Expanded immunopeptide variant searching for hybrid DDA/DIA timsTOF workflows

Savannah Snyder[4], George Rosenberger[1], Gadi Armony[2], Ignacio Jauregui[3], An-phi Nguyen[4], Alexandros Pachos[4], Qixin Liu[5], Bin Ma[5], Josh Elias[6], Jonathan Krieger[7], Tharan Shrikumar[7], Tejas Gandhi[4], Lukas Reiter[4], Dennis Trede[8]

1 Bruker Switzerland AG, Fällanden, Switzerland; 2 Bruker Nederland BV, Leiderdorp, Netherlands; 3 Mestrelab Research S.L., Santiago de Compostela, Spain; 4 Biognosys AG, Schlieren, Switzerland; 5 Rapid Novor Inc., Kitchener ON, Canada; 6 Chan Zuckerberg Biohub, San Francisco CA, USA; 7 Bruker Ltd, Milton ON, Canada; 8 Bruker Daltonics GmbH & Co. KG, Bremen, Germany

Introduction:

Neoantigens—protein variants that differ from canonical sequences—are central to immune recognition and represent important targets for cancer immunotherapy. LC-MS/MS enables identification of HLA-presented peptides, including rare neoantigens, but the immense diversity of potential neoantigens creates significant computational challenges.

Here we introduce a hybrid workflow for DDA/DIA timsTOF datasets that combines de novo and database search strategies, enhancing neoantigen discovery even without genomic or transcriptomic references. Our approach was benchmarked using 33 patient-derived, MHC-I enriched samples and compared directly to a leading RNA-seq guided variant identification method, demonstrating its potential to advance personalized immunotherapy.

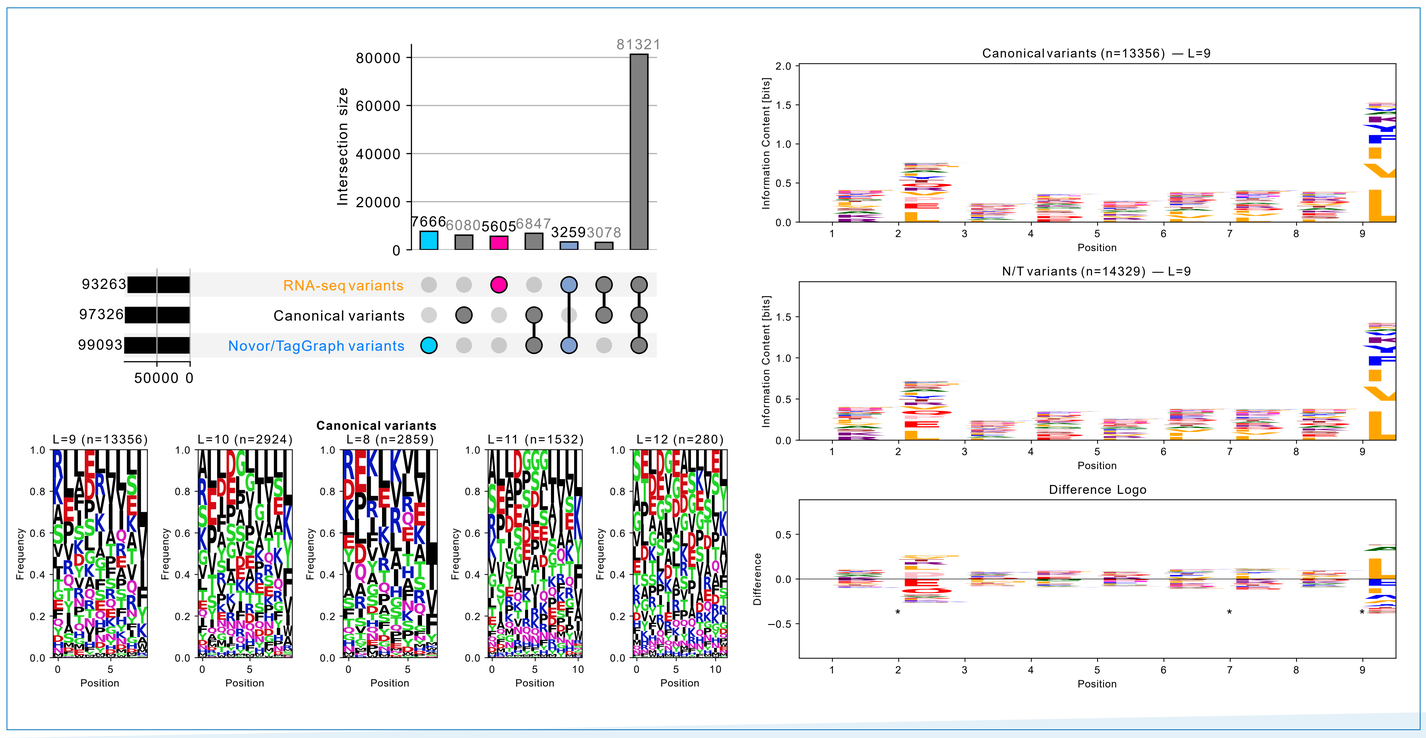
Materials and Methods:

Our workflow (Figure 1), implemented in Bruker ProteoScape (BPS), consists of two stages: (1) variant discovery and (2) validation and quantification.

In the first stage, DDA runs are processed using BPS Novor for de novo sequencing. The resulting spectra and sequences are mapped to canonical databases using TagGraph, generating a context-specific variant search space for each sample separately.

In the second stage, Spectronaut 20, equipped with immunopeptidomics-optimized AI models, uses these search spaces to validate and quantify peptides from DDA and DIA in run-wise fashion.

RNA-seq variant spaces were generated using the corresponding transcriptomic samples and a GATK4 variant calling pipeline.



Results:

We examined three HLA-I-enriched immunopeptidomic samples—primary tumor, metastasis, and normal tissue adjacent to the tumor—from 11 colorectal cancer patients using a timsTOF Ultra 2, along with matching RNA-seq profiles (Poster #XYZ). Each run was analyzed using a) the standard UniProtKB/Swiss-Prot database, b) an additional, RNA-seq-derived variant proteome, and c) an additional, expanded set of peptide sequences identified by BPS Novor and TagGraph.

Figure 2A shows that overall, canonical variants lead in identification rates. The hybrid approach (Novor/TagGraph) detected more peptides than either the canonical or RNA-seq variants alone. Specifically, Novor/TagGraph identified 36.8% of RNA-seq variants and detected 13.5% more variants compared to RNA-seq guided methods.

Figure 2B depicts the motifs for the canonical 8- to 12-mers.

To assess motif differences between the three methods, we used the "DiffLogo" algorithm, which also assesses statistical significance. There are no significant differences between Canonical and RNA-seq variants. However, there are significant (adj. p-value < 0.01) differences between Canonical and Novor/TagGraph variants (Figure 2C) for positions 2, 7 and 9 for the 9-mers.

Figure 1: Hybrid DDA/DIA peptide variant and discovery workflow in BPS. BPS Novor is used for real-time de novo sequencing. TagGraph discovers peptide variants, which are validated by Spectronaut using corresponding DDA/DIA runs.

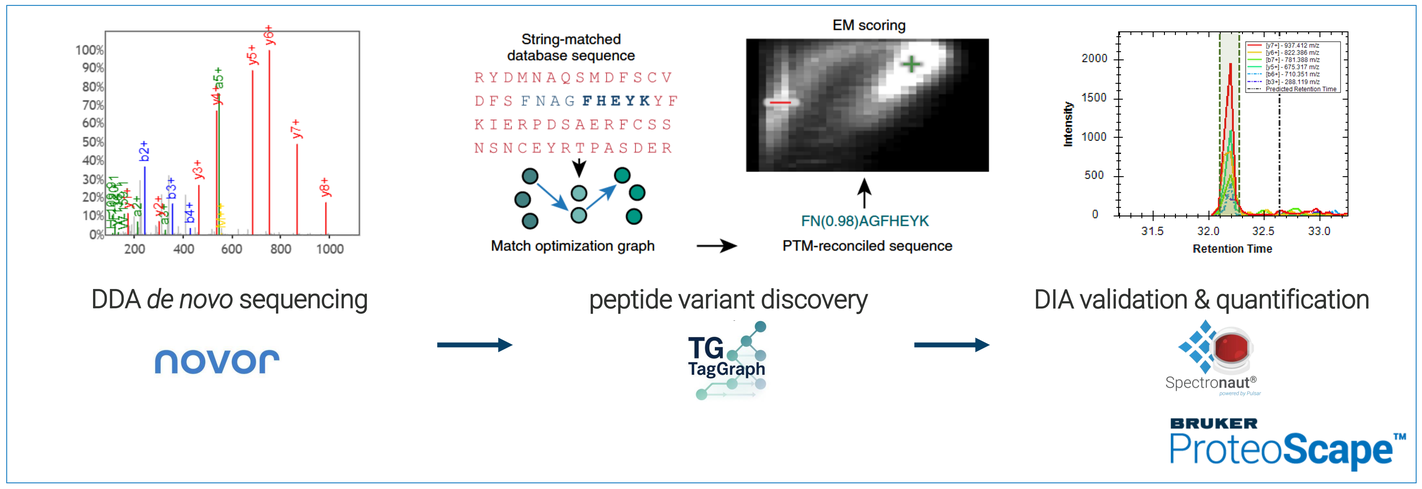


Figure 2. Comparative analysis of peptide identification and motif differences across three search strategies. (A) Peptide identification rates for canonical UniProtKB/Swiss-Prot, RNA-seq-derived variant proteome, and hybrid Novor/TagGraph approaches across HLA-I-enriched immunopeptidomic samples (primary tumor, metastasis, and adjacent normal tissue) from 11 colorectal cancer patients. Canonical variants show the highest overall identification rates, while the hybrid approach detects more peptides than either canonical or RNA-seq-guided methods alone, identifying 36.8% of RNA-seq variants and 13.5% more variants compared to RNA-seq-guided searches. (B) Sequence motifs for canonical 8–12-mer peptides. (C) Motif comparison for 9-mers using DiffLogo. No significant differences were observed between canonical and RNA-seq variants, whereas significant differences (adjusted p-value < 0.01) were found between canonical and Novor/TagGraph variants at positions 2, 7, and 9.

Conclusion:

This comprehensive workflow merges an enhanced DDA-based variant search with reliable DIA-based validation, utilizing AI-powered peptide-centric scoring. It provides a scalable and sensitive method for neoantigen identification, detecting a comparable number of candidates with significant overlap to RNA-seq, demonstrating promising potential for advancing personalized cancer immunotherapy.

Technology

© 2026 Bruker