EMBO reports

Manuscript EMBOR-2011-34844

A genetic in vivo system to detect asymmetrically distributed RNA

N JayaNandanan, Elizabeth R. Gavis, Veit Riechmann and Maria Leptin

Corresponding author: Maria Leptin, Koeln, University

Review timeline:

Submission date:	28 February 2011
Editorial Decision:	25 March 2011
Revision received:	24 June 2011
Editorial Decision:	18 July 2011
Revision received:	10 August 2011
Editorial Decision:	15 August 2011
Accepted:	15 August 2011

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision 25 March 2011

Thank you for the submission of your manuscript to our editorial office. We have now received the evaluation of three expert reviewers on it.

As you will see from the comments pasted below, while referees 2 and 3 agree on the potential usefulness of the method to detect asymmetrically localized mRNAs and support publication of the study in EMBO reports, referee 1 is more negative. However, on balance, we would like to give you the opportunity to revise your study along the lines described below.

All three referees point out technical issues that would need to be addressed in order to further validate the method. For example, referee 2 feels that additional information on the effects of the site of P element insertion should be provided; this should also be taken into account with regard to the negative lines and this referee feels that mRNA expression levels should be analyzed more carefully (reviewer 3 also comments on this aspect). Referee 2 further suggests investigating the localization of the proteins encoded by the asymmetrically distributed mRNAs.

With regard to the criticisms raised by the more negative referee #1, we feel that some of his/her suggestions, while certainly being important to address in the future, go beyond the scope of the current manuscript. Nevertheless, it would certainly strengthen the study if you were able to provide data showing that the asymmetric localization of the mRNAs within the trachea is of biological importance. This would, however, not be an essential prerequisite for publication.
I look forward to seeing a revised form of your manuscript when it is ready.

Yours sincerely,

Editor
EMBO Reports

REFEREE REPORTS:

Referee #1:

This manuscript describes a novel strategy for screening for the subcellular distribution of specific mRNAs in Drosophila by transposon-mediated tagging with a sequence that allows the fusion RNA to be decorated with GFP (using the previously published MS2-GFP method). This is a very nice concept, permitting unbiased identification of novel localizing mRNAs in specific tissues (the system allows expression of the fusion RNAs to be driven with UAS-GAL4). The manuscript describes the use of this tool to identify mRNAs that are present in the processes of the branched terminal cell of the tracheal system or da neurons, as well as one mRNA enriched at the anterior of the oocyte (a smaller number of lines was screened in the oocyte and da neurons than in the tracheal system). There is no overlap in the identity of mRNAs that are scored as localizing in the different cell types, suggesting the existence of distinct targeting mechanisms. In addition, helpful controls are performed to try to understand the reasons why many lines do not show GFP-labelled RNA in the processes of the terminal cells.

As the authors describe in the abstract, the manuscript describes a pilot screen. Unfortunately, the rather small number of lines screened restricts the ability to draw general conclusions. Also, the mechanism of localization is not probed for any of the identified mRNAs. Therefore in my opinion the manuscript falls a long way short of providing a sufficient systematic or mechanistic advance to be published in EMBO Reports. The findings described are interesting, but read like a promising start to a more in depth or systematic study that may be better suited to a longer format journal. Tellingly, the manuscript does not report a key finding of sufficient significance to the field, which is one of the major criteria for acceptance in this journal.

There are also some technical issues, as well as problems with the data presentation. I recommend the authors consider these seriously before resubmitting the paper elsewhere.

Major points:

1. There is not a clear distinction between the cytoplasmic localization of the tagged mRNAs in figure 4A - C. This is particularly the case for B and C, which are classed as negative and positive, respectively. High quality images (as in D and E) should be shown for the lines in A and B. Quantification of representative examples would be welcome, if possible.
2. Figure 4. The authors show that the MS2-GFP puncta in a positive line do not co-localize with Pacman. However, this is a very cursory characterization of the identity of these puncta. Several other components of RNA-associated bodies and transport complexes could be investigated for co-localization.
3. It is mentioned that RNAi-mediated knockdown specifically in the tracheal system of several of the mRNAs found in the screen leads to lethality. Knockdown of one of them leads to tracheal defects. No data or description of the specific phenotype are shown for this particular line. It is not shown that these mRNAs have a selective function in the tracheal system by assaying the requirements in other cell types. Most importantly, the functional significance of the localization process itself is not tested.

Minor points/questions:

1. Figure 2 and 3 can be combined, as the information in figure 2 does not justify its own figure.
2. Figure 1 should be referred to in the results and discussion section, and not the introduction. It seems unnecessary to cite the first author's thesis as the data are shown in figure 1.
3. Figure 4A: the position of the nucleus is not clear. Also, given historical problems with in situ

hybridization in the tracheal system, sense control probes for SRF or Dof could be helpful in providing further confidence in the technique. Figure 4B: the position of the tracheal cell is not clear with the single labelling. The arrows in C are not shown and the arrows in D are not referred to in the legend.
4. Only a very small number of the positives are followed up by in situ hybridization to assay for subcellular localization of the endogenous transcript. This is an important validation of the MS2 technique and the authors should either show the in situ results for more positives, or state in the text which ones were not informative due to expression in the surrounding tissue.
5. Introduction: the MS2-GFP system is described as being developed recently, but the citation is 1998.
6. Figure 5. At least some components of the screen appear to be expressed in egg chambers with the UAST vector, which leads to very low levels of maternal expression. Might a significant proportion of the negatives be scored as such because the level of MS2-RNA expression is not high enough to be visualized (in order to help judge this, an image of MS2-CP alone in the oocyte (without an MS2 RNA) would be very important). Also, not many details of the coat protein constructs are provided. What is the nuclear localization signal? Are there any other modifications compared to other versions?
7. Is cenG1a (tagged in line MS2-152a) endogenously expressed and localized in oogenesis? If so, this would provide further validation of the method.
8. Are the MS2-tagged RNAs translated normally? The presence of secondary structure or start/stop codons in the MS2 cassette may interfere with translation and in some cases this might have consequences on localization.
9. Previously it has been shown by Dr. Gavis and colleagues that a proportion of the MS2 coat protein can exit the nucleus without an MS2-tagged RNA in da neurons. Therefore the appearance of a representative negative EP-MS2 line in da neurons should be shown for comparison to figure 5C and D, as should the appearance of MS2 coat protein without an MS2 tagged mRNA.
10. In figure 5C and (in particular) D, why are there so many green dots outside of the da neuron (where the MS2 RNA is presumably not present with this GAL4 line)?
11. Legend to figure 5: there is a small typo: microns, rather than micromolar, should be referred to.

Referee #2:

This manuscript describes a potentially efficient means to tag genes such that their transcripts can be tracked fluorescently at the subcellular level. A specially designed P-element is mobilized, and those that land in the 5' region of genes can be used to drive expression (GAL4-dependent) of MS2 CP-tagged transcripts. A small pilot screen with over 200 new lines was assessed using a GAL4 driver that expresses in the developing tracheal system. Of these, 11 were detected with GFP, localizing subcellularly to tracheal cell processes. The authors also showed that the system could be used in other tissues, with different types of localization observed and with other gene products. In most cases examined, the endogenous genes were also found to be expressed and localized similarly in the same tisues/cells.

In general, this is a very well done and convincing demonstration of what should prove to be a useful new tool. It should be interesting and usable for a large number of Drosophila researchers. Larger sets of mobilized lines could be shared by numerous groups interested and willing to examine tissues of particular interest to them. The potential scale of these studies, and their results, should also have major implications for other researchers working on different model systems.

There are, however, a few issues that the authors need to address.

Despite the fairly extensive controls, it wasn't clear to me how important the site of P element insertion is, and at what level the negatives occur at. In figure 2, and the first part of the results, insertions in the 5'UTR are described and discussed. However, the proportion of P element insertions in 5'UTRs is generally quite small. Of the 11 positive lines, how many were in the 5'UTR? How often are other insertions (ie upstream? intron etc) expected to yield useful fusion proteins? I was not fully convinced that the majority of negative results are not simply due to lack of expression of full-length tagged mRNA production. Only three "negative" lines were examined for expression levels and 2 of the 3 looked quite weak in the PCR controls shown. These should also be

done quantitatively in order to be convincing. Given that most localization elements appear to be in 3'UTRs, the distance between localization elements and tags may often be quite large, increasing the chances of separation. The PCR controls don't account for this either, as the gene specific primers used appear to be close to the 5' ends of the genes. These analyses should be repeated with larger numbers of genes, and more quantitative assays that also assess transcript integrity. These controls are needed to properly assess the usefulness of the system, and how well it reflects the frequency and types of mRNA localization. Given that many of the negatives observed may occur due to undetectable levels of GFP-tagged mRNAs, the authors should also consider a more sensitive assay on a number of their negatives. Fixing the tissues and staining for GFP immunologically is one possibility (FISH may also be a useful assay). A possibility for improving on the efficiency of this system would be to also add splice donor sites.

More minor points:

I found it somewhat surprising that only one type of localization was observed in this study. Perhaps this is due to the small number of positives and the relative frequency of this type of localization.

The high correlation between expression of endogenous genes and reporters likely points to accessibility of P elements inserted in tracheally expressed genes - little or no expression would be expected if the chromatin were in an inaccessible state. This may also explain, in part, why transcripts that localized in one tissue did not appear to localize in others (p8). They simply may not be inducible in the other tissues. The authors interpretation that localization signals and machineries may be very different in different cell types does not jive well with data published to date, which suggest strong overlaps (at least for microtubule based mechanisms).

The manuscript would be strengthened further if the authors show whether proteins encoded by localized transcripts are also localized.

Referee #3:

JayaNandaram et al report their development of an in vivo genetic system for detecting localized mRNAs in Drosophila which promises to facilitate medium- to large-scale screens. They modified the EP transgenesis vector with the addition of 6 copies of the MS2 hairpin and allowed it to randomly insert in the genome; as P-elements preferentially land within transcribed regions, often in 5' UTRs, this allows mRNAs to be randomly tagged. Co-expression of GFP-tagged MCP then enables detection of the mRNA. Of 250 lines analyzed in developing tracheal cells, 11 showed signal in tracheal branches, 9 of these could be mapped to specific genomic sites, and 7 genes were tagged, proving the feasibility of the technique. Appropriate controls are included to demonstrate that the GFP signal faithfully replicates the distribution of the counterpart endogenous mRNA. This technique promises to be a valuable one for studying mechanisms of mRNA localization and how it is coupled to translational control.

I have a few minor criticisms, however.

page 6: I think it is not valid to extrapolate the data to state that 140/223 lines for 'potentially productive'; the fact is that localized signals were not observed in those lines, but that could be because the detection method is insufficiently sensitive. The MS2-MCP system does not always produce bright signals even for known localized mRNAs, thus this system will produce false negatives, perhaps in significant numbers. This should be made clearer in the narrative.

Fig. 3: The differences between the images in panels B (supposedly unlocalized) and C (localized) are pretty subtle as there is fluorescence in the branches in panel B. How exactly was it decided that B is negative and C is positive? Maybe mRNA localization to tracheal branches is not an all-or-nothing situation?

Typos: Legend Fig. 1, remove apostrophe from ribosome's (l. 4), change µM to µm (micromoles to microns) in the legends to Figs. 1, 3, and 5.

Finally, it is too bad a counterpart UASp vector were not produced as this would be useful for studying RNA localization in early stages of oogenesis.

1st Revision - authors' response 24 June 2011

Referee #1:

Major points:

1. There is not a clear distinction between the cytoplasmic localization of the tagged mRNAs in figure 4A - C. This is particularly the case for B and C, which are classed as negative and positive, respectively. High quality images (as in D and E) should be shown for the lines in A and B. Quantification of representative examples would be welcome, if possible.

We assume this comment is about figure 3. The images were records from the original screen, taken at low resolution. We have now re-photographed similar larvae both at low and high magnification. It is apparent that even though the cell scored as 'negative' has some signal outside the nucleus, at higher magnification the appearance of the signal is dramatically different from that in cells scored as 'positive'. We have included additional images to show the distribution of the GFP-signal throughout a positive cell (Fig. 2H, I and suppl. Fig. S4).

2. Figure 4. The authors show that the MS2-GFP puncta in a positive line do not co-localize with Pacman. However, this is a very cursory characterization of the identity of these puncta. Several other components of RNA-associated bodies and transport complexes could be investigated for co-localization.

We performed these experiments to determine whether the observed puncta were simply stress granules due to overexpression or whether they might represent transport particles. Unfortunately, markers for many of the obvious candidates are GFP-tagged (staufen-GFP, GW182-GFP, Me31B-GFP) , so we cannot use them together with the MS2-GFP. Of the antibodies that exist, we tried staufen, but could not obtain a convincing signal. We were unable to obtain the published antibodies for polyA-binding protein. However, we have now included data for one further component, the decapping enzyme dcp1, for which antibodies exist. We detect Dcp1 in a subset of the particles which may indicate that some of them could be stress granules, as we now comment in the text (page 7).

3. It is mentioned that RNAi-mediated knockdown specifically in the tracheal system of several of the mRNAs found in the screen leads to lethality. Knockdown of one of them leads to tracheal defects. No data or description of the specific phenotype are shown for this particular line.

We have performed an extensive analysis of this gene and the cell biology of the phenotype, which goes far beyond its identification as a localized transcript and will be published elsewhere.

It is not shown that these mRNAs have a selective function in the tracheal system by assaying the requirements in other cell types.

We do not argue that these genes have a *selective* function in the tracheal system, and indeed there is evidence that this is not the case. Btsz is expressed and required in the embryonic and other epithelia. Similarly, lola, ATP-alpha and hsp70 are required in other tissues as well. Part of the rationale of developing this screen was in fact the problem that it is difficult to screen functionally at a late stage of development for genes that are required either earlier or in other places as well. So the point is not that the genes we identify are specific to tracheal cells, but that the RNA is present there and is found in the branches, where it is presumably normally needed.

Most importantly, the functional significance of the localization process itself is not tested.

Testing the functional significance is an important and long-term goal resulting from this screen. This is a major project, as it requires mapping the localization signals of the mRNA in question, which will require construction of numerous transgenic strains, the identification of a mutant phenotype and demonstration of rescue a by wild-type transgene, and finally analysis of both localization and phenotypic rescue by the candidate localization signal mutant transgenes. Such experiments are well beyond the scope of the current paper, which focuses on the screen method itself.

Minor points/questions:

1. Figure 2 and 3 can be combined, as the information in figure 2 does not justify its own figure.

We have done this, and added other modifications, as requested by the referees.

2. Figure 1 should be referred to in the results and discussion section, and not the introduction. It seems unnecessary to cite the first author's thesis as the data are shown in figure 1.

We have moved this part to the results section (page 3).

3. Figure 4A: the position of the nucleus is not clear. Also, given historical problems with in situ hybridization in the tracheal system, sense control probes for SRF or Dof could be helpful in providing further confidence in the technique.

The problems with in situ hybridisations using the histochemical staining in embryos are notorious and we have had much better success with FISH. We had of course nevertheless done the hybridizations with the sense strand control probes in parallel; the images are now included in the supplementary information.
We have now marked the nucleus in panel 4A.

Figure 4B: the position of the tracheal cell is not clear with the single labelling.

We had tried double labeling techniques, but these were not successful. We note that the red appearance of the lumen of the tracheal cells in some images is due to auto-fluorescence, which we observe frequently, but not in all cases. Unfortunately, we did not observe this in our *dof* in situ hybridisation experiments. The borders of the tracheal cell are, however, clearly visible when the intensity of the green channel is increased. This can be used to define the borders of the cell in the figure. We have included a supplementary figure showing this.

The arrows in C are not shown and the arrows in D are not referred to in the legend.

We have corrected this (page 2)

4. Only a very small number of the positives are followed up by in situ hybridization to assay for subcellular localization of the endogenous transcript. This is an important validation of the MS2 technique and the authors should either show the in situ results for more positives, or state in the text which ones were not informative due to expression in the surrounding tissue.

In situ hybridisations on the terminal cells in the larva are technically extremely challenging. The tracheal branches are very thin and most of them are tightly embedded in the surrounding tissue. We have tried various preparation and fixation protocols, but even probes that give reasonable signals in embryos often do not give reproducible and reliable results in the larval terminal cells. This difficulty was in fact one of the reasons that motivated the EP-MS2 screen. While we do not explicitly show that the RNAs we have identified are present in the branches, we present evidence that they are expressed in tracheal cells, either from RT-PCR (btsz, lola, hr39, CG30403), from their functional requirement (all but hr39), or from the presence of their protein product in the terminal cells (ATP-alpha).

5. Introduction: the MS2-GFP system is described as being developed recently, but the citation is 1998.

We have corrected this.

6. Figure 5. At least some components of the screen appear to be expressed in egg chambers with the UAST vector, which leads to very low levels of maternal expression. Might a significant proportion of the negatives be scored as such because the level of MS2-RNA expression is not high enough to be visualized (in order to help judge this, an image of MS2-CP alone in the oocyte (without an MS2 RNA) would be very important).

In the oocyte screen we used an MS2-GFP construct in the UASp vector. We apologize for having omitted this detail from the Materials and Methods section. This has now been corrected. The MS2-tagged RNA is not in the EP vector that was published by P. Rørth, but in another vector she made at the same time but which was not included in the 1996 PNAS publication. This alternative vector does not use the hsp70 minimal promoter, but instead the transposase minimal promoter. We have included information on the sequence of this part of the construct as supplementary information.
We have also included additional images of MS2-CP alone in the egg chamber (it is strongly expressed, see Fig S5G).

Also, not many details of the coat protein constructs are provided. What is the nuclear localization signal? Are there any other modifications compared to other versions?

We used exactly the configuration of the original version published by Rob Singer. This is now explicitly stated in the Materials and Methods.

7. Is cenG1a (tagged in line MS2-152a) endogenously expressed and localized in oogenesis? If so, this would provide further validation of the method.

We have done in situ hybridisations on ovaries with cenG1a probes and do not see staining above background. Thus, the gene is either not expressed, or it is expressed at levels we cannot detect by in situ hybridization.

8. Are the MS2-tagged RNAs translated normally? The presence of secondary structure or start/stop codons in the MS2 cassette may interfere with translation and in some cases this might have consequences on localization.

It is true that this might happen, but in the majority of known cases, translation is not required for mRNA localization.

To test this rigorously, we would need reagents to detect the protein product of the tagged gene, and be able distinguish it from the endogenous product. Or make tagged constructs of genes with and without an MS2 cassette inserted in the 5' end. Whether a single example would allow general conclusions is not clear, so it would be necessary to make such constructs for several genes.

If the screen misses transcripts for which translation is required for localization, this is likely to be a minority. There are many reasons why the screen is unlikely to be saturating screen (in the sense that it is able to test every single gene in the genome), but this is the case for almost all genetic screens. The value of the screen is that it allows us to find large numbers of genes with specific properties, even if we don't find *all* of them.

9. Previously it has been shown by Dr. Gavis and colleagues that a proportion of the MS2 coat protein can exit the nucleus without an MS2-tagged RNA in da neurons. Therefore the appearance of a representative negative EP-MS2 line in da neurons should be shown for comparison to figure 5C and D, as should the appearance of MS2 coat protein without an MS2 tagged mRNA.

The reviewer is correct that the MS2 coat protein fusion is not always confined to the nucleus in the absence of MS2-tagged RNA and can produce particles or a low level diffuse signal on its own. Our previous analysis of *nos* mRNA localization in da neurons established that we can reliably detect localized versus unlocalized transcripts (Brechbiel et al., 2008). Control animals lacking EP-MS2 tagged RNA were included in every imaging session and the line designated as positive showed a clear increase in particle accumulation in dendrites over the controls. We have added additional images to illustrate this (Fig. 4A, B).

10. In figure 5C and (in particular) D, why are there so many green dots outside of the da neuron (where the MS2 RNA is presumably not present with this GAL4 line)?

We presume that the reviewer is referring to the additional red signal present in the surrounding tissue. The *GAL4*^{477} driver we use for expression of MCP-RFP in da neurons is not as specific as previously reported and is also active in many CNS neurons. The semi-intact larval preparations we use for imaging most likely crush the CNS, producing fragments that cause this background. In addition, the *GAL4*^{477} driver is active at lower levels in other da neurons

11. Legend to figure 5: there is a small typo: microns, rather than micromolar, should be referred to.

We have corrected this.

Referee #2:

Despite the fairly extensive controls, it wasn't clear to me how important the site of P element insertion is, and at what level the negatives occur at. In figure 2, and the first part of the results, insertions in the 5'UTR are described and discussed. However, the proportion of P element insertions in 5'UTRs is generally quite small.

This may be a misunderstanding. By "5'" we mean 'at or near the transcription start', rather than "in the 5' UTR". Our diagram was probably responsible for this confusion, because it was inaccurate

(although we state our meaning correctly in the text). We have now modified the diagram to show the insertion of the element immediately upstream to the transcription start site – a site at which P elements are known to insert preferentially.

Of the 11 positive lines, how many were in the 5'UTR?

The information on the insertion sites is included in Table 1 (column 2).
Four of the nine insertions that we were able to map unambiguously were within a few nucleotides of the transcription start site as indicated in Flybase.
All three btsz insertions were in a hotspot that is, according to the current knowledge of the gene structure, at one of the many alternative transcription start sites for the various *btsz* transcripts. This transcription start is *within* exon8 of a transcript that starts several kb upstream. So this insertion is both at a transcription start and within an exon.
The insertion in CG30403 is about 1.7kb upstream of the predicted open reading frame for this gene. No cDNAs or ESTs are reported, so it is not know where the transcription start is. Our PCR data showed that the transcript actually does start close to the insertion site. A short exon is then followed by a long intron, followed by the annotated sequence for this CG. This information is included in the legend to table 1.
roadkill/CG9924 is a very large gene for which several transcription start sites have been identified. Our insertion is in a large intron, at another P-insertion hot spot. Whether an additional start site is present here is not known.

Thus, eight of the nine insertions are at or near known transcription start sites.

How often are other insertions (ie upstream? intron etc) expected to yield useful fusion proteins? I was not fully convinced that the majority of negative results are not simply due to lack of expression of full-length tagged mRNA production. Only three "negative" lines were examined for expression levels and 2 of the 3 looked quite weak in the PCR controls shown. These should also be done quantitatively in order to be convincing. Given that most localization elements appear to be in 3'UTRs, the distance between localization elements and tags may often be quite large, increasing the chances of separation. The PCR controls don't account for this either, as the gene specific primers used appear to be close to the 5' ends of the genes. These analyses should be repeated with larger numbers of genes, and more quantitative assays that also assess transcript integrity. These controls are needed to properly assess the usefulness of the system, and how well it reflects the frequency and types of mRNA localization. Given that many of the negatives observed may occur due to undetectable levels of GFP-tagged mRNAs, the authors should also consider a more sensitive assay on a number of their negatives. Fixing the tissues and staining for GFP immunologically is one possibility (FISH may also be a useful assay). A possibility for improving on the efficiency of this system would be to also add splice donor sites.

The referee raises many important issues which we had also considered during the course of our work. This would be worth discussing at more length in the paper, but since this is a short format, we had reduced the description and discussion to the minimal essential information. Rather than respond to questions point by point, we would like to summarize our reasoning for why we did not include more of our results or conduct further work and perform additional controls to analyse the 'negative' lines.

The central question is whether the negatives are true negatives because they do not localize to the structure under investigation, or whether they are simply not expressed.

'Not expressed' can mean two things:

- No transcript is made at all. This could be because the insertion is not upstream of a gene, or the insertion is near a gene but there may be a problem with transcribing the gene.

- the insertion is upstream of a gene, the targeted gene is expressible *in principle*, but no detectable mRNA is made in the tissue of interest.

It is very likely that there are some cases of the former type, as would be the case for any EP-insertion screen. Our controls were carried out to allow us to exclude these as the most obvious possible reasons for the majority of negative results. The mapping data indicate that insertions that are not near genes do occur, but not at the frequency that would explain all negatives (95% of the screened lines).

If we then assume that a large number of the lines can produce transcripts, the more interesting question is why a gene that is expressible in principle is scored as negative, and whether the reasons are biologically relevant. The biologically interesting reason would be that the mRNA is not transported to the region we score. Alternative reasons might be, as the referee suggests, that transcription levels are too low, or that the transcript cannot be completed so that it would not

include all localization elements. We feel that the strongest, and in our view conclusive, evidence that there is not a general problem with mRNA expression in the negative lines comes from the tissue specificity of our results. Thus, of the lines that were scored negative in the neurons, two (EP-MS2-12 and 67) were positive in tracheal cells (we have stated this more clearly in the revised version of the paper). These tagged genes are therefore expressible in principle, and are detectable in tracheal cells, but they are nevertheless negative in neurons. Conversely, EP-MS2-16 was negative in tracheal cells, but showed a positive signal in neurons. None of the lines that were positive in tracheal cells scored in oocytes, while the positive line EP-MS2-152a from the oocyte screen did not score in tracheal cells. We therefore have, even from this limited screen, evidence from many lines that the 'negative' score in a given tissue is not due to the general failure for a tagged transcript to be produced from the targeted gene.

An important question that arises from this finding is what the basis for this differential behaviour of the RNA may be. One possibility is that the mRNA is present in a part of the cell that we do not score (a likely explanation for some genes, e.g. those where the protein product is nuclear). Another may be that if the mRNA does not find its normal target when expressed in the 'wrong' cell, it is not processed correctly and is degraded. In this case, lower or no expression would be expected. This is consistent with the finding from our RT-PCR results that the 'negative' lines show apparently lower expression.

We do in fact have evidence that the same line that generates detectable transcript in one tissue shows not only low levels but no detectable transcript at all in another (though we did not see such an example in the tracheal cells). We had expressed some of the lines in the mesoderm under the control of Twist-GAL4 and tested for the presence of transcripts (we did not score for MS2-GFP localization, because mesodermal cells are so small that there is insufficient cytoplasm surrounding the nucleus to visualize differential localisation). For three lines we tested (insertions in btsz, lola and bif), we can detect only btsz in the mesoderm, although all three are easily amplifiable from mRNA in tracheal cells. Thus lola and bif GAL4 induced transcripts seem to be degraded in mesodermal but not in tracheal cells. We have included this result with these revision notes (see point 2 of additional data for referees). These findings raise many interesting questions, which will be difficult to address and answer, and will be subjects for future students or postdocs in our labs. It is for these reasons, combined with the evidence of tissue-specific detectability of MS2-tagged transcripts, that we did not continue with a larger scale analysis of the negative lines in tracheal cells. We have rewritten parts of the manuscript to make these points more clear.

Given that many of the negatives observed may occur due to undetectable levels of GFP-tagged mRNAs, the authors should also consider a more sensitive assay on a number of their negatives. Fixing the tissues and staining for GFP immunologically is one possibility (FISH may also be a useful assay).

The oocyte screen was done on fixed and immuno-stained material. We included this information in the revised manuscript. For the other tissues this was not necessary. The tagged genes are over-expressed with the UAS-GAL4 system, so the transcription levels are not related to the expression levels of the endogenous gene. For a large-scale screen, we had to balance efficiency against sensitivity. Filleting, fixing and immuno-staining third instar larvae in order to find possible low-signal positives would reduce the throughput at least five-fold.

A possibility for improving on the efficiency of this system would be to also add splice donor sites.

An interesting idea, and we have discussed this as a modification for carrying out a full-scale screen. But we abandoned it for a number of reasons. The existing vectors for this purpose rely on the endogenous promoter for expression of the tagged gene. We were concerned that we would miss low level transcripts if we did not use the UAS system for high-level expression. If we use the UAS-system together with a splice donor, an insertion upstream of the first exon would lead to skipping the first exon, by splicing the donor sequence in the inserted construct to the first acceptor, i.e. the second exon. If there are localization elements in the 5'UTR or coding region they might be deleted. The same would be true for an internal insertion.

More minor points:

I found it somewhat surprising that only one type of localization was observed in this study. Perhaps this is due to the small number of positives and the relative frequency of this type of localization.

In the tracheal cells, we screened for presence of signals in branches. We would have been excited to find differential expression within the branches, but we do not know enough of the physiology of the system to know whether we should have expected such localization patterns (part of the purpose of the screen was to find this out). In neurons too, rather than being concentrated at a particular subcellular location, the RNA particles are distributed along the length of the processes, giving more

of an asymmetric rather than tightly localized distribution. Similar results have been obtained for ß-actin mRNA in developing axon.

The high correlation between expression of endogenous genes and reporters likely points to accessibility of P elements inserted in tracheally expressed genes - little or no expression would be expected if the chromatin were in an inaccessible state. This may also explain, in part, why transcripts that localized in one tissue did not appear to localize in others (p8). They simply may not be inducible in the other tissues.

This had occurred to us, too, However, this is not consistent with what is generally known about the GAL4 system – the general experience is that pretty much any UAS-driven line can be expressed in any tissue and cell type. Nevertheless, the reviewer's explanation may well account for some of the cases. This will need to be tested in a future large-scale analysis of results from a planned multi-tissue screen.

The authors interpretation that localization signals and machineries may be very different in different cell types does not jive well with data published to date, which suggest strong overlaps (at least for microtubule based mechanisms).

We agree that the core machinery (e.g. transport by microtubules) is almost certainly shared, but how RNAs are packaged so that they can associate with the correct motor is not yet clear in most cases. In particular, not much is known about localization signals and their cell type specificities, which is one of the points of the screen. We have modified our statement in the text as follows: "Thus, localization signals might be recognized in a cell-type specific manner arguing for cell type specific molecules that interact with or form components of the mRNA localization machinery.

The manuscript would be strengthened further if the authors show whether proteins encoded by localized transcripts are also localized.

For most of the proteins the necessary reagents are not available. However, we obtained antibodies against ATP-alpha and now show that the protein is expressed throughout the terminal cell. We also include stainings for Dof, for which we show an in situ hybridization, and for which the protein is also distributed throughout the branches.

Referee #3:

I have a few minor criticisms, however.

page 6: I think it is not valid to extrapolate the data to state that 140/223 lines for 'potentially productive'; the fact is that localized signals were not observed in those lines, but that could be because the detection method is insufficiently sensitive. The MS2-MCP system does not always produce bright signals even for known localized mRNAs, thus this system will produce false negatives, perhaps in significant numbers. This should be made clearer in the narrative.

We have changed this paragraph. It now states : "Thus up to four of the 11 cases may have been unproductive insertions, although it cannot be excluded in these cases that the transposon has inserted upstream of hitherto unannotated genes, or within additional unidentified upstream promoters of known genes. This means that up to 40% of the negatives may be due to non-productive insertions" (page 5). The following paragraph then deals with other possible reasons for negative results (see also discussion in response to referee 2).

Fig. 3: The differences between the images in panels B (supposedly unlocalized) and C (localized) are pretty subtle as there is fluorescence in the branches in panel B. How exactly was it decided that B is negative and C is positive? Maybe mRNA localization to tracheal branches is not an all-or-nothing situation?

We have added high-resolution images that address this point. Also see discussion above (point 1 of referee 1).

Typos: Legend Fig. 1, remove apostrophe from ribosome's (l. 4), change µM to µm (micromoles to microns) in the legends to Figs. 1, 3, and 5.

Thank you, corrected

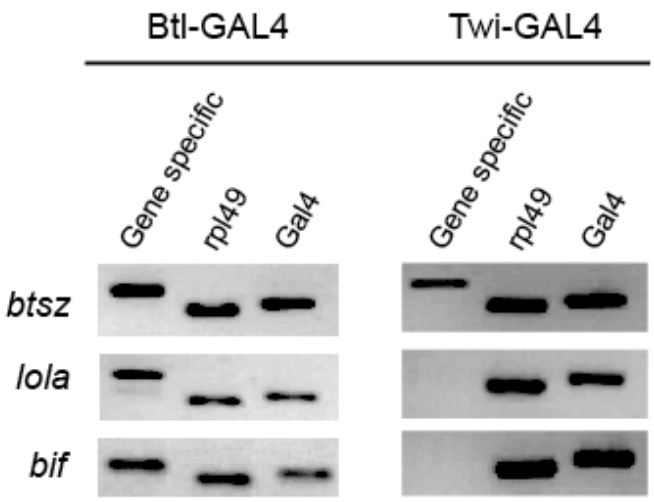
Finally, it is too bad a counterpart UASp vector were not produced as this would be useful for studying RNA localization in early stages of oogenesis.

We have made this construct and are happy to distribute it.

Additional data for the referees

1. RT-PCR analysis of tagged transcripts in the mesoderm

RT-PCR products from the indicated transgenes expressed either in the tracheal system of 3rd instar larvae (btl-GAL4) or the embryonic mesoderm (twi-GAL4). The expression of two lines that scored positive in the trachea (EP-MS2-67: *btsz;* EP-MS2-18: *lola*) and one negative line (EP-MS2-3: *bif*) was compared in trachea and mesoderm. While RNA from all three was detected in the tracheal system, only btsz seemed to be expressed in the mesoderm. RNA localization was not scored in the mesoderm, because the cells are too small and the nuclear staining cannot be distinguished from low level staining in the surrounding cytoplasm.
RT-PCR for *rpl*40 was carried out as a control for endogenous mRNA, and RT-PCR of Gal4 as a control for the expression of the driver.



2nd Editorial Decision 18 July 2011

Many thanks for the submission of your revised manuscript to our editorial office. The study was sent back to two of the original referees and we have now received their comments on it (copied below). Referee #3 only brings up some minor issues that I would kindly ask you to address. Referee #2 acknowledges that most of his/her concerns have been addressed. However, s/he still thinks that two remaining concerns (the sensitivity of the screen and the large number of negative hits) would need to be further discussed and I encourage you to do so, maybe along the lines of your point-by-point response. S/he also points out that the possibility of tissue-specific effects of EP-MS2 insertion sites should be discussed as an alternative explanation for why the same construct is differentially expressed and localized in different tissues.

Another suggestion brought up during the review process was to test whether the positively scored mRNAs are also endogenously expressed in trachea. I appreciate that this was done for five of the seven positive lines (either in this study or in earlier reports), but it looks as if no information is available for the nuclear hormone receptor hr39 and clone CG30403. I would therefore propose to check the expression of these two constructs by RT-PCR and include the results of this experiment in the final version of your manuscript.

As a final point, one of the referees had proposed to show the results on the tracheal defects observed in btsz knock-down cells and you included these data in your point-by-point response as information for the referees. I understand that you are planning to publish these results in a separate

paper. However, I encourage you to include a general description of the phenotype here and show the results, as this would significantly strengthen the study. This would not lead to a significantly longer study, as it would only add one rather small figure panel. If you decide not to include these results, you should not mention them at all, other than saying that this is something that would need to be done in the future, as we discourage the use of 'data not shown'.

Please modify your manuscript accordingly and submit the final version through our website again. On a more formal note, I noticed that you have moved the 'materials and methods' section to the supplementary information. I appreciate that this was done to shorten the text, but I would kindly ask you to include the most relevant methods in the main body of the manuscript before acceptance. To accommodate this and to make space for the slightly extended discussion along the points brought up by referee #2, I can extend the character limit to be slightly above 30,000 characters, as your study is still conceptually short. I also noticed that the arrows that mark the transcriptional start site in figure 2A are almost invisible. May I ask you to enlarge them in the final version of the figure?

I look forward to receiving the final version of your study when it is ready.

Yours sincerely,

Editor
EMBO reports

REFEREE REPORTS:

Referee #2:

The authors have done a good job of dealing with the majority of reviewer critiques and requests. The major remaining issues are the sensitivity of the screen, number of false positives and its relatively small scale. All three reviewers touched on all of these concerns in some form or other. The sensitivity issue is well demonstrated by the lack of variety or discernibility in mRNA localization types, and by the very low frequency of oocyte expression patterns, despite the fact that the majority of genes are maternally expressed and deposited in the oocyte. This means that this approach will only be able to expose a small proportion of localized transcripts. This will be further impacted by the number of genes that are resistant to P element insertions, and those that do not yield full length or abundant messages. That said, it does appear that this approach can identify mRNAs with interesting localization patterns, with some advantages over other approaches. Hence, I think that it is worth publishing in EMBO Reports. The caveats listed here do need to be made clearer though. Certainly, there should be no implication that the types or frequency of localization patterns identified using this approach will be reflective of all those that actually take place.

One more point. Although not generally commented on in the literature, the insertion sites of UAS reporters most definitely have an effect on tissue-specific expression. This is one of the reasons that separate RNAi knockdown lines carrying the same construct have different effectiveness. Furthermore, one construct that works relatively poorly in one tissue, can work beautifully in another. This becomes more common as tissues become more and more differentiated, and more genes become differentially silenced. I still believe this is the best explanation for the differential expression and localization of different EP-MS2 lines in different tissues, and should therefore be suggested prior to suggesting different localization machineries. In support are the CG9924, lola and nos transcripts that localize in more than one tissue, likely using similar cis elements and trans machineries.

Referee #3:

My comments have been adequately addressed in the revised version. Some minor presentational errors remain:

p. 4, l. 3: ensure, not insure
p. 4, l. 26: lens is misspelled
p. 5, l. 10: score, not scores
p. 5, l. 13: should read "inserted in the wrong" (word missing)
p. 8, l. 15: EP MS2-16 not EP-MS2 16

More importantly, the legend to Figure 2 doesn't refer correctly to the panels. I think it should read:

F-K (not F-I). High resolution.....positive lines (H,I) (not I,J).

2nd Revision - authors' response 10 August 2011

Revision notes

Another suggestion brought up during the review process was to test whether the positively scored mRNAs are also endogenously expressed in trachea. I appreciate that this was done for five of the seven positive lines (either in this study or in earlier reports), but it looks as if no information is available for the nuclear hormone receptor hr39 and clone CG30403. I would therefore propose to check the expression of these two constructs by RT-PCR and include the results of this experiment in the final version of your manuscript.

This has been done and the data are included in Figure 3.

As a final point, one of the referees had proposed to show the results on the tracheal defects observed in btsz knock-down cells and you included these data in your point-by-point response as information for the referees. I understand that you are planning to publish these results in a separate paper. However, I encourage you to include a general description of the phenotype here and show the results, as this would significantly strengthen the study. This would not lead to a significantly longer study, as it would only add one rather small figure panel. If you decide not to include these results, you should not mention them at all, other than saying that this is something that would need to be done in the future, as we discourage the use of 'data not shown'.

We no longer mention the tracheal phenotype but only refer to the lethality of all genes. This still makes the point that there is a function, but does not make any statement about a specific subcellular phenotype.

Please modify your manuscript accordingly and submit the final version through our website again. On a more formal note, I noticed that you have moved the 'materials and methods' section to the supplementary information. I appreciate that this was done to shorten the text, but I would kindly ask you to include the most relevant methods in the main body of the manuscript before acceptance.

We have moved the important M and M back into paper, the other parts into table legends or supplementary figure legends.

To accommodate this and to make space for the slightly extended discussion along the points brought up by referee #2, I can extend the character limit to be slightly above 30,000 characters, as your study is still conceptually short.

We are now at 30485 characters.

I also noticed that the arrows that mark the transcriptional start site in figure 2A are almost invisible. May I ask you to enlarge them in the final version of the figure?

We have corrected this.

Referee #2:

The sensitivity issue is well demonstrated by the lack of variety or discernibility in mRNA localization types, and by the very low frequency of oocyte expression patterns, despite the fact that the majority of genes are maternally expressed and deposited in the oocyte. This means that this approach will only be able to expose a small proportion of localized transcripts. This will be further impacted by the number of genes that are resistant to P element insertions, and those that do not yield full length or abundant messages. The caveats listed here do need to be made clearer though. Certainly, there should be no implication that the types or frequency of localization patterns identified using this approach will be reflective of all those that actually take place.

We have included statements to this effect.

One more point. Although not generally commented on in the literature, the insertion sites of UAS reporters most definitely have an effect on tissue-specific expression. This is one of the reasons that separate RNAi knockdown lines carrying the same construct have different effectiveness.

Based on experience in our own labs, we recognize that such tissue-specific influences of sequences near the insertion site can occur but believe that these are likely the minority of cases. Notably, there are many reasons why an RNAi construct might work better in one cell type than another (difference in abundance or accessibility of the endogenous mRNA, tissue-specific differences in RNAi machinery, etc.).

We now mention the possibility that in some cases the observed tissue-specificity of localization could be due to a tissue-specific influence of the insertion site of the EP-MS2 insertion.

Referee #3:

My comments have been adequately addressed in the revised version. Some minor presentational errors remain.

All corrected.

3rd Editorial Decision 15 August 2011

I am very pleased to accept your manuscript for publication in the next available issue of EMBO reports.

Thank you for your contribution to EMBO reports and congratulations on a successful publication. Please consider us again in the future for your most exciting work.

Yours sincerely,

Editor
EMBO Reports

