Published online 17 April 2008

Nucleic Acids Research, 2008, Vol. 36, No. 9 ***e52***
doi:10.1093/nar/gkn174

R-Coffee: a method for multiple alignment of non-coding RNA

Andreas Wilm[1], Desmond G. Higgins[1] and Cédric Notredame[2,*]

[1]The Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Ireland and [2]Centre for Genomic Regulation (CRG), Dr Aiguader, 88, 08003 Barcelona, Spain

Received December 20, 2007; Revised March 14, 2008; Accepted March 25, 2008

ABSTRACT

R-Coffee is a multiple RNA alignment package, derived from T-Coffee, designed to align RNA sequences while exploiting secondary structure information. R-Coffee uses an alignment-scoring scheme that incorporates secondary structure information within the alignment. It works particularly well as an alignment improver and can be combined with any existing sequence alignment method. In this work, we used R-Coffee to compute multiple sequence alignments combining the pairwise output of sequence aligners and structural aligners. We show that R-Coffee can improve the accuracy of all the sequence aligners. We also show that the consistency-based component of T-Coffee can improve the accuracy of several structural aligners. R-Coffee was tested on 388 BRAliBase reference datasets and on 11 longer Cmfinder datasets. Altogether our results suggest that the best protocol for aligning short sequences (less than 200 nt) is the combination of R-Coffee with the RNA pairwise structural aligner Consan. We also show that the simultaneous combination of the four best sequence alignment programs with R-Coffee produces alignments almost as accurate as those obtained with R-Coffee/Consan. Finally, we show that R-Coffee can also be used to align longer datasets beyond the usual scope of structural aligners. R-Coffee is freely available for download, along with documentation, from the T-Coffee web site (www.tcoffee.org).

INTRODUCTION

A number of recent discoveries have cast new light on the importance of RNA, revealing a functional scope much broader than realized only a few years ago. Small non-coding RNAs (ncRNAs) are actively involved in a wide range of cell processes, including gene regulation, cell differentiation, genome maintenance, RNA maturation and protein synthesis (1,2). The ncRNA big picture could change even further in the coming years, as suggested by a recent report of the ENCODE consortium (3) showing an unexpected level of ncRNA transcription across the entire human genome.

While the problem of aligning sequences has been regularly addressed over the last 40 years (4), delivering accurate alignments for ncRNAs remains a challenging task for at least two main reasons. First of all, RNA molecules have a low chemical complexity compared to proteins with just a four-letter alphabet. This limited information content makes it hard to use sequence similarity as an estimator of the biological relevance of RNA alignments. The most notable consequence is the limited sensitivity of RNA alignments, and it is generally accepted that the RNA twilight zone (i.e. the level of identity below which pairwise alignments become uninformative) is close to 70%, as opposed to 25% for proteins (5–7). The second reason for the difficulty in aligning ncRNA comes from their rate and pattern of evolution. In general, functional ncRNAs have a well-defined structure and their evolution seems to be mainly constrained to retain a specific folding, mostly based on Watson and Crick base pairs. Maintaining such a structure can be achieved through compensatory mutations, a phenomenon that explains why sequences can diverge a lot while still coding for the same structure (8). Therefore, sequence identity alone is a very crude measure of biological similarity, as it does not reflect very well the level of structural conservation.

Because of these problems, it is highly desirable to take RNA secondary structure into account while aligning ncRNA sequences, in order to assure optimal usage of the positional interdependence. Sankoff's algorithm, published 20 years ago (9), does exactly this, but suffers from enormous runtime and memory requirements. Given two sequences of length L, the pairwise alignment requires $O(L^6)$ in time and $O(L^4)$ in computer space, while its extension to N sequences is exponential: $O(L^{3N})$ in time and $O(L^{2N})$ in space. Only a few simplified implementations exist, usually constrained to pairwise alignment (10–13). Recently a number of multiple alignment versions

*To whom correspondence should be addressed. Tel: +34 93 316 02 71; Fax: +34 93 316 00 99; Email: cedric.notredame@crg.es

have been published (11,14–18), which employ various techniques to reduce run-time and memory consumption, for example by limiting the length or types of structure motifs or by using banding techniques during the dynamic programming stage [for review, see (19)]. The most accurate of theses heuristics are restricted to sequences shorter than approximately 200 residues, a limitation that explains why it is often more practical to use regular sequence aligners when dealing with larger sequences such as ribosomal RNA, or RNA motifs embedded in long genomic sequences. Most of these aligners treat RNA sequences like regular DNA and rely on an identity-based scoring scheme only suitable for closely related sequences. For instance, ClustalW has long been used for establishing reference collections of ribosomal RNA alignments (20). Following manual curation and visual inspection of the conserved secondary structures, these alignments have been widely used to infer phylogenetic relationships between most living organisms. Taking secondary structures into account may not, however, improve alignment accuracy across the entire spectrum of known ncRNA. For instance, secondary structure-based alignments will not improve the comparison of mature miRNAs or mRNAs that are not structurally conserved.

In this work, we address the problem of RNA multiple sequence alignments by taking advantage of the T-Coffee framework (21). T-Coffee is a multiple sequence alignment method able to combine the output of different sequence alignment packages. It takes as input, a collection of alignments (pairwise or multiple) and outputs a multiple sequence alignment containing all the sequences. The input, which is referred to as a 'library', can consist of alternative and possibly inconsistent alignments of the same sequences. The purpose of the algorithm is to generate a final alignment that is as consistent as possible with the original input alignments. The main advantage of this procedure is its flexibility. For instance, in the original T-Coffee, the library was compiled from pairwise local and global alignments of the sequences. In M-Coffee (22) the compilation was made using alternative multiple sequence alignments while in 3D-Coffee (23) or Expresso (24), the library is derived from structure-based pairwise alignments. This simple protocol can easily be built on top of any existing method, as illustrated by two RNA alignment packages: Marna (25) and T-Lara (19). Both packages focused on the development of a novel pairwise RNA alignment algorithm, which was then used to generate an alignment library fed to T-Coffee in order to produce a multiple alignment. In the present work we decided to go further and modify the library processing/extension algorithm so that it could take advantage of known and predicted secondary structures. This is done when compiling the library and while evaluating the matching score of two residues. This novel strategy forms the core of R-Coffee. Our primary goal was not to produce a stand-alone method, but rather a novel component that can seamlessly be added on top of any existing alignment method. We demonstrate here that it is possible to improve the alignment quality of most existing methods by means of R-Coffee, with only minor computational over-head.

SYSTEMS AND METHODS

Reference alignments and evaluation

We used two different benchmark sets: BRAliBase 2.0 (5), the standard RNA reference alignment dataset and Cmfinder (26), a smaller dataset specifically designed for testing local analysis of long sequences. BRAliBase is collection of RNA reference alignments especially designed for the benchmark of RNA alignment methods. We only used its multiple alignment component made of 388 multiple sequence alignments. These datasets are evenly distributed between 35% and 95% average sequence identity. Each of these datasets was originally produced by extracting sub-alignments from larger seed alignments coming from four RNA families (tRNA, group II intron, 5S rRNA and U5 RNA). Two of these were seed alignments obtained from the Rfam database (27). This procedure may appear slightly circular as it involves comparing sequence-based reference alignments with other sequence-based alignments. In order to address this issue, BRAliScore, the benchmarking scoring scheme, was designed in such a way that it not only depends on the similarity between the reference and the evaluated alignment but also on the intrinsic structural conservation of the target alignment [see also (6)]. This tradeoff illustrates the difficulties in establishing a gold standard for RNA analysis. The main problem comes from the lack of sufficient experimentally validated RNA structures, in contrast to protein sequence analysis where literally hundreds of accurate 3D structures exist.

The BRAliScore combines two measures: the Sum of Pairs Score (SPS) and the Structural Conservation Index (SCI) (28). The SPS is the ratio between the number of residue pairs identically aligned in the target and the reference, divided by the number of pairs in the reference. It was measured using a variant of compalignp [based on Sean Eddy's compalign; see also (6)] adapted to restrict the evaluation to a pre-defined core region. The SCI is a reference-independent measure. It is defined as the ratio between the average free energies of all single sequences of the MSA [as calculated by RNAfold; (29)] and the free energy of the MSA consensus structure [as calculated by RNAalifold; (30)]. A value of 0 indicates a lack of a conserved structure, 1 corresponds to a perfect agreement between the energies of the single sequences and the consensus energy, while values higher than 1 indicate a very good agreement supported by additional co-variation. The BRAliScore is the product of the SCI and the SPS score. This combination can lead to problems when either the SPS or the SCI are close to 0. In practice however, this situation rarely arises, and the combination of these two scores provides a very robust measure, less sensitive than the SPS, to the effective accuracy of the reference alignment. To test for statistical differences between pairs of methods, we applied the Wilcoxon signed rank test as in (6). All analyses were carried out using tools provided from http://www.biophys.uni-duesseldorf.de/bralibase/.

Our second dataset is named after the RNA motif finder program Cmfinder (26). It contains Rfam sequences embedded in 200 nt of their original flanking genomic

Table 1. Programs used for benchmarking and as input for T/R-Coffee

Program	Reference	Version	Structure	Sankoff	Alignment mode	Command line
ClustalW	(33)	1.83	N	N	Multiple	-type = dna
Consan	(10)	1.2	Y	Y	Pairwise	-m mixed80.mod
Dynalign	(12)	Dec-06	Y	Y	Pairwise	Len2-len1 + 5 0.4 5 20 2 1 0
Foldalign	(13)	2.0.3	Y	Y	Pairwise	-global -score_matrix global.fmat
FoldalignM	(15)	1.0.1	Y	Y	Multiple	
Mafft	(35)	5.861	Y	N	Multiple	ginsi/fftns
Marna	(25)	Jan-07	Y	N	Multiple (T-Coffee)	
M-Locarna	(17)	0.99	Y	Y	Multiple	mlocarna-p
Murlet	(14)	Mar-06	Y	Y	Multiple	
Muscle	(32)	3.6	N	N	Multiple	-seqtype rna
Pcma	(45)	2	N	N	Multiple	
Pmcomp	(11)	Jun-04	Y	Y	Pairwise	
Pmmulti	(11)	Jun-04	Y	Y	Multiple	
Poa	(46)	2	N	N	Multiple	blosum80.mat
Proalign	(47)	0.5.a3	N	N	Multiple	
Probcons	(34)	1.1	N	N	Multiple	
Prrn	(48)	SCC 3.0.a	N	N	Multiple	
Rnasampler	(44)	1.3	Y	Y	Multiple	
Stemloc	(16)	Dart 0.19b	Y	Y	Multiple	-multiple –slow -global
Stral	(49)	0.5.4	Y	N	Multiple	
T-Lara	(19)	1.31	Y	N	Multiple (T-Coffee)	-o lara.params
T-Coffee	(21)	5.19	N	N	Multiple	-dp_mode myers_miller_pair_wise

This table lists all the packages evaluated along with their version numbers (or download date). The Structure column indicates whether the consider packages use predicted secondary structures (Y for Yes, N for No). The Sankoff column indicates whether the package is a heuristic implementation of the Sankoff original algorithm. The Alignment Mode column indicates whether the package performs pairwise or multiple alignment or if it's based on the T-Coffee package. The last column gives used command line parameters; most programs were used as in the BRAliBase alignment benchmark publications (5) and (6).

regions, randomly distributed between the 5′ and the 3′ of the ncRNA sequence (i.e. x nucleotides on the 5′-, y nucleotides on the 3′-end with x and y randomly chosen so that $x + y = 200$). The unaligned datasets were kindly provided by the Cmfinder authors. We limited our choice to datasets having less than 40 sequences thus generating 11 reference alignments (9 to 35 sequences, length between 167 and 368 nt). The average level of identity within the core regions of these alignments ranges from 31% to 73%. These characteristics make the Cmfinder dataset a difficult target, especially because of the sequence length and the inclusion of flanking regions. These datasets are also closer to 'real life' situations that often involve discovering an RNA motif within poorly characterized sequences. The presence of flanking genomic regions potentially embedded in the Cmfinder datasets made it impossible to systematically use the SCI component of the BRAliscore. Instead we used the Sum of Pairs score (SPS) and restricted the scoring to the Rfam core region of the alignment.

Note that most available RNA alignment benchmark sets are based on Rfam alignments. These alignments are by no means a gold standard (especially not the 'full' Rfam alignments) and are not based on 3D superposition as most protein alignment benchmarks. For example, the BRAliBase benchmark set was created from four RNA families, two of which were full Rfam alignments (U5, g2intron) and two were Rfam seed alignments (tRNA, 5S) i.e. hand-curated and thus more likely of high quality. The Cmfinder data sets are exclusively based on Rfam seed alignments. The low number of quality alignments suited especially for benchmarking (i.e. equally distributed over a wide sequence identity range etc.) of multiple RNA alignment programs is a notorious problem. New RNA alignment benchmarks with a high number of RNAs using expert, hand-curated alignments, which are based on structural superposition [e.g. from the Comparative RNA web site (31)] would constitute a useful advance in this area.

Alignment programs

To test and benchmark R-Coffee, we compared a variety of programs with different features (Table 1). These programs can be roughly divided in three categories: pairwise structural aligners, multiple structural aligners and regular multiple sequence aligners. The pairwise structural aligners like Consan (10) are heuristic approximations to the original Sankoff algorithm. Their heavy computational requirements limit them to short sequences. The second category includes structural aligners extended to deal with multiple sequences like FoldalignM (15) or Stemloc (16). Like their pairwise counterparts, they use structure and sequence information during the alignment and are therefore restricted to short datasets. The third category is made of regular multiple sequence alignment programs like Muscle (32) or ClustalW (33). These programs do not rely on any kind of structural modeling, although some of them [like Probcons (34) and Mafft (35)] have optimized parameters for BRAliBase i.e. program parameters were trained on BRAliBase alignments by the respective program's authors. These two last categories of packages can either be used to align multiple sequence datasets or pairs of sequences.

Most programs were used as described in (5) and (6). Marna (25), Pmmulti (11) and Stemloc (16) were not able to align all test sets of BRAliBase. In particular, Marna cannot align sequences that contain IUPAC characters and the ability of Stemloc to align sequences seems to depend on the size of the main memory. We therefore had to exclude these packages from the comparison, although it should be noted that they produced accurate alignments on the datasets they could align (data not shown). In case of T-Lara we did not use the pairwise alignments as T-Lara already uses T-Coffee. Instead we used R-Coffee as a drop-in-replacement for T-Coffee. We used the stand-alone versions of all packages to compute multiple alignments for all the reference datasets. We also used them in combination with either T-Coffee or R-Coffee. All programs were run on a Quad-Xeon-3 GHz machine with 6 GB RAM running Red Hat Enterprise Linux.

Original T-Coffee strategy

T-Coffee is a versatile MSA package that allows the combination of many pairwise (or multiple) sequence alignments into one unique final model. The principle is fairly straightforward. Given a set of sequences, a collection of pairwise alignment is computed. This collection can be redundant (several alternative alignments for each pair of sequences) or not, and is compiled into a data structure called the primary library. The primary library contains the list of all the pairs of aligned residues observed in the alignment collection. Each of these pairs receives a weight equal to the score of the alignment it came from (in practice the percent identity is used). These weights are then re-estimated in a process named library extension. The purpose of the new weights (extended weights) is to reflect the level of consistency between each pair of aligned residues and the rest of the library. High-scoring pairs are those in very good agreement with the rest of the pairs and their high score insures that they should easily find their way into the final alignment. R-Coffee uses the Myers and Miller algorithm (command line option: -dp_mode = myers_miller_pair_wise) to align pairs of sequences or profiles rather than the current T-Coffee default (-dp_mode = cfasta_pair_wise) that uses a banded dynamic programming implementation extensively tuned for proteins. The Myers and Miller setting corresponds to the T-Coffee algorithm as described in the original publication (21).

Adaptation of T-Coffee to use RNA structural information

The novel RNA-specific mode of T-Coffee described here has been designed to be able to use secondary structure predictions. The current design supports an arbitrary amount of structural prediction, and each sequence can be associated with one or more secondary structure predictions that do not need to be in agreement. It is also possible not to associate any structural information with some sequences. In practice, however, we expect the best results to be obtained when using at least one secondary structure prediction for each sequence in the dataset. These structural predictions are passed to R-Coffee, using a data structure similar to the T-Coffee primary library and named a structural library. In this structural library, each line indicates a pair of nucleotides predicted to be base-paired. Like its primary sequence counterpart, this structural library can be redundant, contain conflicting pairs or lack data for some pairs.

RNA structures were computed using either a global or a local prediction method. Global structure predictions were obtained with RNAfold (29) which finds a structure with Minimal Folding Energy (MFE). When using a MFE structure as input, each predicted base pair was directly added to the structural library without any further filtering. This global MFE-based prediction has two major limitations: the algorithm is computationally demanding when being applied to very long sequences and its prediction accuracy decreases with sequence length (36,37). When dealing with long sequences, a sensible alternative is to use local RNA structure prediction methods such as RNAplfold (38). RNAplfold predicts local pair probabilities for base pairs within a certain span (default is set to 100 nt). The program outputs base pair probabilities rather than one precise structure and in order to reduce noise, we excluded pairs exhibiting a low thermodynamic probability. We determined a suitable probability threshold by varying the filtering threshold between 0.0 and 0.8 (in steps of 0.1) and estimating the accuracy of the resulting R-Coffee alignments (Figure 2). We found 0.3 to be the optimal threshold, although our results indicate a relative stability of the system around this value (flat peak). The structural pairs thus gathered are then fed to R-Coffee, the version of T-Coffee using the R-Score (see later). The structural libraries used here contain un-weighted structure pairs, although it is, in principle, possible to apply a weighting scheme onto these pairs, possibly reflecting the thermodynamic stability or the likelihood of each considered pair.

For testing purposes we also used random structures as input. These structures were computed by shuffling the input sequences before predicting the structures using RNAfold/RNAplfold as described earlier. For shuffling we used the program shuffle from Sean Eddy's squid package.

The R-score: a novel T-Coffee scoring scheme

The original T-Coffee algorithm was modified in order to incorporate structural information within the library compilation process. This novel evaluation procedure is named the R-score and gives its name to R-Coffee, with the letter R standing for RNA. The R-score requires the secondary structures of the considered sequences to be pre-computed and it also involves two modifications of the original T-Coffee algorithm: one when compiling the pairwise alignment library and the other when evaluating the score for aligning two residues.

The new library compilation procedure involves extending the original T-Coffee library with any residue pair not observed within the pairwise alignments but whose relevance is suggested by the secondary structure predictions (Figure 1). For instance, let $A \sim X$ be two nucleotides predicted to form a secondary structure in sequence 1 and $B \sim Y$ two other paired nucleotides in sequence 2.

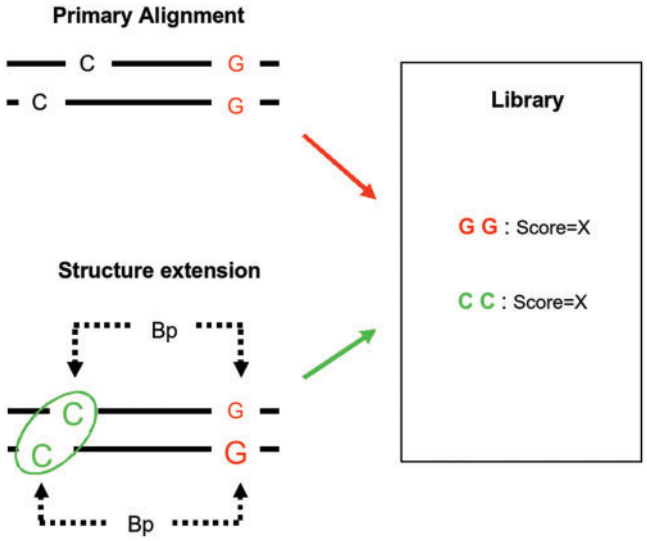

Figure 1. R-Coffee's RNA-extension. The two Gs correspond to a pair of matched residues observed in the input pairwise alignment. This gets incorporated in the library as a constraint. Both of these nucleotides are predicted to be base paired (Bp) with two Cs that have not been found aligned. The RNA extension involves incorporating the associated constraint (C matched to C), based on the information contained in the provided structures. This structure-based extension is one of the two main ingredients of the R-Coffee scoring scheme.

In the standard T-Coffee procedure, if the primary alignment of sequences 1 and 2 contains the aligned pair A–B, this pair will be added as an entry to the library and associated with a weight equal the average identity of the alignment of sequences 1 and 2 (the weights will be added if several alternative alignments contribute the same pair). The R-Coffee library procedure goes further and involves incorporating the pair X–Y into the library (with the weight of A–B) even it was not aligned in any of the input library alignments. The rationale is that if the alignment of A–B is correct and if the predicted structures are correct as well, then the pair X–Y should be aligned and it therefore makes sense to add it to the library (if X–Y is already part of the primary library, its weight is increased by the A–B weight). Whenever more than one structure has been provided for each sequence, the secondary structure library may be ambiguous and provides several alternative base pairings for one or both residues (e.g. A ~ X, A ~ W in one sequence and B ~ Y, B ~ Z in the other). In this case, the update will consider a combination of all the potential structure-induced aligned pairs (i.e. X–Y, X–Z, W–Y, W–Z) and increase their primary weight with that of A–B.

The R-score also uses a new evaluation procedure. The regular T-Coffee scoring scheme computes the matching score of a given residue pair A–B by summing up over the score of all the residue triplets including A, B and a third residue x from a third sequence. This can be formalized as follows:

$$\mathrm{Tscore}(A,B) = \sum_{x} \mathrm{MIN}(\mathrm{Weight}(A,x),\mathrm{Weight}(B,x)) \qquad 1$$

where Weight(A,x) is a primary weight and x is any residue reported aligned both to A and B within the primary library. The R-score of that same pair is then defined as:

$$\mathrm{Rscore}(A,B|A\sim X,B\sim Y) = \mathrm{MAX}\,(\mathrm{Tscore}(A,B),\mathrm{Tscore}(X,Y)) \qquad 2$$

where X pairs with A and Y with B as indicated by the structural library (A ~ X, B ~ Y). Whenever the structural library is ambiguous (i.e. A ~ X, A ~ W, B ~ Y, B ~ W), the final score is estimated by considering all the resulting combinations:

$$\mathrm{Rscore}(A,B|A\sim X,A\sim Z,B\sim Y,B\sim W) = \mathrm{MAX}(\mathrm{Tscore}(A,B),\mathrm{Tscore}(X,Y), \mathrm{Tscore}(X,W),\mathrm{Tscore}(Z,Y),\mathrm{Tscore}(Z,W)) \qquad 3$$

The R-score, like the regular T-Coffee scoring scheme is then used as a position-specific substitution matrix while computing an alignment. R-Coffee uses the progressive alignment strategy described in the original T-Coffee algorithm and inspired from the ClustalW implementation. Sequences are all aligned two by two, using a standard identity based matrix and the Myers and Miller implementation of dynamic programming. These alignments are then used to derive a distance matrix that is turned into a Neighbor-Joining tree (39). This tree is used as a guide tree to define the order in which the sequences are aligned to create the multiple alignment. These alignments are made using the R-score as a position-specific scoring scheme and the Myers and Miller pairwise algorithm. Apart from the use of the Myers and Miller pairwise alignment option, all the other T-Coffee parameters have been left to their original default values.

Availability

R-Coffee is part of the T-Coffee package, an open source freeware distributed under the GPL license and available at no cost along with documentation from www.tcoffee.org. R-Coffee can be compiled on most platforms (UNIX, Mac OS X and Windows).

RESULTS AND DISCUSSION

R-Coffee is an RNA multiple sequence alignment method able to use RNA secondary structure information while computing a multiple sequence alignment. One of the key properties of R-Coffee is its low computational complexity. Given predicted structures, R-Coffee can compute structure-based sequence alignments in a time and space complexity similar to that reported for T-Coffee or Probcons [in the order of $O(N^3L^2)$ for N sequences of length L]. Nonetheless, the computation of the predicted structures can be a limiting factor. For example, for global Minimal Folding Energy methods like RNAfold (29), can be quite demanding with growing sequence length and prediction quality depends on sequence length (36,37). Our first concern was therefore to check whether the replacement of RNAfold with the less-demanding RNAplfold (38) could prove useful. RNAplfold is able to predict the fold of long sequences thanks to its local

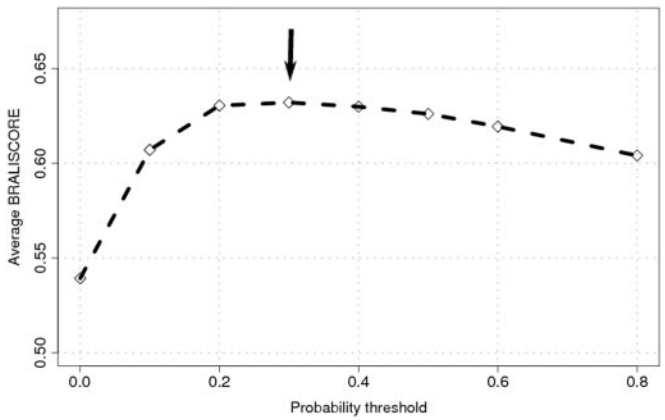

Figure 2. R-Coffee/RNAplfold base pair probability threshold optimization. Base pairs predicted by RNAplfold above a certain probability threshold were used as input for R-Coffee. Then all BRAliBase sets were aligned and the average alignment accuracy (BRAliscore) calculated. The optimal threshold was determined to be 0.3.

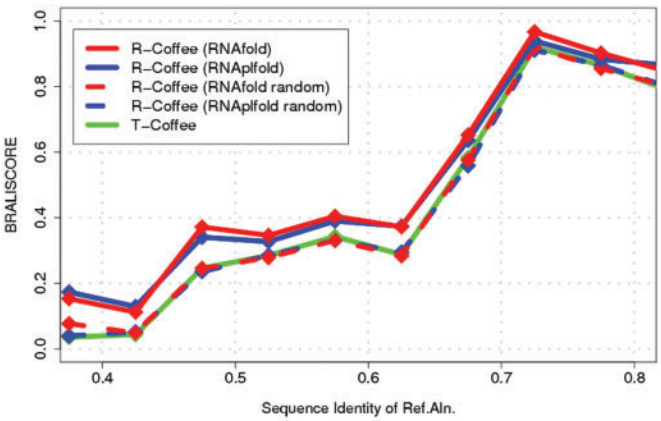

Figure 3. Effect of the RNA-extension on T-Coffee's performance on BRAliBase 2.0. The plot shows the alignment accuracy as function of the sequence identity. Scores are averaged over 5% sequence identity bins. Standard T-Coffee is compared to R-Coffee using structure input from RNAfold and RNAplfold as well as random structures.

structure prediction algorithm. In practice, this result is achieved by restricting the computation to the local neighborhood of each nucleotide (default is a span of 100 nt).

RNAplfold outputs base-pairing probabilities rather than a single secondary structure. We therefore determined an optimal threshold for filtering out unreliable base pairs. We did so by extensive testing on the BRAliBase dataset (see 'Material and methods' section and Figure 2). The cutoff value thus determined (0.3) was used throughout the remaining experiments. Given this cutoff value, we systematically compared the BRAliscore obtained by R-Coffee when using RNAfold and RNAplfold structural libraries. Both structural libraries (RNAfold and RNAplfold) give similar results. Interestingly, this correlation is high regardless of whether the considered sequences are closely or distantly related (Figure 3). The mean BRAliscore for the two methods is the same (0.64) with 53% of the 388 BRAliBase datasets having their BRAliscore within 5% of each other when using the RNAfold or the RNAplfold structural library. We therefore decided to use RNAplfold as the default provider of secondary structure predictions for the rest of the analysis. This allows R-Coffee to deal with sequences of arbitrary size. In order to check the effect of the accuracy of the predicted structures, we also tested R-Coffee using random structure predictions, as input. The performance then returns to the default T-Coffee accuracy (Figure 3), i.e. alignment quality does not get worse as compared to default T-Coffee, but clearly decreases when compared with genuine structure predictions. Altogether these results suggest that it makes little difference in accuracy whether we use RNAfold or RNAplfold for secondary structure prediction in R-Coffee. They also confirm the effectiveness of the incorporation of structural information within the alignment procedure. We wish to note here, that although we limited our analysis to these two approaches, the flexibility of R-Coffee's RNA extension allows incorporating and combining any kind of structure prediction. Alternatives include using RNAfold's partition function and an applied threshold (as done with RNAplfold here) or using methods with a higher selectivity like Contrafold (40). But one could also include, for example sub-optimal structure or pseudoknot predictions (41).

Next, we examined the merits of R-Coffee in comparison with other methods. It should be stressed here that our primary goal was not to produce a stand-alone method, but rather to use R-Coffee as a novel component that can seamlessly be combined with any existing RNA alignment method. We therefore focused our efforts on the evaluation of the combination between R-Coffee and other established methods. In order to determine the baseline of our analysis we ran common sequence alignment methods on the 388 BRAliBase datasets (top part of Table 2). Our results are relatively consistent with previous reports (42,43) of accuracy on protein sequence alignments: Mafft (35), Probcons (34) and Muscle (32) deliver the best alignments. The default T-Coffee is notably inaccurate with RNA (5), most likely because it uses, by default, a banded dynamic programming heavily tuned on protein sequences. The second part of Table 2 (structural aligners) is also consistent with previous reports and confirms that RNA alignment methods making use of structural information have a higher accuracy than sequence aligners. Our results show that FoldalignM (15), Rnasampler (44), T-Lara (19) and Murlet (14) clearly outperform all the regular sequence alignment methods, with more than five points difference between the best structure-based alignment methods (FoldalignM/Rnasampler) and their best non-structure-based counterpart (Mafft ginsi).

The most straightforward way to embed these methods within R/T-Coffee is to use each individual method to generate libraries of pairwise alignments. This protocol merely requires a pairwise alignment for each pair of sequences within a dataset and using the resulting

Table 2. BRAliBase evaluations

Method	BRAliscore			Net improvement	
	Default	+T-Coffee	+R-Coffee	+T-Coffee	+R-Coffee
T-Coffee	0.59	/	0.63	/	125
Poa	0.62	0.65	0.70	48	154
Pcma	0.62	0.64	0.67	34	120
Prrn	0.64	0.61	0.66	−63	45
ClustalW	0.65	0.65	0.69	−7	83
Proalign	0.66	0.68	0.71	30	128
Mafft fftns	0.68	0.68	0.72	17	68
Probcons	0.69	0.67	0.71	−74	51
Muscle	0.69	0.69	0.73	−17	42
Mafft ginsi	0.70	0.68	0.72	−49	39
M-Coffee4	0.71	/	0.74	/	84
M-Locarna	0.66	0.69	0.71	101	133
Stral	0.71	0.70	0.72	−4	19
Murlet	0.73	0.70	0.72	−132	−73
Rnasampler	0.75	0.70	0.71	−101	−95
FoldalignM	0.75	0.76	0.76	72	76
Dynalign	/	0.62	0.62	/	/
Foldalign	/	0.62	0.77	/	/
T-Lara	/	0.74	0.73	/	/
Consan	/	0.79	0.79	/	/

Each line in the table corresponds to the evaluation of the package listed in the Method column. The BRAliscore section indicates the average BRAliscore performance of the package. The default column indicates the score obtained by the considered package. The +T-Coffee indicates the average BRAliscore using the corresponding package combined with T-Coffee. The +R-Coffee column indicates the average BRAliscore of the same package combined with R-Coffee. The slash / indicates values that could not be computed, either because the method only produces pairwise alignments (Dynalign, Foldalign and Consan), or because the method is a derivative of or uses T-Coffee (e.g. T-Lara). The Net Improvement section indicates the net improvement over the stand-alone methods.

alignments as a primary library for either T-Coffee or R-Coffee. The structural libraries were computed once on the entire dataset and then re-used. This protocol was used on all the aligners with the exception of T-Lara for which we followed the combination protocol described by T-Lara's authors. It involves compiling partial T-Coffee libraries with Lara (i.e. libraries restricted to aligned stems) and combining them with the default T-Coffee libraries made of global and local pairwise alignments, that same protocol was used when combining Lara with R-Coffee.

We first evaluated the effect of using the regular T-Coffee to compute an MSA with pairwise libraries generated either with regular sequence or structural aligners. The results are displayed in the +T-Coffee column of Table 2. For each T-Coffee/method ***X*** combination (***X*** being any of the tested methods), we calculated the average BRAliScore and the Net Improvement (NI), which is the absolute improvement induced by combining that method with T-Coffee. It is defined as the number of test cases where a method ***X*** outperforms that method combined with T-Coffee (T-Coffee/***X***) minus the number of times the T-Coffee/***X*** combination outperforms method ***X***:

$$\mathrm{NI}=\left[\frac{\mathbf{T}\text{-Coffee}}{\mathbf{X}\text{outperforms}\mathbf{X}}\right]-\left[\frac{\mathbf{X}\text{outperforms}\mathbf{T}\text{-Coffee}}{\mathbf{X}}\right] \quad \mathbf{4}$$

The NI provides a guide as to whether one of the methods outperforms another. Results in Table 2 are easier to interpret when the regular sequence aligners and the structural aligners are separately considered. The regular aligners show little benefit from the T-Coffee combination of their pairwise output (Column +T-Coffee), probably because these methods already make an efficient use of their sequence information, or at least because they use it as efficiently as T-Coffee could. It is not a surprising result since most of these methods either use a T-Coffee inspired consistency-based scoring scheme (Mafft g/linsi, Probcons) or a sophisticated iterative method (Muscle, Prrn) to improve the original progressive MSA. R-Coffee, on the other hand, provides a clear improvement to all the regular sequence alignment methods tested here (Table 2, +R-Coffee column). This improvement remains regardless of the metrics used (BRAliscore or Net Improvement).

The results obtained when combining R/T-Coffee with structural aligners follow a similar albeit less marked pattern. When added on the top of structural aligners, T-Coffee improves two methods out of five and R-Coffee improves three out of five. These observations are fairly consistent with the underlying principles of the alignment programs (sequence and structural aligners). They suggest that the potential benefits of using R-Coffee come as much from the T-Coffee consistency-based scoring scheme as they do from the R-extension. The relatively small benefit coming from the R-extension in this case also makes sense if one considers that the structural aligners already use structural information and are therefore less likely to benefit from the incorporation of RNAplfold predictions than their sequence-based counterparts. This is especially true when combining T-Coffee with Consan. It is worth mentioning, however, that the use of the R-scoring scheme outperforms similar T-Coffee combinations in most cases with five methods out of nine being improved when switching from the T-Coffee to the R-Coffee combination and four methods remaining unchanged.

Altogether, the data collected in Table 2 strongly suggest that consistency-based scoring schemes provide an efficient framework for making the best out of pairwise alignment methods. T/R-Coffee/Foldalign and T/R-Coffee/Consan provide the best illustration of this concept (bottom of Table 2). Consan is computationally too expensive to be easily extended to MSAs, yet, a straightforward combination with R-Coffee results in a method that outperforms all the other methods analyzed in this work (Tables 2 and 3). Figure 4 shows a detailed performance plot on BRAliBase and compares R-Coffee/Consan with the best sequence alignment method (Mafft ginsi) and FoldalignM. This plot shows, that R-Coffee/Consan performs better than FoldalignM across the full range of sequence identities, even if the difference is not statistically significant (Table 3). It is important to point out that the shape of this curve is a side effect of the two components that comprise BRAliScore (SCI, the structural component and SPS the sequence one). High levels of sequence identity naturally result in high-scoring alignments. At the other side of the spectrum at low identity levels, numerous compensating base pair

Table 3. Net Improvement of R-Coffee/Consan and RM-Coffee4 over programs on BRAliBase

Method	versus R-Coffee-Consan	versus RM-Coffee4
Poa	241***	217***
T-Coffee	241***	199***
Prrn	232***	198***
Pcma	218***	151***
Proalign	216***	150**
Mafft fftns	206***	148*
ClustalW	203***	136***
Probcons	192***	128*
Mafft ginsi	170***	115
Muscle	169***	111
M-Locarna	234***	183**
Stral	169***	62
FoldalignM	146	61
Murlet	130*	−12
Rnasampler	129*	−27
T-Lara	125*	−30

This table indicates the relative performance of the methods listed in the Method column in comparison with the R-Coffee/Consan and RM-Coffee4 combinations, as net improvement. Asterisks indicate statistically significant differences according to Wilcoxon tests ($^{*}P \leq = 0.05$; $^{**}P \leq = 0.01$; $^{***}P \leq = 0.001$). The upper part of the table contains sequence aligners only, the lower part structural alignment programs. Within these sections, programs are sorted by net improvement.

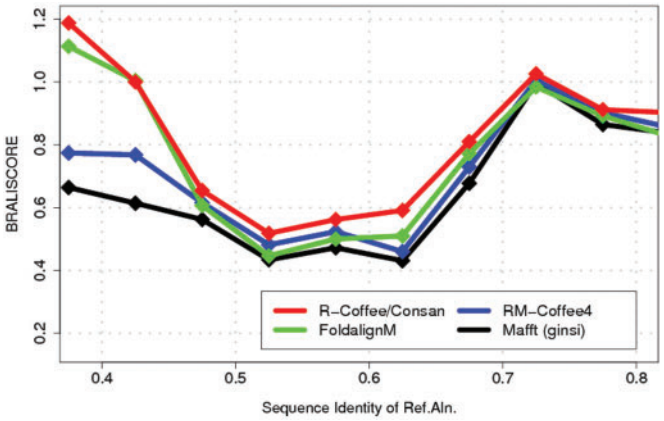

Figure 4. Comparison of R-Coffee/Consan and RM-Coffee with other programs. The plot shows the alignment accuracy on BRAliBase 2.0 as function of the sequence identity. Scores are averaged over 5% sequence identity bins. We included the best stand-alone sequence aligner (MAFFT ginsi), one of the two best structural aligners (FoldalignM), the best R-Coffee combination (R-Coffee/Consan) and RM-Coffee4 that combines the pairwise alignments of Probcons, MAFFT ginsi/fftns and Muscle by means of R-Coffee.

mutations can result in high scores, because they are taken into account by the SCI (see also Reference alignments and Evaluation). Nonetheless, and across the whole identity spectrum, our data supports well the idea that R-Coffee/Consan is probably the most accurate RNA MSA alignment method currently available for the kind of datasets found in BRAliBase (i.e. less than 150 nt).

We next assessed whether R-Coffee is also useful for aligning long sequences. We analyzed the Cmfinder dataset made of Rfam alignments embedded within surrounding genomic sequences of varying lengths. None of the structural aligners except M-Locarna (17), was able to run on all the 11 datasets and the analysis was restricted to regular sequence aligners (Table 4). With the notable exception of Muscle (32), the ranking in this table is not dramatically different from that in Table 2. The behavior of these methods when combined with T- or R-Coffee is also similar. When considering the 10 sequence aligners with T-Coffee, we observed an improvement on 7 methods out of 10. This figure rises to 9 out of 10 when making the combination with R-Coffee. Although these results are based on too small a dataset (11 alignments) to be considered statically significant, they are in very good agreement with those reported on BRAliBase in Table 2 and confirm R-Coffee's ability to improve over most sequence alignment methods.

Table 4. Cmfinder data set comparison

Method	SPS			Net improvement	
	Default	+T-Coffee	+R-Coffee	+T-Coffee	+R-Coffee
ClustalW	0.54	0.57	0.58	5	5
Mafft ginsi	0.64	0.64	0.64	−1	2
Mafft fftns	0.60	0.64	0.64	6	6
Muscle	0.32	0.40	0.42	4	8
Pcma	0.49	0.55	0.58	8	8
Poa	0.31	0.38	0.42	4	8
Proalign	0.40	0.39	0.41	−4	−2
Probcons	0.50	0.45	0.51	−3	2
Prrn	0.43	0.54	0.56	3	4
M-Locarnap	0.53	0.63	0.63	6	5
T-Coffee	0.54	/	0.53	/	2
R/M-Coffee4	/	0.63	0.65	/	0

Each line in the table corresponds to the evaluation of the package listed in the Method column. The SPS section indicates the averaged sum-of-pairs scores (applied to the Rfam core alignment) measured on the considered package; +T-Coffee is the same score measured on the package combined with T-Coffee (+T-Coffee); the +R-Coffee column corresponds to that same package combined with R-Coffee. The slash / indicates values that could not be computed because the method is a derivative of T-Coffee (T-Coffee and M-Coffee). The Net Improvement section indicates the net improvement for similar combinations.

The main practical problem with using R-Coffee is that to reach its highest level of accuracy, it requires the installation of RNA alignment packages, which may be extremely greedy with memory and CPU usage. We therefore checked whether a simpler alternative could be better suited for more modest computational configurations, or for high throughput applications. In a previous paper, Wallace *et al.* reported and characterized a novel mode of T-Coffee named M-Coffee (22). M-Coffee is a meta-aligner that combines alternative multiple sequence alignment methods into one consensus alignment. This combination usually results in an improvement over the constituting methods. We used the M-Coffee approach to combine the four best regular alignment methods (i.e. non-structure based), and tested them on BRAliBase. Following the strategy outlined in the original M-Coffee paper, we incorporated the sequence aligners in order of decreasing performances and kept the combination with the highest average. This protocol resulted in RM-Coffee4, a combination of Muscle, Probcons, Mafft ginsi and Mafft fftns fed to T-Coffee (M-Coffee4) or R-Coffee

(RM-Coffee4). The results (Table 2 and Table 3, Figure 4) are unambiguous and indicate that RM-Coffee4 clearly outperforms all the sequence alignment methods while delivering the best BRAliBase alignments one may obtain without using a structural aligner. These results were not confirmed on the 11 Cmfinder datasets (Table 4), either because this dataset is too small to reveal the trend or because of the negative effect of Muscle on RM-Coffee4 on this specific dataset.

CONCLUSION

We have presented a modified version of the T-Coffee (21) multiple sequence alignment method, named R-Coffee, designed for delivering highly accurate multiple ncRNA alignments. R-Coffee is a heuristic, able to take advantage of secondary structure predictions carried out beforehand. It is best described as an alignment improver and we show in this work that it can effectively improve all sequence alignment packages, taken off the shelf and without tuning. Among all the combinations tested here, one clearly outperformed the alternatives: the combination of R-Coffee and Consan (10). Most of these tests were carried out on the BRAliBase reference datasets (5). We also checked whether R-Coffee was able to deal with datasets of longer sequences, combining a mixture of related and unrelated segments. For that purpose, we used a dataset designed for the Cmfinder algorithm (26). We found that the R-Coffee combination improved, to a greater or lesser extent, all the tested alignment methods. The combined observations made on the BRAliBase and Cmfinder datasets suggest that the R-Coffee scoring scheme is able to make effective use of RNA predicted secondary structures in order to improve accuracy over most regular sequence aligners.

This strategy also works when applied to structural aligners, although less dramatically than when considering regular sequence aligners. These results confirm the strength of consistency-based scoring schemes over regular alignment methods. They suggest that most pairwise alignment methods can usefully be incorporated in a consistency-based framework such as T-Coffee. Our results also indicate that the meta-method approach originally described for M-Coffee (22) can be applied to R-Coffee, and that whenever the computation of highly accurate structure-based RNA pairwise alignments is not feasible, one may obtain alignments of reasonable quality by combining purely sequence-based alignments via R-Coffee. Further progress will also require the assembly of more demanding reference datasets, especially for long sequences. Such datasets are hard to assemble because RNA structural information is scarce (compared to protein structure information).

RNA alignment remains a rapidly developing field. With an increasing number of novel biological functions associated with yet poorly characterized RNA genes, there is an ever growing need for methods allowing accurate comparison of RNA sequences and the identification of distant homologues. Any improvement in alignment accuracy is likely to have a big impact. In this context, R-Coffee can easily be further improved. The flexible way in which secondary structures are fed to the program allows a seamless combination of data from heterogeneous sources. It is important to point out that all the possibilities supported by the current software implementation have not yet been explored. Most notably, we have not yet fully exploited the possibility to associate more than one predicted structure to each sequence. These alternative structures could either be suboptimal structures, or the output of alternative structure prediction programs, such as ContraFold or Rfold. One could also combine structure predictions of any kind, including local, global or even tertiary interactions like pseudoknots, with experimentally verified structures. The possibility of combining data from various sources is, perhaps, the major strength of R-Coffee.

ACKNOWLEDGEMENTS

We thank Iain M. Wallace for useful discussions and all authors for their assistance with using their programs. This work was partly supported by funding from the Science Foundation Ireland. C.N. thanks the centre for genomic regulation for support and funding. Funding to pay the Open Access publication charges for this article was provided by Centro de Regulacio Genomica (CRG).

Conflict of interest statement. None declared.

REFERENCES

1. Zamore,P.D. and Haley,B. (2005) Ribo-gnome: The Big World of Small RNAs. *Science*, **309**, 1519–1524.
2. Costa,F.F. (2007) Non-coding RNAs: lost in translation? *Gene*, **386**, 1–10.
3. Birney,E., Stamatoyannopoulos,J.A., Dutta,A., Guigo,R., Gingeras,T.R., Margulies,E.H., Weng,Z., Snyder,M., Dermitzakis,E.T., Thurman,R.E. *et al.* (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, **447**, 799–816.
4. Levenshtein,V.I. (1966) Binary codes capable of correcting deletions, insertions, and reversals. *Cybern. Control Theory*, **10**, 707–710.
5. Gardner,P.P., Wilm,A. and Washietl,S. (2005) A benchmark of multiple sequence alignment programs upon structural RNAs. *Nucleic Acids Res.*, **33**, 2433–2439.
6. Wilm,A., Mainz,I. and Steger,G. (2006) An enhanced RNA alignment benchmark for sequence alignment programs. *Algorithms Mol. Biol.*, **1**, [Epub ahead of print].
7. Thompson,J.D., Plewniak,F. and Poch,O. (1999) A comprehensive comparison of multiple sequence alignment programs. *Nucleic Acids Res.*, **27**, 2682–2690.
8. van Nimwegen,E., Crutchfield,J.P. and Huynen,M. (1999) Neutral evolution of mutational robustness. *Proc. Natl Acad. Sci. USA*, **96**, 9716–9720.
9. Sankoff,D. (1985) Simultaneous solution of the RNA folding, alignment and protosequence problems. *SIAM J1 Appl1 Math.*, **45**, 810–825.
10. Dowell,R. and Eddy,S. (2006) Efficient pairwise RNA structure prediction and alignment using sequence alignment constraints. *BMC Bioinformatics*, **7**, 400.
11. Hofacker,I.L., Bernhart,S.H.F. and Stadler,P.F. (2004) Alignment of RNA base pairing probability matrices. *Bioinformatics*, **20**, 2222–2227.
12. Mathews,D.H. (2005) Predicting a set of minimal free energy RNA secondary structures common to two sequences. *Bioinformatics*, **21**, 2246–2253.

13. Havgaard,J.H., Lyngso,R.B., Stormo,G.D. and Gorodkin,J. (2005) Pairwise local structural alignment of RNA sequences with sequence similarity less than 40%. *Bioinformatics*, **21**, 1815–1824.
14. Kiryu,H., Tabei,Y., Kin,T. and Asai,K. (2007) Murlet: a practical multiple alignment tool for structural RNA sequences. *Bioinformatics*, **23**, 1588–1598.
15. Torarinsson,E., Havgaard,J.H. and Gorodkin,J. (2007) Multiple structural alignment and clustering of RNA sequences. *Bioinformatics*, **23**, 926–932.
16. Holmes,I. (2005) Accelerated probabilistic inference of RNA structure evolution. *BMC Bioinformatics*, **6**, 73.
17. Will,S., Reiche,K., Hofacker,I.L., Stadler,P.F. and Backofen,R. (2007) Inferring noncoding RNA families and classes by means of genome-scale structure-based clustering. *PLoS Comput. Biol.*, **3**, e65.
18. Meyer,I.M. and Miklos,I. (2007) SimulFold: simultaneously inferring RNA structures including pseudoknots, alignments, and trees using a Bayesian MCMC framework. *PLoS Comput. Biol.*, **3**, e149.
19. Bauer,M., Klau,G.W. and Reinert,K. (2007) Accurate multiple sequence-structure alignment of RNA sequences using combinatorial optimization. *BMC Bioinformatics*, **8**, 271.
20. Wuyts,J., Perriere,G. and Van de Peer,Y. (2004) The European ribosomal RNA database. *Nucleic Acids Res.*, **32**, D101–D103.
21. Notredame,C., Higgins,D.G. and Heringa,J. (2000) T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.*, **302**, 205–217.
22. Wallace,I.M., O'Sullivan,O., Higgins,D.G. and Notredame,C. (2006) M-Coffee: combining multiple sequence alignment methods with T-Coffee. *Nucleic Acids Res.*, **34**, 1692–1699.
23. O'Sullivan,O., Suhre,K., Abergel,C., Higgins,D.G. and Notredame,C. (2004) 3DCoffee: combining protein sequences and structures within multiple sequence alignments. *J. Mol. Biol.*, **340**, 385–395.
24. Armougom,F., Moretti,S., Poirot,O., Audic,S., Dumas,P., Schaeli,B., Keduas,V. and Notredame,C. (2006) Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucleic Acids Res.*, **34**, W604–W608.
25. Siebert,S. and Backofen,R. (2005) MARNA: multiple alignment and consensus structure prediction of RNAs based on sequence structure comparisons. *Bioinformatics*, **21**, 3352–3359.
26. Yao,Z., Weinberg,Z. and Ruzzo,W.L. (2006) CMfinder-a covariance model based RNA motif finding algorithm. *Bioinformatics*, **22**, 445–452.
27. Griffiths-Jones,S., Moxon,S., Marshall,M., Khanna,A., Eddy,S.R. and Bateman,A. (2005) Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Res.*, **33**, D121–D124.
28. Washietl,S., Hofacker,I.L. and Stadler,P.F. (2005) Fast and reliable prediction of noncoding RNAs. *Proc. Natl Acad. Sci. USA*, **102**, 2454–2459.
29. Hofacker,I.L. (2003) Vienna RNA secondary structure server. *Nucleic Acids Res.*, **31**, 3429–3431.
30. Hofacker,I.L., Fekete,M. and Stadler,P.F. (2002) Secondary structure prediction for aligned RNA sequences. *J. Mol. Biol.*, **319**, 1059–1066.
31. Cannone,J.J., Subramanian,S., Schnare,M.N., Collett,J.R., D'Souza,L.M., Du,Y., Feng,B., Lin,N., Madabusi,L.V., Müller,K.M. *et al.* (2002) The Comparative RNA Web (CRW) Site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics*, **3**, 2.
32. Edgar,R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.*, **32**, 1792–1797.
33. Thompson,J.D., Higgins,D.G. and Gibson,T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**, 4673–4680.
34. Do,C.B., Mahabhashyam,M.S.P., Brudno,M. and Batzoglou,S. (2005) ProbCons: probabilistic consistency-based multiple sequence alignment. *Genome Res.*, **15**, 330–340.
35. Katoh,K., Kuma,K.-i., Toh,H. and Miyata,T. (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.*, **33**, 511–518.
36. Doshi,K., Cannone,J., Cobaugh,C. and Gutell,R. (2004) Evaluation of the suitability of free-energy minimization using nearest-neighbor energy parameters for RNA secondary structure prediction. *BMC Bioinformatics*, **5**, 105.
37. Dowell,R. and Eddy,S.R. (2004) Evaluation of several lightweight stochastic context-free grammars for RNA secondary structure prediction. *BMC Bioinformatics*, **5**, 71.
38. Bernhart,S.H., Hofacker,I.L. and Stadler,P.F. (2006) Local RNA base pairing probabilities in large sequences. *Bioinformatics*, **22**, 614–615.
39. Saitou,N. and Nei,M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, **4**, 406–425.
40. Do,C.B., Woods,D.A. and Batzoglou,S. (2006) CONTRAfold: RNA secondary structure prediction without physics-based models. *Bioinformatics*, **22**, e90–e98.
41. Reeder,J., Hochsmann,M., Rehmsmeier,M., Voss,B. and Giegerich,R. (2006) Beyond Mfold: recent advances in RNA bioinformatics. *J. Biotechnol.*, **124**, 41–55.
42. Carroll,H., Beckstead,W., O'Connor,T., Ebbert,M., Clement,M., Snell,Q. and McClellan,D. (2007) DNA reference alignment benchmarks based on tertiary structure of encoded proteins. *Bioinformatics*, **23**, 2648–2649.
43. Edgar,R.C. and Batzoglou,S. (2006) Multiple sequence alignment. *Curr. Opin. Struct. Biol.*, **16**, 368–373.
44. Xu,X., Ji,Y. and Stormo,G.D. (2007) RNA sampler: a new sampling based algorithm for common RNA secondary structure prediction and structural alignment. *Bioinformatics*, **23**, 1883–1891.
45. Pei,J., Sadreyev,R. and Grishin,N.V. (2003) PCMA: fast and accurate multiple sequence alignment based on profile consistency. *Bioinformatics*, **19**, 427–428.
46. Lee,C., Grasso,C. and Sharlow,M.F. (2002) Multiple sequence alignment using partial order graphs. *Bioinformatics*, **18**, 452–464.
47. Löytynoja,A. and Milinkovitch,M.C. (2003) A hidden Markov model for progressive multiple alignment. *Bioinformatics*, **19**, 1505–1513.
48. Gotoh,O. (1996) Significant improvement in accuracy of multiple protein sequence alignments by iterative refinement as assessed by reference to structural alignments. *J. Mol. Biol*, **264**, 823–838.
49. Dalli,D., Wilm,A., Mainz,I. and Steger,G. (2006) STRAL: progressive alignment of non-coding RNA using base pairing probability vectors in quadratic time. *Bioinformatics*, **22**, 1593–1599.