THE EMBO JOURNAL

Manuscript EMBO-2011-78136

Extraordinary transgressive phenotypes of hybrid tomato are influenced by epigenetics and small silencing RNAs

Padubidri V. Shivaprasad, Ruth M. Dunn, Bruno A.C.M. Santos, Andrew Bassett and David C. Baulcombe

Corresponding author: David C. Baulcombe, University of Cambridge

Review timeline:

Submission date:	11 May 2011
Editorial Decision:	16 June 2011
Revision received:	03 August 2011
Editorial Decision:	07 September 2011
Revision received:	09 November 2011
Editorial Decision:	14 November 2011
Accepted:	16 November 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 16 June 2011

Thank you for submitting your manuscript for consideration by The EMBO Journal and I apologise that it has taken longer than usual to have it reviewed. However, it has been now been evaluated two referees and I enclose their reports below. As you will see the referees find the study to be interesting, however, they do have a number of important concerns and require significant further analysis of the transgressive loci. Given the interest in the study should you be able to address these concerns we would be happy to consider a revised version of the manuscript for publication in The EMBO Journal.

I should remind you that it is EMBO Journal policy to allow a single round of revision only and that, therefore, acceptance or rejection of the manuscript will depend on the completeness of your responses included in the next, final version of the manuscript. When you submit a revised version to the EMBO Journal, please make sure you upload a letter of response to the referees' comments. Please note that when preparing your letter of response to the referees' comments that this will form part of the Review Process File, and will therefore be available online to the community. For more details on our Transparent Editorial Process initiative, please visit our website: http://www.nature.com/emboj/about/process.html

Thank you for the opportunity to consider your work for publication. I look forward to your revision.

Yours sincerely,

Editor
The EMBO Journal

REFEREE COMMENTS

Referee #1:

This manuscript presents a very interesting hypothesis, but the evidence for it is not as well developed as one would hope.

The authors propose that siRNAs, by their ability to trigger secondary siRNAs, could lead to a new type of epistatic interaction between genomes, where the original trigger becomes indispensable in subsequent generations. Some of the data are in support of this idea, others are not.

Apart from the individual criticisms below, a major concern is that the authors use a reference genome sequence that lacks an associated publication. Thus, the quality of this not peer-reviewed genome sequence is unknown (plus the link given for download was broken). Moreover, looking through the http://solgenomics.net/ website, it is unclear in how far the authors have coordinated their analysis with the colleagues who generated the genome assembly (they do not appear in the acknowledgement).

Turning to the manuscript at hand, an important concern related to the genome sequence is that two species are being used, but genome sequence information (with the caveat mentioned above) is available for only one of them. While the authors seem to be reasonably cautious in their interpretation, this makes it difficult to properly analyze the data in hybrids. In addition, the authors seem to assume that the introgression lines are indeed "clean", i.e., contain only the known major introgressed segments. However, since these lines were only characterized with relatively low marker density (at least this is what I seem to remember), additional smaller regions of the foreign genome might contribute to the observed effects.

Regarding the Results section, the first major part, on miR395, seems out of place, and I would suggest to leave it out altogether. I believe that it is too uncertain that the observed effects are caused by trans-acting small RNAs. An explanation that seems to be at least as simple is that M82 contains a quantitatively acting transcriptional repressor of miR395, and that the M82 locus in the absence of this repressor is simply more active than the S. pennellii locus. Also, the miR395 analysis would be a lot clearer, if the authors would let us know how many miR395 loci there are in tomato, and whether perhaps miRNA* reads illuminate whether all loci are affected.

The second part of the manuscript presents siRNA loci that show transgressive behaviour in later generations. Unfortunately, even though there are over 100 such loci, the authors focus on only two. There definitely needs to be more analysis - how are they distributed over the genome, how large are they, are the siRNAs normally produced from both strands, how do they overlap with features such as TEs, transcribed mRNAs, etc. etc.

Two loci are analyzed in some detail, but again, important points are not covered. For example, the authors present RT-PCR data for "H06 mRNA". How do they know they are looking at mRNA, if they are not performing strand-specific analysis? Also, real time quantification is essential and the methylation analysis at this locus needs to be supported by BS-sequencing of PCR products.

Similarly, the analyses of the PAL loci is intriguing, but is strongly confounded by not knowing the complete number and arrangement of PAL genes in the two parental species. Furthermore, the authors do not seem to distinguish between siRNAs that map uniquely and ones that map to multiple places in the cluster, which again makes it difficult to interpret the presented data.

Minor points:

Intro: The discussion of epistasis makes it sound like the different scenarios are merely theoretical. They are not; one of the nicest example in plants is that of FLC/FRI, where two early flowering lines

in A. thaliana can produce F1 hybrids and segregate later-generation progeny that flower very late.

The description of the analyzed material belongs at the beginning of the Results (which start without even mentioning the species that have been analyzed).

Small RNA mapping: I assume only perfect mappings?

References are missing (e.g., Ha et al., 2009; Groszmann et al, 2011).

Fig. 3a: give annotation (something like "loci") for the "Y axis" of the heat map. Also, red and green is an unfortunate choice for those of us who are colour blind.

Fig. 3b: red and blue small RNAs are explained, but what are green ones?

Fig. 5a: see comments on Fig. 3a.

Fig. 6a: "significant phasing": this needs to be supported by statistics.

Referee #2:

The manuscript by Shivaprasad et al. contains a potentially interesting set of data concerning changes in abundance of various classes of small RNAs observed in tomato introgression lines (ILs). However rather surprising that these data are not appropriately presented, better characterized and thoroughly discussed. The ILs analyses seem to serve only as controls for populations of small RNAs of F1 hybrids and F2 segregants that potentially should reveal differences that can be assigned to hybrid vigor phenomena. Unfortunately, in the F1 and its progeny nothing interesting could be documented. In consequence the way the manuscript is currently written is inconsistent with the experimental results and very confusing. Focusing on ILs would allow presenting and describing them properly. For example a set of ILs of chromosome 8 cover the same region (e.g. Il8-1-3 and Il8-1-1) however the accumulation of miR395 significantly differs. A thorough explanation or discussion would be necessary. There are a number of related examples throughout the manuscript, and therefore, a very serious refocusing and rewriting of the manuscript to properly consider the interesting ILs data is a prerequisite before resubmission. Since the F1 and F2 data are largely negative, these could serve as controls for the ILs. Speculative suppositions, such as that the observed changes detected in ILs were likely initiated in their F1, should be omitted.

Minor points:

1.Although the introduction is rather long miRNAs are not mentioned.

2.Abstract : "we identified loci from which these sRNAs were more abundant in hybrids than either parents, and show that this accumulation correlated with suppression of the corresponding target genes" is in opposition with the 1st part of the results (See point 3).

3.For the statement "None of the sRNA loci were expressed beyond parental range in the F1 hybrids." data should be presented. However, one of the sRNA loci was shown to be expressed beyond the parental range in the F1 hybrids - levels of siRNAs of PAL5A seem to be transgressive (Supplementary fig2C). This should be properly acknowledged and discussed to avoid confusions.

1st Revision - authors' response 03 August 2011

Referee #1:

This manuscript presents a very interesting hypothesis, but the evidence for it is not as well developed as one would hope.

The authors propose that siRNAs, by their ability to trigger secondary siRNAs, could lead to a new type of epistatic interaction between genomes, where the original trigger becomes indispensable in subsequent generations. Some of the data are in support of this idea, others are not.

Apart from the individual criticisms below, a major concern is that the authors use a reference genome sequence that lacks an associated publication. Thus, the quality of this not peer-reviewed genome sequence is unknown (plus the link given for download was broken). Moreover, looking through the http://solgenomics.net/ website, it is unclear in how far the authors have coordinated their analysis with the colleagues who generated the genome assembly (they do not appear in the acknowledgement).

Author's reply:
The tomato genome is not complete and is not yet published. However, the available genomic sequence as well as EST sequences have been used for mapping small RNAs from tomato by various groups (for example, Mohorianu et al., In Press, Plant J., 2011), as well as for various other genome wide studies. We have provided two references from the colleagues who are generating the tomato sequences. While we have given the source of the BAC sequences (SGN) in our previous manuscript, now we have also acknowledged SGN and the colleagues involved in tomato sequencing in the revised manuscript. As the genome is upcoming and rapidly developing especially in the last few months, we have given the source of sequence and omitted links (Section 'acknowledgements').

Turning to the manuscript at hand, an important concern related to the genome sequence is that two species are being used, but genome sequence information (with the caveat mentioned above) is available for only one of them. While the authors seem to be reasonably cautious in their interpretation, this makes it difficult to properly analyze the data in hybrids. In addition, the authors seem to assume that the introgression lines are indeed "clean", i.e., contain only the known major introgressed segments. However, since these lines were only characterized with relatively low marker density (at least this is what I seem to remember), additional smaller regions of the foreign genome might contribute to the observed effects.

Author's reply:
The absence of completely assembled genome for both parents has limited the extent to which we can interpret our data. However, we have taken a conservative approach by aligning all data to the tomato genome sequence and we have analysed loci that are specific to the tomato genome or that are highly conserved in tomato and *S. pennellii.* At present we cannot analyse loci that are specific to *S. pennellii.* However this limitation and the possibility that the IL lines are not clean does not invalidate our interpretation that some loci are transgressively expressed (ie present at higher levels than in either parent). It means that we underestimate the number of transgressive loci.

Regarding the Results section, the first major part, on miR395, seems out of place, and I would suggest to leave it out altogether. I believe that it is too uncertain that the observed effects are caused by trans-acting small RNAs. An explanation that seems to be at least as simple is that M82 contains a quantitatively acting transcriptional repressor of miR395, and that the M82 locus in the absence of this repressor is simply more active than the S. pennellii locus. Also, the miR395 analysis would be a lot clearer, if the authors would let us know how many miR395 loci there are in tomato, and whether perhaps miRNA* reads illuminate whether all loci are affected.

Author's reply:
The possibility that there would be hybrid specific initiation of trans-acting siRNA was the rationale for our analysis. However, although we do find hybrid-specific siRNAs, our conclusion is that the mechanisms involved go beyond the tasiRNA pathway. The H06 locus illustrates epigenetic mechanisms and the miR395 data illustrate another process that is not fully understood. It could be, for example, that there is a suppressor of miR395 encoded in the introgressed IL8-1-3 locus and that the *S. pennelli* orthologue is less effective than the tomato homologue. This explanation would explain why miR395 is more abundant in *S. pennellii* than in tomato. The transgressive effect would result if the *S. pennellii* repressor is less effective

with the tomato rather than the *S. pennellii* versions of miR395. According to this idea there would be transgressive miR395 in the F2 lines with homozygous repressor loci from the IL8-1-3 region. However there could be other explanations and we thought that there would be interest in the phenomenon as described because there is a link of miR395 to an important trait (salt tolerance).

Arabidopsis has 6 copies of miR395 with 2 unique sequences. They occur in clusters. Tomato has 18 copies that are clustered in Chromosome 5, 3 and 2 as mentioned in the manuscript, but not in Chromosome 8. However we cannot identify whether any one locus is transgressively expressed because the miRNAs are identical from each locus. There are sequence differences in the miRNA* sequences but we have not detected them in our datasets.

Reply to comments Table 1.

Copy	Chromosome region	Start	Stop
1	SL2.40ch05	1703003	1703023
2	SL2.40ch05	1703180	1703200
3	SL2.40ch05	1705348	1705368
4	SL2.40ch05	1708765	1708785
5	SL2.40ch05	1713756	1713776
6	SL2.40ch05	1716672	1716692
7	SL2.40ch05	1716848	1716868
8	SL2.40ch05	1705688	1705708
9	SL2.40ch05	1708941	1708961
10	SL2.40ch05	1713580	1713600
11	SL2.40ch03	30608183	30608203
12	SL2.40ch03	30608715	30608735
13	SL2.40ch03	30608537	30608557
14	SL2.40ch02	27508543	27508523
15	SL2.40ch02	27508725	27508705
16	SL2.40ch02	27526695	27526675
17	SL2.40ch02	27533108	27533088
18	SL2.40ch02	27501989	27501969

The second part of the manuscript presents siRNA loci that show transgressive behaviour in later generations. Unfortunately, even though there are over 100 such loci, the authors focus on only two. There definitely needs to be more analysis - how are they distributed over the genome, how large are they, are the siRNAs normally produced from both strands, how do they overlap with features such as TEs, transcribed mRNAs, etc. etc.

Author's reply:
We have included the following additional analysis: Two new figures (Supplementary Figures S5 and S6) with data on transgressive loci such as their distribution among chromosomes, strand bias among the sRNAs, average length of such transgressive loci in comparison to all loci and their source (Please see Supplementary Figures S5 and S6; last paragraph on Results section 'Transgressive siRNA loci').

Two loci are analyzed in some detail, but again, important points are not covered. For example, the authors present RT-PCR data for "H06 mRNA". How do they know they are looking at mRNA, if they are not performing strand-specific analysis? Also, real time quantification is essential and the methylation analysis at this locus needs to be supported by BS-sequencing of PCR products.

Author's reply:
We have now provided qPCR data for the 'H06 mRNA' (Figure 4B). We have also performed RT-PCR with gene-specific primers and have similar results which confirm the strand-specificity. We do not have bisulphite sequencing data at the moment as we plan for genome-

wide methylation analysis. However, the McrBC methylation analysis has been performed by two individuals with biological replicate samples.

Similarly, the analyses of the PAL loci is intriguing, but is strongly confounded by not knowing the complete number and arrangement of PAL genes in the two parental species. Furthermore, the authors do not seem to distinguish between siRNAs that map uniquely and ones that map to multiple places in the cluster, which again makes it difficult to interpret the presented data.

Author's reply:
We have identified 4 PAL sequences from Chromosome 9 with an unusual genome structure (repeats separated by a transposon). We identified sRNAs matching to unique sequences in each of these four PAL genes (Supplementary Table SV – included in the previous version) indicating that each of them is a source of transgressive siRNAs and silenced correspondingly in the IL8-3. There may be more PAL loci in the tomato genome but our RT-PCR data indicate that they are all suppressed.

Minor points:

Intro: The discussion of epistasis makes it sound like the different scenarios are merely theoretical. They are not; one of the nicest example in plants is that of FLC/FRI, where two early flowering lines in A. thaliana can produce F1 hybrids and segregate later-generation progeny that flower very late.

Author's reply:
We are aware of many examples of epistasis including the one cited. However the scenarios we describe invoke epistasis mediated by RNA that would be novel. We have presented these explanations as plausible explanations of our results. However we have added additional references (Werner et al., 2005, Caicedo et al., 2004 and Rowe et al., 2008) in the appropriate place (Introduction section, 5th paragraph).
The description of the analyzed material belongs at the beginning of the Results (which start without even mentioning the species that have been analyzed).

Author's reply:
We have added names of lines used for sequencing at the start of 'Results' section.

Small RNA mapping: I assume only perfect mappings?

Author's reply:
We have this information in the Bioinformatics analysis section. Indeed they are only perfect matches.

References are missing (e.g., Ha et al., 2009; Groszmann et al, 2011).

These references are in the discussion part (7th paragraph). We have concentrated on transgressive features that arise at later generations and hence did not discuss these interesting papers in the Introduction section.

Fig. 3a: give annotation (something like "loci") for the "Y axis" of the heat map. Also, red and green is an unfortunate choice for those of us who are colour blind.

Author's reply:
Thank you for this suggestion and apologies for overlooking this information. We have changed the colours and also added label for y axis (Figure 3A).

Fig. 3b: red and blue small RNAs are explained, but what are green ones?

Author's reply:
They are 22-nt in size. This information has been added in the appropriate Figure legends (Figure legend 3B).

Fig. 5a: see comments on Fig. 3a.

Author's reply:
We have changed the colours and also added label for y axis.

Fig. 6a: "significant phasing": this needs to be supported by statistics.

Author's reply:
We have provided *p*-values for phasing both in the figure (Fig. 6A) as well as in Figure legend.

Referee #2:

The manuscript by Shivaprasad et al. contains a potentially interesting set of data concerning changes in abundance of various classes of small RNAs observed in tomato introgression lines (ILs). However rather surprising that these data are not appropriately presented, better characterized and thoroughly discussed. The ILs analyses seem to serve only as controls for populations of small RNAs of F1 hybrids and F2 segregants that potentially should reveal differences that can be assigned to hybrid vigor phenomena. Unfortunately, in the F1 and its progeny nothing interesting could be documented. In consequence the way the manuscript is currently written is inconsistent with the experimental results and very confusing. Focusing on ILs would allow presenting and describing them properly. For example a set of ILs of chromosome 8 cover the same region (e.g. Il8-1-3 and Il8-1-1) however the accumulation of miR395 significantly differs. A thorough explanation or discussion would be necessary. There are a number of related examples throughout the manuscript, and therefore, a very serious refocusing and rewriting of the manuscript to properly consider the interesting ILs data is a prerequisite before resubmission. Since the F1 and F2 data are largely negative, these could serve as controls for the ILs. Speculative suppositions, such as that the observed changes detected in ILs were likely initiated in their F1, should be omitted.

Author's reply:
This review has misread our manuscript and we do not believe this is due to poor presentation because the first reviewer understood our points correctly. However we have redrafted parts of the script in an attempt at clarification. What we have done is to present observations based on IL lines. The first question was whether any of these lines exhibit transgressive expression of sRNAs. The answer is yes – as presented – but for reasons associated with the incomplete genome sequence we cannot say how many transgressive loci there are. The next question was as to whether any of these loci were transgressively expressed in F1 or F2 lines. Thus the F1 and the F2 lines are, as suggested, a control for the analysis of ILs. Our analysis is not and was not intended to be an analysis of heterosis.

We are currently attempting to map the determinants of the transgressive effects within the ILs but this is not a trivial task. There is overlap, for example of the introgressed regions between IL8-1-1 and IL8-1-3, but it is several megabases long and the sequence is not available. The available resources therefore do not allow us to map the determinants of transgressivity at high level resolution. We are currently using sub ILs for higher resolution work but this is part of a long term project.

Minor points:

1. Although the introduction is rather long miRNAs are not mentioned.

Author's reply:

We have now included more extensive reference to miRNAs (Introduction Section).

2. Abstract: "we identified loci from which these sRNAs were more abundant in hybrids than either parents, and show that this accumulation correlated with suppression of the corresponding target genes" is in opposition with the 1st part of the results (See point 3).

Author's reply:
The introgression lines are hybrids. We do not say that the overexpression is in all hybrids. To address this point we have reworded the text .....'were more abundant in some hybrids than either parents and we show that this sRNA accumulation correlated with' (abstract).

3. For the statement "None of the sRNA loci were expressed beyond parental range in the F1 hybrids." data should be presented. However, one of the sRNA loci was shown to be expressed beyond the parental range in the F1 hybrids - levels of siRNAs of PAL5A seem to be transgressive (Supplementary Fig. 2C). This should be properly acknowledged and discussed to avoid confusions.

Author's reply:
We provide data that, using the stringent criteria employed in this analysis, there was little or no expression of sRNA loci beyond parental range in the F1s (Supplementary Table SII and Supplementary Table SIII). It remains possible that there are loci that are transgressively expressed in the F1 that would be identified by less stringent criteria and we have revised the text accordingly to indicate that our statements apply only to the most extreme examples of transgressivity (Second paragraph, results section 'Transgressive siRNA loci'; Methods section 'Bioinformatics').

The PAL derived small RNAs in F1 are slightly high in one of the PAL ESTs (Supplementary Figure 2C) as observed by the Reviewer. However, the transgressive index was only +2.8 in the F1 (Supplementary Table SIII and SIV) was much less than in IL8-3 (values varying between +90 to +200). We have been careful to focus this initial report on the few loci with the highest transgressive effect. However the wording is now revised to acknowledge this small effect.

2nd Editorial Decision 07 September 2011

Thank you for submitting your revised manuscript for consideration by the EMBO Journal. I apologise that it has taken longer than usual to have the manuscript reviewed but this was due to one of the referees being unavailable for a significant period of time, however, it has now been re-evaluated by referee #1. Overall the referee remains positive regarding publication but requests that a number of issues are satisfactorily addressed before publication in The EMBO Journal, based in this recommendation I would like to invite you to submit a revised version of the manuscript.

When preparing your letter of response to the referees' comments, please bear in mind that this will form part of the Review Process File, and will therefore be available online to the community. For more details on our Transparent Editorial Process, please visit our website:
http://www.nature.com/emboj/about/process.html

Thank you for the opportunity to consider your work for publication. I look forward to your revision.

Yours sincerely,

Editor
The EMBO Journal

REFEREE COMMENTS

Referee #1:

I have seen the manuscript before and it is improved. However, I continue to have several comments.

I still think that the miR395 part is distracting, because it is likely an siRNA-independent form of epistasis (as the authors now explicitly discuss). Moreover, the additional explanations regarding miR395 loci and isoforms were not particularly helpful. As is, it looks like this part of the manuscript has been retained to pad the story. I very strongly advise to remove this, and instead move some of the SOM figures into the main text.

I had several questions regarding the annotation and distribution of transgressive loci, and these have only been partially answered.

1. Genome-wide distribution: The number of sRNA loci per chromosome is very uneven (Fig. S5A), even though the chromosomes themselves vary less than twofold in length. The good news is that the number of transgressive loci is approximately proportional to the total number of sRNA loci. Still, the authors need to explain why there are such differences in the power to detect sRNA loci.

2. Strand bias: I'm not sure what the authors report. First, Watson and Crick strands are defined as sense and antisense strands for protein coding genes - the terms are inappropriate as used here. What I would like to know is whether there a bias towards one strand for any given locus. To answer this, please report simply the average ratio of reads between the strand with the higher expression and the strand with the lower expression.

3. I had asked for annotation: TEs, mRNAs, non-coding RNA etc. This has not been done.

4. Finally, I would still very much prefer if the methylation analysis of H06 was done by bisulphate sequencing of PCR products. "We do not have bisulphite sequencing data at the moment as we plan for genome-wide methylation analysis." is not a logical argument. (Of course, if the authors would add genome-wide data, this would greatly elevate the level of the manuscript.)

This is an interesting manuscript, and I would like to see it published in EMBO J., but I believe the authors could be more diligent. This also pertains to typos, especially in SOM, and the layout of the figures (I encourage the use of consistent and appropriate font sizes throughout all figures).

2nd Revision - authors' response 09 November 2011

Comments by the Reviewer:

I still think that the miR395 part is distracting, because it is likely an siRNA-independent form of epistasis (as the authors now explicitly discuss). Moreover, the additional explanations regarding miR395 loci and isoforms were not particularly helpful. As is, it looks like this part of the manuscript has been retained to pad the story. I very strongly advise to remove this, and instead move some of the SOM figures into the main text.

*****We presented our rationale for looking at transgressive sRNA loci - it is based on the prediction that primary sRNAs would initiate secondary sRNA production. The miRNA395 part of the results is not consistent with that model but nevertheless it is transgressive and involves an output that is small RNA related. We do not agree that it is a distraction. We feel that by including the miRNA story we reinforce the idea in the manuscript that the transgressive sRNA story is more complicated than we had anticipated originally.**

I had several questions regarding the annotation and distribution of transgressive loci, and these have only been partially answered.

1. Genome-wide distribution: The number of sRNA loci per chromosome is very uneven (Fig. S5A), even though the chromosomes themselves vary less than twofold in length. The good news is that the number of transgressive loci is approximately proportional to the total number of sRNA loci. Still, the authors need to explain why there are such differences in the power to detect sRNA loci.

*****The answer is simply related to the incomplete genome sequence available to us and we have made this clear in the text.**

2. Strand bias: I'm not sure what the authors report. First, Watson and Crick strands are defined as sense and antisense strands for protein coding genes - the terms are inappropriate as used here. What I would like to know is whether there a bias towards one strand for any given locus. To answer this, please report simply the average ratio of reads between the strand with the higher expression and the strand with the lower expression.

***** We have redone the analysis and the results are presented in Supplementary FigureS7A. The text has been modified to take this point into account. There is slight increase in strand bias among transgressive loci when compared to all loci.**

3. I had asked for annotation: TEs, mRNAs, non-coding RNA etc. This has not been done.

*****Please note that we only have access to limited sequence annotations in the public domain. However we have done the best we can and we now provide a summary of features overlapping with transgressive loci (Supplementary Figure S6). The pattern is similar to that of total sRNA loci but there are too few transgressive loci involved for us to say whether this pattern is significantly different from that of the total sRNA loci.**

4. Finally, I would still very much prefer if the methylation analysis of H06 was done by bisulphate sequencing of PCR products. "We do not have bisulphite sequencing data at the moment as we plan for genome-wide methylation analysis." is not a logical argument. (Of course, if the authors would add genome-wide data, this would greatly elevate the level of the manuscript.)

*****We have now presented the bisulphite sequencing data for the H06 loci (Fig 4B and Supplementary Figure S9). Data supports and reinforces the McrBC methylation assay presented before.**

3rd Editorial Decision 14 November 2011

Thank you for submitting your revised manuscript, it has been seen by referee #1 who finds that you have satisfactorily addressed all the remaining concerns. I am happy to accept the manuscript for publication in The EMBO Journal where I believe it will make a great contribution. You will receive the official acceptance letter in the next day or so.

Yours sincerely,

Editor
The EMBO Journal

REFEREE COMMENTS

Referee #1

Kudos to the authors for heeding (most of this) reviewer's advise. Great paper!

