Inaugural dissertation

for

obtaining the doctoral degree

of the

Combined Faculty of Mathematics, Engineering and Natural Sciences

of the

Ruprecht - Karls - University

Heidelberg

Presented by

M.Sc. Alexander Kowar

Born in Dresden, Germany

Oral examination: 10.04.2025

Mitotic-arrest induces aberrant mRNA translation and non-canonical peptide HLA class I presentation

Referees: Prof. Dr. Frank Lyko

Dr. Fabricio Loayza-Puch

„There is
So
Much
More and
Beckons me
To look through to these
Infinite possibilities“
Tool

Table of Contents

Summary

Paclitaxel (PTX) is a widely used chemotherapeutic agent against cancer. The drug disrupts microtubule dynamics, causing mitotic arrest that activates apoptotic pathways. However, the anti-cancer efficacy of PTX is limited to specific cancer types, and its use is often accompanied by significant adverse effects. Thus, further investigation into the cellular effects of PTX is essential.

In this study, I investigated the translational response induced by mitotic arrest. Using ribosome profiling (RiboSeq), I observed increased transcriptome-wide ribosome occupancy at 5' un-translated regions (5'UTRs) and 5' ends of coding sequences in treated cancer cell lines. This effect was independent of the molecular mechanism of mitotic-arrest induction. Further computational analysis revealed an increased prediction rate of non-canonical open-reading frames (ncORFs), specifically upstream and upstream-overlapping ORFs (u/uoORFs) within the 5'UTR. Notably, translation rates for these genomic features were elevated. To enable further research, I compiled uORFs and uoORFs elements from various cancer cells into a comprehensive database.

Building on these findings, I hypothesized that *in vitro* PTX treatment may lead to cell surface presentation of non-canonical peptides (nuPeptides) by HLA-I complexes. To test this hypothesis, state-of-the-art immunopeptidomics were employed to enrich HLA-I ligands and perform quantitative peptide detection via mass spectrometry in PTX- or DMSO-treated cells. This analysis revealed elevated levels of specific peptides originating from uORF or uoORF transcripts. Selected uORFs giving rise to nuPeptides were further validated in a complementary approach using uORF-SIINFEKL-reporter constructs and ex vivo CD8+ T cells. This demonstrated increased cytokine secretion and cancer cell killing capability upon PTX-induced mitotic arrest.

In summary, my results demonstrate that PTX treatment provokes aberrant mRNA translation in the 5'UTR, resulting in nuPeptide synthesis. These ligands bind to HLA-I complexes and may trigger immune responses. My findings provide new insights into treatment-induced peptide biosynthesis from uORF and uoORF sequences to benefit future immunotherapies.

Zusammenfassung

Paclitaxel (PTX) ist ein häufig angewendetes Chemotherapeutikum bei Krebserkrankungen. Der Wirkstoff mindert die Dynamiken der Mikrotubuli, welches zu Mitosearrest führt und Apoptose Signalwege aktiviert. Allerdings ist die Wirksamkeit von PTX auf spezifische Krebsarten beschränkt und die Behandlung ist häufig begleitet von starken Nebenwirkungen. Daher sind weitere Untersuchungen der zellulären Effekte von PTX essentiell.

In dieser Studie untersuchte ich die Reaktion der mRNA-Translation während des Mitosearrest. Mittels Ribosome Profiling (RiboSeq) stellte ich eine stärkere Transkriptom-weite Belegung von Ribosomen an 5'un-translatierten Regionen (5'UTRs) und 5'Enden der codierenden Sequenz in mit PTX behandelten Krebszelllinien fest. Dieser Effekt war unabhängig vom molekularen Mechanismus der Induktion des Mitosearrests. Die computer-gestützte Analyse zeigte erhöhte Prognoseraten von nicht-kanonischen offenen Leserahmen (ncORFs), im speziellen 5'UTR vorgelagert und vorgelagert-überlappende ORFs (u/uoORFs). Bemerkenswerterweise waren die Translationsraten dieser Transkriptbereiche erhöht. Um weitere Untersuchungen zu vereinfachen, stellte ich uORF und uoORF Elemente von verschiedenen Krebzelllinien in eine umfassende Datenbank zusammen.

Auf diesen Ergebnissen aufbauend, stellte ich die Hypothese auf, dass *in vitro* PTX-Behandlung zur Präsentation von nicht-kanonischen Peptiden (nuPeptides) durch HLA-I Komplexe auf der Zelloberfläche führt. Um diese Hypothese zu überprüfen, wurde die moderne Immunopeptidomics-Technik angewandt um HLA-I Liganden anzureichern und die quantitative Detektion von Peptiden mittels Massenspektrometrie von PTX- oder DMSO-behandelten Zellen durchzuführen. Diese Analyse ergab höhere Mengen spezifischer Peptide, welche von uORF und uoORF Transkripten stammen. Ausgewählte uORFs, welche nuPeptides bilden, wurden validiert. In einem ergänzenden Ansatz demonstrierten uORF-SIINFEKL Reporterkonstrukte in Verbindung mit ex vivo CD8+ T Zellen erhöhte Zytokinsekretion und gesteigertes Potential zur Eliminierung von Krebszellen während PTX-induziertem Mitosearrest.

In Zusammenfassung bekräftigen meine Ergebnisse, dass PTX-Behandlung anormale mRNA-Translation im 5'UTR hervorruft, woraus die nuPeptid-Synthese folgt. Diese Liganden binden HLA-I Komplexe und können Immunantworten auslösen. Diese Erkenntnisse gewähren neue Einsicht in die behandlungs-induzierte Peptidbiosynthese von uORF und uoORF Sequenzen, welche zukünftige Immuntherapien begünstigen werden.

Acknowledgement

First, I would like to express my gratitude to Dr. Fabricio Loayza-Puch for giving me the opportunity to pursue my PhD in his lab. The past five years of my PhD and master's thesis have been an incredible journey, allowing me to refine both my scientific motivation and computational skills.

To my committee members, Prof. Dr. Frank Lyko, Prof. Dr. Christiane Opitz, Dr. Andreas Mayer, and Dr. Wilhelm Palm, thank you for your valuable scientific input over the years. Your suggestions were always greatly appreciated and have helped shape my project.

Many thanks also go to Dr. Jonas Becker and Dr. Angelika Riemer. Their expertise in immunopeptidomics and scientific curiosity fostered a truly invaluable collaboration.

I am grateful to the amazing and brilliant students I had the honor to supervise. I truly enjoyed our scientific and non-scientific discussions, as well as all the moments we shared. Thank you, Elisabeth, Fiona, and Hanan— Thank you for your hard work, laughter, and companionship. You were amazing!

Furthermore, I am deeply grateful for the incredible team we have. Each one of you contributes to making our workplace a great environment. A huge shoutout to all my friends and co-workers in Heidelberg. Andres, Yuchong, and Zhiwei, thank you for your calm and easygoing attitudes. Rossella, I truly appreciate your cheerful mood and sharp mind, which bring so much energy to our office. Huge thanks to Rossella and Zhiwei for their effort in this project.

Giuliana, I admire your calm, happy, and resilient personality—you are such a strong person. Keep it on!

Of course, I must also thank the remarkable people who have left the institute over the years but will always be unforgettable. Cinthia, Daniela, Luisa, Bryce, and Taishi—your absence has left a huge gap. Thank you for the scientific and cultural discussions, the non-scientific distractions, and the wonderful times we spent together. Keep on "opai plum plum," and please, let's have one more karaoke night and even more school trips together!

To Daniela, Richard, and Manu—thank you for always being there. Your unconditional love, understanding, and friendship made this journey unforgettable. I will always value the bond we share.

I am also immensely grateful to all my friends scattered across Germany and the world. We have always been there for each other, and I know we always will be.

My heartfelt thanks go to my family, especially my mother and grandmother. Thank you for your support and for reigniting my spark of creativity.

Svenja, the biggest thanks go to you. You were my balance throughout this long journey, bringing joy and calmness to every moment. Now, we will step beyond the path of reason together.

Abbreviations

µl	microlitre
µM	micromolar
25-D1.16	H2-Kb:SIINFEKL detecting antibody clone
4EBP	eukaryotic translation initiation factor 4 e binding protein
AA	amino acid
APC	antigen-presenting cell
ASNS	asparagine synthetase
ADP	adenosine diphosphate
ATP	adenosine triphosphate
β_2M	β_2 microglobulin
BI	BI2536 (polo-like kinase 1 inhibitor)
BLAST	basic local alignment search tool
CD4	cluster of differentiation 4
CD8	cluster of differentiation 8
CDC25	cell division cycle 24
CDK1	cyclin-dependent kinase 1
CDS	coding sequence
CHOP	C/EBP homologous protein (DDIT3 alias)
CITEs	cap-independent translation enhancers
COPII	cytoplasmic coat protein complex II
DAMPs	damage-associated molecular pattern
DDIT	DNA damage-inducible transcript 3
DENR	density regulated re-initiation and release factor
DKFZ	Deutsches Krebsforschungszentrum
DMSO	dimethyl sulfoxid
DRIPs	defective ribosomal products
DTT	dithiothreitol
EDTA	ethylendiaminetetraacetic acid
eIF	eukaryotic translation initiation factor
ELIspot	enzyme linked immunospot assay
ER	endoplasmic reticulum
ERAAP	endoplasmic reticulum aminopeptidase associated with antigen processing
ERp57	endoplasmic reticulum-resident protein 57
GAPDH	glycerinaldehyd-3-phosphate-dehydrogenase
GCN2	general control non-derepressible 2
GO	gene ontology
GrB	granzyme B
GTP	guanosine triphosphate
H3	histone 3
Harr	harringtonine
HLA	human leukocyte antigen
ICD	immunogenic cell death
IFNγ	interferon gamma
IPomics	immunopeptidomics
ISR	integrated stress response
ITAMs	immunoreceptor tyrosine-based activation motif
LARP1	La-related protein 1
LCK	lymphocyte-specific protein tyrosine kinases
LOX	lysyl oxidase
MCTS1	malignant t-cell-amplified sequence 1

MHC	major histocompatibility complex
mRNA	messenger ribonucleic acid
MRPL51	mitochondrial ribosomal protein L51
MS	mass spectrometry
mTORC1	mechanistic target of rapamycin complex1
NCBI	national center for biotechnology information
ncORF	non-canonical open-reading frame
NF-kB	nuclear factor kappa b
NGS	next generation sequencing
NK cells	natural killer cells
Noco	nocodazole
nuORF	non-canonical upstream/upstream-overlapping ORF
nuORFdb	nuORF database
nuPeptide	nuORF-derived peptide
ORF	open reading frame
OT-I	ovalbumin-specific TCR transgenic line
OVA	ovalbumin
PABP	poly-A binding protein
PBS	phosphate buffered saline
PBMCs	peripheral blood mononuclear cell
PCC	Pearson`s correlation coefficient
PDAC	pancreatic ductal adenocarcinoma
PDK1	pyruvate dehydrogenase kinase isoform 1
PI	propidium iodide
PIC	pre-initiation complex
PKM	pyruvate kinase
PLC	peptide-loading complex
PLK1	polo-like kinase 1
PRICE	probabilistic inference of codon activities by an expectation-maximization
PTX	paclitaxel
RefSeq	NCBI reference sequence database
RiboSeq	ribosome profiling
RNA	ribonucleic acid
RNAseq	RNA sequencing
RPFs	ribosomal protected fragments
RPKM	read per kilobase of transcript per million mapped reads
S6K1	ribosomal protein S6 kinase B1
SIINFEKL	single letter amino acid sequence from an ovalbumin peptide
SNVs	single nucleotide variants
STLC	S-Trityl-L-cysteine
TAP	transporter associated with antigen-processing
TCR	T cell receptor
TE	translational efficiency
TME	tumour microenvironment
TMM	trimmed mean of M-values
TNBC	triple-negative breast cancer
TOPmotif	5’ terminal oligo-pyrimidine motif
TPM	transcripts per kilobase million
tRNA	transfer RNA
uoORF	upstream-overlapping ORF
uORF	upstream ORF
uTISs	upstream translation initiation sites
5’UTR	5’ untranslated region
WEG	whole-genome sequencing

WES	whole-exome sequencing
WHO	World Health Organisation
ZAP-70	zeta chain of T cell receptor associated protein kinase 70

List of Figures

List of Tables

1 Introduction

This doctoral thesis describes a phenomenon of aberrant mRNA translation during mitotic-arrest. This altered process gives rise to protein biosynthesis products which are presented on the cell plasma membrane. This introduction provides insights into the related topics of this doctoral project.

1.1 Cancer and the immune system

This chapter aims to provide basic knowledge about cancer, standard-of-care treatments and the immune system, counteracting against this malignant disease.

1.1.1 Epidemiology

Cancer is one of the most abundant diseases globally. On average, every fifth person is likely to develop by a cancerous malignancy throughout their lifespan[1]. This compendium of different diseases was thoroughly studied during the current and last century leading to deeper insights. Based on this, researchers were able to classify cancer-related features into knowledge-based scaffolds that help to understand this disease.

Human cancers are a spectrum of diseases. Based on World Health Organization (WHO) cancer burden statistics, lung, breast and colorectum cancers are the global leading causes in cancer-related mortality (**Figure *1***).

In general, cancer is classified histologically by the organ/ tissue of origin: carcinoma, sarcoma, glioma, blastoma, lymphoma and leukaemia. All of these classes are further subdivided by sublayers of different tissues. According to histologic differentiation, over 200 types of cancers are known, only based on the organ or tissue the malignancy originates from[2]. Based on research and new treatment options of the last decades, general survival rates have improved from 35 % to 69.7 % in the last 60 years[3]. Alongside this strong success, the scientific progress is driven by new classifications of cancers into various subtypes.

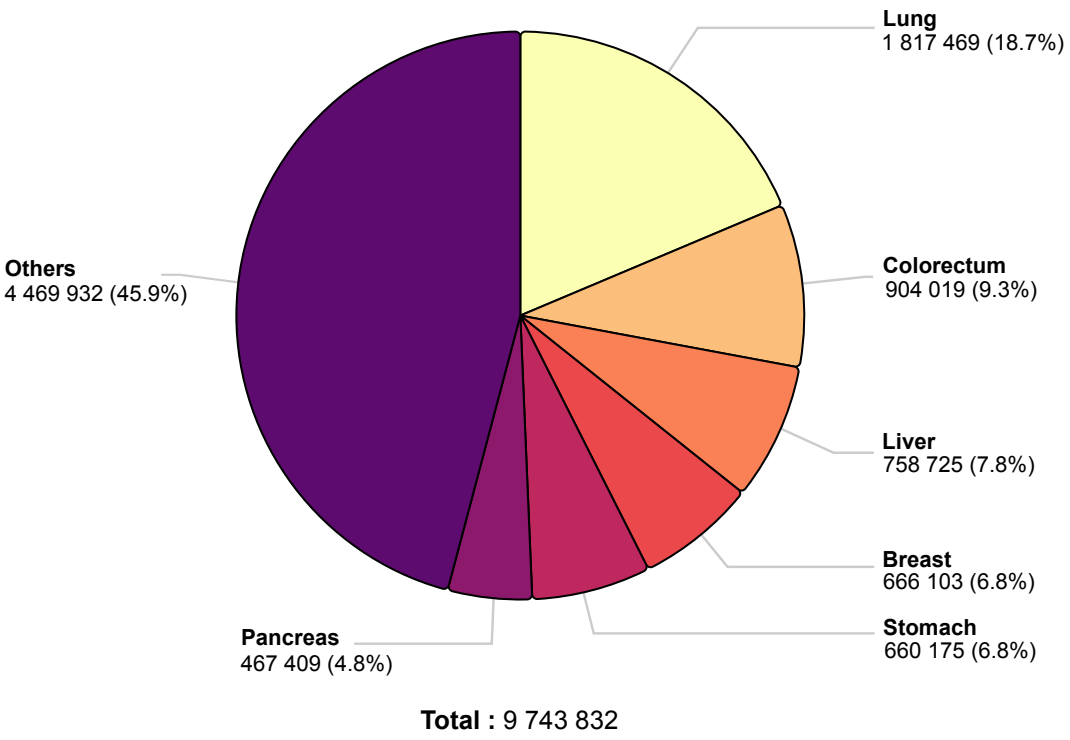

Figure 1: Global Cancer Mortality in 2022 for both sexes. International Agency for Research on Cancer, World Health Organization[1]. Globocan 2022 (version 1.1). https://gco.iarc.who.int

1.1.2 Cancer Heterogeneity

Cancer diseases harbour an extensive biological heterogeneity between patients and within tumours. This is largely mediated by clonal variation in malignant lesions, that is, healthy cells acquire or inherit different genetic mutations. These germline mutations induce oncogenic signalling, supporting cell proliferation[4,5]. If the affected cells are not cleared by the immune system or apoptosis, these cells outgrow the healthy surrounding tissue, whilst deregulating cellular protective mechanisms[6]. Hence, tumours and metastases cannot be perceived as homogenous cell clusters.

Amongst cancer cells, other cell types are present in a tumour. These can include immune cells, fibroblasts, endothelial cells and neurons, in addition to extracellular matrix factors. This combination is defined as the tumour microenvironment (TME)[7], impacting cancer cell proliferation. TME elements can act eradicative to cancer cells, by activation of immune responses and cancer killing[8]. However, TME elements also induce a hostile environment against efficient cancer killing and support a local immunosuppressive environment[8]. The TME composition depends on several factors. The tissue of origin, which determines the grade of vascularization, immune surveillance, metabolic supply and oxygenation. All of these factors shape the TME and produce a unique surrounding for cancer cells, which evolves dynamically and requires therapies that can adapt to these changes.

1.1.3 Chemotherapy

1.1.3.1 General classification

Cancer treatment has changed very strongly in the past decades. Despite modern approaches of patient stratification based on cancer germline mutations and the individual transcriptomic landscape, classical treatment options still imply surgical removal, radiotherapy and chemotherapy.

The latter one involves the combined administration of drugs to tackle tumour heterogeneity. The classic major drug classes involve alkylating agents, antimetabolites and mitosis-arresting drugs[9]. Drugs belonging to these classes are used in adjuvant and neoadjuvant treatment settings. Adjuvant treatments are administered after surgical removal of tumour masses to minimize recurrence and spread of remaining cancer cells. In contrast, neoadjuvant treatment settings are administered before surgery to reduce tumour infiltration in healthy tissue, which supports induces shrinkage and efficient removal, thus, reduced cancer burden of tissue-remaining cells.

1.1.3.2 Mitotic-arrest: Applications and Limitations of Paclitaxel

Mitosis is a sensitive cell state during the cell cycle promoting cell division, in which the nucleus is disrupted and chromosome pairs are separated by the microtubule spindle apparatus. This particular cell cycle phase is tightly regulated by Cyclin-dependent-kinase 1 (CDK1). Before mitosis CDK1 binds cyclin B1, but a inhibitory phosphorylation modulates low activity of CDK1 in G2 cell cycle phase[10,11]. However, activated cell division cycle 25 (CDC25) phosphatases remove inhibitory phosphorylation of CDK1, which promotes its activity. This activity includes the regulation of control of chromosome transmission and coordination of cytokinesis[12,13].

Normal mitotic progression in mammalian cells is considered to take place within less than one hour[14]. Drug-induced mitotic-arrest is a process holding cells in mitosis over the course of hours, which underlines the significant difference between these states.

One sub-class of mitosis-arresting drugs are taxanes. These toxic phytochemicals are alkaloids from yew tree species. The mechanism of action of taxanes is described as binding to microtubule moieties of the spindle apparatus, which blocks depolymerization and leads to a cell cycle arrest in the mitotic metaphase[15].

The lasting stabilization of this structure reduces the cellular ability to separate chromosome pairs. The enduring mitotic-arrest induces apoptotic signalling via caspase-activation[16]. However, the generic effect of microtubule-binding may impact also other cellular processes, which complicates mechanistic research of these compounds [17].

Clinical cancer treatment involves the administration of taxanes, of which one very common compound is Paclitaxel. The compound binds to β-tubulin between strand B9 and B10[18], which

prevents microtubule depolymerization. This drug is implicated for the treatment of HIV-associated Kaposi`s sarcoma, and cancers of lung, ovaries and breast. Paclitaxel is part of the standard of care for neoadjuvant and adjuvant treatment of breast cancer. The subclassification triple-negative breast cancer is treated with anthracyclines and taxanes in a neoadjuvant setting, whereas early stages of this cancer are treated with taxanes in an adjuvant regimen[19]. However, Paclitaxel and Docetaxel are associated with severe adverse effects. The most common effects are the reduction of neutrophil granulocytes (neutropenia) and leukocytes (leukopenia)[20], as well as peripheral neuropathy[21]. These adverse effects immunocompromise patients and drastically reduce quality of life by limb numbness or paraesthesia through neuronal axon degeneration. Also, Paclitaxel is suggested to contribute to hypersensitivity reactions. Accompanying regimen include the administration of dexamethasone with histamine receptor 1/2 antagonists prior to Paclitaxel infusion, which reduces these adverse effects, however, anaphylaxis cases are still reported. These effects are attributed to immunogenic responses.

1.1.4 The immune system

The immune system is an intricate network of cells, certain tissues and signalling molecules that shape the protection against pathogens and abnormal cells. The major arms are the innate immune system, providing non-specific, rapid responses, and the adaptive immune system, aiding with targeted, long-lasting protection. These respective immune cells originate from hematopoietic stem cells in the bone marrow, which form two lineages: the myeloid and lymphoid lineage. Myeloid lineage cells are key components of the innate immune response, such as monocytes, macrophages, dendritic cells and granulocytes. These cells are crucial for antigen presentation and thus, the activation of adaptive immunity. Lymphoid lineage cells drive the adaptive immunity through T cells, B cells and natural killer (NK) cells. Antigen-presenting-cells (APCs), such as macrophages and dendritic cells, can recognize antigens by Toll-like receptors[22], which results in their activation. Mature dendritic cells migrate to lymphoid organs, where they activate naïve CD8+ T cells via cross-presentation. This process is crucial for the initiation of cytotoxic immune responses and immune surveillance.
In-depth specifics about adaptive and innate immunity are excellently reviewed elsewhere[23,24].

1.1.5 HLA Structure and Polymorphism

Antigen presentation is the key step in adaptive immune response. Antigens (peptides) are presented in glycoproteins majorly encoded by the human leukocyte antigen (HLA) gene cluster, which is localized on the short arm of the human chromosome 6 [25]. The HLA definition

only comprises human proteins, the general term for vertebrates is major histocompatibility complex (MHC). In general, HLA glycoproteins are composed as complexes. HLA-I complexes are composed of an α-chain, with three domains, and a β_2-microglobulin (β_2m). The α_1 and α_2 chain form a peptide-binding site, whereas α_3 acts as membrane anchor. These complexes are presented on almost all nucleated cells and thrombocytes. In contrast, HLA-II complex heterodimers are composed of one α-chain, with α_1 and α_2, and of one β-chain, with β_1 and β_2. HLA-II complexes are expressed on antigen-presenting cells, such as macrophages, dendritic cells and B cells.

Both complexes possess subunits from genetic polymorphic regions, which allows for a high diversity, with more than 25,000 known alleles, while every complex can present hundreds or thousands of peptides[26,25]. These alleles possess affinities for peptides of certain lengths (HLA-I: 8-10 amino acids; HLA-II: 13-25 amino acids) and amino acid (AA) sequences or motifs[27]. These facts explain the highly diverse landscape of peptides presented on the plasma membrane amongst different individuals. The HLA polymorphism enables the diverse peptide-binding capabilities for pathogen- or tumour-derived peptides. This is supported by allele-associated increased immune responses against HIV (HLA-B*57)[28] or extended survival of melanoma patients carrying the HLA-B44 supertype[29]. On the contrary, the HLA-B62 supertype is discussed to impair peptide recognition and is associated with poor outcomes in melanoma patients[29]. Thus, the HLA polymorphism and expression strongly impacts the degrees of immunosurveillance and immune responses, which can determine patient survival.

1.1.6 Antigen processing

In order to be presented, MHC ligands need to be processed. For MHC-I complexes, these antigens originate from the cellular proteome, although extracellular proteins are also utilized, which is defined as 'cross-presentation'[30]. However, faulty mRNA translation products, including codon misreading or translational frameshifts, induce defective protein folding. These misfolded proteins are either ubiquitinylated in the cytosol and subjected for proteasomal proteolysis, or are degraded in endolysosomal compartments. The resulting spliced peptides are transported into the endoplasmic reticulum (ER) by the transporter associated with antigen-processing (TAP). MHC class I complexes are loaded by the peptide loading complex (PLC). This complex involves the MHC-I moiety, TAP, Tapasin, ERp57 and a chaperone, such as calreticulin or calnexin (**Figure *2***). First, calreticulin recruits an unloaded MHC-I complex into the asymmetric PLC, forming the symmetric and active PLC. Next, TAP translocates a cytosolic peptide into the PLC in the ER lumen under ATP hydrolysis. Following potential further trimming by ER aminopeptidase associated with antigen processing

(ERAAP), peptides are loaded onto MHC-I complexes and are transferred to the Golgi via COPII, followed by presentation on the cell plasma membrane[25].

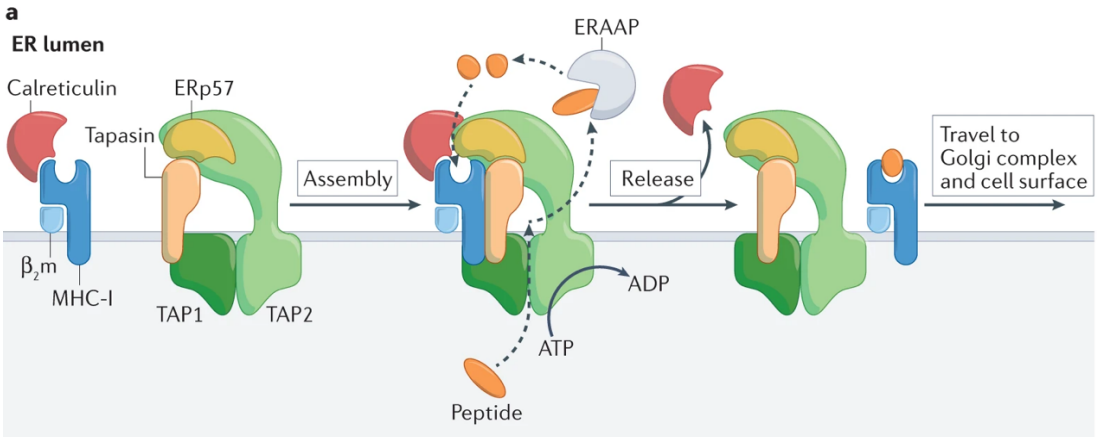

Figure 2: HLA-I peptide loading. Calreticulin recruits the MHC-I complex to the peptide loading complex (PLC; Tapasin, ERp57, TAP1, TAP2) forming a functional PLC. Next, TAP proteins transport a cytosolic peptide into the ER lumen in an ATP-dependent manner. The peptide might be further processed by ERAAP. Upon peptide loading, the MHC-I complex dissociates from the PLC and will be transported to the Golgi complex and cell surface. ATP/ADP: Adenosine tri/di-phosphate, ERAAP: ER aminopeptidase associated with antigen processing, MHC-I: Majyo histocompatibility complex I, TAP: Transporter associated with antigen processing. Reprinted and adapted from Pishesha *et al.* (2022). Reproduced with permission from Springer Nature under License Number 5953030387499.

1.1.7 Antigen presentation and cytotoxic activity

Processed antigens are presented on HLA complexes. These peptides are recognized with T cell receptors (TCRs) on T cells. HLA-I complex ligands are recognized by CD8+ T cells, whereas HLA-II complex ligands are recognized by CD4+ T cells.

Upon peptide engagement the TCR complex recruits lymphocyte-specific protein tyrosine kinases (LCKs), which induces ZAP-70 phosphorylation and interaction, leading to tyrosine phosphorylation on the cytosolic tail of the CD3 proteins (immunoreceptor tyrosine-base activation motifs, ITAMs). This triggers the local assembly of multi-nuclear signalosomes that induce T cell expansion and cytokine transcription. Classical cytokines such as perforin and granzyme B, which promote cytotoxicity by induction of apoptosis. Perforin assembles into transient pores, allowing rapid intracellular access of granzyme b (GrB)[31]. GrB is a serine protease that activates pro-caspases and additional proteins to induce apoptosis[32,33]. Another key cytokine primarily released by $CD8^+$, $CD4^+$ and NK cells is interferon γ (IFNγ). This factor activates the gene expression of HLA proteins and peptide processing proteins, which induces

higher degrees of ligand cell surface display. However, cytokine expression is strongly regulated by RNA-binding proteins during mRNA translation, which allows efficient and rapid T cell responses[34].

Summarizing, all factors of HLA complex disposition, peptide supply and processing for cell surface display and the subsequent activation of T cells for cytokine secretion are essential factors in anti-tumour responses. However, all of these factors are subject to translational control.

1.2 mRNA translation

This chapter provides an in-depth overview of protein biosynthesis, non-canonical mRNA translation and its products.

1.2.1 Translation initiation

mRNA translation is a strongly regulated process in which ribosomes, tRNAs, amino acids and translation factors convert mRNA into proteins. This process is mediated in different process steps: initiation, elongation, termination and ribosome recycling, of which initiation is the rate-limiting step. Translation initiation is divided into cap-independent and cap-dependent initiation. The first one is mainly mediated by internal-ribosome entry sites and other structural elements binding important factors that recruit mRNAs to ribosomes, defined as cap-independent translation enhancers (CITEs)[35].

Cap-dependent initiation involves first the binding of the eIF4F complex to the 7'methylguanosine cap of mRNA 5'ends. This complex includes the cap-binding protein eIF4E, the RNA helicase eIF4A and the scaffold protein eIF4G, also interacting with the poly-A-binding protein PABP of mRNA poly-A 3'ends. Next, the 43S pre-initiation-complex (40 S ribosomal subunit, eIF3, eIF1, eIF1A, eIF5) (PIC) and ternary complex (GTP-loaded eIF2, initiator methionyl $tRNA_i$) are recruited to assemble the 48 S initiation complex. This complex scans the 5' untranslated region (UTR) until reaching an AUG start codon. The AUG recognition elicits GTP hydrolysis in the ternary complex, which induces the 60 S ribosomal subunit to bind the 48 S initiation complex. The assembly of both complexes triggers the release of eIF2-GDP, the hydrolysis of eIF5B-GTP in connection with the subsequent release of eIF5B-GDP and eIF1A. This process forms the active 80 S ribosome, which translates mRNA information from the start codon to stop codon (referred to as open-reading-frame, ORF)[36].

However, ribosomes can also initiate translation outside the CDS, which is determined as non-canonical translation (elaborated in section **1.2.6 ORFs**).

The following steps of elongation involve the addition of amino acids to the growing peptide chain. Anticodon-fitting aminoacylated tRNAs enter the ribosome in the A-site, in which the aminoacyltransferase activity promotes amino acid addition to form the growing peptide chain. When reaching an in-frame stop codon, mRNA translation is terminated, inducing dissociation of ribosomal subunits and deacylated tRNAs. The resulting ribosomal subunits are recycled for further mRNA translation processes.

1.2.2 Translational Control

1.2.2.1 4EBP

Protein biosynthesis is largely mediated by S6K1 (p70S6 Kinase 1) and eIF4E binding protein (4EBP). S6K1 phosphorylation by mTORC1 enables PDK1-mediated activation of S6K1. In turn, S6K1 phosphorylates and activates mRNA translation-promoting targets, such as eIF4B [37]. On the contrary, 4EBP can associate to the cap-binding protein eIF4E, which prevents eIF4F complex formation and hinders cap-dependent translation initiation. 4EBP phosphorylation via mTORC1 supports dissociation from eIF4E, allowing efficient translation initiation[38]. However, other major factors are described as translation initiation modulators by 4EBP phosphorylation, such as CDK1. Despite its well-defined role as cell-cycle modulator, CDK1 was reported to replace mTORC1-mediated phosphorylation of 4EBP1. This enables the selective cap-dependent mRNA translation of certain transcripts[39,40].

1.2.2.2 GCN2

The major mechanism directly connecting amino acid sensing and mRNA translation is tRNA aminoacylation. This highly specific process is catalysed by the class of aminoacyl-tRNA-synthases. If intracellular amino acid levels are reduced, cognate unloaded tRNAs accumulate and bind to the general control non-derepressible 2 (GCN2) kinase supported by GCN1[41]. The binding induces conformational changes in GCN2, which induces its kinase activity. Subsequently, GCN2 phosphorylates the translation initiation factor 2α (eIF2α). Phosphorylated eIF2α blocks the PIC formation by dissociation of eIF2α from the ternary complex, which diminishes 5' cap-dependent translation initiation. The phosphorylation of eIF2α induces the integrated stress response (ISR) supporting cellular survival, stress recovery and apoptosis[42]. This induces the expression of ISR downstream targets such as asparagine synthetase (ASNS) and damage-inducible transcript 3 (DDIT3 or CHOP).

1.2.3 Ribosome Profiling

One major approach to study mRNA translation is Ribosome Profiling (RiboSeq)[43]. This deep-sequencing-based technique enables the detection of ribosome-protected fragments (RPFs) of mRNA during translation.

Throughout this approach, cells are lysed, translation is halted by the addition inhibitors and cytosolic mRNA is digested with RNases (**Figure *3*a**). Translating ribosomes act as protective elements against RNase and shields about 30 nucleotides of mRNA against enzymatic digestion, which is defined as RPF. Hence, any unbound mRNA will be degraded. The resulting suspension is enriched for cytosolic ribosomes by density gradient centrifugation. Next, resulting RPFs are extracted and prepared for Next-Generation-Sequencing (NGS).

RPF alignment to a reference genome allows to determine translated mRNA transcripts, as well as the exact ribosome location on the mRNA transcript. This gives rise to ORF information, which aided the discovery of alternative protein products[44,45]. Initiation at these positions can also be inferred by the usage of certain translation inhibitors. Cycloheximide is commonly used to inhibit translation elongation, but compounds such as Harringtonine inhibit ribosomes directly after initiation. Harringtonine sterically prevents the insertion of an aminoacyl-tRNA to the ribosomal A-site after translation initiation[46]. This block peptide bond formation and allows the detection of local RPF densities originating from translation initiation[47].

RPF periodicity is an important quality measure of RiboSeq data. The ribosome moves along the mRNA transcript in 3-nt increments, thus, the majority of RPFs follow this periodicity (**Figure *3*c**). Furthermore, ribosome sites can be determined, which opens the path for site-specific quantification. In fact, reduced AA availability and subsequently lower tRNA aminoacylation is associated with ribosome pausing at these codons[48]. The detection of this translation snapshot underlines the importance of RiboSeq to capture and study rapid cellular changes. Latest advancements raised bulk RiboSeq to the single-cell level and detected cell cycle-dependent ribosome pausing at certain codons, suggesting AA limitations for these cell states[49]. This elegant concept was very recently further refined and elevated to perform single-cell RiboSeq to capture *in vivo* translational landscapes during aging[50].

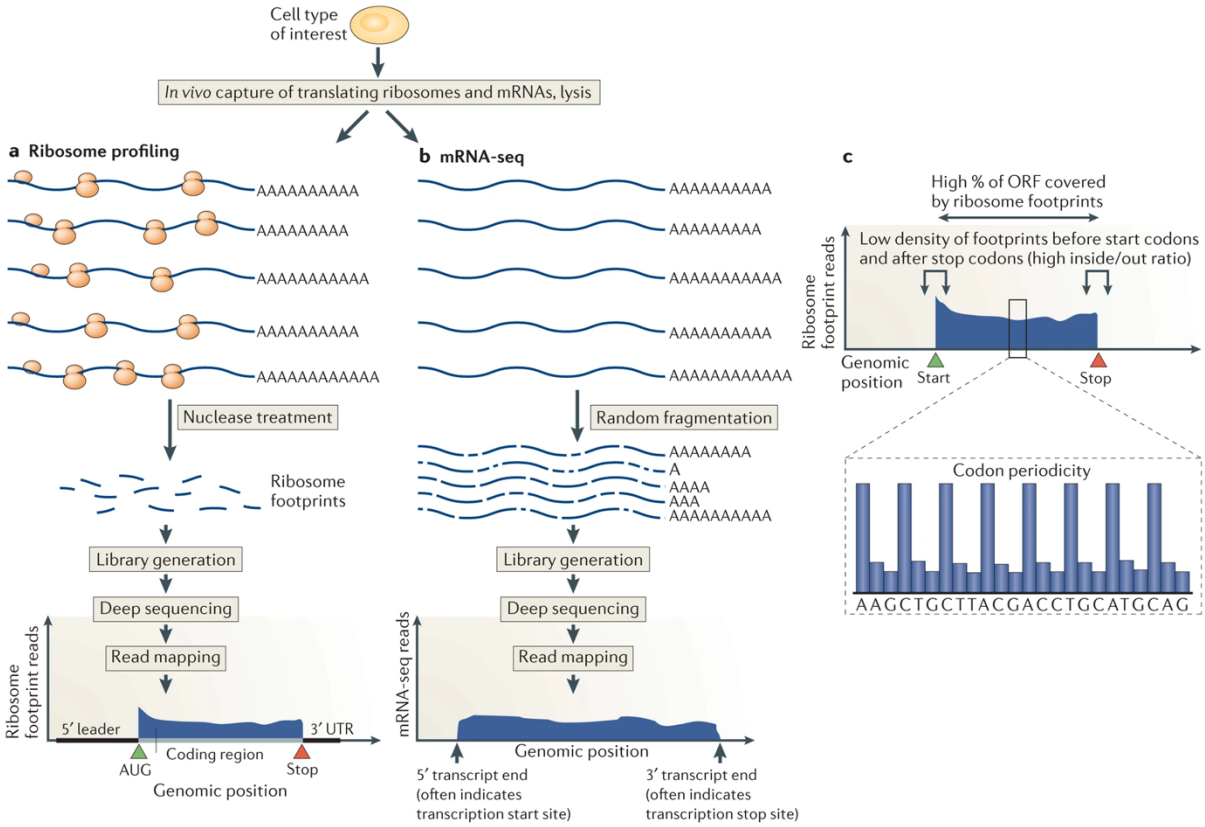

Figure 3: Overview of RiboSeq, RNAseq and periodicity. During Ribosome Profiling (RiboSeq **(a)**) cells are lysed and mRNAs are subjected to nuclease digestion, leaving only small ribosome footprints or protected fragments (RPFs) from where the ribosome was located. During RNAseq **(b)** sample preparation, mRNAs undergo a random fragmentation. After respective size selection, the mRNA fragments are used for NGS library generation and are subjected to deep sequencing. Fragment or read mapping highlights the positional differences between both approaches: RNAseq read cover the whole transcribed mRNA, whereas RiboSeq reads cover the translated area of a particular transcript. **(c)** The majority of RPFs map to the main ORF of a given mRNA, with only low RPF density of in the 5' and 3' UTR. Main ORF reads exhibit a periodic 3 nucleotide (nt) pattern because ribosomes translate in increments of 3 nt. Reprinted and adapted from Brar and Weissmann *et al.* (2017) with Licence Number 5953031299143.

1.2.4 Translational Efficiency

Classically, gene expression studies mainly focus on RNAseq to determine transcript abundance to allow functional hypotheses. However, translational control is a highly regulated step, which impacts protein abundance. That is, high transcript abundance does not imply high rates of translation. Under the assumption that high RPF density per transcript is proportional to translational speed and mRNA abundance, translational efficiency and protein synthesis rates can be predicted[39]. This can be achieved by RNAseq and RiboSeq performance of the same sample. Subsequently, RPFs are normalized to transcript abundance, enabling a differential perspective about efficiently translated mRNA transcripts[51,52].

1.2.5 RiboSeq Bioinformatic Developments

Further developments include full-fetched quality control and processing pipelines like RiboseQC and ribosomeProfilingQC[53], or the interrogation for specific features such as codon content in the ribosomal E,P and A-site[48]. Over the past years, diverse software emerged to predict ORF locations[54–56]. These genomic features can only be inferred with techniques that cover mRNA translation, as sheer transcript abundance is not indicative of translational activity. Thus, these tools apply a fundamental advantage in the prediction of conditional ORFs and their respective protein biosynthesis products, which were shown to be promising in the detection of novel peptides[56–58].

1.2.6 non-canonical ORFs and peptides

mRNA translation is performed in open-reading frames (ORF). The most generic definition of an ORF, is a nucleotide sequence with a length divisible by three, that starts with an mRNA translation start codon and ends at a stop codon. This definition is broad and includes ORFs from coding-sequences (CDSs), named canonical ORFs. But it also includes ORFs from non-CDS origin, such as UTRs and non-coding RNAs[59].

Non-canonical ORFs (ncORFs) and UTR-originating ORFs can have significant impact onto mRNA translation efficiency, by *cis*-acting sequences that induce or inhibit protein biosynthesis of a given transcript[60]. Upstream-ORFs (uORFs), ORFs that start and end within the 5'UTR, are a common feature of mRNAs. A very recent, comprehensive study determined that about 70 % of human coding genes exhibit translation outside the canonical ORF, with approximately 60 % happening uORFs or upstream-CDS-overlapping ORFs (uoORFs)[61].

The functional, *cis*-acting role of uORF-mediated translation is well described for GCN4 in yeast. This protein is a regulator of starvation responses, and exhibits four uORFs. Under nutrient rich conditions, ribosomes initiate in the uORFs 2-4, leading to termination before the translation of the main ORF. However, during nutrient stress conditions, ribosomes scan through uORFs 2-4, leading to initiation at the main ORF[60,62].

Start codon fidelity also determines non-canonical translation initiation. The same study predicted that 56.7 % genes use an AUG start-codon amongst all ncORFs[61]. However, ribosomes may initiate at other low-fidelity initiation sites (CGA, CUG) under special conditions while scanning for coding sequence AUG codons[63]. Furthermore, the impact of start codon fidelity was also shown for the functional absence of the translation elongation factor eIF5A. The absence pauses ribosomes during translation, which induces a ribosome-queue and upstream-initiation of other ribosomes in the 5'UTR[64].

Apart from translational control by *cis*-acting structural interaction, ncORF can give rise to non-canonical peptides (ncPeptides). These peptides can impact translational efficiency in-*trans*. Such as the 93 AA-long peptide/microprotein from the second uORF of the 1A glucocorticoid receptor transcript, which modulates the expression of the same transcript[65]. Furthermore, uORF-derived peptides were also shown to form stable complexes with the downstream-encoded main ORF protein[66].
Strikingly, recent studies found ncPeptides to be presented in HLA-I complexes on the cell membrane of cancer cells[67,68]. However, non-canonical peptides only contributed to only 3.3 % of the HLA-I immunopeptidome in this study. In contrast to the whole proteome, these peptides originate from long non-coding RNAs, 5'UTRs and out-of-frame translation shifts[68]. These antigens are presented for T cell recognition and can elicit an immune response to eliminate the presenting cell[67,69]. Thus, ncPeptides significantly enrich the pool of potential immunogenic epitopes for the development of cancer vaccines or immunotherapies.

Approximately 1 % of the entire human genome is annotated as protein-coding. However, about 75 % is transcribable and maybe subject to protein biosynthesis, which would massively expand the compendium of potential peptide targets[70,67]. Thus, high-throughput peptide detection approaches were developed recently to enable faster epitope detection.

1.3 Immunopeptidomics

This chapter provides a systematic overview about the technique, issues and clinical importance.

1.3.1 General approach

The landscape of peptides presented by the HLA system is defined as immunopeptidome. Immunopeptidomics (IPomics) is a technique that enables the identification of peptides presented in HLA-I and HLA-II complexes. This elegant approach combines HLA complex I or II immunoprecipitation with peptide/ligand isolation and Liquid chromatography (LC)-MS/MS detection. Early pioneering work enabled the detection of tens of peptides from billions of cells, whereas nowadays tens of thousands of peptides are detected routinely[71]. Very recent technical advancements also allow the detection of epitopes from low input material of patient-derived xenografts[72].
The usage of IPomics delivers information about the presence or absence of a peptide, but is also used to quantify a given epitope between samples[73],[74]. However, to determine the identity of a peptide, the sequence needs to be predicted from NGS data to create a database or

'search space' for possible epitopes. Whole-Genome (WEG), Whole-Exome sequencing (WES), RNA seq and RiboSeq data are used to determine tumour-specific peptides. Whereas WEG can provide insights into mutational burden of protein-coding segments and chromosomal allele amplification, RNAseq delivers a more comprehensive pictures of the transcribed RNA landscape and thus, further information about epitope origins can be inferred[75]. However, as mentioned before, only transcript abundance does not support the identification of translated peptides. In contrast, RiboSeq provides a sharper picture of which parts of a transcript are translated and in which quantity. This enables the collection of a more reliable 'search space' by exclusion of untranslated regions[67]. Thus, RiboSeq provides strong evidence of potential peptides from mRNA transcript regions or long non-coding RNAs.

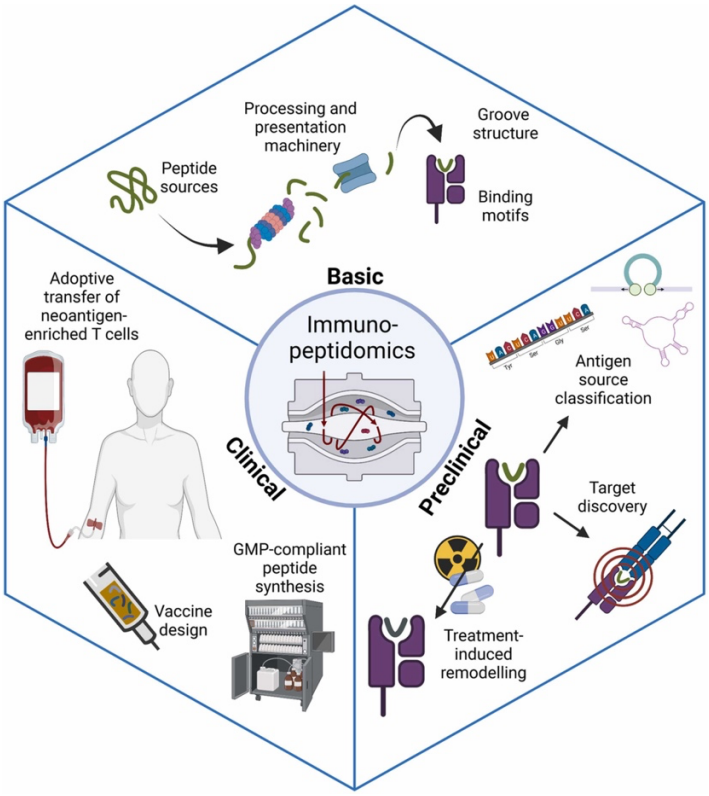

Figure 4: Immunopeptidomics applications. Immunopeptidomics can answer a broad spectrum of scientific hypotheses. Basic research can tackle the question about cell-state dependent peptide sources, alterations peptide processing and peptide properties that ensure efficient ligand presentation. Preclinical research tackles the fields of treatment-induced immunopeptidome changes, the discovery of new target and the classification of the antigen source. Within clinical settings, Immunopeptidomics is utilized to define patient-specific immunogenic peptides, which can be used for adoptive transfer of pre-loaded T cells or peptide vaccines. (Creative Commons CC-BY license)[71]

1.3.2 Chances and Opportunities

The steadily-increasing importance of IPomics and its applications are underlined by the pushing questions in different scientific fields.

The academic and clinical interest in which peptide repertoires are presented in different cell types is very strong. Providing these identified epitopes to public annotation portals[58] is of immense public interest, because it allows to filter benign reference peptides to determine tumour neo-antigens. Furthermore, it allows the investigation of peptide sources, antigen-processing factors, such as peptidases and chaperones[71]. Clinical significance is already reached with the identification of tumour-specific or tumour-associated neo-epitopes that advance personalized immunotherapies against cancer[76,77]. Additionally, pre-clinical stage usage of IPomics elucidates the source classification of an epitope, as well as initial discovery of target peptides (**Figure *4***).

1.3.3 The Drug-Induced Immunopeptidome

The pre-clinical identification of opportunities for immunopeptidome modulation holds promising paths. Specifically, determination of the treatment-induced immunopeptidome remodelling already provided first insights into this new and unexplored field. This scientific avenue will be able to identify conditionally presented peptides, holding a great chance to identify inducers of autoimmune diseases or antigens for therapeutic exploitation.

Altering cellular homeostasis changes the expression of source transcripts and resulting peptides or proteins. In a cancer-related context, this is a result of systemic medication or radiation.

Pioneering work in the field of chemotherapeutic regimen-induced immunopeptidome alterations was conducted with the cytostatic compounds Decitabine and Doxorubicin[78,79]. A 72-h treatment with Decitabine induced a differential upregulation of 1855 peptides, while 23,439 epitopes were detected globally. From the upregulated peptides, only 72 were identified as cancer antigens. For Doxorubicin, 239 differentially upregulated HLA-I peptides were detected following a time course of 24-h in HCT116 cells, while a total of 3349 unique HLA-I ligands were detected.

These findings underline the dynamic adaptation of the immunopeptidome landscape and render an understanding of its conditional change. However, the current state of knowledge in this field still lays in its infancy. Current studies mainly focus on single and defined cancer cell lines, providing basic understanding of the underlying principles of peptide presentation, peptide origin in these cells, as well as common HLA ligands[68,80]. However, a very recent study with clinical focus shed light onto tumour neoantigen heterogeneity, showing differential

peptide presentation between and within distinct melanoma metastases[76]. This strengthens the clinical impact for future applications. Furthermore, personalized vaccines hold strong hopes and expectations, as demonstrated by long-lasting relapse-free survival of pancreatic ductal adenocarcinoma (PDAC) patients with immunogenic response to a personalized RNA vaccine encoding patient-specific tumour antigens[81].

Despite the recent advancements, many more compounds are used as cytostatic drugs in chemotherapy. Thus, there is a high potential in finding new treatment-induced epitopes using standard-of-care compounds. Furthermore, the identification of these peptides will foster our understanding about immunogenic adverse effects and autoimmune diseases associated with certain treatments. These findings could improve anti-cancer therapies and patient quality of life.

1.4 Project outline

This doctoral thesis evaluates the translational origin and downstream HLA localization of ncORF-derived peptides during mitotic-arrest. Mitosis-arresting compounds are a common first-line regimen for different cancers with high prevalence, such as breast cancer.
Thus, this work provides novelty in the fields of mRNA translation and treatment-associated peptide surface display, which will improve our understanding of downstream effects during mitotic-arrest and might provide new therapeutic targets.

Research questions:

1) Translational response to mitotic-arrest
2) Characterization of non-canonical mRNA translation during mitotic-arrest
3) Characterization of HLA class I peptides as product of non-canonical mRNA translation

Firstly, I sequenced and analysed RiboSeq data from asynchronous cancer cells or during mitotic-arrest, showing strong translational comparability between different arresting drugs. These drugs induce higher ribosome counts within the 5'UTR. I inferred RiboSeq data from Harringtonine run-off assays, which shows a strong increase of initiating ribosome during mitotic-arrest compared to asynchronous cell populations, suggesting active, non-canonical mRNA translation in the 5'UTR.
Next, I predicted translational active ORFs from RiboSeq data of diverse cancer types during mitotic-arrest or from asynchronous populations. Next, I obtained highest feature recognition for uORFs and uoORFs. Strongly upregulated translational efficiency of uORF/ uoORF features from U-2 OS cells during mitotic arrest further supports the hypothesis of active

translation in the 5'UTR, which would result in protein biosynthesis products. Thus, I compiled these features into a comprehensive database of nucleotide and peptide sequences.

Next, I analysed Paclitaxel-mediated peptide presentation alterations using IPomics. For this I treated U-2 OS and SUM159 cells with Paclitaxel or DMSO, and HLA-I presented peptides were determined in collaboration. Peptides originating from uORFs and uoORFs were detected. Eventually, I selected uORF sequences for evaluation for PTX-induced non-canonical peptide translation and presentation followed by targeted CD8+ T cell response, which revealed increased cytokine secretion and cancer killing. This underlines the hypothesis and potential of aberrant mRNA translation in the 5'UTR during mitotic arrest, which produces nuPeptides that are presented on the plasma membrane surface (**Figure *5***).

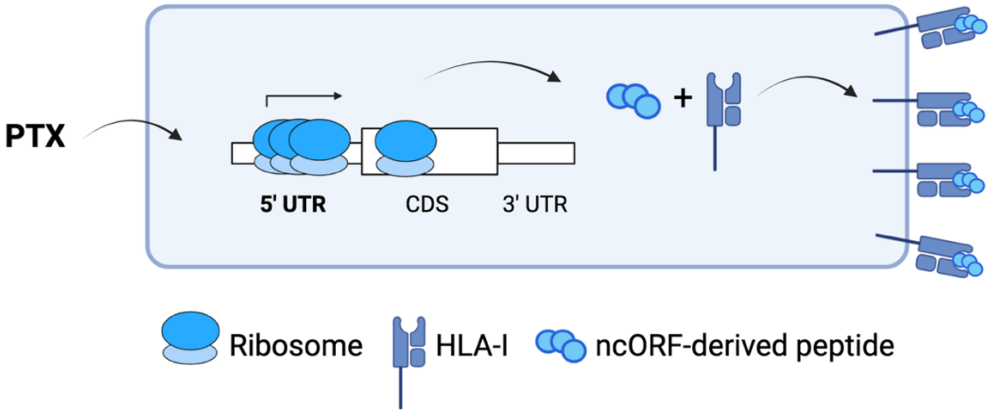

Figure 5: Global model of mitotic-arrest induced aberrant translation and ncPeptide presentation. Paclitaxel or mitotic-arrest induces ribosomes to locate and initiate at 5'UTR sequences. Non-canonical ORF (ncORF) translation gives rise to peptides, that are presented in HLA-I complexes on the plasma membrane.

2 Results

2.1 Active mRNA translation in the 5'UTR during mitotic-arrest

Mitotic-arrest is a prolonged state in mitosis over the course of hours. Experimentally, this can be induced by various agents, which act by different mechanisms. This project explored the effects of mitotic-arrest mainly in U-2 OS cells, an osteosarcoma cell line widely used as model cell line for the study of mitotic-arrest[82–84].

The following section and figures were adapted in parts from the manuscript entitled "Upstream open reading frame translation enhances immunogenic peptide presentation in mitotically arrested cancer cells"[85].

2.1.1 Consistency of mitotic-arrest induction using various drugs

To investigate the translational response during mitotic arrest in cancer cells, I analysed cell state proportions and translation patterns in the osteosarcoma cell line U-2 OS treated with different anti-mitotic compounds to establish models for mitotic-arrest enrichment. Mitotic arrest was induced with inhibitors against BI-2536 (BI), which inhibits Polo-like kinase 1 (PLK1), S-Trityl-L-cysteine (STLC), acting against Eg5, Paclitaxel (PTX) and Nocodazole (Noco), which stabilizes or destabilizes tubulin polymerization, respectively. The RiboSeq data was obtained with Fabricio Loayza-Puch.

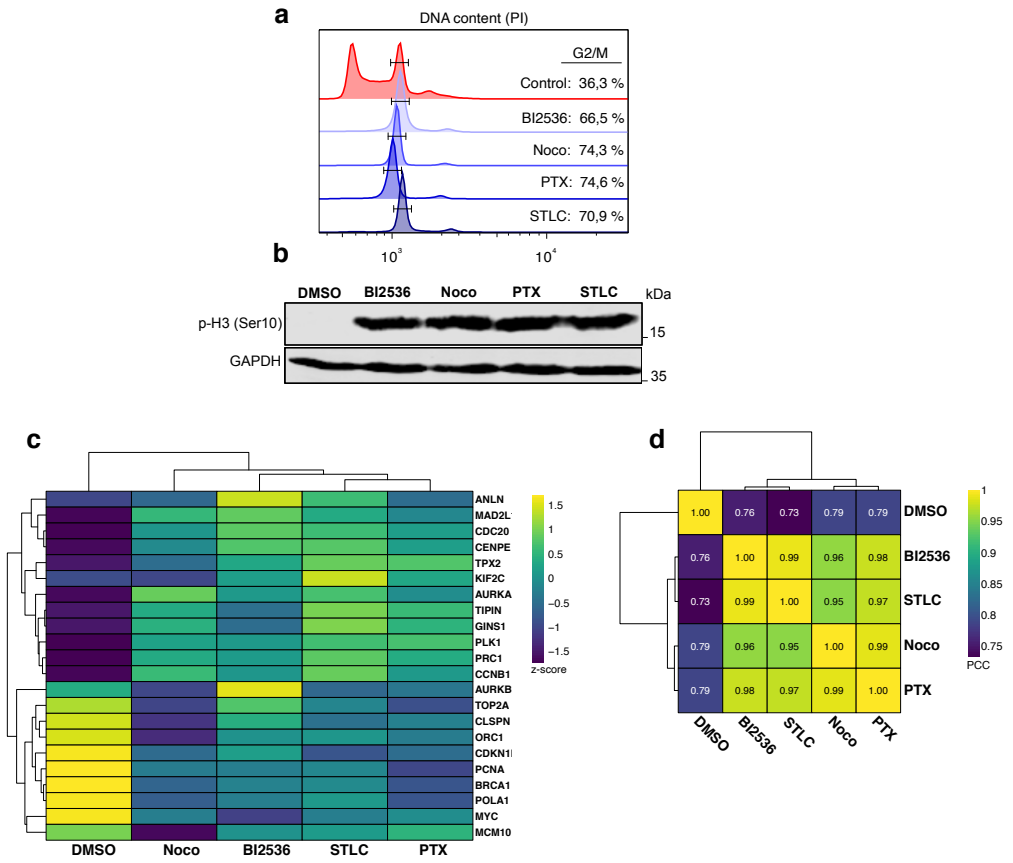

Figure 6: Mitotic-arrest induction correlates between inhibitors with different mode-of-action. U-2 OS osteosarcoma cells were treated for 16 hours with different mitosis-arresting agents (0.1 µM BI2536, 0.5 µM Nocodazole (Noco), 1 µM Paclitaxel (PTX), 1 µM S-Trityl-L-cysteine (STLC)) or DMSO. **(a)** Cell cycle analysis by DNA content with PI incorporation. **(b)** Immunoblot against Histone3 (H3) (Ser10) phosphorylation. **(c)** Clustering of TMM-normalized counts from RiboSeq, highlighting genes relevant for mitosis and S1 phase (**Figure 16 a**). **(d)** Pearson Correlation heatmap of the translational landscape from RiboSeq gene counts comparing control and mitotic-arrest conditions. PCC = Pearson Correlation Coefficient.

All drugs consistently enriched cell population percentages towards mitosis. I determined this with propidium iodide (PI)-staining against DNA content, showing increased G2/M percentages. PI intercalates into DNA in an unbiased way, allowing the discrimination of G2/M-phases with doubled chromosome amount. Additionally, I observed increased phospho-Ser10 Histone 3 (H3) signal by immunoblotting, despite different compound modes-of-actions (**Figure 6 a,b**). H3 Ser10 phosphorylation is correlated with chromosome condensation during mitosis[86]. Hence, increased G2/M population and phospho-Ser10 H3 staining support the finding of mitotic cell enrichment in all drug-treated conditions. Next, I analysed RiboSeq data from the same drug treatments and interrogated the translational landscape during mitotic arrest. For this, I compared trimmed mean of M-values (TMM)-normalized[87,88] counts. TMM-normalization allows reliable cross-sample comparison, better than 'reads per kilobase of transcript per million mapped reads' (RPKM) and 'transcripts per kilobase million' (TPM)[87].

Normalized counts exhibited a treatment-specific clustering pattern for mitotic-arrest at key mitotic transcripts (**Figure *6***; **Supplement: Table**) and in a global manner (**Supplement: Figure 16 a**), in comparison to dimethyl sulfoxid (DMSO) treatment. Furthermore, I determined that global translation patterns strongly correlated between mitotic-arrest mechanisms but not with DMSO-treated asynchronous cell populations (**Figure 6 d**). Mitotic-arrest conditions exhibited a Pearson correlation coefficient (PCC) range between 0.95 – 0.99, whereas PCC range drops to 0.73 - 0.79 in the comparison of DMSO to mitotic-arrest conditions. From this collective data, I conclude a comparable mRNA translation response across different mechanisms of mitosis arrest induction.

2.1.2 Ribosomes distribute towards the 5'UTR during mitotic-arrest

Next, I investigated the global transcript distribution of RPFs from RiboSeq data of asynchronous or mitotically-arrested U-2 OS cells. In asynchronous cell populations, I observed that RPF density was uniform across the CDS. However, during mitotic-arrest, an increased proportion of RPF density was localized within the 5'UTR region and CDS start proximity (**Figure *7* a,c,e**). My quantification of global transcript RPF distribution in the 5'UTR (metagene profile) showed an increase in mitotically-arrested U-2 OS cells, regardless of the mode-of-action (**Figure *7* b**), which is further supported in two other cell lines (**Supplement: Figure *16* b-e**) and by biological replicate data from U-2 OS cells treated with Noco (Fig.2d). During mitosis, mTOR and CDK1 modulate cap-dependent translation initiation through different factors, such as 4EBP1 (**1.2.2.1** 4EBP**; p.16**). Hence, I analysed RiboSeq data from U-2 OS cells treated with DMSO (Control), Noco or Noco + Torin1, a potent mTOR inhibitor. I observed that mTOR inhibition does not change Noco-mediated RPF localization in the 5'UTR (**Figure *6* e,f**). Furthermore, I analysed the phosphorylation state of 4EBP1 during mitotic-arrest with or without Torin1 treatment. In asynchronous U-2 OS cells, mTOR-inhibition ablated 4EBP1 phosphorylation. However, I observed that mitotic-arrest induced a phosphorylation shift of 4EBP1 with a higher molecular weight (**Figure *6* g**) in the presence of Torin1.

Thus, I conclude that ribosomes distribute towards the 5'UTR during mitotic-arrest in a mTOR-independent manner.

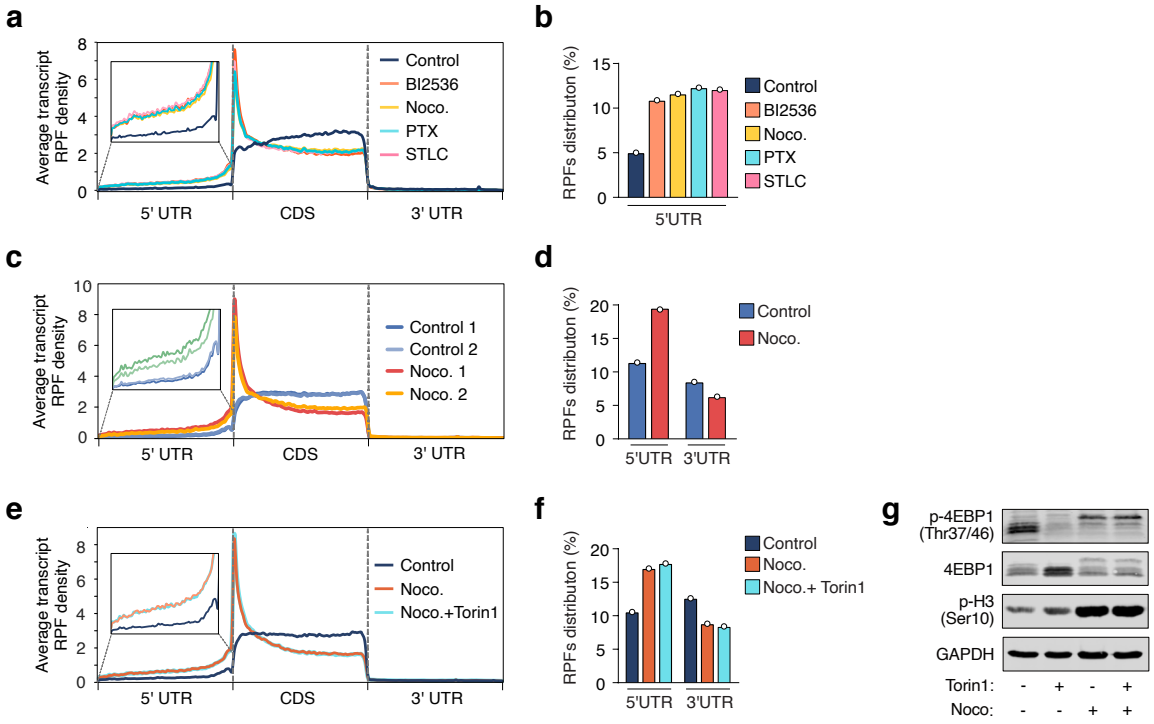

Figure 7: Mitotic-arrest promotes ribosome location towards the 5'UTR. U-2 OS cells were treated for 16 hours with indicated mitosis-arresting agents (0.1 µM BI2536, 0.5 µM Nocodazole (Noco), 1 µM Paclitaxel (PTX), 1 µM S-Trityl-L-cysteine (STLC) and/or for 2 hours with the mTORC1/2-inhbitor Torin1 (250 nM) or DMSO, followed by RiboSeq. **(a)**, **(c)**, **(e)** Metagene profiles of ribosome-protected fragment (RPF) distribution, globally segmented into 5'UTR, CDS and 3'UTR. **(b)**, **(d)**, **(f)** Percent quantification of RPF location. **(g)** Immunoblot against phospho-4EBP1 (Thr37/46), 4EBP1, phospho-H3 (Ser10) and GAPDH during combinational treatment of U-2 OS cells Torin1 and Nocodazole (Noco).

2.1.3 Ribosomes actively initiate translation in the 5'UTR during mitotic-arrest

Next, to determine if the altered RPF distribution during mitotic-arrest is associated with increased initiation rates, I analysed RiboSeq data from global ribosome run-off experiments using Harringtonine with U-2 OS cells. Harringtonine (Harr) associates with the ribosome and hinders the formation of the first peptide bond, hence, Harr treatment times correlate positively with RPF densities at translation initiation sites (TISs) (**Figure *8* a**)[46,47].

My analysis of global transcript RPF density in asynchronous cells showed a time-dependent accumulation of TISs at the main ORF (**Figure *8* b**). However, I found that mitotically-arrested cells showed an increase in upstream TISs (uTISs) with comparable levels to the main ORF under STLC treatment (**Figure *8* c**). I also observed a similar tendency in MDA-MB-231 cells (Supplement: **Figure *16* i,j**). I quantified normalized uTISs RPF counts in a cross-sample comparison, which exhibited a significant increase during STLC treatment (**Figure *8* d**).

Taking together, with my analysis of RiboSeq data, I detected that the mitotic-arrest associated translational landscape was highly comparable between different inducers/inhibitors. In contrast to asynchronous cells, mitotic-arrest induced a global increase in ribosome distribution in the 5'UTR, which arises from actively initiating ribosomes in the 5'UTR.

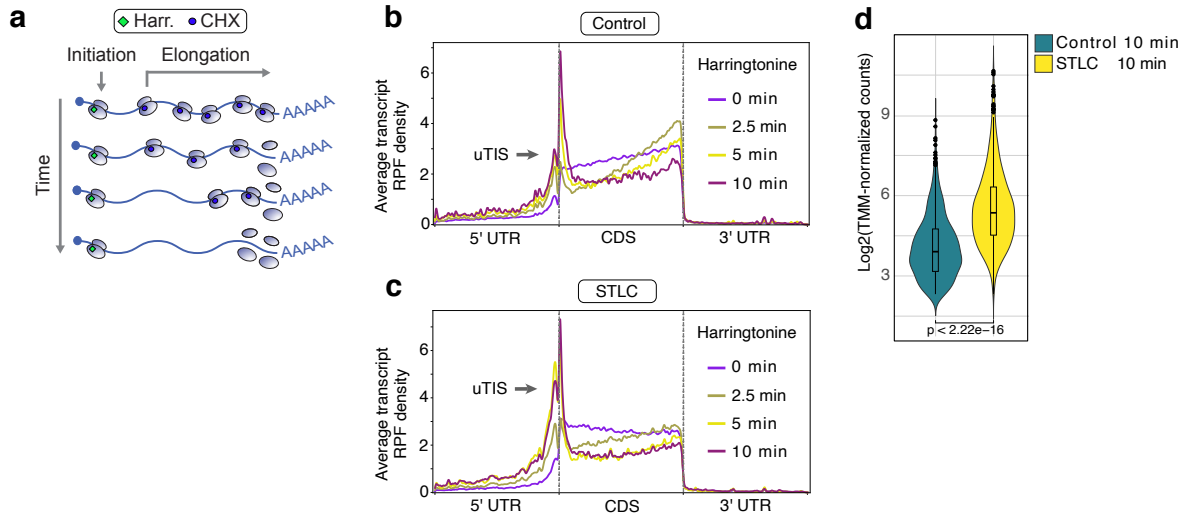

Figure 8: Harringtonine run-off assays determine increased rates of upstream translation initiation sites (uTISs). (a) Harringtonine run-off assay scheme. Cells are lysed in presence of Cycloheximide (CHX) and Harringtonine (Harr), which block ribosomes during elongation and after initiation, respectively. Thus, initiated ribosomes are enriched in a time-dependent manner. U-2 OS cells were treated with DMSO or 1 µM STLC for 16 hours. Cells were lysed in presence of 100 µg/ml CHX and 2 µg/ml Harr. **(b)**, **(c)** Metagene profile of ribosome-protected fragment (RPF) distribution, globally segmented into 5'UTR, CDS and 3'UTR. **(d)** Violin plot of trimmed mean of M-values (TMM)-normalized counts distribution of upstream translation initiation sites (uTISs) comparing DMSO and STLC-treated 10 min timepoint samples. Violin width indicates point density. The box plot center line indicates the median value, whereas whiskers define the 1.5x interquartile range. *p*-value determined by students t-test.

2.2 Elevated uORF/ uoORF translation in cancer cells during mitotic-arrest

Ribosomes actively initiate mRNA translation in the 5'UTR during mitotic-arrest. This process is defined as non-canonical translation, occurring at non-canonical open-reading frames (ncORFs). Thus, I continued to investigate non-canonical translation start sites in RiboSeq data.

2.2.1 Increased detection of uORF and uoORF elements during mitotic-arrest

To define and characterize 5'UTR initiation events, I gathered RiboSeq data from asynchronous U-2 OS cells and during mitotic-arrest induced by various agents. I subjected this data to ORF prediction with PRICE (Probabilistic inference of codon activities by an

expectation-maximization algorithm)[56]. This computational method allows ncORF identification from RiboSeq data and was shown to outperform other bioinformatic tools in the recognition rate of uORFs[89]. Subsequently, I used this process to extract information about predicted ORF moieties and start codons (**Figure *9* a,b**).

With the ORF prediction, I found a strong enrichment for uORFs and uoORFs moieties during mitotic-arrest of U-2 OS, irrespective of drug molecular mechanism (**Figure *9* c**). Notably, this increase was very pronounced with fold changes between 9-14. I observed the highest prediction rates for uORFs and uoORFs with PTX treatment. Additionally, I also observed the increase in uORF and uoORF features in two other cell lines (**Supplement: Figure *17* a**). I found coding sequence (CDS) start site prediction only minorly increased.

Non-canonical translation can initiate at non-CDS ATG and other codons. Thus, I investigated the start codon proportions during mitotic-arrest. Interestingly, predicted non-canonical uORF/uoORF (nuORFs) from U-2 OS initiated majorly from non-ATG codons (~80 %), whereas this number decreased during mitotic-arrest (~70 %). This suggests a shift to canonical ATG translation initiation sites in the 5'UTR (**Figure *9* d**). Next, I identified the nuORFs-containing transcripts to biologically characterize the origin of non-canonical translation (**Figure *9* e**). My comparison of these transcripts revealed heterogeneity between mitotic-arrest inducing agents, however, 411 transcripts were shared in between all treatments. I performed a functional enrichment analysis for these transcripts, which highlighted biological processes such as cytoplasmic translation and protein import (**Figure *9* f**). The same transcripts also belong to cellular components such as cytoskeleton and focal adhesions (**Figure *17* b**). These findings support further evidence for non-canonical translation in the 5'UTR.

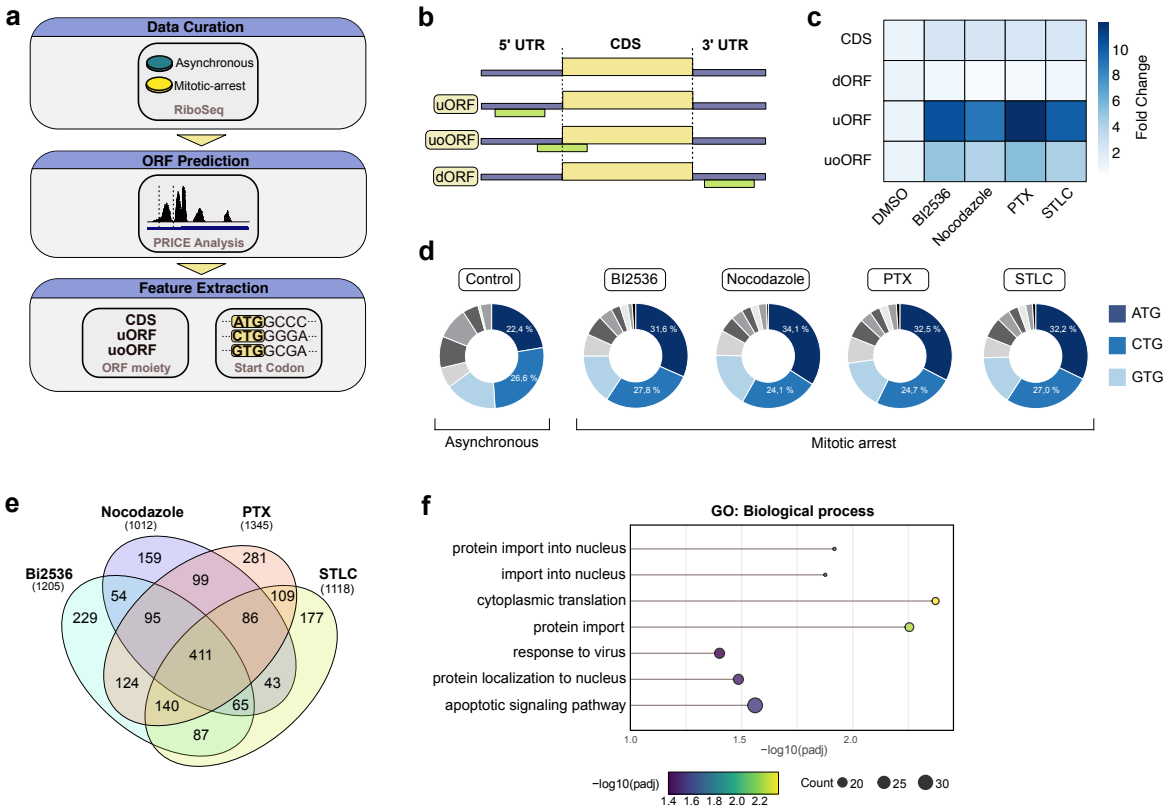

Figure 9: Mitotic-arrest enhances translation of upstream and upstream-overlapping open-reading frames (uORF/uoORF). (a) Experiment scheme of ORF prediction. RiboSeq data from DMSO-treated, asynchronous cells and cells during mitotic-arrest was used to computationally determine ORFs (PRICE, Probabilistic inference of codon activities by an expectation-maximization algorithm) and extract ORF moieties and start codons. **(b)** Scheme of predicted ORF moieties: uORF = upstream ORF, uoORF = upstream-overlapping, dORF = downstream ORF. U-2 OS cells were treated for 16 hours with indicated mitosis-arresting agents (0.1 µM BI2536, 0.5 µM Nocodazole (Noco), 1 µM Paclitaxel (PTX), 1 µM S-Trityl-L-cysteine (STLC)) or DMSO. **(c)** Fold-change heatmap of ORF moiety prediction relative to DMSO condition. **(d)** Pie charts of uORF start codon prediction highlighting the three most common codons ATG, CTG, GTG. **(e)** Venn diagram showing the intersection of identified uORF-containing genes predicted by PRICE. **(f)** Lollipop plot of over-representation analysis (Gene Ontology: Biological process) covering the intersected 411 common uORF-containing genes.

2.2.2 Increased translational levels of nuORF features during mitotic-arrest

After I identified increased non-canonical initiation sites in the 5'UTR, I determined if these features would exhibit elevated translation rates. For this, I calculated translational efficiency (TE) for predicted nuORF sites in U-2 OS cells treated with Nocodazole (**Figure *10* a**). I normalized nuORF-mapping RiboSeq reads to RNAseq reads of the same genomic coordinates.

Strikingly, my analysis resulted in increased TE rates for the majority of nuORF features in these samples (**Figure *10* b**). Examples with increased TE in nuORF elements such as

pyruvate kinase (PKM), mitochondrial ribosomal protein L51 (MRPL51) and caveolea associated protein 1 (CAVIN1) are implicated in critical cellular processes. These include energy metabolism, mitochondrial function and oxidative stress response. These examples exhibit strong RiboSeq read density increases in predicted uORF locations during mitotic-arrest, which I did not observe with RNAseq read densities (**Figure *10* c**). I found upregulated nuORF features (Log2 fold change ≥ 1.5, adjusted *p*-val < 0.05) to be associated with biological processes of translation and mitotic spindle assembly, and cellular compartments of the nucleus, cytosolic large ribosomal subunit and cytoskeleton. This effectively recapitulated the gene set enrichment of mitotic arrest as I observed before (**Figure *9* f**; **Supplement: Figure *17* b**). Notably, I determined that Harringtonine treatment demonstrated increased RPF densities at these predicted uORF start sites (**Supplement: Figure *17* d**).

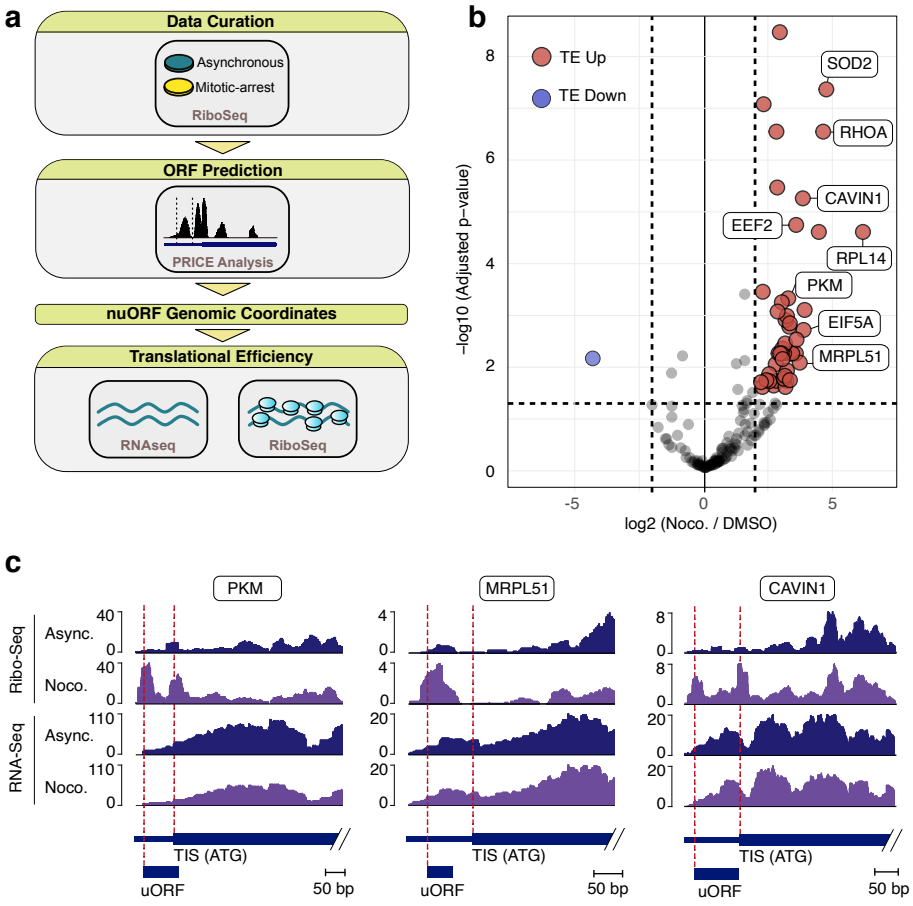

Figure 10: Differential upregulation of nuORF translation during mitotic-arrest. (a) Experimen scheme of translational efficiency calculation. U-2 OS cells were treated with 0.5 µM Nocodazole (Noco) or DMSO for 16 hours. RiboSeq data was used to computationally determine nuORF features (PRICE, Probabilistic inference of codon activities by an expectation-maximization algorithm). nuORF genomic coordinates were extracted and compiled into a reference file suitable for read counting. RNAseq and RiboSeq reads mapping nuORF coordinates were used to calculate translational efficiency (TE). **(b)** TE volcano plot of differentially translated nuORF elements with highlighted gene names. **(c)** Exemplary illustrations of nuORF element genome coverage from asynchronous cells (DMSO-treated) and during mitotic-arrest (Noco). PKM, Pyruvate kinase; MRPL51, Mitochondrial ribosomal

protein L51; CAVIN1, Caveolae associated protein 1. TIS (ATG), canonical CDS translation initiation site using an ATG codon.

2.2.3 Construction of a comprehensive nuORF feature database

Next, to establish a compendium of uORF and uoORF elements, I curated RiboSeq data from various cell lines during asynchronous population or mitotic-arrest.

Following ORF prediction, I extracted nuORF genomic coordinates to create a database (nuORFdB) for the respective DNA and peptide sequences (**Figure *11* a**). This gave rise to 10222 sequences, with ~77.5 % and ~22.5 % originating from uORFs or uoORFs, respectively (**Figure *11* b**). I determined that these sequences have a bias towards increased Guanine/Cytosine (GC) content (**Figure *11* c**). I computationally transformed these DNA sequences to AA sequences, which gave rise to the AA content. I found increased alanine (Ala, A) and arginine (Arg, R) frequencies of ~ 11 % each (**Figure *11* e**), in comparison to AA content of the canonical proteome (Ala: 9 %; Arg: 5.5 %)[90], which posed a Arg frequency doubling. Furthermore, I calculated peptide length of these 10222 peptides, which was majorly covered in the range of 0-40 AA, however, maximum lengths of protein products extended further than 200 AA.

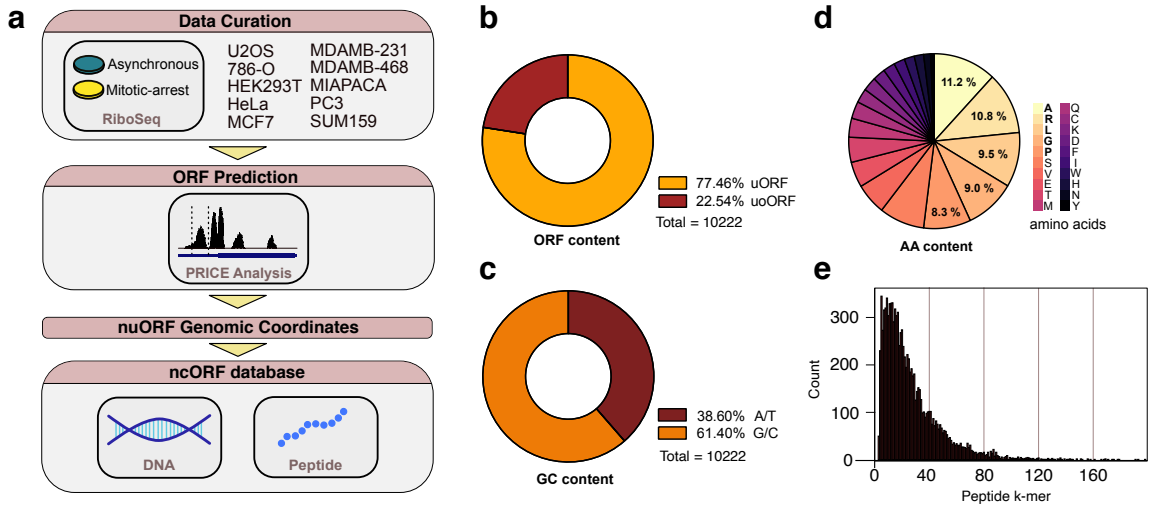

Figure 11: Construction and Characterization of the nuORF database. (a) Experiment scheme for nuORF database construction. First, RiboSeq data from asynchronous cells or during mitotic-arrest was curated. Next, nuORF features were computationally predicted using PRICE (Probabilistic inference of codon activities by an expectation-maximization algorithm). The resulting genomic coordinates were extracted as DNA nucleotide sequences, and peptide sequences. **(b)** Pie chart showing the uORF and uoORF origin quantification of 10,222 sequences. **(c)** Pie chart showing the A/T and G/C content in percent of the predicted 10,222 nuORF DNA sequences. **(d)** Pie chart showing the amino acid composition of the predicted nuORF-derived peptides. **(e)** Bar chart showing the length distribution of the predicted nuORF-derived peptides.

Taking together, I found that mitotic-arrest is associated with increased uORF and uoORF prediction rates and elevated translation rates in hundreds of these sequences. In a comprehensive approach, I compiled DNA and subsequent peptide sequences of nuORF features from RiboSeq data of various cell lines to construct a nuORF database.

2.3 nuORF features give rise to HLA complex I presented peptides

Recently, ncORF-originating peptides were reported to be presented on the cellular plasma membrane in HLA-I complexes[68]. These peptides enrich the immunopeptidome landscape presented to immune cells. Thus, I investigated whether the predicted nuORF-derived peptides from my compiled nuORF database were presented on HLA-I complexes.
For this, I collaborated with Jonas P. Becker from the DKFZ division of Immunotherapy and Immunoprevention. Together, we performed LC-MS/MS-based immunopeptidomics. The analysis required my nuORF database and the annotated human proteome as reference files (**Figure *12* a**). For these experiments, I treated U-2 OS cells and the triple-negative breast cancer (TNBC) cell line SUM-159PT with DMSO or Paclitaxel (PTX) in replicates of $5x10^7$. PTX is a common first-line of treatment regimen in TNBC (**1.1.3.2 Mitotic-arrest: Applications and** Limitations; **p.11**), thus, providing stronger clinical relevance for future applications.

2.3.1 nuORF-derived HLA-presented peptides in U-2 OS cells

The collaborative immunopeptidomics analysis revealed an HLA class I repertoire of 12,904 unique annotated proteome-derived peptides. Notably, we found that 127 unique peptides were derived from nuORFs sequences, representing about 1 % of the HLA-I immunopeptidome in U-2 OS during mitotic-arrest (**Figure *12* b**). HLA allotype binding prediction revealed that 90 % of proteome-derived peptides and 91 % of nuORF-derived peptides bind U-2 OS-specific HLA allotypes. Also, both peptide types exhibited normal length distribution for HLA class I peptides, with a majority of 9-mers (**Figure *12* b,c,e**). Peptide sequence clustering of my samples for proteome-derived and nuORF-derived epitopes revealed similar motifs with dominant aspartic acid moieties in the 2nd position and aromatic AAs in the 9th position (**Figure *12* d**) for the U-2 OS allele HLA-B*44:02. Importantly, we determined that nuORF-derived peptide retention time (RT) showed a strong correlation with the respective observed RT, which was comparable with proteome-derived peptides (**Figure *12* f**).

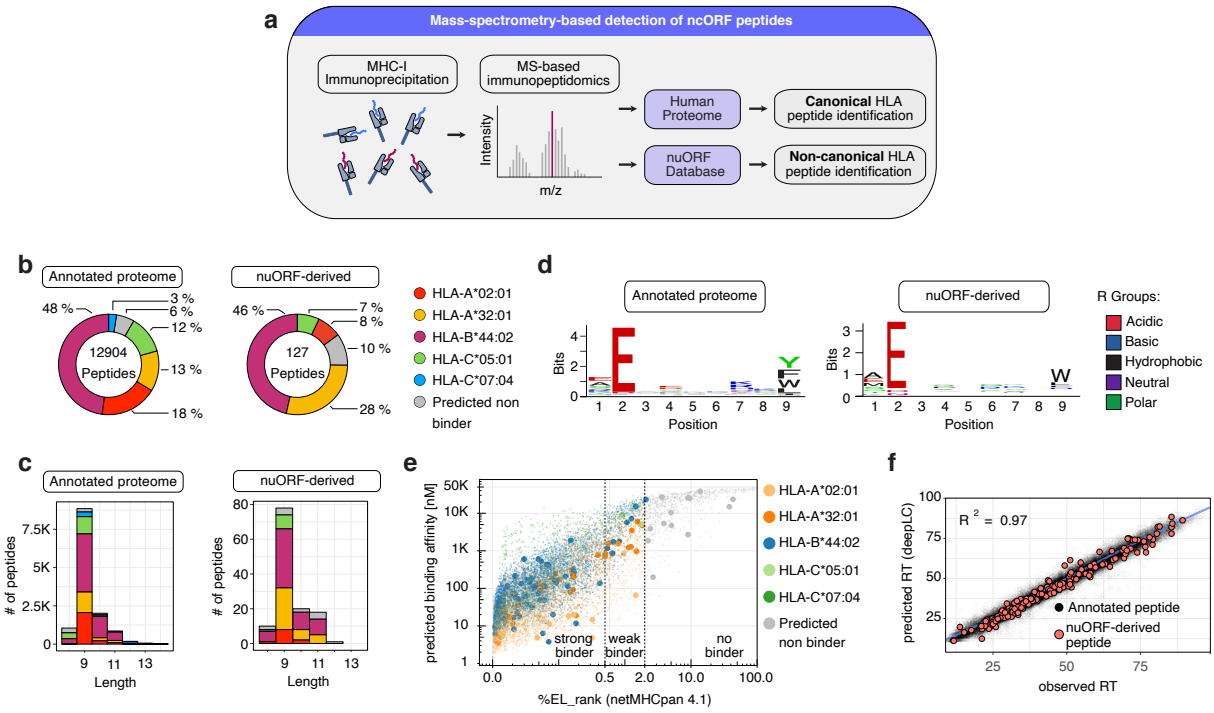

Figure 12: nuORF-derived HLA-presented peptides reveal comparable characteristics to annotated peptides in U-2 OS cells. (a) Experiment scheme of Immunopeptidomics. Cells are lysed and HLA-I complexes are immunoprecipitated and bound peptides are detected with LC-MS/MS. Resulting peptide spectra are scanned against the nuORF database and annotated proteome to identify non-canonical nuORF-derived peptides and canonical peptides. **(b)** HLA allele binding prediction of identified peptide sequences for the annotated proteome and nuORF-derived peptides. **(c)** Peptide ligand length distribution with corresponding predicted HLA allele proportion. **(d)** Sequence motif plots for unique peptides from the annotated proteome (6,624) and nuORF-derived peptides (58) for the HLA-B*44:02 allele. **(e)** Predicted binding affinity plotted against percentage of eluted ligand (EL) peptides predicted by NetMHCpan-4.1. Calculated for peptides derived from the annotated proteome (small spots) and nuORF-derived peptides (big spots) in asynchronous and mitotically arrested U-2 OS. Peptide categorization: Strong binder (%EL rank 0 – 0.5), weak binder (%EL rank 0.5 – 2.0), non-binder (%EL rank 2.0 – 100). **(f)** Observed retention time (RT) plotted against predicted RT of proteome-annotated (black) and nuORF-derived peptides (red). R^2, coefficient of determination.

2.3.2 nuORF-derived HLA-presented peptides in SUM-159PT cells

Our immunopeptidomics analysis of SUM-159PT cells revealed an HLA class I repertoire of 25655 unique proteome-derived peptides. Strikingly, 166 unique nuORF-derived peptides were detected, representing about 0.5 % of the SUM-159PT HLA class I immunopeptidome (**Figure *14* a**).

HLA allotype binding prediction revealed that 92 % of proteome-derived peptides and 82 % of nuORF-derived peptides bind SUM-159PT-specific HLA allotypes. Again, both peptide types exhibited normal length distribution for HLA class I peptides, with a majority of 9-mers (**Figure *14* b**). Peptide sequence clustering of my samples for proteome-derived and nuORF-derived

epitopes revealed similar motifs with dominant glutamine moieties in the 2nd position and aromatic AAs in the 9th position (**Figure 13 c**) for the allele HLA-B*15:01. Also, nuORF-derived peptides are predicted majorly with strong and weak binding affinities (**Figure 13 d).** We determined that ligands exhibited retention times (RT) with strong correlation to the respective observed RT, comparable to proteome-derived peptides (**Figure 13 e**).

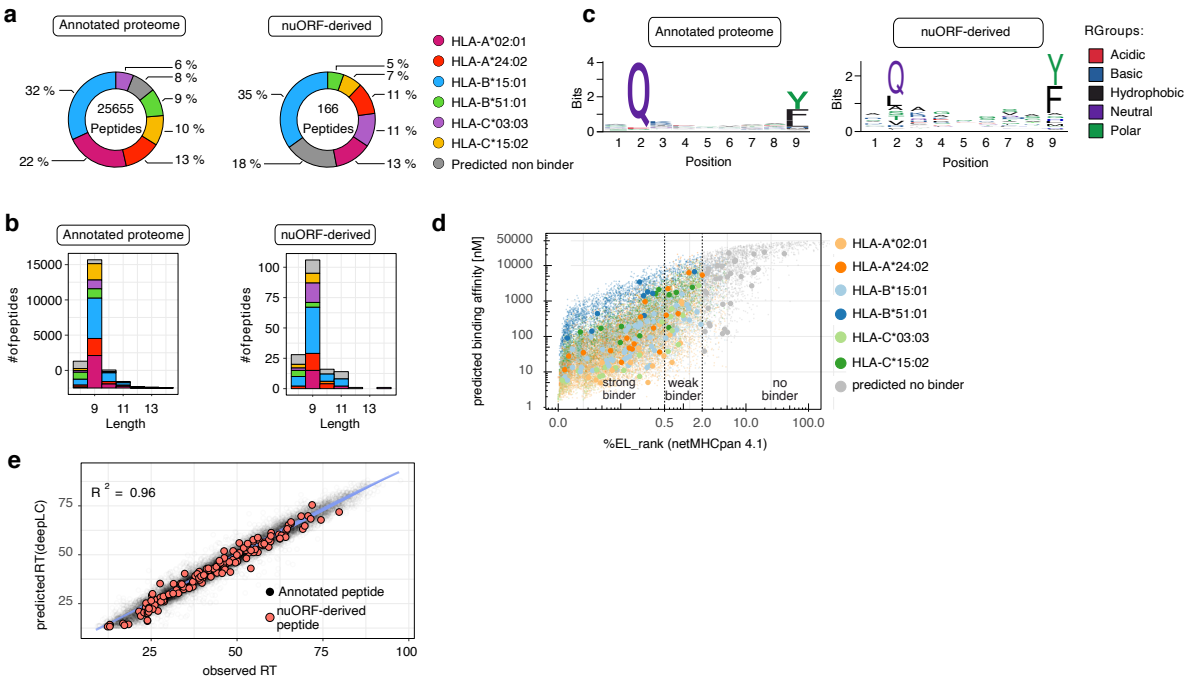

Figure 13: nuORF-derived HLA-presented peptides reveal comparable characteristics to annotated peptides in SUM-159PT cells. (a) HLA allele binding prediction of identified peptide sequences for the annotated proteome and nuORF-derived peptides in SUM-159PT cells. **(b)** Peptide ligand length distribution with corresponding predicted HLA allele proportion. **(c)** Sequence motif plots for unique peptides from the annotated proteome (4,402) and nuORF-derived peptides (59) in the SUM-159PT allele HLA-B*15:01. **(d)** Predicted binding affinity plotted against percentage of eluted ligand (EL) peptides predicted by NetMHCpan-4.1. Calculated for peptides derived from the annotated proteome (small spots) and nuORF-derived peptides (big spots) in asynchronous and mitotically arrested SUM-159PT cells. Peptide categorization: Strong binder (%EL rank 0 – 0.5), weak binder (%EL rank 0.5 – 2.0), non-binder (%EL rank 2.0 – 100). **(e)** Observed retention time (RT) plotted against predicted RT of proteome-annotated (black) and nuORF-derived peptides (red). R^2, coefficient of determination.

2.4 Paclitaxel induces differential nuORF-derived peptide presentation and cancer killing

The identification of nuORF-derived peptides in asynchronous and mitotically arrested cells provides strong novelty for the understanding of peptide sources for cell surface display. This

leads to the question of abundance quantification and how treatment-induced peptides can be exploited for improved cancer therapies.

2.4.1 Differential HLA class I nuORF-derived peptide presentation upon PTX-induced mitotic-arrest

The impact of HLA class I peptide presentation during mitotic-arrest remained uncharacterized. Thus, we performed label-free quantification of HLA-presented peptides following PTX treatment with U-2 OS and SUM-159PT cells. I detected moderately enriched nuORF-derived peptides for both cell lines, with 13 and 25 peptides for U-2 OS and SUM-159PT, respectively (Log2 fold change ≥ 0.5; adjusted *p*-val ≤ 0.05) (**Figure 14 a,b; Supplement: Figure 18 a,b**). Importantly, I observed that these elevated candidates showed increased translation during mitotic-arrest, but mRNA transcript levels did not change (**Figure 14 c**). Furthermore, I determined that Harringtonine treatment increased RPF densities at predicted nuORF initiation start sites (**Figure 14 d**).

One of the detected peptides originates from the EIF4G2 gene, which is implicated for the translation of mRNAs for cell[91]. An additional peptide originates from the HMGA1 gene. The corresponding protein modulates chromatin structure to promote gene expression for G2/M transition[92,93] and was shown to regulate microtubule-destabilization, which supports TNBC motility with reduced PTX sensitivity[93]. To infer the biological context, I performed a functional over-representation analysis of genes expressing increasingly presented nuORF-derived peptides in U-2 OS cells. My analysis revealed an association with cellular components, such as the cytoskeleton and adherens junctions (**Figure 18 c**). Strikingly, I did not find one of the identified peptide sequences in the HLA Ligand Atlas[58], which is a public-accessible compendium of benign human HLA peptides across organs and HLA alleles. Furthermore, I aligned nuPeptide candidates against annotated RefSeq protein sequences using NCBI Protein BLAST, which did not reveal any match. This included Leu/Ile permutations, which cannot be distinguished by mass spectrometry. This provides evidence about the novelty of PTX treatment-associated HLA class I nuORF peptides.

Taking together, I demonstrated that U-2 OS and SUM-159PT cells presented nuORF-derived peptides. I was able to detect that the properties of these new epitopes align with proteome-derived peptides and are moderately-enriched in cells during mitotic-arrest. Thus, these antigenic peptides might be new promising immunotherapy targets.

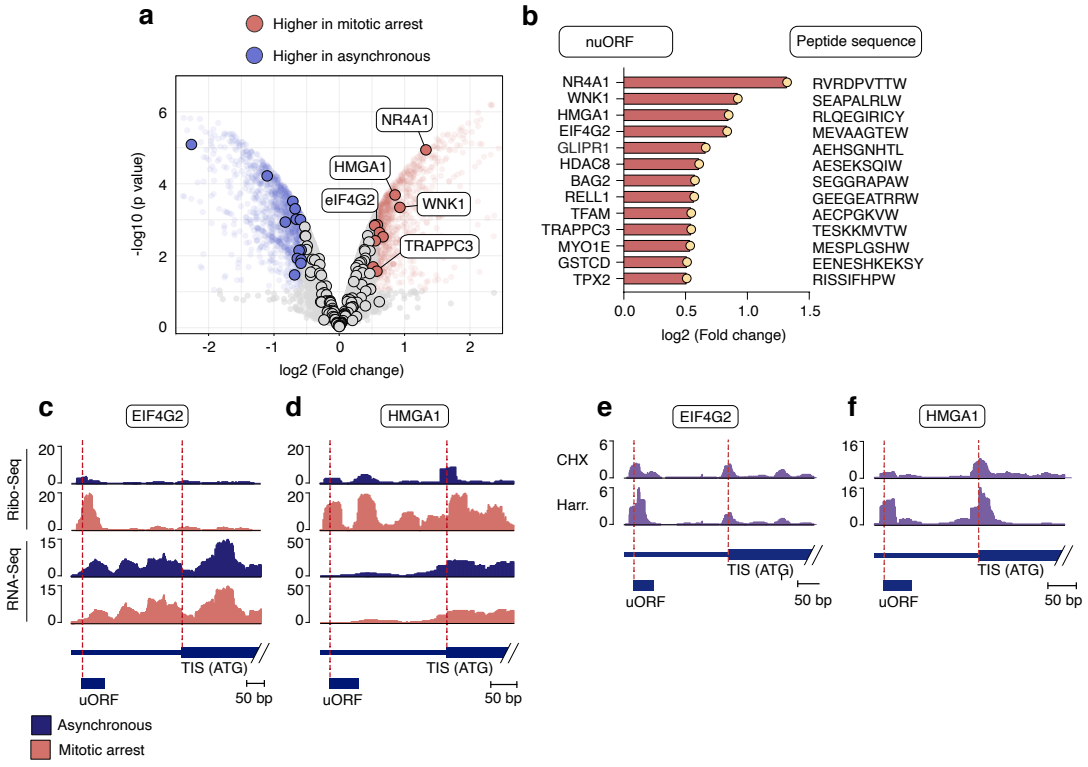

Figure 14: Quantitative analysis of HLA class I presented nuORF-derived peptides. (a) Volcano plot showing the label-free quantification of HLA class I immunopeptidome, highlighting nuORF-derived peptides (big spots) and peptides from the annotated proteome (small spots). Red peptides: Log2 fold change ≥ 0.5 and adjusted *p*-value < 0.05. Blue peptides: Log2 fold change ≤ 0.5 and adjusted *p*-value < 0.05. *P*-values were calculated using a Bayes-moderated *t*-test with two-sided *p*-values. **(b)** nuORF-containing gene with Log2 fold change peptide abundance and corresponding peptide sequence. nuORF element genome coverage plots from asynchronous cells (DMSO-treated) and/or during mitotic-arrest (Noco) showing 5'UTRs and CDS partials **(c-f)**. RiboSeq and RNAseq genome coverage for EIF4G2 **(c)** and HMGA1 **(d)**. RiboSeq genome coverage from Harringtonine (Harr) run-off assays in mitotically arrested U-2 OS cells for EIF4G2 **(e)** and HMGA1 **(f)**. Harringtonine and/ or CHX treatment was performed for 10 min. TIS (ATG), canonical CDS translation initiation site using an ATG codon.

2.4.2 PTX-mediated increased uORF-derived peptide display promotes targeted cancer killing

The presentation of nuORF-derived HLA class I peptides surface upon Paclitaxel treatment holds a promising opportunity for targeted immunotherapy. Thus, chemotherapy-induced display of nuORF-derived peptides in cancer cells can alter the immunopeptidome landscape, which might offer unique immunogenic signatures.

To explore this, F. Loayza-Puch and me generated luciferase reporter constructs, which contain 5'UTRs from transcripts that gave rise to increased peptide presentation during mitotic-arrest. To assure equal peptide stability between conditions, peptide sequences identified by immunopeptidomics were replaced by the SIINFEKL peptide sequence. This

ligand originates from chicken ovalbumin (OVA) and can be presented on murine H-2Kb alleles. The SIINFEKL:H-2Kb complex is recognized by transgenic $CD8^+$ OT-I T cells, as well as the commercial dye-conjugated antibody clone 25-D1.16[94], suitable for flow cytometry applications.

For this validation 5'UTRs of EIF4G2 and TPX2 were chosen. Peptides from these respective uORFs were also detected in an initial IPomics experiment with the group of Reuven Agami at the NKI Amsterdam, Netherlands. (**Figure *14***). Therefore, cross-institutional finding of the same uORF origins supported the decision to validate 5'UTR translation with these sequences. The following experiments were performed and analysed by Z. Tang and F. Loayza-Puch.

To assess the extent of PTX-induced SIINFEKL presentation, Z. Tang and F. Loayza-Puch transfected uORF-SIINFEKL reporters into the murine cancer cells line TC1. Z. Tang determined SIINFEKL presentation by flow cytometry and F. Loayza-Puch analysed the data. Asynchronous cell population did not induce strong 25-D1.16 recognition. In contrast, mitotic-arrest led to a strong increase of SIINFEKL presentation, independent of the agent's mode-of-action (**Supplement: Figure 18 d,e**). Notably, reporter mRNA levels were not altered during mitotic-arrest (**Supplement: Figure 18 f**). Next, we mutated the peptide start sites from ATG to ATA to investigate the start codon importance in these 5'UTRs. Flow cytometry measurement by Z. Tang and analysis by F. Loayza-Puch revealed a loss-of-signal with reduced SIINFEKL presentation in mutated start codon reporters, underlining the impact of active uORF translation on effective antigen presentation (**Supplement: Figure 18 g**).

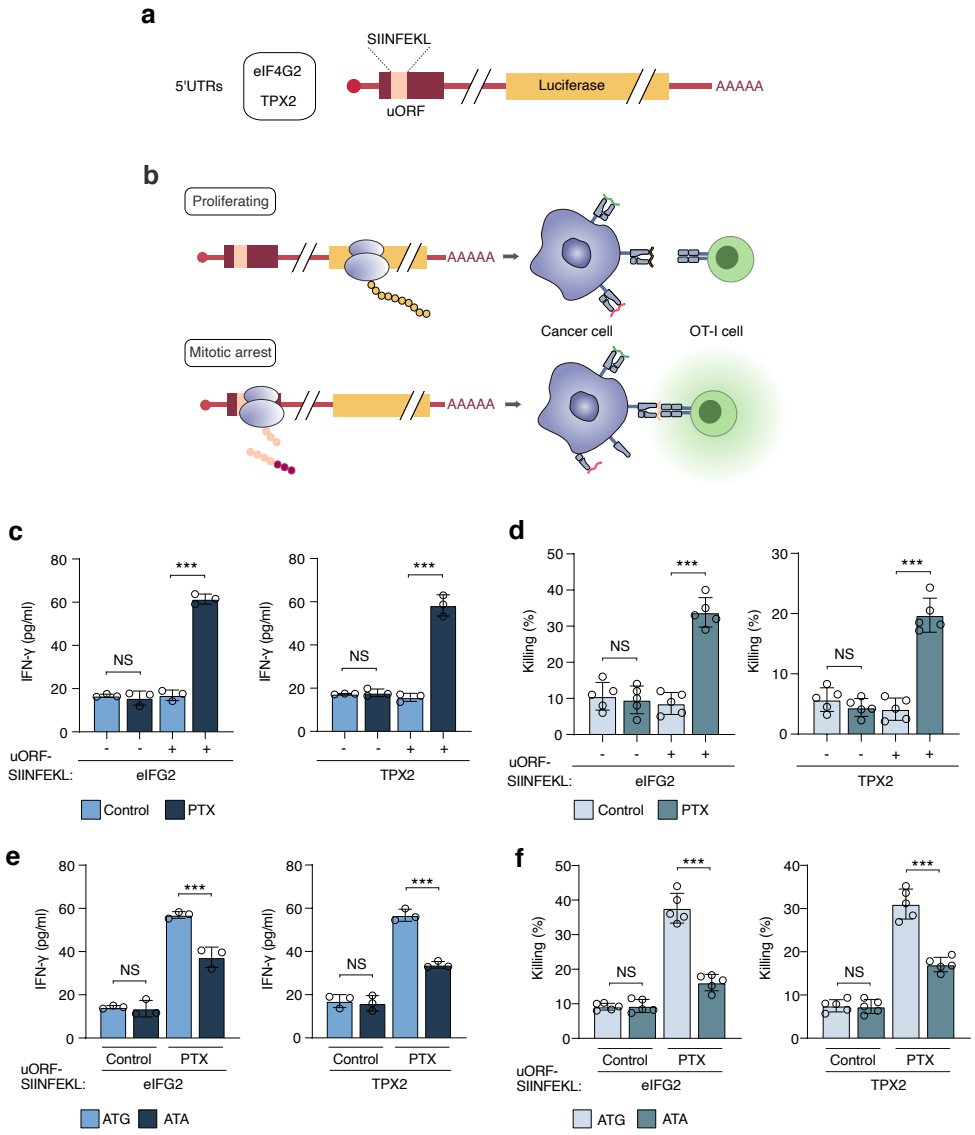

Figure 15: Increased uORF-derived peptide presentation promotes targeted immune response against cancer cells in mitotic-arrest. (a) Model of uORF reporter system. Two 5'UTRs were cloned into the pGL3-luciferase vector, in which the uORF-derived peptide sequence was replaced by the SIINFEKL sequence. **(b)** Functional model of the uORF reporter system in combination with SIINFEKL-specific OT-1 T cells. During normal proliferation, ribosomes do not translate the uORF. However, during PTX-induced mitotic-arrest, ribosomes initiate at the uORF sequence and translate the SIINFEKL peptide. Upon SIINFEKL-MHCI binding, SIINFEKL-specific OT-1 T cells recognize cancer cells and initiate cancer killing. IFN-γ secretion **(c)** and T cell killing **(d)** in murine TC1 cells, transfected with the uORF-SIINFEKL reporters and arrested in mitosis for 16 hours before T cell addition. IFN-γ secretion **(e)** and T cell killing **(f)** in murine TC1 cells, transfected with the start codon-mutated uORF-SIINFEKL reporters and arrested in mitosis for 16 hours before T cell addition. Control cells were treated with equal volumes of DMSO for the same time. Data shown represents the mean ± SD from biological replicates: n = 3 for IFN-γ secretion, n = 5 for T cell killing. Statistical analysis was carried out with a two-tailed unpaired t-test. NS, non-significant. *** *p*-value < 0.001.

Next, I hypothesized if uORF-derived peptides might elicit an enhanced T cell-mediated cytotoxicity, which was performed and analysed by Rossella Del Pizzo and Fabricio Loayza-Puch. For this, R. Del Pizzo activated OT-I $CD8^+$ T cells with αCD3, αCD28 and IL-12 *ex vivo* for 72 hours and co-cultured them *in vitro* with asynchronous or mitotically arrested TC1 cells. The culture medium was harvested for ELISA, while the culture plate was subjected to crystal violet staining and analysis by F. Loayza-Puch.
Strikingly, T cells only secreted higher levels of IFNγ upon co-culture with reporter-expressing, mitotically arrested TC1 cells (**Figure 15 b,c**). Furthermore, we observed enhanced cytotoxic activity only in reporter-expressing, mitotically arrested TC1 cells (**Figure 15 d**). Neither mitotic arrest nor transfection alone induced elevated IFNγ secretion or increased cytotoxicity. Strikingly, start codon mutation of uORF-SIINFEKL reporters exhibited a loss-of-function in terms of IFNγ secretion and antigen-specific $CD8^+$ T cell-mediated cytotoxicity (**Figure 15 e,f).** Taking together, I determined nuORF-derived peptides that are HLA class I presented in an increased manner. uORF-SIINFEKL reporter constructs were created to assess the extent of PTX-induced peptide presentation. In a collective effort we detected that PTX-mediated mitotic-arrest induced increased SIINFEKL presentation, a reporter-specific increase of IFNγ secretion in $CD8^+$ T cells and enhanced cancer killing capacities against cancer cells.

In summary, I demonstrated that the general process of mitotic-arrest induces ribosomes to localize and initiate at the 5'UTR of mRNA transcripts. Here, ribosomes translate uORF and uoORF sequences into peptides or proteins of variable size in a differential manner. My data supports that nuORF-derived peptides are presented on HLA complex I structures in a differential fashion. Furthermore, I was able to show that uORF sequences can give rise to targeted $CD8^+$ T cell cytokine release and cancer killing.

3 Discussion

The aim of this doctoral thesis was to unravel the translational response during mitotic-arrest. In summary, I demonstrated that the general process of mitotic-arrest causes ribosomes to localize and initiate at the 5'UTR of mRNA transcripts. Here, ribosomes translate uORF and uoORF sequences into peptides or proteins of variable size. My data supports that nuORF-derived peptides are presented on HLA complex I structures in a differential manner. Furthermore, I was able to show that uORF sequences can give rise to targeted $CD8^+$ T cell cytokine release and cancer killing.

3.1 Active mRNA translation in the 5'UTR during mitotic-arrest

Initially, I uncovered that cellular responses to mitotic-arrest are similar in U-2 OS cells, treated with drugs that possess different modes-of-action. I observed that all drugs efficiently enriched U-2 OS cells in mitotic-arrest, as seen by H3 Ser10 phosphorylation[86] and PI staining for DNA content (**Figure *6***). While I dissected translational signatures, it was evident that each drug induced the translation of a specific transcript set, which differed from the other agents (**Supplement: Figure 16 a**). Still, I observed that high proportions of translational activation were shared between drug treatments, which allowed for high Pearson correlation coefficients between mitotic-arrest conditions in contrast to asynchronous cells treated with DMSO (**Figure 6 c,d**). Thus, H3 phosphorylation, DNA content and translational signatures support the usage of any tested compound to explore translational responses during mitotic-arrest. Furthermore, all of these agents induce changes in ribosome density, across the CDS and in the 5'UTR of cancer cells (**Figure 7 a-f, Supplement: Figure 16 b-e**).

Notably, I tested if the ISR (**Supplement: Figure *16* f,g,h**) induced increased ribosome density in the 5'UTR by the inhibition of the eIF2α phosphatases with Salubrinal. This treatment induced the downstream transcription of CHOP and ASNS (**Figure *16* f**) but I did not find increased ribosome density levels in the 5'UTR. This suggests active ISR upon Salubrinal treatment, which fails to mimic the observed translational phenotype of aberrant ribosome location. In return, other signalling cascades need to be responsible for this finding.

Cap-dependent translation is highly regulated by mTOR and CDK1, which modulate 4EBP1 phosphorylation. Unphosphorylated 4EBP1 binds eIF4E, also known as the cap-binding protein, that is, translation initiation is hindered. In contrast, phosphorylated 4EBP1 does not bind eIF4E, which allows cap-dependent initiation. Hence, mTOR and CDK1 are considered regulators of translation (**1.2.2.1 4EBP**). I found that mTORC1/2 inhibition with Torin1 did not change ribosome density in the 5'UTR and did not modulate 4EBP1 phosphorylation levels. This data suggests a mTOR-independent modulation of translation. Thus, 4EBP1

phosphorylation is very likely mediated by CDK1. In fact, CDK1-dependent 4EBP1 modulation during mitotic-arrest, independent of mTOR, was already reported[39]. Furthermore, CDK1 is annotated to phosphorylate ribosomal proteins and translation initiation factors, which aids efficient protein biosynthesis during mitotic progression[39,40,95]. Notably, presence of CDK1 increases after about 14 h incubation with Nocodazole[96] comparable to the incubation time of 16 hours used to obtain my data in this thesis. To elucidate the role of CDK1-dependent 4EBP1 phosphorylation on translation initiation in this phenotype, I suggest further experiments with phospho-silent mutants of 4EBP1. Genomic ablation of 4EBP1 with overexpression of a phospho-silent 4EBP1 mutant protein would elucidate the impact on translation initiation in the 5'UTR.

CDK1 was also reported to phosphorylate the DENR:MCTS1 complex[97] (density regulated re-initiation and release factor: malignant T-cell-amplified sequence 1). Interestingly, this complex promotes post-termination recycling of 40 S ribosomes and translation re-initiation to the CDS start sites after uORF sequences with a strong Kozak context[97,98]. Thus, DENR could induce ribosome recycling after uORF and uoORF initiation and induce CDS re-initiation during mitotic-arrest. This would explain 5'UTR ribosome initiation in close proximity to the CDS start, (upstream and downstream) as found by Harringtonine run-assays (**Figure *8* b,c**). Phospho-site mutation of Ser73 to Ala73 would channel a DENR loss-of-function with subsequent proteasomal digestion, where Asp26 to Glu26 would prevent DENR degradation[97]. The functional combination of genetic DENR KO with over-expression of S73A and D26E mutations and RiboSeq would aid to discriminate if the aberrant translational phenotype is induced via the CDK1-DENR axis.

On the contrary, other reported mechanisms of non-canonical translation initiation highlight functional lack of translation elongation factors that induce ribosome queues in the CDS, leading to initiation in the 5'UTR[64]. However, this study pointed out increased translation initiation at sub-optimal codons, that is, non-ATG codons. This is in contrast to my finding of increased ATG start codon utilization in predicted uORF sequences. Thus, the observed translation phenotype in my data might only be explained by multiple factors. Future experiments aiming to understand the mechanism of ribosome initiation in the 5'UTR during mitotic-arrest should first include pharmacological inhibition of CDK1 in combination with immunoblotting for 4EBP1 phosphorylation state, as well as translational efficiency to study quantitative translational changes. However, CDK1 also promotes mitotic progression, hypoxia signalling and apoptosis[99–101]. Because of the broad signalling cascades mediated by CDK1, exploration for other downstream targets would be advantageous, such as 4EBP1 and DENR.

These experiments would demonstrate the possible impact of CDK1 on translation initiation in the 5'UTR.

However, if non-canonical translation initiation is mediated in a cell cycle-dependent manner, it needs to be unravelled if ribosomes localize at the same uTISs elements throughout other cell cycle stages in a time-dependent manner. The exploration will ensure the identification of uTISs features exclusive to mitotic-arrest. Therefore, I suggest to perform further RiboSeq experiments of cells arrested in G1, G2 and M phase using the CDK1 inhibitor RO-3306[102]. This inhibitor prevents the progression from G2 to M, which would serve as G2-enriched cell population. Upon washing out and release from this drug, cells will proceed through mitosis within 45 min[102], which would serve as mitotic cell population. Harvesting and replating mitotic cells will allow cell cycle progression to harvest an G1 fraction after about three hours. Additional PI staining will ensure the cell population enrichment in the respective phase. Further approaches would be cell synchronization followed by release and time-dependent cell harvest using thymidine, as reported for G1 and S phase enrichment[103]. This approach would ensure the normal transition through cell cycle phases compared to arrest by chemical inhibition, which in return would reduce biological variation.

3.2 Elevated uORF/ uoORF translation in cancer cells during mitotic-arrest

The presence of increased RPF density in the 5'UTR prompted me to investigate non-canonical translation. In an exploratory approach I predicted ORF features in RiboSeq data from cancer cells during mitotic-arrest. I determined increased rates of predicted ORFs for upstream ORFs (uORFs) and upstream-overlapping ORFs (uoORFs) in mitotically arrested cells (**Figure 9 c; Supplement: Figure 17 a**). Particularly important is the upregulation in U-2 OS treated with different mitotic-arrest inducing agents. I found a consistent increase in comparison to the DMSO-treated asynchronous cells, regardless of the drug`s mode-of-action. This prompted towards a general cellular response during mitotic-arrest in U-2 OS, PC3 and MDA-MB-231 cells. This finding supports the hypothesis of actively-initiating ribosomes in the 5'UTR.

Notably, I also determined slightly increased prediction rates for CDS ORFs. However, to a much smaller extent when compared to uORF and uoORF features. This CDS ORF increase can be explained by two factors. First, by read amount differences, that is, less reads in the mitotic-arrest samples in comparison to the DMSO condition (DMSO: 12,788,200; BI2536: 8,364,051; Noco: 6,224,505; PTX: 7,850,182; STLC: 6,172,391). This means overall more reads that are associated with CDS ORF features in the DMSO condition. However, lower total read amounts in mitotic-arrest samples support the significance of the detection of

nuORFs in these samples. The second factor would be the mitotic-arrest phenotype, which induced this raise by increased RPF densities at the 5'end of CDSs (**Figure *7* a,c,e**). Elevated read amounts will be detected as CDS ORF in a stronger manner. Thus, the translational phenotype of upstream accumulating ribosomes might be solely responsible for the increased detection rates of CDS ORFs.

ORF prediction is a computational field with strong bioinformatic resources. Most of these algorithms rely on high-quality RiboSeq data to identify ORFs with the characteristic 3-nt periodicity of translating ribosomes. Using these tools, researchers uncovered thousands of human non-canonical ORFs, which shapes our understanding of these transcriptomic features. However, a recent comparison of ORF-predicting tools pointed out differences in the prediction rates of small-upstream ORFs and low cross-overlap[104]. Tool-overlapping features exhibited higher translation levels and amounts of in-frame reads. Unfortunately, this study did not include PRICE although this tool was published 5 years before the report on tool comparison. In contrast, benchmarking of prediction rates per ORF feature revealed the advantage of PRICE in the discovery of high-confidence ORFs in the 5'UTR[89]. However, future annotations should combine the intersected output of different tools to ensure high ORF feature coverage. Still, I conclude that the usage of PRICE for the definition of uORF and uoORF features was advantageous against the use of other resources.

Start codon selection is an important process throughout CDS mRNA translation. However, non-canonical translation initiation underlies different start codon mechanisms. CDS ORFs begin with optimal ATG codons, whereas ribosomes initiate at ncORFs with CTG or ATA codons[64]. Interestingly, my ORF prediction data suggests increased ATG-dependent translation during mitotic-arrest (**Figure *9* d**), because ATG codon percentages increase, whereas CTG and GTG start codon percentages remain stable. In contrast, the quantity of other non-canonical codon percentages decreased (TTG, AAG, ACG, AGG, ATC). This supports a model in which sub-optimal start codons are replaced by optimal start codons, which encourages the hypothesis of active translation-initiation in the 5'UTR.

Despite the highly correlated translational response upon mitotic-arrest with different agents (**Figure *6* d**), my prediction of uORF-containing genes resulted in a treatment-specific scattering. The overlap of all treatments is 411 shared genes, which equals 27 % of the input amount considering approximately 1100 uORF-containing genes per condition (**Figure *9* e**). However, this number raises significantly, when considering at least two mitotic-arrest induction approaches. The difference in uORF-containing genes most likely originates from the distinct mechanisms of mitotic-arrest induction. However, my over-representation analysis of the 411 shared genes revealed biological functions in nuclear transport and cytoplasmic translation (**Figure *9* f**). Notably, I found that several ribosomal proteins are part of the shared uORF-containing genes. All transcripts of cytosolic ribosomal proteins belong to the category

of TOPmotif mRNAs (5' terminal oligo-pyrimidine motif), which starts with an m^7G capped C nucleotide with 4-15 downstream pyrimidines[105]. Translation of TOPmotif mRNAs is proposed to be regulated by CDK1 through phosphorylation of LARP1 and S6K[106,107], which again underlines the potential regulatory function of CDK1 for ncORF translation during mitotic-arrest. The overlap with TOPmotif mRNAs is RPS2, RPS20, RPS24, RPL10, RPL14, RPL17, RPL26, RPL27, RPL39, RPL41 and NAP1L1. This raises the question why only these transcripts were found as overlap between the four mitotic-arrest mechanisms and not more TOPmotif mRNAs. Thus, the overlapping transcripts need to be further characterized in the face of other TOPmotif transcripts (as explained later)

The upregulated translation of nuORF sequences (**Figure *10* b)** shown by my translational efficiency (TE) analysis sets a quantitative perspective to 5'UTR translation initiation. An over-representation analysis of genes belonging to upregulated transcripts containing nuORFs revealed the GO biological terms translation and mitotic spindle organization and cellular compartments of the large cytosolic ribosome subunit. This underlines the prior findings of the uORF-containing gene overlap between the four inductors of mitotic-arrest (**Figure *9* e**).

However, as shown exemplary, other upregulated nuORF-containing genes are associated with energy metabolism, mitochondrial translation and oxidative stress responses.

Corresponding genome coverage data (**Figure *10* c**) strongly supports the global findings, as RNAseq read densities do not change upon treatment. However, I found RiboSeq read densities to be increased locally in uORF genomic coordinates, meaning that mRNA transcript levels do not change, but mRNA translation levels of uORF sequences are elevated upon mitotic-arrest. This suggests increased TE. Notably, I observed a general trend for upregulated TE in U-2 OS cells treated with Noco. However, this result was obtained from pseudo-replicates of matching RNAseq and RiboSeq data to infer general tendencies of uORF translational efficiency. Thus, it is crucial to further validate this finding in U-2 OS with at least three biological replicates of matching RNAseq and RiboSeq data to solidify reliable conclusions of this finding.

Further exploration of this finding should include the repetition of RiboSeq and RNAseq for samples during mitotic-arrest with multiple biological replicates ($n \geq 3$) and increased sequencing depth for RiboSeq. nuORF sequences only comprise a minor fraction of reads obtained from RiboSeq, thus, improving read coverage will help to expand and solidify the finding of increased TE rates. Furthermore, the current RNAseq data was obtained from poly-A enrichment using oligo-dT primers. While this approach allows the enrichment of mRNA, it also reduces upstream read coverage[108]. Therefore, this approach might hamper the differential analysis of 5'UTR reads by TE calculation. Thus, I would recommend to use an

rRNA-depleted total RNAseq approach for further experiments to enrich for 5'UTR read coverage.

Importantly, the current RiboSeq data was exclusively obtained from highly proliferative cancer cells, that harbour genetic driver mutations. Despite the importance for cancer research, the underlying principles of nuORF translation should also be studied in un-transformed or non-malignant cells. Functional uORFs are annotated to be translated during cellular stress responses, such as amino acid starvation[62]. Healthy, non-dividing cells should not be affected by paclitaxel as rapidly-dividing cancer cells are. Thus, to elucidate the functional role of the uORFs that I have found during mitotic-arrest, I suggest RiboSeq experiments of minorly transformed cells, such as RPE-1 cells. Also, human or murine primary cells, such as fibroblasts or keratinocytes, would be suitable model systems. Paclitaxel treatment in these cells would either determine new nuORF elements or features similar to mitotic-arrest. Equal elements would reduce the importance for therapeutic exploitation, because these nuORFs will give rise to peptides in cancer cells and healthy body cells, which is unfavourable. The remaining, mitotic-arrest exclusive features would need to be deeply characterized.

To expand our understanding of translational control of these upregulated elements, I suggest further classification based on multiple parameters: Kozak motifs, nuORF length, codon context, amino acid content, uORF distance to CDS, GC content and structure prediction.

Kozak sequences are eukaryotic motifs that include start codons and adjacent nucleotides[109]. These sequences regulate efficiency of translation initiation[110]. ATG codons provide a strong initiation context, however, adjacent sequence motifs can modulate the efficiency. Next, short nuORF sequences might provide a TE benefit over longer sequences, because start-to-end translation can be performed faster. On the contrary, fast translation of short nuORFs might be able to skew TE data towards false negative elements. TE calculation is based on unnormalized count tables for pre-defined elements, such as a full mRNA transcript or a uORF. These counts serve as input for differential expression analysis. Thus, fast translation of small nuORFs would result in a low read count for this element, which might be misinterpreted as low TE. On the contrary, count normalization by length e.g., TPM (Transcripts per kilobase million), for expression analysis is strictly not recommended[88]. Hence, analysis of nuORF length might reveal a data skew for certain lengths. As another parameter, codon context of nuORFs will determine TE. Slowly-translated codons induce ribosomal pauses during translation[111], which reduces TE. On the contrary, amino acid content might also modulate translational speed. The presence of high arginine levels with a high-arginine content uORF-originating peptide resulted in reduced peptidyl transferase centre activity of ribosomes, based on experiments in *Neurospora crassa* and wheat germ extracts[112]. Elements of the nuORF database exhibit an increased percentage of arginine and alanine (both ~11 %), in contrast to

the canonical proteome content of these amino acids (Arg: 5.5 %, Ala: 9 %)[90], marking differential amino acid content as an important parameter of this database. This feature would pose a substantial issue in the detection of nuPeptide with standard proteomics workflows, because these incorporate protein digestion with trypsin. This enzyme cleaves peptides and proteins at the C-terminal side of Arg and lysine (Lys). Thus, peptide sizes/ molecular weight would dramatically drop and may contribute to low coverage of nuPeptides in standard proteomic workflows.

Also, uORF distance to CDS might reveal supportive data. RPF distribution in the metagene profiles (**Figure 8 b,c; Supplements: Figure 16 i, j**) during Harringtonine treatment showed the strongest read density increase in direct vicinity to the CDS, suggesting uORF-CDS-distance is very likely to impact TE. Furthermore, global mRNA GC content was shown to increase in close vicinity of the translation start site[113]. This is supported by the GC content in the nuORF database (61.4 %), which I have compiled. The average genome GC content is about 41 %[113]. Increased GC contents are associated with complex mRNA secondary structure formation. These structures impact translation of the main ORF[114,115]. Thus, the integration of these parameters with TE data for the classification of nuORF elements would pose a substantial gain for the mechanistic characterization of non-canonical translation during mitotic-arrest.
In summary, I was able to detect uORF and uoORF-mediated translation, which is associated with increased TE, which lead to the compilation of the nuORF database. However, I highly suggest further characterization and expansion of these sequences as explained above.

3.3 nuORF features give rise to HLA complex I presented peptides

Non-canonical peptides can be presented on the cell surface. Recent work reported that 5'UTR and 3'UTR peptides were found in HLA complex I structures[56,68]. Furthermore, other drug treatment, such as Doxorubicin and Decitabine, were reported as inductors of tumour neoepitopes[78,79]. Thus, I sought to determine if 5'UTR nuORF elements would give rise to non-canonical, upstream/upstream-overlapping peptides (nuPeptides) during mitotic-arrest, which would be presented on the cell surface. This would hold a promising strategy in the identification of new treatment-associated targets for immunotherapy. Additionally, the induction of mitotic-arrest with in-clinic chemotherapeutic agent PTX in U-2 OS and the TNBC cell line SUM159-PT underlines the potential clinical relevance.
nuORF-derived peptides were detected for both cell lines using immunopeptidomics in collaboration with the DKFZ division of Immunotherapy and Immunoprevention. This is in line with prior findings that nuORFs encode biologically relevant, HLA-presented peptides in

benign and malignant cells, which suggests a role in cancer cell development and survival[116]. Strikingly, annotated proteome-derived and nuORF-derived peptides for U-2 OS and SUM159 cells exhibited comparable key peptide parameters. Proteome-derived and nuORF peptides cluster to the annotated HLA-alleles of each cell line with high binding affinities (**Figure 12 b,d,c; Figure 13 a,c,d**). Interestingly, nuORF-derived peptides are predicted to bind more to the HLA-A*32:01 allele and less to HLA-A*02:01 in U-2 OS (**Figure 12 b**). Conversely, in SUM-159PT cells, nuORF-derived peptides are predicted to bind less to the HLA-A*02:01 allele, whereas the rate of predicted no-binder slightly increases (**Figure 13 a**). During antigen processing, source proteins/ peptides are degraded by the proteasome into peptides. Thus, enrichment of certain protein substrates enriches a certain peptide pool. These peptides can exhibit different peptide sequences. The resulting possible peptide motifs affect which HLA allele is bound. Thus, the allele binding changes in both cell lines suggest alterations in the source peptide/ protein pool.

For practical considerations, this abundance change could affect the usage of HLA-A*02:01 peptides for screening of immunogenicity, which would need allotype-matching PBMCs (peripheral blood mononuclear cell). HLA-A*02:01 is one of the most common alleles globally [117], thus, screening T cell activation and cytokine release with HLA-A*02:01-binding peptides against this allele would enable a broader and globally-relevant applicability.

HLA complex I peptide length is another key parameter for high quality immunopeptidomics data. Normal HLA complex I peptide lengths vary between 8-11 AA residues, whereas HLA complex II peptides exhibits lengths between 13-25 AA residues. Notably, I found that annotated proteome and nuORF-derived peptides showed equal peptide *k*-mer distributions, peaking at 9-mer sequences (**Figure 12 c; Figure 13 b**). In both cell lines, nuORF-derived peptides exhibit a small increase in 11-mer peptides. Notably, cellular stress conditions have been described to induce the HLA class I binding to peptides with > 10 AA residues. This was shown for 11-mer, 12-mer and 14-mer peptides[118,119,120] due to viral infection or immune responses. Perhaps, the 11-mer nuPeptides will also induce stable HLA binding and immunogenic responses.

Altogether, the comparable allele clustering, binding affinities and peptide length distribution between annotated proteome-derived and nuPeptides supports the authenticity of nuPeptides as HLA class I binders and their comparable quality to annotated proteome-derived peptides. This suggests that nuPeptides may represent an underappreciated source of cancer antigens.

Next, I aimed to quantify nuPeptide presentation in response to mitotic-arrest. For this, I compared peptides from PTX treated cells in comparison to DMSO. With the aim of finding overlapping HLA-A*02:01-binding peptides between U-2 OS and SUM-159PT, I set moderately-low constraints for log2 fold change (LFC ≥ 0.5) and suitable constraints for the

adjusted *p*-value (*p*-adj ≤ 0.05). Hence, I identified 13 and 25 nuPeptides in U-2 OS and SUM-159PT cells during mitotic-arrest, respectively. Corresponding transcripts levels of these candidates during mitotic arrest did not indicate changes, whereas RiboSeq read density markedly increased in the respective uORF sequence, which is further supported by increased read densities at upstream initiation sites using Harringtonine run-off assays (**Figure *8* b,c**). Furthermore, I found none of the listed peptides to be annotated in the HLA ligand atlas or aligned against RefSeq protein using 'National Center for Biotechnology Information' NCBI protein BLAST. This underlines their treatment-associated and novel nature. Thus, we have termed these sequences as "therapy-induced uORF/uoORF-derived epitopes".

Qualitative analysis of presented ligands in U-2 OS and SUM-159PT cells did not show presentation of treatment-specific nuPeptides, meaning that nuPeptides were also found in DMSO-treated asynchronous samples. This effect can potentially be mediated by two factors. First, low levels translation of nuPeptides as part of normal cell homeostasis or low baseline level of cellular stressors, that mediated the translation of nuORF features. Second, asynchronous cell population of highly dividing cancer cells contain mitotic cells in small fractions (**Figure 6 a,b; Figure 7 g**). The PI staining that I have performed in U-2 OS cells showed enrichment of G2/M state cells of about 36 %, contributing to G2 and M state. This is also in line with my immunoblot experiment in **Figure *7* g**, which showed minor abundance of phospho-Ser10 H3 in the DMSO and DMSO+Torin1 condition, that is, mitotic cells in the asynchronous population.

Despite the moderate threshold, I did not detect a peptide overlap between the two cell lines. This might be mediated by several factors: stochastic proteasomal digestion of proteins/ peptides for HLA presentation. This might induce peptide sequence variances. Furthermore, U-2 OS and SUM-159PT exhibit different HLA alleles, which harbour differential peptide binding affinities. Thus, future experiments for the identification of shared peptides should be performed with allele-matching cancer cell lines.

The overall moderate presentation of nuPeptides in comparison to annotated proteome-derived peptides can be explained by the overall low abundance of nuPeptides in the global proteome. Abundant proteins dominate the pool of cleaved peptides suitable for HLA loading. Despite the translational up-regulation of nuORF sequences, I found that canonical proteins of the annotated proteome dominate the peptide pool, leaving only small space for nuPeptides (~1 % in U-2 OS and ~0.5 % in SUM-159PT of total peptides). Recent publications also highlight the overall amount of uORF- and uoORF-derived peptides as 0.75 %, of total mapped peptides[68].

Peptide stability poses an additional factor for peptide presentation. nuPeptides might be classified as 'defective ribosomal products' (DRiPs), which are rapidly cleaved by the proteasome upon translation[121]. Furthermore, nuORF-mediated peptides are described to be

more instable and have more intrinsic disordered regions[122]. This suggests that nuPeptides also exhibit reduced stability, which would further reduce the number of determined HLA class I ligands.

The number of down-regulated peptides upon mitotic-arrest might result from lower translation rates of this feature. The nuORF database was constructed from multiple cell lines using high-confidence features from asynchronous cells and mitotic-arrest. Thus, nuPeptides would appear to be down-regulated in mitotic arrest, if these were highly present in asynchronous cells. These high-confidence features would also be part of the database. Furthermore, PTX treatment might give rise to peptides with stronger HLA binding affinities, which would shift peptides with lower HLA binding affinities towards down-regulated ligands in this condition.
Future experiments should establish the connection of elevated nuPeptide presentation with immunogenicity. First, identified target peptides should be technically validated in LC-MS/MS using isotope-labelled synthetic peptides to assure peptide spectrum authenticity. Next, peptide-induced immunogenicity should be screened against allotype-matching PBMCs from human donors. This can be achieved by ELIspot assays and flow cytometry, to determine T cell activation and cytokine synthesis, such as IFNγ, TNFα and GranzymeB. With this approach immunogenic peptides can be identified, which poses a substantial knowledge gain in the field of immunotherapy.
Post-translational modification, such as phosphorylation, modulate peptide HLA class I and II binding[123,124]. Therefore, this modification might preserve peptides from degradation. Notably, cancer-associated deregulation of protein phosphorylation can give rise to new epitopes[123], which were shown to promote robust $CD8^+$ T cell responses[125]. Interestingly, I observed detection of predicted low-binding affinity or predicted non-binder peptides in the lPomics data from U-2 OS and SUM-159-PT cells (**Figure *12* b**, **Figure *13* b**). This finding might be mediated by peptide phosphorylation anchors, that allow HLA class I binding. Thus, I highly suggest to perform phosphoproteomics on asynchronous cell populations and cells during mitotic-arrest. This would allow further classification of source proteins and modified HLA ligands. Identified phospho-peptides should also be synthesized with this modification and tested for induction of immunogenicity. This approach will significantly enrich the pool of possible peptide targets.

Collectively, my findings provide confident evidence about the differential presentation of novel nuPeptides in HLA complex I structures upon mitotic-arrest with PTX. This emphasizes the role of the therapy-induced immunopeptidome.

3.4 Paclitaxel induces differential nuORF-derived peptide presentation and cancer killing

Next, I investigated the targeted immunogenic potential of PTX-induced uORF translation. I selected two high-confidence uORF elements with upregulated HLA complex I presentation: EIF4G2 and TPX2. I prepared 5'UTR-SIINFEKL reporter constructs as an orthogonal approach to underline the functional role of uORF peptides to elicit immune responses. For this, 5'UTRs were cloned into luciferase reporter vectors by F. Loayza-Puch with the uORF-derived peptide sequence change to SIINFEKL. This peptide originates from chicken ovalbumin and is recognized by murine OT-I $CD8^+$ T cell with transgenic TCRs. Replacing the actual peptide sequence with SIINFEKL allowed for better comparison between different uORFs, because peptide stability is equal and allows the comparison of the presentation of two different uORF sequences. Furthermore, the exchange also allowed the focus on translational control of uORFs during PTX-induced mitotic-arrest. Furthermore, the SIINFEKL peptide has strong binding affinity to the murine H2-Kb allele, which reduces the need of allotyping for fitting human alleles of the original nuPeptide. Thus, for these experiments the murine cancer cell line TC1 was chosen. These experiments aimed to model PTX-induced uORF translation and peptide presentation. These experiments were performed in a collective effort with Rossella Del Pizzo, Zhiwei Tang and Fabricio Loazya-Puch.

First, we transfected uORF-SIINFEKL reporter constructs in TC1 followed by DMSO treatment or mitotic-arrest. SIINFEKL presentation was inferred with the dye-conjugated 25-D1.16 antibody, which specifically detects SIINFEKL-loaded H2-Kb and aids detection of SIINFEKL presentation by flow cytometry. Mitotic-arrest significantly increased SIINFEKL presentation, regardless of the molecular mechanism of arrest induction (**Supplement: Figure *18* d,e**). Start codon mutation from optimal ATG to sub-optimal ATA reduced SIINFEKL-presentation.
Next, we investigated the immune response induction of uORF-induced SIINFEKL presentation. For this, TC1 cells we transfected with the uORF-SIINFEKL constructs, treated with PTX and co-cultured with *ex vivo* activated OT-I T cells. I found that co-culture induced a significant increase in IFNγ secretion and T cell-mediated cancer killing. This effect was only observed in transfected TC1 cells treated with PTX and not with PTX alone (**Figure *15* c,d**). Notably, I also observed reduced IFNγ secretion and cancer killing capabilities when using start codon mutated constructs from optimal ATG to suboptimal ATA.
These findings provide evidence that uORF-mediated peptide presentation can be induced by mitotic-arrest, regardless of the molecular mechanism. Also, uORF-derived peptides can elicit cytokine secretion and cancer killing. Furthermore, from this data I conclude that start codon strength appears to be an important factor for the non-canonical translation of uORFs in 5'UTR

sequences. Importantly, I did not find the increased SIINFEKL presentation to be mediated by differential reporter abundance, as RT-qPCR detection of luciferase transcript levels were comparable for both constructs between DMSO and PTX treatment (**Supplement: Figure *18* f**). This rules out a treatment-induced bias of transcriptional upregulation.
PTX can induce T cell-mediated cytotoxicity in an TCR-independent manner by release of cytotoxic extracellular vesicles[126]. Notably, to eliminate any potential TCR-independent cancer killing, TC1 cells were washed and medium was changed before T cell addition by Rossella Del Pizzo. Thus, preventing OT-I $CD8^+$ T cells from PTX exposition. This is supported by the lack of cancer killing capacity of OT-I against un-transfected TC1 cells treated with PTX.

PTX is a potent compound for the induction of mitotic-arrest and is commonly used as first-line of therapy in the treatment of breast cancer, ovarian cancer and non-small lung cancer. Beyond the annotated function of microtubule stabilization of the mitotic spindle apparatus, PTX is reported to elicit immunomodulatory effects in malignant and non-malignant cells. Apart from mitotic-arrest, PTX was shown to induce immunogenic cell death (ICD). This apoptotic phenomenon induces immune responses, thus, contributes to cancer killing of residual tumour cells[127]. ICD is characterized by the release of damage-associated molecular patterns (DAMPs). Indeed, PTX promotes the exposure calreticulin and ERp57 (protein disulfide isomerase family A member 3) DAMPs, which contribute to T cell activation and antigen-presenting cell (APC) activation[127,128]. Additionally, PTX increases levels of IFNγ secreting $CD8^+$ T cell and IL2 secreting $CD4^+$ T cells. Hence, PTX supports the activation of APCs, cytotoxic and helper T cells which is crucial for efficient anti-tumour immunity.
On the contrary, PTX was shown to contribute to metastasis by promoting lysyl oxidase (LOX) expression in $CD8^+$ T cells, which induces remodelling of the extracellular matrix[129]. This host-mediated mechanism aids cancer cell seeding and invasion.
As described before, PTX also induces very strong adverse effects in patients. These imply immunogenic hypersensitivity reaction, solvent-dependent and independent. The PTX solvent Cremophor EL® was shown to promote aerobic glycolysis and inflammatory gene expression via NF-kB activation[130]. However, adverse effects are still reported for drug formulations without this particular solvent[131]. Thus, the inflammatory induction by PTX might be partially explained with nuPeptide presentation. But this leaves the question if the intra-venous systemic administration of PTX also alters the immunopeptidome landscape in other tissues, such as internal organs, muscle or skin.
This divergence underlines the importance of future research on PTX effects on cancer and immune cells.

3.5 Perspective

This work contributed to the field of PTX research with the discovery of its new role as uORF/uoORF-derived peptide inducer. These findings present a promising strategy to elicit immune responses in chemotherapy-treated cancer.

In perspective, the presented findings should be further supported by the investigation of the potential role of the CDK1-DENR axis, as mentioned above. Comprehensive TE calculation following PTX treatment in cancer cell lines will establish a deeper understanding of nuORF translation. Subsequent immunogenicity screening of upregulated peptides will determine the immunogenic potential of nuPeptides *in vitro*.

However, the transformation of these fundamental biology findings to clinical settings is crucial. As described above, peptide sequences and nuORF sources vary in between cell lines, based on specific allele expression and probably genetic driver mutations, which influence the baseline transcriptional landscape. Only expressed transcripts that contain nuORF sequences will provoke non-canonical translation. Thus, a comprehensive nuORF compendium of solid cancers will help to identify cancer-specific nuORF translation. This can be achieved by quantitative RiboSeq performance from cancer cell lines and patient-derived xenografts of different cancer types. This would allow the collection and characterization of nuORF features, based on the listed parameters above: Kozak motif, nuORF length, codon context, amino acid content, uORF distance to CDS, GC content and structure prediction. This comprehensive collection will aid the identification of cancer type-dependently translated nuORF features. In combination with low-input immunopeptidomics novel nuPeptides can be determined and further classified for immunogenicity screening.

Throughout tumorigenesis, cancer cell harbour genetic mutations, such as single-nucleotide-variants (SNVs). These SNVs are well-studied in the context of transcriptional changes or loss-of-function in the resulting protein[132,133]. In addition, integration of RNAseq into the nuORF characterization will support the identification of mutated 5'UTR regions. Correlating 5'UTR mutations with nuORF translation might reveal mutational hot spots, that alter translation and the resulting peptide sequences. Mutated tumour peptides are more likely to elicit immune responses, thus, in-depth knowledge about 5'UTR mutations will aid the discovery of new immunotherapy targets.

In respective to adverse effects and auto-immunity, performing immunopeptidomics from internal organs, muscle and skin tissue in immunocompetent mouse models following PTX treatment will shape our understanding how cytostatic drugs might provoke severe immunogenic adverse effects, such as anaphylaxis. Eventually, this knowledge will foster refined effective cancer treatments with less adverse effects for the sake of better patient quality of life.

4 Methods

The text of the following chapter has been taken from the manuscript entitled “Upstream open reading frame translation enhances immunogenic peptide presentation in mitotically arrested cancer cells”[85] and was originally written by myself.

4.1 Method descriptions

4.1.1 Cell culture

I cultured all cell lines in DMEM High Glucose (**Table *5***) supplemented with 10 % FBS and 1 % PenStrep at 37 °C and 5 % CO_2. For mitotic shake-off experiments, I cultured cells to a confluency of 70 % and treated with mitosis-arresting compounds for 16 hours (BI: 0.1 µM, Nocodazole: 0.5 µM, Paclitaxel: 1 µM, STLC: 5 µM). I harvested mitotic cells using physical disruption via shake-off. I centrifuged the medium at 600x g for 10 min at 4 °C and washed the cell pellets in ice-cold PBS. All cell lines were regularly checked for contamination.

4.1.2 PI staining

To determine mitotic-arrested population fractions, I harvested cells during mitotic-arrest and stained DNA content with propidium iodide (PI). In brief, I harvested cells as described above and resuspended in them in Flow Wash Buffer (**Table *7***). Next, I counted cells with a Casy Counter system and adjusted to equal cell amounts. I fixed cells in ice-cold absolute ethanol and stored them overnight at -20 °C. I washed cells three times with Flow Wash Buffer until resuspension in PI staining buffer and incubation at 37 °C for 30 min under constant shaking. Next, I washed all samples twice with Flow Wash Buffer, filtered into round-bottom tubes with cell strainer and measured with a BD Canto II system. To select PI-stained single cells the following gating strategy was applied to all samples: I separated cells using FSC and SSC to distinguish cell populations from debris. Next, I identified single cells by combining FSC-H and FSC-A, followed by 488 nm excitation and filtering for BL84/42 signal (**Table *1***, **Table *2***)(**Supplement: Figure 19**).

4.1.3 Immunoblotting

I harvested cell pellets and lysed in whole-cell lysis buffer. Next, I added 1 µl of 0.1 M $MgCl_2$ solution and 2.5 U of Benzonase (**Table *10***) and incubated at 37 °C for 10 min to digest DNA. I used the Bradford assay to determine and level protein input. Next, I diluted lysates with water and 4x Laemmli Buffer (**Table *6***) (supplemented with 10 % β-mercaptoethanol). I used

SDS-PAGE to separate samples and transferred to nitrocellulose membranes using a TurboBlot system with transfer-buffer. Next, I blocked membranes in TBS-T (**Table *7***) plus either 5 % non-fat dry milk or 5 % BSA for 1 hour at room temperature. I diluted primary antibodies (**Table *12***) in TBS-T plus either 5 % non-fat dry milk or 5 % BSA and incubated with the membrane overnight at 4 °C and gentle rotation. Next, I washed the membranes 4 times with TBS-T for 5 min at room temperature before incubating with the secondary antibody in TBS-T for 1 hour at room temperature, protected from light. Eventually, I washed the membranes 4x with TBS-T and one wash with TBS for 5 min each, followed by imaging by an Odyssey Clx system.

4.1.4 RiboSeq

RiboSeq library preparation was performed as published before[43].

I processed RiboSeq samples as the following. At experiment end point, I washed cells with ice-cold PBS supplemented with 100 µg/ml CHX. Next, I resuspended the samples in lysis buffer (20 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 100 mM KCl, 1 % Triton-X 100, supplemented with 100 µg/ml CHX, 2 mM DTT and 1x Complete Protease Inhibitor (Roche)) (**Table *3***,**Table *4***, **Table *7***) and treated with 1300 U/ ml RNaseI (Ambion) (**Table *10***). I layered the samples onto sucrose gradients, prepared from sucrose concentrations of 47 % to 7 % with 20 mM TRIS-HCl pH 7.5, 10 mM $MgCl_2$, 100 mM KCl, 100 µg/ml CHX and 2 mM DTT. Next, samples were centrifuged for 2 hours at 36,000 rpm and 4 °C using a Beckman-Coulter ultracentrifuge (**Table *1***) and SW41-Ti rotor. Subsequently, I harvested monosome-containing fractions and digested with 50 µl recombinant Proteinase K (Roche) and 1 % SDS for 45 min at 45 °C. Next, I extracted total RNA using a standard phenol-chloroform-guanidinium thiocyanate protocol (**Table *8***). I size selected the resulting ribosome protected fragments (RPFs) (**Table *9***) with a denaturing urea-polyacrylamide gel and 3' dephosphorylated these using T4 PNK at 37 °C for 1 hour. Next, I ligated 5' pre-adenylated linkers to the 3' end of RPFs using T4 RNA Ligase 2 truncated KQ at 22 °C for 3 hours, following residual linker degradation with RecJf and 5' yeast de-adenylase for 45 min at 30 °C. Next, I performed rRNA depletion using custom biotinylated rRNA-oligonucleotides and streptavidin-coated magnetic beads. After reverse transcription using the SuperScriptIII First-Strand Synthesis Kit and RNA denaturation with 5 M NaOH for 3 min at 95 °C, I circularized RPFs with the CircLigase ssDNA Ligase (**Table *10***) and submitted these to index integration via PCR with the Q5 High Fidelity 2x Master Mix (**Table *13***,**Table *14***). I size selected the resulting samples with an 8 % non-denaturing PAGE gel. Next, I quantified the DNA via Qubit dsDNA HS kit and adjusted to 2 nM concentration, suitable for NextSeq2000 sequencing.

4.1.5 Harringtonine assay

Harringtonine assay-based RiboSeq data was supplied by Fabricio Loayza-Puch and Reuven Agami. In brief, cells were treated with 2 µg/ml Harringtonine at different time points before harvest. Control cells were scraped into medium, whereas cells arrested in mitosis were harvested by physical disruption and collecting medium. All samples were pelleted and resuspended in lysis buffer (20 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 100 mM KCl, 1 % Triton-X 100, supplemented with 100 µg/ml CHX, 2 mM DTT and 1x Complete Protease Inhibitor) (**Table *3***,**Table *4***). Thus, all samples were exposed to same Harringtonine treatment time.

4.1.6 5'UTR cloning

I amplified a synthetic DNA template containing the 5' UTR and the coding sequence for the SIINKEKL peptide by PCR using primers with overlapping sequences complementary to the pGL3-Promoter vector (Promega). Following PCR, I gel-purified the fragments to remove any non-specific products. I linearized the pGL3-Promoter vector using HindIII- HF. I combined the purified insert and linearized vector with NEBuilder HiFi DNA Assembly Master Mix and incubated the samples at 50°C for 60 minutes. I transformed the assembled product into competent cells for propagation and further analysis.

4.1.7 Mice

OT-I animals were bred in the German Cancer Research Center (DFKZ) animal facility. All mice were kept in a pathogen-free facility and used according to the DFKZ and following permission by the controlling government office (Regierungspräsidium Karlsruhe) according to the German Animal Protection Law, and in compliance with the EU Directive on animal welfare, Directive 2010/63/EU.

4.1.8 $CD8^+$ T cell isolation and culture

I isolated primary naïve $CD8^+$ OT-I T cells using the MojoSort Mouse CD8 Tcell isolation kit and subsequently activated them for 72 hours on plates coated with 2 µg/ml αCD3 and 2 µg/ml αCD28 at 37 °C (**Table *12***). I maintained the T cells in RPMI1640 supplemented with 10 % FBS, 1 % penicillin-streptomycin, 1 mM sodium pyruvate, 20 mM HEPES, 50 µM β-mercaptoethanol and 10 ng/ml murine IL-2.

4.1.9 T cell killing assay

First, TC1 cells were transfected with the uORF-SIINFEKL reporters using Lipofectamine 3000 by Zhiwei Tang. Following 24 hours post-transfection, F. Loayza-Puch seeded uORF-SIINFEKL-reporter transfected mouse cancer cells (TC1) in 12-well plates with DMEM High-Glucose medium and treated these with 1 µM PTX the next day. The cells were washed by F. Loayza-Puch with PBS and he replaced the medium with $CD8^+$ OT-I T cells in RPMI1640 medium at a ratio 1:2. Following 24 hours of incubation at 37 °C in a 5 % CO_2 atmosphere, the cells were washed with PBS and stained with crystal violet by F. Loayza-Puch to evaluate the killing efficiency. Imaging and analysis was conducted by F. Loayza-Puch using the Dual Lens System V850 Pro Scanner (Epson), and colony area was quantified using a previously published ImageJ plugin[134].

4.1.10 Flow cytometry

TC1 cells were transfected with the uORF-SIINFEKL reporters using Lipofectamine 3000 by Zhiwei Tang. The following steps were also performed by Zhiwei Tang. After 24 hours post-transfection, cells were synchronized in mitosis by treatment with 1 µM PTX for an additional 16 hours. Following mitotic arrest, cells were washed with PBS, detached using PBS-EDTA, and then pelleted. The cells were subsequently washed with PBS containing 0.5% BSA and incubated on ice and in the dark with APC-conjugated anti-mouse H-2Kb-SIINFEKL antibodies (**Table 12**) for 30 minutes. After incubation, the cells were washed twice with PBS containing 0.1% BSA and analysed using a FACS Canto II cytometer. Data analysis was conducted by Zhiwei Tag using FlowJo V10.4 software.

4.1.11 IFN-γ quantification

Cytokine release from OT-I CD8+ T cell co-culture was measured from the cell supernatant by Rossella Del Pizzo using the ELISA MAX Deluxe Set Mouse IFN-γ, following the manufacturer's guidelines. Each sample was analysed with the Multiskan FC plate reader (**Table 1**), using absorbance readings at 450 nm and 570 nm for subtraction. Final concentrations were calculated by Rossella Del Pizzo using a 4-parameter logistic curve-fitting algorithm in GraphPad Prism.

4.1.12 Quantitative real-time PCR

I reverse-transcribed a total of 500 ng of RNA using LunaScript RT Supermix and performed quantitative real-time PCR with Luna Universal qPCR Mix. I obtained Ct values using the

QuantStudio 5 RT qPCR System and analysed the data with QuantStudio Design and Analysis Software v2.6.0. Eventually, I calculated mRNA fold change of target genes using the ΔΔCt[135] method, with mRNA expression normalized to GAPDH. Primers are listed in (**Table 15**).

4.1.13 Immunopeptidomics sample processing

I harvested Input cell lines (U-2 OS, SUM159) for immunopeptidomics in triplicates with $5x10^7$ cells per replicate. I treated the cell line U-2 OS with 1 µM PTX and SUM159 with 0.1 µM PTX. For DMSO-treated control conditions, I gently scraped cells in ice-cold PBS. For PTX-treated conditions, I harvested mitotic cells using physical disruption by shake-off, spun them at 600x g for 10 min at 4 °C and washed with ice-cold PBS. I counted all samples of all conditions using a CASY II system and snap-frozen.

The following parts were performed and described by Rebecca Köhler and Jonas P. Becker from the DKFZ Division of Immunotherapy and Immunoprevention[85].
Immunoprecipitation of HLA class I:peptide complexes was performed as previously described[136] with additional steps for the forced oxidation of methionine using H_2O_2 and reduction and alkylation of cysteine using tris(2-carboxyethyl)phosphine (TCEP) and iodoacetamide (IAA). Lyophilized peptides were dissolved in 12 µl of 5% ACN in 0.1% TFA and spiked with 0.5µl of 100 fmol/µl Peptide Retention Time Calibration (PRTC) Mixture (Pierce) and 10 fmol/µl JPTRT 11 (a subset of peptides from the Retention Time Standardization Kit; JPT) and transferred to QuanRecovery Vials with MaxPeak HPS (Waters, Milford, MA, USA). All samples were analyzed using an UltiMate 3000 RSLCnano system coupled to an Orbitrap Exploris 480 equipped with a FAIMS Pro Interface (Thermo Fisher Scientific). For chromatographic separation, peptides were first loaded onto a trapping cartridge (Acclaim PepMap 100 C18 µ-Precolumn, 5µm, 300 µm i.d. x 5 mm, 100 Å; Thermo Fisher Scientific) and then eluted and separated using a nanoEase M/Z Peptide BEH C18 130A 1.7µm, 75µm x 200mm (Waters). Total analysis time was 120 min and separation was performed using a flow rate of 0.3 µl/min with a gradient starting from 1% solvent B (100% ACN, 0.1% TFA) and 99% solvent A (0.1% FA in H_2O) for 0.5 min. Concentration of solvent B was increased to 2.5% in 12.5 min, to 28.6% in 87 min and then to 38.7% in 1.4 min. Subsequently, concentration of solvent B was increased to 80% in 2.6 min and kept at 80% solvent B for 5 min for washing. Finally, the column was re-equilibrated at 1% solvent B for 11 min. The LC system was coupled on-line to the mass spectrometer using a Nanospray-Flex ion source (Thermo Fisher Scientific), a SimpleLink Uno liquid junction (FossilIonTech) and a CoAnn ESI Emitter (Fused Silica 20 µm ID, 365 µm OD with orifice ID 10 µm; CoAnn Technologies). The mass spectrometer was operated in positive mode and a spray voltage of

2400 V was applied for ionization with an ion transfer tube temperature of 275 °C. For ion mobility separation, the FAIMS module was operated with standard resolution and a total carrier gas flow of 4.0 l/min. Each sample was injected twice using either a compensation voltage of -50 V or -70 V for maximal orthogonality and thus increased immunopeptidome coverage. Full Scan MS spectra were acquired for a range of 300 – 1650 m/z with a resolution of 120.000 (RF Lens 50%, AGC Target 300%). MS/MS spectra were acquired in data-independent mode using 44 previously determined dynamic mass windows optimized for HLA class I peptides with an overlap of 0.5 m/z. HCD collision energy was set to 28% and MS/MS spectra were recorded with a resolution of 30.000 (normalized AGC target 3000%).

4.1.14 Immunopeptidomics spectra analysis

This analysis was performed and described by Jonas P. Becker[85].
MS raw data was analysed using Spectronaut software (version 17.6; Biognosys[137]) and searched against the UniProtKB/Swiss-Prot database (retrieved: 21.10.2021, 20387 entries) as well as a database containing protein sequences longer than 7 amino acids predicted from translation of uORFs. Search parameters were set to non-specific digestion and a peptide length of 7-15 amino acids. Carbamidomethylation of cysteine and oxidation of methionine were included as variable modifications. Results were reported with 1% FDR at the peptide level. Peptides identified by Spectronaut were further analyzed using NetMHCpan 4.1 binding predictions[138], Gibbs 2.0 clustering of peptide sequences[139], and retention time prediction by DeepLC[140]. uORF-derived peptide sequences were manually validated using Skyline (version 22[141]) by comparison against spectral libraries *in silico* predicted using PROSIT[142]. Normalized spectral angles (NSAs) were calculated as described previously[143]. Quantification of HLA class I-presented peptides was performed as described previously[74] using the raw output at the MS2 level from Spectronaut 17.6 with cross-run normalization disabled and a custom script in the R programming language. Peptides with an FDR ≤ 0.05 and an foldchange > 2 were defined as “hits” while peptides with an FDR ≤ 0.2 and an foldchange ≥ 1.5 were defined as “candidates”. All results were visualized using in-house developed R scripts.

4.1.15 RiboSeq data processing

Samples were sequenced in cooperation with the group of Reuven Agami at the Netherlands Cancer Institute, Amsterdam, or I sequenced the samples at the Sequencing Open Lab affiliated with the DKFZ Genomics Core Facility. The FASTQ raw data was provided by the DKFZ Genomics Core Facility. In brief, I trimmed sample adapters using cutadapt (v3.4)[144] and demultiplexed with barcode_splitter from FASTX-toolkit (v0.0.6) (Gordon and Hannon,

2010). I dropped fragments smaller than 30 nt. I performed UMIs extraction using umi_tools (v1.1.1)[145] By BLAST-Like Alignment Tool (BLAT) (v36x2), I filtered out and discarded rRNA reads[146]. The rRNA index for RNA18S5, RNA28S5 and RNA5-8S5 was constructed by Ekaterina Stepanova manually from NCBI RefSeq annotation. I aligned the remaining reads with Spliced Transcripts Alignment to a Reference (STAR) (v2.5.3a)[147] to GRCh37/hg19 with --outSAMtype BAM Unsorted --readFilesCommand zcat --quantMode TranscriptomeSAM GeneCounts --outSAMmapqUnique 0. I obtained genome coverage bigwig tracks using samtools (v1.15.1)[148] and bedtools (v2.24.0)[149].

4.1.16 Read clustering

I filtered counting tables from hg19 alignment with ≥ 20 reads per sample and feature. Next, I determined trimmed mean of M-values (TMM) for cross sample comparison using the "calcNormFactors" function from edgeR[150]. I used the resulting TMM-normalized counts per million for heatmap creation using the libraries pheatmap, RColorBrewer and viridis with R Statistical Software (v4.2.0; R Core Team).

4.1.17 Transcript distribution

To analyse ribosome transcript distribution, I used a the transcript distribution software of the Diricore package[48]. The following paragraph describes the rationale of this analysis.
The most representative isoform for each gene was selected by using a hierarchical selection strategy. RPF counts were obtained for each selected transcript and intra-gene normalization was performed by dividing the cumulative read counts for each region (5'UTR, CDS, 3'UTR) by the total RPF counts for that transcript, to allow comparison across different regions of the same transcript. Next, RPF density was computed across regions for every transcript using read count interpolation over a fixed grid of 2000 points. Transcripts with fewer than 50 reads were excluded from the downstream analysis. Next, interpolated RPF densities across transcripts were averaged and subjected to Gaussian smoothing for noise reduction. Finally, resulting RPF densities were plotted with corresponding transcript regions. All analysis parts were conducted using custom Python scripts incorporating numpy, scipy, and matplotlib libraries.

4.1.18 RPF Location Percentage

To obtain percentages of RPFs within the 5'UTR and 3'UTR, I performed an analysis with Ribowaltz (v1.2.0)[151] with R Statistical Software (v4.2.0; R Core Team). I extracted the numeric values from the P-site percentage feature of Ribowaltz.

4.1.19 Over-representation analysis

I calculated enriched gene sets by clusterProfiler[152] (**Figure 9 f**, **Figure *17* b**) with R Statistical Software (v4.2.0; R Core Team) and using a custom background of all genes detected in this RiboSeq experiment. For the over—representation analyses of genes of differentially translated nuORF features (**Figure 10; Supplement: Figure 17 d**) and nuPeptides (**Figure 14; Supplement: Figure 18 c**), I performed an analysis using enrichR[153] without the use of a custom background.

4.1.20 Upstream translation initiation sites quantification

I predicted uORF and uoORF genomic coordinates from U-2 OS RiboSeq data, compiled this information in SAF format and used it for counting using featureCounts[154] from the subread package (v1.5.1). I filtered the resulting counting tables for ≥ 5 reads per sample and feature. Next, I determined trimmed mean of M-values (TMM) for cross sample comparison using the "calcNormFactors" function from edgeR[150]. I subjected normalized counts per million to sample-specific outlier calculation using the Grupps function. Subsequently, I calculated Log2(TMM-normalized counts) using a custom awk script.

4.1.21 ORF Prediction

I predicted ORFs using PRICE[56]. In brief, I re-aligned umi-extracted and rRNA filtered FASTQ files using STAR (v2.5.3a) to GRCh37/hg19 with important outSAMattributes needed by PRICE (--outSAMtype BAM Unsorted --alignEndsType Extend5pOfReads12 --outSAMattributes nM MD NH --readFilesCommand zcat --quantMode TranscriptomeSAM GeneCounts --outSAMmapqUnique 0). I prepared the PRICE reference genome as described using hg19 FASTA and GTF files from Gencode. Next, I ran PRICE with the respective BAM files using '~/Gedi/Gedi_1.0.5/gedi -e Price -D -genomic hg19 -progress -plot'. Subsequently, I quantified all ORF features with *p*-value ≤ 0.05 with standard UNIX-commands in a custom bash script.

Subsequently, I adjusted ORF tables from PRICE to the BED format with chromosome, ORF feature start position, ORF feature end position, ORF feature ID, chromosome strand and

Gene ID. Next, I converted the resulting BED6 files to BED12 format[1], which I used as input for bedtools[149] getfasta with -s -name -split. Eventually, I generated peptide sequences using the faTrans program from the UCSC utility tools[155].

4.1.22 RNAseq Processing

Samples were sequenced in cooperation with the group of Reuven Agami at the Netherlands Cancer Institute, Amsterdam. FASTQ files were provided by FLP. First, I performed adapter trimming with cutadapt (v3.4). Next, I aligned samples to GRCh37/hg19 using STAR[147] (v2.5.3a) with GRCh37/hg19 with --outSAMtype BAM Unsorted --readFilesCommand zcat --quantMode TranscriptomeSAM GeneCounts --outSAMmapqUnique 0. I obtained genome coverage bigwig tracks using samtools[148] (v1.15.1) and bedtools[149] (v2.24.0).

4.1.23 Translational efficiency

I calculated translational efficiency of ncORFs by extracting IDs, start and end positions of predicted uORF and uoORF features (p-value < 0.05) from PRICE ORF tables and arranged these in SAF format, creating a ncORF SAF reference file. Next, I determined read counts in ncORF regions with featureCounts[154] (v1.5.1) and genome-based BAM files from RNAseq and RiboSeq. I subjected the resulting aggregated count matrix to RiboDiff[52] (v0.2.1) calculation. I discarded features with missing calculation. I generated this data using pseudo-replicates prepared with seqkit shuffle and split function[156] from the existing RNAseq data.

4.1.24 Data availability

The sequence data from this study have been submitted to the GEO repository: GSE281253. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD057839.

[1] https://github.com/muhligs/bed6ToBed12

4.2 Tables

Table 1: Technical devices

Product	Supplier
Adventurer Scale	Otthaus
BBD6220	Heraeus
BTD	Grant
Centrifuge 5424R	Eppendorf
Centrifuge 5425	Eppendorf
FACS Canto II	BD
Gel Dox XR+	Bio-Rad
Innova U725	New Brunswick Scientific
Kern EWJ	Kern & Sohn
MediLine LCexv 4010	Liebherr
MediLine LGex 3410	Liebherr
Mini-Vav-Power	Peqlab
Multifuge X312	Thermo Fisher Scientific
NextSeq 2000	Illumina
Nu-437-600E	NuAire
Odyssey	Licor
PowerPac HC	Bio-Rad
Qubit3	Invitrogen
T100 Thermal Cycler	Bio-Rad
Thermomixer Comfort	Eppendorf
Trans-Blot Turbo	Bio-Rad
Waterbath	Memmert
Multiskan FC plate reader	Thermo Fisher Scientific
Dual Lens System V850 Pro Scanner	Epson
QuantStudio 5 RT qPCR System	Thermo Fisher Scientific

Table 2: Software

Product	Supplier
ImageLab 6.1	Bio Rad
Image Studio	Licor
GraphPad Prism	GraphPad Software Inc.
FlowJo	BD
FACSDiva	BD

Table 3: General Chemicals

Product	Supplier	Reference
Triton X-100	Sigma Aldrich	T8787-250ml
TRIS	Acros Organics	167620010
$MgCl_2$	Sigma Aldrich	M2393-500g
KCl	Sigma Aldrich	P5405-250g
CHX	Sigma Aldrich	C7698-5G
SDS 20 % Solution	Sigma Aldrich	05030-1L-F
DTT	Sigma Aldrich	D0631-10G
cOmplete Protease Inhibitor	Merck	11873580001
NaCl	Fisher Chemical	S/3160/60
EDTA	Sigma Aldrich	EDS-1kg
GlycoBlue	Thermo Fisher Scientific	AM9516
SYBR Gold	Thermo Fisher Scientific	S11494
RNaseIn Plus	Promega	N2611
Borate	Sigma Aldrich	31146-2.5kg
Glycerol	Sigma Aldrich	G6279-1L
TEMED	Sigma Aldrich	T7024-25ml
40 % Acrylamide/Bis Solution, 19:1	Bio-Rad	1610144
40 % Acrylamide/Bis Solution, 37.5:1	Bio-Rad	1610148
β-mercaptoethanol	Sigma Aldrich	M6250
Propidiumiodide	Sigma Aldrich	537060

Table 4: General Lab Supplies

Product	Supplier	Reference
0.45 µM Filter	Merck	SLHAR33SS
15 ml Centrifuge tubes	Sarstedt	62.554.502
18 G x 1 1/2'' needle	Terumo	NN-1838S
50 ml Centrifuge tubes	Greiner Bio One	227 261
50 ml Syringe	Terumo	SS+50ES1
Nitrocellulose 0.45 µM	Thermo Fisher Scientific	1620115
Round-Bottom Tubes	Falcon (FACS)	352235
UltraClear Tubes	Beckmann Coulter	344059

Table 5: Cell Culture Media and Reagents

Product	Supplier	Reference
DMEM High-Glucose	Thermo Fisher Scientific	41966-029
FBS	Thermo Fisher Scientific	10270-106
Pen/Strep	Thermo Fisher Scientific	15140-122
DMSO	Thermo Fisher Scientific	D8418-50ml
RPMI1640	Thermo Fisher Scientific	21875-034
Sodium pyruvate	Thermo Fisher Scientific	11360-070
Lipofectamine 3000	Thermo Fisher Scientific	L3000008

Table 6: Pre-formulated buffers

Product	Supplier	Reference
5 M Ammonium Acetate	Thermo Fisher Scientific	AM9070G
PBS	Thermo Fisher Scientific	10977-035
4x Laemmli Buffer	Bio-Rad	1610747
5x Bradford Reagent	SERVA	39222.03
HEPES	Sigma Aldrich	H4034

Table 7: In-house buffers

Buffer	Formulation
Whole-cell lysis buffer	1 M TRIS pH 7.5 10 % Glycerol 2 % SDS
PBS + CHX	100 µg/ml CHX In PBS
Transfer Buffer	240 mM TRIS 195 mM Glycin 0.5 % SDS
Flow Wash Buffer	1 mM EDTA In PBS
Flow PI Buffer	50 µg/ml PI 0.2 mg/ml RNase A 0.4 % Triton X-100 1 mM EDTA In PBS
RiboSeq Lysis Buffer	20 mM TRIS pH 7.5 10 mM $MgCl_2$ 100 mM KCl 1 % Triton X-100
RIPA Lysis Buffer	50 mM TRIS pH 7.5 150 mM NaCl 0.5 % Sodium desoxycholate 0.1 % SDS 1 % Triton X-100
TBS	0.2 M TRIS 1.5 M NaCl Ad pH 7.5
TBST	0.2 M TRIS 1.5 M NaCl 0.1 % SDS Ad pH 7.5
SDS Running Buffer (10x)	248 mM TRIS 1.92 M Glycine 1 % SDS

	Ad pH 8.3
TBE	89 mM TRIS
	89 mM Borate
	2 mM EDTA

Table 8: RNA extraction chemicals

Product	Supplier	Reference
TRI Reagent	Zymo Research	R2050-1-200
Ethanol	Sigma Aldrich	32221-2.5L-M
Propan-2-ol	VWR	20842.330
Chloroform	VWR	22711.260

Table 9: Gel Compositions

Product	Formulation	
12 % PAA Denaturing Gel	Urea	4.8 g
	10x TBE	1 ml
	40 % AA/Bis 19:1	3 ml
	dH2O	2.5 ml
	APS	50 µl
	TEMED	10 µl
8 % PAA Denaturing Gel	Urea	4.8 g
	10x TBE	1 ml
	40 % AA/Bis 19:1	2 ml
	dH2O	3.5 ml
	APS	50 µl
	TEMED	10 µl
8 % PAA Non-Denaturing Gel	10x TBE	1 ml
	40 % AA/Bis 19:1	2 ml
	dH2O	3.5 ml
	APS	50 µl
	TEMED	10 µl
12 % SDS-PAGE Resolving Gel	dH2O	3.1 ml
	TRIS pH 8.8	1.8 ml
		70 µl

	SDS 10 % 40 % AA/Bis 37.5:1 APS TEMED	2.1 ml 70 µl 2.8 µl
SDS-PAGE Stacking Gel	dH2O TRIS pH 8.8 SDS 10 % 40 % AA/Bis 37.5:1 APS TEMED	2.7 ml 0.4 ml 35 µl 0.4 ml 35 µl 3.5 µl

Table 10: Enzyme and kit systems

Enzyme	Supplier	Reference
Benzonase	Merck	70746-4
SuperScript III First-Strand Synthesis System	Thermo Fisher Scientific	18080051
CircLigase II ssDNA ligase	LGC BioSearch	CL9021K
RNaseI	Ambion	AM2294
Proteinase K	Roche	3115828001
Mth RNA Ligase	New England Biolabs	M2611AA
T4 PNK	New England Biolabs	M0201S
T4 Rnl2 K227Q	New England Biolabs	M0351L
5‘Deadenylase	New England Biolabs	M0331S
RecJf Exonuclease	New England Biolabs	M0264S
Q5 Highfidelity 2x Master Mix	New England Biolabs	M0494S
HindIII-HF	New England Biolabs	R3104
NEBuilder HiFi DNA Assembly Master Mix	New England Biolabs	E2621
ELISA MAX Deluxe Set Mouse IFN-γ	BioLegend	430815
Luna Script RT SuperMix	New England Biolabs	M3010L
Luna Universal qPCR Mix	New England Biolabs	M3003X

Table 11: List of cytostatic compounds

Compound	Supplier	Reference
PTX	Santa Cruz	sc-201439
STLC	Torcris	2191
BI2536	Cell Signalling Technology	26744

Nocodazole	Sigma-Aldrich	M1404
Harringtonine	Santa Cruz	sc-204771

Table 12: Antibodies and Cytokines

Antibody/ Cytokine	Supplier	Reference
Phospho-H3 (Ser10)	Cell Signaling	9701
Phospho-4EB1 (Thr37/46)	Cell Signaling	2855
GAPDH	Proteintech	60004-I-Ig
4EBP1	Cell Signaling	9644
Calnexin	Proteintech	66903-1-Ig
Phospho-eIF2α (S51)	Cell Signaling	9721S
eIF2α	Cell Signaling	5324S
IRDye αRabbit 800CW	Licor	926-32211
IRDye αMouse 680RD	Licor	926-32210
αCD3	BioXCell	BE0001-1
αCD28	BioXCell	BE0015-1
IL-2	BioLegend	575404
MojoSort Mouse CD8 Tcell isolation kit	BioLegend	480007
αH2-Kb-SIINFEKL clone 25-D1.16	BioLegend	141606

Table 13: Library PCR temperature profile

Temperature	Time [min:sec]	Cycles
98 °C	0:30	1
98 °C	0:10	
65 °C	0:10	Variable
72 °C	0:05	
72 °C	5:00	1
12 °C	Infinite	1

Table 14: RiboSeq Library Oligonucleotides

Identifier	Sequence

3'linker ATCGT	/5Phos/NNNNNATCGTAGATCGGAAGAGCACACGTCTGAA/3ddC/	IDT
3'linker AGCTA	/5Phos/NNNNNAGCTAAGATCGGAAGAGCACACGTCTGAA/3ddC/	IDT
3'linker CGTAA	/5Phos/NNNNNCGTAAAGATCGGAAGAGCACACGTCTGAA/3ddC/	IDT
3'linker CTAGA	/5Phos/NNNNNCTAGAAGATCGGAAGAGCACACGTCTGAA/3ddC/	IDT
3'linker GATCA	/5Phos/NNNNNGATCAAGATCGGAAGAGCACACGTCTGAA/3ddC/	IDT
3'linker GCATA	/5Phos/NNNNNGCATAAGATCGGAAGAGCACACGTCTGAA/3ddC/	IDT
3'linker TAGAC	/5Phos/NNNNNTAGACAGATCGGAAGAGCACACGTCTGAA/3ddC/	IDT
3'linker TCTAG	/5Phos/NNNNNTCTAGAGATCGGAAGAGCACACGTCTGAA/3ddC/	IDT
RT Primer	/5Phos/NNAGATCGGAAGAGCGTCGTGTAGGGAAAGAG/-iSp18/ GTGACTGGAGTTCAGACGTGTGCTC	IDT
Ligation Control	rUrGrUrUrArGrGrGrArUrArArCrArGrGrGrUrArArUrGrCrGrA-NNNNNCGATCTGATCGGAAGAGCACACGTCTGAArC	IDT
RT Control	/5Phos/NNAGATCGGAAGAGCGTCGTGTAGGGAAAGAG/iSp18/-GTGACTGGAGTTCAGACGTGTGCTCTTC CGATCACAGTC-NNNNNTCGCATTACCCTGTTATCCCTAACAT	IDT
PCR Forward	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTC	IDT
PCR Reverse 1	CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCAGACGTGTG	IDT
PCR Reverse 2	CAAGCAGAAGACGGCATACGAGATACATCGGTGACTGGAGTTCAGACGTGTG	IDT
PCR Reverse 3	CAAGCAGAAGACGGCATACGAGATGCCTAAGTGACTGGAGTTCAGACGTGTG	IDT
34mer Product	rArUrGrUrArCrArCrUrArGrGrGrArUrArArCrArGrGrGrUrArArUrCrArArCrGrCrGrA	IDT

Table 15: RT-qPCR primer

Target	Sequence	Supplier
ASNS forward	CGACCAAAAGAAGCCTTCAG	Sigma
ASNS reverse	CCACTTGGGCATCCAGTAAT	Sigma
CHOP forward	TGGAAGCCTGGTATGAGGAC	Sigma
CHOP reverse	TGTGACCTCTGCTGGTTCTG	Sigma
GAPDH forward	GAGTCAACGGATTTGGTCGT	Sigma
GAPDH reverse	TTGATTTTGGAGGGATCTCG	Sigma

Table 16: Reporter Cloning Oligonucleotides

Target	Sequence	Supplier
eIF4G2	GCAGACGGCAACCGGGCCGCTG ATTGGGCGGCGAAGGAGCCATT CGGGGAGACTCTGGTGGGTTCG	Sigma

uORF SIINFEKL template	GCTGCCCCAAGAGTGATAAGTTC GGCTTCAGACACGCCTTAGCGCC AGCAGTGAGTCGGAGCTCTATGT CCATCATCAACTTCGAAAAGCTC TGGCGGCTGCAGCAGCGACTCC TCTGAGCTGAGTTTGAGGCCGTC CCCGACTCCTTCCTCCCCCTTCC CTCCCCCTTTTTTTTGTTTTCCGT TCCCCTTTCCCCTCCCTTCCCTA TCCCCGACGACCGGATCCTGAG GAGGCAGCTGCGGTGGCAGCTG CTGAGTTCTCGGTGAAGGTATTT CATTTCTCCTGTCCCCTCCCCTC CCCACCCCATCTATTAATATTATT CTTTTGAAGATTCTTCGTTGTCA AGCCGCCAAA	
eIF4G2 *uORF SIINFEKL* Forward primer	AGGCCTAGGCTTTTGCAAAAGCAGACGGCAACCGGGCC	Sigma
eIF4G2 *uORF SIINFEKL* Reverse primer	CCAACAGTACCGGAATGCCATTTGGCGGCTTGACAACGAAGAATCTTC	Sigma
TPX2 uORF SIINFEKL template	TGTAGGCCTGATAGACTGATTAA ACCACAGAAGGTGACCTGCTGA GAAAAGTGGTACAAATACTGGGA AAAACCTGCTCTTCTGCGTTATC CATCATCAACTTCGAAAAGCTCA AAGCTCTTATTCCTATGATGCCC CCTCGGATTTCCATCATCAACTT CGAAAAGCTCATGAAGGAGATA CTCAAAACATAGATTCAT	Sigma
TPX2 *uORF SIINFEKL* Forward primer	AGGCCTAGGCTTTTGCAAAATGTAGGCCTGATAGACTG	Sigma
TPX2 *uORF SIINFEKL* Reverse primer	CCAACAGTACCGGAATGCCAATGAATCTATGTTTTGAGTATCTC	Sigma

5 Supplements

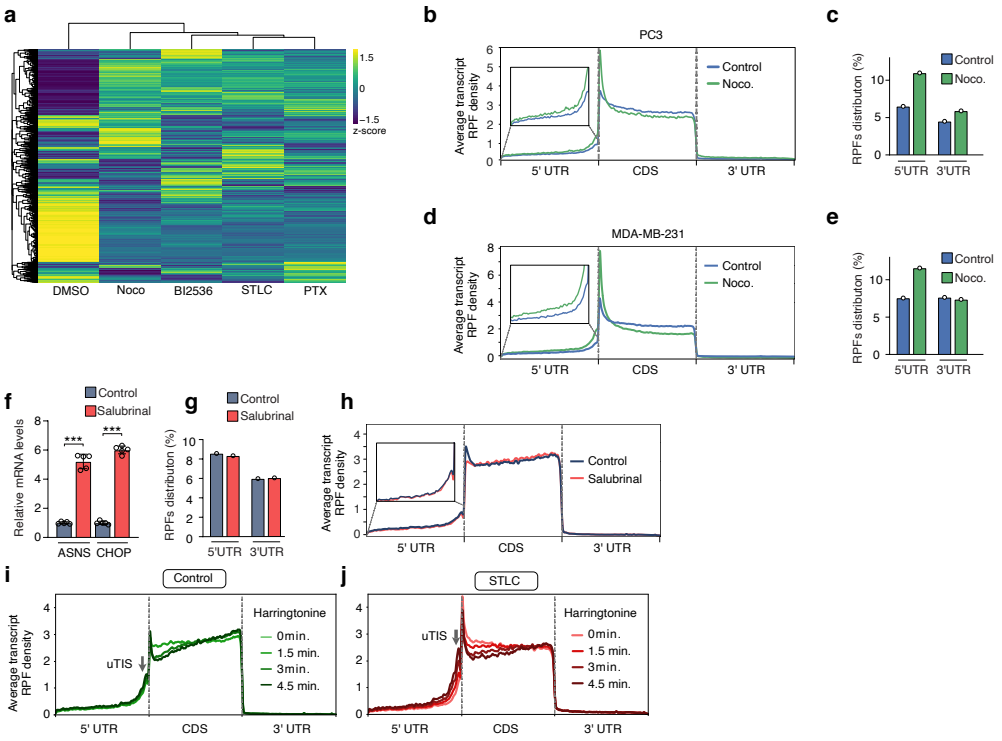

Figure 16: Mitotic-arrest induces active translation in the 5'UTR. (a) Clustered TMM-normalized counts of RiboSeq data obtained from U-2 OS cells treated with DMSO or different mitotic-arrest inducing agents (0.1 µM BI2536, 0.5 µM Nocodazole (Noco), 1 µM Paclitaxel (PTX), 1 µM S-Trityl-L-cysteine (STLC)). Metagene profiles of PC3 **(b)** and MDA-MB-231 **(d)** cells treated with DMSO or 0.5 µM Nocodazole (Noco) for 16 hours. Percent quantification of RPF location for PC3 **(c)** and MDA-MB-231 **(e)** cells treated with DMSO or 0.5 µM Nocodazole (Noco) for 16 hours. **(f)** RT-qPCR bar plot of indicated targets in U-2 OS cells either treated with DMSO of 50 µM Salubrinal for 16 hours. Data represents mean ± SD from technical replicates. Statistical analysis was performed using a two-tailed unpaired t-test. *** *p*-value < 0.001. **(h)** Metagene profile of U-2 OS cells either treated with DMSO or 50 µM Salubrinal for 16 hours. Metagene profiles of Harringtonine run-off assays in MDA-MB-231 cells treated with DMSO **(i)** or 1 µM STLC **(j)**.

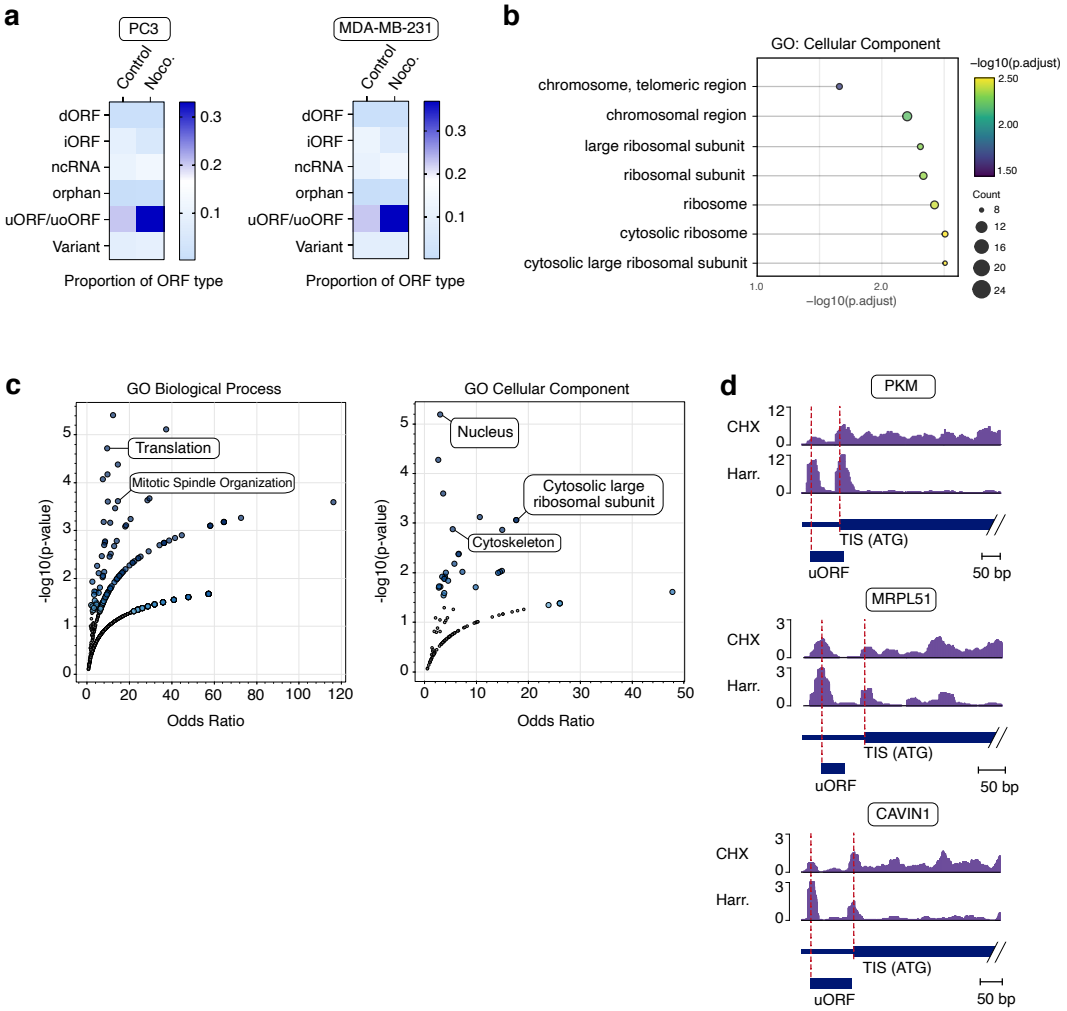

Figure 17: Mitotic-arrest enhances translation of upstream and upstream-overlapping open-reading frames (uORF/uoORF) in various cancer cells. (a) Proportion of ORF types from RiboSeq data in PC3 and MDA-MB-231 cells either treated with DMSO or 0.5 µM Nocodazole (Noco). **(b)**,**(c)** Volcano plot of over-representation analysis (Gene Ontology: Biological Process and/ or Cellular Component) from common uORF-containing genes listed in **Figure 9 e (b)** or from nuORF-bearing transcripts **(c)** with increased translational efficiency in **Figure 10 b**. Volcano plots represents single terms as points with corresponding odds ratio and -log10(*p*-value). Larger and darker-coloured points represent more significantly enriched gene sets for the particular term. **(d)** RPF genome coverage from Harringtonine run-off assays in U-2 OS cells treated with 1 µM STLC for 16 hours and additional CHX or CHX and Harringtonine for 10 min. CHX, Cycloheximide. Harr, Harringtonine.

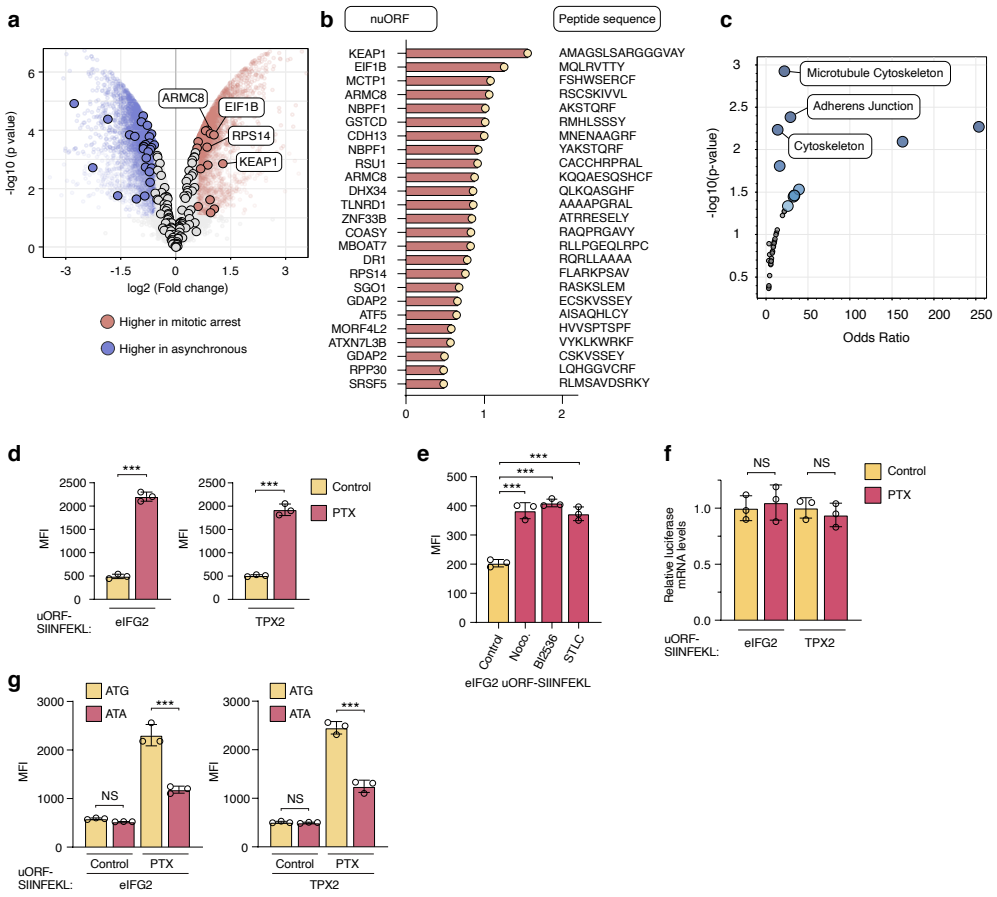

Figure 18: Paclitaxel induces differential presentation of nuORF-derived peptides and immune responses. **(a)** Volcano plot showing the label-free quantification of HLA class I immunopeptidome in SUM-159-PT cells, highlighting nuORF-derived peptides (big spots) and peptides from the annotated proteome (small spots). Red peptides: Log2 fold change ≥ 0.5 and adjusted *p*-value < 0.05. Blue peptides: Log2 fold change ≤ -0.5 and adjusted *p*-value < 0.05. *P*-values were calculated using a Bayes moderated t-test with two-sided *p*-values. **(b)** nuORF-containing gene with Log2 fold change peptide abundance and corresponding peptide sequence from SUM-159-PT cells. **(c)** Over-representation analysis (Gene Ontology: Biological Process and Cellular Component) from upregulated nuORF-derived peptides in U-2 OS cells from **Figure 12 b**. **(d-g)** Detection of SIINFEKL:H-2K by flow cytometry, shown as bar plot of Mean Fluorescence Intensity (MFI) from murine TC1 cells transfected with respective uORF-reporter: **(d)** Treatment with 1 µM Paclitaxel (PTX) for 16 hours. **(e)** Treatment with DMSO, 0.1 µM BI2536, 0.5 µM Nocodazole (Noco), or 1 µM S-Trityl-L-cysteine (STLC) for 16 hours. **(f)** RT-qPCR of firefly luciferase in TC1 cells transfected with respective uORF-reporter constructs. Cells were treated with 1 µM PTX for 16 hours. **(g)** TC1 cells were transfected with start codon mutated uORF-reporter constructs and treated with 1 µM PTX for 16 hours. **(d-g)** All data represent mean ± SD from biologically independent experiments (n = 3). Statistical analysis was conducted using a two-tailed unpaired t-test. *** *p*-value < 0.001.

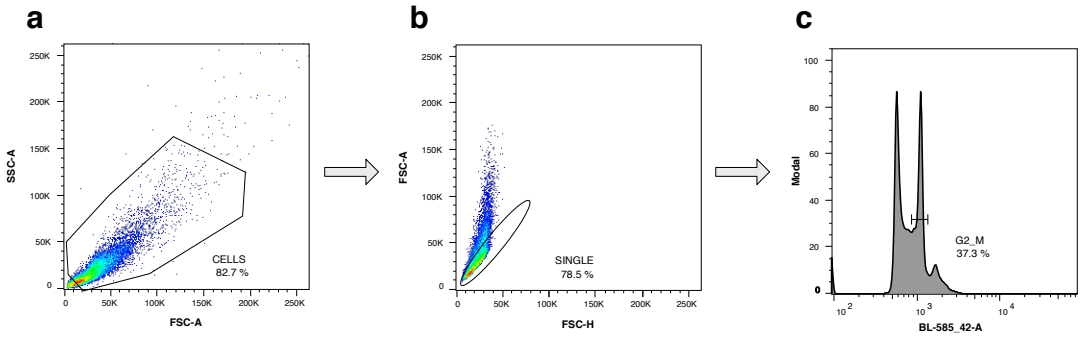

Figure 19: Cell cycle analysis gating strategy. (a) Flow cytometry data was gated against FSC-A and SSC-A to separate cells from cellular debris. Next, single cell populations were obtained by gating FSC-H vs FSC-A **(b)**. Eventually, propidium iodide signal **(c)** was obtained by excitement with the BD Canto II 488 nm laser and detection with a 584/42 bandpass filter.

Table 17: Supplementary descriptive table. Selected genes from **Figure 6 c**

Gene Symbol	Full gene name
ANLN	Anillin, Actin-binding protein
MAD2L1	Mitotic Arrest Deficient 2-Like Protein 1
CDC20	Cell Division Cycle 20
CENPE	Centromere Protein P
TPX2	Targeting-protein for Xklp2, Microtubule Nucleation factor
KIF2C	Kinesin-like protein 2C
AURKA	Aurora kinase A
TIPIN	TIMELESS interacting protein
GINS1	GINS complex subunit 1
PLK1	Polo-like kinase 1
PRC1	Protein regulator of cytokinesis
CCNB1	Cyclin B1
AURKB	Aurora kinase B
TOP2A	DNA Topoisomerase II A
CLSPN	Claspin
ORC1	Origin recognition complex subunit 1
CDKN1B	Cyclin dependent kinase inhibitor 1B
PCNA	Proliferating cell nuclear antigen
BRCA1	Breast cancer type 1 susceptibility protein
POLA1	DNA polymerase alpha 1
MYC	Myc proto-oncogene protein
MCM10	Minichrosome maintenance 10 replication initiation factor

References

1. Bray, F. *et al.* Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA A Cancer J Clinicians* **74**, 229–263 (2024).
2. Song, Q., Merajver, S. D. & Li, J. Z. Cancer classification in the genomic era: five contemporary problems. *Hum Genomics* **9**, 27 (2015).
3. Swanton, C. *et al.* Embracing cancer complexity: Hallmarks of systemic disease. *Cell* **187**, 1589–1616 (2024).
4. Dalfovo, D., Scandino, R., Paoli, M., Valentini, S. & Romanel, A. Germline determinants of aberrant signaling pathways in cancer. *npj Precis. Onc.* **8**, 57 (2024).
5. Sanchez-Vega, F. *et al.* Oncogenic Signaling Pathways in The Cancer Genome Atlas. *Cell* **173**, 321-337.e10 (2018).
6. Neophytou, C. M., Trougakos, I. P., Erin, N. & Papageorgis, P. Apoptosis Deregulation and the Development of Cancer Multi-Drug Resistance. *Cancers* **13**, 4363 (2021).
7. Hanahan, D. Hallmarks of Cancer: New Dimensions. *Cancer Discovery* **12**, 31–46 (2022).
8. De Visser, K. E. & Joyce, J. A. The evolving tumor microenvironment: From cancer initiation to metastatic outgrowth. *Cancer Cell* **41**, 374–403 (2023).
9. Anand, U. *et al.* Cancer chemotherapy and beyond: Current status, drug candidates, associated risks and progress in targeted therapeutics. *Genes & Diseases* **10**, 1367–1401 (2023).
10. Deibler, R. W. & Kirschner, M. W. Quantitative Reconstitution of Mitotic CDK1 Activation in Somatic Cell Extracts. *Molecular Cell* **37**, 753–767 (2010).
11. Pellarin, I. *et al.* Cyclin-dependent protein kinases and cell cycle regulation in biology and disease. *Sig Transduct Target Ther* **10**, 11 (2025).
12. Nugroho, T. T. & Mendenhall, M. D. An Inhibitor of Yeast Cyclin-Dependent Protein Kinase Plays an Important Role in Ensuring the Genomic Integrity of Daughter Cells. *Molecular and Cellular Biology* **14**, 3320–3328 (1994).
13. McCusker, D. *et al.* Cdk1 coordinates cell-surface growth with the cell cycle. *Nat Cell Biol* **9**, 506–515 (2007).
14. Araujo, A. R., Gelens, L., Sheriff, R. S. M. & Santos, S. D. M. Positive Feedback Keeps Duration of Mitosis Temporally Insulated from Upstream Cell-Cycle Events. *Mol Cell* **64**, 362–375 (2016).
15. Schiff, P. B., Fant, J. & Horwitz, S. B. Promotion of microtubule assembly in vitro by taxol. *Nature* **277**, 665–667 (1979).

16. Mielgo, A., Torres, V. A., Clair, K., Barbero, S. & Stupack, D. G. Paclitaxel promotes a caspase 8-mediated apoptosis through death effector domain association with microtubules. *Oncogene* **28**, 3551–3562 (2009).
17. Hamm-Alvarez, S. F. *et al.* Coordinate depression of bradykinin receptor recycling and microtubule-dependent transport by taxol. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 7812–7816 (1994).
18. Nogales, E., Wolf, S. G. & Downing, K. H. Structure of the αβ tubulin dimer by electron crystallography. *Nature* **391**, 199–203 (1998).
19. Loibl, S., Poortmans, P., Morrow, M., Denkert, C. & Curigliano, G. Breast cancer. *The Lancet* **397**, 1750–1769 (2021).
20. Marupudi, N. I. *et al.* Paclitaxel: a review of adverse toxicities and novel delivery strategies. *Expert Opinion on Drug Safety* **6**, 609–621 (2007).
21. Mielke, S. *et al.* Comparative neurotoxicity of weekly non-break paclitaxel infusions over 1 versus 3 h: *Anti-Cancer Drugs* 785–792 (2003) doi:10.1097/00001813-200311000-00003.
22. Blander, J. M. & Medzhitov, R. Toll-dependent selection of microbial antigens for presentation by dendritic cells. *Nature* **440**, 808–812 (2006).
23. Pradeu, T., Thomma, B. P. H. J., Girardin, S. E. & Lemaitre, B. The conceptual foundations of innate immunity: Taking stock 30 years later. *Immunity* **57**, 613–631 (2024).
24. Lam, N., Lee, Y. & Farber, D. L. A guide to adaptive immune memory. *Nat Rev Immunol* **24**, 810–829 (2024).
25. Pishesha, N., Harmand, T. J. & Ploegh, H. L. A guide to antigen processing and presentation. *Nat Rev Immunol* **22**, 751–764 (2022).
26. Robinson, J. *et al.* IPD-IMGT/HLA Database. *Nucleic Acids Research* gkz950 (2019) doi:10.1093/nar/gkz950.
27. Kulski, J. K., Suzuki, S. & Shiina, T. Human leukocyte antigen super-locus: nexus of genomic supergenes, SNPs, indels, transcripts, and haplotypes. *Hum Genome Var* **9**, 49 (2022).
28. Lobos, C. A., Downing, J., D'Orsogna, L. J., Chatzileontiadou, D. S. M. & Gras, S. Protective HLA-B57: T cell and natural killer cell recognition in HIV infection. *Biochemical Society Transactions* **50**, 1329–1339 (2022).
29. Chowell, D. *et al.* Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy. *Science* **359**, 582–587 (2018).
30. Cruz, F. M., Colbert, J. D., Merino, E., Kriegsman, B. A. & Rock, K. L. The Biology and Underlying Mechanisms of Cross-Presentation of Exogenous Antigens on MHC-I Molecules. *Annu. Rev. Immunol.* **35**, 149–176 (2017).
31. Lopez, J. A. *et al.* Perforin forms transient pores on the target cell plasma membrane to facilitate rapid access of granzymes during killer cell attack. *Blood* **121**, 2659–2668 (2013).

32. O'Connell, A. R. & Stenson-Cox, C. A more serine way to die: Defining the characteristics of serine protease-mediated cell death cascades. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1773**, 1491–1499 (2007).
33. Goping, I. S. *et al.* Granzyme B-Induced Apoptosis Requires Both Direct Caspase Activation and Relief of Caspase Inhibition. *Immunity* **18**, 355–365 (2003).
34. Popović, B. *et al.* Time-dependent regulation of cytokine production by RNA binding proteins defines T cell effector function. *Cell Reports* **42**, 112419 (2023).
35. Shatsky, I. N., Terenin, I. M., Smirnova, V. V. & Andreev, D. E. Cap-Independent Translation: What's in a Name? *Trends in Biochemical Sciences* **43**, 882–895 (2018).
36. Brito Querido, J., Díaz-López, I. & Ramakrishnan, V. The molecular basis of translation initiation and its regulation in eukaryotes. *Nat Rev Mol Cell Biol* **25**, 168–186 (2024).
37. Holz, M. K., Ballif, B. A., Gygi, S. P. & Blenis, J. mTOR and S6K1 mediate assembly of the translation preinitiation complex through dynamic protein interchange and ordered phosphorylation events. *Cell* **184**, 2255 (2021).
38. Gingras, A.-C. *et al.* Regulation of 4E-BP1 phosphorylation: a novel two-step mechanism. *Genes & Development* **13**, 1422–1437 (1999).
39. Shuda, M. *et al.* CDK1 substitutes for mTOR kinase to activate mitotic cap-dependent protein translation. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 5875–5882 (2015).
40. Haneke, K. *et al.* CDK1 couples proliferation with protein synthesis. *Journal of Cell Biology* **219**, e201906147 (2020).
41. Sattlegger, E. Separate domains in GCN1 for binding protein kinase GCN2 and ribosomes are required for GCN2 activation in amino acid-starved cells. *The EMBO Journal* **19**, 6622–6633 (2000).
42. Tian, X. *et al.* Targeting the Integrated Stress Response in Cancer Therapy. *Front Pharmacol* **12**, 747837 (2021).
43. Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. S. & Weissman, J. S. Genome-Wide Analysis in Vivo of Translation with Nucleotide Resolution Using Ribosome Profiling. *Science* **324**, 218–223 (2009).
44. Andreev, D. E. *et al.* Translation of 5′ leaders is pervasive in genes resistant to eIF2 repression. *eLife* **4**, e03971 (2015).
45. Smith, J. E. *et al.* Translation of Small Open Reading Frames within Unannotated RNA Transcripts in Saccharomyces cerevisiae. *Cell Reports* **7**, 1858–1866 (2014).
46. Fresno, M., Jiménez, A. & Vázquez, D. Inhibition of Translation in Eukaryotic Systems by Harringtonine. *European Journal of Biochemistry* **72**, 323–330 (1977).
47. Tomuro, K. *et al.* Calibrated ribosome profiling assesses the dynamics of ribosomal flux on transcripts. *Nat Commun* **15**, 7061 (2024).

48. Loayza-Puch, F. *et al.* Tumour-specific proline vulnerability uncovered by differential ribosome codon reading. *Nature* **530**, 490–494 (2016).

49. VanInsberghe, M., van den Berg, J., Andersson-Rolf, A., Clevers, H. & van Oudenaarden, A. Single-cell Ribo-seq reveals cell cycle-dependent translational pausing. *Nature* **597**, 561–565 (2021).

50. Duré, C. *et al. In vivo* single-cell ribosome profiling reveals cell-type-specific translational programs during aging. Preprint at https://doi.org/10.1101/2024.11.02.621639 (2024).

51. Ingolia, N. T. Ribosome profiling: new views of translation, from single codons to genome scale. *Nat Rev Genet* **15**, 205–213 (2014).

52. Zhong, Y. *et al.* RiboDiff: detecting changes of mRNA translation efficiency from ribosome footprints. *Bioinformatics* **33**, 139–141 (2017).

53. Calviello, L., Sydow, D., Harnett, D. & Ohler, U. Ribo-seQC: comprehensive analysis of cytoplasmic and organellar ribosome profiling data. Preprint at https://doi.org/10.1101/601468 (2019).

54. Fields, A. P. *et al.* A Regression-Based Analysis of Ribosome-Profiling Data Reveals a Conserved Complexity to Mammalian Translation. *Molecular Cell* **60**, 816–827 (2015).

55. Ji, Z. RibORF: Identifying Genome-Wide Translated Open Reading Frames Using Ribosome Profiling. *CP Molecular Biology* **124**, e67 (2018).

56. Erhard, F. *et al.* Improved Ribo-seq enables identification of cryptic translation events. *Nat Methods* **15**, 363–366 (2018).

57. Erhard, F., Dölken, L., Schilling, B. & Schlosser, A. Identification of the Cryptic HLA-I Immunopeptidome. *Cancer Immunology Research* **8**, 1018–1026 (2020).

58. Marcu, A. *et al.* HLA Ligand Atlas: a benign reference of HLA-presented peptides to improve T-cell-based cancer immunotherapy. *J Immunother Cancer* **9**, e002071 (2021).

59. Ingolia, N. T. *et al.* Ribosome Profiling Reveals Pervasive Translation Outside of Annotated Protein-Coding Genes. *Cell Reports* **8**, 1365–1379 (2014).

60. Lin, Y. *et al.* Impacts of uORF codon identity and position on translation regulation. *Nucleic Acids Research* **47**, 9358–9367 (2019).

61. Yang, H., Li, Q., Stroup, E. K., Wang, S. & Ji, Z. Widespread stable noncanonical peptides identified by integrated analyses of ribosome profiling and ORF features. *Nat Commun* **15**, 1932 (2024).

62. Hinnebusch, A. G. TRANSLATIONAL REGULATION OF *GCN4* AND THE GENERAL AMINO ACID CONTROL OF YEAST. *Annu. Rev. Microbiol.* **59**, 407–450 (2005).

63. Hanson, G. & Coller, J. Codon optimality, bias and usage in translation and mRNA decay. *Nat Rev Mol Cell Biol* **19**, 20–30 (2018).

64. Manjunath, H. *et al.* Suppression of Ribosomal Pausing by eIF5A Is Necessary to Maintain the Fidelity of Start Codon Selection. *Cell Reports* **29**, 3134-3146.e6 (2019).
65. Diba, F., Watson, C. S. & Gametchu, B. 5'UTR sequences of the glucocorticoid receptor 1A transcript encode a peptide associated with translational regulation of the glucocorticoid receptor. *J Cell Biochem* **81**, 149–161 (2001).
66. Chen, J. *et al.* Pervasive functional translation of noncanonical human open reading frames. *Science* **367**, 1140–1146 (2020).
67. Chong, C. *et al.* Integrated proteogenomic deep sequencing and analytics accurately identify non-canonical peptides in tumor immunopeptidomes. *Nat Commun* **11**, 1293 (2020).
68. Ouspenskaia, T. *et al.* Unannotated proteins expand the MHC-I-restricted immunopeptidome in cancer. *Nat Biotechnol* **40**, 209–217 (2022).
69. Bartok, O. *et al.* Anti-tumour immunity induces aberrant peptide presentation in melanoma. *Nature* **590**, 332–337 (2021).
70. International Human Genome Sequencing Consortium *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
71. Shapiro, I. E. & Bassani-Sternberg, M. The impact of immunopeptidomics: From basic research to clinical implementation. *Seminars in Immunology* **66**, 101727 (2023).
72. Salek, M. *et al.* optiPRM: A Targeted Immunopeptidomics LC-MS Workflow With Ultra-High Sensitivity for the Detection of Mutation-Derived Tumor Neoepitopes From Limited Input Material. *Molecular & Cellular Proteomics* **23**, 100825 (2024).
73. Stopfer, L. E., Mesfin, J. M., Joughin, B. A., Lauffenburger, D. A. & White, F. M. Multiplexed relative and absolute quantitative immunopeptidomics reveals MHC I repertoire alterations induced by CDK4/6 inhibition. *Nat Commun* **11**, 2760 (2020).
74. Becker, J. P. *et al.* NMD inhibition by 5-azacytidine augments presentation of immunogenic frameshift-derived neoepitopes. *iScience* **24**, 102389 (2021).
75. Becker, J. P. & Riemer, A. B. The Importance of Being Presented: Target Validation by Immunopeptidomics for Epitope-Specific Immunotherapies. *Front. Immunol.* **13**, 883989 (2022).
76. Huber, F. *et al.* A comprehensive proteogenomic pipeline for neoantigen discovery to advance personalized cancer immunotherapy. *Nat Biotechnol* (2024) doi:10.1038/s41587-024-02420-y.
77. Bassani-Sternberg, M. *et al.* Direct identification of clinically relevant neoepitopes presented on native human melanoma tissue by mass spectrometry. *Nat Commun* **7**, 13404 (2016).
78. Shraibman, B., Kadosh, D. M., Barnea, E. & Admon, A. Human Leukocyte Antigen (HLA) Peptides Derived from Tumor Antigens Induced by Inhibition of DNA Methylation for

Development of Drug-facilitated Immunotherapy. *Molecular & Cellular Proteomics* **15**, 3058–3070 (2016).

79. Murphy, J. P. *et al.* Multiplexed Relative Quantitation with Isobaric Tagging Mass Spectrometry Reveals Class I Major Histocompatibility Complex Ligand Dynamics in Response to Doxorubicin. *Anal. Chem.* **91**, 5106–5115 (2019).

80. Goyal, A. *et al.* DNMT and HDAC inhibition induces immunogenic neoantigens from human endogenous retroviral element-derived transcripts. *Nat Commun* **14**, 6731 (2023).

81. Rojas, L. A. *et al.* Personalized RNA neoantigen vaccines stimulate T cells in pancreatic cancer. *Nature* **618**, 144–150 (2023).

82. Giotti, B., Joshi, A. & Freeman, T. C. Meta-analysis reveals conserved cell cycle transcriptional network across multiple human cell types. *BMC Genomics* **18**, 30 (2017).

83. Jansen, K. I., Iwanski, M. K., Burute, M. & Kapitein, L. C. A live-cell marker to visualize the dynamics of stable microtubules throughout the cell cycle. *Journal of Cell Biology* **222**, e202106105 (2023).

84. Mutso, M. *et al.* The methyltransferase N6AMT1 participates in the cell cycle by regulating cyclin E levels. *PLoS ONE* **19**, e0298884 (2024).

85. Kowar, A. *et al.* Upstream open reading frame translation enhances immunogenic peptide presentation in mitotically arrested cancer cells. Preprint at https://doi.org/10.1101/2025.01.08.631915 (2025).

86. Hendzel, M. J. *et al.* Mitosis-specific phosphorylation of histone H3 initiates primarily within pericentromeric heterochromatin during G2 and spreads in an ordered fashion coincident with mitotic chromosome condensation. *Chromosoma* **106**, 348–360 (1997).

87. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* **11**, R25 (2010).

88. Zhao, Y. *et al.* TPM, FPKM, or Normalized Counts? A Comparative Study of Quantification Measures for the Analysis of RNA-seq Data from the NCI Patient-Derived Models Repository. *J Transl Med* **19**, 269 (2021).

89. Prensner, J. R. *et al.* What Can Ribo-Seq, Immunopeptidomics, and Proteomics Tell Us About the Noncanonical Proteome? *Molecular & Cellular Proteomics* **22**, 100631 (2023).

90. Poznański, J. *et al.* Global pentapeptide statistics are far away from expected distributions. *Sci Rep* **8**, 15178 (2018).

91. Marash, L. *et al.* DAP5 Promotes Cap-Independent Translation of Bcl-2 and CDK1 to Facilitate Cell Survival during Mitosis. *Molecular Cell* **30**, 447–459 (2008).

92. Olan, I. *et al.* HMGA1 orchestrates chromatin compartmentalization and sequesters genes into 3D networks coordinating senescence heterogeneity. *Nat Commun* **15**, 6891 (2024).

93. Sgubin, M. *et al.* HMGA1 positively regulates the microtubule-destabilizing protein stathmin promoting motility in TNBC cells and decreasing tumour sensitivity to paclitaxel. *Cell Death Dis* **13**, 429 (2022).
94. Dersh, D., Yewdell, J. W. & Wei, J. A SIINFEKL-Based System to Measure MHC Class I Antigen Presentation Efficiency and Kinetics. in *Antigen Processing* (ed. Van Endert, P.) vol. 1988 109–122 (Springer New York, New York, NY, 2019).
95. Imami, K. *et al.* Phosphorylation of the Ribosomal Protein RPL12/uL11 Affects Translation during Mitosis. *Molecular Cell* **72**, 84-98.e9 (2018).
96. Choi, H. J., Fukui, M. & Zhu, B. T. Role of Cyclin B1/Cdc2 Up-Regulation in the Development of Mitotic Prometaphase Arrest in Human Breast Cancer Cells Treated with Nocodazole. *PLoS ONE* **6**, e24312 (2011).
97. Clemm Von Hohenberg, K. *et al.* Cyclin B/CDK1 and Cyclin A/CDK2 phosphorylate DENR to promote mitotic protein translation and faithful cell division. *Nat Commun* **13**, 668 (2022).
98. Bohlen, J. *et al.* DENR promotes translation reinitiation via ribosome recycling to drive expression of oncogenes including ATF4. *Nat Commun* **11**, 4676 (2020).
99. Petrone, A., Adamo, M. E., Cheng, C. & Kettenbach, A. N. Identification of Candidate Cyclin-dependent kinase 1 (Cdk1) Substrates in Mitosis by Quantitative Phosphoproteomics. *Molecular & Cellular Proteomics* **15**, 2448–2461 (2016).
100. Warfel, N. A., Dolloff, N. G., Dicker, D. T., Malysz, J. & El-Deiry, W. S. CDK1 stabilizes HIF-1α via direct phosphorylation of Ser668 to promote tumor growth. *Cell Cycle* **12**, 3689–3701 (2013).
101. Matthess, Y., Raab, M., Sanhaji, M., Lavrik, I. N. & Strebhardt, K. Cdk1/Cyclin B1 Controls Fas-Mediated Apoptosis by Regulating Caspase-8 Activity. *Molecular and Cellular Biology* **30**, 5726–5740 (2010).
102. Tanenbaum, M. E., Stern-Ginossar, N., Weissman, J. S. & Vale, R. D. Regulation of mRNA translation during mitosis. *eLife* **4**, e07957 (2015).
103. Stumpf, C. R., Moreno, M. V., Olshen, A. B., Taylor, B. S. & Ruggero, D. The Translational Landscape of the Mammalian Cell Cycle. *Molecular Cell* **52**, 574–582 (2013).
104. Tong, G., Hah, N. & Martinez, T. F. Comparison of software packages for detecting unannotated translated small open reading frames by Ribo-seq. *Briefings in Bioinformatics* **25**, bbae268 (2024).
105. Cockman, E., Anderson, P. & Ivanov, P. TOP mRNPs: Molecular Mechanisms and Principles of Regulation. *Biomolecules* **10**, 969 (2020).
106. Papst, P. J. *et al.* Cdc2-Cyclin B Phosphorylates p70 S6 Kinase on Ser411at Mitosis. *Journal of Biological Chemistry* **273**, 15077–15084 (1998).

107. Tcherkezian, J. *et al.* Proteomic analysis of cap-dependent translation identifies LARP1 as a key regulator of 5′TOP mRNA translation. *Genes Dev.* **28**, 357–371 (2014).
108. Cui, P. *et al.* A comparison between ribo-minus RNA-sequencing and polyA-selected RNA-sequencing. *Genomics* **96**, 259–265 (2010).
109. Kozak, M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucl Acids Res* **15**, 8125–8148 (1987).
110. Acevedo, J. M., Hoermann, B., Schlimbach, T. & Teleman, A. A. Changes in global translation elongation or initiation rates shape the proteome via the Kozak sequence. *Sci Rep* **8**, 4018 (2018).
111. Lyu, X., Yang, Q., Zhao, F. & Liu, Y. Codon usage and protein length-dependent feedback from translation elongation regulates translation initiation and elongation speed. *Nucleic Acids Research* **49**, 9404–9423 (2021).
112. Wei, J., Wu, C. & Sachs, M. S. The Arginine Attenuator Peptide Interferes with the Ribosome Peptidyl Transferase Center. *Molecular and Cellular Biology* **32**, 2396–2406 (2012).
113. Qiu, Y. *et al.* The GC-content at the 5′ ends of human protein-coding genes is undergoing mutational decay. *Genome Biol* **25**, 219 (2024).
114. Pelletier, J. & Sonenberg, N. Insertion mutagenesis to increase secondary structure within the 5′ noncoding region of a eukaryotic mRNA reduces translational efficiency. *Cell* **40**, 515–526 (1985).
115. Weber, R. *et al.* DAP5 enables main ORF translation on mRNAs with structured and uORF-containing 5′ leaders. *Nat Commun* **13**, 7510 (2022).
116. Jürgens, L. & Wethmar, K. The Emerging Role of uORF-Encoded uPeptides and HLA uLigands in Cellular and Tumor Biology. *Cancers* **14**, 6031 (2022).
117. Sanchez-Mazas, A. *et al.* The most frequent HLA alleles around the world: A fundamental synopsis. *Best Practice & Research Clinical Haematology* **37**, 101559 (2024).
118. Tynan, F. E. *et al.* A T cell receptor flattens a bulged antigenic peptide presented by a major histocompatibility complex class I molecule. *Nat Immunol* **8**, 268–276 (2007).
119. Probst-Kepper, M. *et al.* Conformational Restraints and Flexibility of 14-Meric Peptides in Complex with HLA-B*3501. *The Journal of Immunology* **173**, 5610–5616 (2004).
120. Benlalam, H. *et al.* Identification of Five New HLA-B*3501-Restricted Epitopes Derived from Common Melanoma-Associated Antigens, Spontaneously Recognized by Tumor-Infiltrating Lymphocytes. *The Journal of Immunology* **171**, 6283–6289 (2003).
121. Komov, L., Melamed Kadosh, D., Barnea, E. & Admon, A. The Effect of Interferons on Presentation of Defective Ribosomal Products as HLA Peptides. *Molecular & Cellular Proteomics* **20**, 100105 (2021).
122. Ruiz Cuevas, M. V. *et al.* Most non-canonical proteins uniquely populate the proteome or immunopeptidome. *Cell Reports* **34**, 108815 (2021).

123. Mohammed, F. *et al.* Phosphorylation-dependent interaction between antigenic peptides and MHC class I: a molecular basis for the presentation of transformed self. *Nat Immunol* **9**, 1236–1243 (2008).
124. Depontieu, F. R. *et al.* Identification of tumor-associated, MHC class II-restricted phosphopeptides as targets for immunotherapy. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 12073–12078 (2009).
125. Cobbold, M. *et al.* MHC Class I–Associated Phosphopeptides Are the Targets of Memory-like Immunity in Leukemia. *Sci. Transl. Med.* **5**, (2013).
126. Vennin, C. *et al.* Taxanes trigger cancer cell killing in vivo by inducing non-canonical T cell cytotoxicity. *Cancer Cell* **41**, 1170-1185.e12 (2023).
127. Lau, T. S. *et al.* Paclitaxel Induces Immunogenic Cell Death in Ovarian Cancer via TLR4/IKK2/SNARE-Dependent Exocytosis. *Cancer Immunology Research* **8**, 1099–1111 (2020).
128. Wanderley, C. W. *et al.* Paclitaxel Reduces Tumor Growth by Reprogramming Tumor-Associated Macrophages to an M1 Profile in a TLR4-Dependent Manner. *Cancer Research* **78**, 5891–5900 (2018).
129. Haj-Shomaly, J. *et al.* T Cells Promote Metastasis by Regulating Extracellular Matrix Remodeling following Chemotherapy. *Cancer Research* **82**, 278–291 (2022).
130. Dai, Q. *et al.* Excipient of paclitaxel induces metabolic dysregulation and unfolded protein response. *iScience* **24**, 103170 (2021).
131. He, F. *et al.* Adverse Event Profile for Nanoparticle Albumin-Bound Paclitaxel Compared With Solvent-Based Taxanes in Solid-Organ Tumors: A Systematic Review and Meta-Analysis of Randomized Clinical Trials. *Ann Pharmacother* **56**, 898–909 (2022).
132. Rodriguez-Meira, A. *et al.* Unravelling Intratumoral Heterogeneity through High-Sensitivity Single-Cell Mutational Analysis and Parallel RNA Sequencing. *Molecular Cell* **73**, 1292-1305.e8 (2019).
133. Marot-Lassauzaie, V. *et al.* Identifying cancer cells from calling single-nucleotide variants in scRNA-seq data. *Bioinformatics* **40**, btae512 (2024).
134. Guzmán, C., Bagga, M., Kaur, A., Westermarck, J. & Abankwa, D. ColonyArea: An ImageJ Plugin to Automatically Quantify Colony Formation in Clonogenic Assays. *PLoS ONE* **9**, e92444 (2014).
135. Livak, K. J. & Schmittgen, T. D. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods* **25**, 402–408 (2001).
136. Chong, C. *et al.* High-throughput and Sensitive Immunopeptidomics Platform Reveals Profound Interferonγ-Mediated Remodeling of the Human Leukocyte Antigen (HLA) Ligandome. *Molecular & Cellular Proteomics* **17**, 533–548 (2018).

137. Bruderer, R. *et al.* Extending the Limits of Quantitative Proteome Profiling with Data-Independent Acquisition and Application to Acetaminophen-Treated Three-Dimensional Liver Microtissues. *Molecular & Cellular Proteomics* **14**, 1400–1410 (2015).
138. Reynisson, B., Alvarez, B., Paul, S., Peters, B. & Nielsen, M. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Research* **48**, W449–W454 (2020).
139. Andreatta, M., Alvarez, B. & Nielsen, M. GibbsCluster: unsupervised clustering and alignment of peptide sequences. *Nucleic Acids Research* **45**, W458–W463 (2017).
140. Bouwmeester, R., Gabriels, R., Hulstaert, N., Martens, L. & Degroeve, S. DeepLC can predict retention times for peptides that carry as-yet unseen modifications. *Nat Methods* **18**, 1363–1369 (2021).
141. Pino, L. K. *et al.* The Skyline ecosystem: Informatics for quantitative mass spectrometry proteomics. *Mass Spectrometry Reviews* **39**, 229–244 (2020).
142. Gessulat, S. *et al.* Prosit: proteome-wide prediction of peptide tandem mass spectra by deep learning. *Nat Methods* **16**, 509–518 (2019).
143. Toprak, U. H. *et al.* Conserved Peptide Fragmentation as a Benchmarking Tool for Mass Spectrometers and a Discriminating Feature for Targeted Proteomics. *Molecular & Cellular Proteomics* **13**, 2056–2071 (2014).
144. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet j.* **17**, 10 (2011).
145. Smith, T., Heger, A. & Sudbery, I. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res.* **27**, 491–499 (2017).
146. Kent, W. J. BLAT —The BLAST -Like Alignment Tool. *Genome Res.* **12**, 656–664 (2002).
147. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
148. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
149. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
150. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR : a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
151. Lauria, F. *et al.* riboWaltz: Optimization of ribosome P-site positioning in ribosome profiling data. *PLoS Comput Biol* **14**, e1006169 (2018).

152. Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation* **2**, 100141 (2021).

153. Chen, E. Y. *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128 (2013).

154. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).

155. Kuhn, R. M., Haussler, D. & Kent, W. J. The UCSC genome browser and associated tools. *Briefings in Bioinformatics* **14**, 144–161 (2013).

156. Shen, W., Le, S., Li, Y. & Hu, F. SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. *PLoS ONE* **11**, e0163962 (2016).