Author's response to reviews

Title: Comprehensive Expressional Analyses of Antisense Transcripts in Colon Cancer Tissues Using Artificial Antisense Probes

Authors:

Rintaro Saito (rsaito@sfc.keio.ac.jp)
Keisuke Kohno (c0730177@md.tsukuba.ac.jp)
Yuki Okada (okayan@gmail.com)
Yuko Osada (yuko.osada@gmail.com)
Koji Numata (numata@brc.riken.jp)
Chihiro Kohama (s0921075@u.tsukuba.ac.jp)
Kazufumi Watanabe (watanabe@hssnet.co.jp)
Hajime Nakaoka (nakaoka@cslab.co.jp)
Naoyuki Yamamoto (yamamoto@cslab.co.jp)
Akio Kanai (akio@sfc.keio.ac.jp)
Hiroshi Yasue (hyasue@affrc.go.jp)
Soichiro Murata (soichiro@md.tsukuba.ac.jp)
Kuniya Abe (abe@rtc.riken.jp)
Masaru Tomita (mt@sfc.keio.ac.jp)
Nobuhiro Ohkohchi (nokochi3@md.tsukuba.ac.jp)
Hidenori Kiyosawa (kiyosawa@rtc.riken.jp)

Version: 2 Date: 8 October 2010

Author's response to reviews: see over
Response to referees’ comments

1. Reviewer #1 (Version: 1 Date: 28 May 2010, Reviewer: Andreas Werner)

Comment 1-1: However, the line of argument of the manuscript is difficult to follow; the results should be better structured.

Response 1-1: The manuscript has been revised accordingly. We re-structured most of the manuscript, including the discussion, so that readers could better understand the results (Discussion).

Comment 1-2: I have two major concerns that question the main outcome of the study. The first one relates to the 10% cutoff for expression changes. I can see at least two experimental parameters that are likely to influence signal intensity by more than 10%. First, the tissue samples in the study are heterogeneous. Whereas the cancer tissue may be of clonal origin, the cells will differentiate quickly resulting in variations in gene expression. The control samples will certainly be polyclonal and probably contain different cell types (smooth muscle cells, interstitial cells, epithelial cells). This heterogeneity may contribute significantly to apparent changes in gene expression. We note that we are focusing on antisense transcripts which are expressed in colon cancer tissues regardless of their stages.

Response 1-2: We agree that variations in gene expression might be attributed to polyclonal control cells and differentiated cancer cells. We included a description in the manuscript (Page 19 lines 11-14) to address this point. We still suggest that genes in cancer tissues are up-regulated or down-regulated, despite this “blurring effect” caused by these random variations in gene expressions. In fact, the set of up-regulated and down-regulated sense genes in our results is consistent with those presented in a previous study (Page 18 lines 23-24, Page 19 line 1). Furthermore, in our procedure, we chose sense-AFAS pairs satisfying all of the following four conditions: (1) The expression level of the sense transcript is increased by >10% in cancer tissue. (2) The expression level of the AFAS transcript is increased by >10% in normal tissue. (3) The expression level of the sense transcript is 10% greater than that of the AFAS transcript in cancer tissue. (4) The expression level of the AFAS transcript is 10% greater than that of the sense transcript in normal tissue. Thus, our thresholds for the detection of expression changes between normal tissues and cancer tissues are rigorous. We added this
This is supported by the random shuffling test showing that the number of sense-AFAS pairs satisfying the above criteria (i.e., the observed number, 44) is statistically significant (i.e., the expected number $12.07 \pm 3.11$ SD).

We also changed the threshold of 10% to 20% and 30% (Page 15 lines 1-4). Our main result is the expression changes in antisense transcripts detected by the AFAS probes (described in the abstract). We have identified 172 antisense transcripts showing a change of expression of greater than 100% between normal tissues and cancer tissues.

**Comment 1-3:** Second, the authors will be aware of the fact that second strand synthesis generates (estimated) 15-20% misprimed cDNA. The resulting probes will detect the opposite strand of the original RNA and therefore skew the readout of the experiment. Taking these drawbacks into consideration, a 10% change in expression is hardly significant.

**Response 1-3:** For random priming, we conducted only first-strand cDNA synthesis, thus avoiding the misprimed cDNA synthesis inherent in second-strand cDNA synthesis. For oligo dT priming, we performed second-strand cDNA synthesis. However, we followed the standard protocol of Agilent, which is widely accepted by the researcher community. Furthermore, we took into account the mispriming associated with second-strand cDNA synthesis by changing the threshold from 10% to 20% and 30% (Page 15 lines 1-4).

**Comment 1-4:** The second concern relates to two first strand priming methods used in the study, oligo dT or random primers. Quite generally, polyadenylated transcripts are more stable and can be exported to the cytoplasm. The polyadenylated antisense transcripts usually derive from autosomes (i.e. they are under-represented on the X chromosome). Poly A- transcripts are short lived and nuclear. To a certain extent these NATs may relate to the transcriptional activity around promoter sites that has recently been demonstrated (by RNAseq). The characteristics of the two groups of antisense transcripts indicate that they may serve rather different biological roles. I assume that different biological stimuli will lead to expression changes of poly A+ or poly A- antisense RNA fractions, respectively. As a consequence, data derived from the two cDNA priming methods cannot be compared.
**Response 1-4**: It is true that the physiological functions of polyA+ transcripts and polyA- transcripts are usually different, and a comparison of their expression patterns, assuming that they function in the same cellular process, is meaningless. We changed the text to reflect the comments of the reviewer in the revised manuscript where expression patterns obtained by oligo-dT and random priming are directly compared (Page 15 lines 14-16). However, we still suggest that it is meaningful to compare expression patterns detected by oligo-dT and random priming without making such an assumption. Firstly, we found two types of transcripts, those that can be detected only by random priming, and those that can be detected by both priming methods. It is meaningful to discuss differences in characteristics between these two types of transcripts. Also, there are transcripts whose expression patterns detected by the two priming methods are similar and those whose patterns are different. The functions of transcripts that are specifically detected by random priming remain to be elucidated. According to our previous work (Kiyosawa et al. 2005) the detection of transcripts by random priming did not change greatly upon stimulation. Secondly, differences between polyA+ transcripts and polyA- transcripts are sometimes not clear. For example, PolyA- transcripts such as Xist transcripts are stable and expressed at a steady rate.

**Comment 1-5**: The different platforms should be better explained in the results section and details should be included in the methods section.

**Response 1-5**: The manuscript has been revised accordingly. We briefly described our platform at the beginning of the results section (Page 8 lines 14-15). Note that we used the same custom microarray platform throughout this work.

**Comment 1-6**: A table that summarizes the key results would help (for example)
platform/priming
sense x positive calls
antisense y positive calls

**Response 1-6**: The manuscript has been revised accordingly. We included a table to summarize our key results (Supplemental table 3).
Comment 1-7: The figure legends should give more details. For example, figures 1 and 3 contain a panel A and B which is not mentioned in the legend.

Response 1-7: The manuscript has been revised accordingly. We added descriptions of panels A and B in the legends of Figures 1 and 3 (Figure 1 legend, Figure 3 legend).

Comment 1-8: There are a few parts where the text could be clarified (Abstract/Background, last sentence: Abstract/Results middle sentence, for example). In addition, wordings like “exemplary example”, “may play an important role” “may be highly related to cancer” should be avoided. The latter two examples express pure speculations and the terms “important” and “highly” seem pretentious.

Response 1-8: The manuscript has been revised accordingly (Page 3 lines 5-8, lines 13-16, line 22 (Abstract), Page 10 line 6).

Comment 1-9: How is protein coding defined?

Response 1-9: We included a definition of protein coding in the revised manuscript (Page 8 line 24, Page 9 lines 1-3).

2. Reviewer #2 (Version: 1 Date: 3 June 2010, Reviewer: Claes Wahlestedt)

Comment 2-1: First and foremost, the concept of antisense RNA being non-polyadenylated is dominating the manuscript. This is, however, not a novel finding as the senior author and some other researchers have reported on this finding before. Thus, this finding has to be significantly downplayed in the Abstract, Introduction, Discussion, Figure legends and Methods section.

Response 2-1: The manuscript has been revised accordingly. In the revised manuscript, we emphasized that antisense transcripts without poly(A) tails have been described in previous studies (Page 3 lines 16-17 (Abstract), Page 17 lines 5-7).
Comment 2-2: Page 5: “we found 66.1% of 635 genes targeted...”. It is not clear whether this sentence refers to current findings presented in this present manuscript or if it is related to previously published work.

Response 2-2: In the revised manuscript we have clarified that the sentence in question relates to our previously published work (Page 6 lines 17-19).

Comment 2-3: On page 6, second paragraph, the authors bring up 68 sense-antisense transcripts. This sentence is not clear and it is not clear how the authors quantified expression patterns between cancerous and normal tissue.

Response 2-3: In the revised manuscript we changed the sentence in question to make the point clear (Page 7 lines 15-18). In addition, we briefly described how we quantified the expression patterns in the results section (Page 9 lines 12-20).

Comment 2-4: One of the fundamental alterations in cancerous cells is the process of de-differentiation. This process can generate alterations in sense or antisense RNA expression. Therefore, it is not clear which fractions of the observed alterations are causes or consequences of cancerous transformation. Please discuss more carefully.

Response 2-4: We agree with the comment and therefore incorporated the above mentioned fact in the revised manuscript (Page 19 lines 20-22).

Comment 2-5: In the first part of the Results section, page 7, the authors describe DNA chips, which include 2,358 probes that are only detecting protein coding genes as well as 250 probes that are paired between coding and non-coding. Are these 2,358 probes detecting coding-coding sense-antisense pairs? On the following page (Page 8) the authors have mentioned that most of sense and antisense genes that are showing altered expression were both protein coding. Is this conclusion simply a reflection of greater presence of probes that detect coding-coding pairs? Please discuss as thoroughly as possible.
Response 2-5: The 2358 probes included on the DNA chips are intended to detect coding-coding sense-antisense pairs. The statement “most of sense and antisense genes that are showing altered expression were both protein coding” is not important for this work and was removed from the manuscript (Page 10 line 3).

Comment 2-6: How is DNA contamination ruled out? In particular, a reader might wish to know if signals are stronger with random hexamers than oligoDT probes.

Response 2-6: RNA was extracted using the standard protocol (as described in “Preparation of tissue samples”). We conducted electrophoresis of samples and observed no visible bands corresponding to DNA. If samples were contaminated with small amounts of DNA, we suggest that it would have had very little effect on our results because there are only two copies of DNA per cell, which is much less than the amount of RNA.

Comment 2-7: The authors have started with designing probes for 501 cancer related genes, and then on page-14 they have concluded that the altered expression pattern is enriched with oncogenes; “there appear to be a relationship between expression pattern of sense-AFAS genes and Oncogenesis”. This seems to be a biased conclusion and needs further attention.

Response 2-7: We agree that this sentence gives a biased conclusion. In the revised manuscript, we changed this sentence so that it is clear that our results are related to the design of the AFAS probes rather than to the relationship between AFAS probe expression and oncogenesis (Page 16 lines 3-7).

Comment 2-8: Three AFAS for validation seem insufficient. Please expand or provide a better justification.

Response 2-8: RT-PCR for three AFAS probes is shown as an example of expression patterns and not for complete validation. Complete validation of the AFAS probes used in our present study was conducted in our previous work. We revised our manuscript to make this point clear (Page 18 lines 20-24, Page 19 line 1).
Comment 2-9: Page 15, second paragraph, is confusing: are those four probes designed for AFAS or for the sense transcript?

Response 2-9: The four probes in question are AFAS probes and were designed to detect antisense transcripts. In the revised manuscript, we have changed the relevant sentence to make this point clear (Page 17 lines 5-8).

Comment 2-10: The first paragraph of page 21 in method section is more of discussion and needs editing.

Response 2-10: The manuscript has been revised accordingly. In the revised manuscript, we included a discussion comparing our microarray data with that of previously published data (Page 18 lines 23-34, Page 19 line 1).

Comment 2-11: Page 22: It is not clear whether validation studies were performed on three commercial RNA samples or on the original RNA samples. If this is the result that is presented in figure 4, how can the authors show the error bars on that graph? Please reexamine.

Response 2-11: We performed RT-PCR on commercially available RNA samples (Page 24 lines 17-18). The error bars represent the standard deviation for the technical duplicates. In the revised manuscript, we included this information in the legend of Figure 4 (Figure 4 legend).

Comment 2-12: The figure legends are not complete and needs more attention.

Response 2-12: The manuscript has been revised accordingly. We addressed this issue by expanding and completing the legends of Figures 1, 3, and 4 (Figure 1 legend, Figure 3 legend, Figure 4 legend).
3. Reviewer #3 (Version: 1 Date: 7 June 2010, Reviewer: eric adriaenssens)

**Comment 3-1:** Authors exploit their previously established microarray platform and referred often to their previous works (ref 7 and 10) especially for control experiments (WB, RT-PCR to validate data of microarray). Some of these controls can be included because this report appears as "more of the same".

**Response 3-1:** We are willing to use previously published figures in this manuscript, but this is an editorial issue. We will receive an opinion on this issue from the editorial office.

**Comment 3-2:** For fig 1, authors identify genes in which expression of sense-antisense transcripts changes by 10% (between normal and cancer) for at least 3 among 6 patients. This threshold seems weak. The variations of expression of numerous genes between 2 cells or 2 patients are often superior in 10%. In this line, what is the physiological relevance of this report?

**Response 3-2:** We addressed this issue in the revised manuscript by changing the threshold from 10% to 20% and 30% (Page 15 lines 1-4). We also note the following points. Firstly, in our procedure, we chose sense-AFAS pairs satisfying all of the following four conditions: (1) The expression level of the sense transcript is increased by >10% in cancer tissue. (2) The expression level of the AFAS transcript is increased by >10% in normal tissue. (3) The expression level of the sense transcript is 10% greater than that of the AFAS transcript in cancer tissue. (4) The expression level of the AFAS transcript is 10% greater than that of the sense transcript in normal tissue. Thus, our thresholds for the detection of expression changes between normal tissues and cancer tissues are rigorous. This is supported by the random shuffling test showing that the number of sense-AFAS pairs satisfying the above criteria (i.e. the observed number, 44) is statistically significant (i.e., the expected number 12.07 ± 3.11 SD). We added this description (Page 14 lines 21-24). Our main result is the expression changes of antisense transcripts detected by the AFAS probes (described in the abstract). We have identified 172 antisense transcripts showing a change in expression of more than 100% between normal tissues and cancer tissues.
Comment 3-3: As indicated, most of the sense and antisense genes showing altered expression balances in cancer tissues were both protein-coding genes. So, a simple deregulation of gene expression (frequent in cancer) can cause this result. It is more interesting to examine the variations of expression of non-coding RNA leading perturbation of protein-coding gene expression.

Response 3-3: In our analysis, we found that of the 68 cDNA-based sense-antisense transcripts in which the expression balance was altered in colon cancer tissues the, only six protein-coding – non-coding pairs showed remarkable perturbations. For this reason, we designed artificial probes for the antisense strand of cancer-related genes to detect such non-coding RNAs.

Comment 3-4: Authors compare 6 tumor samples with adjacent normal tissues. This number is very too small to give sound conclusions. Tissues adjacent to tumor can be quite different from normal ones (inflammation, change in cell types). By addition, cancer samples are often exclusively composed of epithelial cells, but normal tissue is composed of different types of cells (epithelial, fibroblast and others). Authors must evaluate the composition of their biopsies. In addition, no clinical data (stage, grade, ...), or molecular features (DCC, ret, APC genes loss, aneuploidy, ...) are presented but commonly searched and available in anatomopathological labs.

Response 3-4: The normal tissues we extracted were ~10 cm away from the cancer tissues. The normal tissue we extracted was mucosa, which is mostly composed of epithelial cells. Pathological examinations demonstrated that all tumors included in the study were adenocarcinomas. We added this information to the revised manuscript (Page 22 lines 14-18). We also included the clinical data of the patients (Supplemental Table 6). It is possible that some antisense transcripts are specifically expressed at a single stage in the cancer process. However, it is difficult to extract such stage-specific antisense transcripts as the number of patients was too small. Therefore, we identified antisense transcripts that were expressed in colon cancer tissues, regardless of the stage in which they were expressed, and compared their expression levels in colon cancer tissues with those in normal tissues.

Comment 3-5: As indicated by the authors, antisense transcripts may be
“transcriptional noises”. And this feature can be increased in cancer cells in which transcriptional control are often loss. The antisense detected in this report could have biological functions, probably by acting on sense mRNA and protein realized. But, nothing in this report allows a characterization of this type of antisense even if this point is well discussed. The fact that antisenses are often non polyA and that their border are not defined are in favour of transcriptional noises.

Response 3·5: We agree that many of the antisense transcripts detected by our AFAS probes might be the result of transcriptional noise and this possibility has been discussed in the revised manuscript (Page 19 lines 11-14). We also pointed out in the abstract that these transcripts are potential targets for further experimental validation. We still suggest that the identified transcripts include several functional antisense RNAs, some of which might be associated with transcription. For example, our results showed that the positions of antisense RNAs we detected are biased toward the 5’ end of sense genes. This finding is consistent with previously published results (He Y et al. Science 2008; Taft RJ et al. Nat Genet. 2009). We included this information in the revised manuscript (Page 19 lines 14-17).

We suggest that a discussion of the regulatory roles of antisense transcripts is important because some antisense transcripts appear to be specific to cancer tissues or anti-correlated with the expression pattern of the corresponding sense transcript.

Comment 3·6: Several data shown possible antisense artefacts in transcriptome microarray experiments due to non desired priming in reverse transcription (Perrochi et al., 2007). Authors have been examined this possibility?

Response 3·6: The paper by Perocchi F et al. Nucleic Acids Res. 35(19):e128, 2007 describes artifacts that might arise during target preparation by reverse transcription, confounding the interpretation of array data when both genomic strands are interrogated. This paper warns that reverse transcription is likely to generate spurious second-strand cDNAs whose sequence pattern will be identical to that of expected antisense transcripts.

In our procedure, for oligo-dT priming, we first make constructs consisting of the T7 promoter + dT, which targets the poly-A tail of transcripts, and then we transcribe the constructs so that cRNAs are synthesized. We label these cRNAs and do not use the products of reverse transcription. For this reason, we believe that for
oligo-dT priming the concerns raised by the reviewer will not eventuate. However, the problem pointed by the reviewer could be a potential problem for our random priming approach since we labeled cDNA generated by reverse transcription. If the synthesis of such artifacts was frequent, we would observe a positive correlation between expression levels of sense and antisense transcripts. However, we observed no such correlation in random priming. Therefore, we suggest that the level of artifacts arising from random priming is minimal. We observed many cases that cannot be attributed to artifacts where the expression level of the antisense transcripts is much greater than that of the corresponding sense transcripts (Figure 5). These findings cannot be explained by the undesired synthesis of second-strand cDNA. We included this information in the revised manuscript (Page 19 lines 2-8).

**Comment 3-7:** Legend and presentation of the figures could be greatly ameliorated. By example, fig 1 described sense-antisense transcripts showing altered expression balances but what is a) and b)?? it is not indicated in the legend (idem for fig 3). It is impossible to compare the fig 1 with table 1. There is 68 sense-antisense transcripts with altered expression in table 1, but less in fig 1, Why?

**Response 3-7:** According to the reviewer’s suggestions, we have revised the figure legends of Figures 1, 3, and 4, so that readers can better understand the figures (Figure 1 legend, Figure 3 legend, Figure 4 legend). We note that Figure 1 and Table 1 cannot be compared because Figure 1 refers to sense-antisense transcripts whereas Table 1 refers to sense-AFAS pairs. The number of pairs shown in Figure 1 is less than 68 because Figure 1 only refers to expression data obtained by the random priming approach. In the revised manuscript, we added this information to the legend of Figure 1 (Figure 1 legend).