Long-Read RNA-Sequencing for identification of novel isoforms and allelic context of driver oncogene mutations in the treatment-resistant setting

Kevin M. Levine, MD, PhD[1,2], Colette Felton PhD[3], Christina S. Baik MD, MPH[1,4], Angela N. Brooks PhD[3], Alice H. Berger, PhD[2],
1) Division of Hematology and Oncology, Department of Medicine, University of Washington, Seattle, WA
2) Human Biology Division, Fred Hutchinson Cancer Center, Seattle, WA
3) Department of Biomolecular Engineering, University of California Santa Cruz, Santa Cruz, CA
4) Clinical Research Division, Fred Hutchinson Cancer Center, Seattle, WA

Background
A recurrent clinical problem with targeted therapies for patients with advanced NSCLC is the development of resistance to these therapies. Some mechanisms of resistance have been identified; for example, for the EGFR inhibitor osimertinib, both on-target mutations, such as *EGFR* C797X, and off-target alterations such as MET amplification have been detected in ctDNA and tumor tissue. However, for the majority of patients, the mechanisms of resistance to newer targeted therapies are unknown. Long-read RNA-Sequencing, with 1kb and longer reads, can identify novel isoforms, differential expression of alternative splicing events, and phasing of mutations with respect to isoform usage. This newer technology may be able to overcome this gap in knowledge and inform better treatment strategies in the recurrent setting.

Methods
Through an IRB-approved research protocol at Fred Hutch Cancer Center titled MoRe (Mechanisms of Resistance to targeted therapies), we have consented over 40 patients eligible for targeted therapies for the prospective collection of serial blood draws and extra cores of tissue biopsies to be used for research sequencing. To date, we have performed long-read and short-read RNA-Sequencing on 4 treatment-resistant biopsy samples, all on tumors with *EGFR* driver mutations. We have also performed RNA-Sequencing on 32 matched tumor/normal pairs of surgically-resected NSCLC. RNA was isolated from fresh frozen tissue specimens and cDNA libraries prepared using the PacBio Kinnex full-length isoform method. Sequencing was performed at the University of Washington Long Reads Sequencing Center or at UC Davis DNA Technologies Core. Libraries were sequenced on the PacBio Revio system to a read depth of >10M HiFi reads per sample. FLAIR3 was used to identify and quantify isoforms.

Results
For the treatment-resistant biopsy samples, long-read RNA-Sequencing identified novel rearrangements in key oncogenes such as *EGFR*, that cannot be detected on short-read RNA sequencing or ctDNA sequencing. It also allowed for identification of the allelic context of treatment-resistant mutations in *EGFR*. For the paired tumor / normal samples, alternative splicing and differential isoform usage were identified in oncogenes such as *KRAS* and *BRAF*, which may impact sensitivity to targeted therapies. Ongoing work includes sequencing of additional tumor samples to try to identify a broader landscape of somatic alterations that are missed by earlier sequencing technologies.

Conclusion
Overall, long-read RNA-Sequencing is a technology that has significant promise for identifying novel rearrangements, alternative splicing, and phasing of mutations that warrant further study for their role in resistance to targeted therapy.