

Response to reviewers

General Note: The reviewers' comments were copied from the Genome Research website without changes. Our responses addressing individual comments and critiques are intercalated below.

Reviewer 1 Comments for the Author...

Tizard and colleagues describe cloning and deep sequencing of small RNAs from three stages of chicken development. The work follows the standard outline for "miRNA cloning" papers, and although it does not add novel insights into miRNA or small RNA biology, it expands the repertoire of chicken small RNAs and as such is a useful resource.

I have the following comments on the manuscript:

1. The title is misleading, since "atlas" implies information on spatial expression. My first reaction upon reading the title was that there will be a lot of in situ patterns or cloning from specific tissues that will give information where and when certain miRNAs are expressed. Unfortunately this work provides information only about relative abundance of small RNAs in total embryo extracts – not as useful as a real atlas would be. It's rather a catalogue.

R: In the hindsight we recognized that the 'atlas' in the manuscript title might have been misleading. We revised the main title and the running title to avoid ambiguity and to indicate the experimental approach described in the manuscript.

2. Accession numbers. Authors do not mention whether the raw sequencing data will be submitted into a public database. It is a standard procedure to submit the raw data (where only adapters are trimmed) to GEO database, and should be done for this work as well.

R: As indicated in the submitted manuscript, the most frequently sequenced tags derived from known and new miRNAs will be submitted to the miRBase, which generally requires a confirmation of manuscript acceptance prior to the data submission (see communication with Sam Griffith-Jones at the end). Sequences proposed for miRBase submission are now provided in Supplemental Tables S9-S12. Following the reviewer's request, the raw sequence data files for the 3 embryonic chicken libraries were submitted to the GEO database. The GEO ID for the data series is GSE10686. The link providing anonymous read-only access to the data for reviewers is as follows:

<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=tfmvhweyaekqatc&acc=GSE10686>.

The data will be fully accessible to public upon the final acceptance of the manuscript or after 1st of June 2008 whichever happens earlier. In addition, these data will be mirrored at our research website. The link to this website is provided in the manuscript.

3. Throughout the text authors inappropriately refer to the miRBase database as RFAM, which is a different database, and miRBase is independent of RFAM these days.

R: The RFAM references were replaced with miRBase references throughout the text to correct this error. We also updated the reference from miRBase 10.0 release (August 2007) to more recent miRBase 10.1 release (December 2007). There were no new chicken miRNAs added to the database since the previous release. The change was made simply to indicate that the most recent source of data was used.

4. Novel miRNAs. Authors report 449 unique sequences identified as putative novel chicken miRNAs. It is not clear how many of these can be classified as confident novel miRNAs (and thus submitted to miRBase), and how many as candidates. At the given sequencing depth having only two reads for a hairpin locus does not seem to be sufficient evidence to classify the locus as miRNA. E.g. how many of the putative miRNAs also have the star sequence cloned? In how many location of the mature sequence relative to the stem loop/base is in accordance with known requirements for Dicer/Drosha processing? How many of the novel miRNAs are in clusters and families?

R: The comments regarding the problem of discrimination between the real functional miRNAs and other small RNAs or RNA fragments present in small RNA libraries were brought up in different ways by all three reviewers and are highly relevant. Here we provide detailed response to all related comments, explanations of our analyses and revisions made to the manuscript to address this issue.

I. predicted RNA-hairpin properties and relative positions of sequence tags

We confirm that the known properties of the miRNA-hairpin precursors that satisfy basic criteria for Dicer/Drosha processing as well as relative positions of the sequence tags within the pre-miRNA hairpin were incorporated into our bioinformatics pipeline for miRNA identification. This methodology is relatively well established and has been described in several publications addressing the problem of accurate miRNA identification (for recent examples see (Morin et al. 2008; Ruby et al. 2007)). We confirm that to our best knowledge all new miRNAs and miRNA precursors presented in the manuscript satisfy general sequence and structural properties of authentic miRNAs. Supplemental File S3 contains sequences and predicted secondary structure data for all 449 new miRNA precursors described in the manuscript. We amended the corresponding section of the Supplemental Material to clearly indicate this fact.

II. Cut off value for the number of reads to satisfy minimal inclusion criteria.

Reviewer #1 and reviewer #2 have a rightful concern regarding the numbers of reads necessary for reliable identification of a sequence originating from the RNA hairpin precursor as real miRNA. As reviewer #2 correctly points out, this is an evolving consensus. Here we detail the rationale underlying our approach.

The first issue to be discussed is what can be accepted as an independent detection of a sequence by Solexa's deep sequencing method. Similarly to other high-throughput sequencing methods, this method includes a PCR-based amplification step that means that a random RNA fragment can be amplified and subsequently detected multiple times in the same sample; clearly this does not constitute an independent detection and may be referred to as 'noise'. In contrast to that, detections of the same sequence in two or more different RNA libraries represent truly independent events. In this case the probability of detecting an exactly the same 'random' fragment can be roughly estimated by multiplication of the individual observed probabilities in each library. In our case, the probability of detecting the same sequence once in all three small RNA libraries, based on a presumption that these sequences are detected by random chance, would be $(1/3 \times 10^6) \times (1/3 \times 10^6) \times (1/3 \times 10^6) = 2.7 \times 10^{-19}$.

The second important consideration is that the low counts of the detected sequence reads do not necessarily mean that a miRNA has an extremely low expression level. Instead, low sequence read counts may reflect a restricted tissue-specific expression pattern which is 'diluted' in the whole-embryo extract samples. Therefore simply lifting cut-off value for sequence read counts may have the undesirable effect of increasing the false discovery rate by the increasing the numbers of false negatives.

Out of the 449 new miRNAs presented in the manuscript 267 were detected at least once in all three chicken embryonic libraries (CE579), another 93 were detected at least once in two out of the three libraries, and the remaining 89 were detected at least twice in one out of the three libraries. In the latter subset 1 miRNA candidate has detectable miRNA* clones. Following the rationale presented above we are confident that 361 (267+93+1=361) out of 449 new miRNAs represent genuine miRNAs and the remaining 88 may be considered as miRNA candidates awaiting additional independent validation.

We revised relevant parts of the manuscript and methods section to reflect these changes. The sequences and predicted structures for the 361 new miRNAs and 88 miRNA candidates are supplied in the Supplemental File S3. The mature miRNA sequences for the 361 new miRNAs and 85 miRNA* are supplied in the

Supplemental Table S11. The mature sequences for the 88 miRNA candidates are supplied in the Supplemental Table S13.

III. miRNA* derived sequence tags

We agree that sequence tags originating from both arms of the RNA hairpin precursor serve as a strong indicator of the DICER-dependent RNA processing and increase confidence in the newly discovered miRNAs. Although the presence of miRNA*-derived sequence tags may be used as an additional criteria to validate new miRNAs it should be applied with caution. Most of the miRNAs* have much lower abundance than their respective miRNA (He and Hannon 2004). Thus the failure to detect miRNA* does not necessarily mean that the detected miRNA-derived sequence tags are false positives. The best example illustrating this point is the present state of the miRBase 10.1 annotations for chicken. If the known chicken miRNAs in the miRBase 10.1 were filtered for presence of miRNA*, then only 3 (!) out of 149 known chicken miRNA would have passed this filter. This example demonstrates that presence of miRNA* sequence is merely an issue of the sensitivity of the existing methods and stability of miRNA*.

The results presented in our manuscript re-validate and refine the annotation of known chicken miRNAs by providing information about miRNA* for most of the known chicken miRNAs.

To answer the question from the reviewer #1 of how many putative miRNAs have the star sequence cloned, we performed additional analyses. We found that 85 out of 449 new miRNAs presented in the manuscript have at least one sequence read derived from miRNA* detectable in at least one RNA library. The same 85 miRNAs are present in our revised stringent set of 361 miRNAs described above. All representative sequence tags derived from new miRNAs and new miRNA* are now supplied in the Supplemental Table S11. Following the acceptance of the manuscript these sequence tags will be submitted to miRBase.

IV. miRNA clusters

Only a small fraction of known miRNAs are located in close proximity (< 3Kb) on the genomic strand; these are known as miRNA clusters (see Ruby 2007, Figure 6 for an example (Ruby et al. 2007)).

Using 3 Kb as a distance cut-off we analyzed genomic distribution of known chicken miRNAs together with 361 new miRNAs. We identified 42 clusters consisting of 2 or more miRNAs. Of these, only 10 had 3 or more miRNAs and another 32 had 2 miRNAs. In total there were 100 miRNAs (66 known and 34 new) located within clusters, which is approximately 20% of all analyzed miRNAs (149

known + 361 new). All miRNAs found in clusters has been provided in the Supplemental Table S7.

5. Atypical mirtrons. Authors found cases where mature sequence is adjacent to the splice site but the host intron is large and the hairpin constitutes only part of the intron. They call these “atypical mirtrons” since “they appear to be processed in a similar way to other mirtrons”. However, mirtrons are defined as miRNAs where Drosha processing is substituted by splicing. There is no evidence that it happens in these atypical mirtrons and indeed Drosha seems to be still required to process at least the other, non-splice site arm of the hairpin. Thus the provided cases are not much different from other intron-residing miRNAs and should not be called mirtrons unless there is experimental evidence that splicing is an essential part of their biogenesis.

R: Mirtrons are a recently identified sub-class of miRNA that uncovered a link between miRNA processing and mRNA splicing machinery.

In our manuscript we present 18 cases which we call “atypical mirtrons”. Although the decision to call these mirtrons “atypical” is mostly symbolic (there are only 3 publications available to date on mirtrons and one of those doesn’t contain any experimental data), we felt that it’s important to point out the fact that these mirtrons are distinct from the previously described cases of mirtrons and intronic miRNAs. The reviewer correctly points out that ‘typical’ mirtrons do not require Drosha for processing and suggests that ‘atypical’ mirtrons would be at least partially dependent on Drosha and therefore should be called intronic miRNAs. This may be true. However, a ‘typical’ intronic miRNA does not overlap any major intronic landmarks that are utilized by the splicing machinery such as donor splice site, acceptor splice site and the branch-point. The overlap of the miRNA precursor sequence with any of these intronic features indicates close interaction between the spliceosome and Drosha complex.

Comments from the reviewer #1 and reviewer #2 made us realize that these points did not come across clearly in the text and Table 3 of the submitted manuscript. We have revised Table 3 which now includes an additional column that indicates position of sequence tags within the predicted RNA hairpin structures. We also performed an additional analyses of sequence tag locations relative to the splice sites; only the mirtrons that had sequence tags *directly* adjacent to either acceptor or donor splice sites have been kept in the revised Table 3; the remaining mirtron candidates were moved to Supplemental Table S8. Supplemental Figures S2a-S2r illustrating mirtrons listed in the Table 3 have now been included in the Supplemental Materials. In addition, we performed sequence analyses of the mirtron-derived sequence tags to

compare our results with the data presented by Berezikov et al., (Berezikov et al. 2006). The results of this analysis were added to the revised manuscript and provided in the Supplemental Figure S5.

Reviewer 2 Comments for the Author...

Glazov et al describe an extensive effort to catalog the small RNA (particularly microRNA) component of the chick embryo, using Solexa next-generation sequencing. As the authors say, the number of validated miRNAs in chicken is low compared with other vertebrates, in particular primates, and this study does much to redress this balance. A complete catalog of chicken miRNAs is vital to our understanding of gene regulation in birds, as well as enabling systematic and fundamental comparisons of the characteristics of small RNA species in birds with other vertebrates. The science is sound and the manuscript well-written and reasonably concise.

1. My major point of discussion relates to evolving consensus as to what constitutes a miRNA -- or pragmatically, what must be done to be confidently annotate a short RNA as a miRNA, rather than another small RNA or fragment:

R: This comment has been addressed as part of the detailed response to the reviewer #1 (comment #4).

1(a). The authors state that they have discarded all sequences seen only once, in line with previous studies. However, Solexa sequence is expected to yield an order of magnitude more data than 454 sequencing (on which many recent large-scale miRNA efforts are based). We might therefore expect the "noise" level of Solexa sequencing to be a little higher, and therefore require >2 independent observations for each sequence. I would therefore like to see the proportion of novel miRNAs that are discarded when more independent sequences are required. The observation of miR and miR* sequences, with the expected relationship across the hairpin, is compelling evidence of a real miRNA, and should therefore allow lower sequence counts to be used. In the absence of miR* sequences, it may be appropriate to be a little more conservative.

R: This comment has been addressed as part of the detailed response to the reviewer #1 (comment #4).

1(b). The authors should provide sequencing counts for their novel miRNAs in table S7, as they have done for known miRNAs in table S1.

R: Representative sequence read counts for new miRNAs has now been provided in the Supplemental Table S11. Note that only the sequences identical to the ones proposed for the miRBase submission were counted; the total read counts including sequences varying at the 5' and 3' ends are generally higher.

1(c). The authors should confirm that all of their novel sequences meet the miRNA annotation rules, published in Ambros et al, RNA, 2003. In particular, extensive base-pairing between mature miRNA and the opposite arm of the precursor hairpin has strongly discriminated miRNAs from other small RNAs and fragments.

R: This comment has been addressed as part of the detailed response to the reviewer #1 (comment #4).

1(d). The authors cite Berezikov 2006a as an example of primate miRNA counts, in support of the number of sequences identified here. Many potential miRNAs in that paper were not considered strong enough to be called true miRNAs, and remain "candidates". Many of those sequences have not been confirmed by several subsequent next-generation sequencing studies.

R: This comment has been addressed as part of the detailed response to the reviewer #1 (comment #4) in the section detailing the rationale of the underlying inclusion criteria for the confident miRNA set. In addition we note that the citation mentioned by the reviewer deals with very specialized type of tissue – brain; given that many miRNAs display tissue-specific expression pattern it is not surprising that different studies identify different subsets of miRNAs. We are not aware of any other study of a similar scale and tissue characteristics that would allow to make direct comparisons and to draw conclusions about rate of false positives in that study. Besides, as we mentioned above, the work published by Berezikov and colleagues is not a one-of-a-kind type of study, similar approaches have been used in a number of recent publications cited in our manuscript. More recent references include (Morin et al. 2008; Ruby et al. 2007).

2. The authors only very briefly discuss the observation that the small RNA counts decrease substantially from CE5 to CE9. I'd like to see more conjecture on the meaning of this, and the relevance of particular miRNAs that buck this trend.

R: As we discuss on the page 5, in our interpretation this observation reflects the underlying changes in the composition of steady-state small RNA populations between the 3 developmental stages. To avoid misinterpretation of the data, it is crucial to remember that the *detected* diversity in the small RNA populations is limited by the size of sampling (in our case ~3 million reads per sample), and represent variations in *relative abundance* of various small RNA classes (not limited to miRNAs). For example, if hypothetically, expression of a single small RNA changes such that its *absolute* abundance changes from one to one million copies, this would substantially re-balance *detected* diversity of other small RNAs in the *sampled* 3 million reads. This is why all comparative analyses on our and other similar datasets are performed using relative units such as read counts per million (see Supplemental Table S1). To

summarize, *unique* sequence read counts reflect the detected diversity of the steady-state small RNA populations, while *total* sequence counts reflect *relative* expression of the individual small RNAs and/or small RNA classes. Analyses of the *relative* changes in expression of known chicken miRNAs have been presented in the Table 2 and Supplemental Table S1.

3. The authors find 18 atypical mirtrons, in longer introns. Do they claim that these sequences have only one end of the precursor hairpin defined by the splicing machinery, or that the hairpin itself is longer? Do they observe miR* sequences in the correct place in the predicted hairpin?

R: The 'atypical mirtrons' issue has been addressed in detail in response to reviewer #1 (comment #5). The miR* part of the comment has been covered in the detailed response to the reviewer #1 (comment #4-III). The precursors for 'atypical mirtrons' do not span the entire intron from donor to acceptor site, therefore only one of the sites is defined by the spliciosome. We have amended the relevant part of the text to clarify this point.

4. The importance of the section on rasiRNAs seems to hang on whether the proportion of small RNAs that are rasiRNAs in chicken is low as a result of fewer repeat families, but this result isn't reported. Otherwise, I suggest this section doesn't add much, and is dropped.

R: We agree with the reviewer that the section on rasiRNAs is not of a crucial importance to the presented manuscript and might detract attention from the main topic of the paper, which is focused on microRNAs. Following the reviewer's suggestions we excluded this section from the revised version of the manuscript.

Minor points:

5. The miRBase database is variously referred to as miRBase, Rfam, and the miRNA Registry. The data used by the authors comes from miRBase v10.0.

R: This issue has been addressed above in the response to the reviewer #1. The miRBase references were consistently renamed throughout the text.

6. Page 7: I don't understand the relevance of the sentence about mir-18a and the splicing repressor protein.

R: The purpose of this sentence was to highlight the connection between the unusual distribution of sequence tags derived from the 3 reported loci and the alternative pathway of miRNA processing described by Guil and Caceres (Guil and Caceres 2007). We agree with the reviewer that the indicated place in the manuscript was not well suited to bring up this point. In the revised manuscript the sentence was moved to the discussion section.

7. Page 10, line 15: If what is true?

8. Page 10, line 16: I would think that many miRNAs (eg species or clade-specific) are likely to be very fast evolving.

R: Comments 7 and 8 refer to one sentence. We recognized that this sentence wasn't worded well and failed to deliver the point we were trying to make. The sentence has now been revised.

9. Page 12, line 10: "While true" -- I don't think we can say that this is true. Indeed, the rest of sentence claims that it probably isn't.

R: This sentence has been revised.

10. Page 12, line 16: "largely responsible" -- this sounds like you are claiming that the most important controllers of basic cellular and developmental pathways are conserved miRNAs, which I don't think you mean!

R: Following the reviewer's comment we recognized that the sentence in question could be misinterpreted. We revised the sentence to avoid ambiguity and to clarify our point.

11. The chicken genome consortium paper is strangely cited in the references and text.

R: This citation has been corrected to comply with the standard referencing format.

Reviewer 3 Comments for the Author...

In their manuscript Glazov et al have used a deep sequencing approach (Solexa) to obtain 9.5 million short sequence reads from three small RNA libraries, which were prepared from different developmental stages of the chicken embryo. These sequences were analysed with the aim to identify novel microRNAs. The results of this work are of interest to the readership of Genome Research and particularly to researchers working with the chicken.

However, in the opinion of this reviewer the manuscript is not of sufficient quality to warrant publication by this journal. The principle shortcoming is that it is not clear whether the sequences identified are true microRNAs/mirtrons. No other experimental approach other than sequencing was employed to confirm this.

Prior to publication a number of issues need to be addressed.

The title does not accurately reflect the content of the manuscript. Only three time points during chick development were analysed and this does not constitute a 'microRNA expression atlas'. There is no information regarding the different tissues or cell types which may be expressing these putative microRNAs.

R: This issue was addressed above in the response to the reviewer #1. The title of the manuscript has been revised.

449 putative new chicken miRNAs were discovered, of which 430 are said to be specific to the avian lineage. This seems like a lot and without further information regarding the

bioinformatics pipeline and the criteria used to define these sequences as true microRNAs confidence in this data has to be low.

R: This comment has been addressed as part of the detailed response to the reviewer #1 (comment #4). In addition we revised the corresponding section in the Supplemental Methods file to provide more detail about sequence analysis and miRNA identification.

For some of the sequences the mir* was sequenced and this provides additional evidence. However, additional approaches (QPCR, Northern, in situ hybridization) are required to confirm the sequences as microRNAs.

R: The miR* part of the comment has been covered in the detailed response to the reviewer #1 (comment #4). In regards of the additional experimental validation of the individual miRNAs we feel that this task, while important, goes beyond the scope of this publication; especially given the fact that the general validity of the deep sequencing approaches for miRNA discovery has already been demonstrated in several independent studies (Cummins et al. 2006; Landgraf et al. 2007; Morin et al. 2008; Ruby et al. 2006; Ruby et al. 2007).

The frequency with which a sequence is obtained should, in my opinion, not be used to speculate about relative miRNA abundance. The quality of the different libraries is most likely variable and unless the experiment is performed in triplicate no direct comparisons can be made.

R: The only place in the manuscript that describes comparison of known miRNA abundance between samples is the paragraph at the end of the section ‘Known miRNAs’. This section clearly states that “our experiment was not designed for direct comparison of miRNA abundance between the three different developmental stages”, and thus no claims of ‘direct comparisons’ have been made. What we did do, however, is to analyze *potentially significant* ($P < 0.01$) changes in *relative abundance* of known miRNAs using an accepted statistical approach described in Supplemental Methods and illustrated in the Supplemental Table S1. Moreover, an alternative statistical approach developed by Audic and Claverie (Audic and Claverie 1997), was successfully used for a relative quantification of the two deep sequencing libraries in the recent paper by Morin et al. published by the Genome Research (Morin et al. 2008). Thus the reviewer’s opinion on this matter appears to be misinformed. We therefore maintain that the paragraph in question does not contain any false statements and adds an informative point to the manuscript.

In general the results section contains too much discussion/speculation, e.g. top of page 7 'developmental switching', page 10 end of mirtron section 'suggest that mirtrons are rapidly evolving', etc.

R: Most of these have now been moved to the Discussion section of the revised manuscript. In few remaining cases we decided that it is appropriate to provide interpretation of the specific points of the results where discussing them separately would inevitably lead to repetitions in the manuscript.

It is not acceptable to provide supporting online material through the author's own research website. This material cannot form part of the review process unless it is deposited with the journal.

R: This issue has been addressed in part in the response to the reviewer #1 (comment #2). The research website has been established as a supplemental source of data to assist the review process and the dissemination of data rather than a substitute to public databases or journal's supplemental materials online. Additional Supplemental Materials have now been provided together with the revised manuscript.

There is some published work regarding the expression of microRNAs during chick development and the relevant papers should be included in the bibliography.

R: The citation to a publication by Darnell et al. (Darnell et al. 2006) describing patterns of miRNA expression in chicken embryo has been added to the discussion section.

Additional information:

-----Original Message-----

From: Sam Griffiths-Jones [mailto:sam.griffiths-jones@manchester.ac.uk]

Sent: Monday, 3 December 2007 10:21 PM

To: Glazov, Evgeny (LI, St Lucia)

Cc: microrna@sanger.ac.uk

Subject: Re: miRNA registry submission

Hi Evgeny

I'm sorry for the delayed reply. I've interspersed comments below:

On Wed, 21 Nov 2007, Evgeny.Glazov@csiro.au wrote:

> Dear Sam,

> We are preparing a set of novel chicken miRNAs for submission to the RFAM. Before we proceed, I want to make sure that this is done in the most efficient way for us and for the miRBase support team. I would greatly appreciate if you could provide answers to the following questions.

> 1. What is an expected date for the next release of the RFAM?

A miRBase release is overdue, so I hope to get one out in two weeks or so.

> 2. How long does it usually take to assign official names and IDs for novel miRNA candidates?

We usually do this when you have an acceptance letter for the journal for your manuscript. Then you should email me (or better microrna@sanger.ac.uk) will the final sequence set you need names for, and I can usually turn them around in a couple of days. You can then include those names in the final manuscript version.

> 3. We analysed RNA secondary structure predictions and evolutionary conservation for our novel miRNA candidates. How can we submit this information to RFAM and will it be useful?

I guess the conservation data can go in to the database in the form of orthologous sequences. We can add obvious orthologs of experimentally validated miRNAs without requiring additional experimental data. Those sequences then get the same treatment as everything else -- defined families and the like. Please include those sequences when you send your experimentally validated set.

I hope that helps

Sam

Additional references

- Audic, S. and J.M. Claverie. 1997. The significance of digital gene expression profiles. *Genome Research* **7**: 986-995.
- Berezikov, E., F. Thummler, L.W. van Laake, I. Kondova, R. Bontrop, E. Cuppen, and R.H.A. Plasterk. 2006. Diversity of microRNAs in human and chimpanzee brain. *Nature Genetics* **38**: 1375-1377.
- Cummins, J.M., Y.P. He, R.J. Leary, R. Pagliarini, L.A. Diaz, T. Sjoblom, O. Barad, Z. Bentwich, A.E. Szafranska, E. Labourier et al. 2006. The colorectal microRNAome. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 3687-3692.
- Darnell, D.K., S. Kaur, S. Stanislaw, J.H. Konieczka, T.A. Yatskievych, and P.B. Antin. 2006. MicroRNA expression during chick embryo development. *Dev Dyn* **235**: 3156-3165.
- Guil, S. and J.F. Caceres. 2007. The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. *Nature Structural & Molecular Biology* **14**: 591-596.
- He, L. and G.J. Hannon. 2004. MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* **5**: 522-531.
- Landgraf, P., M. Rusu, R. Sheridan, A. Sewer, N. Iovino, A. Aravin, S. Pfeffer, A. Rice, A.O. Kamphorst, M. Landthaler et al. 2007. A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **129**: 1401-1414.
- Morin, R.D., M.D. O'Connor, M. Griffith, F. Kuchenbauer, A. Delaney, A.L. Prabhu, Y. Zhao, H. McDonald, T. Zeng, M. Hirst et al. 2008. Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res*.
- Ruby, J.G., C. Jan, C. Player, M.J. Axtell, W. Lee, C. Nusbaum, H. Ge, and D.P. Bartel. 2006. Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. *Cell* **127**: 1193-1207.
- Ruby, J.G., A. Stark, W.K. Johnston, M. Kellis, D.P. Bartel, and E.C. Lai. 2007. Evolution, biogenesis, expression, and target predictions of a substantially expanded set of *Drosophila* microRNAs. *Genome Res* **17**: 1850-1864.