Genome **Biology**

EDITORIAL

The RBPome: where the brains meet the brawn

Naomi Attar

Old friends

Recent celebrations of the double helix, on the 60th anniversary of its discovery, heralded DNA as the key to life: the universal set of instructions for all organisms, the ubiquitous stuff of inheritance, without which our planet would be sterile. Although *Genome Biology* was among the culprits of this DNA-centric effervescence [1,2], our focus on DNA within the context of life as we know it belied the widely supported hypothesis that DNA has not always been integral to life, and that life's earliest forms existed in a pre-DNA world. Francis Crick, co-discoverer of DNA's double helical structure, devised a schema - with which you will undoubtedly be familiar - known as the Central Dogma, in which information flows from DNA to RNA to protein (but see [3] for an updated version). Of these three molecules, it is very likely that DNA is the new kid on the block, a crouton freshly added to the primordial soup. By contrast, the 'RNA world' hypothesis holds that RNA is the oldest extant genetically encoded macromolecule, and that it coexisted with proteins - and, prior to that, amino acids and peptides - before DNA arrived on the scene.

And so there we have it, RNA and proteins are old friends, who have learned to live intertwined with one another for longer than either has with DNA. By the nature of evolution, the co-dependency of these two molecules at the very early stages of life means that many of their interactions have become embedded in the fabric of the cell's most critical processes, inherited by the DNA world from its RNA predecessor. Famous examples of course include protein translation, through the ribosome, and RNA splicing, through the spliceosome (but see [4] for a dissenting, or at least more nuanced, view).

For such old friends, however, the interface where RNA and RNA-binding proteins (RBPs) meet - 'the RBPome' - remains remarkably unexplored, even though the more we search for RNA–protein interactions, the more we find them. This thematic issue of *Genome Biology*, dedicated to the RBPome, humbly sets out to take a small step toward bringing this neglected component of Crick's Central Dogma out into the light. We are assisted in this endeavor by Jernej Ule (University College London, UK) and John Rinn (Broad Institute, USA), the RBPome issue's talented, dedicated and altogether delightful Guest Editors.

Why the RBPome remained unloved: methods!

The neglect of the RBPome is not so much due to the disfavor of scientists as it is due to the paucity of methods available for studying RNA–protein interactions. However, a number of assays have been developed in recent years that are beginning to make the study of the RBPome more attractive, as reviewed in this issue by Mitchell Guttman and colleagues [5].

CLIP

Guest Editor Jernej Ule has led the way in RBPome assay development with CLIP (sometimes known as HITS-CLIP when combined with high-throughput sequencing) and iCLIP (compared in [6]), which are immunoprecipitation-based methods for transcriptome-wide profiling of RNA-binding sites for a given RNA-binding protein. Another popular, and related, method is PAR-CLIP [7], published by the groups of Mihaela Zavolan and Thomas Tuschl, who both contribute to this issue ([8] and [9,10], respectively).

The emergence of these high-throughput RBPome-mapping methods awaits a bioinformatics catch-up, with few tools available for data analysis. In this issue, three new user-friendly CLIP bioinformatics methods are described: the comparative tool dCLIP [11], and the two Galaxy-based tools pyCRAC [12] and PIPE-CLIP [13]. One of the most popular existing tools for PAR-CLIP data analysis is PARalyzer [14], from the groups of Jack Keene and Uwe Ohler, who both contribute to this issue ([15] and [10], respectively).

An important but specialized application of CLIP and related methods is the study of microRNA (miRNA) binding sites, through the immunoprecipitation of Argonaute 2, a protein that forms complexes with miRNAs and their RNA targets. For example, such an approach is used in

Correspondence: naomi.attar@genomebiology.com
Genome Biology, BioMed Central, 236 Gray's Inn Road, London WC1X 8HB, UK

this issue to guide the examination of miRNA targets in breast cancer subtypes [9]. The use of RBPome methods to map small RNA binding to the transcriptome is reviewed by Mihaela Zavolan and Nitish Mittal [8].

As immunoprecipitation-based assays, the CLIP suite of methods might be expected to suffer from the technical limitation of background binding. A study from Matthew Friedersdorf and Jack Keene clearly illustrates the potential for a background problem in PAR-CLIP data, but also shows how correcting for it can enhance the power of the assay to accurately map the RBPome [15].

Footprinting
A second approach for studying the RBPome looks at the global RBP footprint on the transcriptome; that is, the totality of RNA sequences bound by RBPs, rather than the binding profile of a single RBP. This strategy was pioneered in mammalian cells by the groups of Markus Landthaler, Christoph Dieterich, Matthias Hentze and Jeroen Krijgsveld ([16,17]; reviewed in [18]), and was explored last year in *Genome Biology* in a yeast setting [19]. In this issue, Landthaler and Dieterich extend their previous study by assaying a second mammalian cell line and developing a bioinformatics approach, POPPI, for analyzing these datasets and the differences between them [20]. In addition, Brian Gregory and Guest Editor John Rinn describe a new method, PIP-seq, that uses RNase digestion to globally profile the RBP footprint and that has the power to identify RBP-binding motifs [21].

The mRNA RBPome: a new frontier in gene regulation
Although all RNAs are relevant to the RBPome, the interaction of RBPs with mRNAs is of most interest for the large portion of the genomics community concerned with mechanisms of gene regulation. Armed with improved methods, many researchers are now trying to understand more about exactly how these binding events enable RBPs to regulate the mRNA component of the transcriptome. The many mechanisms by which they do so include promoting or repressing the splicing, stabilizing or degrading of mRNA molecules, as well as editing or modifying individual RNA bases. In this way, the muscle of the proteome bashes into shape the intelligentsia of the cell (in the form of the transcriptome). One might say that the mRNA RBPome is very much where the RNA brain meets the RBP brawn.

Profiling individual RBPs
Smaug
Smaug is a *Drosophila* RBP that was previously known to regulate two target RNAs through two different mechanisms: in one case it promoted destabilization, while in the other it repressed translation. In this issue,

Craig Smibert, Howard Lipshitz and colleagues use RIP-chip to profile transcriptome-wide Smaug binding targets in *Drosophila* early embryos, and perform experiments to identify mRNAs whose translation is repressed or whose stability is reduced by Smaug [22]. They show that Smaug employs a two-pronged approach for many of its targets, by using both its mechanisms to achieve down-regulation. Smaug binds a restricted set of mRNAs, with genes targeted including those functioning in the control of protein folding and degradation, lipid droplets and metabolism. Elmar Wahle and Michael Götze's Research Highlight discusses the study in more detail [23].

Sm proteins
An instructive example of the ancient nature of the RBPome is the protein family made up of the Sm and Sm-like RBPs, which are conserved throughout all three domains of life. These proteins form heterogeneous nuclear ribonucleoproteins (hnRNPs) best known for their function in RNA splicing. In this issue, Gregory Matera and colleagues profile the RNAs bound to a number of Sm and Sm-like proteins in two very different eukaryotic settings: *Drosophila* ovaries and HeLa cells [24]. The detection of mature mRNAs in RIP-seq experiments suggests splicing-independent functions for these hnRNPs.

ZFP36
ZFP36 is an RBP known to interact with AU-rich sequences and to have an antagonistic function to the better characterized transcript-stabilizing RBP ELAV1. In this issue, Neelanjan Mukherjee, Uwe Ohler and colleagues use PAR-CLIP and overexpression experiments to further probe the role of ZFP36 in negatively regulating the transcriptome, finding that among its targets are transcripts related to immune function and cancer, as well as those encoding other RBPs [10]. A comparison with ELAV1 data makes clear the large number of overlapping bindings sites between the two RBPs, and points toward the similar but non-identical motif preferences underlying their partial overlap.

Splicing
Intron retention
Rather than directly probing binding sites, RBPs that regulate splicing can instead be investigated through experiments that perturb or enhance their function. In this issue, such an approach is used both to study hnRNPLL in T lymphocytes [25] and the Sm-like protein LSm5 in salt-stressed *Arabidopsis* [26]. Both RBPs are found to be linked to the particular form of splicing known as intron retention, but in opposite directions.

Whereas alternative splicing is popularly thought to amount to either a different choice in exons or a shift in exon boundaries, more and more reports are emerging

of alternative splicing events where a transcript's exons are unchanged but an intron is retained. The purpose of intron retention is not altogether clear - in fact, whether or not it is even a deliberate regulatory process, rather than leaky RNA metabolism, is itself a matter of debate. For this reason, studies providing additional data on how RBPs relate to intron retention are very welcome.

Depletion of LSm5 in *Arabidopsis*, for example, increases intron retention, suggesting that LSm5 functions to promote splicing fidelity, and that intron retention in these instances is an unwanted result of cellular error [26]. However, another plant study in this issue, which examines light-responsive global splicing changes in the moss *Physcomitrella patens* - the first report to our knowledge of light-regulated splicing in plants - contains results consistent with intentional intron retention, in that these events are restricted to transcripts with a narrow set of functions, including light signaling and splicing [27]. Similarly, and contrary to the effect of LSm5, the hnRNPLL study finds that this developmentally regulated RBP actually induces intron retention, and that it does so in specific transcripts [25]. Moreover, the specificity of intron retention by hnRNPLL is not limited to target transcripts, but is also apparent in the location of the affected introns within these transcripts, with a preference for those flanking alternative exons.

Methods for computational analyses of splice sites

Bioinformatics methods for studying the RBPome, including those described in this issue, are not limited to tools designed for the analysis of CLIP data. Gene expression and other omics data can be exploited to make indirect inferences about RBPs, perhaps with even greater functional clarity than CLIP analysis, as can be seen in a number of research studies presented in this issue [10,22,25-29]. Two methods for making such inferences about RBPs are also published in the issue, and both relate to splice site analysis.

Matteo Cereda, Guest Editor Jernej Ule and colleagues present RNAmotifs, a bioinformatics method that identifies alternative splicing-related tetramer motifs from gene expression data [30]. At the heart of RNAmotifs is the principle that the binding of splice-regulatory RBPs to RNA is position dependent, which enables maps of enriched motifs to be generated where data are available for an RBP's differentially regulated transcriptome. Applying RNAmotifs maps to a number of RBPs shows that the positions at which they operate within transcripts tend to be very similar, but that the effect on splicing varies according to the RBP.

As reviewed by Jeremy Sanford and Timothy Sterne-Weiler [31], aberrant splicing has been linked to a number of human diseases. Sanford contributes to a Software article in this issue, from Matthew Mort and Sean Mooney, that describes MutPred Splice, a machine-learning tool for predicting those disease-associated variants that are likely to disrupt splicing [32]. Analyzing a dataset of disease-associated exonic variants with MutPred Splice suggests that different types of splicing defects are more common in inherited disease and cancer, respectively.

mRNA abundance

In parallel to the progress made in methods for experimental mapping of RNA-protein interactions, the development of bioinformatics tools has improved the *in silico* prediction of these interactions. One such tool, catRAPID [33], is deployed in this issue by Gian Gaetano Tartaglia and colleagues in an effort to relate RBPome interactions to gene expression. Using knockdown data for two RBPs, this approach finds that the propensity for a target RNA and an RBP to interact is predictive of the strength of target RNA expression change upon RBP depletion [28]. Discussed in Bojan Zagrovic's Research Highlight [34] is the subsequent global analysis, which uses RBPome-wide data to show that correlation with target RNA expression is a general trend for RBPs, and that this correlation can be either positive or negative. Target transcripts with positive correlations are associated with different cellular functions to those with negative correlations.

The abundance of those mRNAs encoding RBPs themselves is the focus of a study from the Janga laboratory [29], who find that these transcripts are more likely to be highly expressed in cancers than are other classes of genes. A network analysis considers what the consequences of this high expression might be for the RBPome.

The editosome and epitranscriptome

In addition to regulating splicing and mRNA abundance, RBPs can also modulate the transcriptome through editing and modifying bases. The epitranscriptome is not addressed in this issue, but has been elsewhere in *Genome Biology* [3,35,36]. However, RNA editing - a function of the ADAR family of RBPs [37] - is the subject of a comparative genomics computational study by Erez Levanon and colleagues [38]. An analysis of public human and mouse data, together with zebrafish as an outlier, shows that only a very small number of A-to-I editing sites are conserved in mammalian transcriptomes - just 59.

Levanon and colleagues note that conserved mammalian RNA editing sites tend to be located in transcripts with a synaptic function, suggesting that synapses may have a particular tendency to use RNA editing as a regulatory mechanism. An accompanying Research Highlight by Robert Reenan and Yiannis Savva discusses the study's findings in the wider context of the RNA editing field [39].

RBP specificity for RNA structure

Whereas transcription factors largely bind to the genome by forming contacts specific to the primary DNA sequence, RBPs have much more scope to achieve specificity through secondary structure, thanks to the large number of intramolecular bonds that twist RNA molecules into hairpins, stem-loops and various other bumps and bulges.

Rolf Backofen and colleagues use both structure and sequence features of RNA to develop a machine learning-based approach for predicting RBP binding sites from CLIP data [40]. In doing so, they demonstrate the importance of RNA secondary structure for conferring binding specificity in a subset of RBPs, while noting that structure does not make a strong impact on binding preference for some other RBPs.

A second structure-centered method to appear in the issue is that described by Hisanori Kiryu and colleagues, termed CapR [41]. Able to operate at high-throughput, CapR uses energy calculations to determine the probability of secondary structure throughout an RNA molecule. When applied to CLIP data, the software can be used to consider the secondary structure preferences of various RBPs, by calculating probability values at CLIP-determined binding sites. For example, the human RBP Pumilio-2 is shown by CapR to have a preference for hairpin loop structures.

The RBPome in reverse: RNAs regulate protein

While *Genome Biology*'s RBPome issue focuses on RNA–protein actions in which proteins are thought to regulate RNAs, the opposite can also be true, such as when long non-coding RNAs (lncRNAs) act as scaffolds for complex assembly. For the most part, lncRNA regulation of proteins is thought to occur in the nucleus [42], where the cell's lncRNA population is believed to be concentrated [43,44], although ribosomal profiling studies have also observed ribosome-associated lncRNAs [45,46].

Edwin Cuppen and colleagues take a new approach to studying lncRNA subcellular localization by sequencing transcriptome samples subjected to ribosomal fractionation (which yields separate nuclear, cytosolic, monosomal and polysomal fractions) [47]. The results of this experiment suggest that, although a small number of lncRNAs are indeed enriched in the nucleus, the majority are not - contrary to expectations. Instead, many lncRNAs appear to reside outside of the nucleus, including a large number that are associated with the proteins that make up ribosomes (both in monosomes and polysomes).

The finding that lncRNA association with ribosomes is commonplace begs the question: what purpose do untranslatable RNAs have in binding to the translation machinery? One possibility is that lncRNAs regulate ribosomes in some way, just as they have been shown to do for the various other cellular components that enact Crick's Central Dogma.

The architects of the RBPome issue

Scaling the heights of the RBPome would have been an impossible task without the contributions of many scientists from the genomics and wider research communities. The editors of *Genome Biology* are very grateful to everyone who supported this project by submitting a manuscript and to the many uncredited referees who were very generous with their time. Most of all, we are grateful for the invaluable and very significant assistance provided by our hands-on Guest Editors, Jernej Ule and John Rinn, who span both the Atlantic and the broad scope of the RBPome field. Please do read the Guest Editors' Editorial [48], in which their thoughts on the past, present and future of the RBPome field are conveyed.

We hope that, together, the authors, reviewers and Guest Editors of the issue, with a little help from *Genome Biology*, have put the brains, the brawn and the beauty of the RBPome firmly on the genomics map.

Abbreviations

CLIP: Crosslinking and immunoprecipitation; HITS-CLIP: High-throughput sequencing CLIP; hnRNPs: Heterogeneous nuclear ribonucleoproteins; iCLIP: Individual-nucleotide resolution CLIP; lncRNA: Long non-coding RNA; miRNA: microRNA; PAR-CLIP: Photoactivatable-ribonucleoside-enhanced CLIP; qPCR: Quantitative polymerase chain reaction; RBP: RNA-binding protein; RIP: RNA immunoprecipitation; RIP-chip: microarray RIP; RIP-seq: High-throughput sequencing RIP.

References

1. Attar N: **Raymond Gosling: the man who crystallized genes.** *Genome Biol* 2013, **14**:402.
2. Doolittle WF, Fraser P, Gerstein MB, Graveley BR, Henikoff S, Huttenhower C, Oshlack A, Ponting CP, Rinn J, Schatz M, Ule J, Weigel D, Weinstock G: **Sixty years of genome biology.** *Genome Biol* 2013, **14**:205.
3. Saletore Y, Meyer K, Korlach J, Vilfan ID, Jaffrey S, Mason CE: **The birth of the Epitranscriptome: deciphering the function of RNA modifications.** *Genome Biol* 2012, **13**:175.
4. Doolittle WF: **The spliceosomal catalytic core arose in the RNA world… or did it?** *Genome Biol* 2013, **14**:141.
5. McHugh CA, Russell P, Guttman M: **Methods for comprehensive experimental identification of RNA–protein interactions.** *Genome Biol* 2014, **15**:203.
6. Sugimoto Y, König J, Hussain S, Zupan B, Curk T, Frye M, Ule J: **Analysis of CLIP and iCLIP methods for nucleotide-resolution studies of protein-RNA interactions.** *Genome Biol* 2012, **13**:R67.
7. Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M, Jungkamp A-C, Munschauer M, Ulrich A, Wardle GS, Dewell S, Zavolan M, Tuschl T: **Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP.** *Cell* 2010, **141**:129–141.
8. Mittal N, Zavolan M: **Seq and CLIP through the miRNA world.** *Genome Biol* 2014, **15**:202.
9. Farazi TA, Hoeve JJ T, Brown M, Mihailovic A, Horlings HM, Vijver MJ VD, Tuschl T, Wessels LF: **Identification of distinct miRNA target regulation between breast cancer molecular subtypes using AGO2-PAR-CLIP and patient datasets.** *Genome Biol* 2014, **15**:R9.

10. Mukherjee N, Jacobs NC, Hafner M, Kennington EA, Nusbaum JD, Tuschl T, Blackshear PJ, Ohler U: **Global target mRNA specification and regulation by the RNA-binding protein ZFP36.** *Genome Biol* 2014, **15**:R12.

11. Wang T, Xie Y, Xiao G: **dCLIP: a computational approach for comparative CLIP-seq analyses.** *Genome Biol* 2014, **15**:R11.

12. Webb S, Hector RD, Kudla G, Granneman S: **PAR-CLIP data indicate that Nrd1-Nab3-dependent transcription termination regulates expression of hundreds of protein coding genes in yeast.** *Genome Biol* 2014, **15**:R8.

13. Chen B, Yun J, Kim MS, Mendell JT, Xie Y: **PIPE-CLIP: a comprehensive online tool for CLIP-seq data analysis.** *Genome Biol* 2014, **15**:R18.

14. Corcoran DL, Georgiev S, Mukherjee N, Gottwein E, Skalsky RL, Keene JD, Ohler U: **PARalyzer: definition of RNA binding sites from PAR-CLIP short-read sequence data.** *Genome Biol* 2011, **12**:R79.

15. Friedersdorf MB, Keene JD: **Advancing the functional utility of PAR-CLIP by quantifying background binding to mRNAs and lncRNAs.** *Genome Biol* 2014, **15**:R2.

16. Baltz AG, Munschauer M, Schwanhäusser B, Vasile A, Murakawa Y, Schueler M, Youngs N, Penfold-Brown D, Drew K, Milek M, Wyler E, Bonneau R, Selbach M, Dieterich C, Landthaler M: **The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts.** *Mol Cell* 2012, **46**:674–690.

17. Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann BM, Strein C, Davey NE, Humphreys DT, Preiss T, Steinmetz LM, Krijgsveld J, Hentze MW: **Insights into RNA biology from an atlas of mammalian mRNA-binding proteins.** *Cell* 2012, **149**:1393–1406.

18. Sibley CR, Attig J, Ule J: **The greatest catch: big game fishing for mRNA-bound proteins.** *Genome Biol* 2012, **13**:163.

19. Freeberg MA, Han T, Moresco JJ, Kong A, Yang YC, Lu ZJ, Yates JR, Kim JK: **Pervasive and dynamic protein binding sites of the mRNA transcriptome in *Saccharomyces cerevisiae*.** *Genome Biol* 2013, **14**:R13.

20. Schueler M, Munschauer M, Gregersen LH, Finzel A, Loewer A, Chen W, Landthaler M, Dieterich C: **Differential protein occupancy profiling of the mRNA transcriptome.** *Genome Biol* 2014, **15**:R15.

21. Silverman IM, Li F, Alexander A, Goff L, Trapnell C, Rinn JL, Gregory BD: **RNase-mediated protein footprint sequencing reveals protein-binding sites throughout the human transcriptome.** *Genome Biol* 2014, **15**:R3.

22. Chen L, Dumelie JG, Li X, Cheng MH, Yang Z, Laver JD, Siddiqui NU, Westwood JT, Morris Q, Lipshitz HD, Smibert CA: **Global regulation of mRNA translation and stability in the early *Drosophila* embryo by the Smaug RNA-binding protein.** *Genome Biol* 2014, **15**:R4.

23. Götze M, Wahle E: **Smaug destroys a huge treasure.** *Genome Biol* 2014, **15**:103.

24. Lu Z, Guan X, Schmidt CA, Matera AG: **RIP-seq analysis of eukaryotic Sm proteins identifies three major categories of Sm-containing ribonucleoproteins.** *Genome Biol* 2014, **15**:R7.

25. Cho V, Sanny A, Chan S, Enders A, Bertram EM, Tan A, Goodnow CC, Andrews TD: **The RNA-binding protein hnRNPLL induces a T cell alternative splicing program delineated by differential intron retention in polyadenylated RNA.** *Genome Biol* 2014, **15**:R26.

26. Cui P, Zhang S, Ding F, Ali S, Xiong L: **Dynamic regulation of genome-wide pre-mRNA splicing and stress tolerance by the Sm-like protein LSm5 in *Arabidopsis*.** *Genome Biol* 2014, **15**:R1.

27. Wu HP, Su Y, Chen HC, Chen YR, Wu CC, Lin WD, Tu SL: **Genome-wide analysis of light-regulated alternative splicing mediated by photoreceptors in *Physcomitrella patens*.** *Genome Biol* 2014, **15**:R10.

28. Cirillo D, Marchese D, Agostini F, Livi CM, Botta-Orfila T, Tartaglia GG: **Constitutive patterns of gene expression regulated by RNA-binding proteins.** *Genome Biol* 2014, **15**:R13.

29. Kechavarzi B, Janga SC: **Dissecting the expression landscape of RNA-binding proteins in human cancers.** *Genome Biol* 2014, **15**:R14.

30. Cereda M, Pozzoli U, Rot G, Juvan P, Schweitzer A, Clark T, Ule J: **RNAmotifs: prediction of multivalent RNA motifs that control alternative splicing.** *Genome Biol* 2014, **15**:R20.

31. Sterne-Weiler T, Sanford JR: **Exon identity crisis: disease-causing mutations that disrupt the splicing code.** *Genome Biol* 2014, **15**:201.

32. Mort M, Sterne-Weiler T, Li B, Ball EB, Cooper D, Radivojac P, Sanford JR, Mooney SD: **MutPred Splice: machine learning-based prediction of exonic variants that disrupt splicing.** *Genome Biol* 2014, **15**:R19.

33. Agostini F, Zanzoni A, Klus P, Marchese D, Cirillo D, Tartaglia GG: **CatRAPID omics: a web server for large-scale prediction of protein-RNA interactions.** *Bioinformatics* 2013, **29**:2928–2930.

34. Zagrovic B: **Of RBPs and their targets: interaction determines expression.** *Genome Biol* 2014, **15**:102.

35. Hussain S, Aleksic J, Blanco S, Dietmann S, Frye M: **Characterizing 5-methylcytosine in the mammalian epitranscriptome.** *Genome Biol* 2013, **14**:215.

36. Roberts RJ, Carneiro MO, Schatz MC: **The advantages of SMRT sequencing.** *Genome Biol* 2013, **14**:405.

37. Savva YA, Rieder LE, Reenan RA: **The ADAR protein family.** *Genome Biol* 2012, **13**:252.

38. Pinto Y, Cohen HY, Levanon EY: **Mammalian conserved ADAR targets comprise only a small fragment of the human editosome.** *Genome Biol* 2014, **15**:R5.

39. Savva YA, Reenan RA: **Identification of evolutionarily meaningful information within the mammalian RNA editing landscape.** *Genome Biol* 2014, **15**:103.

40. Maticzka D, Lange SJ, Costa F, Backofen R: **GraphProt: modeling binding preferences of RNA-binding proteins.** *Genome Biol* 2014, **15**:R17.

41. Fukunaka T, Ozaki H, Terai G, Asai K, Iwasaki W, Kiryu H: **CapR: revealing structural specificities of RNA-binding protein target recognition using CLIP-seq data.** *Genome Biol* 2014, **15**:R16.

42. Geisler S, Coller J: **RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts.** *Nat Rev Mol Cell Biol* 2013, **14**:699–712.

43. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG, Lagarde J, Veeravalli L, Ruan X, Ruan Y, Lassmann T, Carninci P, Brown JB, Lipovich L, Gonzalez JM, Thomas M, Davis CA, Shiekhattar R, Gingeras TR, Hubbard TJ, Notredame C, Harrow J, Guigo R: **The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression.** *Genome Res* 2012, **22**:1775–1789.

44. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi AM, Tanzer A, Lagarde J, Lin W, Schlesinger F, Xue C, Marinov GK, Khatun J, Williams BA, Zaleski C, Rozowsky J, Roder M, Kokocinski F, Abdelhamid RF, Alioto T, Antoshechkin I, Baer MT, Bar NS, Batut P, Bell K, Bell I, Chakrabortty S, Chen X, Chrast J, Curado J, *et al*: **Landscape of transcription in human cells.** *Nature* 2012, **489**:101–108.

45. Ingolia NT, Lareau LF, Weissman JS: **Ribosome profiling of mouse embryonic stem cells reveals the complexity of mammalian proteomes.** *Cell* 2011, **147**:789–802.

46. Guttman M, Russell P, Ingolia NT, Weissman JS, Lander ES: **Ribosome profiling provides evidence that large non-coding RNAs do not encode proteins.** *Cell* 2013, **154**:240–251.

47. van Heesch S, van Iterson M, Jacobi J, Boymans S, Essers PB, de Bruijn E, Hao W, MacInnes AW, Cuppen E, Simonis M: **Extensive localization of long noncoding RNAs to the cytosol and mono- and polyribosomal complexes.** *Genome Biol* 2014, **15**:R6.

48. Rinn JL, Ule J: **'Oming in on RNA–protein interactions.** *Genome Biol* 2014, **15**:401.