PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Pair-wise comparison analysis of differential expression of mRNAs in early and advanced stage primary colorectal adenocarcinomas
AUTHORS	Lau, Tze Pheng; Roslani, April Camilla; Lian, Lay Hoong; Chai, Hwa Chia; Lee, Ping Chin; Hilmi, Ida; Goh, Khean Le; Chua, Kek Heng

VERSION 1 - REVIEW

REVIEWER	Prof RAVINDRAN ANKATHIL Human Genome Center, Universiti Sains Malaysia , Malaysia
REVIEW RETURNED	01-Apr-2014

GENERAL COMMENTS	There are few grammatical and syntax errors which needs to be edited. Few clarifications to be sought from the authors. 1.Did all the patients studied, belong to same ethnic group or different ethnic groups in Malaysian Population (Malays, Chinese, Indians) ? please clarify If belonged to different ethnic groups, don't you think it would be worthwhile to analyze separately based on ethnicity of the patients also , and see the difference, if any? 2. Based on the distinctive expression signatures of genes, do you have any postulations to make on the pathways which might be involved ? I think your clarifications or added contributions on the above aspects will give more strength to your data.

REVIEWER	Dr Adrian C. Bateman University Hospital Southampton NHS Foundation Trust, UK
REVIEW RETURNED	11-May-2014

GENERAL COMMENTS	This script describes a study of 27 cases of colorectal cancer (CRC), in which the mRNA expression profile of several genes was compared between tumour and normal tissue – with the cases being divided into ‘early’ and ‘advanced’ cancers. The authors found that some genes were over-expressed in early and advanced CRC, while one (ARPC2) appeared to be underexpressed in early CRC and another (C6orf173) was overexpressed in advanced CRC, compared to normal colonic tissue. This is potentially interesting but I have the following more specific comments:

	[1] The authors raise several issues with current histopathology-based staging methods and they state that 'current clinicohistopathological parameters are inadequate for accurate individual prognostic prediction'. While there are difficulties with the application of certain aspects of pTNM staging e.g. interpretation of the criteria for assessing subserosal tumour deposits, TNM staging is widely used during multidisciplinary discussions regarding patients with CRC and this system is able to stratify patients for prognostication and treatment decision-making. Therefore, while gene expression studies may in the future help to refine histopathological assessment, I do not accept that the latter can be regarded as inadequate. Furthermore, the information derived from histopathological assessment is likely to be very cost-effective compared to molecular analyses. [2] The authors use a system (i.e. clinicohistopathological staging) that they claim to be inadequate for clinical management purposes, to divide up the cases into groups for molecular analysis – which seems at odds with their earlier statement regarding the utility of clinical/pathological staging. [3] The numbers of cases are small, especially when they have been divided into early and advanced tumours, and the cases within each of these groups are heterogeneous in terms of pTNM stage. [4] Were the tumour samples taken for analysis from the primary tumours in all cases, or from nodal or distant metastases? [5] This is a recent patient series, with no inclusion of clinical follow-up and therefore the study cannot determine whether the findings represent epiphenomena or independent prognostic factors. [6] The method for recording and therefore comparing the mRNA expression levels would benefit from clarification as potential readers may well not be experts in the field of mRNA expression studies. [7] Have the data been tested to ensure that they are normally distributed, as a parametric test has been use during the analysis? [8] Have the p values obtained been corrected for the number of comparisons made – to reduce the chance of finding a significant result purely due to the multiple comparisons undertaken? [9] Could it be that the difference in mRNA expression profiles between early and advanced CRC may purely reflect the age of the tumours rather than differences in stage? In summary, I feel that this script presents interesting data that merit publication, but that the nature of these data is very preliminary and that they should not be overinterpreted. For example, these data would be best used for 'hypothesis generation' purposes (but they could still be published for this purpose).

VERSION 1 – AUTHOR RESPONSE

Reviewer's Comments (Prof. Ravindran Ankathil)
1. Did all the patients studied, belong to same ethnic group or different ethnic groups in Malaysian population (Malays, Chinese, Indians)? Please clarify.
Answer: The patients' group comprised of the three main ethnic groups in the Malaysian population (i.e., Chinese, Malays and Indians). The number of patients recruited was in concordance with the ethnic-specific incidence rates reported in Malaysia, where the Chinese scored the highest CRC incidence compared to the Malays and Indians.

If belonged to different ethnic groups, don't you think it would be worthwhile to analyze separately based on ethnicity of the patients also, and see the difference, if any?
Answer: Yes, we do agree that it would be of significant value if we could analyse the gene expression patterns of colorectal tumours specific to each ethnic group. Unfortunately, our small sample size limits the merit of patients' stratification based on ethnicity. Thus, it would be our next immediate step to replicate our findings in another larger sample cohort, which will enable us to analyse based on ethnicity with adequate statistical power.

2. Based on the distinctive expression signatures of genes, do you have any postulations to make on the pathways which might be involved?
Answer: It was postulated that the actin cytoskeleton might be involved in determining the dysplastic cell morphology during the early development of CRC. On the other hand, the abnormality in the assembly of functional kinetochore was postulated to be critical in determining the aneuploidic characteristic of advanced stage colorectal tumours.

Reviewer's Comments (Dr. Adrian C. Bateman)
1. The authors raise several issues with current histopathology-based staging methods and they state that 'current clinicohistopathological parameters are inadequate for accurate individual prognostic prediction'. While there are difficulties with the application of certain aspects of pTNM staging e.g. interpretation of the criteria for assessing subserosal tumour deposits, TNM staging is widely used during multidisciplinary discussions regarding patients with CRC and this system is able to stratify patients for prognostication and treatment decision-making. Therefore, while gene expression studies may in the future help to refine histopathological assessment, I do not accept that the latter can be regarded as inadequate. Furthermore, the information derived from histopathological assessment is likely to be very cost-effective compared to molecular analyses.
Answer: Despite it is widely used in current patient prognostication and treatment decision-making, the application of the latest TNM staging system is challenged with its complexity and inter-observer subjectivity in criteria assessment. Moreover, the TNM staging criteria that based on the anatomical and histological features of primary tumours are still insufficient to address the great tumour heterogeneity of CRC. Thus, there is a need to integrate molecular markers to complement current TNM staging system in order to improve the patients' management in future. However, the author does not imply that the TNM staging system should be aborted and completely replaced with molecular analyses.
*The sentences were re-phrased at lines 26 and 27 (page 4), as well as lines 1 and 2 (page 5) under the section "Introduction" (red fonts).
*The complexity of the latest TNM staging system and challenging issues associated with its application were reviewed in Puppa et al., 2010 (Arch Pathol Lab Med, Vol. 134, pp.837-852), Doyle and Bateman, 2012 (J Clin Pathol, Vol.65, pp.372-374).

2. The authors use a system (i.e. clinicohistopathological staging) that they claim to be inadequate for clinical management purposes, to divide up the cases into groups for molecular analysis – which

seems at odds with their earlier statement regarding the utility of clinical / pathological staging.
Answer: It was not the author's initial intention to abort the existing TNM staging method and replace it with molecular analyses. The author's main objective is to characterise the mRNA expression patterns specific to early- and advanced stage CRC, which might be exploited for future prognostication and treatment response prediction, complementary to the histological features.

3. The numbers of cases are small, especially when they have been divided into early and advanced tumours, and the cases within each of these groups are heterogeneous in terms of pTNM stage.
Answer: The findings of this study were considered preliminary owing to the small sample size. Our next immediate step is to investigate the expression of all the 16 differentially expressed genes in different pTNM stage group within a larger sample cohort.

4. Were the tumour samples taken for analysis from the primary tumours in all cases, or from nodal or distant metastases?
Answer: Yes, all tumour samples were excised from the primary colorectal tumours in all cases.

5. This is a recent patient series, with no inclusion of clinical follow-up and therefore the study cannot determine whether the findings represent epiphenomena or independent prognostic factors.
Answer: Since these findings on different mRNA expression profiles between the early- and advanced stage primary colorectal tumours were considered preliminary, we cannot over-interpret that these molecular markers can be targeted as prognostic factors specific to our Malaysian CRC patients. As mentioned in the last paragraph of this manuscript, continuous efforts in validating the differential or predictive ability of these gene expression patterns on CRC staging or prognostication, in conjunction with other histopathological and imaging parameters, are needed.

6. The method for recording and therefore comparing the mRNA expression levels would benefit from clarification as potential readers may well not be experts in the field of mRNA expression studies.
Answer: The means of comparing the mRNA expression levels were further elaborated in pg.10 & 14, as highlighted in yellow.

7. Have the data been tested to ensure that they are normally distributed, as a parametric test has been used during the analysis?
Answer: Yes, the distribution of the ΔCT values obtained for each DEGs within each CRC and control group were tested for their normality via the Shapiro-Wilk test. For the early stage CRC group, the data set for ARPC2, RPL35, RPS23 and TIMP1 were in normal distribution for both CRC and control groups, except in the CRC group of the RPL35. As for the advanced CRC group, the data set for C6orf173, RPL35, RPS23 and TIMP1 were in normal distribution for both CRC and control groups, except in the control group of the TIMP1.

8. Have the p values obtained been corrected for the number of comparisons made – to reduce the chance of finding a significant result purely due to the multiple comparisons undertaken?
Answer: We think that the multiple comparisons problem does not exist in current study as only one attribute was compared for each DEGs, i.e. the ΔCT values between the CRC and control group. The p values were obtained to assess the statistical significance of the differences in ΔCT values between two groups. Therefore, no correction was made on the p values obtained, but the median absolute deviation was calculated to determine the within-group correlation of these ΔCT values, and the group outliers were excluded from the statistical analysis.

9. Could it be that the difference in mRNA expression profiles between early and advanced CRC may purely reflect the age of the tumours rather than differences in stage?
Answer: The early CRC group was comprised of patients aged between 40 to 85 years, while the

CRC patients grouped under the advanced CRC were of 52 to 86 years old. Since patients recruited for both early- and advanced stage CRC group were of comparable age group, it is unlikely that our reported differences in mRNA expression profiles were merely reflective of the age of tumours rather than differences in cancer stage. Furthermore, all recruited patients were newly-diagnosed with sporadic CRC and the cancer stage at the time of diagnosis was used to divide them into either the early- (Stages I & II) or advanced (Stages III & IV) CRC group. The comparison on mRNA expression levels was then made on these primary tumours and their pair-wise normal colonic mucosa within each CRC group. Therefore, in our opinion, the observed differences in mRNA expression profiles between the early- and advanced stage CRC were due to the differences in cancer stage rather than the age of tumours.