Responses to reviewers' comments

Reviewer No.2546300

Comment: The manuscript is well written and easy to follow.

Response: Thank you for your review and the positive comments about the manuscript.

Comment: But the aim of the study is not clearly indicated in the title and also in the manuscript. It is not known exactly for what purpose the cell lines were used for NGS.

Response: Thank you very much for pointing this out. The Title, Abstract, Core tip, and Introduction sections were rewritten in the revised version of the paper to clearly indicate the aim of the study. The title was changed to "Next-generation sequencing traces human iPSC lines clonally generated from heterogeneous cancer tissue".

Comment: If the aim of authors was to characterize the obtained cell lines, then the authors had to first do karyotype analysis to establish a karyogram of the cell lines, then doing molecular analysis.

Response: Thank you for your comments. Karyotyping of the iPSC lines was not performed in this study because the karyotype of each single starting cell could not be technically analyzed and then compared to those of other iPSC lines. In addition, karyotyping is no longer performed as a routine analysis, as several iPSC lines have shown no karyotype abnormalities in a similar culture experiment. Therefore, the current study focused on the next-generation sequencing of iPSC lines. The heteroallelic depth (Table 7) of the sequencing was expected to reflect the genotype of each single starting cell present within the heterogeneous cancer tissue.

Comment: If they were about to find new or relevant mutations to colon cancer, it is nonsense because they had to firstly prove that the cell lines are in fact cancer cell lines. Other important concern is even if the cell lines are established and proved, functional analysis should be done to prove the involvement of mutations in pathogenesis of cancer. NGS might reveal many mutations which have nothing to do with the disease.

Response: Thank you for your insightful comments and suggestions. Although

I am interested in new or relevant mutations in colon cancer and the involvement of such mutations in the etiology of cancer, this was not the focus of the current study. To my knowledge, there are currently no data on human iPSC lines clonally generated from heterogeneous primary cells of cancer tissues and analyzed by next-generation sequencing. Heterogeneous cancer tissues constitute not only pre-cancer (stem) cells and pre-metastatic cancer cells but also stromal cells (such as mesenchymal stem cells, cancer-associated fibroblasts and tumor endothelial cells) and immune cells (such as tumor-associated macrophages, dendritic cells and tumor-infiltrating T cells). The genotypes of mature cancer cells would be expected to be identical to those of cancer tissues; however, the genotypes of the other cells might be different from those of the cancer tissues. Next-generation sequencing and iPSC technology were used to resolve genotype variations among single cells present within heterogeneous cancer tissues. The genomic DNA of ten iPSC lines, which were clonally generated from human colon cancer tissue, was analyzed and compared with the genotypes of the starting cancer tissues and the matched adjacent non-cancerous tissues.

Reviewer No. 02446041

Comment: The author produced colon cancer tissue-derived iPSCs by retroviral gene transfer (OCT3/4, SOX2, and KLF4) and characterized 10 iPSCs lines with NGS. It's of great interest to compare both at the genome level.

Response: Thank you very much for your review and the positive comments on the paper.

Comment: Page 5: "Primary cells from cancer tissue were cultured for one day at approximately 5%–10% confluency and then incubated with the pantropic retrovirus vector solution (OCT3/4, KLF4, and SOX2) at 37 °C for further one day." What does it mean to have "Primary cells?" What's the nature of such Primary cells? Given the heterogeneity of tumor, Primary cells might be normal stroma cells, not cancer cells.

Response: Thank you very much for the critical comments. "Primary cells" was rewritten to "Heterogeneous primary cells" in the Material and Methods section. It is generally conceivable that primary solid tumors constitute not only (pre-) cancer (stem) cells and pre-metastatic cancer cells but also stromal cells (such as mesenchymal stem cells, cancer-associated fibroblasts and tumor endothelial cells) and immune cells (such as tumor-associated macrophages, dendritic cells and tumor-infiltrating T cells). The genotypes of mature cancer

cells would be identical to those of cancer tissues; however, the genotypes of other cells might differ from those of cancer tissues.

Comment: "Thirteen non-synonymous SNVs were compared among the genotypes of the ten iPSC lines and both the tissues. A missense mutation in *EIF2AK2, TTN, ULK4, MCC, FLT4, STK19, STK31, TRRAP, WNK1, PLK1* or *PIK3R5* of the respective iPSC lines was not identical to the genotypes of both the tissues. However, the genomes of all iPSC lines did not have mutations in *ERBB2*and *MKNK2* that were the genotypes of the cancer tissues, suggesting the preference of their starting cells without such mutations." These are not written with Standard English. Please re-write. E.g., what's "both tissues"?
Response: Thank you very much for pointing this out. "Both tissues" was rewritten to "the cancerous or non-cancerous tissues". In addition, "*MCC*" was corrected to "*TSSK1B*". The Results section of the Abstract section was rewritten accordingly.

Comment: How could you "suggesting the preference of their starting cells without such mutations?" "It was expected that the genomes of the non-cancer tissues had genotypes identical to that of their germline whereas the genomes of the cancer tissues had somatic mutation. In detail, as the cancer tissues would consist of heterogeneous cell populations, their genomes must be heterogeneous genotypes. Meanwhile, it was also expected that the genome sequences of the respective iPSC lines play a role as the tracer of their starting single cells from the cancer tissues. It was likely that the genomic mutations of each iPSC line originated from those of each single cell present within the cancer tissues."
Response: Thank you for pointing this out. These points have been addressed in the revised Discussion section.

Comment: "Ten iPSC lines were clonally generated from primary cells of cancer tissues, and were subjected to next-generation sequencing. Genome mutations of iPSCs were different from genotypes of cancer tissues." Why is that? What's the driver mutation in iPSCs? Many sections are not written with Standard English, such as "In contrast, genomes of all iPSC lines did not have some mutations that were genotypes of cancer tissues," (core tip).
Response: Thank you for pointing this out. In this study, it remains unknown

whether non-synonymous SNVs of the iPSC lines were driver mutations or not. Further studies will be performed to investigate the involvement of these SNVs in the etiology of cancer. The "Core tip" section was rewritten to be more easily understood.

Comment: All the figure legends should be written in a concise manner.
Response: Thank you for this suggestion. All figure legends were rewritten in a concise manner.