGTAGACTGCAGAAGGGGAACTCTGCTGGAGGGGTAAGTACTATTACGGCATACACGTC<br>GAAGACCAAGGGGACCTTCCGGGCTTCGCACTCGATGCTGCCCAAGAGGAGCTGAGTATGCTCCACAAAGAAAGAACCAACACAC<br>16S-A199T;T218A| TTCTAATACGACTGACTCATATTGGAAGACGACTCGAGTAGACTGCAGAAGGGGAACTCTGCTGGAGGGGTAAGTACTATTACGGCATACACGTC<br>GAAGACCAAGGGGACCTTCCGGGCTTCGCACTCGATGCTGCCCAAGAGGAGCTGAGTATGCTCCACAAAGAAAGAACCAACACAC<br>16S-A199C      | TTCTAATACGACTGACTCATATTGGAAGACGACTCGAGTAGACTGCAGAAGGGGAACTCTGCTGGAGGGGTAAGTACTATTACGGCATACACGTC<br>GAAGACCAAGGGGACCTTCCGGGCTTCGCACTCGATGCTGCCCAAGAGGAGCTGAGTATGCTCCACAAAGAAAGAACCAACACAC<br>16S-T219G      | TTCTAATACGACTGACTCATATTGGAAGACGACTCGAGTAGACTGCAGAAGGGGAACTCTGCTGGAGGGGTAAGTACTATTACGGCATACACGTC<br>GAAGACCAAGGGGACCTTCCGGGCTTCGCACTCGATGCTGCCCAAGAGGAGCTGAGTATGCTCCACAAAGAAAGAACCAACACAC<br>16S-T218G      | TTCTAATACGACTGACTCATATTGGAAGACGACTCGAGTAGACTGCAGAAGGGGAACTCTGCTGGAGGGGTAAGTACTATTACGGCATACACGTC<br>GAAGACCAAGGGGACCTTCCGGGCTTCGCACTCGATGCTGCCCAAGAGGAGCTGAGTATGCTCCACAAAGAAAGAACCAACACAC<br>16S-A199C;T219G| TTCTAATACGACTGACTCATATTGGAAGACGACTCGAGTAGACTGCAGAAGGGGAACTCTGCTGGAGGGGTAAGTACTATTACGGCATACACGTC<br>GAAGACCAAGGGGACCTTCCGGGCTTCGCACTCGATGCTGCCCAAGAGGAGCTGAGTATGCTCCACAAAGAAAGAACCAACACAC<br>16S-A199C;T218G| TTCTAATACGACTGACTCATATTGGAAGACGACTCGAGTAGACTGCAGAAGGGGAACTCTGCTGGAGGGGTAAGTACTATTACGGCATACACGTC<br>GAAGACCAAGGGGACCTTCCGGGCTTCGCACTCGATGCTGCCCAAGAGGAGCTGAGTATGCTCCACAAAGAAAGAACCAACACAC<br>16S-G201T      | TTCTAATACGACTGACTCATATTGGAAGACGACTCGAGTAGACTGCAGAAGGGGAACTCTGCTGGAGGGGTAAGTACTATTACGGCATACACGTC<br>GAAGACCAAGGGGACCTTCCGGGCTTCGCACTCGATGCTGCCCAAGAGGAGCTGAGTATGCTCCACAAAGAAAGAACCAACACAC<br>16S-G201T;C217G| TTCTAATACGACTGACTCATATTGGAAGACGACTCGAGTAGACTGCAGAAGGGGAACTCTGCTGGAGGGGTAAGTACTATTACGGCATACACGTC<br>GAAGACCAAGGGGACCTTCCGGGCTTCGCACTCGATGCTGCCCAAGAGGAGCTGAGTATGCTCCACAAAGAAAGAACCAACACAC<br>16S-G201T;T216A| TTCTAATACGACTGACTCATATTGGAAGACGACTCGAGTAGACTGCAGAAGGGGAACTCTGCTGGAGGGGTAAGTACTATTACGGCATACACGTC<br>GAAGACCAAGGGGACCTTCCGGGCTTCGCACTCGATGCTGCCCAAGAGGAGCTGAGTATGCTCCACAAAGAAAGAACCAACACAC<br>16S-G201C      | TTCTAATACGACTGACTCATATTGGAAGACGACTCGAGTAGACTGCAGAAGGGGAACTCTGCTGGAGGGGTAAGTACTATTACGGCATACACGTC<br>GAAGACCAAGGGGACCTTCCGGGCTTCGCACTCGATGCTGCCCAAGAGGAGCTGAGTATGCTCCACAAAGAAAGAACCAACACAC<br>16S-G217G      | TTCTAATACGACTGACTCATATTGGAAGACGACTCGAGTAGACTGCAGAAGGGGAACTCTGCTGGAGGGGTAAGTACTATTACGGCATACACGTC<br>GAAGACCAAGGGGACCTTCCGGGCTTCGCACTCGATGCTGCCCAAGAGGAGCTGAGTATGCTCCACAAAGAAAGAACCAACACAC<br>16S-T216G      | TTCTAATACGACTGACTCATATTGGAAGACGACTCGAGTAGACTGCAGAAGGGGAACTCTGCTGGAGGGGTAAGTACTATTACGGCATACACGTC<br>GAAGACCAAGGGGACCTTCCGGGCTTCGCACTCGATGCTGCCCAAGAGGAGCTGAGTATGCTCCACAAAGAAAGAACCAACACAC<br>16S-G201C;C217G| TTCTAATACGACTGACTCATATTGGAAGACGACTCGAGTAGACTGCAGAAGGGGAACTCTGCTGGAGGGGTAAGTACTATTACGGCATACACGTC<br>GAAGACCAAGGGGACCTTCCGGGCTTCGCACTCGATGCTGCCCAAGAGGAGCTGAGTATGCTCCACAAAGAAAGAACCAACACAC
16S-G201C/T216G
TTCTAATACGACTCTATAGAGGAAACCTGATGAGCAGATTGCTGAAAGGAGGGGGATACTCGCTGAGAAGGGGAGTAGCTTATATACGCTAAGGCTG
GCAAGACCAAAGGGGCGACTCTGGCCCTGCCTGCCTGACTGAGTGGCCAAAAGGAGTCTGAGTACTGCTCAAAAAAAACAAACACACACACAC
16S-G203T
TTCTAATACGACTCTATAGAGGAAACCTGATGAGCAGATTGCTGAAAGGAGGGGGATACTCGCTGAGAAGGGGAGTAGCTTATATACGCTAAGGCTG
GCAAGACCAAAGGGGCGACTCTGGCCCTGCCTGCCTGACTGAGTGGCCAAAAGGAGTCTGAGTACTGCTCAAAAAAAACAAACACACACACAC
16S-C215A
TTCTAATACGACTCTATAGAGGAAACCTGATGAGCAGATTGCTGAAAGGAGGGGGATACTCGCTGAGAAGGGGAGTAGCTTATATACGCTAAGGCTG
GCAAGACCAAAGGGGCGACTCTGGCCCTGCCTGCCTGACTGAGTGGCCAAAAGGAGTCTGAGTACTGCTCAAAAAAAACAAACACACACACAC
16S-C214A
TTCTAATACGACTCTATAGAGGAAACCTGATGAGCAGATTGCTGAAAGGAGGGGGATACTCGCTGAGAAGGGGAGTAGCTTATATACGCTAAGGCTG
GCAAGACCAAAGGGGCGACTCTGGCCCTGCCTGCCTGACTGAGTGGCCAAAAGGAGTCTGAGTACTGCTCAAAAAAAACAAACACACACACAC
16S-G203T/C215A
TTCTAATACGACTCTATAGAGGAAACCTGATGAGCAGATTGCTGAAAGGAGGGGGATACTCGCTGAGAAGGGGAGTAGCTTATATACGCTAAGGCTG
GCAAGACCAAAGGGGCGACTCTGGCCCTGCCTGCCTGACTGAGTGGCCAAAAGGAGTCTGAGTACTGCTCAAAAAAAACAAACACACACACAC
16S-G203T/C214A
TTCTAATACGACTCTATAGAGGAAACCTGATGAGCAGATTGCTGAAAGGAGGGGGATACTCGCTGAGAAGGGGAGTAGCTTATATACGCTAAGGCTG
GCAAGACCAAAGGGGCGACTCTGGCCCTGCCTGCCTGACTGAGTGGCCAAAAGGAGTCTGAGTACTGCTCAAAAAAAACAAACACACACACAC
16S-C215G
TTCTAATACGACTCTATAGAGGAAACCTGATGAGCAGATTGCTGAAAGGAGGGGGATACTCGCTGAGAAGGGGAGTAGCTTATATACGCTAAGGCTG
GCAAGACCAAAGGGGCGACTCTGGCCCTGCCTGCCTGACTGAGTGGCCAAAAGGAGTCTGAGTACTGCTCAAAAAAAACAAACACACACACAC
16S-C214G
TTCTAATACGACTCTATAGAGGAAACCTGATGAGCAGATTGCTGAAAGGAGGGGGATACTCGCTGAGAAGGGGAGTAGCTTATATACGCTAAGGCTG
GCAAGACCAAAGGGGCGACTCTGGCCCTGCCTGCCTGACTGAGTGGCCAAAAGGAGTCTGAGTACTGCTCAAAAAAAACAAACACACACACAC
16S-G203G/C215G
TTCTAATACGACTCTATAGAGGAAACCTGATGAGCAGATTGCTGAAAGGAGGGGGATACTCGCTGAGAAGGGGAGTAGCTTATATACGCTAAGGCTG
GCAAGACCAAAGGGGCGACTCTGGCCCTGCCTGCCTGACTGAGTGGCCAAAAGGAGTCTGAGTACTGCTCAAAAAAAACAAACACACACACAC
16S-G203G/C214G
TTCTAATACGACTCTATAGAGGAAACCTGATGAGCAGATTGCTGAAAGGAGGGGGATACTCGCTGAGAAGGGGAGTAGCTTATATACGCTAAGGCTG
GCAAGACCAAAGGGGCGACTCTGGCCCTGCCTGCCTGACTGAGTGGCCAAAAGGAGTCTGAGTACTGCTCAAAAAAAACAAACACACACACAC
16S-A143C
TTCTAATACGACTCTATAGAGGAAACCTGATGAGCAGATTGCTGAAAGGAGGGGGATACTCGCTGAGAAGGGGAGTAGCTTATATACGCTAAGGCTG
GCAAGACCAAAGGGGCGACTCTGGCCCTGCCTGCCTGACTGAGTGGCCAAAAGGAGTCTGAGTACTGCTCAAAAAAAACAAACACACACACAC
16S-A143T
TTCTAATACGACTCTATAGAGGAAACCTGATGAGCAGATTGCTGAAAGGAGGGGGATACTCGCTGAGAAGGGGAGTAGCTTATATACGCTAAGGCTG
GCAAGACCAAAGGGGCGACTCTGGCCCTGCCTGCCTGACTGAGTGGCCAAAAGGAGTCTGAGTACTGCTCAAAAAAAACAAACACACACACAC
16S-A143G
TTCTAATACGACTCTATAGAGGAAACCTGATGAGCAGATTGCTGAAAGGAGGGGGATACTCGCTGAGAAGGGGAGTAGCTTATATACGCTAAGGCTG
GCAAGACCAAAGGGGCGACTCTGGCCCTGCCTGCCTGACTGAGTGGCCAAAAGGAGTCTGAGTACTGCTCAAAAAAAACAAACACACACACAC
16S-C176A
TTCTAATACGACTCTATAGAGGAAACCTGATGAGCAGATTGCTGAAAGGAGGGGGATACTCGCTGAGAAGGGGAGTAGCTTATATACGCTAAGGCTG
GCAAGACCAAAGGGGCGACTCTGGCCCTGCCTGCCTGACTGAGTGGCCAAAAGGAGTCTGAGTACTGCTCAAAAAAAACAAACACACACACAC
16S-G143T
TTCTAATACGACTCTATAGAGGAAACCTGATGAGCAGATTGCTGAAAGGAGGGGGATACTCGCTGAGAAGGGGAGTAGCTTATATACGCTAAGGCTG
GCAAGACCAAAGGGGCGACTCTGGCCCTGCCTGCCTGACTGAGTGGCCAAAAGGAGTCTGAGTACTGCTCAAAAAAAACAAACACACACACAC
16S-G143C
TTCTAATACGACTCTATAGAGGAAACCTGATGAGCAGATTGCTGAAAGGAGGGGGATACTCGCTGAGAAGGGGAGTAGCTTATATACGCTAAGGCTG
GCAAGACCAAAGGGGCGACTCTGGCCCTGCCTGCCTGACTGAGTGGCCAAAAGGAGTCTGAGTACTGCTCAAAAAAAACAAACACACACACAC
16S-C176G
TTCTAATACGACTCTATAGAGGAAACCTGATGAGCAGATTGCTGAAAGGAGGGGGATACTCGCTGAGAAGGGGAGTAGCTTATATACGCTAAGGCTG
GCAAGACCAAAGGGGCGACTCTGGCCCTGCCTGCCTGACTGAGTGGCCAAAAGGAGTCTGAGTACTGCTCAAAAAAAACAAACACACACACAC
16S-G143C/C176G
TTCTAATACGACTCTATAGAGGAAACCTGATGAGCAGATTGCTGAAAGGAGGGGGATACTCGCTGAGAAGGGGAGTAGCTTATATACGCTAAGGCTG
GCAAGACCAAAGGGGCGACTCTGGCCCTGCCTGCCTGACTGAGTGGCCAAAAGGAGTCTGAGTACTGCTCAAAAAAAACAAACACACACACAC
16S-G143C/C217G
TTCTAATACGACTCTATAGAGGAAACCTGATGAGCAGATTGCTGAAAGGAGGGGGATACTCGCTGAGAAGGGGAGTAGCTTATATACGCTAAGGCTG
GCAAGACCAAAGGGGCGACTCTGGCCCTGCCTGCCTGACTGAGTGGCCAAAAGGAGTCTGAGTACTGCTCAAAAAAAACAAACACACACACAC
16S-G146C
TTCTAATACGACTCTATAGAGGAAACCTGATGAGCAGATTGCTGAAAGGAGGGGGATACTCGCTGAGAAGGGGAGTAGCTTATATACGCTAAGGCTG
GCAAGACCAAAGGGGCGACTCTGGCCCTGCCTGCCTGACTGAGTGGCCAAAAGGAGTCTGAGTACTGCTCAAAAAAAACAAACACACACACAC
16S-C175G
TTCTAATACGACTCTATAGAGGAAACCTGATGAGCAGATTGCTGAAAGGAGGGGGATACTCGCTGAGAAGGGGAGTAGCTTATATACGCTAAGGCTG
GCAAGACCAAAGGGGCGACTCTGGCCCTGCCTGCCTGACTGAGTGGCCAAAAGGAGTCTGAGTACTGCTCAAAAAAAACAAACACACACACAC
16S-G146C;C175G
TTCTAATACGACTCACTATAAGGAACGCTGAGTGAGTTCGGAAGGGGAAGGATCCTGCTATGAGGGGATGATGACTTAATAGGGCATAAACGTC
GCAAGACCAAGGGGGACCTTGGGCTTCTGCATCGAGTGCTCCCAGGGAGGTCAGTACCTGACCACAAAAGAAACAACACACAC

16S-G146A
TTCTAATACGACTCACTATAAGGAACGCTGAGTGAGTTCGGAAGGGGAAGGATCCTGCTATGAGGGGATGATGACTTAATAGGGCATAAACGTC
GCAAGACCAAGGGGGACCTTGGGCTTCTGCATCGAGTGCTCCCAGGGAGGTCAGTACCTGACCACAAAAGAAACAACACACAC

16S-C175T
TTCTAATACGACTCACTATAAGGAACGCTGAGTGAGTTCGGAAGGGGAAGGATCCTGCTATGAGGGGATGATGACTTAATAGGGCATAAACGTC
GCAAGACCAAGGGGGACCTTGGGCTTCTGCATCGAGTGCTCCCAGGGAGGTCAGTACCTGACCACAAAAGAAACAACACACAC

16S-G146A;C175G
TTCTAATACGACTCACTATAAGGAACGCTGAGTGAGTTCGGAAGGGGAAGGATCCTGCTATGAGGGGATGATGACTTAATAGGGCATAAACGTC
GCAAGACCAAGGGGGACCTTGGGCTTCTGCATCGAGTGCTCCCAGGGAGGTCAGTACCTGACCACAAAAGAAACAACACACAC

16S-G146C
TTCTAATACGACTCACTATAAGGAACGCTGAGTGAGTTCGGAAGGGGAAGGATCCTGCTATGAGGGGATGATGACTTAATAGGGCATAAACGTC
GCAAGACCAAGGGGGACCTTGGGCTTCTGCATCGAGTGCTCCCAGGGAGGTCAGTACCTGACCACAAAAGAAACAACACACAC

16S-G212A
TTCTAATACGACTCACTATAAGGAACGCTGAGTGAGTTCGGAAGGGGAAGGATCCTGCTATGAGGGGATGATGACTTAATAGGGCATAAACGTC
GCAAGACCAAGGGGGACCTTGGGCTTCTGCATCGAGTGCTCCCAGGGAGGTCAGTACCTGACCACAAAAGAAACAACACACAC

16S-G212C
TTCTAATACGACTCACTATAAGGAACGCTGAGTGAGTTCGGAAGGGGAAGGATCCTGCTATGAGGGGATGATGACTTAATAGGGCATAAACGTC
GCAAGACCAAGGGGGACCTTGGGCTTCTGCATCGAGTGCTCCCAGGGAGGTCAGTACCTGACCACAAAAGAAACAACACACAC

16S-A199C;T209G
TTCTAATACGACTCACTATAAGGAACGCTGAGTGAGTTCGGAAGGGGAAGGATCCTGCTATGAGGGGATGATGACTTAATAGGGCATAAACGTC
GCAAGACCAAGGGGGACCTTGGGCTTCTGCATCGAGTGCTCCCAGGGAGGTCAGTACCTGACCACAAAAGAAACAACACACAC

16S-A199T;T209A
TTCTAATACGACTCACTATAAGGAACGCTGAGTGAGTTCGGAAGGGGAAGGATCCTGCTATGAGGGGATGATGACTTAATAGGGCATAAACGTC
GCAAGACCAAGGGGGACCTTGGGCTTCTGCATCGAGTGCTCCCAGGGAGGTCAGTACCTGACCACAAAAGAAACAACACACAC

16S-T209A
TTCTAATACGACTCACTATAAGGAACGCTGAGTGAGTTCGGAAGGGGAAGGATCCTGCTATGAGGGGATGATGACTTAATAGGGCATAAACGTC
GCAAGACCAAGGGGGACCTTGGGCTTCTGCATCGAGTGCTCCCAGGGAGGTCAGTACCTGACCACAAAAGAAACAACACACAC

16S-G144C
TTCTAATACGACTCACTATAAGGAACGCTGAGTGAGTTCGGAAGGGGAAGGATCCTGCTATGAGGGGATGATGACTTAATAGGGCATAAACGTC
GCAAGACCAAGGGGGACCTTGGGCTTCTGCATCGAGTGCTCCCAGGGAGGTCAGTACCTGACCACAAAAGAAACAACACACAC

16S-C175G;G213C
TTCTAATACGACTCACTATAAGGAACGCTGAGTGAGTTCGGAAGGGGAAGGATCCTGCTATGAGGGGATGATGACTTAATAGGGCATAAACGTC
GCAAGACCAAGGGGGACCTTGGGCTTCTGCATCGAGTGCTCCCAGGGAGGTCAGTACCTGACCACAAAAGAAACAACACACAC

16S-T209A
TTCTAATACGACTCACTATAAGGAACGCTGAGTGAGTTCGGAAGGGGAAGGATCCTGCTATGAGGGGATGATGACTTAATAGGGCATAAACGTC
GCAAGACCAAGGGGGACCTTGGGCTTCTGCATCGAGTGCTCCCAGGGAGGTCAGTACCTGACCACAAAAGAAACAACACACAC
<table>
<thead>
<tr>
<th>Mutation</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>16S-G147C</td>
<td>TTCTAATACGAATCTACTATTAGGAAGACCTTGAGATGTGGAAAGGAAGAGGAAGCGAGTGAATTGAGGGGTATAGTTGAATAGCTAAATACGGGATAAACGTC</td>
</tr>
<tr>
<td>16S-G147C;C175G</td>
<td>TTCTAATACGAATCTACTATTAGGAAGACCTTGAGATGTGGAAAGGAAGAGGAAGCGAGTGAATTGAGGGGTATAGTTGAATAGCTAAATACGGGATAAACGTC</td>
</tr>
<tr>
<td>16S-G177C</td>
<td>TTCTAATACGAATCTACTATTAGGAAGACCTTGAGATGTGGAAAGGAAGAGGAAGCGAGTGAATTGAGGGGTATAGTTGAATAGCTAAATACGGGATAAACGTC</td>
</tr>
<tr>
<td>16S-G177A</td>
<td>TTCTAATACGAATCTACTATTAGGAAGACCTTGAGATGTGGAAAGGAAGAGGAAGCGAGTGAATTGAGGGGTATAGTTGAATAGCTAAATACGGGATAAACGTC</td>
</tr>
<tr>
<td>16S-G145T;G177A</td>
<td>TTCTAATACGAATCTACTATTAGGAAGACCTTGAGATGTGGAAAGGAAGAGGAAGCGAGTGAATTGAGGGGTATAGTTGAATAGCTAAATACGGGATAAACGTC</td>
</tr>
<tr>
<td>16S-T208C</td>
<td>TTCTAATACGAATCTACTATTAGGAAGACCTTGAGATGTGGAAAGGAAGAGGAAGCGAGTGAATTGAGGGGTATAGTTGAATAGCTAAATACGGGATAAACGTC</td>
</tr>
<tr>
<td>16S-T218C</td>
<td>TTCTAATACGAATCTACTATTAGGAAGACCTTGAGATGTGGAAAGGAAGAGGAAGCGAGTGAATTGAGGGGTATAGTTGAATAGCTAAATACGGGATAAACGTC</td>
</tr>
<tr>
<td>16S-T208C;T218C</td>
<td>TTCTAATACGAATCTACTATTAGGAAGACCTTGAGATGTGGAAAGGAAGAGGAAGCGAGTGAATTGAGGGGTATAGTTGAATAGCTAAATACGGGATAAACGTC</td>
</tr>
<tr>
<td>16S-T219C</td>
<td>TTCTAATACGAATCTACTATTAGGAAGACCTTGAGATGTGGAAAGGAAGAGGAAGCGAGTGAATTGAGGGGTATAGTTGAATAGCTAAATACGGGATAAACGTC</td>
</tr>
<tr>
<td>16S-G145C;T219C</td>
<td>TTCTAATACGAATCTACTATTAGGAAGACCTTGAGATGTGGAAAGGAAGAGGAAGCGAGTGAATTGAGGGGTATAGTTGAATAGCTAAATACGGGATAAACGTC</td>
</tr>
<tr>
<td>16S-G198C</td>
<td>TTCTAATACGAATCTACTATTAGGAAGACCTTGAGATGTGGAAAGGAAGAGGAAGCGAGTGAATTGAGGGGTATAGTTGAATAGCTAAATACGGGATAAACGTC</td>
</tr>
<tr>
<td>16S-C210G</td>
<td>TTCTAATACGAATCTACTATTAGGAAGACCTTGAGATGTGGAAAGGAAGAGGAAGCGAGTGAATTGAGGGGTATAGTTGAATAGCTAAATACGGGATAAACGTC</td>
</tr>
<tr>
<td>16S-G198C;C210G</td>
<td>TTCTAATACGAATCTACTATTAGGAAGACCTTGAGATGTGGAAAGGAAGAGGAAGCGAGTGAATTGAGGGGTATAGTTGAATAGCTAAATACGGGATAAACGTC</td>
</tr>
<tr>
<td>16S-G200T</td>
<td>TTCTAATACGAATCTACTATTAGGAAGACCTTGAGATGTGGAAAGGAAGAGGAAGCGAGTGAATTGAGGGGTATAGTTGAATAGCTAAATACGGGATAAACGTC</td>
</tr>
<tr>
<td>16S-T208A</td>
<td>TTCTAATACGAATCTACTATTAGGAAGACCTTGAGATGTGGAAAGGAAGAGGAAGCGAGTGAATTGAGGGGTATAGTTGAATAGCTAAATACGGGATAAACGTC</td>
</tr>
<tr>
<td>16S-G200T;T208A</td>
<td>TTCTAATACGAATCTACTATTAGGAAGACCTTGAGATGTGGAAAGGAAGAGGAAGCGAGTGAATTGAGGGGTATAGTTGAATAGCTAAATACGGGATAAACGTC</td>
</tr>
<tr>
<td>16S-G200T;T208A;T218C</td>
<td>TTCTAATACGAATCTACTATTAGGAAGACCTTGAGATGTGGAAAGGAAGAGGAAGCGAGTGAATTGAGGGGTATAGTTGAATAGCTAAATACGGGATAAACGTC</td>
</tr>
<tr>
<td>16S-G200C</td>
<td>TTCTAATACGAATCTACTATTAGGAAGACCTTGAGATGTGGAAAGGAAGAGGAAGCGAGTGAATTGAGGGGTATAGTTGAATAGCTAAATACGGGATAAACGTC</td>
</tr>
<tr>
<td>16S-T208G</td>
<td>TTCTAATACGAATCTACTATTAGGAAGACCTTGAGATGTGGAAAGGAAGAGGAAGCGAGTGAATTGAGGGGTATAGTTGAATAGCTAAATACGGGATAAACGTC</td>
</tr>
<tr>
<td>16S-G200C;T208G</td>
<td>TTCTAATACGAATCTACTATTAGGAAGACCTTGAGATGTGGAAAGGAAGAGGAAGCGAGTGAATTGAGGGGTATAGTTGAATAGCTAAATACGGGATAAACGTC</td>
</tr>
<tr>
<td>16S-G200T;T218G</td>
<td>TTCTAATACGAATCTACTATTAGGAAGACCTTGAGATGTGGAAAGGAAGAGGAAGCGAGTGAATTGAGGGGTATAGTTGAATAGCTAAATACGGGATAAACGTC</td>
</tr>
<tr>
<td>16S-G144T</td>
<td>TTCTAATACGAATCTACTATTAGGAAGACCTTGAGATGTGGAAAGGAAGAGGAAGCGAGTGAATTGAGGGGTATAGTTGAATAGCTAAATACGGGATAAACGTC</td>
</tr>
<tr>
<td>16S-G144T;T218C</td>
<td>TTCTAATACGAATCTACTATTAGGAAGACCTTGAGATGTGGAAAGGAAGAGGAAGCGAGTGAATTGAGGGGTATAGTTGAATAGCTAAATACGGGATAAACGTC</td>
</tr>
</tbody>
</table>
SUPPORTING REFERENCES


