Reviewer's report

Reviewer number: 1

Reviewed by 02451547

Comments To Authors

The authors mainly focus on verifying gene expression profiles for colorectal cancer using 12 internet public microarray data sets. The results suggest a novel gene expression profile was associated with CRC and can potentially be applied to blood-based detection assays. However, there are some concerns raised from this paper.

1. Table 1 should be one to list the characteristics of the included papers.

- We thank for the Reviewer pointed out an important issue. We have added and revised the texts in the manuscript as well as the following,

Added on table 1 and table 2.

…. Furthermore, as table 1, we validated the 17 CRC-associated genes from Quyun, et al., 2010 (Model 1: 5 genes), Marshall et. al. (Model 2: 7 genes) and Han et. al. (Model 3: 5 genes) ….

2. Why do the authors select the CRC-associated genes from 3 studies (Ref. 7-9) to verify external validation rather than from other studies.

- We thank for the Reviewer's remind. We have provided the details in the manuscript as well as the following,

Many studies develop an accurate, reliable and less invasive test for detecting CRC. Han et al and Marshall et al with similar screening approaches used different gene sets to detect CRC. The two gene sets were obtained by direct selection from differentially expressed genes in peripheral blood samples using microarray techniques followed by real-time PCR. The biomarkers they selected may more or less reflect the static and dynamic changes of the immune system in response to cancer. In our study, genes clinically confirmed to be cancer-associated in tumor tissues were chosen for selection and validation in peripheral blood samples. However these two studies even with similar

approaches and some overlapped samples, but reported respective profiles cover no genes in common with the profile of 5 genes from Qunyun et al. The absence of concordant genes also exists in Xu et al(Xu, et al., 2013) that could be related to differences in studying samples and genes coming out from the upstream or downstream of on- and anti-cogenasis pathways, supposedly they all perform perfect gene quantification and statistical analysis to develop particular CRC gene expression profiles. The present study intends to stand on and demonstrates to rapidly converge and verify these promising biomarkers using pooling external validation and public microarray GSE datasets in GEO of NCBI before the further practical uses and clinical implementation. In general, this is an alternative effort to establish a standard testing procedure and to confirm the profile performance.

3. Although this study included 12 public pooling microarray datasets, the sample size is relatively small, only including 519 cases of adenocarcinoma and 88 controls of normal mucosa. Do the authors think whether the statistical power is enough.

- We thank for the Reviewer's remind. We have provided the details in the manuscript as well as the following,

In general, this is an alternative effort to establish a standard testing procedure and to confirm the profile performance.

4. Have the authors considered to collect new sample by your lab to verify external validation of the selected gene.

- We thank for the Reviewer's remind. We would collect new samples via lab to verify external validation of the selected genes in the future work and that is why we do this present study. Furthermore, we have provided the details in the manuscript as well as the following,

In general, this is an alternative effort to establish a standard testing procedure and to confirm the profile performance. According to the results of this present study, the selected genes can be verified via lab

for collecting new samples in the future work.

Reviewed by 00555672

Manuscript Number 4538

Manuscript Title To verify gene expression profiles for colorectal cancer using 12 internet public microarray data sets-

Review Time 2013-09-09 20:39

Comments To
Authors

The manuscript by Huang and colleagues is a dataset analysis of three reported signature of colorectal cancer in a pooled dataset comprising 12 microarray studies. The study identifies a common signature which should be related to colorectal cancer. The premise of the study makes no sense. Authors want to develop a signature to detect CRC cells in peripheral blood, but they use microarray studies obtained from normal and cancer colonic tissue. In peripheral blood it would be expected a different background of gene expression ascribed to blood components which makes the study not useful for the purpose. Last and not least without qPCR validation it is not given that the differences noticed at the probe level can be reproduced with qPCR. Indeed, a microarray will never be used as a tool to detect CRC cells in blood. So definitely a qPCR validation step would be required.

- We thank for the Reviewer's remind. This present study is an alternative effort to establish a standard testing procedure and to confirm the profile performance. We would collect new samples via lab to verify external validation of the selected genes in the future work and that is why we do this present study. Furthermore, we have provided the details in the manuscript as well as the following,

Many studies(Rosenthal, 2006; Simi, et al., 2008; Wu, Chung, Chang, & Wang, 2013; Xu, et al., 2013; Yip, et al., 2010) develop an accurate, reliable and less invasive test for detecting CRC using tissues or blood samples from microarray to qPCR validation. In general, this present study is an alternative effort to establish a standard testing procedure and to confirm the profile performance. Genes clinically confirmed to

be cancer-associated in tumor tissues are chosen for selection and validation in peripheral blood samples. According to the results of this present study, the selected genes can be verified via lab for collecting new samples in the future work.

Han et al and Marshall et al with similar screening approaches used different gene sets to detect CRC(Han, et al., 2008; Marshall, et al., 2010). The two gene sets were obtained by direct selection from differentially expressed genes in peripheral blood samples using microarray techniques followed by real-time PCR. The biomarkers they selected may more or less reflect the static and dynamic changes of the immune system in response to cancer. However these two studies even with similar approaches and some overlapped samples, but reported respective profiles cover no genes in common with the profile of 5 genes from Qunyun et al. The absence of concordant genes also exists in Xu et al(Xu, et al., 2013) that could be related to differences in studying samples and genes coming out from the upstream or downstream of on- and anti-cogenasis pathways, supposedly they all perform perfect gene quantification and statistical analysis to develop particular CRC gene expression profiles. The present study intends to stand on and demonstrates to rapidly converge and verify these promising biomarkers using pooling external validation and public microarray GSE datasets in GEO of NCBI before the further practical uses and clinical implementation.

Han, M., Liew, C. T., Zhang, H. W., Chao, S., Zheng, R., Yip, K. T., et al. (2008). Novel blood-based, five-gene biomarker set for the detection of colorectal cancer. *Clin Cancer Res, 14*(2), 455-460.

Marshall, K. W., Mohr, S., Khettabi, F. E., Nossova, N., Chao, S., Bao, W., et al. (2010). A blood-based biomarker panel for stratifying current risk for colorectal cancer. *Int J Cancer, 126*(5), 1177-1186.

Rosenthal, L. S. (2006). Getting involved in gastroenterology beyond the borders of your fellowship program. *Gastrointest Endosc, 64*(4), 602-603.

Simi, L., Pratesi, N., Vignoli, M., Sestini, R., Cianchi, F., Valanzano, R., et al. (2008). High-resolution melting analysis for rapid detection of KRAS, BRAF, and PIK3CA gene mutations in colorectal cancer. *Am J Clin Pathol, 130*(2), 247-253.

Wu, C.-H., Chung, F.-Y., Chang, J.-Y., & Wang, J.-Y. (2013). Rapid detection of gene expression by a colorectal cancer Enzymatic Gene Chip Detection Kit. *Biomarkers and Genomic Medicine, 5*(3), 87-91.

Xu, Y., Xu, Q., Yang, L., Ye, X., Liu, F., Wu, F., et al. (2013). Identification and validation of a blood-based 18-gene expression signature in colorectal cancer. *Clin Cancer Res, 19*(11), 3039-3049.

Yip, K. T., Das, P. K., Suria, D., Lim, C. R., Ng, G. H., & Liew, C. C. (2010). A case-controlled validation study of a blood-based seven-gene biomarker panel for colorectal cancer in Malaysia. *J Exp Clin Cancer Res, 29*, 128.