Title: MONTE enables serial immunopeptidome, ubiquitylome, proteome, phosphoproteome, acetylome analyses of sample-limited tissues

Authors

Jennifer G. Abelin[1,2]*, Erik J. Bergstrom[1,2], Hannah B. Taylor[1,3], Keith D. Rivera[1,3], Susan Klaeger[1,3], Meagan E. Olive[1,3], Karl R. Clauser[1], C. Jackson White[1], Myranda Maynard[1], M. Harry Kane[1], Cadence Pearce[1], Suzanna Rachimi[1], Michael A. Gillette[1], Shankha Satpathy[1], Namrata D. Udeshi[1]* and Steven A. Carr[1]*

1 Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge, MA, 02142, USA
2 These authors contributed equally
3 These authors contributed equally

*Correspondence: jabelin@broadinstitute.org (J.G.A.), udeshi@broadinstitute.org (N.D.U.), scarr@broad.mit.edu (S.A.C.)

Abstract

Multiomic characterization of patient tissues provides insights into the function of different biological pathways in the context of disease. Much work has been done to serialize proteome and post-translational modification (PTM) analyses to conserve precious patient samples. However, characterizing clinically relevant tissues with multi-ome workflows that have distinct sample processing requirements remains challenging. To overcome the obstacles of combining enrichment workflows that have unique input amounts and utilize both label free and chemical labeling strategies, we developed a highly-sensitive **m**ulti-**o**mic **n**etworked **t**issue **e**nrichment (MONTE) workflow for the full analysis of HLA-I and HLA-II immunopeptidome, ubiquitylome, proteome, phosphoproteome and acetylome all from the same tissue sample. The MONTE workflow enables identification of a median of 9,000 HLA-I peptides, 6,000 HLA-II peptides, 10,000 Ub sites, 12,000 proteins, 20,000 phosphorylation sites and 15,000 acetylation sites from patient LUAD tumors. Because all omes are generated from the exact same tissue sample, there is less biological variability in the data enabling more robust integration. The information available in MONTE datasets facilitates the identification of putative immunotherapeutic targets, such as CT antigens and neoantigens presented by HLA complexes, as well as reveal insights into how disease-specific changes in protein expression, protein degradation, cell signaling, metabolic, and epigenetic pathways are involved in disease pathology and treatment.

Introduction

Mass spectrometry-based proteomics is a proven technology for characterization of cell surface immunopeptidomes along with cellular proteins and their post-translational modifications (PTMs)(Abelin et al., 2017; Admon and Bassani-Sternberg, 2011; Bassani-Sternberg et al., 2016; Choudhary et al., 2009; Kim et al., 2011; Marcu et al., 2021; Mertins et al., 2013; Sarkizova et al., 2020; Svinkina et al., 2015; Udeshi et al., 2020). We and others have made technological advances to measure each of these omes to deepen our understanding of tumor biology, infectious diseases, autoimmunity, and other diseases. However, patient tissues are

often limited in quantity. One of the main ways to overcome this limitation has been to serialize sample processing such that the flow-through of one enrichment step (e.g., phosphopeptide enrichment) is used as the input for the next enrichment step (e.g., acetylpeptide enrichment). Current serial enrichment workflows for measuring omes at high multiplex leverage isobaric reagents such as TMT(Dou et al., 2020; Gillette et al., 2020; Krug et al., 2020; Satpathy et al., 2021; Wang et al., 2021) or iTRAQ(Mertins et al., 2016, 2018). However, to date it has not been possible to include either immunopeptidome or ubiquitylome (i.e. anti-K-ε-GG antibody enrichment) in serial, multiplexed analyses as these 'omes' require that enrichment steps be carried out prior to or without labeling. In this scenario, conventional immunopeptidomics and ubiquitylomics protocols require separate sample aliquots as they cannot be enriched after chemical labeling because the immunopeptidome enrichment occurs pre-sample digestion(Weingarten-Gabbay et al., 2021) and enrichment of ubiquitylated peptides occurs prior to TMT labeling(Udeshi et al., 2013, 2020). These limitations impede the identification of cancer driver signatures and the detection of signaling network adaptations, changes in molecular complex formation and protein localization by concordant readout of the immunopeptidome, proteome and PTM-omes (ubiquitylome, phosphoproteome, acetylome).

There are two main complications precluding addition of immunopeptidome analysis in a serial processing strategy. The first is that much more sample input has typically been used for immunopeptidomics than for proteomic and PTM-omics in order to enable detection of low abundant, clinically relevant antigens such as neoantigens(Bassani-Sternberg et al., 2016; Marcu et al., 2021). For immunopeptidomics, a separate aliquot of tissue, usually 500 – 1000 mg of wet weight tissue is needed compared to 25-50 milligrams for serial, multiplexed proteomics and PTM-omics(Gillette et al., 2020; Krug et al., 2020). Moreover, sample preparation for immunopeptidomics is distinct from that used for conventional proteomics. Immunopurification of HLA molecules requires the use of native lysis buffer containing mild detergent to maintain protein conformations and solubilize membrane bound HLA proteins. In contrast, current serial proteome and PTM-ome enrichment protocols denature proteins using urea or SDS prior to tryptic digestion precluding upstream HLA peptide complex enrichment.

Historically, the ubiquitylome has also been difficult to analyze in high-multiplex for primary tissue samples due to the inability of the antibody used for enrichment of di-glycyl remnant (K-ε-GG) of formerly ubiquitylated peptides to recognize these peptides once the newly generated N-terminal glycine is labeled with TMT or iTRAQ reagents(Udeshi et al., 2013). To address this issue, we recently developed UbiFast, a highly-sensitive, rapid and multiplexed protocol for quantifying >10,000 ubiquitylation sites from cells or tissue in a TMT10plex(Udeshi et al., 2020). UbiFast labels K-ε-GG peptides with TMT reagents while the modified peptides are still bound to the anti-K-ε-GG antibody. The key to the UbiFast approach is the on-antibody labeling step that allows TMT labeling of peptide N-terminal amine groups and the ε-amine groups of lysine residues, but not the primary amine of the di-glycyl remnant which is protected from labeling by the antibody. Because the UbiFast method requires enrichment of K-ε-GG peptides prior to TMT labeling and sample mixing, the method has only been used to enrich ubiquitylated peptides from a separate sample aliquot in parallel to samples processed for whole proteome and other PTM-omes(Satpathy et al., 2021).

To overcome the challenges of serializing deep-scale immunopeptidome and ubiquitylome with proteome, phosphoproteome and acetylome profiling from a single tissue sample of limited quantity, we developed an integrated proteomics workflow that we term Multi-Omic Networked Tissue Enrichment (MONTE). This new method integrates our published methods for isolation of immunoprecipitated HLA-peptide complexes from as few as 50 million cells and 0.2g of clinical specimens(Abelin et al., 2017; Chong et al., 2020; Keskin et al., 2015; Sarkizova et al., 2020), in combination with recent improvements in MS instrumentation, off-line fractionation, and separation in the gas phase(Klaeger et al., 2021) to increase immunopeptidome yield. Together with an optimized version of a published semi-automated serial HLA-I and HLA-II enrichment protocol(Chong et al., 2018), these changes provided a high-throughput immunopeptidome workflow that we could put upfront of proteomic and PTM-omic workflows. The flow-through of the HLA immunopeptidome purification contains the intact cellular proteome that we make compatible with the current multiplexed, serialized multi-proteomics-workflow by use of SDS based lysis and tryptic digestion post HLA enrichment(HaileMariam et al., 2018). The resulting digest is then processed by the UbiFast workflow for multiplexed ubiquitylation profiling using anti-K-ε-GG antibodies and on-antibody TMT labeling(Rivera et al., 2021; Udeshi et al., 2020). The peptide flow-throughs of the UbiFast enrichment step containing unlabeled, non-K-ε-GG peptides are further processed for deep-scale and highly-multiplexed measurement of the proteome, phosphoproteome and acetylome using previously described methods(Dou et al., 2020; Gillette et al., 2020; Krug et al., 2020; Mertins et al., 2013; Satpathy et al., 2021; Svinkina et al., 2015; Wang et al., 2021).

MONTE generated datasets can be integrated to provide valuable information relating to the direct identification of disease specific HLA presented peptides that can be targeted with immunotherapeutic approaches as well as the status of protein expression, protein degradation, cell signaling, metabolic, and epigenetic pathways regulated by post-translational modification. By enabling multiple, key omes to be obtained on limiting amounts of exactly the same human tissue samples, MONTE overcomes prior limitations that have prevented concordant readout of the immunopeptidome, proteome and PTM-omes (ubiquitylome, phosphoproteome, acetylome) thereby enabling new insights into cancer and other disease biology.

Results

Networked serial enrichment of the HLA immunopeptidome, proteome, and post-translational modifications by MONTE

To address the challenge of deeply characterizing clinically relevant samples with limited cellular input, we serialized HLA-I and HLA-II immunopeptidomics with ubiquitylome, proteome, phosphoproteome, and lysine acetylome profiling workflows. The full Multi-Omic Networked Tissue Enrichment (MONTE) is represented in **Figure 1**. Three major changes were made to previously reported serial multi-omic enrichment protocols(Dou et al., 2020; Mertins et al., 2013, 2016, 2018; Udeshi et al., 2020; Wang et al., 2021) that were evaluated to ensure that each proteomic data type was not significantly biased. These changes included i) adding a protein

level, serial HLA-II and HLA-I immunopeptidome enrichment prior to the downstream multi-omics analyses listed above ii) switching from an 8M urea cell lysis to an SDS cell lysis and digestion on an S-Trap to facilitate removal of detergents present in the native lysis buffer used for HLA immunopurification and iii) putting UbiFast K-ε-GG peptide enrichment before serial, multiplexed proteome, phosphoproteome, and acetylome analysis. We also incorporated a modified version of a semi-automated, 96-well plate based HLA immunopeptidomics workflow (Chong et al., 2018) along with improved automated phosphopeptide enrichment(Abelin et al., 2016) and automated UbiFast K-ε-GG peptide enrichment(Rivera et al., 2021) workflows that enable higher throughput and reproducibility. The MONTE workflow allows for integrated proteomic analysis of the HLA immunopeptidome, proteome and post-translational modifications using a single sample of limited quantity.

UbiFast followed by serial multi-omic sample processing shows expected coverage, reproducibility and retainment of biological information

To conserve precious clinical and biological material for multi-omic measurements in the same samples, we first sought to integrate the UbiFast workflow with our well established TMT multiplexed proteome, phosphoproteome and acetylome workflow(Mertins et al., 2013, 2016, 2018). For this, we created the new workflow shown in **Figure 2A** that starts with the UbiFast method for enrichment and on-antibody TMT labeling of K-ε-GG peptide(Rivera et al., 2021; Udeshi et al., 2020). After UbiFast processing, flow-throughs from the antibody enrichment step contain non-K-ε-GG peptides that are not yet labeled with TMT. To enable serialization, these flow-through peptides are subsequently TMT labeled and used as input to generate proteome, phosphoproteome and other PTMome datasets.

Using tumors isolated from breast cancer patient-derived xenograft (PDX) models, representing Basal (WHIM2) and Luminal (WHIM16) subtypes of breast cancer(Li et al., 2013; Mertins et al., 2018; Udeshi et al., 2020) we previously enriched ubiquitylated peptides from each of WHIM2 and WHIM16 using 0.5 mg of peptide per sample. Here, we evaluated use of UbiFast peptide flow-throughs from these same samples using 0.3 mg peptides per state for serial proteome and phosphoproteome analysis to confirm that implementing UbiFast on samples prior to serial processing of the other “omes” has no significant effect on depth of coverage and biological information content of the experiments(Mertins et al., 2018; Udeshi et al., 2020). Analysis of UbiFast flow-through samples showed expected coverage of both the proteome and phosphoproteome when compared to previous data collected by Zecha *et al.* on these same PDX models (**Figure 2B**)(Zecha et al., 2019). We identified and quantified 8,954 human proteins and 39,328 human phosphorylation sites from UbiFast flow-through samples. The overlap of proteins and phosphorylation sites between experiments with and without serial UbiFast processing was high (83.9% proteome and 59.1% phospho sites) (**Figure 2B**)(Mertins et al., 2018). Pearson correlations of TMT ratios between intraplex replicates for UbiFast flow-throughs were high with median correlation of 0.98 for Basal samples and 0.97 for Luminal samples, indicating that UbiFast pre-processing does not negatively affect reproducibility (**Figure 2C**). Scatter plots of Basal/Luminal protein and phosphosite TMT ratios measured in UbiFast flow-through samples and data from Zecha et al. correlated well (R2 = 0.84 proteome

and R2 = 0.79 phosphosites). Finally, pathway enrichment of regulated proteins and phosphorylation sites correlated well with the top regulated gene sets in data acquired without pre-UbiFast processing(Zecha et al., 2019). Taken together, these results support the feasibility of incorporating multiplexed ubiquitylation profiling using UbiFast up-front and serially with multiplexed proteome and PTM profiling workflows.

The MONTE workflow achieves similar data depth when compared to a serial multi-omic enrichment workflow

We evaluated the impact of adding serial HLA-II and HLA-I enrichment prior to serial multi-omic enrichment workflow and compared the ubiquitylomes, proteomes, phosphoproteomes, and acetylomes of ten cryopreserved primary lung adenocarcinoma (LUAD) tumors from the CPTAC cohort(Gillette et al., 2020). Human tumor samples are more relevant than tumor samples derived from immunocompromised mice, and in testing, we found that the yield of HLA-I and HLA-II immunopeptidomes from the PDX models was very low (**Table S1**). LUAD tumors were selected because lung tissue is known to have HLA-I and HLA-II expression (He et al., 2017; Wosen et al., 2018) and one LUAD primary tumor has been profiled successfully using serial HLA-I and HLA-II immunopeptidomics(Chong et al., 2020). The human LUAD tumors (50-70 mg wet weight tissue, each) were lysed in SDS and processed using MONTE (**Figure 1**) with and without serial HLA enrichment (**Figure 3A)**. S-Trap-based protein digestion(HaileMariam et al., 2018) was used instead of 8M urea digestion post HLA enrichment because we have previously shown that serial HLA immunopeptidome and downstream whole proteome analysis required the removal of detergents present in native lysis buffer used for HLA enrichment(Weingarten-Gabbay et al., 2021).

Label free, antibody-based, serial HLA enrichment identified a median of ~12,000 HLA-I (8,750-14,645) and ~6,000 HLA-II (1,273-11,283) bound peptides from each of these ten LUAD tumors (**Figure 3B**). In comparison, a published serial HLA immunopeptidome analysis of a LUAD tumor reported 8,340 HLA-I and 4,098 HLA-II bound peptides(Chong et al., 2020), but the amount of sample used as input was not provided preventing a direct comparison. Regardless, our depth of >9,000 HLA-I peptides from as little as 50mg cryopulverized tumor corresponding to ~2 mg protein lysate was encouraging, and clearly indicated that the method could be used with even smaller amounts of input tumor material as needed, with some expected loss in depth. We confirmed that the observed HLA-I and HLA-II peptides had the expected length distributions (**Figure 3C, D**) and HLA-I binding characteristics (**Figure 3E, F**) using a motif analysis and the HLA-I presentation predictor HLAthena(Abelin et al., 2017; Sarkizova et al., 2020). We did note that Patient C3N-01416 had a larger representation of 8mers in the HLA-I immunopeptidome, which was expected because of the known preference for 8mers presented by HLA-B*18 alleles. Although not observed in our data, we believe HLA-I presented peptide allele assignments via epitope prediction algorithms could be used to estimate differences or losses of HLA-A, -B, and C allele expression. Loss of heterozygosity (LOH) is a known tumor immune evasion mechanism that may be directly accessible by HLA-I immunopeptidomics.

The intact protein flow-throughs from HLA immunopeptidome enrichments were next digested with Lys-C and trypsin using S-Traps in parallel with half of each LUAD tumor that was not HLA enriched. A summary of the resulting depths of these head-to-head proteomes, ubiquitylomes, phosphoproteomes, and lysine acetylomes are shown in **Figure 4A**. The proteome and ubiquitylome results demonstrate that similar numbers of proteins (11,089 vs. 10,621) and K-ε-GG peptides (10,694 vs. 10,594) were identified between the non-HLA enriched ("No HLA") and HLA enriched ("HLA FT") samples, respectively. A 19% decrease in the total number of phospho-sites (-9% phosphorylated proteins) was observed when using the HLA enriched samples (No HLA: 29,703 vs. HLA FT: 24,084 phosphosites), suggesting that the phosphatase inhibitors added to our lysis buffer may be losing their activity during the protein-level, HLA immunopeptidome enrichment. The addition of phosphatase inhibitors 3hrs into the 6hr serial HLA-IP step may improve the recovery of phosphopeptides.

We unexpectedly observed a >50% increase in lysine acetylated peptides and significantly higher enrichment efficiency using HLA enriched samples. Given that the protein lysates are incubated at 4°C for 6 hr during HLA enrichment, we sought to rule out the impact of non-enzymatic acetylation(James et al., 2018). Thus, we performed acetylome analysis of A375 melanoma cells with and without HLA immunopurification prior to S-Trap-based digestion by varying different experimental conditions, such as incubation time, the inclusion of KDAC inhibitors, and lysis methods (**Table S2**). Our label-free single shot acetylome analysis demonstrated the expected increase in recovered acetylpeptides by using HDAC inhibitors and that non-enzymatic acetylation during the 6hr serial HLA-IP step was likely not the cause for such a significant boost in acetylated peptides, as the sample lysed in the native lysis buffer for a short period of time resulted in a similar number of acetylated peptides when compared to the HLA enriched samples. We also observed a small increase in the numbers of acetylated peptides between the conventional S-Trap (SDS/S-Trap Lysis) and HLA IP lysis buffer (HLA lysis) conditions, but the magnitude of the difference was relatively small, likely due in part to the use of unfractionated single-shot used for the A375 experiment vs. four fractions for the LUAD samples. Another possible explanation for the large increase in LUAD acetylation sites could be due to pre-clearing of non-specifically binding components in the complex tissue lysates by HLA- and K-ε-GG antibodies, resulting in increased enrichment efficiency and total identified acetylated peptides.

The MONTE workflow recapitulates biological signals observed in a serial multi-ome enrichment workflow without HLA enrichment

To assess potential differences between the information content of the HLA enriched and non HLA enriched samples, we analyzed ten LUAD tumor proteomes, ubiquitylomes, phosphoproteomes, and acetylomes using a principal component analysis (PCA) (**Figure 4B**). PCA analysis shows that samples cluster by LUAD tumor, not by the processing method used, demonstrating that biological differences among the samples are stronger than technical variation between these serial workflows. The acetylomes appeared to cluster the least, likely due to the significant increase in acetylated peptides observed when using the flow-through from the HLA enrichment step. The total number of proteins identified and quantified from HLA

enriched and non-HLA enriched samples were shown to have an 84% overlap (**Figure 4C**). Slightly fewer proteins (4%) were identified from the HLA enrichment flow-throughs due to the depletion of HLA complexes and associated binding partners, including a minor depletion in the proteins involved in the microtubule organizing center (MTOC). This observation may relate to the direct binding of MTOC proteins with HLA peptide complexes embedded in membranes, as he HLA-II presentation pathway requires endocytic compartments that contain motor proteins like dynein and kinesin(Rocha and Neefjes, 2008; Vyas et al., 2007).

The overlap between HLA enriched and unenriched protein lysates was 55% for ubiquitylation sites, 62% for ubiquitylated proteins, 57% for phosphorylation sites and 64% for phosphorylated proteins, which is an expected result using multiplexed, data-dependent LC-MS/MS methods for highly similar processing workflows(Mertins et al., 2018) (**Figure 4C, D**). The lowest overlap across experiments was observed for acetylome data because more than twice as many acetylated peptides were observed in the HLA enriched samples. The summed TMT reporter ion intensities of the proteins and modified peptides identified with and without HLA enrichment (**Figure 4E**) suggests that the acetylated peptides identified without pre-HLA enrichment are lower in abundance than corresponding non-HLA enriched samples. This observation requires further investigation to fully understand if the HLA enrichment step removes non-specific binding sample components that make the acetylated peptide enrichment more efficient. Overall, the HLA enriched samples capture the same depth of coverage observed in non-HLA enriched samples and adding this enrichment step upfront in a serial workflow does not introduce detectable bias in downstream proteome ubiquitylome, phosphoproteome, and acetylome.

Immunopeptidome data from the MONTE workflow enables the identification of HLA peptides from mutated and cancer-testis antigen source proteins

The addition of HLA-I and HLA-II immunopeptidomics to the serial PTM enrichment workflow enables access to patient tumors not previously available for analysis, as tumor tissues were dedicated to either immunopeptidome or proteomic analysis because of the large input requirements. Initially, we looked in our LUAD immunopeptidome data for peptides derived from CT antigen source proteins reported in the cancer-testis database (CTdatabase) that have been associated with non-small-cell lung cancer tumors(Almeida et al., 2009; Djureinovic et al., 2016). Across the set of LUAD tumors, peptides from ten CT antigen source proteins were detected in our immunopeptidome data (**Figure 5A**), with most peptides being presented by HLA-I. Surprisingly, ten unique peptide:patient combinations, from the bromodomain testis-specific protein (BRDT) were presented by our LUAD tumor set. In fact, 6 of the patients presented at least one peptide from BRDT, making it a candidate for future immunogenicity investigations.

Next, we sought to determine if immunopeptidome data could be used to assess if a tumor is either “hot” or “cold” in terms of immune infiltration because primary tumors are a mixture of cell types that include tumor, stromal, and immune cells. To do this, we evaluated the total number of HLA-I and HLA-II peptides derived from proteins annotated as being part of the interferon gamma (IFNγ) signaling pathway (**Figure 5B**). Unlike HLA-I that is expressed by almost all

nucleated cell types in the body, HLA-II is constitutively expressed by antigen presenting cells and can be induced on other cells by IFNγ(Taylor et al., 2021). Therefore, higher overall HLA-II peptide presentation combined with the presence of peptides from IFNγ signaling proteins may together be indicative of a tumor with high immune infiltration. The tumors from Patient C3N-00579 and C3N-01416 demonstrated the highest overall numbers of HLA-II peptides derived from IFNγ signaling pathway proteins when compared to HLA-I, and these two tumors had the highest ESTIMATE ImmuneScores(Yoshihara et al., 2013) (**Table S3**). The tumor from Patient C3L-01632 had a similar behavior in terms of higher HLA-II peptides derived from IFNγ signaling pathway proteins, yet this tumor was not annotated as having an extraordinarily high ESTIMATE ImmuneScore. Instead, this tumor had the third highest StromalScore, suggesting that fully understanding the contributions of HLA-II presentation by stromal cells vs. infiltrating immune cells in the tumor microenvironment via immunopeptidomics will require more data from a larger patient cohort.

Reported data on the detection of HLA-I binding neoantigens from mostly high mutation tumors such as melanoma suggests that neoantigens are presented at levels as low as 0.01%(Bassani-Sternberg et al., 2016; Olsson et al., 2021; Yadav et al., 2014). Therefore, an immunopeptidome depth of ~10,000 peptides, similar to the depth of HLA-I peptidomes observed across the ten LUAD tumors, may yield at least one detectable neoantigen. In order to assess if our data depth enables the detection of HLA peptides containing mutations specific to each patient, we first mined our LUAD immunopeptidome data for HLA peptides derived from mutations detected across multiple tumor types reported in TCGA(Liu et al., 2018) (**Figure 5C**). Seven of the ten patients (70%) had at least one mutation containing HLA-I peptide mapping to a shared tumor mutation detected in TCGA data, while no HLA-II peptides containing shared mutations were identified. The most frequently presented frequent TCGA mutation was a point mutation (H33Q) in IDUA, an enzyme found in lysosomes, that was presented by three of the ten patients on either HLA-B*55:02 or -B*37:01. We also observed HLA-I peptides containing a frequent TCGA mutation presented by at least two of the patients from SEMA3A (point mutation) and VHL (frameshift).

Although these HLA-I peptides cover mutations, they are not automatically candidate neoantigens because sequencing data is needed to confirm these mutations are somatic. To do this, we obtained previously published whole exome sequencing (WES) data from both blood (normal) and tumor from these LUAD patients(Gillette et al., 2020). By comparing tumor and the blood normal, we determined that the mutations in IDUA (10/10 patients) and VARS (8/10 patients) are germline, and therefore not neoantigens. Interestingly, ATP6AP2 and SEMA3A mutations were not called in any patient and will require further validation before they can be confirmed as putative neoantigens. Next, we reanalyzed the HLA-I immunopeptidome data using databases generated with QUILTS that contained somatic mutations unique to each LUAD patient(Cibulskis et al., 2013; Ellrott et al., 2018; Ruggles et al., 2016) (**Figure 5D**). From this reanalysis, we observed that patients with high mutation burden were most likely to have LC-MS/MS detectable neoantigens, as three of the ten patients (30%) had at least one neoantigen present in their HLA-I immunopeptidome. Patient C3N-00547 with highest mutation burden (~1400 somatic single amino acid variants (SAAVs)) had three detectable neoantigens.

Alternatively, patient C3N-00199 was the patient with the lowest number of SAAVs (492) that resulted in at least one detectable neoantigen. These data demonstrate that a combination of mutation burden and HLA immunopeptidome depth are critical considerations when developing methods that leverage LC-MS/MS for neoantigen detection across large patient cohorts. Overall, the identification of HLA presented neoantigens and peptides derived from CT antigen source proteins provides data to better understand the immunopeptidome landscape of LUAD tumors, confirm the accuracy of neoantigen predictions, and provide putative targets for future immunotherapies.

Discussion

Patient tissue samples with limited input amounts create obstacles for deep proteomic characterization. To overcome this challenge, we developed the MONTE workflow, which enables serial HLA-I and HLA-II immunopeptidome profiling followed by ubiquitylome, proteome, phosphoproteome, and acetylome analysis of each single sample. Previously, both HLA immunopeptidome and ubiquitylome analyses required a dedicated tissue aliquot, preventing their application to large-scale proteomic efforts that utilized serial enrichment workflows(Gillette et al., 2020; Satpathy et al., 2021). With the combination of the MONTE workflow and improved sensitivity of mass spectrometry instrumentation, we were able to profile a set of primary LUAD tumors starting with as little as 50mg of cryopulverized tissue, and further reductions in the amounts of input tissue required are likely in the near future. In addition, genomic analyses including Whole Exome Sequencing and RNA-Seq were previously performed on aliquots of the cryopulverized tissue to prevent tumor microheterogeneity from confounding subsequent proteogenomic data integration(Gillette et al., 2020). Because the exact same sample is used for all HLA and multi-ome analyses, MONTE datasets enable comprehensive investigations into how changes in protein abundance, phosphorylation signaling, ubiquitination mediated degradation, and acetylation involved in epigenetic regulation and metabolism pathways are related to antigen presentation in the context of cancer and other disease states.

Two technical hurdles had to be addressed to fully serialize the MONTE process. First, the protein level HLA immunopeptidomics was performed prior to the serial proteome and PTM-ome enrichment workflow, which required intact proteins to be captured and digested on S-Traps. Second, the enrichment of K-ε-GG peptides by UbiFast was completed before TMT labeling of peptides (at lysines and free N-termini) and other PTM enrichments because the anti-K-ε-GG antibody will not capture TMT-GG-ε-K. After implementing these changes into the serial workflow, we observed high correlation between the proteomes, ubiquitylomes, phosphoproteomes and acetylomes in both our breast cancer xenograft and LUAD datasets, and we were able to recapitulate expected biological signals. These datasets demonstrated that putting HLA and UbiFast enrichment upfront of the serial enrichment workflow does not negatively affect downstream proteome and PTMome data. The most striking difference was observed in our LUAD acetylome analysis, where the HLA enriched samples appear to have an increased total number of recovered acetylated peptides along with increased overall intensities. We believe this improvement may be due to enhanced sample preparation via the removal of

non-specific binders prior to enrichment. Although this may be a promising result, more experiments in tissue are needed to fully elucidate the cause of the boost in acetylated peptide recovery and determine if a similar impact is observed across diverse tissue types and cell lines.

Integration of proteome and PTM datasets has deepened our understanding of the biological underpinnings of disease(Gillette et al., 2020; Krönke et al., 2014; Mertins et al., 2016; Satpathy et al., 2020, 2021), but to date HLA peptidomics has not been routinely employed and ubiquitylomics is only now being used to profile human tumor samples. Ubiquitylome (K-ε-GG) profiling done in serial on the same samples should, for example, enable crosstalk between lysine ubiquitylation and acetylation to be discovered and, combined with outlier analyses(Blumenberg et al., 2019; Gillette et al., 2020) identification of activated ligases and deubiquitylases, potentially lead to new targets for therapeutic intervention. Layering of HLA-I and HLA-II immunopeptidomes on these other data types provides a window into the antigen landscape of cancer and other disease types. The combination can help guide the selection of immunotherapy targets, such as patient specific neoantigens and tumor associated antigens and their application to personalized vaccines and T cell therapies, as well as provide insights into how cell signaling, protein degradation, metabolic and epigenetic protein pathways impact the likelihood a specific source protein is processed and presented by HLA-I and HLA-II. These datasets will continue to improve our understanding of the rules that govern antigen processing presentation, and advance algorithms used for HLA epitope predictions(Abelin et al., 2017, 2019; Bassani-Sternberg and Gfeller, 2016; Jurtz et al., 2017; Sarkizova et al., 2020). MONTE datasets may also enable PTM-containing HLA epitope predictions, as both the modified source protein and presented HLA peptide can be directly identified from the exact same sample. Moreover, these fully integrated datasets may provide information regarding tumor immune cell infiltration status and dysregulation of signaling, degradation, and epigenetic pathways that may affect disease progression and treatment options.

High-throughput multi-omic data generation has proven to be a useful resource for understanding disease biology and the identification of therapeutic targets. By combining serial multi-ome enrichments with HLA-I and HLA-II immunopeptidomics into a single workflow, we have provided a path forward to understanding connections between antigen presentation and protein expression, signaling, protein degradation, and epigenetic regulation from a single sample, which was not previously possible. Further improvements to the MONTE workflow that address its current limitations will likely include decreasing the sample input further by incorporating low-input proteomic sample processing advances(Hartlmayr et al., 2021; Myers et al., 2019; Satpathy et al., 2020; Yi et al., 2019) and on-line, gas-phase fractionation technologies such as ion mobility(Hebert et al., 2018; Schweppe et al., 2019), and the incorporation of fully automated sample processing steps for all -omes. In addition to cancer biology, the MONTE workflow is also fully applicable to the study of other disease states, such as autoimmune and infectious diseases to address questions related to disease pathology and treatment options.

Methods

Tissue Aliquoting

Ten cryo-pulverized LUAD patient tumors were weighed and split into two approximately equal aliquots, one of which would go through the HLA immunoprecipitation.

Serial Immunoprecipitation of HLA-I & HLA-II complexes from tumors

Half of each of the ten cryo-pulverized LUAD patient tumors went through the HLA serial immunoprecipitation prior to multi-omic analysis. Each tumor was weighed and lysed with cold lysis buffer (20mM Tris, pH 8.0, 100mM NaCl, 6mM MgCl2, 1mM EDTA, 60mM Octyl β-d-glucopyranoside, 0.2mM Iodoacetamide, 1.5% Triton X-100, 50xC0mplete Protease Inhibitor Tablet-EDTA free, 1mM PMSF, 10mM NaF, 1:100 dilution of Protease Inhibitor Cocktail 2 (Sigma-Aldrich, P5726), 1:100 dilution of Protease Inhibitor Cocktail 3 (Sigma-Aldrich, P0044), 50 uM PR-619, 10mM Sodium Butyrate, 2uM SAHA, 10mM Nicotinamide) obtaining a total of 1.2 ml lysate per tumor. Each lysate was moved into an eppendorf tube, incubated on ice for 30 min with 2ul of Benzonase (Thomas Scientific, E1014-25KU) and inverted after 15 min to degrade nucleic acid. The lysates were then centrifuged at 15,000 rcf for 20 min at 4ºC and the supernatants were transferred to another set of eppendorf tubes containing ~37.5uL pre-washed beads. The beads and lysate are rotated at 4ºC for one hour in order to remove hydrophobic molecules and non-specifics that may interfere with the HLA-IP. The beads are centrifuged at 1500 rcf for 1 min at 4ºC and the lysate is transferred to a mixture of ~37.5uL pre-washed beads (Millipore Sigma, GE17-0886-01) and 15 ul of HLA-II antibody mix (9uL TAL-1B5 (abcam, ab20181), 3uL EPR11226 (abcam, ab157210), 3uL B-K27 (abcam, ab47342)).

The HLA complexes were captured on the beads by incubating on a rotor at 4ºC for 3hr. Following the incubation all tubes were centrifuged at 1,500 rcf for 1 min at 4ºC and the lysates were transferred from to new eppendorfs containing prewashed beads and 15uL of HLA-I antibody (W6/32) (Santa Cruz Biotechnology, sc-32235). The HLA-I antibody-bead-lysate mixture rotated for 3hr at 4ºC and was spun at 1500 rcf for 1 min at 4ºC. The unbound lysates were transferred to new eppendorfs and flash-frozen with liquid nitrogen for multi-omic downstream analysis. During HLA complex capture a 10uM PE fritted plate (Agilent, S7898A) was sawed in half, placed on a Waters Positive Pressure Manifold, and washed using 1mL acetonitrile and 3x 1mL room-temperature PBS. After each liquid addition PSI of less than 5 was applied to the plate to achieve liquid movement. Immediately following each HLA capture, beads were resuspended in 1mL cold PBS and transferred to one half of the pre-washed 10uM PE fritted plate. Each eppendorf was then rinsed with 500uL cold PBS and remaining beads were transferred to the correct well. The beads were washed to remove nonspecifically bound material. In total, four wash steps were performed; two washes with 2mL of cold complete wash buffer (20mM Tris, pH 8.0, 100mM NaCl, 1mM EDTA, 6mM Octyl β-d-glucopyranoside, 0.2mM Iodoacetamide), and two washes with 2mL of 10mM Tris pH 8.0 buffer. The 10uM PE fritted plate with dry HLA-II beads was wrapped with parafilm and stored at 4ºC until all HLA-I beads were washed on the other half of the plate and all samples were prepared for mass spectrometry analysis.

96-Well Plate Desalt of HLA Immunoprecipitation using a Positive Pressure Manifold

HLA peptides were eluted and desalted from beads as follows: 20 wells of the tC18 40mg Sep-Pak desalting plate (Waters, Milford, MA) were activated two times with 1mL of methanol (MeOH), 500uL of 99.9% acetonitrile (ACN)/0.1% formic acid (FA), then washed four times with 1mL of 1% FA. The two halves of the 10uM PE fritted filter plate containing the beads were put together and placed on top of the Sep-Pak plate. To dissociate peptides from HLA molecules and facilitate peptides binding to the tC18 solid-phase 200uL of 3% ACN/5% FA was added to the beads in the filter plate. Then 100fmol Biognosys internal retention time (iRT) standards (SKU: Ki-3002-2) was spiked into each sample as a loading control. One wash of 400uL of 1% FA was pushed through the 10uM filter plate. The beads were then incubated with 500uL of 10% acetic acid (AcOH) three times for 5 min to further dissociate bound peptides from the HLA molecules. The beads were rinsed once with 1mL 1% FA and the filter plate was removed. The Sep-Pak desalt plate was rinsed with 1mL 1% FA an additional three times. The peptides were eluted from the Sep-Pak desalt plate using 250uL of 15%ACN/1% FA and 2x 250uL of 50% ACN/1% FA. HLA-peptides were eluted into unused 1.5mL micro tubes (Sarstedt, Nümbrecht, Germany), frozen, and dried down via vacuum centrifugation. Dried peptides were stored at -80ºC until being subjected to microscaled basic reverse phase separation.

Briefly, peptides were loaded on Stage-tips with 2 punches of SDB-XC material (Empore 3M). HLA-I and HLA-II peptides were eluted in three fractions with increasing concentrations of ACN (HLA-I: 5%, 10% and 30% in 0.1% NH4OH, pH 10, HLA-II: 5%, 15%, and 40% in 0.1% NH4OH, pH 10) (Klaeger et al., 2021). Peptides were reconstituted in 3% ACN/5% FA prior to loading onto an analytical column (35 cm, 1.9µm C18 (Dr. Maisch HPLC GmbH), packed in-house PicoFrit 75 µm inner diameter, 10 µm emitter (New Objective)). Peptides were eluted with a linear gradient (EasyNanoLC 1200, Thermo Fisher Scientific) ranging from 6-30% Solvent B (0.1%FA in 90% ACN) over 84 min, 30-90% B over 9 min and held at 90% B for 5 min at 200 nl/min. MS/MS were acquired on a Thermo Orbitrap Exploris 480 equipped with (HLA-I) and without (HLA-II) FAIMS (Thermo Fisher Scientific) in data dependent acquisition. FAIMS CVs were set to -50 and -70 with a cycle time of 1.5s per FAIMS experiment. MS2 fill time was set to 100ms, collision energy was 30CE for HLA-I and 34CE for HLA-II.

Immunopeptidome LC-MS/MS data interpretation

Peptide sequences were interpreted from MS/MS spectra using Spectrum Mill (v 7.08, proteomics.broadinstitute.org) to search against a protein sequence database containing human Ensembl release 83 (all non-redundant protein coding transcript biotypes mapped to the human reference genome GRCh37/hg19) (ftp://ftp.ensembl.org/pub/grch37/current/fasta/homo_sapiens/pep/) with the additional appending of 602 common laboratory contaminant proteins and bovine proteins, 553 human non-canonical small open reading frames, common human viruses from Uniprot, TCGA shared mutations from 26 tumor types (https://www.cancer.gov/tcga), and nuORF DB v1.0 (Ouspenskaia et al., 2020) for a total of 345,062 entries.

MS/MS spectra were excluded from searching if they did not have a precursor MH+ in the range of 600-4000, had a precursor charge >5, or had a minimum of <5 detected peaks. Merging of similar spectra with the same precursor m/z acquired in the same chromatographic peak was disabled. Prior to searches, all MS/MS spectra had to pass the spectral quality filter with a sequence tag length >1 (i.e., minimum of 3 masses separated by the in-chain masses of 2 amino acids) based on HLA v3 peak detection. MS/MS search parameters included: ESI-QEXACTIVE-HCD-HLA-v3 scoring parameters; no-enzyme specificity; fixed modification: carbamidomethylation of cysteine; variable modifications: cysteinylation of cysteine, oxidation of methionine, deamidation of asparagine, acetylation of protein N-termini, and pyroglutamic acid at peptide N-terminal glutamine; precursor mass shift range of -18 to 81 Da; precursor mass tolerance of ±10 ppm; product mass tolerance of ± 10 ppm, and a minimum matched peak intensity of 30%. Peptide spectrum matches (PSMs) for individual spectra were automatically designated as confidently assigned using the Spectrum Mill auto-validation module to apply target-decoy based FDR estimation at the PSM level of <1.0% FDR for peptides of minimum length 6 with a backbone cleavage score minimum of 5.

Neoantigens were detected using a separate search with the above parameters except for use of a different protein sequence database. Instead to maintain a reference proteome consistent with the reference genome used for variant calls we created a base reference protein sequence database containing human Ensembl release 100 (all non-redundant protein coding transcript biotypes mapped to the human reference genome GRCh38) (ftp://ftp.ensembl.org/pub/release-100/fasta/homo_sapiens/pep/) with the additional appending of 602 common laboratory contaminant proteins and bovine proteins, 2043 human non-canonical small open reading frames,and TCGA shared mutations from 26 tumor types (https://www.cancer.gov/tcga) for a total of 79, 727 entries. To that base sequence database we appended personalized sequence entries prepared using QUILTS v3 containing somatic and germline variants and indels for each of the 10 LUAD patients ((Cibulskis et al., 2013; Ellrott et al., 2018; Ruggles et al., 2016).

HLA peptide prediction using HLAthena

HLA peptide prediction was performed using HLAthena (Sarkizova et al., 2020). Unless otherwise specified, peptides were assigned to an allele using a percentile rank cutoff ≤ 0.5.

Serial Processing of Proteome and Phosphoproteome from UbiFast flow-through samples Enrichment (CompRef)

WHIM2 and WHIM16 patient-derived xenografts (PDX) were processed for proteomic analysis as described previously(Mertins et al., 2018). Enrichment of K-ε-GG peptides was performed using UbiFast(Udeshi et al., 2020) and non-K-ε-GG peptides were collected in the flow-through for proteome and phosphoproteome analysis. For proteome analysis peptides were acidified and desalted by SPE using a 100 mg tC18 SepPak cartridge. Eluates were frozen and a vacuum centrifuge was used to dry peptides. Peptides were reconstituted in 30% ACN, peptides concentration was determined using a BCA assay and peptides were dried again. 300 ug peptides from each sample (5 replicates of WHIM2 and five replicates of WHIM16), were labeled with 300ug TMT in 20 % ACN, 50 mM HEPES for 1 hour. TMT labeling was quenched

by adding 4 uL 5 % hydroxylamine for 15 minutes at room temperature while shaking. Samples were combined into a 15 mL conical tube, frozen at -80°C and dried in a vacuum centrifuge. The combined sample was desalted using a 200 mg tC18 SepPak cartridge and the eluate was snap frozen then dried in a vacuum centrifuge. Offline bRP fractionation was performed as previously described (Mertins et al., 2018). Briefly, peptides were separated over a 96 minute gradient with a flow rate of 1 ml/min. Solvent A was 5 mM ammonium formate, 2% ACN and solvent B was 5 mM ammonium formate, 90% ACN. 96 fractions were concatenated into 24 fractions for proteome analysis. 5% of each of the 24 fractions were transferred into HPLC vials, frozen and dried in a vacuum centrifuge for analysis. The remaining 95 % of each fraction were concatenated into 13 fractions for phosphopeptide enrichment. For IMAC enrichment of phosphopeptides Ni-NTA Superflow Agarose Beads were stripped and incubated with 10 mM FeCl3 (451649, Sigma) for 30 minutes. Peptides were reconstituted in 80% ACN, 0.1% TFA and added to 10 uL of prepared beads for 30 minutes at room temperature. Beads were spun at 1000 x g for 1 minute and the supernatant was removed. Beads were added to stage-tips and washed with 50 uL 80% ACN, 0.1% TFA twice and 50 uL of 1% TFA once. Peptides were eluted from beads onto the C18 stage-tip discs with three washes of 75 uL 500 mM potassium phosphate, pH 7. Peptides were further eluted from the C18 discs with 60 uL 50% ACN, 0.1% FA, snap frozen and dried with a vacuum centrifuge. For LC/MS analysis peptides were reconstituted in 9 uL of 3% ACN, 0.1% FA and 4 uL were injected into the MS.

Serial Processing of Ubiquitylome, Proteome, Phosphoproteome, and Acetylome from cryopulverized tissue vs HLA IP flow-through (LUAD)

Lysis and tryptic digestion of LUAD cryopulverized tissue and HLA IP flow-through

Each set of 10 replicate tumors underwent denaturing lysis in SDS to prepare for S-Trap digestion. Flow-through of the HLA-I IP, at this point in native HLA lysis buffer and stored as flash-frozen unbound lysates, were briefly thawed on ice for ~15 min. Once thawed, 10% SDS was added for a final concentration of 2.5% SDS to denature the lysate, resulting in a final volume of ~1.5 mL lysate which was prepared for S-Trap digestion.

Replicates of the HLA-depleted samples were lysed from cryopulverized tissue in 1 mL 5% SDS buffer (5% SDS, 50 mM TEAB pH 8.5, 2 mM $MgCl_2$). The samples were disrupted by pipette mixing and gentle vortexing and incubated at room temperature for ~10 min. Samples were treated with 2 uL benzonase to shear DNA, mixed again, and incubated at room temperature for another ~20 min. Finally, non-HLA-depleted lysates were homogenized with a probe sonicator for 30 sec and left to lyse again for ~10 min. The lysates were cleared by centrifugation for 15 min at 15000 x g and the supernatant was prepared for S-Trap digestion.

In both sets of LUAD tumors, all further processing steps were executed identically. Protein concentration was estimated using a BCA assay for scaling of digestion enzymes. Disulfide bonds were reduced in 5 mM DTT for 30 min at 25°C and 1000 rpm shaking, and cysteine residues alkylated in 10 mM IAA in the dark for 45 min at 25°C and 1000 rpm shaking. Lysates were then transferred to a 15 mL conical tube to prepare for protein precipitation. 27% phosphoric acid was added at a 1:10 ratio of lysate volume to acidify, and proteins were

precipitated with 6X sample volume of ice cold S-Trap buffer (90% methanol, 100 mM TEAB). The precipitate was transferred in successive loads of 3 mL to a Protifi S-Trap Midi and loaded with 1 min centrifugation at 4000 x g, mixing the remaining precipitate thoroughly between transfers. The precipitated proteins were washed 4x with 3 mL S-Trap buffer at 4000 x g for 1 min. To digest the deposited protein material, 350 uL digestion buffer (50 mM TEAB) containing both trypsin and endopeptidase C (LysC), each at 1:50 enzyme:substrate, was passed through each S-Trap column with 1 min centrifugation at 4000 x g. The digestion buffer was then added back atop the S-Trap and the cartridges were left capped overnight at 25°C.

Peptide digests were eluted from the S-Trap, first with 500 uL 50 mM TEAB and next with 500 uL 0.1% FA, each for 30 sec at 1000 x g. The final elution of 500 uL 50% ACN / 0.1% FA was centrifuged for 1 min at 4000 x g to clear the cartridge. Peptide concentration of the pooled elutions was estimated with a BCA assay, divided into 750 ug peptide aliquots for K-ε-GG enrichment, snap frozen and dried in a vacuum centrifuge.

Semi-automated UbiFast K-ε-GG enrichment

Peptides containing the K-ε-GG tryptic remnant of ubiquitin/ubiquitin-like small protein modifications were enriched using an adaptation of the UbiFast protocol for the Thermo KingFisher automation platform. Briefly, 750 ug peptide aliquots were reconstituted in 250 uL CST HS bind buffer w/ 0.01% CHAPS. All following steps for UbiFast enrichment excluding labeling and final bead collection contained 0.01% CHAPS. Reconstituted peptides were added to 5 uL PBS-washed CST HS anti-K-ε-GG antibody beads and incubated at 4°C for 1 hour in a foil sealed KingFisher plate with rotation. Following removal of the beads from the incubation by the KingFisher robot, the incubation plate containing non-TMT labeled, K-ε-GG-depleted peptide flow-through was sealed and frozen for downstream proteome, phosphoproteome, and acetylproteome processing. Briefly, bead-bound enriched peptides were washed with 50% ACN / 50% CST HS wash buffer and washed again with PBS. K-ε-GG peptides were labeled on-bead with 400 ug TMT11 in 100 mM HEPES (prepared immediately before run) for 20 minutes and labeling was quenched with 2% hydroxylamine. Finally, the beads were washed with CST HS wash buffer before being deposited into 100 uL PBS containing no CHAPS buffer. Each well containing each TMT channel was combined by 11-plex, the supernatant was removed, and enriched peptides were eluted from the beads with 2 x 10 min 0.15% TFA. The eluate was desalted with a C18 stagetip, frozen, and dried in a vacuum centrifuge. For LC-MS/MS analysis, the unfractionated K-ε-GG peptides were reconstituted in 9 uL 3% ACN / 0.1% FA and 4 uL was injected twice back-to-back for each sample.

Semi-automated UbiFast K-ε-GG enrichment

Non-TMT labeled, K-ε-GG-depleted peptide flow-through of the K-ε-GG IPs were acidified with neat formic acid to a final concentration of 1% and desalted with 100 mg tC18 SepPak cartridges. Eluates were frozen and dried in a vacuum centrifuge. Peptides were reconstituted in30% ACN / 0.1% FA, peptide concentration was estimated using a BCA assay and peptides were aliquoted for downstream processing and dried again.

TMT labeling of UbiFast flow-through and bRP fractionation

Desalted, non-labeled, non-K-ε-GG peptide IP flow-through was processed for further serial enrichment of proteome, phosphoproteome, and acetylproteome as described(Mertins et al., 2018). 300 ug of each sample was reconstituted in 60 uL 50 mM HEPES and labeled with 300 ug TMT10 reagent at a final concentration of 20% ACN for 1 hour at 25°C and 1000 rpm. Each tumor replicate was assigned the same TMT channel in its corresponding TMT 10-plex for an identical experimental design. Labeling reactions were diluted to 2.5 mg/mL with 50 mM HEPES. Complete labeling and balancing of input material were confirmed. TMT labeling was quenched with 3 uL 5% hydroxylamine for 15 minutes and each TMT 10-plex was combined, frozen, and dried. Dried, labeled, combined peptides were reconstituted with 3 mL 1% FA and desalted with a 200 mg tC18 SepPak, and the eluate was snap frozen and dried in a vacuum centrifuge.

Offline bRP fractionation was performed as described previously and above (PMID: 29988108). Briefly, peptides were separated over a 96 minute gradient with a flow rate of 1 ml/min. Solvent A was 5 mM ammonium formate, 2% ACN and solvent B was 5 mM ammonium formate, 90% ACN. 96 fractions were concatenated into 24 fractions for proteome analysis. 5% of each of the 24 fractions were transferred into HPLC vials, frozen and dried in a vacuum centrifuge for analysis. The remaining 95% of each fraction were concatenated into 13 fractions for phosphopeptide enrichment. Proteome fractions were reconstituted in 3% ACN / 0.1% FA and 500 ng at 0.25 ug/uL from each of the 24 fractions were injected for LC-MS/MS analysis.

Automated IMAC phosphopeptide enrichment

IMAC enrichment of phosphopeptides was performed using AssayMap Fe(III)-NTA cartridges (Agilent, G5496-60085). Concatenated fractions were solubilized with 80 uL 50% ACN / 0.1% TFA in a bath sonicator for 5 min, followed by addition of 120 uL 100% ACN / 0.1% TFA for a final concentration of 80% ACN. Insoluble peptides were pelleted by centrifugation at 6000 x g for 5 min and 160 uL was transferred to 96 well plate for enrichment. The remaining 40 uL were set aside for re-enrichment. The Agilent “Phosphopeptide Enrichment” protocol was used. Briefly, the syringes were rinsed with HPLC water and primed with 50% ACN / 0.1% TFA. Cartridges were equilibrated with 80% ACN / 0.1% TFA.160 uL of sample was loaded at 5 uL/min, and the phosphopeptide-depleted flow-through was collected and frozen for downstream acetyllysine enrichment. The cartridges were washed 3x with 80% ACN / 0.1% TFA to remove non-specific peptides. Enriched phosphopeptides were eluted from the cartridges with 20 uL fresh 1% ammonium hydroxide at 5 uL/min into a plate containing 2.5 uL neat FA. Phosphopeptide-enriched eluates were transferred to HPLC vials, frozen and dried in a vacuum centrifuge. For LC/MS-MS analysis, peptides were reconstituted in 9 uL 3% ACN / 0.1% FA and 4 uL were injected from each of the 12 fractions.

Acetyl-lysine immunoaffinity enrichment

Acetyl peptide enrichment was performed in concordance with(Gillette et al., 2020) with minor variations described below. Acetylated lysine peptides were enriched using protein A agarose beads conjugated to an acetyl-lysine motif antibody (CST PTM-SCAN Catalogue No. 13416). Phosphopeptide-depleted IMAC flow-through were concatenated from 12 to 4 fractions (~750 µg peptides per fraction) and dried down using a SpeedVac apparatus. Prior to enrichment, antibody beads were washed 4x with IAP buffer (5 mM MOPS pH 7.2, 1 mM sodium phosphate [dibasic], 5 mM NaCl). Peptides were reconstituted with 1.4ml of IAP buffer per fraction, added to washed beads, and incubated for 2 h at 4°C. Bead-bound acetyl-enriched peptides were washed 4 times with ice-cold PBS followed by two elutions with 100ul of 0.15% TFA. Eluents were desalted using C18 stage tips, eluted with 50% ACN / 0.1% FA, and dried down. Acetylpeptides were reconstituted in 7ul of 3% ACN / 0.1% FA and 4ul were injected from each of the 4 fractions for LC-MS/MS analysis.

Acknowledgements

This work was supported by the National Cancer Institute (NCI) grants U24CA210986, U01CA214125, and U24CA210979 to S.A.C., Swiss National Science Foundation (SNF) Sinergia grant CRSII5_186405 to S.A.C., Dr. Miriam and Sheldon G. Adelson Medical Research Foundation to S.A.C. and N.D.U. and by a SPARC Award to N.D.U. from the Broad Institute of MIT & Harvard (#800373). This work was made possible by a grant from BroadIgnite at the Broad Institute of MIT and Harvard to S.K. S.K was supported by the Cancer Research Institute as a Hearst Foundation fellow.

Competing Interest Statement

S.A.C. is a member of the scientific advisory boards of Kymera, PTM BioLabs and Seer and an ad hoc scientific advisor to Pfizer and Biogen.

Figures & Legends

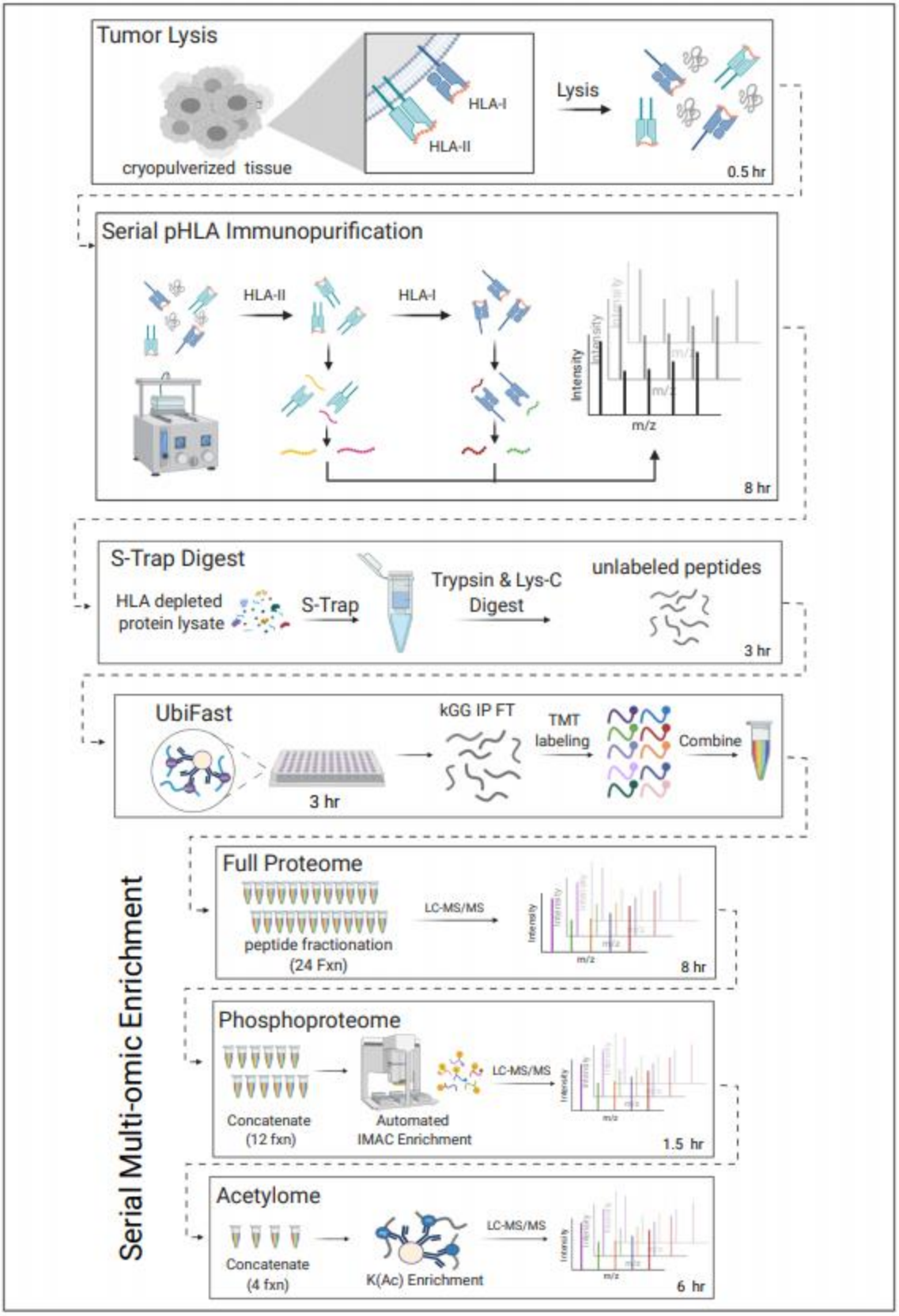


Figure 1: Schematic Overview of MONTE. The **M**ulti-**O**mic **N**etworked **T**issue **E**nrichments (MONTE) workflow for serial HLA immunopeptidome, ubiquitylome, proteome, phosphoproteome, and lysine acetylome with annotations indicating estimated timing for the different enrichment steps.

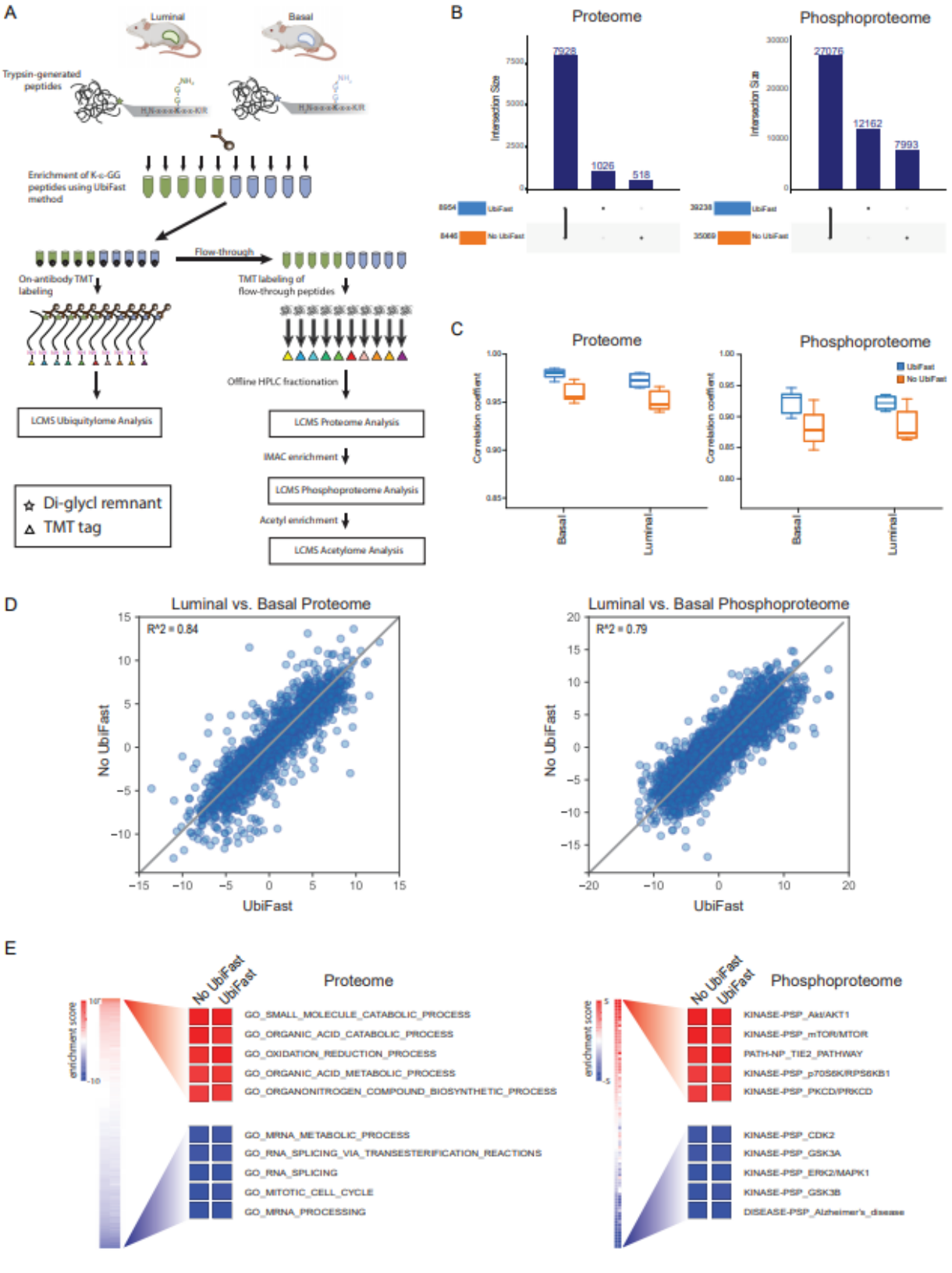
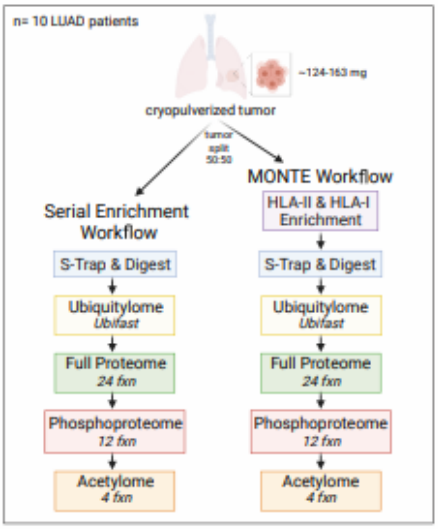


Figure 2: Comparison of the proteome and phosphoproteome of PDX samples with and without UbiFast pre-enrichment A) Schematic diagram showing experimental design used to analyze K-ε-GG peptides, proteins and phosphopeptides from the same PDX samples. The ubiquitylome was previously acquired by Udeshi *et al.(Udeshi et al., 2020)* using 500 ug of peptide input per TMT channel. The flow-through of the UbiFast enrichment was used for proteome and phosphoproteome using 300 ug peptide input per TMT state. B) The total number of quantified proteins and phosphorylation sites that could be assigned to human proteins are illustrated as 'UpSet' plots(Lex et al., 2014). Horizontal bars indicate the total number of proteins or phosphorylation sites detected by each experiment; vertical bars depict the number of uniquely or jointly detected features, as indicated by the layout matrix below. C) Pearson correlation between Luminal and Basal PDX replicates (n = 5 process replicates). Boxplot depicts upper and lower quartiles, with the median shown as a solid line. Whiskers show 1.5 interquartile range. D) Scatter plot showing log2 fold-change between Basal and Luminal PDXs, for proteins and phosphorylation sites. E) Heatmaps of all gene sets present in both experiments with the top and bottom 5 most enriched gene sets highlighted for both the proteome and phosphoproteome.

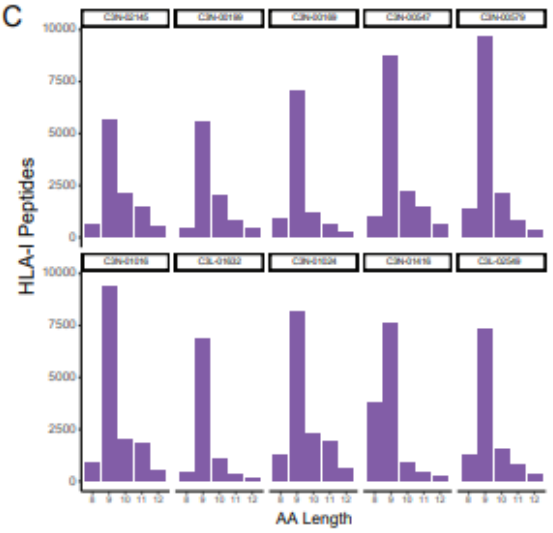
A



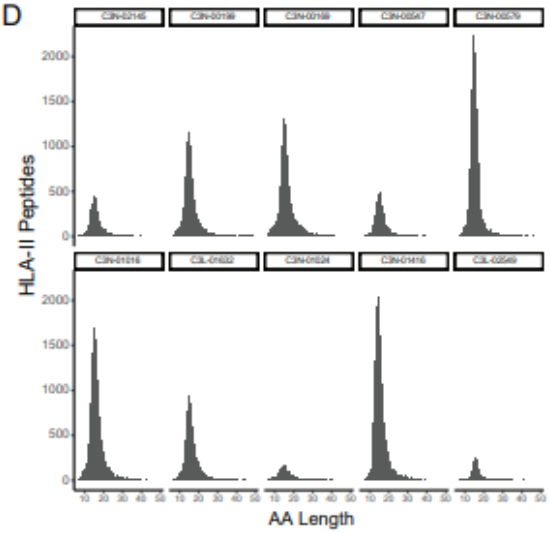
B

Participant ID	Tumor Purity	Sample Input (mg)	HLA-I peptides	HLA-II peptides
C3N-02145	62%	74.4	10315	2600
C3N-00199	75%	50	9203	6640
C3N-00169	68%	69.2	9892	7807
C3N-00547	75%	79.2	13833	2811
C3N-00579	62%	85.8	14196	10858
C3N-01016	70%	84.9	14645	9809
C3L-01632	72%	67.4	8750	5337
C3N-01024	75%	90.7	14226	1264
C3N-01416	48%	83.2	13014	11283
C3L-02549	70%	66.6	11196	1273

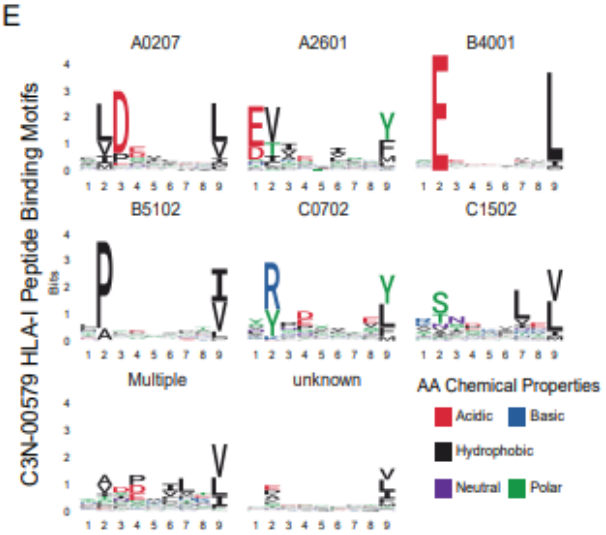
C



D



E



F

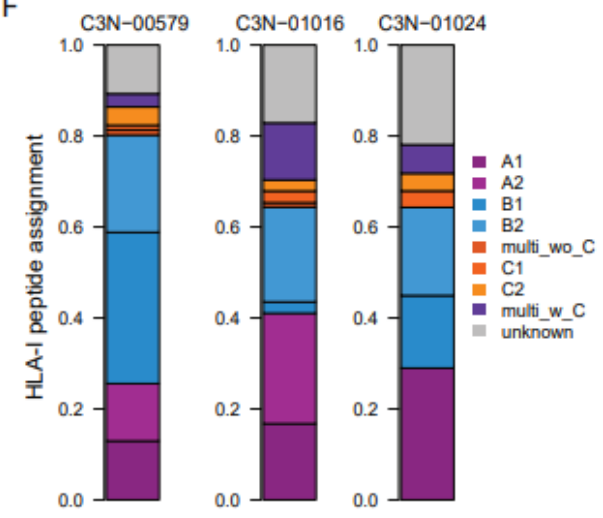
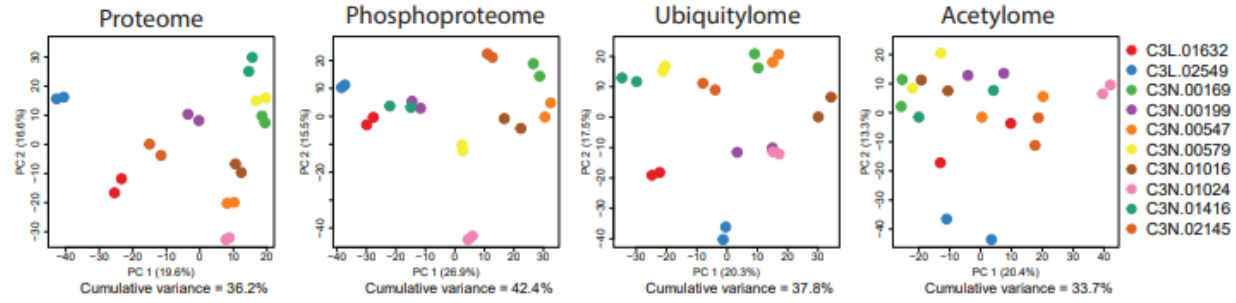


Figure 3: LUAD tumor profiling using serial multi-omic enrichment vs. MONTE workflows. A) Schematic overview of the head-to-head serial multi-omic enrichment vs. MONTE workflows used to characterize ten cryopulverized primary LUAD tumors. B) Table summarizing the input and resulting HLA-I and HLA-II peptides detected from the LUAD tumor analysis that leveraged the MONTE workflow. C) Length distributions of HLA-I peptides from the LUAD cohort. D) Length distributions of HLA-II peptides from the LUAD cohort. E) Example HLA-I peptide motifs and the allele expressed by LUAD Patient C3N-00579. F) Example HLA-I peptide:allele assignments obtained using the presentation predictions from HLAthena for three LUAD patients(Abelin et al., 2017; Sarkizova et al., 2020).

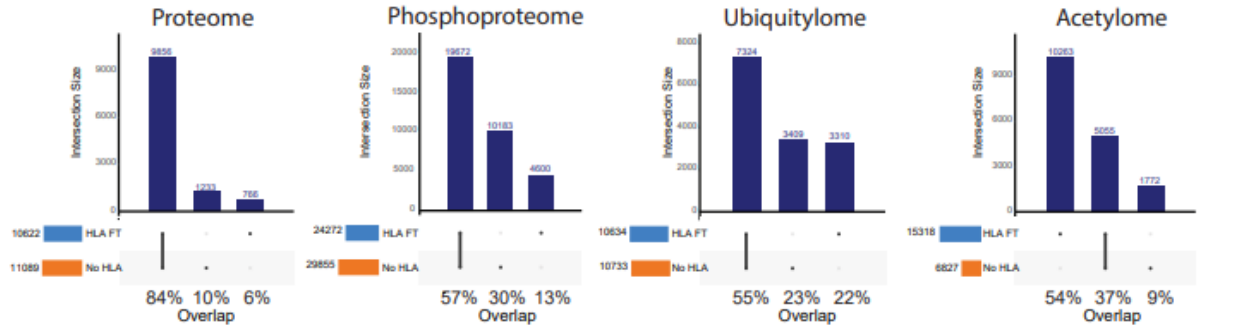
A

"Ome" in serial enrichment order		# unique proteins/sites, Serial Enrichment (no HLA IP)	# unique proteins/sites, MONTE (HLA IP FT)	% change in IDs (no IP vs HLA IP FT)
Ubiquitylome (K-GG)	Total human KGG sites	14895		--
	Fully quantified human KGG sites	10733	10634	-0.9%
	Fully quant. and localized KGG sites	10694	10594	-0.9%
Global proteome	Total human proteins	14972		--
	Fully quantified human proteins	13311	12983	-2.5%
	Fully quant. human proteins w/ > 1 unique peptide	11089	10621	-4.2%
Phosphoproteome (phospho-STY)	Total human phospho-sites	35790		--
	Fully quantified human phospho-sites	29703	24084	-18.9%
	Fully quant. and localized phospho-sites	20652	16847	-18.4%
Acetylome (acetyllysine, AcK)	Total human AcK sites	19301		--
	Fully quantified human AcK sites	6824	15316	+124.4%
	Fully quant. and localized AcK sites	6748	15173	+124.9%

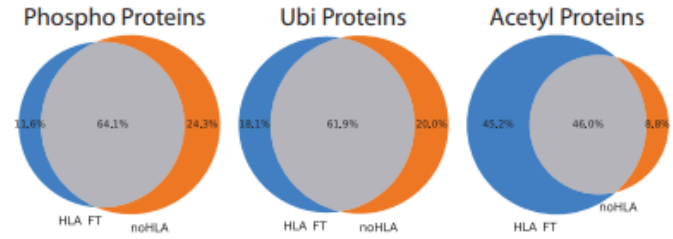
B



C



D



E

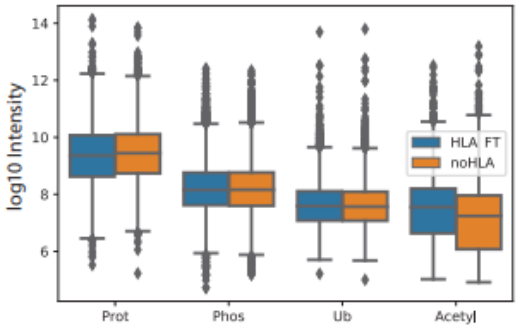


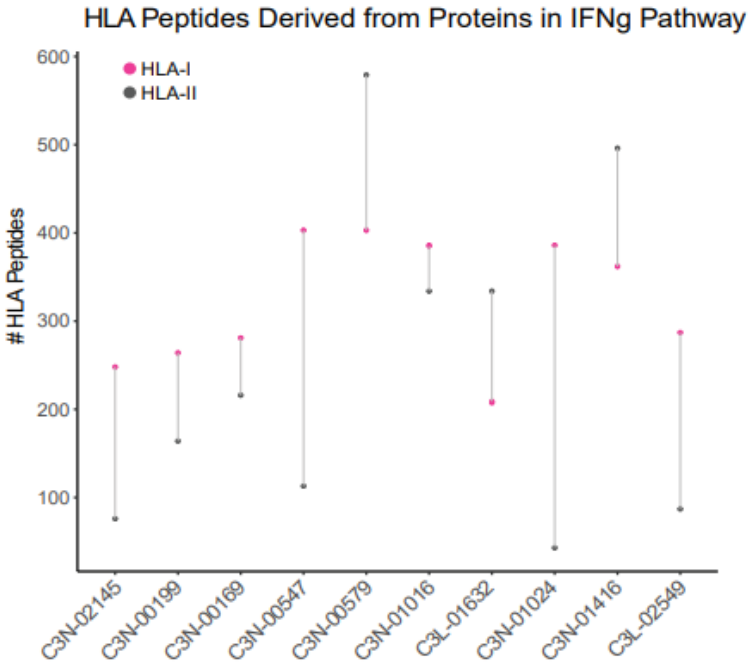
Figure 4: Evaluation of data depth between serial multi-omic enrichment with and without HLA enrichment

A) Summary table reporting the numbers of proteins and PTM containing peptides detected from non HLA enriched (No HLA) and HLA enriched (HLA FT) LUAD samples. **B**) Principal component analysis of LUAD tumors (n=10) for ~2000 regulated proteins and PTM sites. **C**) The total number of quantified proteins and PTM sites illustrated as 'UpSet' plots(Lex et al., 2014). Horizontal bars indicate the total number of proteins or phosphorylation sites detected by each experiment; vertical bars depict the number of uniquely or jointly detected features, as indicated by the layout matrix below. **D**) Venn diagrams showing the overlap of proteins and modified proteins for No HLA and HLA FT samples. **E**) Log10 total intensity distributions of proteins and PTM sites from No HLA and HLA FT samples. Boxplot depicts upper and lower quartiles, with the median shown as a solid line. Whiskers show 1.5 interquartile range.

A

CT Antigen Source Protein	# unique HLA peptides
MAGEA3	2
MAGEA4	1
PRAMEF12	1
TEX101	1
BRDT	10
CABYR	2
DDX53	1
MAGEA6	2
PRAMEF10	1
XAGE1D	3

B



C

Patient #	Length	HLA-I peptide	Sequence Coverage Map	ppm delta	TCGA mutation & UCSC annotation	Predicted HLA-I Allele	HLAthena % Rank
C3N-00199	9	qETPVVLQL	q E\T\P\V\V\L/Q/L	-1.8	ATP6AP2_p.E119Q_uc004det.2	B4403	0.053089
C3N-02145	9	APHLVQVDA	A P H\L-V\Q\V/D-A	-0.2	IDUA_p.H33Q_uc003gby.2	B5502	0.1359
C3N-02145	10	APHLVQVDAA	A P H\L\V\Q\V D\A-A	2.6	IDUA_p.H33Q_uc003gby.2	B5502	0.0414
C3L-02549	9	AEAPHLVQV	A E\A\P/H\L\V\Q/V	1.9	IDUA_p.H33Q_uc003gby.2	B3701	0.05813
C3N-00547	9	APHLVQVDA	A P H\L-V\Q\V\D-A	0	IDUA_p.H33Q_uc003gby.2	B5502	0.0147
C3N-00547	10	APHLVQVDAA	A P H\L\V\Q\V\D\A-A	0.1	IDUA_p.H33Q_uc003gby.2	B5502	0.0414
C3L-01632	10	IFSFNLVnIK	I/F S F N\L-V\n\I/K	0.5	SEMA3A_p.D81N_uc003uhz.2	A2402	14.057
C3N-01416	10	IFSFNLVnIK	I/F S F N\L-V\n\I/K	1.2	SEMA3A_p.D81N_uc003uhz.2	A2410	7.400515
C3N-01024	10	QPPPTSRTSF	Q P/P/P/T/S R T S/F	2.5	VARS_p.P51S_uc003nxe.2	B0705	0.2305
C3N-02145	9	HTmGFWLTK	H/T\m G F\W\L/T/K	1.2	VHL_p.L116fs_uc003bvc.2	A1101, A1102	0.1789, 0.0835
C3N-00547	9	HTMGFWLTK	H/T\M/G F\W\L/T/K	-1	VHL_p.L116fs_uc003bvc.2	A1101, A1102	0.1789, 0.0835

D

Patient	HLA-A1	HLA-A2	HLA-B1	HLA-B2	HLA-C1	HLA-C2	# HLA-I neoantigens	Somatic SAAVs	Germline SAAVs
C3L-01632	A*02:05	A*24:02	B*44:02	B*50:01	C*05:01	C*06:02	0	365	25571
C3L-02549	A*01:01	A*24:02	B*07:02	B*37:01	C*07:02	C*06:02	0	61	24237
C3N-00169	A*24:03	A*11:01	B*15:02	B*15:12	C*03:03	C*08:01	0	386	25329
C3N-00199	A*01:01	A*01:01	B*44:03	B*27:05	C*02:02	C*16:01	1	492	24527
C3N-00547	A*11:02	A*11:01	B*55:02	B*15:02	C*12:03	C*08:01	3	1405	24537
C3N-00579	A*02:07	A*26:01	B*51:02	B*40:01	C*07:02	C*15:02	0	238	24916
C3N-01016	A*02:07	A*33:03	B*58:01	B*46:01	C*03:02	C*01:02	2	963	24663
C3N-01024	A*33:03	A*33:03	B*58:01	B*07:05	C*03:02	C*15:05	0	175	25041
C3N-01416	A*24:10	A*24:02	B*46:01	B*18:02	C*01:02	C*07:04	0	186	25101
C3N-02145	A*11:01	A*11:02	B*54:01	B*55:02	C*01:02	C*12:03	0	68	25282

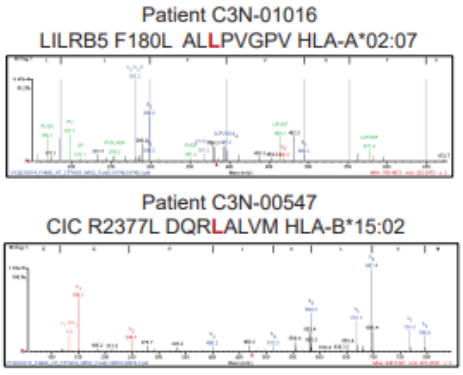


Figure 5: Analysis of HLA peptides presented by LUAD tumors derived from CT antigen and mutation containing source proteins

A) A summary table reporting the number of unique HLA peptide:patient combinations observed across the small LUAD cohort that are derived from CT antigen source proteins. **B**) A dot plot showing the number of unique HLA-I and HLA-II peptides derived from proteins annotated as being part of the IFNγ signaling pathway presented by LUAD patient tumors. **C**) Summary table showing antigens detected in the LUAD cohort that are peptides containing shared somatic mutations observed across cancer tissue cohorts in TCGA. **D**) Summary table of HLA neoantigens detected using personalized databases containing patient-specific somatic mutations accompanied by two example MS/MS spectra of neoantigens identified from two different LUAD patients with high mutation burdens.

Table S1: **HLA-I and HLA-II immunopeptidomes from the PDX models of breast cancer.** Tables reporting the unique HLA-I and HLA-II peptide and their associated source proteins identified from serial HLA immunopurification from Basal (n=3, P93_241, P93_242, P93_243) and Luminal (n=3; P94_163, P94_164, P94_165) breast cancer PDX tumors.

Table S2: Comparison of experimental conditions and effects of lysis buffer conditions and sample preparation steps on label-free A375 acetylome coverage. A summary table describing the different conditions used to test how native lysis buffer conditions, HDAC inhibitors, and empty Protein G sepharose bead pre-clear impact acetylpeptide immunopurification. The table describes how different sample preparation and lysis conditions were varied. These variables included different lysis buffer bases (native vs. SDS), the inclusion of HLA enrichment, timing, temperature, lysis buffer additives, the inclusion of a bead-based pre-clear step prior to digestion, and the inclusion of HDAC inhibitors in the lysis buffer. The resulting acetylpeptide (AcK) enrichment specificities (% spectra containing AcK) and total acetyl peptides (distinct peptides containing AcK) are shown for each sample and associated condition. The average number of distinct AcK peptides are also reported in terms of the major variables being tested.

Table S3: LUAD tumor characteristics and HLA immunopeptidome results. A description of the ten LUAD patient tumors characterized using the MONTE workflow along with HLA immunopeptidome results.

References

Abelin, J.G., Patel, J., Lu, X., Feeney, C.M., Fagbami, L., Creech, A.L., Hu, R., Lam, D., Davison, D., Pino, L., et al. (2016). Reduced-representation Phosphosignatures Measured by Quantitative Targeted MS Capture Cellular States and Enable Large-scale Comparison of Drug-induced Phenotypes. Mol. Cell. Proteomics *15*, 1622–1641.

Abelin, J.G., Keskin, D.B., Sarkizova, S., Hartigan, C.R., Zhang, W., Sidney, J., Stevens, J., Lane, W., Zhang, G.L., Eisenhaure, T.M., et al. (2017). Mass Spectrometry Profiling of HLA-Associated Peptidomes in Mono-allelic Cells Enables More Accurate Epitope Prediction. Immunity *46*, 315–326.

Abelin, J.G., Harjanto, D., Malloy, M., Suri, P., Colson, T., Goulding, S.P., Creech, A.L., Serrano, L.R., Nasir, G., Nasrullah, Y., et al. (2019). Defining HLA-II Ligand Processing and Binding Rules with Mass Spectrometry Enhances Cancer Epitope Prediction. Immunity *51*, 766–779.e17.

Admon, A., and Bassani-Sternberg, M. (2011). The Human Immunopeptidome Project, a suggestion for yet another postgenome next big thing. Mol. Cell. Proteomics *10*, O111.011833.

Almeida, L.G., Sakabe, N.J., deOliveira, A.R., Silva, M.C.C., Mundstein, A.S., Cohen, T., Chen, Y.-T., Chua, R., Gurung, S., Gnjatic, S., et al. (2009). CTdatabase: a knowledge-base of high-throughput and curated data on cancer-testis antigens. Nucleic Acids Res. *37*, D816–D819.

Bassani-Sternberg, M., and Gfeller, D. (2016). Unsupervised HLA Peptidome Deconvolution Improves Ligand Prediction Accuracy and Predicts Cooperative Effects in Peptide–HLA Interactions. J. Immunol. *197*, 2492.

Bassani-Sternberg, M., Bräunlein, E., Klar, R., Engleitner, T., Sinitcyn, P., Audehm, S., Straub, M., Weber, J., Slotta-Huspenina, J., Specht, K., et al. (2016). Direct identification of clinically relevant neoepitopes presented on native human melanoma tissue by mass spectrometry. Nat. Commun. *7*, 13404–13404.

Blumenberg, L., Kawaler, E., Cornwell, M., Smith, S., Ruggles, K., and Fenyö, D. (2019). BlackSheep: A Bioconductor and Bioconda package for differential extreme value analysis.

Chong, C., Marino, F., Pak, H., Racle, J., Daniel, R.T., Müller, M., Gfeller, D., Coukos, G., and Bassani-Sternberg, M. (2018). High-throughput and sensitive immunopeptidomics platform reveals profound interferonγ-mediated remodeling of the Human Leukocyte Antigen (HLA) Ligandome. Mol. Cell. Proteomics *17*, 533–548.

Chong, C., Müller, M., Pak, H., Harnett, D., Huber, F., Grun, D., Leleu, M., Auger, A., Arnaud, M., Stevenson, B.J., et al. (2020). Integrated proteogenomic deep sequencing and analytics accurately identify non-canonical peptides in tumor immunopeptidomes. Nat. Commun. *11*, 1293.

Choudhary, C., Kumar, C., Gnad, F., Nielsen, M.L., Rehman, M., Walther, T.C., Olsen, J.V., and Mann, M. (2009). Lysine acetylation targets protein complexes and co-regulates major cellular functions. Science *325*, 834–840.

Cibulskis, K., Lawrence, M.S., Carter, S.L., Sivachenko, A., Jaffe, D., Sougnez, C., Gabriel, S., Meyerson, M., Lander, E.S., and Getz, G. (2013). Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat. Biotechnol. *31*, 213–219.

Djureinovic, D., Hallström, B.M., Horie, M., Mattsson, J.S.M., La Fleur, L., Fagerberg, L., Brunnström, H., Lindskog, C., Madjar, K., Rahnenführer, J., et al. (2016). Profiling cancer testis antigens in non-small-cell lung cancer. JCI Insight *1*, e86837.

Dou, Y., Kawaler, E.A., Cui Zhou, D., Gritsenko, M.A., Huang, C., Blumenberg, L., Karpova, A., Petyuk, V.A., Savage, S.R., Satpathy, S., et al. (2020). Proteogenomic Characterization of Endometrial Carcinoma. Cell *180*, 729–748.e26.

Ellrott, K., Bailey, M.H., Saksena, G., Covington, K.R., Kandoth, C., Stewart, C., Hess, J., Ma, S., Chiotti, K.E., McLellan, M., et al. (2018). Scalable Open Science Approach for Mutation

Calling of Tumor Exomes Using Multiple Genomic Pipelines. Cell Syst *6*, 271–281.e7.

Gillette, M.A., Satpathy, S., Cao, S., Dhanasekaran, S.M., Vasaikar, S.V., Krug, K., Petralia, F., Li, Y., Liang, W.-W., Reva, B., et al. (2020). Proteogenomic Characterization Reveals Therapeutic Vulnerabilities in Lung Adenocarcinoma. Cell *182*, 200–225.e35.

HaileMariam, M., Eguez, R.V., Singh, H., Bekele, S., Ameni, G., Pieper, R., and Yu, Y. (2018). S-Trap, an Ultrafast Sample-Preparation Approach for Shotgun Proteomics. J. Proteome Res. *17*, 2917–2924.

Hartlmayr, D., Ctortecka, C., Seth, A., Mendjan, S., Tourniaire, G., and Mechtler, K. (2021). An automated workflow for label-free and multiplexed single cell proteomics sample preparation at unprecedented sensitivity.

He, Y., Rozeboom, L., Rivard, C.J., Ellison, K., Dziadziuszko, R., Yu, H., Zhou, C., and Hirsch, F.R. (2017). MHC class II expression in lung cancer. Lung Cancer *112*, 75–80.

Hebert, A.S., Prasad, S., Belford, M.W., Bailey, D.J., McAlister, G.C., Abbatiello, S.E., Huguet, R., Wouters, E.R., Dunyach, J.-J., Brademan, D.R., et al. (2018). Comprehensive Single-Shot Proteomics with FAIMS on a Hybrid Orbitrap Mass Spectrometer. Anal. Chem. *90*, 9529–9537.

James, A.M., Smith, C.L., Smith, A.C., Robinson, A.J., Hoogewijs, K., and Murphy, M.P. (2018). The Causes and Consequences of Nonenzymatic Protein Acylation. Trends Biochem. Sci. *43*, 921–932.

Jurtz, V., Paul, S., Andreatta, M., Marcatili, P., Peters, B., and Nielsen, M. (2017). NetMHCpan-4.0: Improved Peptide-MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. J. Immunol. *199*, 3360–3368.

Keskin, D.B., Reinhold, B.B., Zhang, G.L., Ivanov, A.R., Karger, B.L., and Reinherz, E.L. (2015). Physical detection of influenza A epitopes identifies a stealth subset on human lung epithelium evading natural CD8 immunity. Proc. Natl. Acad. Sci. U. S. A. *112*, 2151–2156.

Kim, W., Bennett, E.J., Huttlin, E.L., Guo, A., Li, J., Possemato, A., Sowa, M.E., Rad, R., Rush, J., Comb, M.J., et al. (2011). Systematic and quantitative assessment of the ubiquitin-modified proteome. Mol. Cell *44*, 325–340.

Klaeger, S., Apffel, A., Clauser, K.R., Sarkizova, S., Oliveira, G., Rachimi, S., Le, P.M., Tarren, A., Chea, V., Abelin, J.G., et al. (2021). Optimized liquid and gas phase fractionation increase HLA-peptidome coverage for primary cell and tissue samples.

Krönke, J., Udeshi, N.D., Narla, A., Grauman, P., Hurst, S.N., McConkey, M., Svinkina, T., Heckl, D., Comer, E., Li, X., et al. (2014). Lenalidomide causes selective degradation of IKZF1 and IKZF3 in multiple myeloma cells. Science *343*, 301–305.

Krug, K., Jaehnig, E.J., Satpathy, S., Blumenberg, L., Karpova, A., Anurag, M., Miles, G., Mertins, P., Geffen, Y., Tang, L.C., et al. (2020). Proteogenomic Landscape of Breast Cancer Tumorigenesis and Targeted Therapy. Cell.

Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R., and Pfister, H. (2014). UpSet: Visualization of Intersecting Sets. IEEE Trans. Vis. Comput. Graph. *20*, 1983–1992.

Li, S., Shen, D., Shao, J., Crowder, R., Liu, W., Prat, A., He, X., Liu, S., Hoog, J., Lu, C., et al.

(2013). Endocrine-therapy-resistant ESR1 variants revealed by genomic characterization of breast-cancer-derived xenografts. Cell Rep. *4*, 1116–1130.

Liu, J., Lichtenberg, T., Hoadley, K.A., Poisson, L.M., Lazar, A.J., Cherniack, A.D., Kovatich, A.J., Benz, C.C., Levine, D.A., Lee, A.V., et al. (2018). An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. Cell *173*, 400–416.e11.

Marcu, A., Bichmann, L., Kuchenbecker, L., Kowalewski, D.J., Freudenmann, L.K., Backert, L., Mühlenbruch, L., Szolek, A., Lübke, M., Wagner, P., et al. (2021). HLA Ligand Atlas: a benign reference of HLA-presented peptides to improve T-cell-based cancer immunotherapy. J Immunother Cancer *9*.

Mertins, P., Qiao, J.W., Patel, J., Udeshi, N.D., Clauser, K.R., Mani, D.R., Burgess, M.W., Gillette, M.A., Jaffe, J.D., and Carr, S.A. (2013). Integrated proteomic analysis of post-translational modifications by serial enrichment. Nat. Methods *10*, 634–637.

Mertins, P., Mani, D.R., Ruggles, K.V., Gillette, M.A., Clauser, K.R., Wang, P., Wang, X., Qiao, J.W., Cao, S., Petralia, F., et al. (2016). Proteogenomics connects somatic mutations to signalling in breast cancer. Nature *534*, 55–62.

Mertins, P., Tang, L.C., Krug, K., Clark, D.J., Gritsenko, M.A., Chen, L., Clauser, K.R., Clauss, T.R., Shah, P., Gillette, M.A., et al. (2018). Reproducible workflow for multiplexed deep-scale proteome and phosphoproteome analysis of tumor tissues by liquid chromatography-mass spectrometry. Nat. Protoc. *13*, 1632–1661.

Myers, S.A., Rhoads, A., Cocco, A.R., Peckner, R., Haber, A.L., Schweitzer, L.D., Krug, K., Mani, D.R., Clauser, K.R., Rozenblatt-Rosen, O., et al. (2019). Streamlined Protocol for Deep Proteomic Profiling of FAC-sorted Cells and Its Application to Freshly Isolated Murine Immune Cells. Mol. Cell. Proteomics *18*, 995–1009.

Olsson, N., Heberling, M.L., Zhang, L., Jhunjhunwala, S., Phung, Q.T., Lin, S., Anania, V.G., Lill, J.R., and Elias, J.E. (2021). An Integrated Genomic, Proteomic, and Immunopeptidomic Approach to Discover Treatment-Induced Neoantigens. Front. Immunol. *12*, 662443.

Ouspenskaia, T., Law, T., Clauser, K.R., Klaeger, S., Sarkizova, S., Aguet, F., Li, B., Christian, E., Knisbacher, B.A., Le, P.M., et al. (2020). Thousands of novel unannotated proteins expand the MHC I immunopeptidome in cancer.

Rivera, K.D., Olive, M.E., Bergstrom, E.J., Nelson, A.J., Lee, K.A., Satpathy, S., Carr, S.A., and Udeshi, N.D. (2021). Automating UbiFast for High-throughput and Multiplexed Ubiquitin Enrichment. bioRxiv 2021.04.28.441860.

Rocha, N., and Neefjes, J. (2008). MHC class II molecules on the move for successful antigen presentation. EMBO J. *27*, 1–5.

Ruggles, K.V., Tang, Z., Wang, X., Grover, H., Askenazi, M., Teubl, J., Cao, S., McLellan, M.D., Clauser, K.R., Tabb, D.L., et al. (2016). An Analysis of the Sensitivity of Proteogenomic Mapping of Somatic Mutations and Novel Splicing Events in Cancer. Mol. Cell. Proteomics *15*, 1060–1071.

Sarkizova, S., Klaeger, S., Le, P.M., Li, L.W., Oliveira, G., Keshishian, H., Hartigan, C.R., Zhang, W., Braun, D.A., Ligon, K.L., et al. (2020). A large peptidome dataset improves HLA

class I epitope prediction across most of the human population. Nat. Biotechnol. *38*, 199–209.

Satpathy, S., Jaehnig, E.J., Krug, K., Kim, B.-J., Saltzman, A.B., Chan, D.W., Holloway, K.R., Anurag, M., Huang, C., Singh, P., et al. (2020). Microscaled proteogenomic methods for precision oncology. Nat. Commun. *11*, 532.

Satpathy, S., Krug, K., Jean Beltran, P.M., Savage, S.R., Petralia, F., Kumar-Sinha, C., Dou, Y., Reva, B., Harry Kane, M., Avanessian, S.C., et al. (2021). A proteogenomic portrait of lung squamous cell carcinoma. *accepted in principle*.

Schweppe, D.K., Prasad, S., Belford, M.W., Navarrete-Perea, J., Bailey, D.J., Huguet, R., Jedrychowski, M.P., Rad, R., McAlister, G., Abbatiello, S.E., et al. (2019). Characterization and Optimization of Multiplexed Quantitative Analyses Using High-Field Asymmetric-Waveform Ion Mobility Mass Spectrometry. Anal. Chem. *91*, 4010–4016.

Svinkina, T., Gu, H., Silva, J.C., Mertins, P., Qiao, J., Fereshetian, S., Jaffe, J.D., Kuhn, E., Udeshi, N.D., and Carr, S.A. (2015). Deep, Quantitative Coverage of the Lysine Acetylome Using Novel Anti-acetyl-lysine Antibodies and an Optimized Proteomic Workflow. Mol. Cell. Proteomics *14*, 2429–2440.

Taylor, H.B., Klaeger, S., Clauser, K.R., Sarkizova, S., Weingarten-Gabbay, S., Graham, D.B., Carr, S.A., and Abelin, J.G. (2021). Mass spectrometry-based HLA-II peptidomics combined with multi-omics will aid the development of future immunotherapies. Mol. Cell. Proteomics.

Udeshi, N.D., Mertins, P., Svinkina, T., and Carr, S.A. (2013). Large-scale identification of ubiquitination sites by mass spectrometry. Nat. Protoc. *8*, 1950–1960.

Udeshi, N.D., Mani, D.C., Satpathy, S., Fereshetian, S., Gasser, J.A., Svinkina, T., Olive, M.E., Ebert, B.L., Mertins, P., and Carr, S.A. (2020). Rapid and deep-scale ubiquitylation profiling for biology and translational research. Nat. Commun. *11*, 359.

Vyas, J.M., Kim, Y.-M., Artavanis-Tsakonas, K., Love, J.C., Van der Veen, A.G., and Ploegh, H.L. (2007). Tubulation of class II MHC compartments is microtubule dependent and involves multiple endolysosomal membrane proteins in primary dendritic cells. J. Immunol. *178*, 7199–7210.

Wang, L.-B., Karpova, A., Gritsenko, M.A., Kyle, J.E., Cao, S., Li, Y., Rykunov, D., Colaprico, A., Rothstein, J.H., Hong, R., et al. (2021). Proteogenomic and metabolomic characterization of human glioblastoma. Cancer Cell.

Weingarten-Gabbay, S., Klaeger, S., Sarkizova, S., Pearlman, L.R., Chen, D.-Y., Gallagher, K.M.E., Bauer, M.R., Taylor, H.B., Dunn, W.A., Tarr, C., et al. (2021). Profiling SARS-CoV-2 HLA-I peptidome reveals T cell epitopes from out-of-frame ORFs. Cell.

Wosen, J.E., Mukhopadhyay, D., Macaubas, C., and Mellins, E.D. (2018). Epithelial MHC Class II Expression and Its Role in Antigen Presentation in the Gastrointestinal and Respiratory Tracts. Front. Immunol. *9*, 2144.

Yadav, M., Jhunjhunwala, S., Phung, Q.T., Lupardus, P., Tanguay, J., Bumbaca, S., Franci, C., Cheung, T.K., Fritsche, J., Weinschenk, T., et al. (2014). Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing. Nature *515*, 572–576.

Yi, L., Tsai, C.-F., Dirice, E., Swensen, A.C., Chen, J., Shi, T., Gritsenko, M.A., Chu, R.K., Piehowski, P.D., Smith, R.D., et al. (2019). Boosting to Amplify Signal with Isobaric Labeling (BASIL) Strategy for Comprehensive Quantitative Phosphoproteomic Characterization of Small Populations of Cells. Analytical Chemistry *91*, 5794–5801.

Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-Garcia, W., Treviño, V., Shen, H., Laird, P.W., Levine, D.A., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. Nat. Commun. *4*, 2612–2612.

Zecha, J., Satpathy, S., Kanashova, T., Avanessian, S.C., Kane, M.H., Clauser, K.R., Mertins, P., Carr, S.A., and Kuster, B. (2019). TMT Labeling for the Masses: A Robust and Cost-efficient, In-solution Labeling Approach. Mol. Cell. Proteomics *18*, 1468–1478.