

# IntaRNA 2.0: enhanced and customizable prediction of RNA–RNA interactions

Martin Mann<sup>1</sup>, Patrick R. Wright<sup>1</sup> and Rolf Backofen<sup>1,2,\*</sup>

<sup>1</sup>Bioinformatics, Computer Science, University of Freiburg, Georges-Koehler-Allee 106, 79110 Freiburg, Germany and

<sup>2</sup>Centre for Biological Signalling Studies (BIOSS), University of Freiburg, Schaezenlestr. 18, 79104 Freiburg, Germany

Received February 27, 2017; Revised March 31, 2017; Editorial Decision April 05, 2017; Accepted May 02, 2017

## ABSTRACT

The INTARNA algorithm enables fast and accurate prediction of RNA–RNA hybrids by incorporating seed constraints and interaction site accessibility. Here, we introduce INTARNAv2, which enables enhanced parameterization as well as fully customizable control over the prediction modes and output formats. Based on up to date benchmark data, the enhanced predictive quality is shown and further improvements due to more restrictive seed constraints are highlighted. The extended web interface provides visualizations of the new minimal energy profiles for RNA–RNA interactions. These allow a detailed investigation of interaction alternatives and can reveal potential interaction site multiplicity. INTARNAv2 is freely available (source and binary), and distributed via the conda package manager. Furthermore, it has been included into the Galaxy workflow framework and its already established web interface enables *ad hoc* usage.

## INTRODUCTION

The interaction of RNA molecules is an important factor for regulatory processes in all organisms. The *in silico* modeling and prediction of such RNA–RNA interactions is thus a central aspect in current research projects. In the last decades, various approaches to solve this problem have been proposed (see reviews (1,2)). To enable highly accurate predictions, state-of-the-art tools not only take the stability (energy) of possible RNA–RNA interactions into account but also consider the accessibility of the interacting subsequences (3). The latter is often incorporated by adding a pseudo-energy penalty proportional to the subsequences' probability to be involved in intramolecular structure formation (4). Prediction quality can be further improved by addressing additional features known from experimentally validated interactions. For instance, bacterial small RNAs (sRNA) and eukaryotic microRNAs form highly stable

subinteractions, which are considered to seed the formation of extended RNA–RNA duplexes (5,6).

INTARNA, first introduced in (7), is one of the most widely used state-of-the-art RNA–RNA interaction prediction approaches (2,3). It is employed in prokaryotic (8–10) and eukaryotic (11–13) systems, and its web server (14,15) received ~16 500 external jobs in 2016. Its fast heuristic prediction mode enables genome wide target prediction while its high prediction accuracy is based on the incorporation of interaction site accessibility and seed constraints.

However, INTARNA has become technically outdated for the following reasons. Firstly, INTARNA was fixed to an aged Vienna RNA library (v1.\*) (16) and its energy parameters Turner99 (17). Secondly, it featured a strongly restricted user interface and was hard to maintain, update or extend.

Here, we introduce INTARNAv2, a complete open-source reimplementation of the previous approach (from now on referred to as INTARNAv1) with a strong focus on efficiency, modularity and flexibility. The implementation closely follows the software development guidelines suggested in (18) and enables enhanced and customizable RNA–RNA interaction prediction.

INTARNAv2 enforces the seed interaction to be energetically favorable. Furthermore, a slightly changed treatment of dangling end energy contributions, compared to INTARNAv1 and other approaches, has been incorporated. This, in concert with the usage of recent RNA energy parameters (from the Vienna RNA package v2.\*) (19), enables an improvement of prediction quality. The newly performed benchmark uses a data set of 160 experimentally verified enterobacterial sRNA–mRNA interactions and evaluates genome wide screens for each interaction pair.

The extended INTARNAv2 web interface within the *Freiburg RNA tools* [Freiburg RNA tools: <http://rna.informatik.uni-freiburg.de/IntaRNA>.] v4.4.2 (14,15) enables full control over energy parameters, seed constraints and accessibility computation. The user has to provide either lists of interacting RNAs (all versus all predictions) or can perform a genome wide target screen (providing the according NCBI accession number) for a given input RNA. Using an efficient dynamic programming scheme, the minimum free energy interactions (or suboptimals if requested)

\*To whom correspondence should be addressed. Tel: +49 761 2037461; Fax: +49 761 2037462; Email: backofen@informatik.uni-freiburg.de

are listed and additional details can be retrieved for each predicted interaction pair. Furthermore, the new version provides minimal energy profiles for interacting RNA pairs that enable a detailed investigation of interaction alternatives.

## METHODS

INTARNAv2 reimplements and extends the original INTARNA approach (7) and closely follows established rules for software development in systems biology, e.g. see (18).

Besides the heuristic prediction mode of INTARNAv1 that allows fast, genome-wide predictions (7), INTARNAv2 also features exact interaction prediction by reimplementing and extending the RNAUP approach (4). Since seed constraints are optional for all prediction modes, INTARNAv2 can emulate RNAUP predictions (by disabling the seed requirement). Furthermore, the user interface enables full control of accessibility usage within RNA–RNA interaction prediction. Per default, unpaired probabilities are computed using the Vienna RNA package v2.\* routines. Here, both global (RNAUP-like) or local unpaired probabilities (as computed by RNAPL FOLD (20)) can be selected. INTARNAv2 also supports the loading of pre-computed accessibility data. This can significantly reduce the run time since the computation of the accessibility is highly time consuming (7). This feature is especially useful when screening a large target sequence (21). By disabling the integration of accessibility information, INTARNAv2 can be used to emulate predictions from TARGETSCAN (22) or RNAHYBRID (23).

**Seed stability constraints** are a new feature of INTARNAv2, which enable direct control over the stability of the seed interaction. To this end, an upper energy bound can be provided. This defaults to 0 kcal/mol and thus requires all seed interactions to be energetically favorable. Furthermore, the user can optionally provide a lower bound on the unpaired probability of the subsequences forming the seed to define whether highly accessible regions are to be favored.

**Dangling end contributions** incorporate a stabilizing effect due to the stacking of unpaired nucleotides directly neighboring base pairs (17). Since the interaction model of INTARNA incorporates the accessibility, i.e. ‘unpairedness’, of the involved subsequences, an explicit dangling end case treatment was implemented in INTARNAv1, i.e. dangling end contributions were only incorporated if the according nucleotide was also part of the accessible subsequence considered for interaction formation (see Supplementary material for further details).

INTARNAv2 simplifies the dangling end treatment for interactions. Here, the conditional probability of the dangling end to be free when the interaction is formed is used to weight the according dangling end contributions (see Supplementary material for details). These conditional probabilities can be directly computed from the respective accessibility terms already available for the interaction scoring.

**Minimal energy profiles**, either position wise for each RNA (4) or intermolecular index pairwise for an RNA pair, can provide deeper insights into RNA–RNA interaction alternatives, since as for single RNA structure prediction, the optimal interaction is not necessarily the only or bio-

logically correct option. With INTARNAv2 both types of profiles can be generated by linking their generation with the respective dynamic programming computations. That is, whenever an interaction site (a cell in the according table) is computed, the respective minimal energy profiles are updated. This feature comes with a linear/quadratic overhead in computation time for single/pairwise profiles, respectively, since all indices within the according subsequences have to be updated. For instance, the run time of the heuristic prediction mode of  $O(n^2)$  is increased to  $O(n^3)$  for single sequence profiles and  $O(n^4)$  for pairwise minimal energy profiles. While computationally demanding, the investigation of these profiles provides meaningful additional information about possible interaction patterns, as discussed in the Results section.

**Suboptimal interaction enumeration** also enables the investigation of interaction alternatives. In INTARNAv1, only query interactions were allowed to overlap, since interaction alternatives in the target were of interest. INTARNAv2 now enables explicit control where suboptimal interactions are allowed to overlap. Besides the (so far limited) possibility to constrain overlaps to one sequence, this enables the enumeration of variants of the minimum free energy interaction (overlap in both) as well as the suppression of any overlap in both.

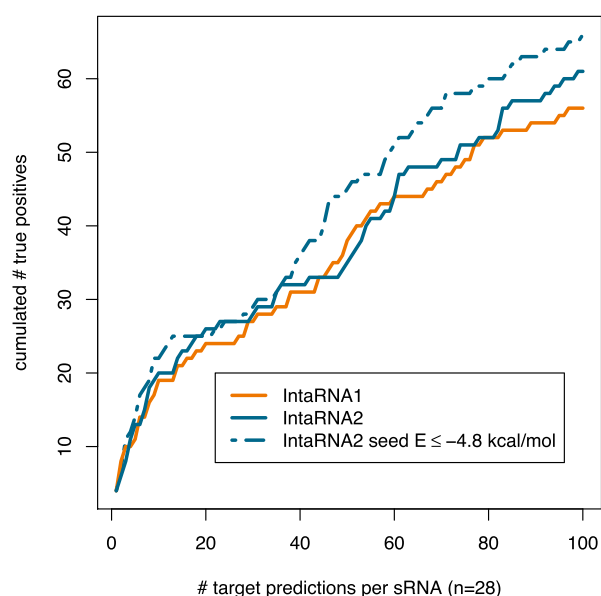
**Customizable output in CSV format** complements the normal or detailed ASCII chart output of the interaction details. The CSV output provides for the first time a machine-readable output format of all interaction details. To this end, the user can specify all details of interest (columns of the CSV table) that will be reported for each predicted interaction.

For further options of INTARNAv2 to constrain the interaction, seed site, accessibility computation, energy model, multi-threading, output formats, installation, docker/conda/galaxy packages, etc. we refer to its documentation [INTARNA at github: <https://github.com/BackofenLab/IntaRNA/>].

## RESULTS AND DISCUSSION

In order to evaluate the impact of recent energy parameters and the altered energy computation (see Methods), we performed a benchmark on genome-wide interaction predictions. In recent comprehensive benchmarks (2,24,25), INTARNA and RNAUP were shown to be the best performing tools that are not based on conservation. Thus, we limited our comparison to INTARNAv1 and INTARNAv2. To assess performance, we followed the procedure previously applied in (25). The dataset of verified RNA–RNA interactions was extended by several candidates that have been reported in recent literature (26–36). Together, this is a dataset of 160 verified sRNA–target interactions (see Supplementary Table S1).

For interaction prediction, we extracted 200 nucleotides upstream and 100 nucleotides downstream of the annotated start codon for each protein coding gene in *Escherichia coli* (NC.000913) and *Salmonella* (NC.003197). This gives rise to 4319 putative targets in *E. coli* and 4553 putative targets in *Salmonella* all of length 300 nucleotides with the start codon located at positions 201, 202 and 203. Tar-



**Figure 1.** The figure shows the whole genome target prediction performance of INTARNAv1 (orange) and INTARNAv2 (blue) on a benchmark set of 160 experimentally verified enterobacterial (*E. coli*, *Salmonella*) RNA–RNA interactions including 28 different sRNAs (see Supplementary Table S1 and Supplementary FASTA). Furthermore, the performance of INTARNAv2 was assessed with the same data set under enforcement of a seed energy  $E \leq -4.8$  kcal/mol (dashed blue). The x-axis represents the amount of predictions per sRNA at a given rank and the y-axis shows the cumulated number of true positives for all sRNA whole genome target predictions.

get predictions were performed for sRNAs (see Supplementary FASTA) from both species individually with INTARNAv1 and INTARNAv2 (parameters see Supplementary material). The resulting predictions were ranked by energy score with the lowest energy ranking on position 1.

A comparison is provided in Figure 1, which shows a slightly increased performance of INTARNAv2 compared to INTARNAv1 using comparable parameterization for both tools with equivalent run time (data not shown). The memory requirement of INTARNAv2 was drastically reduced by a factor of 6 (from an average of 79MB to 13MB).

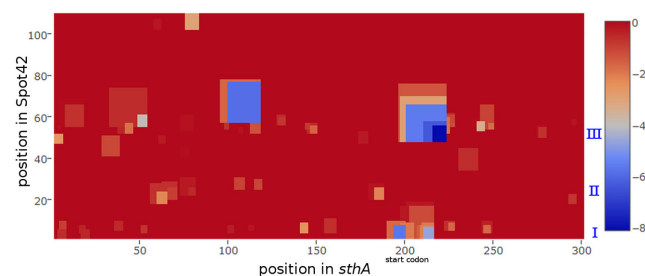
### Seed evaluation

To investigate the impact of more restrictive seed constraints, we first evaluated the energy distribution of seed interactions (without ED and dangling end contributions). To this end, we exhaustively generated all seeds (without bulges) for 5–8 bp, which revealed the seed energies to be normally distributed (see Table 1 for details). We observe that the mean value follows the formula  $(bp - 4) * -1.6$ , where  $-1.6$  is the average contribution of an additional stacking base pair for the Turner04 energy parameters.

For testing more tight constraints, we enforced the seed's stability to be above average ( $P$ -value: 0.5) and reran the benchmark. Since the number of seed base pairs was set to 7, only interactions with a seed that shows an energy below  $-4.8$  kcal/mol are reported. The results are shown in Figure 1 and reveal that a noticeable performance increase can

**Table 1.** Distribution of seed energies  $E$  for different seed lengths (= number of base pairs; no bulges allowed) with  $\mu\sigma$  = mean/standard deviation. The highlighted row corresponds to the seed constraints of the benchmark

| Seed length | min( $E$ )   | max( $E$ ) | $\mu(E)$    | $\sigma(E)$ |
|-------------|--------------|------------|-------------|-------------|
| 5           | −9.2         | 8.3        | −1.6        | 2.5         |
| 6           | −12.5        | 9.6        | −3.2        | 2.8         |
| <b>7</b>    | <b>−15.8</b> | <b>9.9</b> | <b>−4.8</b> | <b>3.1</b>  |
| 8           | −19.1        | 11.2       | −6.4        | 3.4         |



**Figure 2.** Minimal energy profile for all intermolecular index pairs covered by any predicted interaction of Spot42 with the *sthA* mRNA (with  $E < 0$ ). Conserved accessible regions I, II and III of Spot42 known to interact are tagged on the right.

be obtained. Furthermore, the run time is heavily reduced (halved) since less than half the number of sequence pairs that showed a stable interaction without constraint enable such a stable seed (only 55 227 of 121 988).

While this looks promising, such boundaries have to be chosen carefully. For instance, the prediction for the verified interaction of DsrA with the *rpoS* mRNA (37) is completely lost (no seed with  $E \leq -4.8$  kcal/mol) while it is top ranked (rank 1) for both INTARNAv1 as well as INTARNAv2 without seed constraint (see Supplementary Table S1). Thus, further research is needed on whether or not general suggestions for upper seed energy bounds can be provided or if such bounds are to be chosen in a sequence and organism dependent manner.

### Minimal energy profiles

The extended web interface of INTARNAv2 visualizes minimal energy profiles for interaction predictions that enable a sophisticated overview of interaction alternatives and their relative positioning. Figure 2 shows an example for the interaction of Spot42 with a region around the translation start site of the *sthA* (b3962) mRNA, which encodes a pyridine nucleotide transhydrogenase. Spot42 is known to interact with its targets via three accessible regions (I, II and III) (38), and the profile highlights sites I and III near the start codon. In fact, it has been shown that both regions I and III are important for regulation of the *sthA* mRNA (39), and this can also be deduced by investigating the profiles for the according Spot42 mutants (see Supplementary material).

Interestingly, while region II does not seem to be centrally involved in the interaction, its mutation still slightly reduces regulation by Spot42 (39). This might stem from a minor stability reduction of both sites I and III (predicted minimal energy increase of about 0.5 kcal/mol) when region II is mutated.



## CONCLUSIONS

The reimplementa-tion of INTARNA enables state-of-the-art energy parameters as well as seed constraint and accessibility incorporation for fast and accurate RNA–RNA interaction prediction. Increased prediction quality for genome-wide target predictions are also beneficial for comparative interaction prediction approaches like COPRANA (25,40).

Based on an extended benchmark set, we show the enhanced prediction quality of INTARNAv2 and highlight the possibility for further improvements, e.g. using more restrictive seed constraints. While promising, such constraints need careful tuning since they can lead to false negative predictions.

One of the new features of the INTARNA web server is the visualization of minimal energy profiles for interacting RNAs. This enables the detailed study of alternative RNA–RNA interactions. This way, it is possible to see whether multiple interaction sites are likely and if these can occur in conjunction or only exclusively. Furthermore, mutational effects can be analyzed and thus guide wet-lab experiments, which attempt in depth validation of predicted target sites (33,39,41).

The new flexible framework is the foundation for further upcoming extensions of INTARNA. One direction is to enable multi-site RNA–RNA interaction predictions. This way, hypotheses from the minimal energy profile investigations can be extended and further consolidated for successive experimental validation. Another direction is to enable further established input and output formats for an even more general embedding of INTARNA into high-throughput workflow systems like Galaxy (42).

In summary, the reimplementa-tion lays the groundwork for INTARNA to remain among the state-of-the-art RNA–RNA interaction prediction tools in future. Thus, it will continue to be useful and available for both bioinformaticians and wet-lab experimentalists.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

## ACKNOWLEDGEMENTS

We thank A.S. Richter, M. Uhl and S. Will for various ideas how to improve INTARNA and Z. Wanli for its integration into the Galaxy tool shed.

## FUNDING

German Research Foundation (DFG) [BA 2168/14-1 and BA 2168/16-1]. Funding for open access charge: Budget resources.

*Conflict of interest statement.* None declared.

## REFERENCES

- Backofen, R. and Hess, W.R. (2010) Computational prediction of sRNAs and their targets in bacteria. *RNA Biol.*, **7**, 33–42.
- Lai, D. and Meyer, I.M. (2016) A comprehensive comparison of general RNA–RNA interaction prediction methods. *Nucleic Acids Res.*, **44**, e61.
- Umu, S.U. and Gardner, P.P. (2016) A comprehensive benchmark of RNA–RNA interaction prediction tools for all domains of life. *Bioinformatics*, **32**, btw728.
- Mückstein, U., Tafer, H., Hackermüller, J., Bernhart, S.H., Stadler, P.F. and Hofacker, I.L. (2006) Thermodynamics of RNA–RNA binding. *Bioinformatics*, **22**, 1177–1182.
- Wagner, E. G.H. and Romby, P. (2015) Small RNAs in bacteria and archaea: who they are, what they do, and how they do it. *Adv. Genet.*, **90**, 133–208.
- Fabian, M.R., Sonenberg, N. and Filipowicz, W. (2010) Regulation of mRNA translation and stability by microRNAs. *Annu. Rev. Biochem.*, **79**, 351–379.
- Busch, A., Richter, A.S. and Backofen, R. (2008) IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. *Bioinformatics*, **24**, 2849–2856.
- Overloper, A., Kraus, A., Gurski, R., Wright, P.R., Georg, J., Hess, W.R. and Narberhaus, F. (2014) Two separate modules of the conserved regulatory RNA AbcR1 address multiple target mRNAs in and outside of the translation initiation region. *RNA Biol.*, **11**, 624–640.
- Klahn, S., Schaal, C., Georg, J., Baumgartner, D., Knippen, G., Hagemann, M., Muro-Pastor, A.M. and Hess, W.R. (2015) The sRNA NsiR4 is involved in nitrogen assimilation control in cyanobacteria by targeting glutamine synthetase inactivating factor IF7. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, E6243–E6252.
- Tobias, N.J., Heinrich, A.K., Eresmann, H., Wright, P.R., Neubacher, N., Backofen, R. and Bode, H.B. (2016) Phototrophic symbiosis is dependent on hfq-mediated regulation of secondary metabolites. *Environ. Microbiol.*, doi:10.1111/1462-2920.13502.
- Starczynowski, D.T., Morin, R., McPherson, A., Lam, J., Chari, R., Wegrzyn, J., Kuchenbauer, F., Hirst, M., Tohyama, K., Humphries, R.K. *et al.* (2011) Genome-wide identification of human microRNAs located in leukemia-associated genomic alterations. *Blood*, **117**, 595–607.
- Villamizar, O., Chambers, C.B., Riberdy, J.M., Persons, D.A. and Wilber, A. (2016) Long noncoding RNA Saf and splicing factor 45 increase soluble Fas and resistance to apoptosis. *Oncotarget*, **7**, 13810–13826.
- Wang, F., Song, W., Zhao, H., Ma, Y., Li, Y., Zhai, D., Pi, J., Si, Y., Xu, J., Dong, L. *et al.* (2017) The RNA-binding protein QKI5 regulates primary miR-124-1 processing via a distal RNA motif during erythropoiesis. *Cell Res.*, **27**, 416–439.
- Smith, C., Heyne, S., Richter, A.S., Will, S. and Backofen, R. (2010) Freiburg RNA Tools: a web server integrating IntaRNA, ExPaRNA and LocARNA. *Nucleic Acids Res.*, **38**(Suppl), W373–W377.
- Wright, P.R., Georg, J., Mann, M., Sorescu, D.A., Richter, A.S., Lott, S., Kleinkauf, R., Hess, W.R. and Backofen, R. (2014) CopraRNA and IntaRNA: predicting small RNA targets, networks and interaction domains. *Nucleic Acids Res.*, **42**, W119–W123.
- Hofacker, I.L., Fontana, W., Stadler, P.F., Bonhoeffer, S., Tacker, M. and Schuster, P. (1994) Fast Folding and Comparison of RNA Secondary Structures. *Monatsh. Chem.*, **125**, 167–188.
- Turner, D.H. and Mathews, D.H. (2010) NNDB: the nearest neighbor parameter database for predicting stability of nucleic acid secondary structure. *Nucleic Acids Res.*, **38**, D280–D282.
- List, M., Ebert, P. and Albrecht, F. (2017) Ten simple rules for developing usable software in Computational Biology. *PLoS Comput. Biol.*, **13**, e1005265.
- Lorenz, R., Bernhart, S.H., Höner Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F. and Hofacker, I.L. (2011) ViennaRNA Package 2.0. *Algorithms Mol. Biol.*, **6**, 26.
- Bernhart, S.H., Mückstein, U. and Hofacker, I.L. (2011) RNA Accessibility in cubic time. *Algorithms fMol. Biol.*, **6**, 3.
- Alkan, F., Wenzel, A., Palasca, O., Kerpedjiev, P., Rudebeck, A.F., Stadler, P.F., Hofacker, I.L. and Gorodkin, J. (2017) RIssearch2: suffix array-based large-scale prediction of RNA–RNA interactions and siRNA off-targets. *Nucleic Acids Res.*, doi:10.1093/nar/gkw1325.
- Lewis, B.P., Shih, I.-h., Jones-Rhoades, M.W., Bartel, D.P. and Burge, C.B. (2003) Prediction of mammalian microRNA targets. *Cell*, **115**, 787–798.
- Rehmsmeier, M., Steffen, P., Höchsmann, M. and Giegerich, R. (2004) Fast and effective prediction of microRNA/target duplexes. *RNA*, **10**, 1507–1517.

24. Pain,A., Ott,A., Amine,H., Rochat,T., Boulloc,P. and Gautheret,D. (2015) An assessment of bacterial small RNA target prediction programs. *RNA Biol.*, **12**, 509–513.
25. Wright,P.R., Richter,A.S., Papenfort,K., Mann,M., Vogel,J., Hess,W.R., Backofen,R. and Georg,J. (2013) Comparative genomics boosts target prediction for bacterial small RNAs. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, E3487–E3496.
26. Cayrol,B., Fortas,E., Martret,C., Cech,G., Kloska,A., Caulet,S., Barbet,M., Trepout,S., Marco,S., Taghbalout,A. *et al.* (2015) Riboregulation of the bacterial actin-homolog MreB by DsrA small noncoding RNA. *Integr. Biol. (Camb.)*, **7**, 128–141.
27. Wang,J., Rennie,W., Liu,C., Carmack,C.S., Prevost,K., Caron,M.-P., Masse,E., Ding,Y. and Wade,J.T. (2015) Identification of bacterial sRNA regulatory targets using ribosome profiling. *Nucleic Acids Res.*, **43**, 10308–10320.
28. Bos,J., Duverger,Y., Thouvenot,B., Chiaruttini,C., Branlant,C., Springer,M., Charpentier,B. and Barras,F. (2013) The sRNA RyhB regulates the synthesis of the *Escherichia coli* methionine sulfoxide reductase MsrB but not MsrA. *PLoS One*, **8**, e63647.
29. Benjamin,J.-A.M. and Masse,E. (2014) The iron-sensing aconitase B binds its own mRNA to prevent sRNA-induced mRNA cleavage. *Nucleic Acids Res.*, **42**, 10023–10036.
30. Moon,K., Six,D.A., Lee,H.-J., Raetz,C. R.H. and Gottesman,S. (2013) Complex transcriptional and post-transcriptional regulation of an enzyme for lipopolysaccharide modification. *Mol. Microbiol.*, **89**, 52–64.
31. De Lay,N. and Gottesman,S. (2012) A complex network of small non-coding RNAs regulate motility in *Escherichia coli*. *Mol. Microbiol.*, **86**, 524–538.
32. Coornaert,A., Chiaruttini,C., Springer,M. and Guillier,M. (2013) Post-transcriptional control of the *Escherichia coli* PhoQ-PhoP two-component system by multiple sRNAs involves a novel pairing region of GcvB. *PLoS Genet.*, **9**, e1003156.
33. Beisel,C.L., Updegrove,T.B., Janson,B.J. and Storz,G. (2012) Multiple factors dictate target selection by Hfq-binding small RNAs. *EMBO J.*, **31**, 1961–1974.
34. Chen,J. and Gottesman,S. (2016) Spot 42 sRNA regulates arabinose-inducible araBAD promoter activity by repressing synthesis of the high-affinity low-capacity arabinose transporter. *J. Bacteriol.*, doi:10.1128/JB.00691-16.
35. Yang,Q., Figueroa-Bossi,N. and Bossi,L. (2014) Translation enhancing ACA motifs and their silencing by a bacterial small regulatory RNA. *PLoS Genet.*, **10**, e1004026.
36. Holmqvist,E., Wright,P.R., Li,L., Bischler,T., Barquist,L., Reinhardt,R., Backofen,R. and Vogel,J. (2016) Global RNA recognition patterns of post-transcriptional regulators Hfq and CsrA revealed by UV crosslinking in vivo. *EMBO J.*, doi:10.15252/embj.201593360.
37. Lease,R.A., Cusick,M.E. and Belfort,M. (1998) Riboregulation in *Escherichia coli*: DsrA RNA acts by RNA:RNA interactions at multiple loci. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 12456–12461.
38. Møller,T., Franch,T., Udesen,C., Gerdes,K. and Valentin-Hansen,P. (2002) Spot 42 RNA mediates discoordinate expression of the *E. coli* galactose operon. *Genes Dev.*, **16**, 1696–1706.
39. Beisel,C.L. and Storz,G. (2011) The base-pairing RNA Spot 42 participates in a multioutput feedforward loop to help enact catabolite repression in *Escherichia coli*. *Mol. Cell*, **41**, 286–297.
40. Wright,P.R. (2016) Predicting small RNA targets in prokaryotes - a challenge beyond the barriers of thermodynamic models. PhD thesis, Albert-Ludwigs-University Freiburg, doi:10.6094/UNIFR/11472.
41. Robledo,M., Frage,B., Wright,P.R. and Becker,A. (2015) A stress-induced small RNA modulates alpha-rhizobial cell cycle progression. *PLoS Genet.*, **11**, e1005153.
42. Afgan,E., Baker,D., van den Beek,M., Blankenberg,D., Bouvier,D., Cech,M., Chilton,J., Clements,D., Coraor,N., Eberhard,C. *et al.* (2016) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res.*, **44**, W3.