IDENTIFICATION OF RECURRENT FUSION GENES IN ESOPHAGEAL SQUAMOUS CELL CARCINOMA PATIENTS FROM KAZAKHSTAN

Aibek Karimov[1], Rakhimbek Bektayev[1], Ruslan Kalendar[1], Aigul Sharip[1], Saule Rakhimova[2], Nazerke Satvaldina[2], Ainur Seisenova[1], Ainur Akilzhanova[2] and Ulykbek Kairov[1]*

[1]Laboratory of Bioinformatics and Systems Biology, Center for Life Sciences, National Laboratory Astana, Nazarbayev University, Kazakhstan.
[2]Laboratory of Genomic and Personalized Medicine, Center for Life Sciences, National Laboratory Astana, Nazarbayev University, Kazakhstan.

**Corresponding author (s): ulykbek.kairov@nu.edu.kz*

Background: Esophageal squamous cell carcinoma (ESCC) is one of the most lethal cancers in Central Asia, yet its molecular landscape in Kazakhstan remains poorly explored. However, the fusion-gene landscape in Kazakhstan remains undercharacterized. Fusion transcripts can serve as diagnostic biomarkers or therapeutic targets, motivating a focused survey in a local cohort.

Materials and methods: We analyzed paired tumor and adjacent normal tissues from ESCC patients collected in Kazakhstan (34 normal, 33 tumor samples). Tumor specimens were obtained from patients who had undergone Ivor-Lewis esophagectomy without receiving prior chemotherapy or radiotherapy. Total RNA was sequenced (paired-end), reads aligned to GRCh38 (Gencode v44 / CTAT genome library). Fusion calling was performed with two orthogonal tools (STAR-Fusion v1.13.0 and Arriba v2.4.0). Callsets were intersected to prioritize concordant events; candidate junctions were manually inspected in IGV to remove artifacts. Primers for prioritized fusions were designed and PCR followed by Sanger sequencing is underway for orthogonal validation.

Results: Intersection of STAR-Fusion and Arriba reduced noise and produced a concise set of high-confidence fusion candidates detected in multiple tumors and absent from matched normals. Manual IGV review confirmed strong split- and spanning-read support at exon boundaries for several recurrent events. Notable recurrent partners include MAP4K5-L2HGDH, ADAMTS2-ENSG00000253652, RAB8A-CIB3, BAZ2B-WDSUB1, HUWE1-SUPT3H, and PLEKHA5-KRAS.

Conclusion: Using complementary fusion callers and rigorous manual curation, we generated an assay-ready shortlist of recurrent fusion transcripts in Kazakh ESCC. Pending PCR/Sanger confirmation, these candidates warrant follow-up as potential biomarkers or mechanistic leads for ESCC in Central Asia.

Acknowledgement:

We thank all the participants in this study. This research has been funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grants No. AP23490594, BR18574184, BR24993023, BR24992841, BR27199879), Nazarbayev University funding CRP grants 021220CRP2222, 211123CRP1608.

Key words: esophageal squamous cell carcinoma, gene fusions, RNA-seq, Kazakhstan