



REVIEW

RNA interactomics: recent advances and remaining challenges

[version 1; referees: 2 approved]

Brigitte Schönberger, Christoph Schaal, Richard Schäfer, Björn Voß 

Institute of Biochemical Engineering, Computational Biology Group, University of Stuttgart, Stuttgart, 70569, Germany

v1 **First published:** 20 Nov 2018, 7(F1000 Faculty Rev):1824 (<https://doi.org/10.12688/f1000research.16146.1>)**Latest published:** 20 Nov 2018, 7(F1000 Faculty Rev):1824 (<https://doi.org/10.12688/f1000research.16146.1>)**Abstract**

Tight regulation of cellular processes is key to the development of complex organisms but also vital for simpler ones. During evolution, different regulatory systems have emerged, among them RNA-based regulation that is carried out mainly by intramolecular and intermolecular RNA–RNA interactions. However, methods for the transcriptome-wide detection of these interactions were long unavailable. Recently, three publications described high-throughput methods to directly detect RNA duplexes in living cells. This promises to enable in-depth studies of RNA-based regulation and will narrow the gaps in our understanding of RNA structure and function. In this review, we highlight the benefits of these methods and their commonalities and differences and, in particular, point to methodological shortcomings that hamper their wider application. We conclude by presenting ideas for how to overcome these problems and commenting on the prospects we see in this area of research.

Keywords

RNA, RNA-RNA interactions, Regulation

Open Peer Review**Referee Status:**  

	Invited Referees	
	1	2
version 1 published 20 Nov 2018		

F1000 Faculty Reviews are commissioned from members of the prestigious F1000 Faculty. In order to make these reviews as comprehensive and accessible as possible, peer review takes place before publication; the referees are listed below, but their reports are not formally published.

- 1 **Roland Hartmann**, Institute of Pharmaceutical Chemistry, Germany
- 2 **Lennart Randau**, Max-Planck-Institute for Terrestrial Microbiology, Germany

Discuss this article

Comments (0)

Corresponding author: Björn Voß (bjoern.voss@ibvt.uni-stuttgart.de)

Author roles: **Schönberger B:** Investigation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Schaal C:** Investigation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Schäfer R:** Data Curation; **Voß B:** Conceptualization, Investigation, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: The authors receive funding from the German Ministry of Education and Research through the e:Bio project “inteRNAct” (grant 031A310 to BV).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2018 Schönberger B *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Schönberger B, Schaal C, Schäfer R and Voß B. **RNA interactomics: recent advances and remaining challenges** [version 1; referees: 2 approved] *F1000Research* 2018, 7(F1000 Faculty Rev):1824 (<https://doi.org/10.12688/f1000research.16146.1>)

First published: 20 Nov 2018, 7(F1000 Faculty Rev):1824 (<https://doi.org/10.12688/f1000research.16146.1>)

Introduction

RNA-based regulation is an important regulatory layer in all domains of life, carried out by a multitude of non-coding RNAs (ncRNAs). They interact with proteins, RNA, and DNA and thereby interconnect different regulatory systems. The elucidation of the underlying networks is thus of central importance for a holistic understanding of cellular regulation. Computational prediction based on sequence alone turns out to be unsatisfactory, and, with the availability of second- and third-generation sequencing technologies, data-based inference of regulatory networks involving ncRNAs is the more promising approach. RNA–protein interactions are more or less routinely studied using crosslinking immunoprecipitation-sequencing (CLIP-seq)¹, RNA immunoprecipitation-sequencing (RIP-seq)², mapping RNA interactome *in vivo* (MARIO)³, or related methods, which rely on ultraviolet-induced crosslinking or affinity of proteins to RNA; specific complexes are then enriched by immunoprecipitating the RNA-binding protein of interest, followed by sequencing of the RNA content using RNA sequencing (RNA-seq) (see recent reviews^{4,5}).

A large fraction of ncRNAs carry out their function through complementary base pairing to other RNAs, often mRNAs. This may involve proteins, as in the case of the RISC complex for microRNA (miRNA)- or small interfering RNA (siRNA)-mediated post-transcriptional gene silencing, but the ncRNA is the main factor that specifies the target RNAs via base complementarity. Thus, unbiased methods for the transcriptome-wide discovery of direct RNA–RNA interactions are needed to decipher RNA regulatory networks. Approaches developed in 2016, described as Direct Duplex Detection (DDD) methods⁶, represent an important first step toward this goal. In the following, we will discuss their pros and cons and present our view on what future improvements may look like and what their impact might be.

State of the art

RNA–RNA interactions pose two basic problems when analyzing them. First, they are commonly disrupted during the process of RNA extraction and cannot be reconstituted to their native state *in vitro*. Second, there is no technology available that can sequence interacting RNA strands and maintain their relationship. The first approach to tackle these problems (in a rather coarse way) was described by Ramani *et al.*⁷. They presented RNA proximity ligation (RPL), in which interacting RNA strands are ligated to each other in a standard ligation reaction using RNase inhibitor-treated crude cell extracts followed by RNA-seq. The resulting sequencing reads contained chimeric reads but to a rather low fraction (0.28%). Nevertheless, the proximity ligation approach presents a solution to the abovementioned second problem because it allows one to trace back RNA–RNA interactions in data sets obtained by next-generation sequencing technologies.

The remaining problem of losing RNA–RNA interactions due to the harsh RNA extraction protocols is as long-standing as its solution: direct RNA crosslinking via psoralen derivatives.

Following irradiation, psoralens covalently and reversibly connect two pairing stretches of RNA. Thus, duplexes are maintained during RNA extraction, which enables their enrichment, for example, by nucleases that digest single-stranded RNAs. The psoralen-mediated *in vivo* crosslinking of RNAs has been known since the 1970s and successfully used to interrogate the structure of ribosomal RNAs, interactions between small nuclear RNAs, and several other RNA structures and interactions^{8–20}.

The combination of the two aforementioned strategies is the hallmark of the currently available DDD methods, namely ligation of interacting RNA followed by high-throughput sequencing (LIGR-seq)²¹, psoralen analysis of RNA interactions and structures (PARIS)²², and sequencing of psoralen-crosslinked, ligated, and selected hybrids (SPLASH)²³. Although they differ in several steps of their experimental protocols, they all share the same principal design, which is *in vivo* RNA crosslinking followed by RNA extraction, enrichment of crosslinked RNAs, proximity ligation, and sequencing via RNA-seq. A comparison of the individual methods is shown in [Figure 1](#).

A common bias that occurs when psoralen is used to crosslink double-stranded RNA (dsRNA) is its preference to intercalate into adjacent opposite pyrimidine bases, preferably uracil²⁴, meaning that interactions in GC-rich regions may be under-represented. Nevertheless, the chance that uracil residues are neighboring another base pair is 25%, assuming a uniform distribution of all possible pairs of bases.

The major problem of all of the abovementioned methods is that the proximity ligation step is highly inefficient and that it is not possible to perform an RNA-seq library preparation that enriches for or specifically targets the ligated RNAs. Furthermore, the proximity ligation step does not exclusively ligate interacting RNA strands but also all kinds of single-stranded RNAs in a random fashion. As a result, the yield of RNA–RNA interaction informative reads is very low, as shown in [Table 1](#).

Nevertheless, the presented approaches provided interesting and deep insights into the RNA–RNA interaction networks of the studied organisms and the respective growth conditions. With the help of these DDD methods, already experimentally verified RNA structures could be accurately recapitulated; in addition, novel RNA structures as well as new RNA–RNA interactions were revealed. For instance, LIGR-seq identified novel interactions between small nucleolar RNAs (snoRNAs) and mRNAs, including the snoRNA SNORD83B that controls the steady-state levels of its target mRNAs²¹. Data from PARIS support the results of Lin *et al.*²⁵, who discovered that the interactions within the long ncRNA NEAT1 have an important architectural function in paraspeckle formation and thus impact the regulation of gene expression. In addition, PARIS could detect inter-repeat duplexes in the long ncRNA XIST, which is essential for X chromosome silencing²². Furthermore, the use of SPLASH has uncovered the RNA–RNA interactome of influenza A viruses required for virus growth²⁶; this might facilitate the prediction of the emergence of new pandemic influenza strains.

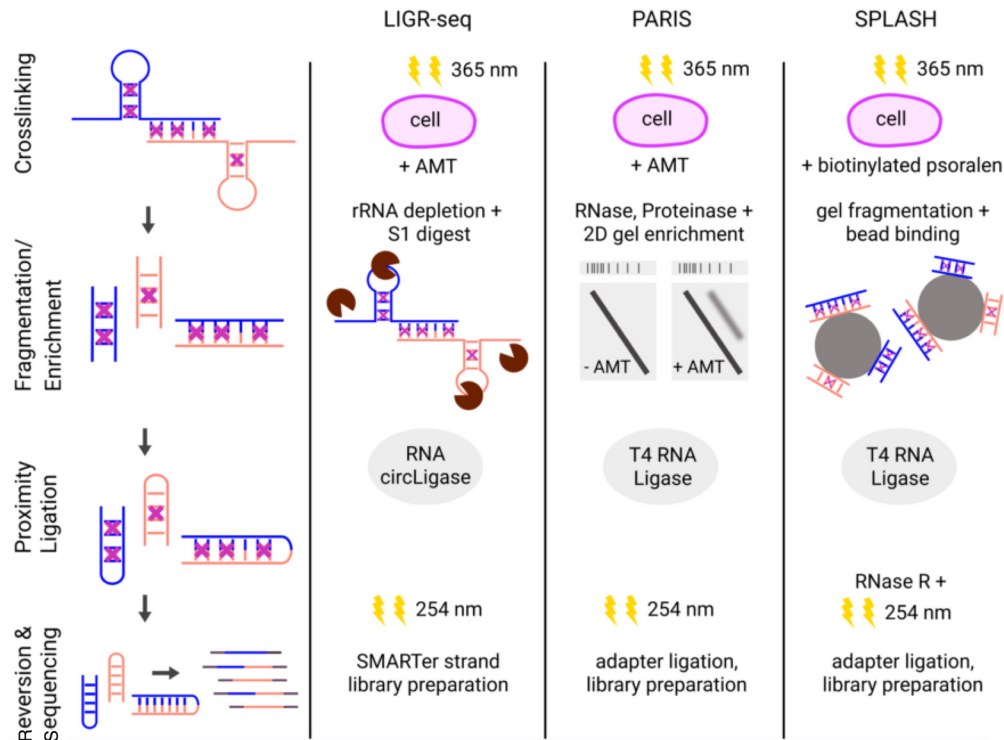


Figure 1. Schematic comparison of Direct Duplex Detection methods. Overview of the protocols for LIGR-seq²¹, PARIS²², and SPLASH²³. The first column shows the principal steps of the three experimental procedures: crosslinking (violet crosses) was conducted *in vivo* by 365 nm irradiation and 4'-aminomethyltrioxsalen (AMT) or biotinylated psoralen treatment. Fragmentation was performed enzymatically (LIGR-seq and PARIS) or chemically (SPLASH). Crosslinked RNAs were additionally enriched either by size separation using two-dimensional (2D) gel electrophoresis (crosslinked RNAs above the main diagonal were eluted; PARIS) or by biotin-streptavidin binding to magnetic beads (SPLASH). Proximity ligation was carried out using different ligases. Crosslinks were reverted by 254 nm irradiation. For sequencing, different library preparation strategies were performed. Colors of RNA strands (blue and orange) indicate different RNA molecules. LIGR-seq, ligation of interacting RNA followed by high-throughput sequencing; PARIS, psoralen analysis of RNA interactions and structures; SPLASH, sequencing of psoralen-crosslinked, ligated, and selected hybrids.

Table 1. Comparison of read statistics.

	LIGR-seq	PARIS	SPLASH
Total number of sequencing reads	171,239,817	99,698,824	189,340,955
Chimeric reads	6,614,251 (~3.9%)	2,077,743 (~2%)	1,038,801 (~0.5%)
RNA-RNA interactions	1,029	232,031 ^a	4,026

Data derived from psoralen-treated human cell line samples (LIGR-seq and PARIS: total RNA isolated from HEK293T human embryonic kidney cells; SPLASH: total RNA from GM12892 human lymphoblastoid cells); all replicates included. Values are taken from the supplementary information of the respective publication (LIGR-seq²¹, PARIS²², and SPLASH²³). In the case of LIGR-seq, the number of chimeric reads was determined by the corresponding analysis pipeline Aligater²¹. ^aSo-called Duplex Groups, representing gapped reads with interacting RNA sites. LIGR-seq, ligation of interacting RNA followed by high-throughput sequencing; PARIS, psoralen analysis of RNA interactions and structures; SPLASH, sequencing of psoralen-crosslinked, ligated, and selected hybrids.

Conclusion and perspective

Even based on the highly inefficient proximity ligation, the currently available DDD methods have been able to capture and analyze crucial RNA-RNA interactions in a transcriptome-wide

fashion and *in vivo*. Nevertheless, there is still room for improvement. A stricter enrichment of crosslinked RNAs by combining enzymatic digestions (LIGR-seq) with two-dimensional gel-electrophoretic separation (PARIS) or biotinylated-psoralen

selection (SPLASH) or both could eliminate the undesired fortuitous ligation of single-stranded or non-crosslinked RNA duplexes.

The efficiency of the proximity ligation could be enhanced with non-complementary overhangs. These may be either left over by a limited nuclease digestion or introduced synthetically before ligation, for example, by polyA tailing or by using different terminal transferases. Such duplex overhangs are more likely to be ligated and could lead to improved yields of interaction-informative sequencing reads.

The most straightforward approach to turn RNA duplexes into sequencing templates would be a direct dsRNA ligation method. So far, only one report has proposed using RNA and DNA ligases for the enzymatic joining of two or more dsRNA molecules²⁷. However, to the best of our knowledge, the results have never been reproduced, nor was the method used in further studies. In addition, future discoveries might provide engineered DNA ligases with substrate promiscuity to increase the dsRNA ligation efficiency. Eventually, a direct ligation of dsRNA adapters would pave the way for novel and hopefully more-efficient DDD protocols.

Likewise, the bioinformatics pipelines for basic data analysis and statistical inference of interactions need to be improved. Currently, none of the existing software packages considers the actual complementarity of the putative interaction partners but

relies only on statistical measures. Furthermore, RNA–RNA interactions are driven by base-pair formation, for which thermodynamic parameters and kinetic models exist. Integrating these aspects into the analyses may additionally improve the reliability and robustness of data analysis.

In conclusion, the currently available DDD methods pave the way to a deeper understanding of the so-far little-studied RNA–RNA interactome. Crucial for their future acceptance and success will be to increase the yield of informative reads, which would allow RNA regulatory networks to be deciphered in greater depth. Results from such studies could have a great impact on many research areas. For example, a lot of human diseases are associated with ncRNAs (for example, Alzheimer's disease²⁸, schizophrenia²⁹, and various cancer types³⁰ such as leukemia³¹, breast cancer³², or lung cancer³³). Thus, there is and will be a wide application range for first- and next-generation DDD methods that have the potential to provide important contributions in areas of basic and applied sciences.

Grant information

The authors receive funding from the German Ministry of Education and Research through the e:Bio project “inteRNAct” (grant 031A310 to BV).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References



1. Licatalosi DD, Mele A, Fak JJ, *et al.*: **HITS-CLIP yields genome-wide insights into brain alternative RNA processing.** *Nature*. 2008; 456(7221): 464–9. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
2. Zhao J, Ohsumi TK, Kung JT, *et al.*: **Genome-wide identification of polycomb-associated RNAs by RIP-seq.** *Mol Cell*. 2010; 40(6): 939–53. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
3. Nguyen TC, Cao X, Yu P, *et al.*: **Mapping RNA–RNA interactome and RNA structure *in vivo* by MARIO.** *Nat Commun*. 2016; 7: 12023. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
4. Saliba AE, Santos S, Vogel J: **New RNA-seq approaches for the study of bacterial pathogens.** *Curr Opin Microbiol*. 2017; 35: 78–87. [PubMed Abstract](#) | [Publisher Full Text](#) | [F1000 Recommendation](#)
5. Ryder SP: **Protein–mRNA interactome capture: cartography of the mRNP landscape [version 1; referees: 3 approved].** *F1000Res*. 2016; 5: 2627. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
6. Weidmann CA, Mustoe AM, Weeks KM: **Direct Duplex Detection: An Emerging Tool in the RNA Structure Analysis Toolbox.** *Trends Biochem Sci*. 2016; 41(9): 734–6. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
7. Ramani V, Qiu R, Shendure J: **High-throughput determination of RNA structure by proximity ligation.** *Nat Biotechnol*. 2015; 33(9): 980–4. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
8. Rimoldi OJ, Raghu B, Nag MK, *et al.*: **Three new small nucleolar RNAs that are psoralen cross-linked *in vivo* to unique regions of pre-rRNA.** *Mol Cell Biol*. 1993; 13(7): 4382–90. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
9. Calvet JP, Pederson T: **Base-pairing interactions between small nuclear RNAs and nuclear RNA precursors as revealed by psoralen cross-linking *in vivo*.** *Cell*. 1981; 26(3 Pt 1): 363–70. [PubMed Abstract](#) | [Publisher Full Text](#)
10. Rabin D, Crothers DM: **Analysis of RNA secondary structure by photochemical reversal of psoralen crosslinks.** *Nucleic Acids Res*. 1979; 7(3): 689–703. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
11. Isaacs ST, Shen CK, Hearst JE, *et al.*: **Synthesis and characterization of new psoralen derivatives with superior photoreactivity with DNA and RNA.** *Biochemistry*. 1977; 16(6): 1058–64. [PubMed Abstract](#) | [Publisher Full Text](#)
12. Wollenzien PL, Youvan DC, Hearst JE: **Structure of psoralen-crosslinked ribosomal RNA from *Drosophila melanogaster*.** *Proc Natl Acad Sci U S A*. 1978; 75(4): 1642–6. [PubMed Abstract](#) | [Free Full Text](#)
13. Wassarman DA: **Psoralen crosslinking of small RNAs *in vitro*.** *Mol Biol Rep*. 1993; 17(2): 143–51. [PubMed Abstract](#)
14. Rinke J, Appel B, Digweed M, *et al.*: **Localization of a base-paired interaction between small nuclear RNAs U4 and U6 in intact U4/U6 ribonucleoprotein particles by psoralen cross-linking.** *J Mol Biol*. 1985; 185(4): 721–31. [PubMed Abstract](#) | [Publisher Full Text](#)
15. Wassarman DA, Steitz JA: **Interactions of small nuclear RNA's with precursor messenger RNA during *in vitro* splicing.** *Science*. 1992; 257(5078): 1918–25. [PubMed Abstract](#) | [Publisher Full Text](#)
16. Thompson JF, Hearst JE: **Structure of *E. coli* 16S RNA elucidated by psoralen crosslinking.** *Cell*. 1983; 32(4): 1355–65. [PubMed Abstract](#) | [Publisher Full Text](#)
17. Lipson SE, Cimino GD, Hearst JE: **Structure of M1 RNA as determined by psoralen cross-linking.** *Biochemistry*. 1988; 27(2): 570–5. [PubMed Abstract](#) | [Publisher Full Text](#)
18. Hausner TP, Giglio LM, Weiner AM: **Evidence for base-pairing between mammalian U2 and U6 small nuclear ribonucleoprotein particles.** *Genes Dev*. 1990; 4(12A): 2146–56. [PubMed Abstract](#) | [Publisher Full Text](#)
19. Lustig Y, Wachtel C, Safo M, *et al.*: **'RNA walk' a novel approach to study**

- RNA-RNA interactions between a small RNA and its target.** *Nucleic Acids Res.* 2010; **38**(1): e5.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
20. Wagner R, Garrett RA: **A new RNA-RNA crosslinking reagent and its application to ribosomal 5S RNA.** *Nucleic Acids Res.* 1978; **5**(11): 4065–75.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
21. **F** Sharma E, Sterne-Weiler T, O'Hanlon D, *et al.*: **Global Mapping of Human RNA-RNA Interactions.** *Mol Cell.* 2016; **62**(4): 618–26.
[PubMed Abstract](#) | [Publisher Full Text](#) | [F1000 Recommendation](#)
22. **F** Lu Z, Zhang QC, Lee B, *et al.*: **RNA Duplex Map in Living Cells Reveals Higher-Order Transcriptome Structure.** *Cell.* 2016; **165**(5): 1267–79.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
23. **F** Aw JG, Shen Y, Wilm A, *et al.*: **In Vivo Mapping of Eukaryotic RNA Interactomes Reveals Principles of Higher-Order Organization and Regulation.** *Mol Cell.* 2016; **62**(4): 603–17.
[PubMed Abstract](#) | [Publisher Full Text](#) | [F1000 Recommendation](#)
24. Cimino GD, Gamper HB, Isaacs ST, *et al.*: **Psoralens as photoactive probes of nucleic acid structure and function: organic chemistry, photochemistry, and biochemistry.** *Annu Rev Biochem.* 1985; **54**: 1151–93.
[PubMed Abstract](#) | [Publisher Full Text](#)
25. Lin Y, Schmidt BF, Bruchez MP, *et al.*: **Structural analyses of NEAT1 lncRNAs suggest long-range RNA interactions that may contribute to paraspeckle architecture.** *Nucleic Acids Res.* 2018; **46**(7): 3742–52.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
26. Dadonaite B, Barilaite E, Fodor E, *et al.*: **The structure of the influenza A virus genome.** *bioRxiv.* 2017.
[Publisher Full Text](#)
27. Faridani OR, McInerney GM, Gradin K, *et al.*: **Specific ligation to double-stranded RNA for analysis of cellular RNA:RNA interactions.** *Nucleic Acids Res.* 2008; **36**(16): e99.
[PubMed Abstract](#) | [Free Full Text](#)
28. Faghihi MA, Modarresi F, Khalil AM, *et al.*: **Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase.** *Nat Med.* 2008; **14**(7): 723–30.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
29. Barry G, Briggs JA, Vanichkina DP, *et al.*: **The long non-coding RNA Gomafu is acutely regulated in response to neuronal activation and involved in schizophrenia-associated alternative splicing.** *Mol Psychiatry.* 2014; **19**(4): 486–94.
[PubMed Abstract](#) | [Publisher Full Text](#)
30. **F** Gupta RA, Shah N, Wang KC, *et al.*: **Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis.** *Nature.* 2010; **464**(7291): 1071–6.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#) | [F1000 Recommendation](#)
31. Wojcik SE, Rossi S, Shimizu M, *et al.*: **Non-codingRNA sequence variations in human chronic lymphocytic leukemia and colorectal cancer.** *Carcinogenesis.* 2010; **31**(2): 208–15.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
32. Mourtada-Maarabouni M, Pickard MR, Hedge VL, *et al.*: **GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer.** *Oncogene.* 2009; **28**(2): 195–208.
[PubMed Abstract](#) | [Publisher Full Text](#)
33. Ji P, Diederichs S, Wang W, *et al.*: **MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer.** *Oncogene.* 2003; **22**(39): 8031–41.
[PubMed Abstract](#) | [Publisher Full Text](#)

Open Peer Review

Current Referee Status:  

Editorial Note on the Review Process

F1000 Faculty Reviews are commissioned from members of the prestigious F1000 Faculty and are edited as a service to readers. In order to make these reviews as comprehensive and accessible as possible, the referees provide input before publication and only the final, revised version is published. The referees who approved the final version are listed with their names and affiliations but without their reports on earlier versions (any comments will already have been addressed in the published version).

The referees who approved this article are:

Version 1

- 1 **Lennart Randau** Max-Planck-Institute for Terrestrial Microbiology, Marburg, Germany
Competing Interests: No competing interests were disclosed.
- 2 **Roland Hartmann** Philipps-Universität Marburg, Institute of Pharmaceutical Chemistry, Marburg, Germany
Competing Interests: No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research