Peer Review File

Article information: http://dx.doi.org/10.21037/tlcr-21-48

Reviewer Comments

Krause et al examined 5 cases of lung squamous cell carcinoma, using enrichment via nuclei-flow sorting and whole exome sequencing to interrogate the genetic features of the primary tumors and corresponding metastases. I have the following comments:

Comment 1: In the discussion section, the authors stated that "this study is the first to investigate spatial and temporal genomic heterogeneity of metastatic LUSC over the disease course". It is unclear to me if that statement is true, given that the authors seemed to leave out other relevant related studies, for example this study (PMID: 30348992) that used whole-genome sequencing in a series of lung carcinomas (including 2 lung squamous cell carcinomas) similarly to determine relationships between primaries and metastases. Another study (PMID: 32676329) also examined one lung squamous cell carcinoma with similar approach. A better literature search and more balanced discussion are needed.

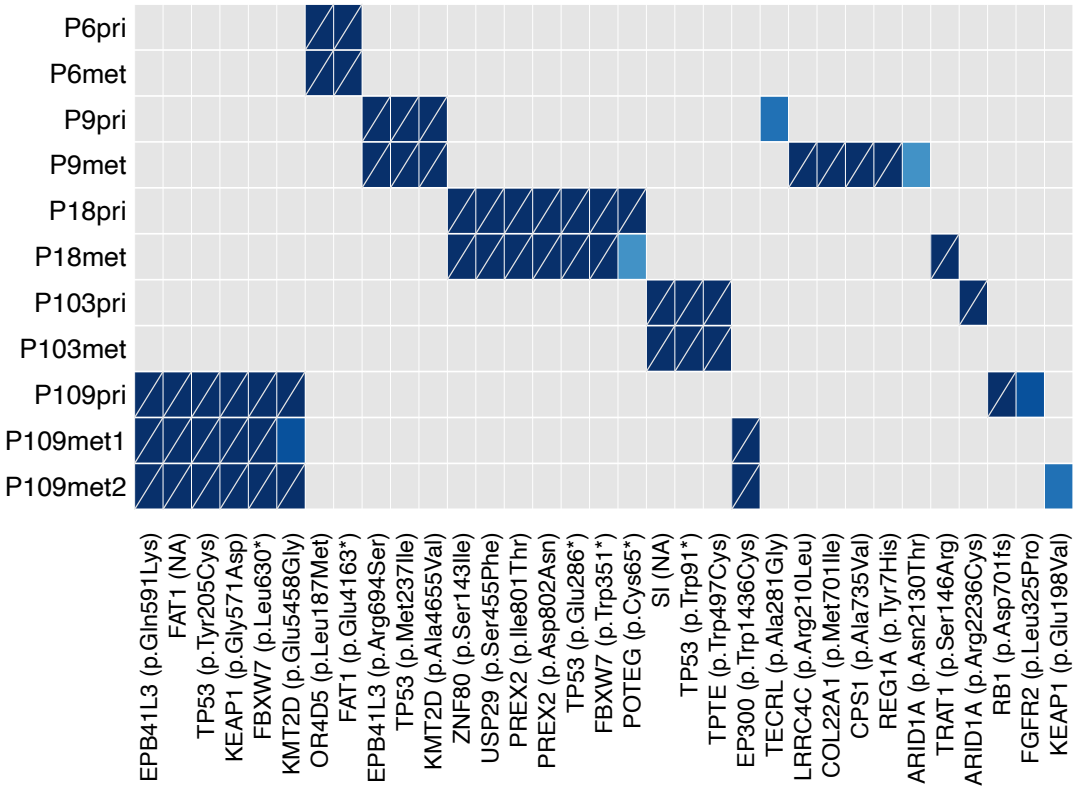
Reply: We thank the reviewer for drawing attention to the recently published papers from Peng et al. and Leong et al. which were not aware of at the time of submission. We changed the discussion accordingly.

Changes in the text:
(Please, see Page 15, lines 8 -13; Page 25, lines 4 - 9)

~~To the best of our knowledge, this study is the first to investigate spatial and temporal genomic heterogeneity of metastatic LUSC over the disease course~~.
Two other studies have recently been published on spatial and temporal heterogeneity of metastatic LUSC from a small number of patients (39,40). Leong et al. found that the interactions of the DNA repair genes with the tumor microenvironment may play a role in the acquisition of somatic mutations following the metastasis during lung cancer progression (39). Another recent study on metastatic LUSC, demonstrated that large structural variants may play a crucial role in intratumoral genetic heterogeneity (40).

Comment 2: Figure 1B primarily focuses on the shared mutations. Are there unique and pathogenic mutations not listed? Why did the authors use the dataset from Bailey et al to restrict the gene list displayed?
Reply: We thank the reviewer for this question. We selected the genes from Bailey et al., because we wanted to demonstrate that the most important oncogenic driver genes are clonally related in primary LUSCs and matched metastases. As suggested, we expanded the figure by also including other genes significantly mutated in TCGA dataset based on 511 LUSC samples.

New figure:



Changes in the text:

(Please, see Page 8 lines 2- 5; Page 33, lines 6 – 7)

(**b**) Heatmap depicts a comparison between the non-synonymous genes in the dataset of Bailey et al. and the significant mutated genes in TCGA dataset (MutSig Q-value < 0.05 based on 511 LUSC samples) that represent the most significantly mutated genes in LUSC compared to the primary tumor and the matched metastases of the presented patients. Heatmap illustrates the cancer cell fraction of selected mutations. Clonal mutations are illustrated with a diagonal line.

Comment 3: In Figure 2, for P109, since the Met2 is after treatment, this should ideally be annotated on the figure to alert the readers. Furthermore, for any other samples that are post-treatment, they can be similarly noted here as well so it is easier for readers to follow.

Reply: We agree with the reviewer and we annotated this information in Figure2 including an asterisk and changing the text of Figure 2. P109met2 was the unique sample for which we performed the sequencing after the treatment.

Changes in the text:
(Please, see Page 33, lines 17 - 18)

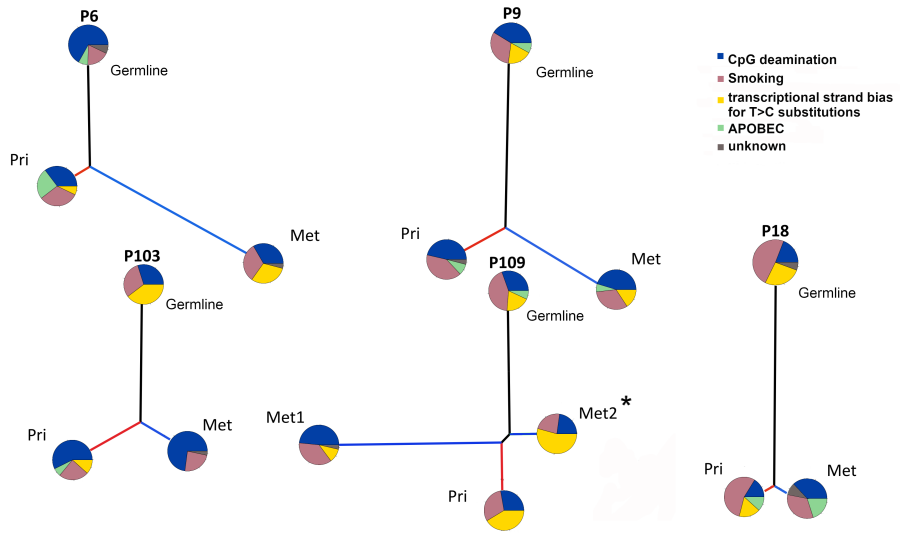


Figure 2. Evolutionary change of mutational signatures in primary tumor and metastases. Evolution of the somatic genetic alterations illustrates the changes in mutational processes. The pie chart delineates the proportion of COSMIC mutational signatures. Black, red, and blue lines represent the trunk, the primary LUSC and the metastasis branch, respectively. Branch lengths are proportional to the number of mutations. * This patient received a palliative radiotherapy and chemotherapy (Cisplatin/Vinorelbine and Cisplatin/Gemcitabine) prior to resection of Met 2.

Comment 4: Line 304: "LUSC" should be changed to "tumor"
Reply: We thank the reviewer and we apologize for this error. We changed the text accordingly.
Changes in the text:
(Please, see Page 17, line 8)
In breast cancer, Ng et al found significant differences in mutational signatures comparing private mutations of the primary ~~LUSC~~ tumor and metastasis (27).

Comment 5: Given that the mean coverage is ~50-70 for many cases and that any calls with <3 reads were excluded, this means that this assay would be expected to fail to detect any subclones with allele frequencies of <4-6% even in the ideal circumstances. This should be explicitly acknowledged and discussed.

Reply: We thank the reviewer for this question.
In order to remove possible sequencing and FFPE artefacts, we discarded all mutations validated by less than 3 reads and/or with an allele frequency lower than 1%. If a mutation was found in both, primary tumor and matched metastasis, we applied a cut-off of two reads. This is an important filter for FFPE samples, that reduces the false positive mutations and allowed us to calculate a more precise tumor mutation burden. We appreciated the reviewer's suggestion and we acknowledged and discussed this aspect accordingly in the discussion section.

Changes in the text:
(Please, see Page 18, lines 15 - 20)

In fact, one of the limitations of this study, due the possible presence of sequencing and FFPE artefacts, is to remove the mutations with an allele frequency lower then 1% and/or validated by at least 3 reads. This means, especially for the samples with a mean coverage of ~50-70 x, failing to detect any subclone with an allele frequency lower than 4-6% even in the ideal circumstances.

Comment 6: Line 481: reference #46 does not look correct

Reply: We apologize for this error. We changed the reference #46.

Changes in the text:
(Please, see Page 26, lines 1-3)
~~https://wwwcbioportalcom.~~ Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012;2:401-404.