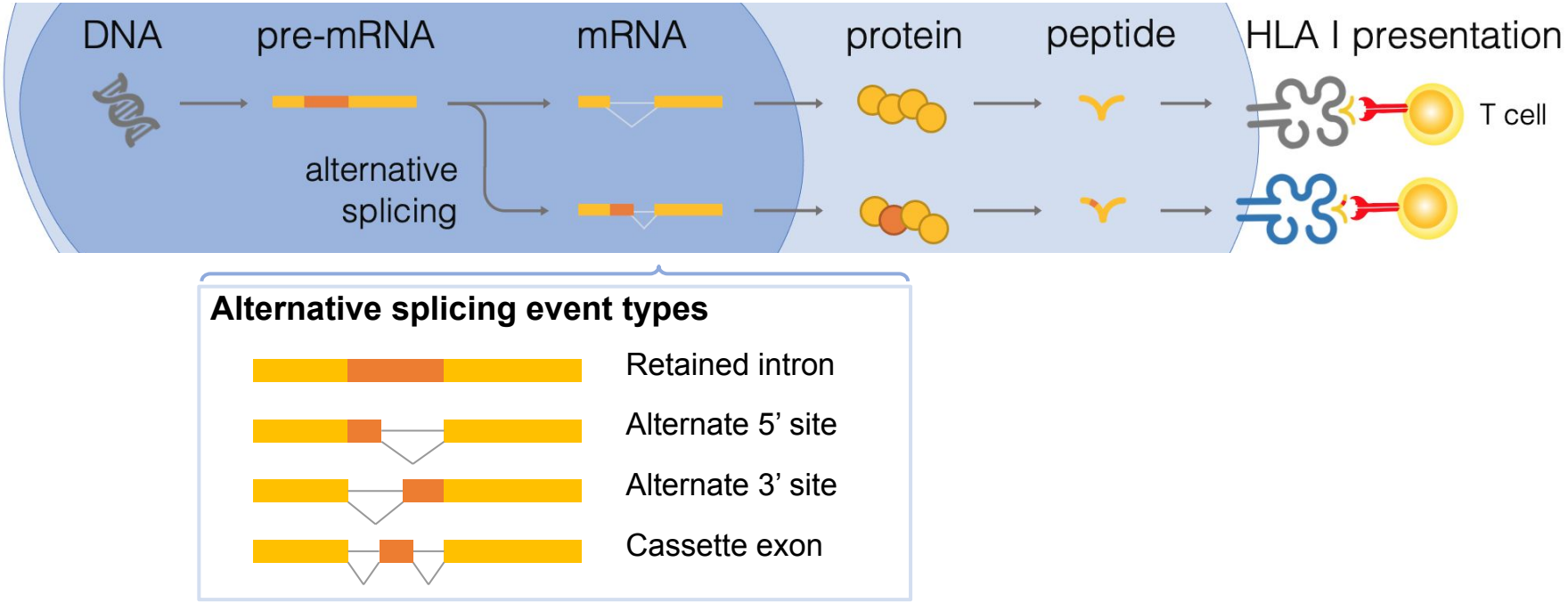
In silico prediction of alternative splicing-derived neoantigens in leukemia

Sarah Chen

MIT PRIMES Computational Biology
Under the direction of Nicoletta Cieri and Kari Stromhaug

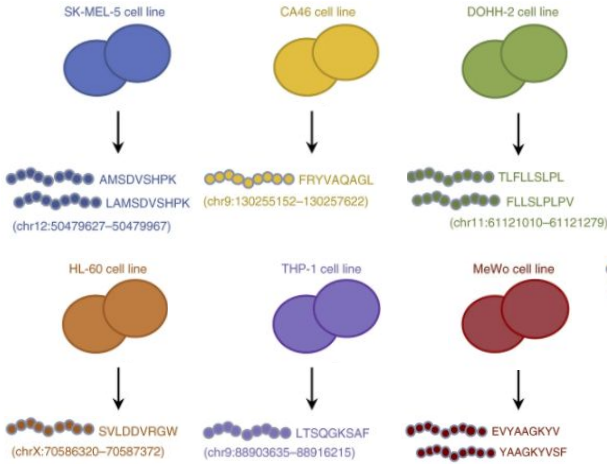
Introduction

Abnormal alternative splicing (AS) in cancer cells can yield tumor-specific isoforms, which are a potential source of neoantigens.



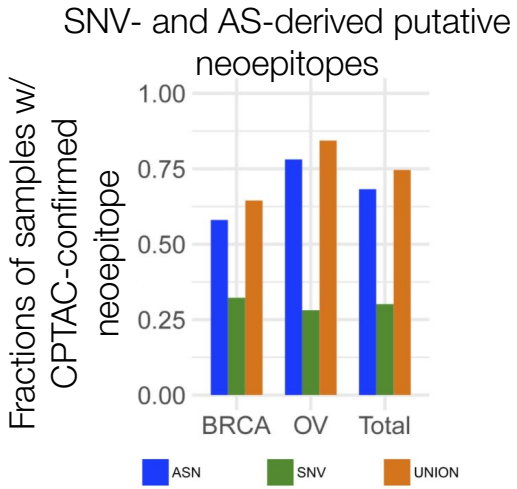
Alternative splicing in cancer is a potential source of neoantigens

Intron retention:



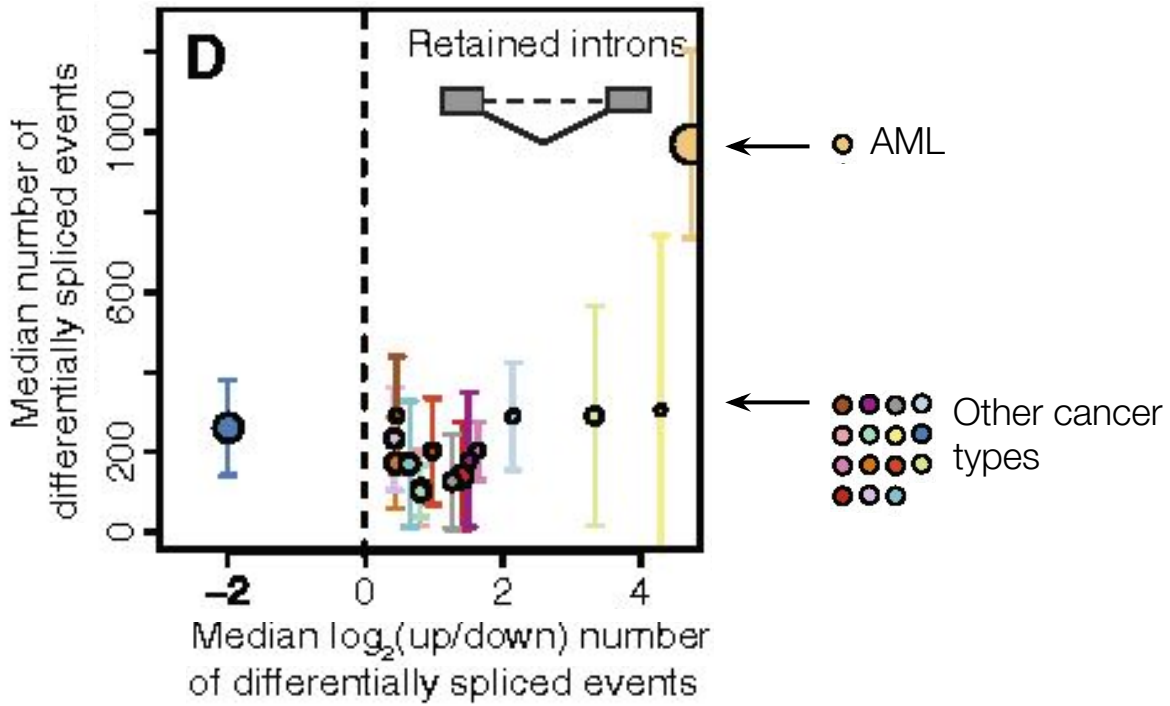
Smart et al. Nature Biotechnology. 2018

Alternative splicing:



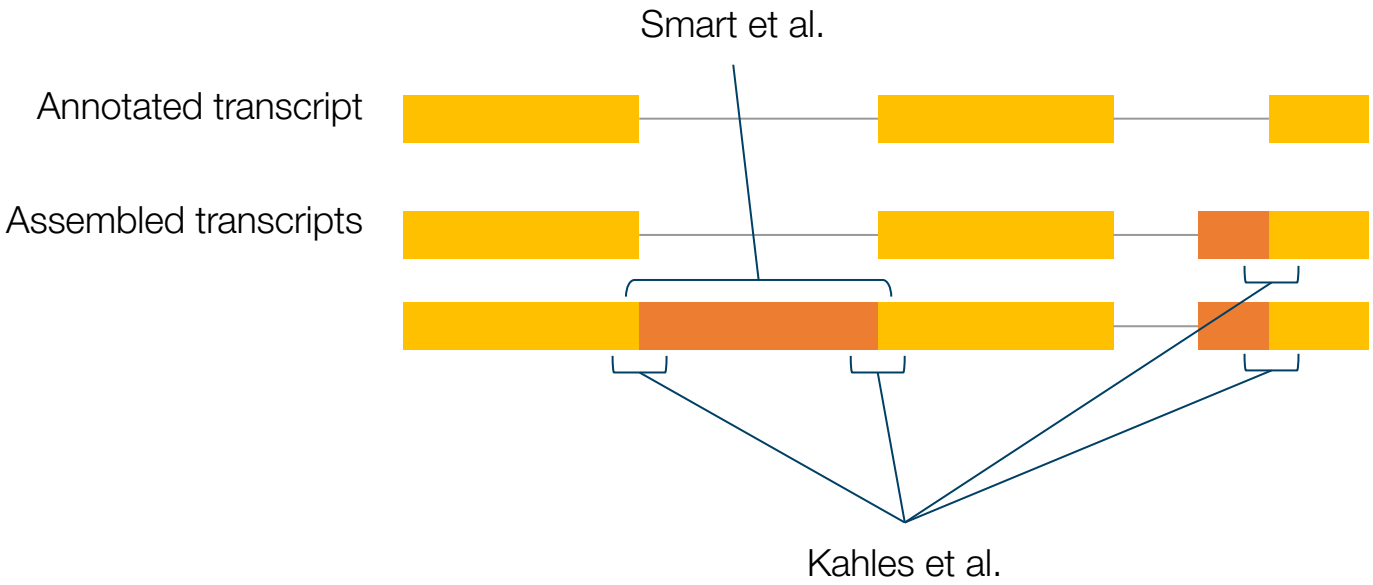
Kahles et al. Cancer Cell. 2018

Upregulation of intron retention is widespread in cancer and in AML in particular

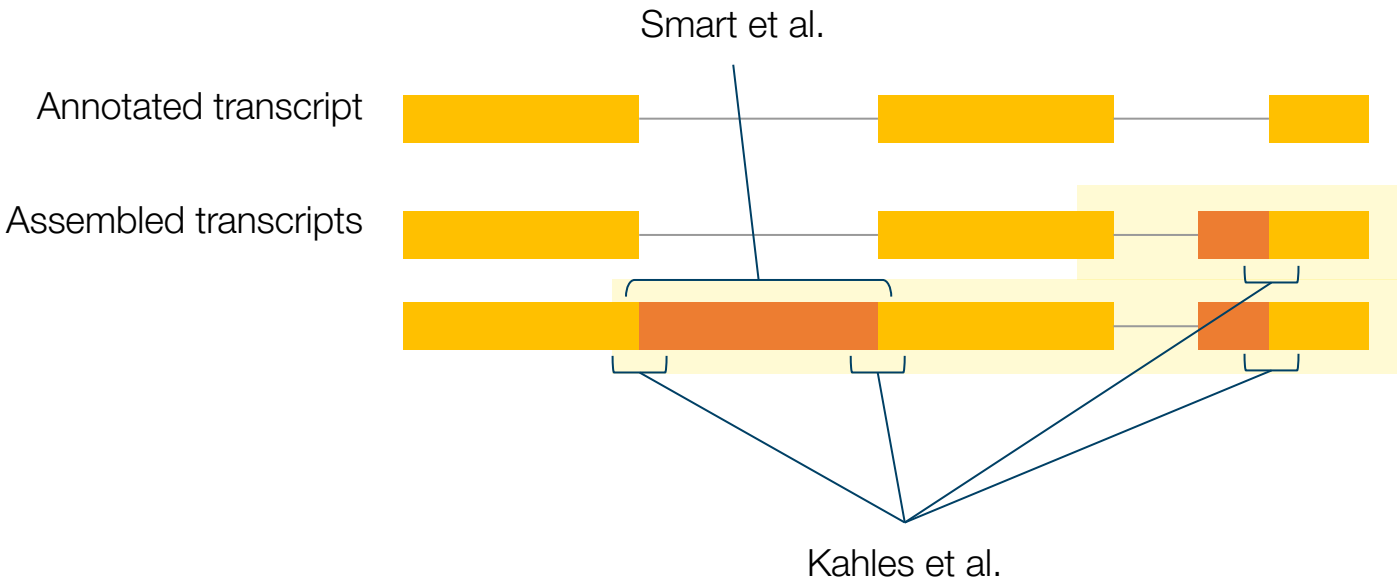


Dvinge and Bradley. *Genome Medicine* 2015

Past work failed to consider the full scope of potential derived neoantigens

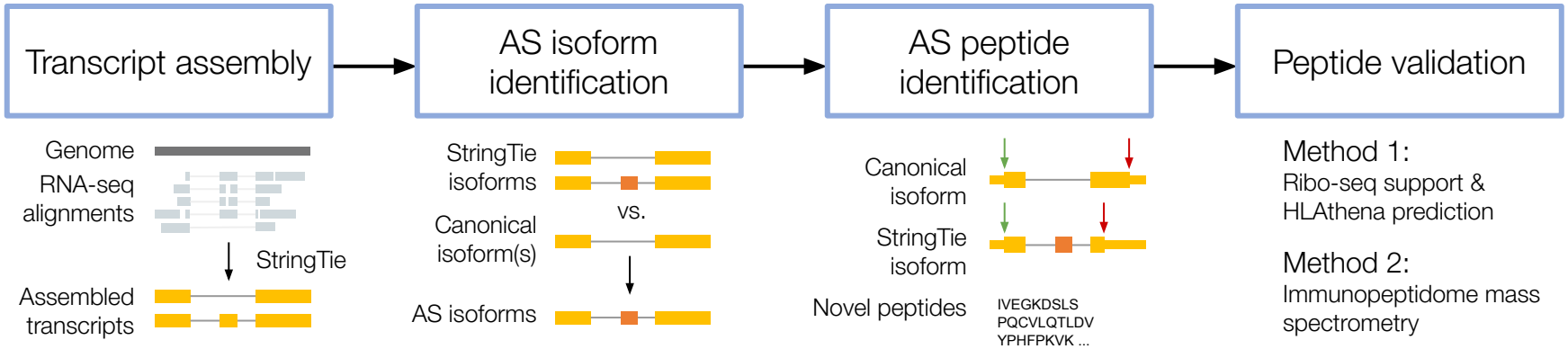


Past work failed to consider the full scope of potential derived neoantigens

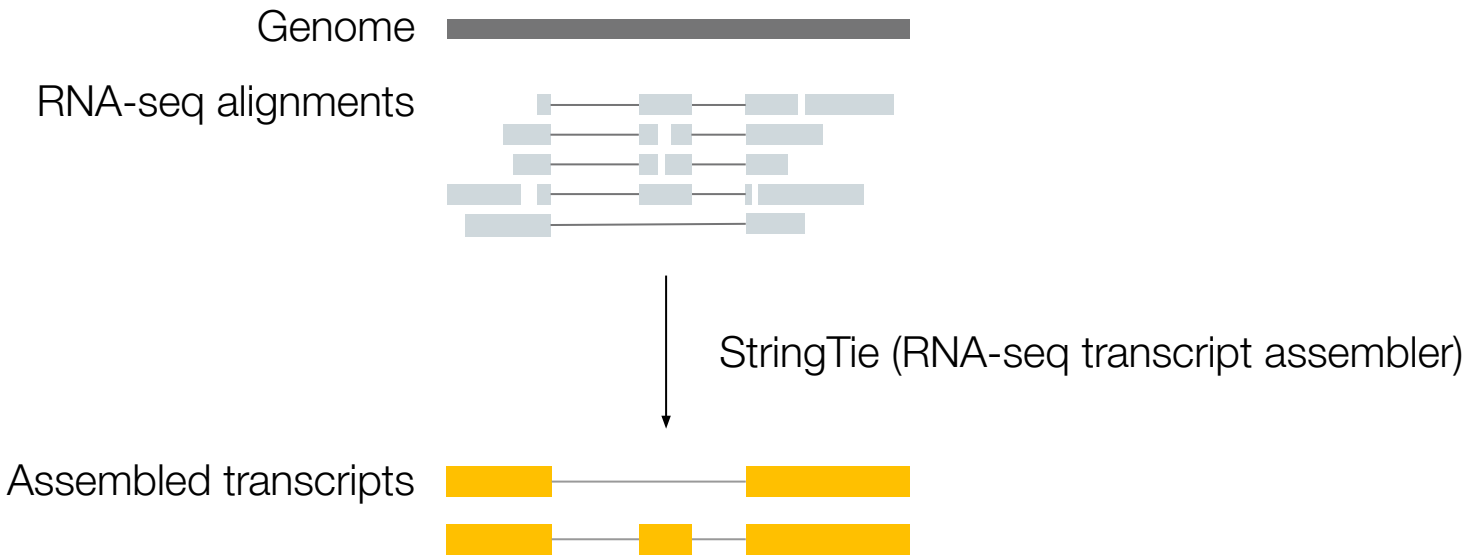


Pipeline Overview

We developed a computational pipeline to predict AS-derived neoantigens from RNA-seq data and validate them using ribosome profiling and immunoproteomics.



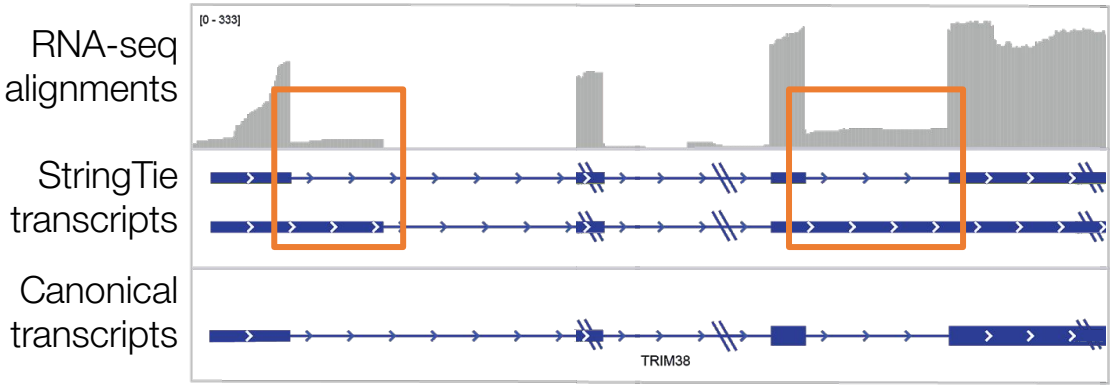
Transcript assembly via StringTie

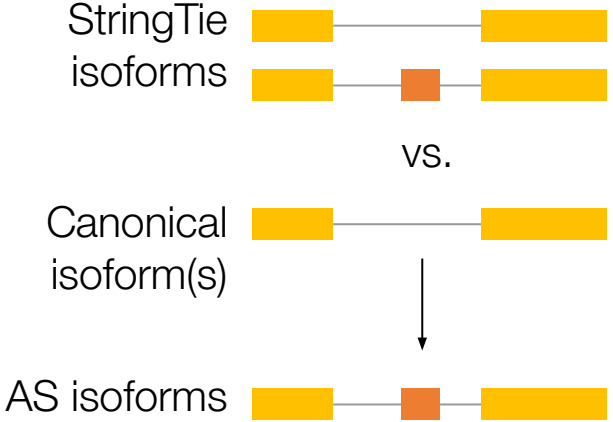


Pertea et al, Nat Biotech 2015

Identifying alternatively spliced isoforms

Example of alt 5’ site and retained intron:

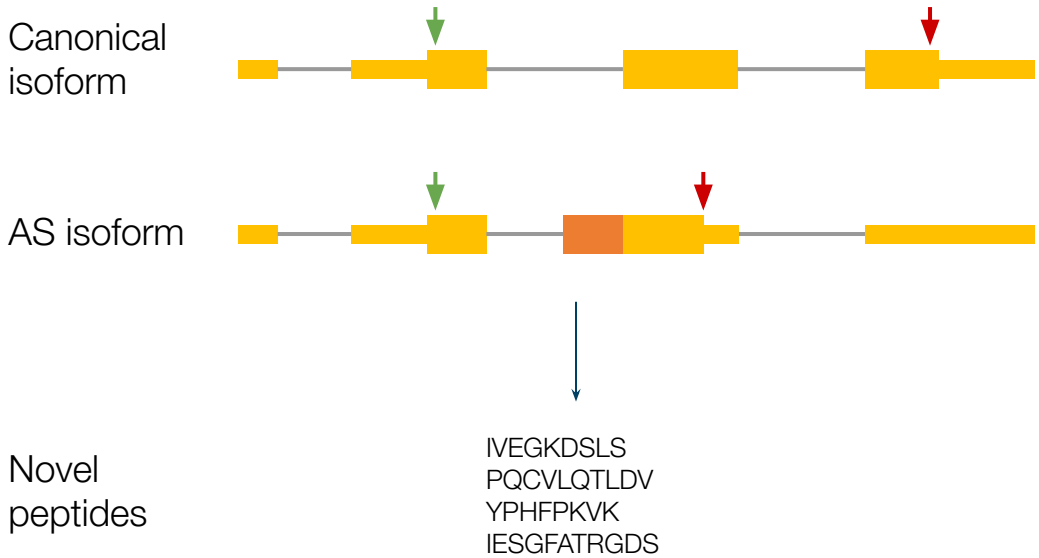




AS isoform identification

Translating alternatively spliced isoforms

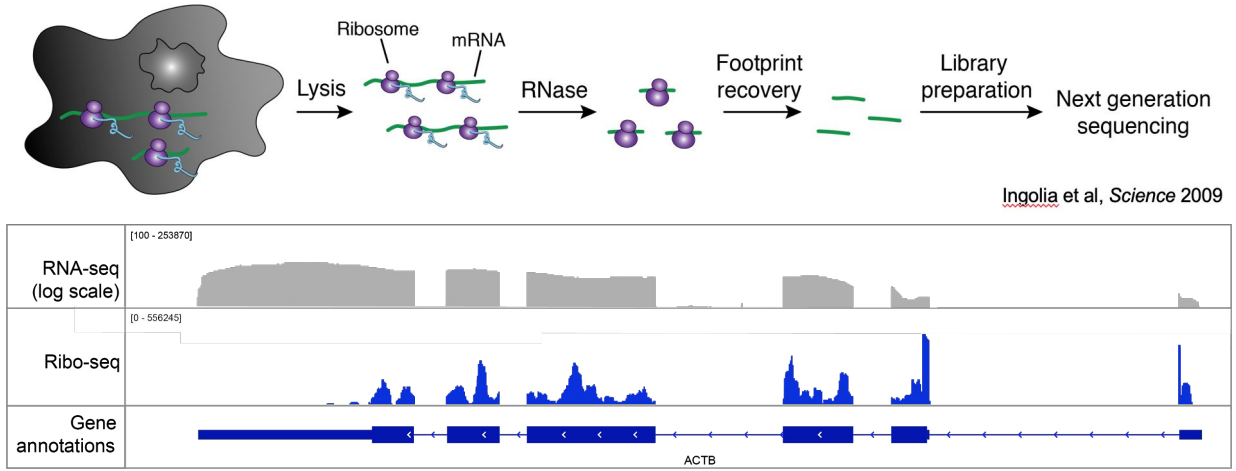
AS isoforms are translated from canonical start codons to the first downstream in-frame stop codon



AS peptide identification

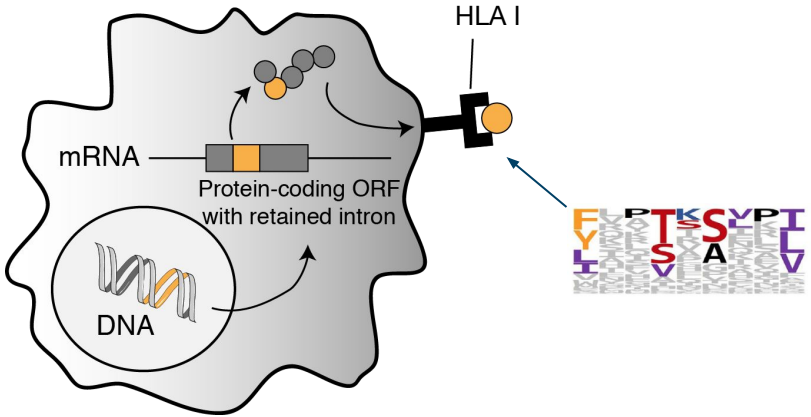
Validating predictions with Ribo-seq

Ribo-seq provides evidence that a sequence is **translated**



Validating predictions as HLA I binders

HLAthena predicts the likelihood that peptides bind to **HLA I**



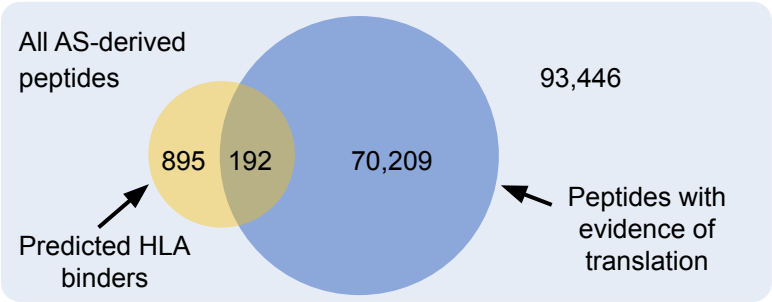
We considered peptides with a binding score the top 0.1% and 0.5% of HLAthena's background decoys

Figure adapted from Abelin et al, *Proteomics* 2019

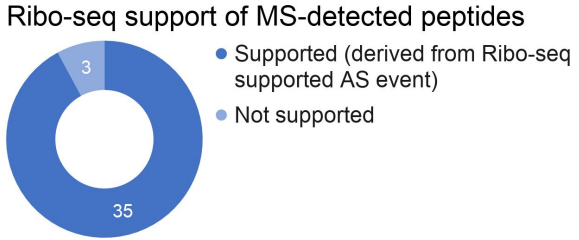
Results in the B721.221 model system

We identified 192 AS-derived peptides supported by Ribo-seq and predicted as HLA binders. Preliminary immunoproteomic analysis has validated 38 peptides (91% w/ Ribo-Seq support).

Ribo-seq and HLAthena validation



Immunopeptidome MS validation

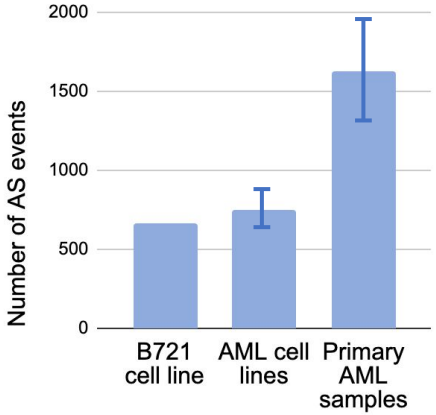


Across 40 HLA alleles, 38 peptides were detected, mapping to 65 AS events

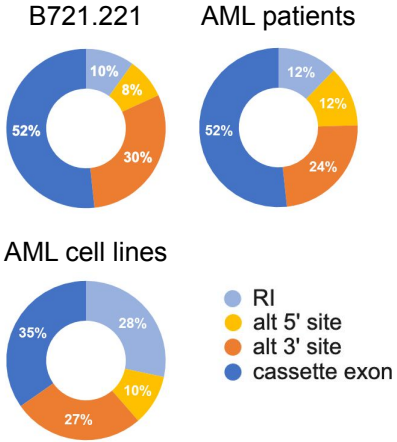
Results in AML cell lines and patients

We analyzed AML cell lines (n=8) and primary samples (n=7) to generate a patient-specific AS database to mine for potential neoantigens

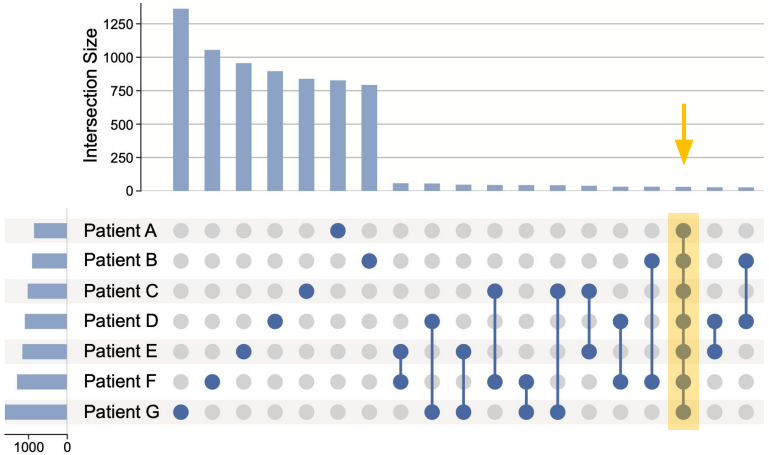
AS events in B721.221, AML cell lines, and AML patients



AS type distribution



31 AS events are shared across all patients

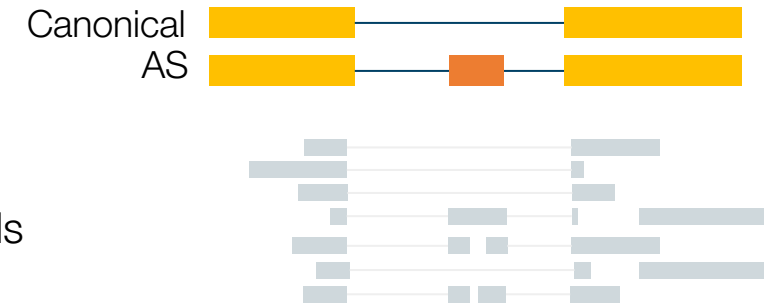


AS events shared across AML samples, omitting bins with less than 25 AS events.

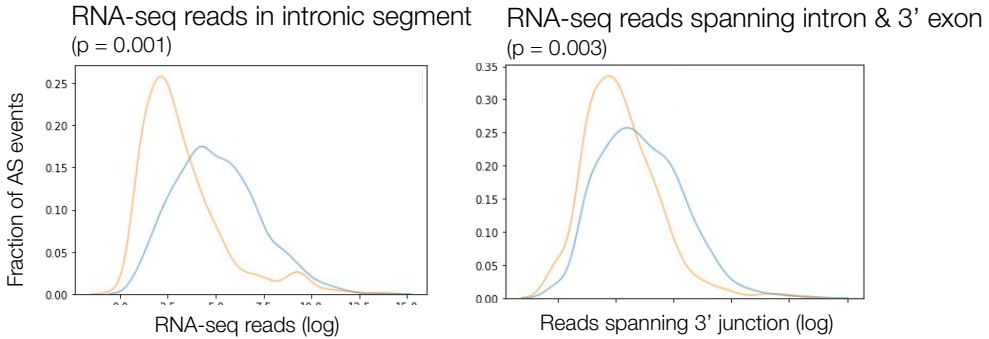
AS features may improve pipeline accuracy when Ribo-seq/MS data are unavailable

Features include:

- Read support of introns vs. adjacent exons
- Number of reads spanning intron-exon boundary
- Proportion of multi-mapping and/or indel reads in introns vs. adjacent exons

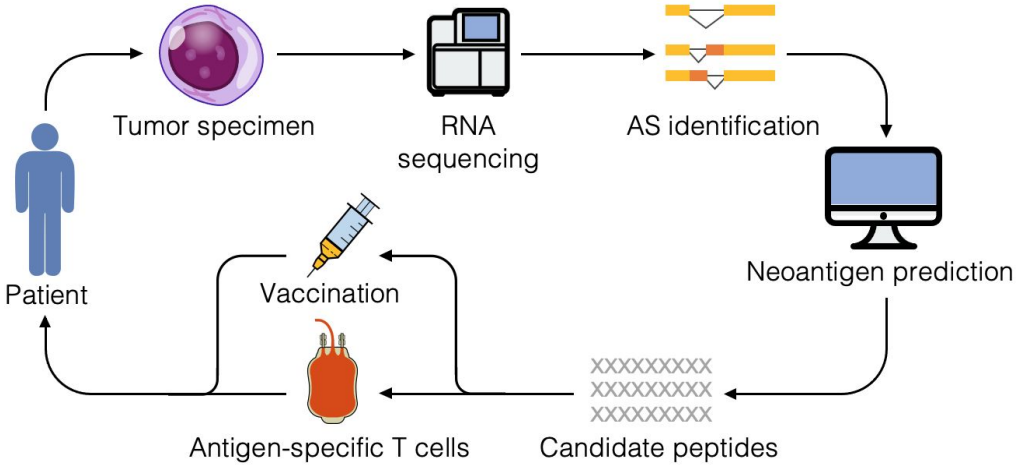


Filtering predictions based on RNA-seq features may enrich for true positives



Conclusion

- Alternative splicing is a promising source of neoantigens, especially for cancers with low mutation burden
- Current work focuses on increasing pipeline accuracy with RNA-seq features
- Future work will validate the cancer-specificity and immunogenicity of predicted AS-derived neoepitopes



Acknowledgements


Cathy Wu


Aviv Regev


Slava Gerovitch


Brian Haas


Karl Clauser


Nicoletta Cieri


Tamara Ouspenskaia


Kari Stromhaug


Travis Law