

Improved statistical methods reveal direct interactions between 16S and 23S rRNA

Scott T. Kelley*, Viatcheslav R. Akmaev¹ and Gary D. Stormo²

Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, CO 80309-0347, USA, ¹Genzyme Corporation, 1 Mountain Road, PO Box 9322, Framingham, MA 01701, USA and ²Department of Genetics, Washington University Medical School, 660 South Euclid, Box 8232, St Louis, MO 63110, USA

Received August 4, 2000; Revised October 11, 2000; Accepted October 25, 2000

ABSTRACT

Recent biochemical studies have indicated a number of regions in both the 16S and 23S rRNA that are exposed on the ribosomal subunit surface. In order to predict potential interactions between these regions we applied novel phylogenetically-based statistical methods to detect correlated nucleotide changes occurring between the rRNA molecules. With these methods we discovered a number of highly significant correlated changes between different sets of nucleotides in the two ribosomal subunits. The predictions with the highest correlation values belong to regions of the rRNA subunits that are in close proximity according to recent crystal structures of the entire ribosome. We also applied a new statistical method of detecting base triple interactions within these same rRNA subunit regions. This base triple statistic predicted a number of new base triples not detected by pair-wise interaction statistics within the rRNA molecules. Our results suggest that these statistical methods may enhance the ability to detect novel structural elements both within and between RNA molecules.

INTRODUCTION

The two main RNA components of the ribosome, 16S and 23S rRNA, have long been thought to directly interact across the interface of the two ribosomal subunits (1–3). The two subunits of the ribosome perform protein translation in a highly coordinated manner, suggesting that many intermolecular interactions occur between the subunits. Several molecular studies have uncovered regions of the 16S or 23S subunit that are directly connected (from cross-linking studies) or are protected from chemical modification by association with the opposite subunit (1,2,4–6). The most direct connection was discovered by Mitchell *et al.* (5), who cross-linked nucleotides 1408–1411 and 1518–1520 of 16S rRNA to nucleotides 1912–1920 of 23S rRNA. More recently, Merryman *et al.* (7,8) completed a comprehensive study on the effects of subunit association on

the accessibility of nucleotides in 16S and 23S rRNA. Using a series of chemical probes the authors produced an extensive map of sites in 16S and 23S of *Escherichia coli* that are chemically protected by association with the opposite ribosomal subunit (i.e. 16S with the 50S subunit and 23S with the 30S subunit).

This broad evidence of subunit association, along with previous evidence of direct contact between the 16S and 23S rRNA, suggests that interactions between the ribosomal RNA molecules are an important aspect of ribosome structure and function. In order to determine which specific nucleotides might be involved in such intermolecular interactions, we utilized recently developed comparative analysis techniques to identify bases that show significant co-variation between the two rRNA molecules (9,10). Comparative analyses have proven very effective at uncovering secondary and tertiary structure within numerous RNA molecules (11–15) and similar principles should apply to the study of co-variation between RNA molecules (16).

In this study we employed improved methods of comparative analysis that directly incorporate phylogenetic information in the prediction of interactions. Several authors have pointed out that phylogenetic information should help improve the prediction of interactions because it increases the signal to noise ratio by identifying the number of truly independent evolutionary changes (17–19). In a previous paper we demonstrated that incorporation of phylogenetic information enhanced the prediction of interactions over standard comparative analyses, which are already quite effective (9). Using these refined techniques we examined the nucleotides identified by Merryman *et al.* (7,8) as being at the interface of the two subunits and therefore potentially interacting. Such predictions not only indicate which particular nucleotides co-vary between the rRNAs, but also which regions of the two molecules might interact in three-dimensional space. Indeed, our statistical analysis predicted a number of nucleotides that may interact across the subunit interface and recent crystallography studies of the entire ribosome allowed us to verify that several of these co-varying nucleotides lie in close proximity of one another (20). Finally, the new computational techniques also enabled us to predict novel base triples within both 16S and 23S rRNA.

*To whom correspondence should be addressed. Tel: +1 303 735 1808; Fax: +1 303 492 7744; Email: scott.kelley@colorado.edu

MATERIALS AND METHODS

Data set

Sequence alignments of 16S and 23S rRNA were downloaded from the Internet (21,22). The database of these sequences can be accessed at URL <http://rrna.uia.ac.be/>. We collected 16S and 23S sequences from 144 different species belonging to a diverse array of bacteria, eukarya and archaea. We excluded the mitochondrial sequences from the analyses, because they were missing several of the regions identified at the interface by Merryman *et al.* (7,8). From each of the sequences we selected 10 regions, five from 16S and five from 23S rRNA, that were identified as protected by subunit association. We included not only the nucleotides they identified directly, but also some surrounding nucleotides, so that in most cases the regions corresponded to known stem structures. This gave us a 'positive control' to test whether our techniques were able to recover known interactions in the stem regions of 16S and 23S rRNA. In 16S rRNA the regions we used corresponded to the following nucleotide positions in *E.coli*: 240–280, 670–730, 764–820, 890–910 and 1408–1500. In 23S rRNA the regions corresponded to the following *E.coli* nucleotide positions: 640–680, 700–740, 870–910, 1680–1750 and 1910–1960.

Phylogenetic methods

Details of the statistical methods used in this paper were presented in two previous publications (9,10). In those papers we developed two statistics to predict interacting bases, H_{ij} and R_{ij} , and showed that, because they incorporate phylogenetic information, they perform better than standard mutual information (MI) methods. H_{ij} is the more robust of the two statistics and would be the method of choice except that it is computationally intensive and requires too much time to calculate for all of the potentially interacting pairs. Thus, we used the easily calculated R_{ij} statistic to identify the most likely interacting pairs, which are then tested more rigorously using H_{ij} .

We also used a parsimony approach, using the program MacClade v.3.01 (23), to verify that the positions at which our methods predicted interactions had, in fact, changed multiple times on the phylogenetic tree. Given a phylogenetic tree and the character states for each taxon (leaf) of the tree (in this case the nucleotide bases at each position in the sequence), the program reconstructs the ancestral states at each node of the tree using the parsimony criterion (23). One of the outputs in the program is the 'tree length' for each position in the sequence, which is a calculation of the minimal number of changes that must have occurred at every position given the tree. In this way we demonstrate that interactions with significant H_{ij} values have also changed multiple times. This is similar to the approach of other researchers who verify that two positions have changed together at least twice before accepting the correlation.

The third statistic we employed detected simultaneous correlations among three nucleotides. Base triples have been predicted and experimentally verified in a number of RNA molecules, including 23S rRNA, type I and II tRNA and group I introns, showing that they are an important and common RNA structural motif (24). Previous studies showed that S_{ijk} was adept at predicting known triple interactions and distinguishing these correlations from other types of interactions (10).

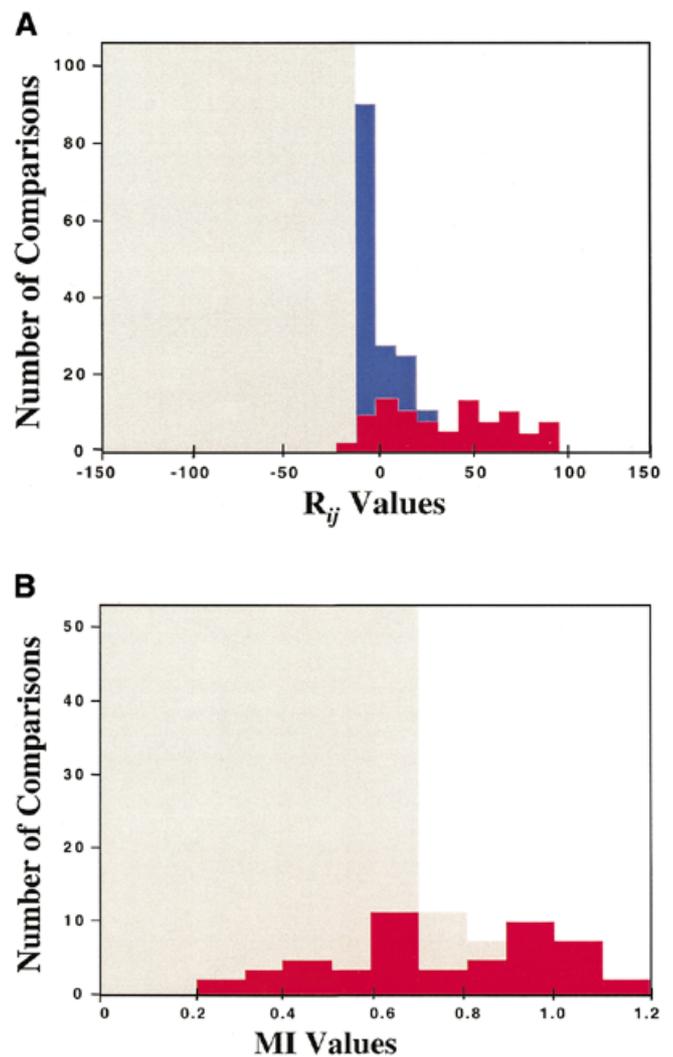


Figure 1. Distribution of pair-wise correlation values. (A) Histogram of R_{ij} statistic results for all pair-wise combinations of nucleotides in potentially interacting regions of 16S and 23S rRNA, with the number of comparisons truncated at 100. The histogram includes values for combinations both within and between molecules. Values for known base pairs within 16S and 23S are shown in red. Other potentially interacting positions with R_{ij} values less than -10 are shown in blue. (B) MI values for the same set of pair-wise combinations (truncated at 50) with values for known base pairs shown in red. For both figures the gray values represent other pair-wise combination scores (from the nearly 47 000 comparisons made).

RESULTS AND DISCUSSION

R_{ij} statistic results

For the statistical analysis we removed positions that were $>95\%$ conserved or where $\geq 15\%$ of the sequences had deletions. This left 143 and 164 variable positions in 16S and 23S rRNA, respectively. R_{ij} statistics were computed for all possible pairs of these positions using the program CgHKY (9). The total number of pair-wise comparisons was 46 971, of which 23 452 were intermolecular. A histogram of the results is presented in Figure 1A.

To test the reliability of our methods we first determined how many of the known interactions within 16S and 23S rRNA were accurately predicted. Previous work indicated that an R_{ij} value of 0 was a reasonable threshold with which to detect truly correlated positions (9). There are 105 known pairs (not counting the positions that we ignored because they are too highly conserved or contain too many gaps) and 92 of them (88%) have $R_{ij} > 0$ (Fig. 1A). If we use a threshold of -10 for R_{ij} , all but nine of the known pairs are found (i.e. there were nine false negatives). The nine 'true pairs' we did not predict did not form pairs in many of the sequences in the data set, a fact which explains their low correlation values. There are 102 additional pairs with $R_{ij} > -10$ (dark blue in Fig. 1A), of which 94 are intermolecular and represent potential interactions between the subunits. The other eight intramolecular pairs ($R_{ij} > -10$) are not found in the basic secondary structure, but may represent parts of base triples (Table 2) or possible false positives. The fact that we were able to predict most of the recognized stem structure using only 144 sequences, with most of the values greatly exceeding 0, indicated the general reliability of this approach for predicting nucleotide interactions. Compared with the R_{ij} statistic MI does not perform nearly as well (Fig. 1B).

H_{ij} statistic results

After finding the most likely candidate nucleotide positions for intermolecular interaction using the R_{ij} statistic, we then tested the significance of these results with the more rigorous H_{ij} statistic. H_{ij} utilizes phylogenetic information more effectively than R_{ij} and also approximates a χ^2 distribution with 9 degrees of freedom (9). Given 23 452 intermolecular pair-wise comparisons, we set the H_{ij} significance threshold at <1 in 25 000 ($P < 0.00004$). If the H_{ij} values exceeded this conservative significance threshold we would expect to see such a result less than one time by chance. [At the standard significance value of 1 in 20 ($P < 0.05$) we would expect to see more than 1000 significant values by chance with 23 452 pair-wise comparisons.] Because H_{ij} has 9 degrees of freedom (9), we needed H_{ij} values greater than 36 in order to expect no occurrences by chance (i.e. $P < 0.00004$). We considered this a conservative threshold because even some of the known pairs do not have H_{ij} values greater than 36 (Fig. 2). This reasoning also underlies the importance of using the Merryman *et al.* (7,8) data for our analyses. By limiting the analysis to regions where we expect to see interactions between the subunits we tested $\sim 10^4$ potential pairs. On the other hand, if we had searched for interacting positions throughout all the 16S and 23S rRNAs, that would have resulted in $>10^6$ comparisons and the threshold for significance would have been much more stringent. With typical secondary structure prediction one can apply additional information to detect intramolecular interactions, such as the expectation that base pairs will usually exist in antiparallel helices. However, for the intermolecular interactions we did not have such expectations and could not apply such constraints. If there were reasonably long helices formed between the subunits, such obvious motifs would most likely have already been identified. Instead, we expected to find isolated pairs, perhaps interacting as base triples or other complex structures.

The 12 intermolecular pairs with highest values of H_{ij} are shown in Table 1. Three of these exceed the value of 36 and

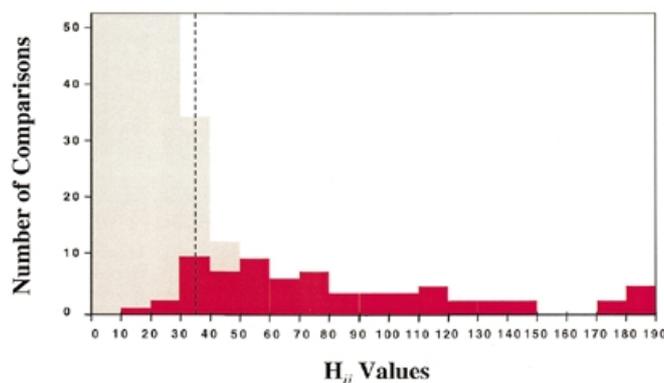


Figure 2. Histogram of H_{ij} statistic values for a set of ~ 300 intra- and intermolecular interactions, all with R_{ij} values greater than -10 . The dotted line indicates the H_{ij} critical value of 36 for $P < 0.00004$ (see text for explanation). Values for both inter- and intramolecular base pairs are shown in red.

have $P < 10^{-5}$ of occurring by chance. A few others have somewhat lower H_{ij} values in the range where false positives might be expected, but they still remain plausible because they have P values in the range occupied by several of the known base pairs (Fig. 2).

Although we have strong statistical evidence of intermolecular correlations, we acknowledge the possibility that the high H_{ij} values are artifacts of the rRNA alignments that we used. However, after manually constructing alternative alignments, particularly for the 1450 region of 16S rRNA, we were unable to find any better alignments. The correlation values of the alternative alignments were always less and there was no obviously better way to construe the alignment given the current structural models. Therefore, we trust the alignments we used and are confident that the H_{ij} statistic has identified truly correlated positions.

Nature of predicted intermolecular interactions

Figure 3 illustrates the six best candidates for intermolecular interactions according to our methods (Table 1). The highest co-variation scores we uncovered in our analyses corresponded to nucleotides on the penultimate stem of the 16S rRNA (Fig. 3). Recently published crystal structure information reveals that this stem region runs vertically along the 30S subunit body at the interface of the subunits and is certainly the most prominent feature of the subunit interface (20). The crystal structure information also places the helical region around position 1700 of 23S rRNA in close proximity to the 16S penultimate stem (20). Thus, we find significance in the fact that the two highest values in our co-variation analyses predict interactions between the 16S penultimate stem and the 1700 region of 23S rRNA (Table 1 and Fig. 3). The current crystal structure is not of high enough resolution to determine whether these nucleotides directly interact, but the predictions are at least plausible given this information.

The types of interactions shown in Figure 3 do not appear to be standard Watson–Crick base pairs. Instead, most of them appear to be more complicated motifs between sets of base pairs. Although models of such interactions are not currently

Table 1. R_{ij} and H_{ij} values for predicted intermolecular interactions between nucleotides in 16S and 23S rRNA

Number	Interactions		R_{ij}	H_{ij}	Steps		P value	False positives
	16S	23S			Pos 1	Pos 2		
1	1460	1706	0.52	51.21	26	13	<1.0 e ⁻⁵	<1
2	1459	1748a	25.12	48.73	21	29	<1.0 e ⁻⁵	<1
3	1459	660a	5.27	43.25	21	13	<1.0 e ⁻⁵	<1
4	1459	665 ^a	-4.35	32.56	21	27	<0.0002	5
5	263	678	1.63	31.68	13	4	<0.0003	7
6	276 ^a	1711 ^a	-1.99	31.36	10	35	<0.0003	7
7	1484 ^a	665 ^a	0.31	27.88	12	27	<0.0014	35
8	1408	1940	0.13	23.27	2	6	<0.0062	155
9	684	736	7.12	22.33	8	8	<0.009	225
10	252 ^a	649 ^a	-1.31	22.04	6	20	<0.009	225
11	1416 ^a	739	-2.72	21.87	15	11	<0.013	325
12	716	718	1.47	21.04	4	11	<0.013	325

Also included are the minimal number of mutations (steps) calculated by parsimony for each position on the tree.

The P value is the significance value for the H_{ij} statistic assuming a χ^2 distribution with 9 degrees of freedom. The false positives are the number of correlations expected by chance at each given P value with 25 000 total comparisons.

^aPositions that are statistically correlated with another intramolecular position and form a Watson-Crick base pair.

available, similar motifs have been discovered in other analyses of RNA structure. For instance, previous studies of interactions in RNase P uncovered strong correlations between sets of Watson-Crick paired nucleotides within the molecule (N.Pace, personal communication). Secondary structure diagrams of 16S rRNA (e.g. *Bacillus subtilis*) also indicate possible interactions occurring between groups of paired nucleotides (R.Gutell, S.Subashchandran, M.Schnare, Y.Du, N.Lin, L.Madabusi, K.Muller, N.Pande, N.Yu, Z.Shang, S.Date, D.Konings, V.Schweiker, B.Weiser and J.Cannone, manuscript in preparation).

Alternatively, these motifs may be similar to the base triple type interactions that have been discovered in a number of RNA molecules (18). A closer look at the statistics of the interactions indicates that, in some cases, at least one of the interacting pair groups is not present in many of the RNA molecules. For instance, Figure 3 shows 2 nt in 16S rRNA (positions 1459 and 1460) which are apparently base paired with other nucleotides in the stem region (positions 1443 and 1442, respectively). However, mutational change at 1459 is not correlated with change at 1443 nor is change at 1460 correlated with change at 1442.

If these pairs were truly interacting in most of the sequences one would expect them to be highly correlated. In fact, many of the sequences (~16% of those we utilized) did not appear to have a stem region at the end the helix pictured in Figure 3 and had more of a bulge or loop region instead. In this case these nucleotides would have the free hydrogen bonds necessary to form base triple type interactions. Further statistical tests using the S_{ijk} 'triple statistic' did not reveal substantial three-way co-variation at these nucleotide positions, which indicates that the main co-variation occurs between pairs of nucleotides.

The fact that these correlations do not seem to be standard base pairs or, perhaps, even base triples suggests that these

intermolecular interactions may be uncommon RNA motifs. This is particularly true of position 1459, which is statistically correlated with changes at three different nucleotide positions (Fig. 3 and Table 1). In this case we suggest that these nucleotides might be associated with a metal ion (25) or comprise some type of complicated RNA structure. Alternatively, the nature of the interactions at these positions may change as the conformation of the ribosome changes. Models of ribosome function suggest that in the process of translation the subunits rapidly change conformational states (26). In some states these positions may be base paired as part of a stem region, while in other states they may interact with helical regions in the 23S rRNA. Finally, there is the possibility that the nucleotide co-variations we predicted may be affected by concerted changes across the ribosome. If so, the nucleotides need not interact directly for meaningful co-variation to occur and the co-variations we observed would be indicative of change over a large section of the rRNA. As evidence of this, we note that the 16S penultimate helix is remarkably variable between phylogenetic groups in length and structure, especially given its important position at the subunit interface. The intermolecular co-variations we discovered in this helix may, therefore, reflect this variation. However, if that were the case we might also expect there to be many correlated positions at the end of the helix, rather than just the two we observe.

Base triple predictions

Along with the analysis of pair-wise interactions, we also applied a new 'triple statistic', S_{ijk} , to predict three-way correlations between nucleotides (10). Again, we did not include conserved positions or positions with numerous gaps. We applied this method to detect triples both within and between 16S and 23S rRNA using the regions identified by Merryman *et al.* as being on the interface of the subunits (7,8). Similar to

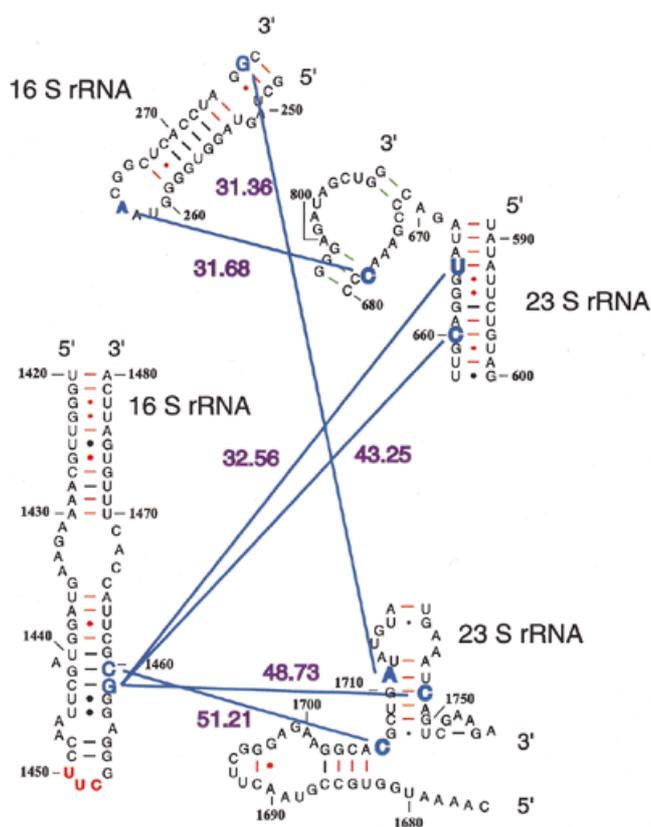


Figure 3. Diagram of the six most significant intermolecular correlations between regions of 16S and 23S rRNA. The helical models are based on the secondary structural models of Konings and Gutell (28) and Fields and Gutell (29). Red lines between bases indicate known interactions predicted by the R_{ij} method. Blue lines and nucleotides indicate predicted intermolecular correlations predicted by the R_{ij} statistic that were significant, or nearly significant, under the full model H_{ij} statistic (see Table 1). H_{ij} values for the correlations are shown in purple next to the lines. Green lines indicate known base pairs that were not tested for interaction.

the results with known triple interactions (10), we found a skewed and dichotomous statistical distribution of triple values (Fig. 4). All of the highest triple correlations we detected were within either 16S or 23S rRNA and none of them have been previously described to our knowledge. These are presented in Table 2 (along with the R_{ij} and H_{ij} statistics values for all the pair-wise combinations) and graphically in Figure 5. Interestingly, in some cases the triple combinations with high S_{ijk} values did not have high H_{ij} values, although they were always positive. This suggests that S_{ijk} is able to uncover triple interactions even when the pair-wise interaction statistics are not significant.

The most promising potential base triple discovered by S_{ijk} was in 16S rRNA between the paired positions 248–276 and the unpaired position 260 (Table 2 and Fig. 5). This particular interaction conforms well to standard base triple interactions found by other researchers, where an unpaired nucleotide connects to a paired group (24,27). The correlation between changes at the different positions is supported by all three statistical measures (R_{ij} , H_{ij} and S_{ijk} statistics; Table 2). We also discovered that one of the 23S nucleotides (position 1706)

predicted to be involved in an intermolecular interaction may also interact with the paired positions 1708 and 1750. And we found high S_{ijk} correlation values for an unusual interaction between two base pairs within 16S rRNA (Table 2 and Fig. 5). Similar to some of the predicted intermolecular interactions, these last two groups of three-way co-variations may represent some unusual RNA motifs of the types mentioned above.

Table 2. The six highest S_{ijk} value base triples within 16S and 23S rRNA

	Position 1	Position 2	Position 3	S_{ijk}
16S triples				
	247	260	277	1.02
	774	805	902	15.28
	774	805	897	1.65
	774	897	902	-1.02
	805	897	902	-1.02
23S triples				
	1708	1750	1706	1.92
				R_{ij}
				H_{ij}
16S triples				
	260	247	19.37	37.84
	260	277	18.36	39.95
Pair	247	277	50.99	50.39
Pair	774	805	32.69	31.6
	774	897	9.96	17.91
	774	902	12.66	20.52
	805	897	9.96	18.09
	805	902	12.66	19.52
Pair	897	902	31.18	37.83
23S triples				
Pair	1708	1750	14.07	15.87
	1708	1706	6.21	11.02
	1750	1706	7.21	8.29

R_{ij} and H_{ij} values for the pair-wise interactions are also presented.

CONCLUSION

In this paper we have used a suite of novel statistical and phylogenetic approaches to predict potential nucleotide interactions between the 23S and 16S rRNA molecules. By focusing on regions of the rRNA molecules experimentally identified as being at the interface of the two ribosomal subunits we were able to detect a number of highly significant co-variations between nucleotides in 16S and 23S rRNA. The predictions with the strongest statistical support included nucleotides belonging to the penultimate stem of the 16S rRNA that lies at the interface of the two ribosomal subunits (20). The corresponding 23S rRNA nucleotides also belong to regions at the ribosomal interface in close proximity to this 16S stem region.

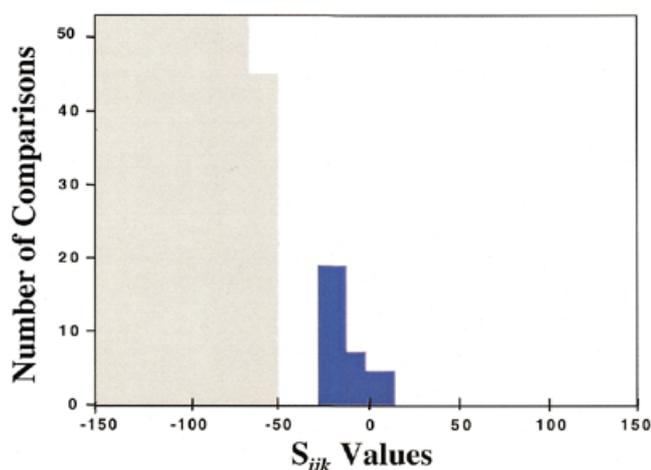


Figure 4. Histogram of S_{ijk} statistic results for base triples within 16S and 23S rRNA (see Table 2), with the number of comparisons cut at 50. The distribution of values is similar to the separation achieved with known base triples (10). The predicted interactions illustrated in Figure 5 are the three highest S_{ijk} values. The gray values represent other pair-wise combination scores.

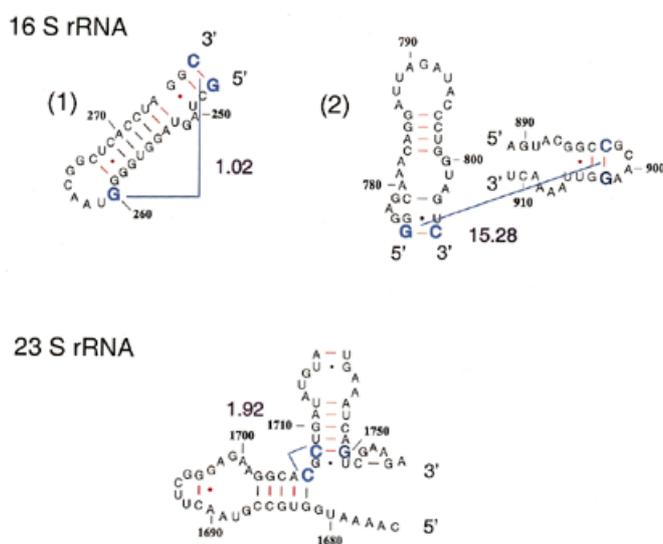


Figure 5. Novel interactions within 16S and 23S rRNA predicted using the S_{ijk} 'triple statistic' (see Table 2). The helical models are based on the secondary structural models of Konings and Gutell (28) and Fields and Gutell (29). Red lines indicate known interactions predicted by the R_{ij} statistic and blue lines and nucleotides show the predicted interactions. S_{ijk} values are shown in purple.

The types of interactions we discovered, if verified, represent rather unusual RNA motifs. However, similar types of interactions have been predicted in other RNA molecules, such as RNase P, which suggests that these types of motifs may be common in RNA molecules, although they are still not well

understood in a structural sense. Finally, using the same statistical approaches we also predicted several interesting new interactions within both 16S and 23S rRNA, suggesting that our methods may be generally useful in elucidating novel structural motifs in RNA molecules.

ACKNOWLEDGEMENTS

We would like to thank H. Noller for comments on some early results, N. Pace for access to unpublished versions of an earlier manuscript, E. Jabri for discussion and references on unusual RNA structures and motifs and two anonymous reviewers for helpful comments and questions. We also thank Jan Gorodkin and Shawn Strickland for help exploring and comparing alternative alignments of the rRNA sequences. This work was supported by NIH grant HG00249.

REFERENCES

- Chapman, N.M. and Noller, H.F. (1977) *J. Mol. Biol.*, **109**, 131–149.
- Herr, W., Chapman, N.M. and Noller, H.F. (1979) *J. Mol. Biol.*, **130**, 433–449.
- Moazed, D. and Noller, H.F. (1989) *Nature*, **342**, 142–148.
- Herr, W. and Noller, H.F. (1979) *J. Mol. Biol.*, **130**, 421–432.
- Mitchell, P., Osswald, M. and Brimacombe, R. (1992) *Biochemistry*, **31**, 3004–3011.
- Tappich, W.E. and Hill, W.E. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 556–560.
- Merryman, C., Moazed, D., Daubresse, G. and Noller, H.F. (1999) *J. Mol. Biol.*, **285**, 107–113.
- Merryman, C., Moazed, D., McWhirter, J. and Noller, H.F. (1999) *J. Mol. Biol.*, **285**, 97–105.
- Akmaev, V.R., Kelley, S.T. and Stormo, G.D. (1999) *Intelligent Systems for Molecular Biology*, pp. 10–17.
- Akmaev, V.R., Kelley, S.T. and Stormo, G.D. (2000) *Bioinformatics*, **16**, 501–512.
- Fox, G.W. and Woese, C.R. (1975) *Nature*, **256**, 505–507.
- Nishikawa, K. and Takemura, S. (1974) *FEBS Lett.*, **40**, 106–109.
- Noller, H.F. and Woese, C.R. (1981) *Science*, **212**, 403–411.
- Zucker-Aprison, E., Thomas, J.D. and Blumenthal, T. (1988) *Nucleic Acids Res.*, **16**, 7188.
- Gutell, R.R., Power, A., Hertz, G.Z., Putz, E.J. and Stormo, G.D. (1992) *Nucleic Acids Res.*, **20**, 5785–5795.
- Thanaraj, T.A. (1994) *Nucleic Acids Res.*, **22**, 3936–3942.
- Gulko, B. and Haussler, D. (1996) In Hunter, L. and Kleon, T. (eds), *Biocomputing: Proceedings of the 1996 Symposium*. World Scientific Publishing Co., Singapore, pp. 350–367.
- Gautheret, D., Damberger, S.H. and Gutell, R.R. (1995) *J. Mol. Biol.*, **248**, 27–43.
- Winker, S., Overbeek, R., Woese, C.R., Olsen, G.J. and Pfluger, N. (1990) *Comput. Appl. Biosci.*, **6**, 365–371.
- Cate, J.H., Yusupov, M.M., Yusupova, G.Z., Earnest, T.N. and Noller, H.F. (1999) *Science*, **285**, 2095–2104.
- Van de Peer, Y., Robbrecht, E., de Hoog, S., Caers, A., De Rijk, P. and De Wachter, R. (1999) *Nucleic Acids Res.*, **27**, 179–183.
- De Rijk, P., Robbrecht, E., de Hoog, S., Caers, A., Van de Peer, Y. and De Wachter, R. (1999) *Nucleic Acids Res.*, **27**, 174–178.
- Maddison, W.P. and Maddison, D.R. (1992) *MacClade 3.01*. Sinauer Associates, Sunderland, MA.
- Conn, G.L., Gutell, R.R. and Draper, D.E. (1998) *Biochemistry*, **37**, 11980–11988.
- Cate, J.H., Gooding, A.R., Podell, E., Zhou, K., Golden, B.L., Kundrot, C.E., Cech, T.R. and Doudna, J.A. (1996) *Science*, **273**, 1678–1685.
- Burkhardt, N., Junemann, R., Spahn, C.M. and Nierhaus, K.H. (1998) *Crit. Rev. Biochem. Mol. Biol.*, **33**, 95–149.
- Gowers, D.M. and Fox, K.R. (1999) *Nucleic Acids Res.*, **27**, 1569–1577.
- Konings, D.A. and Gutell, R.R. (1995) *RNA*, **1**, 559–574.
- Fields, D.S. and Gutell, R.R. (1996) *Fold Des.*, **1**, 419–430.