Towards HERV-H Envelope-Based Tumor Stratification and Intervention Strategies

DISSERTATION ZUR ERLANGUNG DES
DOKTORGRADES DER NATURWISSENSCHAFTEN (DR. RER. NAT.)
DER FAKULTÄT FÜR BIOLOGIE UND VORKLINISCHE MEDIZIN
DER UNIVERSITÄT REGENSBURG

vorgelegt von
Jasmin Gille
aus
Bad Saarow
im Jahr
2024

Towards HERV-H Envelope-Based Tumor Stratification and Intervention Strategies

DISSERTATION ZUR ERLANGUNG DES
DOKTORGRADES DER NATURWISSENSCHAFTEN (DR. RER. NAT.)
DER FAKULTÄT FÜR BIOLOGIE UND VORKLINISCHE MEDIZIN
DER UNIVERSITÄT REGENSBURG

vorgelegt von
Jasmin Gille
aus
Bad Saarow
im Jahr
2024

Das Promotionsgesuch wurde eingereicht am:

18.12.2024

Die Arbeit wurde angeleitet von:
Prof. Dr. Ralf Wagner

Unterschrift:

Jasmin Gille

Table of contents

Table of Contents

List of Abbreviations

%	percent
Δ	delta
°C	degree Celsius
µg	microgram
µl	microliter
µm	micrometer
µM	micromolar
Amp	Ampicillin
BN PAGE	blue native polyacrylamide gel electrophoresis
bp	base pair
BSA	Bovine serum albumin
CAR	Chimeric antigen receptor
CIP	Calf intestinal alkaline phosphatase
cm^2	Square centimeter
CMV	Cytomegalovirus
CRC	Colorectal cancer
C_T	Threshold cycle
CV	Column volume
Cyt	Cytoplasmic tail
DAB	3,3'-diaminobenzidine
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide Solution Mix (equimolar)
DPBS	Dulbecco's phosphate-buffered saline
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EBOV	Ebola virus
EC50	Half-maximal effective concentration
EDTA	Ethylenediaminetetraacetic acid
EMT	Epithelial-to-mesenchymal transition
env	Envelope glycoprotein
FACS	Fluorescence-activated cell sorting
FC	Flow cytometry
FCS	Fetal calf serum
FFPE	Formalin-fixed paraffin-embedded
FP	Fusion peptide
fwd	forward
Gag	Group-specific antigen
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
gDNA	Genomic DNA
GP	Glycoprotein
h	hour(s)
H&E	hematoxylin and eosin
HA	hemagglutinin
HEK293 T	Human Embryonic Kidney cell line T antigen
HERV	Human endogenous retrovirus
HIER	Heat-induced epitope retrieval
HNSCC	Head and neck squamous cell carcinoma
HRP	Horseradish peroxidase
ICC	Immunocytochemistry
IgG	Immunoglobulin G
IgM	Immunoglobulin M
IHC	Immunohistochemistry
IMAC	Immobilized metal affinity chromatography
int	Integrase

IS	Immune staining
ISD	Immunosuppressive domain
ISDmut	Mutated immunosuppressive domain
Kan	Kanamycin
kb	kilobase
kDa	Kilo Dalton
LASV	Lassa virus
LB	Lysogeny Broth
LNP	Lipid nanoparticle
LTR	Long terminal repeat
M	molar
mA	milliampere
mAb	Monoclonal antibody
MARV	Marburgvirus
MFI	Mean fluorescence intensity
min	minute(s)
ml	milliliter
mM	millimolar
MMTV	Mouse mammary tumor virus
MS	Multiple sclerosis
MSRV	Multiple sclerosis-associated retrovirus
MVA	Modified vaccinia Ankara
ng	nanogram
nm	nanometer
nM	nanomolar
nt	nucleotide
OD	Optical density
ORF	Open reading frame
pAb	Polyclonal antibody
PBMC	Peripheral blood mononuclear cell
PBS	Phosphate-buffered saline
U	Units
PBS	Primer binding site
PCR	Polymerase chain reaction
PE	Phycoerythrin
PEI	Polyethyleneimine
Pen/Strep	Penicillin/streptomycin
pH	*potentia hydrogenii*
POD	Peroxidase
pol	polymerase
PPT	Polypurine tract
pro	protease
rev	reverse
Rn	Normalized reporter signal
rpm	revolutions per minute
RPMI	Roswell Park Memorial Institute
RT	Room temperature
RT-PCR	Reverse transcriptase polymerase chain reaction
RT-qPCR	Reverse transcriptase quantitative polymerase chain reaction
SDS PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
sec	second(s)
SEC	Size exclusion chromatography
shRNA	Small hairpin RNA
SP	Signal peptide
SU	Surface subunit
TBE	Tris-borate-EDTA buffer
TBS	Tris-buffered saline
TCA	Trichloroacetic acid
TE	Transposable elements
TM	Transmembrane domain
TMA	Tissue microarray
TMB	3,3',5,5'-Tetramethylbenzidin
tPA	Tissue plasminogen activator

List of Abbreviations

UV	Ultraviolet
V	Volt
v/v	volume per volume
V_H	variable heavy
V_L	variable light
VLP	Virus-like particle
VSV	Vesicular stomatitis virus
VSV-G	Vesicular stomatitis virus glycoprotein
w/v	weight per volume
WB	Western blot
WT	Wild type
x g	relative centrifugal force

1 Abstract

Human endogenous retroviruses (HERV) are retroviral fossils encoded within the human genome. They constitute for about 8 % of all sequences in the genome, and even though most HERV elements are highly depleted and inactive, HERV re-activation has been repeatedly associated with the development and progression of different kinds of cancer. Especially enhanced transcription and overexpression of their envelope proteins (env) has been observed in tumor tissues and human tumor cell lines, making HERV env an interesting new target for cancer immunotherapy.

Env protein sequences of 5 HERV families, namely HERV-W, -FRD, -H, -K and ERV3.1 have been characterized for their protein expression properties and showed stable expression for all proteins. Except for ERV3.1 and HERV-W, modifications within the signal peptide and transmembrane-anchored domain including the cytoplasmic tail did not affect protein expression levels. Furthermore, cell-cell fusion activity was observed for HERV-FRD and HERV-W, which could be abolished for HERV-FRD by introducing mutations knocking out specific N-glycosylation sites previously described in the literature. Therefore, protein expressing env sequences could be provided for further vaccination strategies to our cooperation partners for all HERV families tested. Within this thesis, all the following work was solely focused on HERV-H, the most abundant HERV family in the human genome.

In the second part of this work, a suite of env-based antigens from the HERV-H family was designed as membrane-anchored and soluble proteins. Therefore, modifications of the transmembrane- and the putative immunosuppressive domain were introduced to the env protein. The antigens were administered to BALB/c mice in either a heterologous DNA-prime, protein-boost, or a homologous protein prime-boost regimen to assess their immunogenic properties. All env variants elicited specific antibody responses and almost no differences were seen between the groups. An effect of the ISD, or a presumably inactivating mutation thereof, on the humoral immune response could not be observed in the mouse model. Overall, it could be demonstrated that the HERV-H env protein is a suitable antigen that induces a strong, specific humoral response in mice.

Furthermore, while it was not possible to generate HERV-H env specific monoclonal antibodies (mAbs) via the hybridoma approach, polyclonal antibodies (pAbs) purified from immunized

rabbits were able to specifically recognize the env protein in ELISA, western blot and immunocytochemistry (ICC) assays.

The last part of this thesis was dedicated to establishing diagnostic tools screening human tumor cells for HERV-H env transcription and expression. RT-PCR and RT-qPCR assays could detect an approximately 600-bp fragment of the *env* 3' part in all human cell lines tested. Furthermore, the combination of western blot and an ICC assay with formalin-fixed paraffin-embedded (FFPE) human tumor cell lines could prove HERV-H env protein expression to some extent. However, all assays showed limitations regarding the proof of actual cell-surface protein expression. A sequencing approach from isolated genomic DNA of selected cell lines showed, that at least one of the cell lines (MCF-7) had the genetic potential to express a complete full-length env protein. However, the overall ratio of incomplete-to-complete *env* sequences was very high. Additionally, treatment with the proteasome-inhibitor Bortezomib led to enhanced HERV-H env protein levels, suggesting that most env proteins get degraded after expression, presumably due to incorrect folding.

The established ICC assay was further applied to tumorous and normal colorectal tissues to show HERV-H env protein expression in colorectal cancer (CRC) tissue, but the results remained inconclusive.

Overall, these data demonstrate that while there is strong evidence of HERV-H env transcription and expression in tumor cells on both mRNA and protein levels, it remains unclear if these proteins are structural integer and presented on the cell surface of tumor tissue, which would be a prerequisite for an antibody-based cancer immunotherapy targeting cell-surface displayed HERV-H env.

2 Zusammenfassung

Humane endogene Retroviren (HERV) sind retrovirale Fossilien, die im menschlichen Genom kodiert sind. Sie machen etwa 8 % aller Sequenzen im Genom aus, und obwohl die meisten HERV-Elemente stark verändert und inaktiv sind, wurde die Reaktivierung von HERV wiederholt mit der Entstehung und dem Fortschreiten verschiedener Krebsarten in Verbindung gebracht. Insbesondere eine verstärkte Transkription und Überexpression ihrer Hüllproteine (env) wurde in Tumorgeweben und menschlichen Tumorzelllinien beobachtet, was HERV-env zu einem interessanten neuen Ziel für die Krebsimmuntherapie macht.

Env-Proteinsequenzen von 5 HERV-Familien, nämlich HERV-W, -FRD, -H, -K und ERV3.1, wurden hinsichtlich ihrer Proteinexpressionseigenschaften charakterisiert und zeigten eine stabile Expression für alle Proteine. Mit Ausnahme von ERV3.1 und HERV-W hatten Modifikationen innerhalb des Signalpeptids und der transmembranverankerten Domäne einschließlich des zytoplasmatischen Schwanzes keinen Einfluss auf die Proteinexpressionsniveaus. Darüber hinaus wurde für HERV-FRD und HERV-W eine Zell-Zell-Fusionsaktivität beobachtet, die für HERV-FRD durch die Einführung von Mutationen, die bestimmte in der Literatur beschriebene N-Glykosylierungsstellen ausschalten, aufgehoben werden konnte. Daher konnten unseren Kooperationspartnern für alle getesteten HERV-Familien Protein-exprimierende env-Sequenzen für weitere Impfstrategien zur Verfügung gestellt werden. Im Rahmen dieser Arbeit konzentrierten sich alle folgenden Arbeiten ausschließlich auf HERV-H, die im menschlichen Genom am häufigsten vorkommende HERV-Familie.

Im zweiten Teil dieser Arbeit wurde eine Reihe env-basierter Antigene aus der HERV-H-Familie als membranverankerte und lösliche Proteine entwickelt. Dafür wurden Modifikationen der transmembran- und der vermeintlich immunsuppressiven Domäne in das env-Protein eingeführt. Die Antigene wurden BALB/c-Mäusen entweder in einem heterologen DNA-Prime-Protein-Boost- oder einem homologen Protein-Prime-Boost-Schema verabreicht, um ihre immunogenen Eigenschaften zu beurteilen. Alle env-Varianten lösten spezifische Antikörperreaktionen aus und es wurden fast keine Unterschiede zwischen den Gruppen festgestellt. Eine Auswirkung der ISD oder einer vermutlich inaktivierenden Mutation davon auf die humorale Immunreaktion konnte im Mausmodell nicht beobachtet werden. Insgesamt

konnte gezeigt werden, dass das HERV-H-env-Protein ein geeignetes Antigen ist, das bei Mäusen eine starke, spezifische humorale Reaktion hervorruft.

Während es nicht möglich war, HERV-H-env-spezifische monoklonale Antikörper (mAbs) über den Hybridomansatz zu erzeugen, konnten polyklonale Antikörper (pAbs), die aus immunisierten Kaninchen isoliert wurden, das env-Protein in ELISA-, Western Blot- und Immunzytochemie-(ICC)-Tests spezifisch erkennen.

Der letzte Teil dieser Arbeit widmete sich der Entwicklung diagnostischer Plattformen zum Screening menschlicher Tumorzellen auf HERV-H-env-Transkription und -Expression. RT-PCR- und RT-qPCR-Tests konnten in allen getesteten menschlichen Zelllinien ein etwa 600-bp-Fragment des env-3'-Teils nachweisen. Darüber hinaus konnte die Kombination aus Western Blot und einem ICC-Test mit Formalin-fixierten, in Paraffin eingebetteten (FFPE) menschlichen Tumorzelllinien die Expression des HERV-H-env-Proteins bis zu einem gewissen Grad nachweisen. Alle Tests zeigten jedoch Einschränkungen hinsichtlich des Nachweises der tatsächlichen Expression von Zelloberflächenproteinen. Ein Sequenzierungsansatz aus isolierter genomischer DNA ausgewählter Zelllinien zeigte, dass mindestens eine der Zelllinien (MCF-7) das genetische Potenzial hatte, ein vollständiges env-Protein in voller Länge zu exprimieren. Das Gesamtverhältnis unvollständiger zu vollständigen env-Sequenzen war jedoch sehr hoch. Darüber hinaus führte die Behandlung mit dem Proteasom-Inhibitor Bortezomib zu erhöhten HERV-H-env-Proteinspiegeln, was darauf hindeutet, dass die meisten env-Proteine nach der Expression abgebaut werden, vermutlich aufgrund falscher Faltung. Der etablierte ICC-Test wurde außerdem auf tumoröses und normales kolorektales Gewebe angewendet, um die Expression des HERV-H-env-Proteins in kolorektalem Krebsgewebe (CRC) nachzuweisen. Die Ergebnisse blieben jedoch unschlüssig.

Insgesamt zeigen diese Daten, dass es zwar starke Hinweise auf die Transkription und Expression von HERV-H-env in Tumorzellen sowohl auf mRNA- als auch auf Proteinebene gibt, es jedoch unklar bleibt, ob diese Proteine strukturell intakt sind und auf der Zelloberfläche des Tumorgewebes vorhanden sind, was eine Voraussetzung für eine antikörperbasierte Krebsimmuntherapie wäre, die auf ein auf der Zelloberfläche vorhandenes HERV-H-env abzielt.

3 Introduction

3.1 Human Endogenous Retroviruses

When the Human Genome Project unraveled the complete sequence of the human genome, it was shown that around 45 % of the genome consists of transposable elements (TEs) with a total of 8 % belonging to the class of LTR retro-transposons, also called human endogenous retroviruses (HERV) (Lander *et al.*, 2001). These elements were first integrated into the human genome several million years ago through multiple infections of the germ line by now-extinct exogenous retroviruses (Bannert and Kurth, 2024; Katzourakis, Rambaut and Pybus, 2005). Retroviruses, which are single-stranded positive-sense RNA viruses utilize two enzymes to integrate their viral DNA into the host genome upon infection. First, the viral RNA is converted to DNA via a reverse transcriptase, then the DNA is integrated into the host genome via a viral integrase. The integration into the germ line of primates allowed for Mendelian inheritance of these HERV proviruses through the offspring, enabling stable integration into the human genome (Bannert and Kurth, 2024).

Since HERVs derive from these ancient retroviral infections, they show a similar genomic structure as exogenous retroviruses (Figure 1). A structurally complete HERV consists of two long terminal repeats (LTR), constituting the 5’ and the 3’ parts of the genome. The internal part between the LTRs comprises 4 viral genes: *gag*, *pro*, *pol* and *env*. *Gag*, short for group-specific antigen, encodes the structural proteins matrix, capsid and nucleocapsid. *Pro* and *pol* encode the viral enzymes protease, the ribonuclease H and the before-mentioned reverse transcriptase and integrase. Lastly, *env* encodes the envelope surface protein (env), which will be discussed in more detail in section 3.3. Additionally, between the 5’LTR and *gag* a primer binding site (PBS) can be found as well as a polypurine tract (PPT) between *env* and the 3’LTR. During the reverse transcription, the PBS serves as binding site for the cellular tRNA priming the (-)strand DNA synthesis, while the PPT is used as a primer for the (+)strand DNA synthesis (Grandi and Tramontano, 2018a).

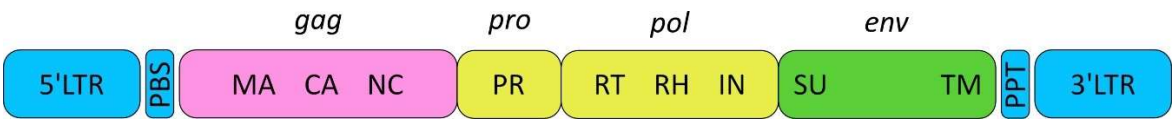


Figure 1: General genomic HERV structure.

Schematic view of the HERV genomic structure. The 4 viral genes *gag* (encoding for matrix (MA), capsid (CA) and nucleocapsid (NC)), *pro* (encoding for the protease (PR)), *pol* (encoding for the reverse transcriptase (RT), ribonuclease H

(RH) and integrase (IN)) and *env* (encoding for the surface (SU) and transmembrane (TM) subunits of the envelope protein) are flanked by a 5' and 3' long terminal repeat (LTR). A primer binding site (PBS) is located between the 5'LTR and *gag,* whereas the polypurine tract (PPT) is located between *env* and the 3'LTR. Adapted from Grandi and Tramontano, 2018a. Illustration created using Affinity Designer 2.5.2

Due to the occurrence of HERVs in the human genome for several million years, many sequences acquired inactivating mutations caused by the cellular editing system and the prolonged exposure to the host genome substitution rate (Grandi and Tramontano, 2018a). Additionally, homologous recombination can occur between the two LTRs of proviruses in different loci, leading to large deletions within the HERV sequence up to the formation of so-called solitary LTRs, where the complete internal HERV part is deleted (Sverdlov, 1998; Stoye, 2001). Most of the still intact HERV elements are transcriptionally silenced through CpG methylation and are therefore also inactive (Lavie *et al.*, 2005; Alcazer, Bonaventura and Depil, 2020). Nevertheless, some HERVs retain protein-coding functions and get expressed under certain circumstances. The reactivation of HERV expression is most likely induced by several environmental stimuli such as hormones, cytokines, epigenetic modulation through for example UV radiation and infections with microorganisms like herpesviruses or *Toxoplasma gondii* (Matteucci *et al.*, 2018). The expression products of HERVs partly play important roles in human physiology and pathology (as reviewed by Griffiths, 2001).

3.2 HERV Classification

The nomenclature and classification of HERVs have been following various methodologies over the years leading to inconsistency and confusion in the field. For example, HERV families have been named according to the cellular tRNA recognized by their PBS (HERV-H for histidine, HERV-W for tryptophan, HERV-K for lysine), by neighboring genes like HERV-ADP or a particular motif like in the case of HERV-FRD (as reviewed by Grandi and Tramontano, 2018a). In the most recent attempts of classification, HERVs are broadly divided into three main classes based on their similarity to existing exogenous retroviruses. In that matter, class I harbors HERVs showing similarity to *Gamma*- and *Epsilonretroviruses*, class II consists of *Betaretrovirus*-like HERVs and class III HERVs show similarity to *Spumaretroviruses* (Figure 2)(Jern, Sperber and Blomberg, 2005). This classification is now mainly based on phylogenetic relationships, considering above all the highly conserved *pol* gene (Grandi and Tramontano,

2018b). Additionally, a comprehensive analysis using the RetroTector software developed by Sperber *et al.* in 2007 (Sperber *et al.*, 2007) allowed for further classification of HERVs in 39 “canonical” groups and 31 “non-canonical” clades, displaying mosaic structures that emerged from recombination events and secondary integrations (Vargiu *et al.*, 2016).

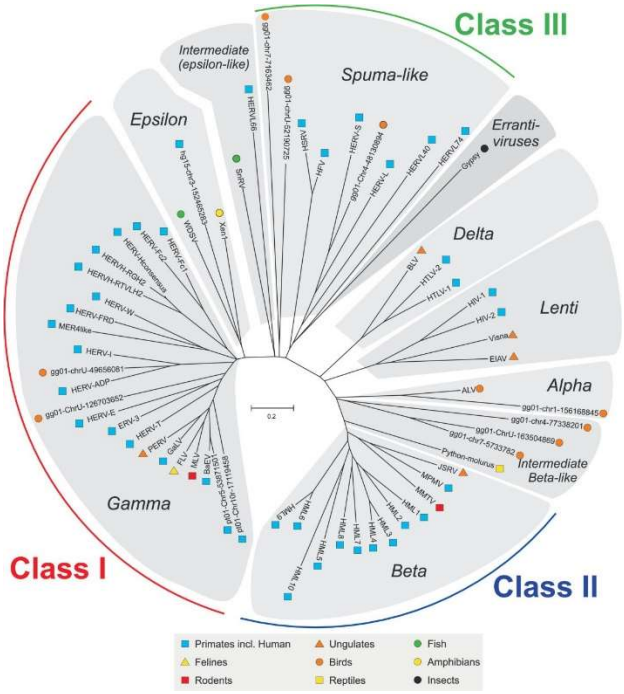


Figure 2: Classification of HERVs based on similarity to exogenous retroviruses.

Unrooted Pol neighbor joining dendogram of the following seven retroviral genera: *alpha-, beta-, gamma-, delta-, epsilon-, lenti- and spuma*-like retroviruses. The defined retroviral classes are indicated in the periphery with class I shown in red, class II in blue and class III in green. Adapted from Jern, Sperber and Blomberg, 2005.

3.3 The HERV Envelope Protein

The envelope protein (env) is the sole surface glycoprotein of retroviruses and consists of two subunits forming a heterodimer. The two subunits, namely the extracellular surface subunit (SU) and the membrane-anchored transmembrane subunit (TM) are either covalently linked through a disulfide bond in the case of *Gammaretrovirus*-like HERVs, or non-covalently associated as seen in *Beta*- and *Spumaretrovirus*-like HERVs (Henzy and Coffin, 2013; Hogan and Johnson, 2023) (Figure 3). These heterodimers are then further assembled to form a trimeric env protein (Grandi and Tramontano, 2018a). In the context of exogenous, infectious retroviruses the SU serves as ligand binding to the respective cellular receptor on the host cell membrane. This mediates the insertion of a fusion peptide (FP), which is located within the TM, into the cellular membrane, and triggers the fusion of the viral membrane with the host

cell membrane, thereby releasing the nucleocapsid into the host cell cytoplasm (Grandi and Tramontano, 2018a).

A further distinct feature of the env protein is the putative immunosuppressive domain (ISD) located within the TM. This domain was first described by Cianciolo *et al.* in 1985 who designed a 17 amino acid long synthetic peptide termed CKS-17 based on a highly homologous region found in the TM of several murine, feline and human retroviruses. This isolated peptide not only inhibited the proliferation of an interleukin-2-dependent murine cytotoxic T-cell line but was also able to inhibit the alloantigen-stimulated proliferation of murine and human lymphocytes (Cianciolo *et al.*, 1985). In 2007, Mangeney *et al.* identified a 20 amino acid motif comprising the CKS-17 peptide as the minimal ISD through truncation and mutation experiments in the murine retrovirus Moloney murine leukemia virus (MoMLV) (Mangeney *et al.*, 2007). For exogenous retroviruses, this highly conserved 20 amino acid long peptide is thought to counteract the host's immune responses upon infection (Mangeney and Heidmann, 1998; Blaise, Mangeney and Heidmann, 2001). In the context of HERVs, it has been hypothesized that the immunosuppressive properties of the env protein of HERV-W and HERV-FRD are involved in maintaining the maternal tolerance towards the semi-allogeneic fetoplacental unit during pregnancy (Mangeney *et al.*, 2007; Lavialle *et al.*, 2013).

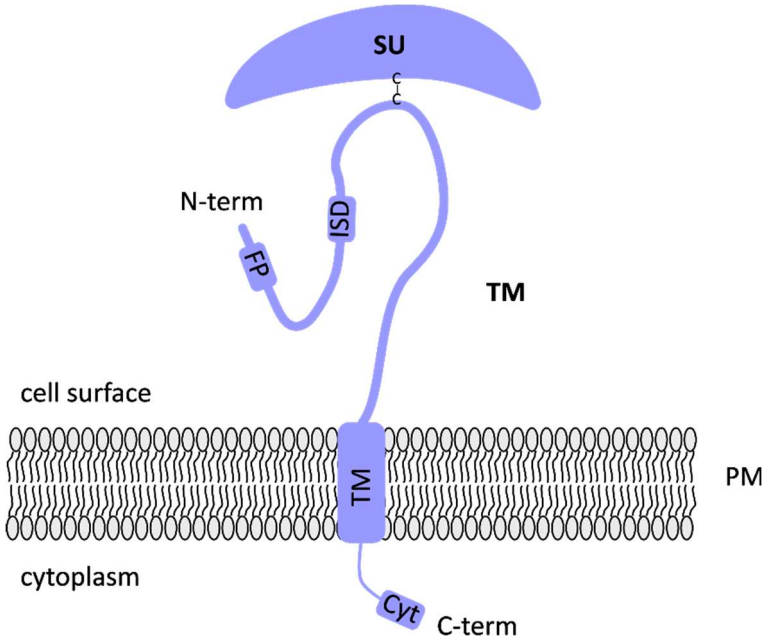


Figure 3: Schematic depiction of a monomeric env protein.

Depicted is a simplified schematic view of an env monomer anchored in a cell membrane. SU: surface subunit; TM: transmembrane domain; FP: fusion peptide; ISD: immunosuppressive domain; PM: plasma membrane; Cyt: cytoplasmic tail. The illustration was created using Affinity Designer 2.5.2 and is adapted from Grandi and Tramontano, 2018a.

3.4 The Role of HERVs in Human Physiology and Pathology

Although the vast majority of HERVs are thought to be inactive, some still contain coding capacities for different HERV elements, and their expression products are linked to a variety of physiological but also pathological functions. An example for HERVs playing a major role in physiological processes are the env proteins of HERV-W and HERV-FRD, also called syncytin-1 and syncytin-2 respectively. Both proteins possess fusogenic properties and play a crucial part in pregnancy by contributing to the formation of the placental syncytiotrophoblast (Blaise *et al.*, 2003; Blond *et al.*, 2000; Mi *et al.*, 2000). Additionally, both proteins are proposed to contribute to the maternal immune tolerance towards the fetal allograft through their ISDs, although studies from Mangeney and colleagues showed a lack of immunosuppressive activity by syncytin-1 caused by two amino acid substitutions within the ISD motif (Mangeney *et al.*, 2007).

Another example of HERVs associated with human physiological processes is HERV-H, where upregulation of transcripts has been shown to play an important role in maintaining human embryonic stem cell identity (Lu *et al.*, 2014; Wang *et al.*, 2014).

On the other hand, HERVs have been shown to be associated with auto-immune diseases and different types of cancer. One of the most prominent examples is the association of a HERV-W env protein with multiple sclerosis (MS), which is also referred to as multiple sclerosis-associated retrovirus (MSRV). Several studies could show increased expression levels of MSRV env in peripheral blood mononuclear cells (PBMCs) and brain samples of MS patients (Antony *et al.*, 2006; Mameli *et al.*, 2007; Nevalainen, Autio-Kimura and Hurme, 2024). These findings inspired the development of a therapeutic monoclonal antibody targeting the MSRV env protein called temelimab, which recently completed a clinical phase IIb trial. Even though no significant effects on acute inflammation were observed, the data demonstrated a potential positive effect on neurodegeneration and repeated administration of temelimab was well tolerated. (Curtin *et al.*, 2015; Hartung *et al.*, 2022).

Other studies show an association of HERV transcription or expression with the development and progression of cancer. For example, full-length and spliced HERV-K env transcripts of the subfamily HML-2 were not only found in breast cancer tissues, but the protein expression levels were also negatively correlated with survival (Zhou *et al.*, 2016; Johanning *et al.*, 2017). Furthermore, HERV-K *env* transcripts could be observed in pancreatic cancer tissues and the

knock-down of HERV-K *env* via shRNA led to a reduced tumor growth in a xenograft mouse model (Li *et al.*, 2017). Additionally, HERV-K transcripts was also detected in tumor samples from prostate, ovarian, colon or lung cancer, among others as reviewed by Bermejo *et al.* (Bermejo *et al.*, 2020).

Aside from the before-mentioned role of the HERV-W-derived syncytin-1 protein, it was found to be expressed in breast and endometrial cancers, where it is involved in cell-cell fusion with other tumor or endothelial cells. Similarly, syncytin-2 overexpression was detected in endometrial tumors (Bjerregaard *et al.*, 2006; Strick *et al.*, 2007; Strissel *et al.*, 2012).

There is also evidence for HERV-H transcripts and expression in several types of tumors (detailed description in 3.6). Non-coding spliced transcripts were found to be present in colon cancer (Strissel *et al.*, 2012), HERV-H env binding antibodies could be detected in the plasma of prostate cancer patients (Strissel *et al.*, 2012), and HERV-H RNA was significantly enriched in tumor tissues of patients with head and neck squamous cell carcinoma (HNSCC) (Kolbe *et al.*, 2020).

Even though a lot of research has been dedicated to the subject of HERVs and their association with cancer in recent years, it still remains unclear whether HERVs are a causative agent or just a byproduct of cancer. As reviewed by Kassiotis and Stoye, several hypotheses exist on how HERVs could contribute to the development and progression of cancer. For example, identical proviruses or solitary LTRs could trigger non-allelic homologous recombination events leading to chromosomal rearrangement, a common feature in the genomic alterations of cancer cells (Kassiotis and Stoye, 2017). It has been further hypothesized, that HERVs can influence gene expression through the promoter activity of their LTRs since many HERV LTRs have shown to be targets for DNA binding proteins like the tumor suppressor protein p53 (Wang *et al.*, 2007). Additionally, HERV-encoded proteins like env could play a role in tumorigenesis. As discussed before, HERV env proteins like syncytin-1 and syncytin-2 possess cell-fusion properties which are also suspected to contribute to neoplastic transformation or metastasis (Bjerregaard *et al.*, 2006; Strick *et al.*, 2007). Furthermore, there is evidence that the immunosuppressive activity mediated by the ISD is able to suppress the immune response against tumors in an *in vivo* mouse model (Mangeney and Heidmann, 1998; Mangeney *et al.*, 2001).

3.5 The HERV-H Family

The HERV-H family belongs to the class I HERVs and was first described by Mager and Henthorn in 1984 under the name RTVL-H for retrovirus-like element with a PBS homologous to the histidine tRNA (Mager and Henthorn, 1984). This family is of particular interest in the field of HERV research as it is one of the most abundant HERV families with about 1000 copies per haploid genome (Mager and Henthorn, 1984; Wilkinson, Mager and Leong, 1994; Jern, Sperber and Blomberg, 2004). More importantly, even though most of the HERV-H proviruses are partially deleted, around 100 copies per haploid genome still contain a putative *env* gene (Hirose, Takamatsu and Harada, 1993). Eventually, a study by De Parseval and colleagues from 2001 narrowed the HERV-H sequences down to three integration sites containing a complete, full-length *env* sequence using DNA isolated from human peripheral blood leukocytes from healthy donors. Corresponding to the transcripts' theoretical molecular weights, the three associated proviruses were termed HERV-H/env62, HERV-H/env60 and HERV-H/env59 (De Parseval *et al.*, 2001). In addition to that, Mangeney *et al.* could prove in an *in vivo* mouse model that the HERV-H env62 protein not only possesses an ISD, but that it was also active and enabled transduced tumor cells to escape immune rejection in immunocompetent mice (Mangeney *et al.*, 2001).

3.6 HERV-H and Cancer

As briefly mentioned earlier, HERV-H transcription can be detected in a variety of different cancers. Kong *et al.* investigated the transcription of TEs using a computational method called *REdiscoverTE,* screening RNA sequencing data from The Cancer Genome Atlas database. They could show that the proviral part of the HERV-H sequence was transcribed in different cancer types including stomach adenocarcinomas, bladder carcinoma, liver hepatocellular carcinoma, HNSCC, colon adenocarcinoma and lung squamous cell carcinoma. HERV-H was not only one of the most frequently transcribed TEs within their study, but their data also suggest a possible expression of full-length HERV-H sequences. Furthermore, Kong and colleagues could also show that the loss of DNA methylation, especially on CpG sites proximal to the respective TE, was associated with the observed overexpression of TEs in cancer (Kong *et al.*, 2019).

Two other studies performed by one research group looked at HERV-H *env* RNA transcripts in different human tumor cell lines as well as healthy and tumor tissues via RT-PCR. Within the first study, HERV-H *env* RNA was detected in a variety of human tumor cell lines, including cell lines from bladder carcinoma, breast and mammary gland carcinomas, pancreas, cervix, lung and liver carcinoma and several more. (Yi, Kim and Kim, 2005). Notably, Yi *et al.* also sequenced the PCR fragments and found clones from three cell lines encoding for the complete 3' part of the env peptide sequence without any frameshifts or premature stop codons. The cell lines containing the intact sequence were the mammary gland adenocarcinoma cell line MCF-7, the pancreas carcinoma cell line MIA-PaCa-2 and the cervix carcinoma cell line C-33A (Yi, Kim and Kim, 2005). In the second study, partial HERV-H *env* RNA transcription could be detected in liver, lung and testis tumor tissues via real-time RT-PCR amplification (Ahn and Kim, 2009).

While the above-mentioned studies looked at HERV-H transcription in a broad range of different cancer types, other studies focused on the association of HERV-H with individual types of cancer. This is most prominently exemplified by studies regarding colorectal cancer (CRC), where findings of Jern *et al.* show that especially HERV-H *env* transcripts were most prevalent in colon tumors. (Jern *et al.*, 2005). One of the earliest evidences of HERV-H association with colorectal adenomas and cancers was found by Wentzensen *et al.* in 2004 when they discovered a yet uncharacterized HERV-H *env* sequence transcription upregulated in tumor tissue (Wentzensen *et al.*, 2004). In 2007, Wentzensen and his colleagues identified and analyzed another HERV-H sequence that was strongly transcribed in a subset of gastrointestinal cancers, including colorectal cancers and adenomas. Furthermore, they could link the increased transcription to chromosomal demethylation in a region surrounding the 5'LTR (Wentzensen *et al.*, 2007). These findings could be validated by Alves *et al.* in 2008, who also detected the same HERV-H sequence in CRC tissue (Alves *et al.*, 2008). Additionally, Pérot and colleagues showed increased HERV-H transcription of *gag*, *env* and *pol* elements in CRC tissues via RT-qPCR analysis. They could also show a correlation of HERV-H transcription with lymph node invasion of tumor cells, thereby suggesting a functional role of HERV-H in tumorigenesis (Pérot *et al.*, 2015). A study conducted by Kudo-Saito *et al.* further underlined a functional role of HERV-H in tumor progression. Here, elevated HERV-H *env* transcription could be detected in metastatic tumor cells undergoing epithelia-to-mesenchymal transition (EMT). Kudo-Saito *et al.* linked this effect to the ISD of the HERV-H env protein since they could

induce EMT in tumor cells transcribing only low levels of HERV-H by treating the cells with the isolated ISD peptide. Furthermore, they were able to show, that treatment with the HERV-H ISD peptide stimulated the expression of CCL19, a chemokine that is involved in the recruitment and expansion of immunoregulatory $CD271^+$ cells. CD271 has been described as marker for cancer stem cells and also mesenchymal stem cells, which are suspected to contribute to tumor growth and progression by inducing immunosuppression, EMT and survival of cancer stem cells among other (Bühring *et al.*, 2007; Boiko *et al.*, 2010; Imai *et al.*, 2013; Yang *et al.*, 2013). Notably, an increased level of $CD271^+$ cells could be detected in HERV-H and CCL19-positive tumor tissues and siRNA-mediated knock-down of HERV-H *env* in *in vitro* and *in vivo* experiments improved the immune response and led to tumor regression (Kudo-Saito *et al.*, 2014).

Another cancer type showing an association with HERV-H transcription is the afore-mentioned HNSCC. Concerning this, a study using the Telescope software to screen tumor and normal tissue samples from The Cancer Genome Atlas program showed enrichment of HERV-H sequences in HNSCC tumor tissues (Kolbe *et al.*, 2020).

There is also evidence that HERV-H, and again especially the env protein are associated with prostate and breast cancer. Manca *et al.* detected antibodies targeting the HERV-H SU in plasma samples of patients with prostate cancer and Rhyu *et al.* showed increased HERV-H *env* mRNA transcription in blood samples of breast cancer patients, which were decreased again in patients undergoing chemotherapy (Rhyu *et al.*, 2014; Manca *et al.*, 2022).

3.7 Using HERVs as a Target for Cancer Stratification and Immunotherapy

Due to the frequent association of HERV transcription with cancer and the implication of their role in metastasis and immune evasion, HERVs have become of interest as a potential target for tumor immunotherapy to the extent that expression of HERV proteins has been proven. The use of HERVs as novel tumor targets, not only in a therapeutic but also in a diagnostic context was already proposed in 2006 by Wang-Johanning *et al.* (Wang-Johanning *et al.*, 2006). In 2011, the group developed monoclonal antibodies (mAbs) targeting the HERV-K env protein. They could show that these mAbs were able to inhibit growth of breast tumor cell

lines while simultaneously inducing apoptosis within these cells. Furthermore, treatment with one mAb led to tumor reduction in a xenograft mouse model (Wang-Johanning *et al.*, 2011).

Kraus *et al.* used a different approach to target HERV-K env proteins by designing a recombinant modified vaccinia Ankara (MVA)-based vaccine expressing the env protein. Their vaccine not only showed a therapeutic effect by reducing pulmonary tumor nodules in a mouse model, but also proved to be prophylactic when challenging immunized mice with HERV-K env expressing tumor cells (Kraus *et al.*, 2013).

The development of chimeric antigen receptor (CAR) T-cell therapy led to new promising treatment options in the field of cancer immunotherapy. Zhou *et al.* used this technology by fusing the single-chain variable fragment of an anti-HERV-K mAb to the T-cell receptor signaling domain. Their HERV-K CAR T-cells not only showed growth inhibition of breast cancer cells, but also a cytotoxic effect. Additionally, treatment with the HERV-K CAR T-cells reduced breast tumor growth and weight in a xenograft mouse model. Notably, the treatment also prevented tumor metastasis to other organs (Zhou *et al.*, 2015). The same group could also show similar effects on a metastatic melanoma mouse model (Krishnamurthy *et al.*, 2015).

So far, the research on targeting HERVs for cancer immunotherapy has been primarily focused on the HERV-K env protein. Nevertheless, the association of HERV-H with several types of cancer as described in section 3.6 marks another potential target candidate. While some research has been conducted regarding the HERV-H LTR-associating proteins 1 and 2, genes that obtain their polyadenylation signal through HERV-H LTR regions (Mager *et al.*, 1999), it seems reasonable to also focus on the HERV-H env protein for both, diagnostic and therapeutic options.

4 Objective

As outlined above, HERVs and especially their env proteins have the potential to serve as a new form of tumor-associated antigen. Therefore, HERV env proteins should be considered for new therapeutic and diagnostic approaches in tumor immunotherapy, including for example therapeutic vaccinations or monoclonal antibody treatment targeting these proteins.

Within this work, four main work packages should be elaborated to address this form of env-targeted tumor immunotherapy approach, focusing especially on the HERV-H family:

I. Antigen-design and characterization of five potential HERV env candidates
II. Assessing the immunogenic properties of different HERV-H env variants in mice
III. Generation of monoclonal and polyclonal antibodies targeting the HERV-H env protein
IV. Establishing diagnostic platforms to screen tumor tissues for HERV-H env transcription and protein expression.

Within the first work package, env sequences of the five HERV families HERV-W, -FRD, -H, -K and ERV3.1 should be recombinantly produced and assessed for their protein expression capacities and cell-cell fusion activity.

In the second part, BALB/c mice should be immunized with different variants of the membrane-anchored and soluble HERV-H env protein. The humoral immune reaction to different antigen combinations as DNA or recombinant protein vaccines should be assessed in mice.

The third work package will build on the previous immunization experiment. Monoclonal antibodies targeting HERV-H env should be generated from immunized mice and resulting hybridoma clones should be characterized. Additionally, polyclonal antibodies purified from rabbits immunized with recombinant HERV-H env protein should be characterized.

Lastly, diagnostic platforms based on RT-(q)PCR, sequencing and immunocyto-/-histochemistry (ICC/IHC) staining should be implemented to screen tumor tissues for HERV-H env transcription and protein expression. These methods will be first established using HERV-H env positive human tumor cell lines and then applied to tumor tissue of colorectal cancer patients.

5 Material and Methods

5.1 Antibodies

Table 1: Antibodies used in this work.

Antibody	Conjugate	Supplier	Specification	Application/Dilution)
Alpha Tubulin antibody (DM1A)	-	Santa Cruz Biotechnology (Cat. #: sc-32293)	Mouse monoclonal IgG1 κ	WB/1:1000
Alpha Tubulin antibody (DM1A)	Alexa Fluor™ 488	Santa Cruz Biotechnology (Cat. #: sc-32293 AF488)	Mouse monoclonal IgG1 κ	ICC/1:100
Anti-HA tag antibody	-	SinoBiological (Cat. #: 100028-MM10)	Mouse monoclonal IgG1	FC/1:150; IS/1:500; WB/1:5000
Anti-HERV-H env Affinity-purified	-	Davids Biotechnologie	Rabbit polyclonal HERV-H affinity-purified fraction	WB/(5 µg/ml)
Anti-HERV-H env Protein-A-purified	-	Davids Biotechnologie	Rabbit polyclonal Protein-A-purified fraction	ICC/(10 µg/ml); WB/(1 µg/ml)
Goat anti-mouse IgG	PE	BioLegend (Cat. #: 405307)	Polyclonal goat anti-mouse IgG	FC/1:100
Goat anti-mouse IgG (H+L)	HRP	Jackson Immuno Research (Cat. #: 115-036-146)	Goat anti-mouse F(ab')2 fragment	WB/1:10000
Goat anti-mouse IgG1	HRP	Southern Biotech (Cat. #: 1071-05)	Polyclonal goat anti-mouse IgG1	ELISA/1:8000
Goat anti-mouse IgG2a	HRP	Southern Biotech (Cat. #: 1081-05)	Polyclonal goat anti-mouse IgG2a	ELISA/1:8000
Goat anti-mouse IgG2b	HRP	Southern Biotech (Cat. #: 1091-05)	Polyclonal goat anti-mouse IgG2b	ELISA/1:8000
Goat anti-mouse IgG3	HRP	Southern Biotech (Cat. #: 1101-05)	Polyclonal goat anti-mouse IgG3	ELISA/1:8000
Goat anti-mouse IgM	HRP	Southern Biotech (Cat. #: 1021-05)	Polyclonal goat anti-mouse IgM	ELISA/1:8000
Goat Anti-Rabbit Ig	HRP	Dako (Cat. #: P044801-2)	Polyclonal goat anti-rabbit Ig affinity isolated	ICC/1:500; WB/1:2000; ELISA/1:2000
Goat anti-rabbit IgG	PE	Cell Signaling Technology (Cat. #: 79408)	Goat anti-rabbit F(ab')2 fragment	FC/1:500
Goat anti-rabbit IgG (H + L)	Alexa Fluor™ 647	Invitrogen (Cat. #: A21244)	Polyclonal goat anti-rabbit IgG	ICC/1:500

Mouse IgG1 isotype control	-	BioLegend (Cat. #: 401401)	Mouse monoclonal IgG1, κ isotype control antibody	ELISA/0.1 µg
Mouse IgG2a isotype control	-	BioLegend (Cat. #: 401501)	Mouse monoclonal IgG2a, κ isotype control antibody	ELISA/0.1 µg
Mouse IgG2b isotype control	-	BioLegend (Cat. #: 401201)	Mouse monoclonal IgG2b, κ isotype control antibody	ELISA/0.1 µg
Mouse IgG3 isotype control	-	BioLegend (Cat. #: 401301)	Mouse monoclonal IgG3, κ isotype control antibody	ELISA/0.1 µg
Mouse IgM isotype control	-	BioLegend (Cat. #: 401601)	Mouse monoclonal IgM, κ isotype control antibody	ELISA/0.1 µg
Rabbit anti-mouse IgG	HRP	Dako (Cat. #: P0260)	Polyclonal rabbit anti-mouse Ig purified Ig fraction	IS/1:1000; ELISA/1:2000

ELISA: Enzyme-linked immunosorbent assay; FC: flow cytometry; ICC: immunocytochemistry; IS: immune staining; WB: western blot

5.2 Molecular Biology Methods

5.2.1 General Molecular Biology

DNA constructs were cloned using common molecular biology protocols (Lessard, 2013). HERV *env* inserts were gene- and codon-optimized for human codon usage utilizing the GeneArt GeneOptimizer (GeneArt/Thermo Fisher Scientific). For the amplification of inserts, attachment of restriction sites or protein tags and the insertion of single nucleotide exchanges, polymerase chain reaction (PCR) was performed using the S7 Fusion Polymerase. For this, suitable primers were designed as listed in Table S 1. Restriction digestions were performed with restriction enzymes purchased from New England Biolabs according to the manufacturer's instructions. After restriction digest or PCR, the DNA fragments were separated by agarose gel electrophoresis and the DNA from the bands were extracted using the QIAquick Gel Extraction kit. The inserts were ligated into the expression vectors pURVac

or pLib1.1 using the Quick Ligation™ kit. Chemically competent *E. coli* DH5α were transformed with the cloned DNA according to standard protocols (Hanahan, 1983) and positive colonies were identified by colony PCR using the GoTaq Green MasterMix.

Table 2: Materials used for general cloning.

Material	Supplier
S7 Fusion Polymerase	Mobidiag (Cat. #: MD-S7-500)
QIAquick Gel Extraction Kit	Qiagen (Cat. #: 28706)
Quick Ligation™ Kit	NEB (Cat. #: M2200L)
GoTaq Green MasterMix	Promega (Cat. #: M7123)

5.2.2 Plasmid Preparation

For plasmid preparation, *E. coli* DH5α were transformed with the plasmid of interest and cultured overnight in the desired volume of LB medium, depending on the amount of plasmid DNA needed after preparation. The plasmid DNA was then extracted from the bacteria using DNA plasmid preparation kits according to the manufacturer's instructions. The DNA concentration was determined spectrophotometrically using a NanoDrop ND-1000 photometer. To check for the correct sequence, the plasmids were sent to Sanger sequencing at SeqLab Sequence Laboratories GmbH (Göttingen, Germany).

Table 3: Material used for plasmid preparation.

Material	Supplier
GeneJET Plasmid Miniprep Kit	Thermo Fisher Scientific (Cat. #: K0503)
Plasmid Plus Midi Kit	QIAGEN (Cat. #: 12945)
EndoFree Plasmid Mega Kit	QIAGEN (Cat. #: 12381)

5.2.3 HERV-H *env* Sequencing from Human Tumor Cell Lines

To sequence the HERV-H *env* reading frame from human tumor cell lines, genomic DNA (gDNA) was extracted using a gDNA isolation kit according to the manufacturer's manuals. From the gDNA, the complete *env* reading frame was amplified by PCR using the Q5 High-Fidelity DNA Polymerase. PCR products were separated by agarose gel electrophoresis and correctly sized DNA bands were extracted from gel using the Zymoclean Gel DNA Recovery Kit. The purified PCR fragments were either sent directly to Sanger sequencing, or after blunt-end cloning into a pJET sequencing vector using the CloneJET PCR cloning kit.

Table 4: Material used for the sequencing of the HERV-H env reading frame from human tumor cell lines.

Material	Supplier
Quick-DNA Miniprep	Zymo Research (Cat. #: D3025)
Q5® High-Fidelity DNA Polymerase	NEB (Cat. #: M0491L)
Zymoclean Gel DNA Recovery Kit	Zymo Research (Cat. #: D4001)
CloneJET PCR cloning kit	Thermo Fisher Scientific (Cat. #: K1231)

5.2.4 RT-(q)PCR

RT-PCR and RT-qPCR assays were conducted with RNA extracted from different cell lines or FFPE tissue. RNA extraction from fresh cell lines was performed using the RNeasy Mini Kit, while the Pinpoint Slide RNA Isolation System II was used to extract RNA from FFPE samples. The RT-PCR was performed using the LyoPrime Luna One-Step RT-qPCR mix and the PCR products were analyzed by agarose gel electrophoresis. For the RT-qPCR, the Luna Universal One-Step RT-qPCR Kit was used. The reaction and quantification were carried out in the StepOnePlus Real-Time PCR cycler. $\Delta\Delta C_T$ values were calculated by using the following formula:

1. $\Delta CT = CT(reference\ gene) - CT\ (target\ gene)$
2. $\Delta\Delta CT = \frac{1.29^{\Delta CT(target\ gene)}}{1.29^{\Delta CT(reference\ gene)}}$

Table 5: Material used for RT-(q)PCR.

Material	Supplier
RNeasy Mini Kit	QIAGEN (Cat. #: 74104)
Pinpoint Slide RNA Isolation System II	Zymo Research (Cat. #: R1007)
LyoPrime Luna® Probe One-Step RT-qPCR Mix with UDG	NEB (Cat. #: L4001S)
Luna Universal One-Step RT-qPCR Kit	NEB (Cat. #: E3005S)
StepOnePlus System	Thermo Fisher Scientific

5.3 Cell Biology Methods

5.3.1 Cultivation and Transfection of Cell Lines

5.3.1a Cultivation of soluble cells

Expi293 cells were cultivated in Expi293 expression medium at 37 °C, 125 rpm and 8 % CO_2. A cell density between 0.3 and 3x10^6 cells/ml was maintained by passaging the cells 3 times per week.

Sp2/0-Ag14 cells were cultivated in RPMI-K medium and passaged every 2-3 days in a 1:2 ratio.

Stable hybridoma cell lines were passaged in a 1:2 ratio 3 times per week. The cells were cultivated in ClonaCell™-HY AOF expansion medium.

Except for the Expi293 cells, all soluble cells were cultivated at 37 °C and 5 % CO_2.

5.3.1b Cultivation of adherent cells

Adherent cells were cultured in T-75 flasks at 37 °C and 5 % CO_2. Cells were passaged before reaching confluence by splitting them between 1:5 and 1:25, according to the respective cell line (Table 7).

Table 6: Media and material used for cell culture.

Medium	Composition
DMEM-0	Dulbecco's Modified Eagle Medium, Thermo Scientific (Cat. #: 41966-029)
DMEM-K	DMEM-0; 10 % FCS; 1 % Pen/Strep
DMEM-K for MIA-PaCa-2	DMEM-0; 10 % FCS; 2.5 % horse serum; 1 % Pen/Strep
DMEM selection + Doxy	DMEM-K; 2.5 µg/ml Puromycin; 1 µg/ml Doxycycline
DMEM selection - Doxy	DMEM-K; 2.5 µg/ml Puromycin
Expi293 expression medium	Thermo Fisher Scientific (Cat. #: A1435101)
ClonaCell™-HY AOF expansion medium	StemCell technologies (Cat. #: 03835)
Penicillin/Streptomycin (Pen/Strep)	10,000 U/ml Penicillin / 10 mg/ml Streptomycin PAN Biotech (Cat. #: P06-07100)
RPMI-K medium	Pan™ Biotech (Cat. #: P04-16500); 10 % FCS; 1 % Pen/Strep
TC Flask T75, Standard	Sarstedt (Cat. #: 83.3911)

Table 7: List of adherent cell lines.

Cell line	supplier	Culture medium	Split ratio when confluent
HEK293 T	ATTC, Cat. #: CRL-11268	DMEM-K	1:10
HEK293	Kindly provided by InProTher	DMEM-K	1:10
MCF-7	Kindly provided by InProTher	DMEM-K	1:5
C-33A	ATCC, Cat. #: HTB-31	DMEM-K	1:10
DF-1	ATCC, Cat. #: CRL-3586	DMEM-K	1:10
MIA-PaCa-2	DSMZ, Cat. #: ACC 733	DMEM-K for MIA-PaCa-2	1:5
BHK-21	ATCC, Cat. #: C-13	DMEM-K	1:25
NIH-3T3 HERV-H TetOFF	Kindly generated and provided by Sirion	DMEM selection with or without doxycycline	1:10

5.3.1c Transfection of soluble cells

Expi293 cells were transfected in a volume of 300 ml at a cell density of $2.5x10^6$ cells/ml. For this, 300 µg plasmid DNA and 800 µl ExpiFectamine were diluted in 30 ml OptiMEM and subsequently incubated for 20 min at RT. After incubation, the transfection mix was added to the cells and the cells were incubated at 37 °C, 125 rpm and 8 % CO_2 for 16-18 h. Then, 1.5 ml Enhancer I, 15 ml Enhancer II and 0.5 % Penicillin/Streptomycin were added, and the cells were further incubated for 5 days.

5.3.1d Transfection of adherent cells

Adherent cells were transfected with either PEI or Lipofectamine™ 2000 in 6-well, 24-well, or round-bottom 96-well cell culture plates. For this, cells were seeded 24 h before transfection. The transfection mix was prepared as shown in Table 8. For PEI transfection the medium was exchanged to DMEM-0 right before transfection and 5-6 h post-transfection the transfection mix was replaced with DMEM-K. The Lipofectamine transfection mix was added directly to the cells in DMEM-K. The transfected cells were incubated at 37 °C and 5 % CO_2 for 48 h.

Table 8: Transfection mixtures for the different plate formats and transfection reagents.

Material for transfection	96-well plate	24-well plate	6-well plate
Cells seeded per well	$0.35x10^5$	$1x10^5$	$5x10^5$
PEI transfection			
Plasmid DNA	-	0.75 µg	2.5 µg
PEI	-	2.25 µg	7.5 µg
Lipofectamine 2000 transfection			
Plasmid DNA	0.1 µg	-	2.5 µg
Lipofectamine 2000 reagent	1.5 µl	-	12 µl

Table 9: Material used for the transfection of cells.

Material	Supplier
ExpiFectamine 293 Transfection Kit	Thermo Fisher Scientific (Cat. #: A14524)
OptiMEM™	Thermo Fisher Scientific (Cat. #: 31985062)
Lipofectamine™ 2000 Transfection reagent	Thermo Fisher Scientific (Cat. #: 11668030)
PEI MAX 4000	Polyscience Inc. (Cat. #: 24765-1)

5.3.1e Paraffin-embedding of cell pellets

For immunocytochemistry staining, cell pellets were embedded in paraffin and subsequently, sections were cut using a microtome and placed onto microscope slides. For this, the cells were cultivated in T-75 cell culture flasks until they were confluent, then the cells were detached, washed three times with PBS and resuspended in formaldehyde. Embedding in paraffin and cutting of microscope sections was kindly performed at the pathology department of the University Hospital in Regensburg by Doris Gaag using standard protocols.

5.3.2 Flow Cytometry

Cells were stained for flow cytometry readout either intracellularly or via surface staining. For the intracellular staining, the cells were transfected in a 6-well format. After transfection, the cells were transferred to FACS tubes. Washing steps were carried out at 300 x g and 4 °C for 5 min. The cells were first washed one time with PBS, then one time with FACS buffer. Next, the cells were resuspended in Cytofix/Cytoperm and incubated on ice for 20 min. After two following wash steps with Perm/Wash, the anti-HA tag antibody (Table 1) was added to the cells diluted in Perm/Wash and incubated on ice for 25 min. The cells were washed three times with Perm/Wash and as a secondary antibody, a rat anti-mouse IgG1 PE antibody (Table 1)

was added to the cells. The samples were again incubated on ice for 25 min. After three more wash steps, the cells were resuspended in 500 µl FACS buffer and analyzed by flow cytometry using the Attune NxT flow cytometer.

The surface staining was performed in a round-bottom 96-well plate. For this, the transfected cells were centrifuged at 500 x g and 4 °C for 5 min and washed with FACS buffer two times. Afterwards, the primary antibody was added and incubated on ice for 1.5 h. The cells were washed again two times with FACS buffer, and a fluorescence-labeled secondary antibody was added and incubated for 30 min. Following two more wash steps, the cells were resuspended in 200 µl FACS buffer per well and analyzed in the Attune NxT flow cytometer.

Table 10: Material used for flow cytometry staining.

Material	**Composition/Supplier**
FACS buffer	1x PBS; 1 % heat-inactivated FCS; 0.1 % NaN_3
Cytofix/Cytoperm	1x PBS; 4 % PFA; 1 % saponin
Perm/Wash	1x PBS; 0.1 % saponin
Attune NxT flow cytometer	Thermo Fisher Scientific

5.3.3 Immunostaining of Cells

HEK293 T cells were transfected with the different HERV-env plasmids in a 24-well plate. The transfected cells were washed with PBS and then fixed with ice-cold fixing solution for 3 min followed by overnight blocking. The next day, the cells were first stained with an anti-HA tag antibody (Table 1) for 1 h at RT. Afterwards, the cells were washed three times with blocking buffer and incubated with an anti-mouse IgG/HRP conjugate (Table 1) for 45 min. The cells were washed again three times with blocking buffer and the colorimetric reaction was developed using the TrueBlue substrate until a blue color was visible. Subsequently, the substrate was exchanged with H_2O and the cells were imaged using the Keyence microscope.

Table 11: Materials for immunostaining.

Material	**Composition/Supplier**
Fixing solution	1:1 mixture of acetone and methanol
Blocking buffer/antibody dilution buffer	PBS; 3 % BSA
TrueBlue™ Peroxidase Substrate	SeraCare (Cat. #: 5510-0030)
Keyence BZ-9000 fluorescence microscope	Keyence

5.3.4 Immunocytochemistry Staining

For immunocytochemistry (ICC) and immunohistochemistry (IHC) staining, formalin-fixed paraffin-embedded (FFPE) tumor cell lines or colon tissues were either stained using a fluorescence staining protocol or by 3,3'-diaminobenzidine (DAB) staining. The slides were first deparaffinated for both ICC staining methods by treating them with Xylene followed by wash steps in decreasing Ethanol concentrations. The slides were shortly rinsed in tap water and then boiled in heat-induced epitope retrieval (HIER) buffer at 95 °C for 10 min. Afterwards, the slides were washed two times with wash buffer and were then incubated in blocking buffer for 1-2 h at RT. The primary Protein-A-purified anti-HERV-H env antibody (Table 1) was diluted in antibody dilution buffer and incubated on the slides overnight at 4 °C. The next day, the slides were washed again two times. For the fluorescence staining the cells were then co-stained with a secondary anti-rabbit IgG/Alexa647 antibody, an anti-tubulin/Alexa488 antibody and a Hoechst 33342 stain (Table 1) for 2 h at RT. Afterwards, the slides were washed three times with PBS and subsequently incubated in a $CuSO_4$ + NH_4Cl solution for 10 min to prevent autofluorescence. After rinsing the slides with tap water and letting them dry completely the slides were mounted using the FluorSave Reagent mounting medium.

For the DAB staining, the slides were treated with 0.3 % H_2O_2 for 15 min after the primary antibody incubation. After washing the cells once with PBS, an anti-rabbit IgG/HRP conjugate (Table 1) diluted in dilution buffer was added and the samples were incubated for 2 h at RT. The slides were washed with PBS again three times and the colorimetric reaction was developed using a DAB detection solution. Afterwards, the slides were counterstained with hematoxylin and eosin (H&E staining) using the H&E staining kit according to the manufacturer's instructions. The slides were dehydrated by sending them through the deparaffinization steps in the opposite order and mounted using the Limonene Mounting Medium.

After both staining methods, the slides were kept in the dark for at least 24 h, letting the mounting medium completely solidify before imaging. All incubation steps were done in a humidified chamber and the slides were imaged using the Keyence microscope.

Table 12: Materials used for fluorescence and DAB ICC/IHC staining.

Material	Composition/Supplier
Formaldehyde	
Superfrost plus slides	R.Langenbrinck GmbH Labor- und Medizintechnik (Cat. #: 03-0060)
Xylene	Roth (Cat. #: 9713.5)
Ethanol	Sigma Aldrich/Merck (Cat. #: 1.00983.2500)
Wash buffer	PBS; 0.1 % Triton X-100
Blocking buffer	PBS; 5 % heat-inactivated FCS
Dilution buffer	PBS; 1 % BSA
HIER buffer	10 mM Sodium citrate; 0.05 % Tween-20, pH 6.0
Hoechst33342	Cell Signaling Technologies (Cat. #: 4082)
$CuSO_4$ + NH_4Cl solution	10 mM $CuSO_4$; 50 mM NH_4Cl
FluorSave-Reagent	Merck (Cat. #: 345789-20ML)
PBS	Thermo Scientific (Cat. #: 12559069)
3,3'-Diaminobenzidin Tetrahydrochlorid (DAB reagent)	Roth (Cat. #: CN75.1)
DAB detection solution	92 % PBS; 4 % 25x DAB; 4 % 25x H_2O_2
H&E staining kit	Roth (Cat. #: 9194)
Limonene Mounting Medium	Abcam (Cat. #: ab104141)
Keyence BZ-9000 fluorescence microscope	Keyence

5.3.5 Bortezomib treatment

To treat cells with the proteasome-inhibitor Bortezomib, $0.5x10^6$ cells were seeded 24 h before treatment either in DMEM-K if cellular proteins were to be obtained via cell lysis, or in DMEM-0 if proteins were to be precipitated from the cell supernatant. The cells were treated with 1000, 500, 100, 50, 10 and 1 nM Bortezomib for 24 h.

Table 13: Material used for Bortezomib treatment of cells.

Material	Supplier
Bortezomib	Cell Signaling Technology (Cat. #: 2204)

5.3.6 Generation of Mouse Hybridoma Cell Lines

Mouse hybridoma cell lines were generated using the hybridoma technology which was first described by Köhler and Milstein (Köhler and Milstein, 1975) by fusing B-cells isolated from the spleen of immunized mice with the myeloma cell line Sp2/0-Ag14. The preparation of

splenocytes and myeloma cells as well as the hybridoma fusion were performed using the ClonaCell-HY hybridoma kit according to the manufacturer's instructions.

Table 14: Material for the generation of hybridoma cells.

Components	Supplier
ClonaCell-HY Hybridoma kit	StemCell (Cat. #: 03800)
Sp2/0-Ag14 myeloma cell line	Kindly provided by InProTher

5.4 Protein Biochemistry Methods

5.4.1 Protein Purification Methods

5.4.1a Immobilized metal ion affinity chromatography

For protein purification, the supernatant of transfected Expi293™ cells was cleared by two centrifugation steps first at 1000 x g and 4 °C for 10 min and then at 4000 x g and 4 °C for 15 min. Subsequently, the supernatant was filtered through a 0.2 µm filter unit and the His_6-tagged soluble proteins were purified using immobilized metal ion affinity chromatography (IMAC). After loading the protein onto a Ni^{2+} ion matrix of a HisTrap excel column using a peristaltic pump (flow rate of 2 ml/min), the column was first washed with 5 CV washing buffer and subsequently, the protein was eluted with elution buffer following a linear gradient from 30 mM to 500 mM imidazole using the ÄKTA pure™ purification system and a flow rate of 1 ml/min. After elution, all fractions containing the protein of interest were pooled and buffer-exchanged to PBS using PD-10 columns following the manufacturer's protocol. Lastly, the protein was concentrated with an Amicon® Ultra-15 centrifugal filter unit to a concentration of approximately 0.8-1 mg/ml.

Table 15: Materials used for IMAC protein purification.

Material	Supplier
HisTrap™ Excel	Merck (Cat. #: GE17-3712-06)
PBS	Thermo Fisher Scientific (Cat. #: 14190-094)
Washing buffer	PBS, 30 mM Imidazole
Elution buffer	PBS, 500 mM Imidazole
PD-10 Desalting column	Cytiva (Cat. #: 17085101)
Amicon™ Ultra-15 centrifugal filter unit	Merck Millipore (Cat. #: UFC903024)

5.4.1b Size exclusion chromatography

The soluble recombinant protein was further purified and analyzed by size exclusion chromatography (SEC). For this, 500 µl purified protein was injected to a Superdex 200 Increase 10/300 GL column which had been equilibrated before with 1.5 CV PBS. Fractions of 500 µl were taken and fractions within the protein peak were further analyzed on a Blue Native PAGE (BN-PAGE).

Table 16: Material used for SEC.

Material	Supplier
Superdex® 200 Increase 10/300 GL	Sigma-Aldrich (Cat. #: GE28-9909-44)
PBS	Thermo Fisher Scientific (Cat. #: 14190-094)

5.4.2 Blue Native PAGE

Blue Native PAGE (BN-PAGE) was performed using 4-16 % gradient native gels (Vertikal-Nativgel 4-16 %). From each fraction of the SEC run 20 µl were mixed with a 2X native loading dye and loaded onto the gel. As a protein standard, 6 µl of a marker mix for blue/clear native PAGE was used. The electrophoresis was done in two steps; first at 50 V for 20 min followed by 200 V for 2 h. The gels were stained with Coomassie Blue staining solution for 15-20 min and subsequently destained with 7 % acetic acid until the background appeared clear.

Table 17: Material for BN PAGE.

Material	Composition/Supplier
SERVAGel™ N 4-16, Vertikal-Nativgel 4-16 %	Serva Electrophoresis GmbH (Cat. #: 43252.01)
10X BN/CN-cathode buffer	Serva Electrophoresis GmbH (Cat. #: 42536) 500 mM tricine, 150 mM BisTRIS, 1 x solution was supplemented with 0.002 % Serva Blue G solution (1 x) (Cat. #: 42538)
10X BN/CN-anode buffer	Serva Electrophoresis GmbH (Cat. #: 42535) 500 mM BisTRIS HCl
2X BN loading dye	Serva Electrophoresis GmbH (Cat. #: 42533)
Coomassie Blue staining solution	0.125 % (w/v) Coomassie brilliant blue R-250, 50 % (v/v) ethanol, 7 % (v/v) acetic acid
Destaining solution	7 % (v/v) acetic acid

5.4.3 SDS-PAGE

SDS-PAGE was performed under reducing and non-reducing conditions. For this, 3 µg protein or 20-50 µg cell lysate was diluted in PBS with either a 6X reducing or a 5X non-reducing buffer in a total volume of 20 µl. After incubation at 95 °C for 5 min, the protein was loaded onto a 12.5 % SDS gel. As a protein standard, 6 µl of a PageRuler Prestained Protein Ladder were loaded. The electrophoresis was carried out first for 20 min at 90 V followed by 140 V for 55 min. The SDS gels were stained as described in 5.4.2 Blue Native PAGE.

Table 18: Material used for SDS PAGE.

Material	Composition/Supplier
SDS gel	5 % stacking gel and 12.5 % running gel
SDS running buffer	25 mM TRIS, 192 mM glycine, 0.1 % (w/v) SDS
5X non-reducing loading buffer	312,5 mM TRIS, 50 % (v/v) glycerin, 10 % (w/v) SDS, 0.005 % bromophenol blue, pH 6.8
6X Reducing loading buffer	375 mM TRIS, 60 % (v/v) glycerol, 12 % (w/v) SDS, 30 % (v/v) β-mercaptoethanol, 0.006 % bromophenol blue, pH 6.8
PageRuler™ Prestained Protein ladder	Thermo Fisher Scientific (Cat. #: 26616)

5.4.4 Preparation of Cell Lysates

Cell lysates were prepared from transfected and untransfected cells. For this, $4x10^6$ cells were harvested from untransfected cells, whereas two 6-wells with transfected cells were harvested and pooled. The cell pellets were washed twice in PBS and were then resuspended in 200 µl TDLB+PI buffer. After 15 min incubation on ice, the cells were sonicated in 30 sec intervals for 7 min in the Bioruptor and finally centrifuged for 5 min at max. rpm at 4 °C. The supernatants were stored at -20 °C.

Table 19: Material used for the preparation of cell lysates.

Material	Composition/Supplier
PBS	Thermo Fisher Scientific (Cat. #: 14190-094)
Triple-detergent lysis buffer (TDLB)	50 mM Tris-HCl (pH 8.0); 150 mM NaCl; 0.1% (w/v) SDS; 1% (v/v) Nonidet P-40; 0.5% (w/v) sodium deoxycholate; 1 mM EDTA
complete™ Protease Inhibitor Cocktail tablets	Roche (Cat. #: 04693124001)

TDLB-PI	5 ml TDLB buffer + ½ complete™ Protease Inhibitor Cocktail tablet

5.4.5 TCA Precipitation from Cell Supernatants

Secreted protein was precipitated from the supernatant of cell cultures using 100 % trichloroacetic acid (TCA). For this, one volume of TCA was incubated with 4 volumes from the supernatant for 10 min on ice. After centrifugation at 14000 x g for 5 min, the pellet was washed twice with ice-cold acetone, then the pellet was resuspended in 2x reducing SDS buffer and incubated at 95 °C for 10 min.

Table 20: Material used for TCA precipitation.

Material	Composition/Supplier
Trichloroacetic acid solution	100 % (w/v) TCA dissolved in H_2O
2X reducing SDS buffer	6X Reducing loading buffer diluted 1:3 in H_2O

5.4.6 Bradford Assay

Protein concentrations within the cell lysates were determined with a Bradford assay using the "Bio-Rad Protein Assay" following the Microassay procedure for microtiter plates. BSA in different concentrations was used as a protein standard.

Table 21: Materials for the Bradford assay.

Material	Composition/Supplier
Bovine serum albumin (BSA)	Biomol (Cat. #: 01400.100)
Protein Assay Dye Reagent Concentrate	Bio-Rad (Cat. #: 5000006)
Micorplate reader 680	Bio-Rad

5.4.7 Western Blot

For western blot analysis, 20-50 µg cell lysates were loaded under reducing conditions onto a 12.5 % SDS gel. After electrophoresis, the proteins were blotted onto a 0.2 µm nitrocellulose membrane using the BlueFlash-L semi-dry blotting unit with 1.5 mA/cm^2 for 1.5 h. Gel and membranes were shortly incubated in transfer buffer before blotting. Successful blotting was checked by reversible Ponceau S staining. Afterwards, the membranes were blocked in blocking buffer at 4 °C overnight. The next day, the membranes were washed 4 times for 5 min with TBS-T. The primary antibodies were diluted at the appropriate concentration (Table 1) in dilution buffer and added to the membrane for an incubation at RT for 2 h. Then the

membranes were washed again 4 times with TBS-T and the secondary antibodies (Table 1) were applied. The membrane was incubated with the secondary antibody for 1 h at RT. Then the membranes were washed again with TBS-T for 4 times and 1 ml SuperSignal West or ECL substrate was applied and incubated on the membrane for 5 min. Subsequently, the chemiluminescent signal of the secondary antibody was detected using a ChemoStar Imager.

Table 22: Material used for western blotting.

Material	Composition/Supplier
BlueFlash-L semi-dry blotting unit	Serva Electrophoresis
Nitrocellulose membrane, Amersham™ Protran 0.2 µm	GE Healthcare (Cat. #: 10600001)
WhatmanTM 3MM Chr Blotting Paper	GE Healthcare (Cat. #: 3030-917)
Transfer buffer	25 mM Tris, 150 mM glycine, 10 % (v/v) methanol
Ponceau S staining solution	0.2% (w/v) Ponceau S; 1% (v/v) acetic acid
Skim milk powder	Heirler (Cat. #: 4010318030305)
TBS	150 mM NaCl, 50 mM Tris/HCl, pH 7.4
TBS-T	TBS, 0.1 % Tween-20 (v/v)
Blocking buffer	TBS, 0.1 % Tween-20 (v/v), 5 % skim milk powder (w/v)
Dilution buffer	TBS, 0.1 % Tween-20 (v/v), 1 % skim milk powder (w/v)
ECL substrate solution	100 mM Tris/HCl, pH 8.5; 12.5 mM luminol, 1.98 mM coumaric acid, 305 µl/l 30 % H_2O_2
SuperSignal™ West Femto Chemiluminescent substrate	Thermo Scientific™ (Cat. #: 3094)

5.4.8 ELISA

5.4.8a Serum ELISA

Sera from immunized mice were tested for anti-HERV-H env antibodies by either single-point or titration ELISA. For both assays, Nunc MaxiSorp™ flat-bottom plates were coated with 0.25 µg soluble recombinant protein diluted in 100 µl PBS per well at 4 °C overnight. Subsequently, the plates were washed with PBS-T and then blocked with 200 µl serum ELISA blocking buffer for 2 h at RT. Afterwards, the plates were washed again and 50 µl of the serum dilution was added to each well. For the single point ELISA, the serum was diluted 1:50 in dilution buffer, whereas for the serum titration, the serum was titrated in a 4-fold serial dilution starting at a 1:50 dilution. The serum was added to the plates that were then

incubated for 1 h at RT. After incubation and a subsequent washing step, an anti-mouse IgG/HRP conjugated secondary antibody (Table 1) was diluted in dilution buffer and 50 µl were added per well. The plate with the secondary antibody was incubated for 30 min, washed again, and the colorimetric reaction of the HRP-conjugated bound secondary antibody was developed by applying 50 µl TMB substrate solution. The reaction was stopped after 8 min with 50 µl 1 M H_2SO_4. The absorption was measured at 450 nm in a microplate reader.

5.4.8b Hybridoma and Polyclonal Anti-HERV-H Antibody ELISA

For the validation of the polyclonal anti-HERV-H antibodies (pAbs) and to screen hybridoma cells for anti-HERV-H antibody expression, Nunc MaxiSorp™ flat-bottom plates were coated with HERV-H recombinant protein or other non-HERV proteins as described above (5.4.8a Serum ELISA). In both cases, the plates were blocked with a hybridoma ELISA blocking buffer for 1-2 h at RT, washed with PBS-T and then, 100 µl of hybridoma supernatant or 50 µl pAbs diluted in hybridoma ELISA blocking buffer were applied to the wells. The plates with the supernatants and pAbs were incubated for 1.5 h at RT and washed again. As secondary antibodies for the hybridoma screening, different anti-mouse Ig class HRP-conjugated antibodies (Table 1) were used while for the pAbs validation, an anti-rabbit IgG/HRP conjugate (Table 1) was used. After incubation for 1 h at RT and a subsequent washing step, the signals of the secondary antibodies were developed and detected as previously described (5.4.8a Serum ELISA).

Table 23: Materials used for ELISA.

Material	**Components/Supplier**
Nunc MaxiSorp™ flat-bottom	Thermo Fisher Scientific (Cat. #: 44-2404-21)
PBS-T	1x PBS, 0.1 % Tween-20
Serum ELISA blocking buffer	1x PBS, 0.1 % Tween-20, 5 % milk powder
Serum ELISA dilution buffer	1x PBS, 1 % BSA
Hybridoma ELISA blocking buffer	1x PBS, 0.1 % Tween-20, 0.5 % BSA
TMB-A	30 mM potassium citrate, 10 % (w/v) citric acid, pH 4.1
TMB-B	10 mM TMB, 10 % (v/v) acetone, 90 % ethanol, 80 mM of 30 % H_2O_2
TMB substrate solution	TMB-A and TMB-B mixed at a 20:1 ratio
Tecan Hydroflex	Tecan

Microplate reader	Bio-Rad

Table 24: Recombinant proteins used in ELISA.

Protein	Characteristics	Supplier
solHERV-H_His6	soluble recombinant HERV-H env protein with C-terminal His6 tag	Produced and purified from Expi293 cells
solHERV-H_SUonly_His6	soluble recombinant HERV-H env surface subunit protein with C-terminal His6 tag	Produced and purified from Expi293 cells
HERV-H_WT-TM	soluble recombinant HERV-H env ectodomain protein	InProTher
solStabHA30_His6	Soluble, stabilized hemagglutinin protein from Influenza A/England/195/2009 with C-terminal His6 tag	Wagner lab (kindly provided by Michael Schachtner)
BG505_His6	Soluble, recombinant env protein form HIV clade BG505 with C-terminal His6 tag	Wagner lab (kindly provided by Michael Schachtner)
BG505_noTag	Soluble, recombinant env protein form HIV clade BG505 without C-terminal tag	Wagner lab (kindly provided by Michael Schachtner)

5.5 Animal Experiments

All animal experiments have been performed in collaboration with Hervolution therapeutics (former InProTher, Copenhagen/Denmark) at the University of Copenhagen according to the Danish national guidelines and experimental protocols approved by the Danish National Animal Experiments Inspectorate (Dyreforsøgstilsynet). Adult female BALB/c mice of uniform age were randomly distributed into 9 groups with 6 animals each. The mice were immunized with 50 µg HERV-H env DNA in a volume of 30 µl or 10 µg recombinant soluble HERV-H env protein in a volume of 35 µl on 4 time points via the subcutaneous route into the lower limb. The protein was mixed with MPLA as adjuvant directly before immunization in a protein-to-MPLA ratio of 1:2.5. Before the initial immunization, two weeks after the second immunization, and two weeks after the 4th immunization, blood samples were collected and the serum was extracted by centrifugation at 800 x g for 8 min at 8 °C. Serum samples were then tested for HERV-H env binding antibodies as described in 5.4.8a Serum ELISA. The mice with the highest antibody titers were immunized again with MPLA-conjugated protein and 2-

4 days after the 5th immunization these mice were euthanized, the spleens were removed and transferred to 10 ml RPMI-K medium. Each spleen was mashed through a cell strainer using the plunger of a 12 ml syringe and was further processed as described in 5.3.6 Generation of Mouse Hybridoma Cell Lines.

5.6 Software and Statistics

HERV env protein sequences were collected from the UniProtKB database (Bateman *et al.*, 2023) and similar DNA sequences were searched using the NCBI Basic Local Alignment Search Tool (BLAST; Altschul *et al.*, 1990).

The HERV env protein sequences were analyzed for different protein features using SignalP-5.0 (http://www.cbs.dtu.dk/services/SignalP/; Almagro Armenteros *et al.*, 2019), TMHMM Server 2.0 (http://www.cbs.dtu.dk/services/TMHMM/; Sonnhammer, Von Heijne and Krogh, 1998), FP_predict.exe (FP_predict.exe; Wu *et al.*, 2016) and the ISDTool 2.0 (ISDTool-2.0; Lv *et al.*, 2014).

Graphs and statistical analyses were done using GraphPad Prism 8 (GraphPad Software, Inc., La Jolla, USA) while illustrations were created using Affinity Designer (version 5.2.5, Serif (Europe) Ltd, Nottingham, UK) or BioRender (Toronto, Canada).

Fluorescence intensities of microscopic images were processed and measured using ImageJ (version 1.54j, NIH, USA).

In silico cloning and analysis of sequencing results were performed with the CLC Main Workbench (Version 22.0.2; QIAGEN, Aarhus, Denmark).

6 Results

6.1 Production and Characterization of Recombinant HERV Env Proteins

In order to generate HERV env-based antigens for vaccination strategies and antibody production, env sequences of different HERV strains were collected and assessed for their protein expression capacities. Furthermore, different modifications within the env sequences were tested and compared to the respective wild type (WT) variants to detect the best expressing variant.

6.1.1 Design and *in silico* Characterization of HERV Env Sequences

For the HERV strains HERV-H, HERV-FRD and ERV3.1 (also known as HERV-R), env protein sequences were collected from various DNA and protein datasets. First, at least one protein sequence per HERV was collected from the UniProtKB database. These sequences were then blasted in a protein to translated-nucleotide BLAST (tblastn) and similar sequences were searched for, from which, if necessary, a consensus sequence was calculated.

For HERV-FRD, the sequence P60508 (SYCY2_HUMAN) was retrieved from UniProtKB. The tblastn search resulted in 8 additional sequences, all identical to the input sequence (Figure S 1). Therefore, no consensus calculation was needed.

One sequence was found for ERV3.1 with the accession number Q14264 (ENR1_HUMAN) at UniProtKB. The tblastn search led to 4 further sequences. An alignment of all 5 sequences showed an overall high similarity. Nevertheless, one major mismatch was found at amino acid position 192. Two out of 5 sequences contained a tyrosine residue at this position, while the other three had a cysteine residue at the same position (Figure S 2). Therefore, two different consensus sequences were calculated: ERV3.1_192Y and ERV3.1_192C.

The search at UniProtKB for HERV-H resulted in two different sequences: Q9N2K0 (ENH1_HUMAN) and Q9N2J8 (ENH3_HUMAN). A tblastn with Q9N2K0 (ENH1_HUMAN) resulted in 8 additional sequences, including one sequence identical to Q9N2J8 (ENH3_HUMAN). All 10 sequences were aligned (Figure S 3). Since the alignment showed severe differences within the sequences regarding various amino acid positions, clusters of highly similar sequences were searched. A total of three clusters were found, the largest one containing 5 sequences including the BLAST input sequence. The finding of these three clusters is in accordance with the findings from De Parseval *et al.* who identified three large

HERV-H env open reading frames (ORFs) named HERV-H/env62, HERV-H/env60, and HERV-H/env59, referring to the respective env protein`s molecular weight (De Parseval *et al.*, 2001). A consensus sequence was calculated from the largest cluster, which resembles the HERV-H/env62 sequence. This sequence was used in all further assays and will be referred to as HERV-H env. For later analysis, the HERV-H variants HERV-H/env60 and HERV-H/env59 were also used.

Ready-to-use sequences for HERV-W and HERV-K were kindly provided by InProTher (now Hervolution Therapeutics; Copenhagen/Denmark). For HERV-K, two sequences were available: one WT sequence (HERV-K-WT) and one containing a substitution at the amino acid position 525 from glutamine to alanine. This mutation lies within the ISD and was inspired by the finding, that a similar mutation in the ISD of the HIV-1 env abrogated the immunosuppressive properties of the protein. This HERV-K variant was termed HERV-K-ISDmut (Morozov *et al.*, 2012).

All HERV env sequences were analyzed for typical retroviral env protein features including signal peptide (SP), transmembrane domain (TM), furin cleavage site, fusion peptide and immunosuppressive domain (ISD) using suitable prediction tools and literature (Figure S 4-Figure S 8).

For HERV-K, algorithms failed to predict an SP and an ISD. Yet, Ruggieri *et al.* showed in 2009, that the HERV-K env contains an unusually long 96 aa SP, which differs from SPs of class I HERV envs and was therefore not recognized by the SP prediction tool (Ruggieri *et al.*, 2009). The presence of an ISD was proven by Morozov *et al.* in 2013 who showed that the ISD of HERV-K env contains amino acids from gammaretroviruses, lentiviruses as well as the betaretrovirus mouse mammary tumor virus (MMTV) (Morozov, Dao Thi and Denner, 2013). Therefore, the SP and ISD of HERV-K were annotated according to these two publications.

Both ERV3.1 sequences were the only sequences where neither a TM nor a fusion peptide could be found. The absence of a TM is in line with the findings of Cohen *et al.* and De Parseval *et al.* Both publications showed, that the ERV3 env encodes a truncated TM subunit due to a premature stop codon and is therefore not membrane-bound (Cohen *et al.*, 1985; De Parseval *et al.*, 2003). Therefore, a TM was manually added to design a membrane-bound env protein for ERV3.1. This TM was derived from one of the sequences found in the tblastn search that contained a continuing sequence after the env stop codon. According to the TM prediction

tool, this sequence did include a TM. This is also in line with the findings of Cohen *et al.* who noticed a typical retroviral transmembrane protein sequence C-terminal of the env stop codon (Cohen *et al.*, 1985). Consequently, the first stop codon after the env sequence was replaced by an alanine, and the TM-containing part of the open-reading frame was added at the C-terminus. The absence of the fusion peptide was also seen by De Parseval *et al.* (de Parseval *et al.*, 2003).

A C-terminal HA-tag for protein detection and quantification in later experiments was added to all sequences. The sequences were RNA- and codon-optimized for efficient protein expression in human cells and ordered for subcloning into the pURVac expression vector.

6.1.2 HERV Env Sequence Optimization and Characterization

Considering the HERV env proteins as immunogens for the mAb generation or as a basis for vaccine candidates, a sufficient protein expression is important. This could possibly be enhanced by modifying the env sequences regarding the SP and TM. For this, first, the autologous SP was exchanged with a minimal version of the tissue plasminogen activator SP (mini-tPA). The use of the tPA-SP has been shown to be advantageous by enhancing the surface expression of membrane-bound proteins and consequently improving their immunogenic properties (Costa *et al.*, 2006; Luo *et al.*, 2008; Wen *et al.*, 2011).

To assess the impact on the HERV env proteins, HEK293 T cells were transiently transfected with the WT and mini-tPA DNA constructs and subsequently analyzed for protein expression using western blot and flow cytometry assays targeting the C-terminal HA-tag. In flow cytometry, no significant differences could be detected between the respective pairs of WT and the modified variant regarding protein expression (Figure 4A). In the western blot, all env proteins exhibited bands between 55 and 70 kDa, which were higher than the expected calculated molecular weight. This most likely resulted from post-translational modifications of the proteins like glycosylation. Except for both HERV-K constructs, the band intensities were comparable between the WT and the modified variant. The HERV-K constructs exhibited qualitatively stronger bands for the WT-SP variants than the variants encoding the mini-tPA SP (Figure 4B). Overall, expression levels between the different HERV env proteins, and especially between HERV-W and all other HERVs differed a lot. HERV-W exhibited the highest expression levels, followed by HERV-FRD and HERV-H (Figure 4).

In the next step, the membrane-integral part of the TM subunit and the cytoplasmic tail (Cyt) was exchanged for the TM-anchor and Cyt of the *vesicular stomatitis virus* glycoprotein (VSV-TM). This exchange has been proven to be beneficial in HIV-1 envelope proteins, especially in a viral vector vaccine context (Rabinovich *et al.*, 2014; Bresk *et al.*, 2019). Again, all variants were tested for protein expression with flow cytometry and western blot. For HERV-W and ERV3.1_192Y, a significant difference could be detected in flow cytometry between the WT and the VSV-TM. For HERV-W, a higher signal was seen for the WT-TM ($p = 0.002$), whereas for ERV3.1_192Y, a significantly higher MFI was measured for the VSV-TM variant ($p = 0.004$). Slight, but not significant tendencies towards the VSV-TM variant could be shown for HERV-H and ERV3.1_192C (Figure 4C). These results also correspond to the western blot assay, where stronger bands could be detected for HERV-H and both ERV3.1 variants comprising the VSV-TM, as well as HERV-W with the WT-TM (Figure 4D). However, the western blot also showed qualitatively stronger bands for HERV-FRD and HERV-K compared to their VSV-TM counterparts. These differences were not observed in the flow cytometry. However, the partially scattered data points in the flow cytometry, which resulted from three independent measurements, indicate that the transfection efficiency can differ between different experiment time points. Therefore, the differences seen in the western blot not corresponding to the flow cytometry could be a product of low transfection efficiency in the respective sample.

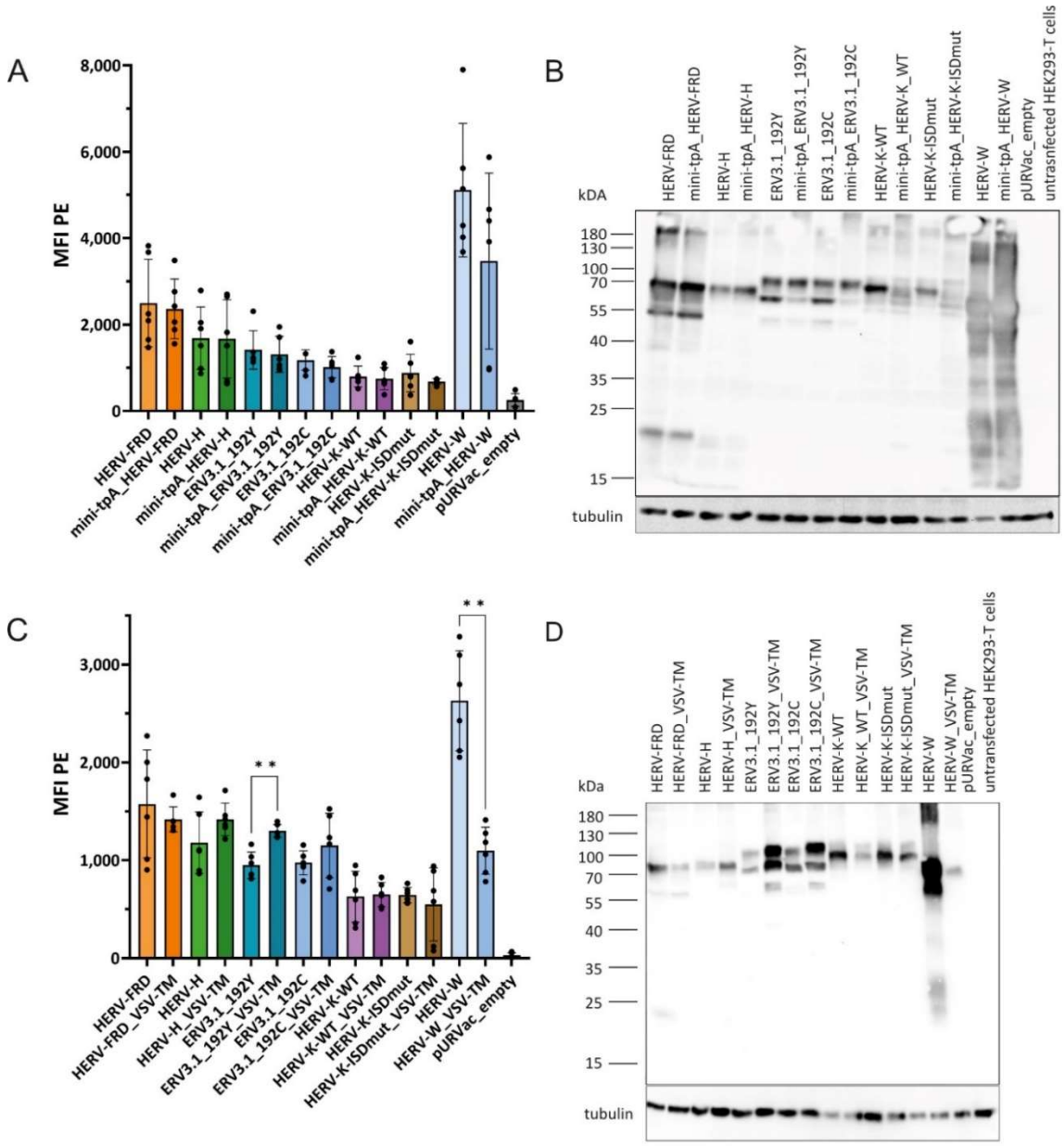


Figure 4: Flow cytometry and western blot of WT and modified HERV env constructs.

HEK293 T cells were transiently transfected with either the WT HERV env constructs or the respective SP (A + B) or TM (C + D) modified variants. An empty pURVac vector (pURVac_empty) and untransfected cells were used as negative controls. A + C: Transfected cells were intracellularly stained with an anti-HA tag antibody. The mean fluorescence intensity (MFI) of the secondary anti-mouse IgG/PE antibody was measured via flow cytometry. The MFI of untransfected cells was subtracted from each measurement. The measurements were performed in three independent experiments, each comprising technical duplicates. Each dot represents one technical replicate. The asterisks indicate significant differences between the MFI values of the different variants determined with a Brown-Forsythe and Welch ANOVA test (*: $p \leq 0.05$; **: $p \leq 0.01$; ***: $p \leq 0.001$). B + D: Cell lysates from transfected cells were generated and loaded onto a SDS gel under reducing conditions. The western blot was stained with an anti-HA tag antibody and an anti-mouse IgG/HRP secondary antibody. An anti-tubulin blot was used as loading control and the PageRuler Prestained Protein Ladder as marker for the molecular weight.

6.1.3 Characterization of Fusogenic Properties of HERV Env Proteins

Some HERV env proteins are known to fuse cells upon binding to their respective receptor (Blaise *et al.*, 2003). An immunostaining assay was performed to test the different HERV envelope proteins for fusogenicity. For this, transiently transfected HEK293 T cells were stained using the C-terminal HA-tag of the env proteins, and the signal of the HRP-coupled secondary antibody was visualized using a TrueBlue substrate. The cell morphology was imaged using a light microscope. Fusion of cells could be detected for the env proteins of HERV-W and HERV-FRD, but not for the other HERV strains (Figure 5).

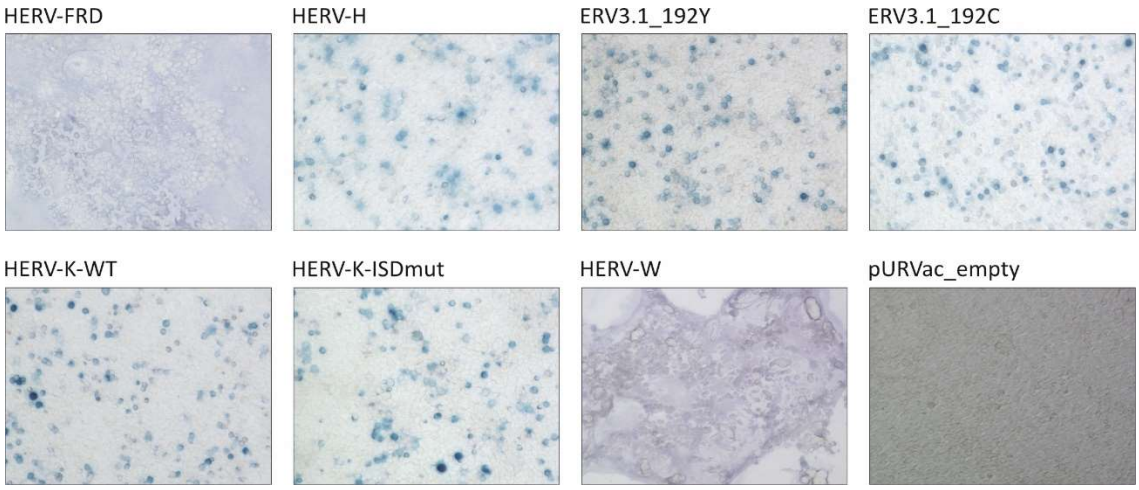


Figure 5: Immunostaining of HERV env proteins.

Transiently transfected HEK293 T cells were fixed and then stained with an anti-HA tag primary and an anti-mouse IgG/HRP secondary antibody. The signal of the HRP-coupled secondary antibody was visualized using TrueBlue substrate. The images were taken under a light microscope with a 100x magnification.

Since the fusion activity is most likely initiated by the TM part of the env protein, HERV-W and HERV-FRD variants containing the heterologous VSV-TM were also tested for fusogenicity. In the immunostaining assay, it could be shown that the exchange of the TM led to a loss of the fusogenic properties of HERV-W and HERV-FRD (Figure 6).

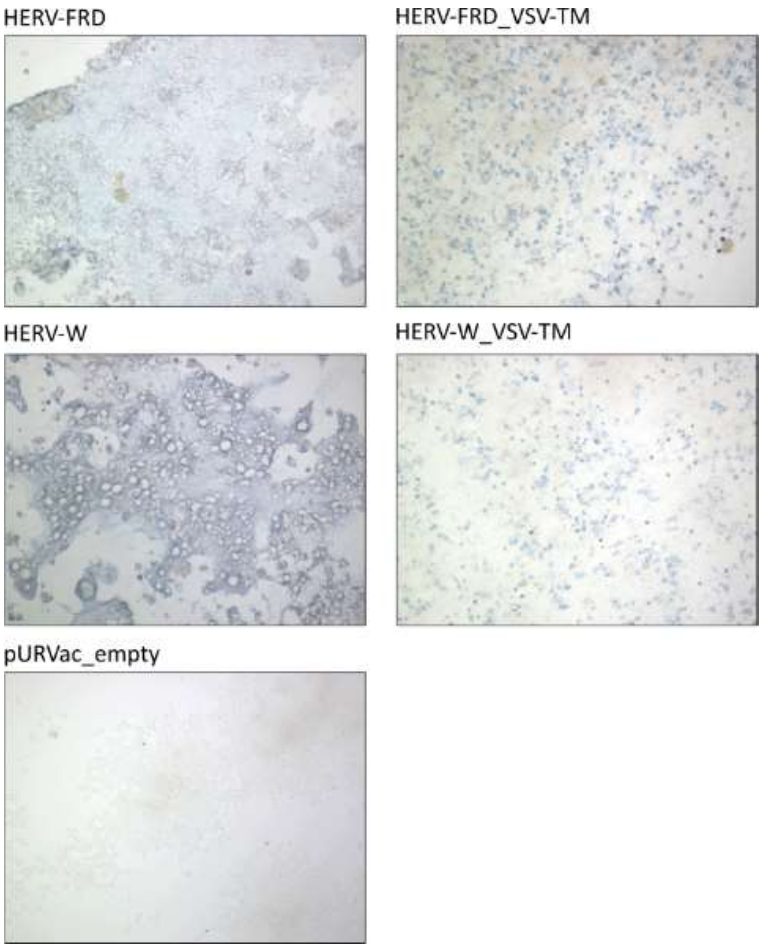


Figure 6: Immunostaining of HERV-W and HERV-FRD WT-TM and VSV-TM variants.

HEK293 T cells were transiently transfected with HERV-W and HERV-FRD plasmids encoding either the variant with the WT or the VSV-TM. Cells were treated as described in Figure 5.

Considering that the fusogenic properties of HERV-W and HERV-FRD could pose a problem for further immunization experiments or the production of recombinant viral vectors, further mutations abrogating the fusogenic activity should be introduced to the env sequence without having an impact on protein structures or needing to exchange the complete transmembrane-anchored and cytoplasmic region. For HERV-FRD, fusion-incompetent mutations were designed by knocking out several N-glycosylation sites following the design and results of Cui *et al.* (Cui *et al.*, 2016). Three variants, which proved to be fusion-incompetent in this study, were generated: HERV-FRD_N133Q, HERV-FRD_N312Q and HERV-FRD_N443Q. For HERV-W, no published mutations with limited structural impact were found. Therefore, the design of the HERV-FRD mutations was adapted to HERV-W by aligning the three N-glycosylation sites of HERV-FRD and their neighboring amino acid sequences with the HERV-W env sequence. For HERV-FRD_N443Q a homologous motif could be found in HERV-W at position 409 (Figure 7A).

Accordingly, the variant HERV-W_N409Q was generated. All mutants were tested for protein expression in flow cytometry and western blot, as well as for their fusogenic properties in an immunostaining assay comparing them to their corresponding WT env proteins. For HERV-FRD, all mutants exhibited significantly reduced protein expression compared to the WT variant in the flow cytometry read-out (Figure 7B). In the western blot, these differences were not as distinct (Figure 7C). Regarding the fusogenic properties, all mutations led to a reduction of the fusogenicity (HERV-FRD_N133Q) or completely abrogated this function (HERV-FRD_N312Q and HERV-FRD_N443Q) (Figure 7D). HERV-W_N409Q showed a similar protein expression level as the WT but did not have reduced fusogenic properties as compared with HERV-W (Figure 7C + D). Since no effect on the fusogenicity was seen for the HERV-W mutant, and no differences could be observed in the western blot, the HERV-W variants were not included in the flow cytometry assay.

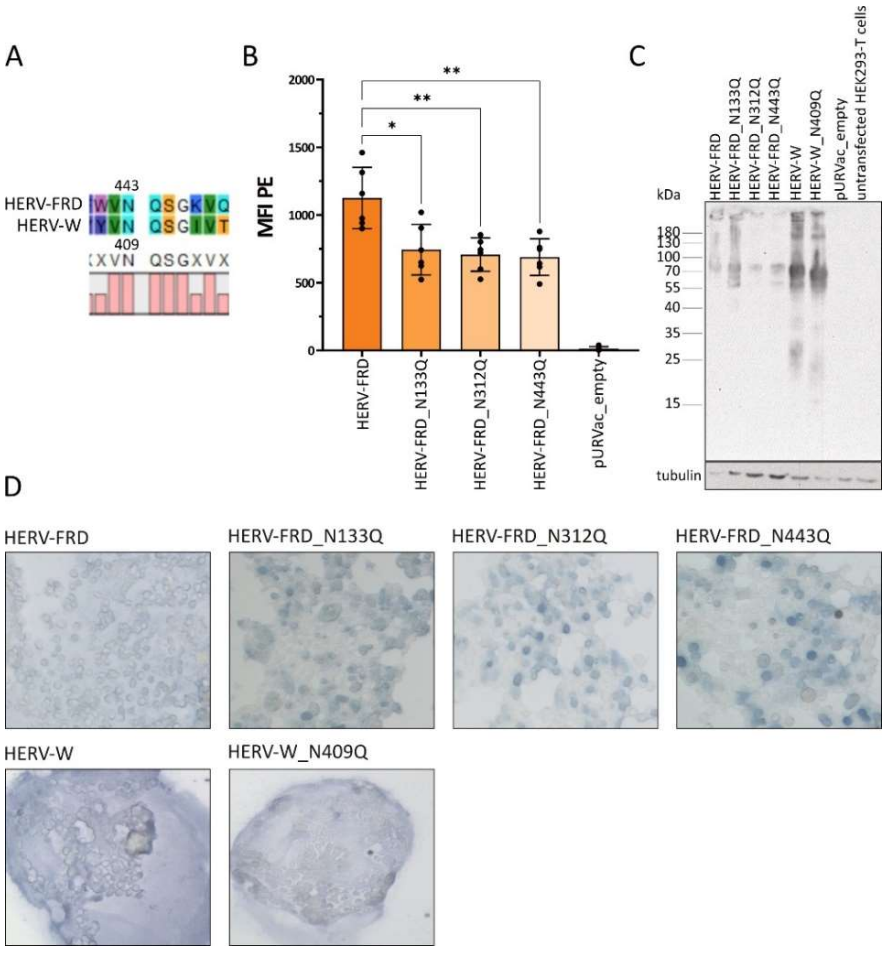


Figure 7: In vitro characterization of HERV-W and HERV-FRD fusion mutants.

HEK293 T cells were transiently transfected with HERV-W and HERV-FRD WT and potential fusion mutants. A: Alignment of HERV-FRD amino acid sequences from position 441-449 and HERV-W 407-415. The numbers indicate the position of the N-glycosylation sites that were knocked out. The alignment was performed using the alignment tool of the CLC Main Workbench 22. B: Cells were analyzed for protein expression using flow cytometry. The

measurements were performed in three independent experiments, each containing technical duplicates. Each dot represents one technical replicate. The asterisks indicate significant differences between the MFI values of the different variants determined using a Brown-Forsythe and Welch ANOVA test (*: $p \leq 0.05$; **: $p \leq 0.01$; ***: $p \leq 0.001$). C: Cell lysates were analyzed in a western blot for protein expression. D: Cells were tested for cell fusion with an immunostaining assay using TrueBlue substrate. In all assays, cells were stained with an anti-HA tag primary antibody and a suitable anti-mouse IgG PE- (A) or HRP- (B and C) labeled secondary antibody.

In summary, the results of the characterization and optimization experiments of HERV env proteins showed, that all tested HERV strains encode for an env protein that can be successfully expressed upon transfection. Furthermore, modifications within the env protein did not lead to an enhanced protein expression for most env variants, indicating that for all HERV envelope proteins, the WT sequences are optimal variants for the use in future vaccination strategies. Additionally, cell-fusion activity could only be observed in the env proteins of HERV-W and HERV-FRD, which could be abrogated for HERV-FRD.

Since the WT HERV-H env protein had one of the higher protein expression levels, and unlike HERV-W and HERV-FRD did not show cell-cell fusion activity, it was decided to focus all further work in this thesis on HERV-H. This family is not only the most abundant HERV family in the human genome, but also one of the HERVs most frequently expressed in different tumor tissues as reviewed by Bermejo *et al* (Bermejo *et al.*, 2020). This makes HERV-H an interesting candidate for further research, including immunization experiments, the generation of antibodies targeting HERV-H env and efforts toward tumor stratification.

6.2 HERV-H Env Immunization Study in BALB/c Mice

In order to assess the immunogenic properties of the HERV-H env protein, an immunization study in BLAB/c mice was conducted. For this study, different variants of the env protein were produced either as soluble recombinant protein or as DNA, encoded in the pURVac expression vector. All variants were characterized for their protein expression level prior to immunization, and the soluble proteins were biochemically characterized. This experiment served the purpose, to find the best HERV-H env candidate for further vaccination strategies, while simultaneously laying the foundation for the generation of monoclonal antibodies via the hybridoma approach (see section 6.3).

6.2.1 Generation and Characterization of Soluble HERV-H Env Antigens

Soluble proteins are useful tools not only as antigens for immunization but also as reagents for immunological read-outs like ELISA. Thus, two DNA constructs were designed and generated encoding for secreted versions of the HERV-H env protein. The first construct was designed by deleting the transmembrane-anchor and cytoplasmic domain (TM + Cyt) resulting in the secretion of a potentially trimeric protein termed solHERV-H. For the second variant, the entire TM subunit was deleted, thereby this variant encoded only the surface subunit (SUonly) leading to the secretion of a monomeric protein. To facilitate protein purification, a C-terminal His6-tag was added to all soluble variants (Figure 8).

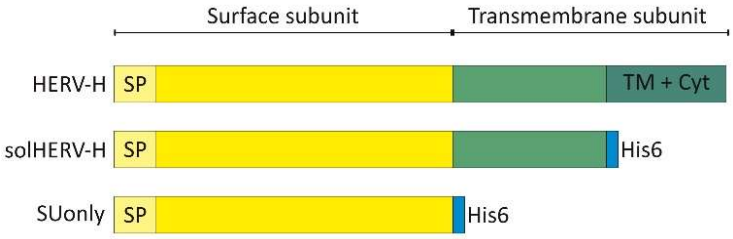


Figure 8: Construct design for soluble HERV-H env proteins.

Design of two soluble HERV-H env constructs (solHERV-H and SUonly) compared to the full-length HERV-H env. SP: signal peptide (bright yellow); TM + Cyt: transmembrane domain + intracytoplasmic tail (dark green); His6: 6x Histidine-tag (blue).

Soluble HERV-H env proteins were purified from the supernatant of transiently transfected Expi293 cells via IMAC and were characterized using SDS PAGE under reducing and non-reducing conditions. Both proteins could be expressed and purified from Expi293 cells with yields in the mg range (Table 25).

Table 25: Protein yields from the purification of solHERV-H and SUonly protein.

protein	Yield/(mg/300 ml transfection)	Yield/(mg/$l_{culture}$)
solHERV-H	1.02	3.4
SUonly	2.76	9.1

In the SDS gel analysis, the solHERV-H protein exhibited a molecular weight of approximately 70 kDa under reducing conditions, whereas the monomeric SUonly protein had a molecular weight of about 55 kDa (Figure 9A). For both proteins, the observed molecular weight was comparable to the expected, theoretical molecular weight when including possible post-

translational modifications like glycosylation. Under non-reducing conditions, both proteins exhibited a protein band above the highest band of the standard, indicating a molecular weight over 180 kDa in the non-reduced form (Figure 9B). Since the SU and TM subunit in gammaretroviral env proteins are covalently linked through a disulfide bond (see 3.3), which would not be cleaved under non-reducing conditions, a higher molecular weight would have been expected for the solHERV-H env, but the observed band exceeded the expected band size. Due to its monomeric nature, the SUonly was not expected to exhibit a higher molecular weight under non-reducing conditions, compared to the reduced SDS PAGE. Therefore, the high molecular weight bands hint at a possible aggregation rather than multimerization in both proteins. The one distinct protein band and the presence of only few weak background bands showed, that both env proteins exhibited a high purity already after the IMAC-purification.

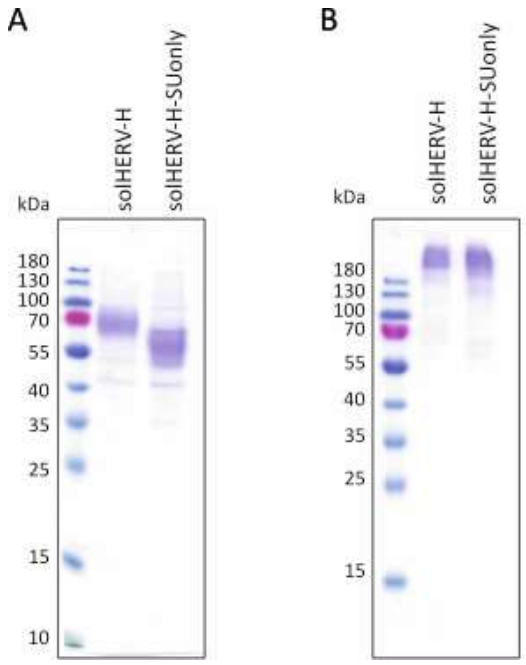


Figure 9: Reducing and non-reducing SDS PAGE of IMAC-purified soluble HERV-H env proteins.

Soluble proteins were purified from the supernatant of transfected Expi293 cells via IMAC and analyzed on an SDS gel. A: SDS PAGE under reducing conditions. B: SDS PAGE under non-reducing conditions. As molecular weight standard the PageRuler Prestained Protein Ladder was used.

To analyze the conformation of the proteins, analytical and preparative SEC analysis were performed. From the preparative SEC, protein fractions were collected and characterized on a blue native PAGE. Comparing the SEC analysis of the solHERV-H and the SUonly, it could be observed that the SUonly exhibited only one peak at an elution volume between 9-10 ml,

whereas the solHERV-H protein had two peaks at approximately 9 ml and between 10-11 ml (Figure 10A). The partial overlap of the first peak from solHERV-H with the peak from the SUonly protein confirms the hypothesis from the non-reducing SDS PAGE, indicating aggregation in both proteins. The second peak of the solHERV-H protein indicates a protein fraction with a smaller molecular weight than the presumably aggregated fraction. This could be confirmed in the blue native PAGE. Here, all fractions within both peaks (fractions 1-25) were analyzed. Protein bands could be detected from fraction 6-21, with the bands becoming more distinct from fraction 15 onwards, corresponding to the start of the second SEC peak. All protein fractions exhibited a molecular weight of approximately 400 kDa (Figure 10B). From the preparative SEC analysis of the SUonly protein, the fractions 1-25 were also analyzed on a blue native PAGE. Here, protein bands could be detected from fractions 5-24. Similar to solHERV-H, the bands became more distinct from fraction 16 onwards, running at a molecular weight of around 272 kDa. Fractions 20-22 showed two distinct bands between 150 and 272 kDa (Figure 10C).

Since solHERV-H exhibited two peaks in the SEC analysis, fractions 10-14 (corresponding to the first SEC peak) and 15-22 (corresponding to the second SEC peak) were pooled and analyzed by reducing and non-reducing SDS PAGE. Under reducing conditions, both protein pools exhibited the earlier-seen protein band at 70 kDa (Figure 11A). In the non-reducing gel, smeared bands above the 180 kDa band were observed for both protein pools. Additionally, a band at approximately 130 kDa was seen, which was more distinct for the pool of fraction 15-22 (Figure 11B). Regarding the molecular weight, this band most likely represents a kind of dimeric protein.

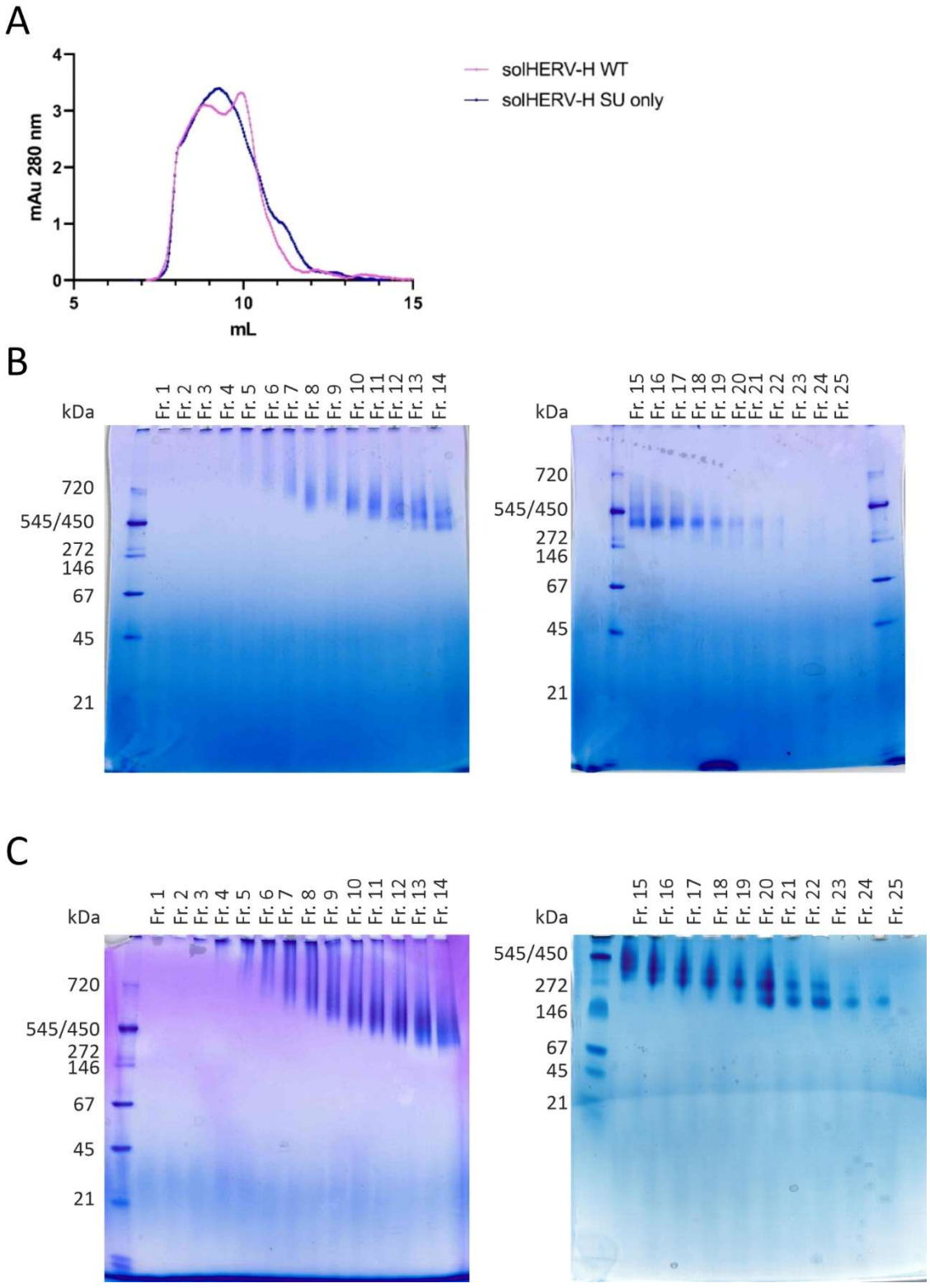


Figure 10: SEC run and corresponding Blue Native PAGE of soluble HERV-H env proteins.

The soluble proteins were analyzed via SEC using a Superdex 200 Increase 10/300 GL column followed by a blue native PAGE of the collected fractions. A: Overlay of the analytical SEC curves of solHERV-H and SUonly. B: Blue native PAGE of the fractions 1-25 of solHERV-H. C: Blue native PAGE of fractions 1-25 of the SUonly. As protein standard for both blue native PAGEs, the Serva Native Marker was used. The data displayed in this figure were collected during the master's thesis of Jan Grebner under my experimental supervision.

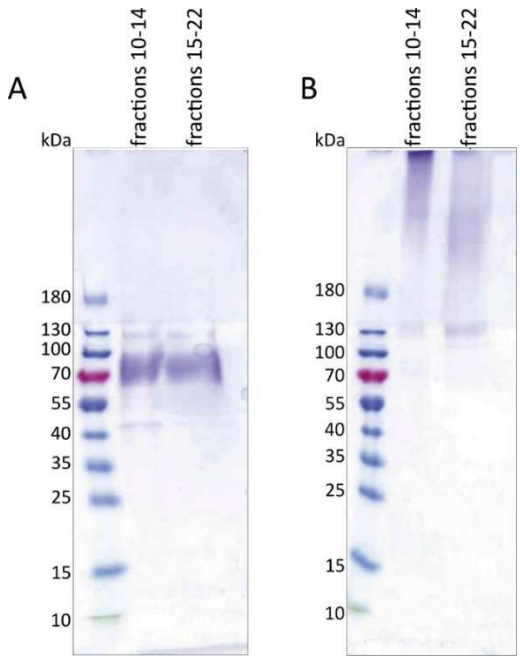


Figure 11: Reducing and non-reducing SDS PAGE of pooled protein fractions after SEC analysis.

Fractions 10-14 and 15-22 from the SEC-purified solHERV-H protein (see Figure 10B) were pooled and analyzed on a SDS PAGE under A: reducing and B: non-reducing conditions. As molecular weight standard the PageRuler Prestained Protein Ladder was used. The data displayed in this figure were collected during the master's thesis of Jan Grebner under my experimental supervision.

In conclusion, two soluble recombinant HERV-H env proteins could be produced in sufficient amounts for the immunization study and the read-out assays. The biochemical characterization via SEC and blue native PAGE showed, that both proteins, solHERV-H and SUonly, exhibit an aggregated protein form after purification. In order to stabilize the solHERV-H env in its presumed trimeric form, an artificial trimerization domain, originating from the T4 bacteriophage fibritin, was added to the C-terminus of the protein. SEC and blue native analysis showed that this resulted in a potential trimeric protein fraction (Figure S 9). However, the extensive purification steps led to a highly reduced protein yield. Since protein yields and purity of the IMAC-purified env protein were sufficient, and additional purification steps did not lead to increased purity or substantial separation of aggregated protein, it was decided to use the IMAC-purified proteins for the immunizations and read-outs.

6.2.2 Generation and Characterization of ISD-Mutated Membrane-Bound and Soluble HERV-H Env Antigens

Like env proteins of other retroviruses, the HERV-H env contains an ISD which was shown to suppress anti-tumoral immune responses in mice (Mangeney *et al.*, 2001). These immunosuppressive properties could hamper the immune reaction resulting from a therapeutic vaccination. Therefore, an inactivation of the ISD activity would be beneficial.

A study conducted on the ISD of the HERV-W syncytin-1 and the HERV-FRD syncytin-2 showed, that exchanging the amino acids at position 14 and 20 of the 20 amino acid-long ISD of syncytin-2, with that of the corresponding non-immunosuppressive syncytin-1, led to an abrogation of the immunosuppressive activity in syncytin-2 (Mangeney *et al.*, 2007). The same results were seen in the murine homologs of syncytin-1 and -2, called syncytin-A and -B. Here, the ISDs only differ in amino acid position 14, and an exchange of this single amino acid was sufficient to abrogate the immunosuppressive activity in syncytin-2. Since the ISD of HERV-H env has an 80 % sequence identity and 90 % sequence similarity to the murine syncytin-B ISD, showing the same amino acid at position 14 (Figure 12A), an ISD-mutated HERV-H env variant was generated based on the ISD mutation design for syncytin-B, replacing the lysine on position 14 by an arginine (Figure 12B).

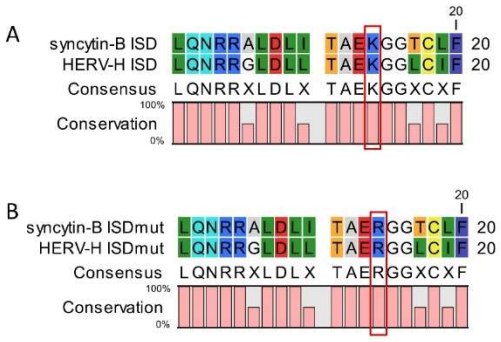


Figure 12: Alignment of the HERV-H and syncytin-B ISDs.

The ISD sequences of HERV-H and syncytin-B were aligned using the alignment tool of the CLC Main Workbench 22. A: Alignment of the WT-ISD sequences. B: Alignment of the ISDmut sequences. The red boxes mark the mutated amino acid at position 14.

Based on this design, membrane-bound and soluble HERV-H-ISDmut env protein variants were generated. Two constructs were generated for the membrane-bound variants, one containing the WT-TM and the second one containing the modified VSV-TM. Both variants were

transiently transfected in HEK293 T cells, and the protein expression was compared with the corresponding ISD-WT variants using flow cytometry and western blot assay as described in 6.1.2. Both assays showed no significant differences in the protein expression levels between the different constructs (Figure 13).

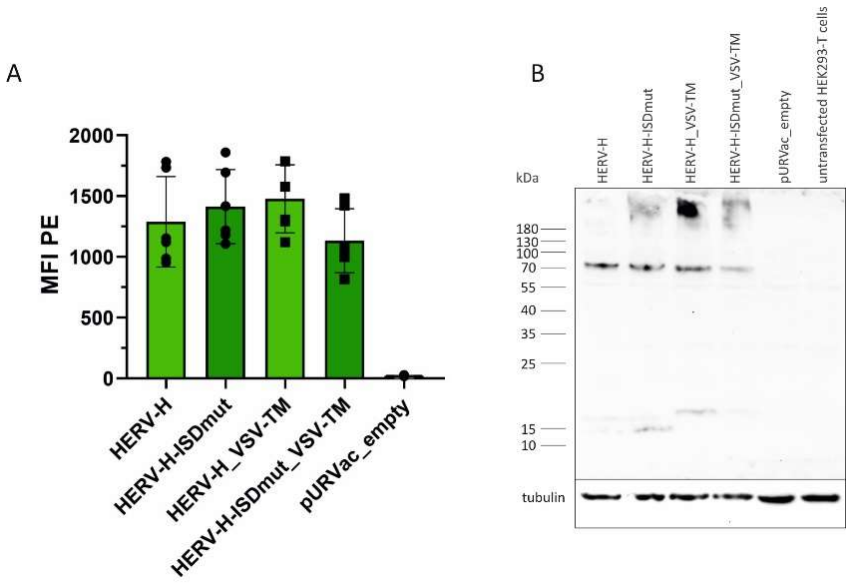


Figure 13: Characterization of membrane-bound HERV-H env ISD-WT and ISDmut variants.

HEK293 T cells were transiently transfected with the different HERV-H env constructs containing either the WT or the mutated ISD. A: Flow cytometry of transfected HEK293 T cells. Depicted is the MFI of the PE-labeled secondary antibody. The measurements were performed in three independent experiments, each comprising technical duplicates. Each dot represents one technical replicate. B: Western blot of cell lysates. Anti-tubulin was used as a loading control. For flow cytometry and western blot, cells were stained using an anti-HA tag antibody as primary antibody and a PE-labeled (A) or HRP-labeled (B) anti-mouse Ig secondary antibody.

The soluble HERV-H-ISDmut protein was purified from the supernatant of transiently transfected Expi293 cells via IMAC and characterized and compared to the WT soluble protein on a reducing and non-reducing SDS PAGE. Under both conditions, the solHERV-H-ISDmut protein showed a similar protein band as the solHERV-H WT protein at approximately 70 kDa for reduced (Figure 14A) and above the highest band of the protein ladder at 250 kDa for non-reduced samples (Figure 14B) with high purity after IMAC purification. Both protein purifications resulted in comparable yields of approximately 1 mg per 300 ml transfection.

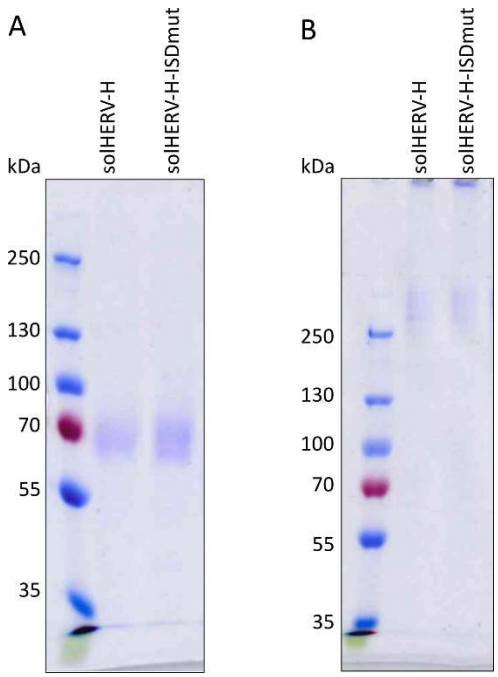


Figure 14: SDS PAGE of ISD-WT and ISDmut solHERV-H env proteins.

The purified soluble proteins solHERV-H and solHERV-H-ISDmut were analyzed on an SDS PAGE with Coomassie staining. A: Proteins were subjected to the SDS PAGE under reducing conditions. B: SDS PAGE of soluble HERV-H proteins under non-reducing conditions. As a marker for the molecular weight, the PageRuler Prestained Plus Protein Standard was used.

6.2.3 Characterization of the HERV-H Env Proteins HERV-H/env62, HERV-H/env60 and HERV-H/env59

Among the approximately 1000 genomic HERV-H copies found in the human genome, only three complete HERV-H env ORFs encompassing an ISD are known so far. These ORFs encode the three HERV-H env proteins HERV-H/env62, HERV-H/env60 and HERV-H/env59 (De Parseval *et al.*, 2001). Since these are naturally occurring variants of HERV-H env, a generated HERV-H-specific diagnostic or therapeutic antibody should be able to recognize all three of these variants. Therefore, it was assessed if all three env proteins get expressed equally to ensure in later antibody testing if potential binding differences are a result of the antibody property or different protein expression level. The three sequences were collected from annotated database protein sequences (HERV-H/env62: NCBI GenBank number CAB94192.1; HERV-H/env60: NCBI GenBank number CAB94193.2; HERV-H/env59: NCBI GenBank number CAB94194.1), RNA- and codon-optimized for efficient protein expression in human cells and subsequently ordered for subcloning into the pURVac expression vector. Two variants of each construct were designed, one containing the C-terminal HA-tag and one without a tag. An alignment of all three env protein sequences is shown in Figure S 10.

The protein expression of the variants was characterized by flow cytometry and western blot assay staining against the HA-tag. In the flow cytometry assay, the HERV-H/env62 protein exhibited a significantly higher surface expression than the other two variants with a p-value of 0.014 when compared to HERV-H/env60 and p = 0.04 when compared to HERV-H/env59 (Figure 15A). This was qualitatively confirmed by the western blot analysis, where a stronger and more distinct band at 70 kDa could be seen for HERV-H/env62 than for HERV-H/env60 and HERV-H/env59 (Figure 15B). Between HERV-H/env60 and HERV-H/env59, no significant difference could be detected.

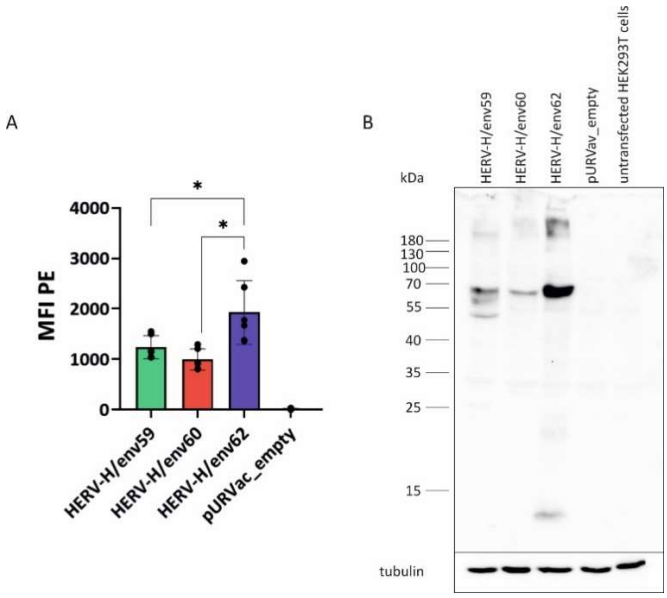


Figure 15: In vitro characterization of the three HERV-H env ORFs.

The three env ORFs of HERV-H were characterized using flow cytometry and western blot. A: Transiently transfected HEK293 T cells were intracellularly stained for HERV-H env protein expression. Depicted is the MFI of the PE-labeled secondary antibody. The measurements were performed in three independent experiments, each comprising technical duplicates. Each dot represents one technical replicate. The asterisks indicate significant differences between the MFI values of two variants determined using a Brown-Forsythe and Welch ANOVA test (*: $p \leq 0.05$; **: $p \leq 0.01$; ***: $p \leq 0.001$). B: Western blot of transiently transfected HEK293 T cells. Anti-tubulin staining was used as a loading control. For both assays, an anti-HA tag antibody was used as a primary antibody which was detected by a PE- (A) or HRP-coupled (B) anti-mouse secondary antibody. The data were generated by Anna Niebauer during her internship under my experimental supervision.

Concluding, these results show that the two proteins HERV-H/env60 and HERV-H/env59 get expressed in lower amounts than the HERV-H/env62 protein. This should be kept in mind when testing generated antibodies targeting HERV-H env, since higher binding to the env62 will most likely be ascribed to the overall protein expression level, and not the binding capacity of the antibody.

6.2.4 Immunization of BALB/c Mice with HERV-H Env Antigens

To assess the immunogenicity of different HERV-H env antigens, BALB/c mice were immunized, and antibody titers were compared. With this experiment, the best HERV-H env variant for future vaccination strategies should be assessed. Furthermore, it served the purpose to test, whether the HERV-H ISD affects antibody production in mice. So far, for HERV-H env, the immunosuppressive effect in mice was shown for anti-tumoral immune responses, which are most likely induced by a cellular immune response (Mangeney *et al.*, 2001). Effects on a humoral immune level have not yet been analyzed for HERV-H. Additionally, spleens of immunized mice having high antibody titers should be used for the generation of mAb using the hybridoma approach by Köhler and Milstein (Köhler and Milstein, 1975).

For the immunization experiment, adult female BALB/c mice were randomly distributed into 9 groups with 6 mice per group. The mice were immunized on four time points subcutaneously in the lower limb, either with a DNA vector encoding different HERV-H env variants or the soluble HERV-H env protein. The DNA vaccine candidates comprised the HERV-H ISD-WT env, either with the WT-TM or the modified VSV-TM, as well as the HERV-H-ISDmut variants, also with both TM variations. Additionally, three plasmids encoding for soluble env proteins were administered. These variants included the solHERV-H WT env, solHERV-H-ISDmut and the SUonly construct (Figure 16). All DNA constructs were encoded in the DNA vaccine vector pURVac containing a CMV promoter and an intron to enhance expression levels.

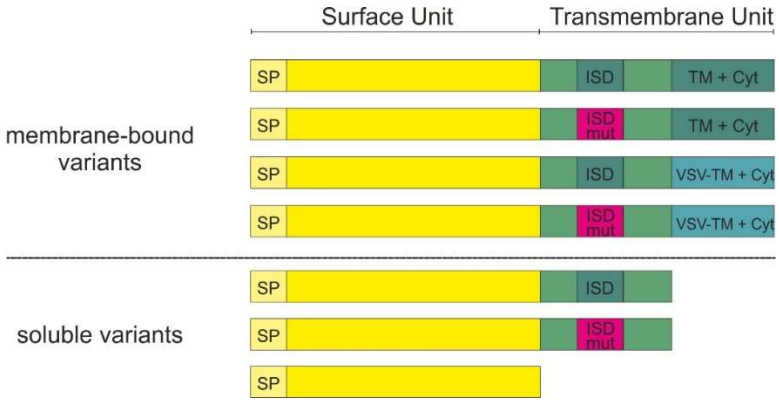


Figure 16: Schematic overview of the DNA vaccine constructs.

Four membrane-bound and three soluble DNA variants were used as DNA vaccines. SP: signal peptide; ISD: immunosuppressive domain; ISDmut: ISD with an amino acid substitution at position 14 from lysine to arginine; TM + Cyt: WT transmembrane-anchor domain and cytoplasmic tail; VSV-TM + Cyt: modified TM + Cyt of the *vesicular stomatitis virus* glycoprotein.

Except for groups 5 and 6, all groups were primed with two DNA vaccinations, followed by two protein boost immunizations. The protein for the booster immunization was adjuvanted with MPLA and chosen according to the DNA variant used for the prime. Thus, immunization with an ISD-WT DNA vector was followed by a WT env protein boost, while immunizations with the ISDmut variants were followed by the soluble ISDmut protein. Group 5 was immunized with the WT env protein on all 4 occasions, and group 6 served as a control group and was immunized two times with the empty DNA vector (pURVac_empty), followed by two immunizations with PBS (Table 26). Blood samples were taken on three occasions: one day before the first immunization (prebleed), two weeks after the second immunization (post DNA prime, 2nd bleed), and two weeks after the 4th immunization (3rd bleed). Selected mice having a high antibody titer were immunized a 5th time with protein and the spleen was extracted 3-4 days later for the hybridoma fusion (Figure 17).

Due to health issues of mice in form of spontaneous tumor growth, three mice could only receive one protein boost (mice #6 from groups 1B and 3B and mouse #3 from group 2B). Nevertheless, blood samples from these mice were taken at the time point of euthanasia.

Table 26: Immunization overview of the HERV-H immunization experiment (n = 6 BALB/c mice per group).

Group	1st immunization Day 0	2nd immunization Day 14	3rd immunization Day 42	4th immunization Day 70
1a	DNA HERV-H-WT	DNA HERV-H-WT	Protein solHERV-H	Protein solHERV-H
1b	DNA HERV-H-ISDmut	DNA HERV-H-ISDmut	Protein solHERV-H-ISDmut	Protein solHERV-H-ISDmut
2a	DNA HERV-H_VSV-TM	DNA HERV-H_VSV-TM	Protein solHERV-H	Protein solHERV-H
2b	DNA HERV-H-ISDmut_VSV-TM	DNA HERV-H-ISDmut_VSV-TM	Protein solHERV-H-ISDmut	Protein solHERV-H-ISDmut
3a	DNA solHERV-H	DNA solHERV-H	Protein solHERV-H	Protein solHERV-H
3b	DNA solHERV-H-ISDmut	DNA solHERV-H-ISDmut	Protein solHERV-H-ISDmut	Protein solHERV-H-ISDmut
4	DNA solHERV-H-SUonly	DNA solHERV-H-SUonly	Protein solHERV-H	Protein solHERV-H
5	Protein solHERV-H	Protein solHERV-H	Protein solHERV-H	Protein solHERV-H
6	pURVac_empty	pURVac_empty	PBS	PBS

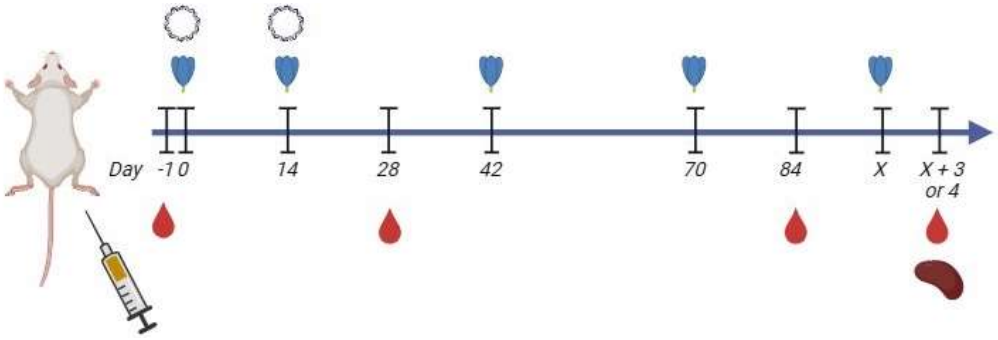


Figure 17: Schematic immunization schedule of the HERV-H immunization experiment.

Mice were immunized on days 0 and 14 either with DNA or protein and on days 42 and 70 with protein. Blood samples were taken on day 0 (prebleed), day 28 (2nd bleed) and day 84 (3rd bleed). High titer mice were immunized a 5th time with protein (day x) and the mice were sacrificed 3-4 days later, extracting the spleen and optionally taking another blood sample (x + 3 or 4). Created with BioRender.com.

From the blood samples, the serum was extracted and screened for antibodies binding HERV-H env using a direct ELISA with the IMAC-purified recombinant solHERV-H env protein (see 6.2.1a) as read-out antigen. A single-point measurement from the 2nd bleed sera showed the presence of HERV-H env antibodies in all groups, but not all mice, after the first two DNA immunizations (Figure 18, 2nd bleed). Compared to the DNA groups, where only few mice per group had antibodies varying in strength, group 5, immunized with protein instead of DNA, exhibited uniformly high antibody responses for all mice in the second bleed. The protein boost elevated the antibody responses for all mice to a comparable level, abolishing the difference in overall antibody levels that has been seen before between the protein group and the DNA groups (Figure 18, 3rd bleed).

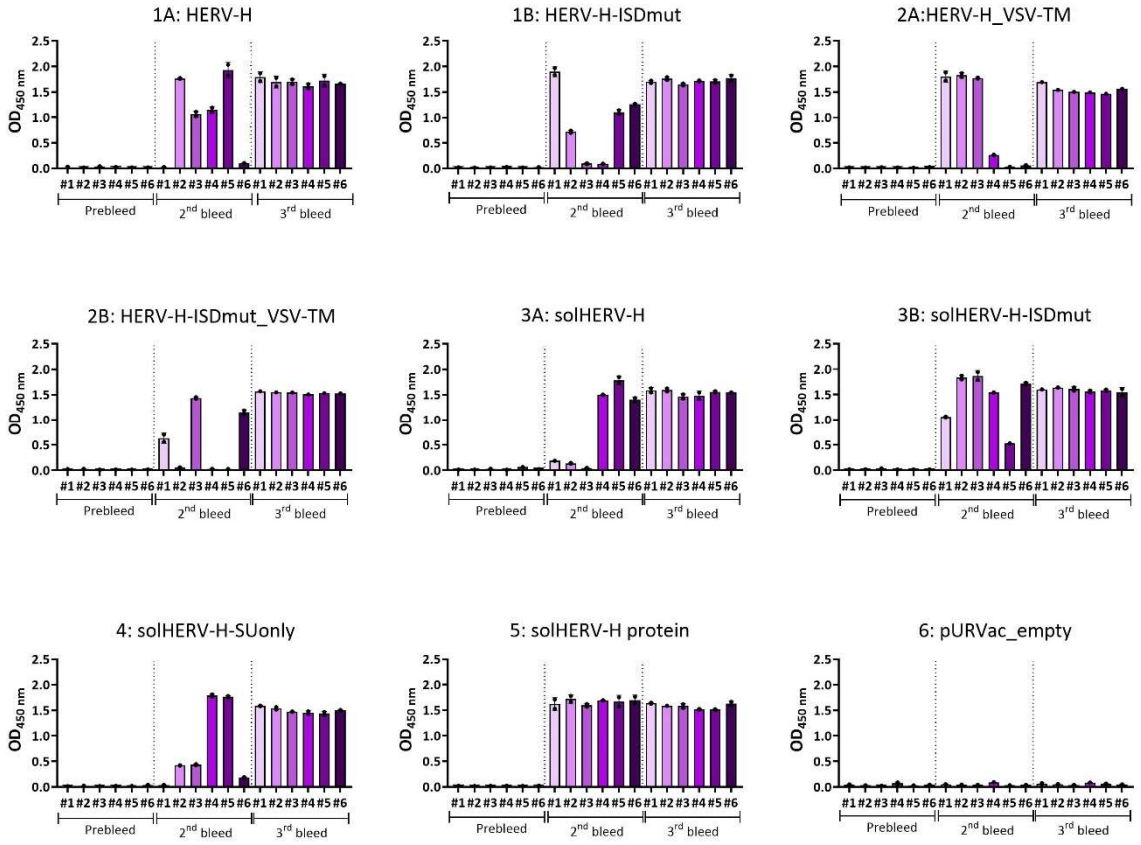


Figure 18: Single-point ELISA of sera from all time points.

Sera of all mice from all three time points (prebleed, 2nd bleed, and 3rd bleed) were analyzed for HERV-H env binding antibodies in a single-point ELISA in a 1:50 dilution. The recombinant solHERV-H env protein was directly coated to the Nunc MaxiSorp plates and an HRP-coupled anti-mouse Ig antibody was used as a secondary antibody. Each bar represents one mouse with the dots showing the values of technical ELISA duplicates.

To get a better insight into the antibody binding titer, a titration ELISA was performed with the 3rd bleed sera starting at a 1:50 serum dilution and following a 4-fold serial dilution (Figure S 11). The EC50 values for all mice were calculated from the titration curves. Comparing the EC50 values of all individual mice within one group, it is noticeable that, while the titers in the protein immunized group 5 are uniformly high, the titers in the DNA groups are more distributed with some mice having higher and some lower antibody titers. Especially mice receiving only one protein boost (mice #6 from groups 1B and 3B and mouse #3 from group 2B) had lower antibody titers than the other mice in their respective group. When comparing the grouped EC50 values between all immunization groups, group 4 exhibited the highest antibody titers. However, the differences between the other groups and group 4 were not significant, except for group 2A. Also, no differences between the ISDmut and the corresponding WT-ISD groups could be observed (Figure 19).

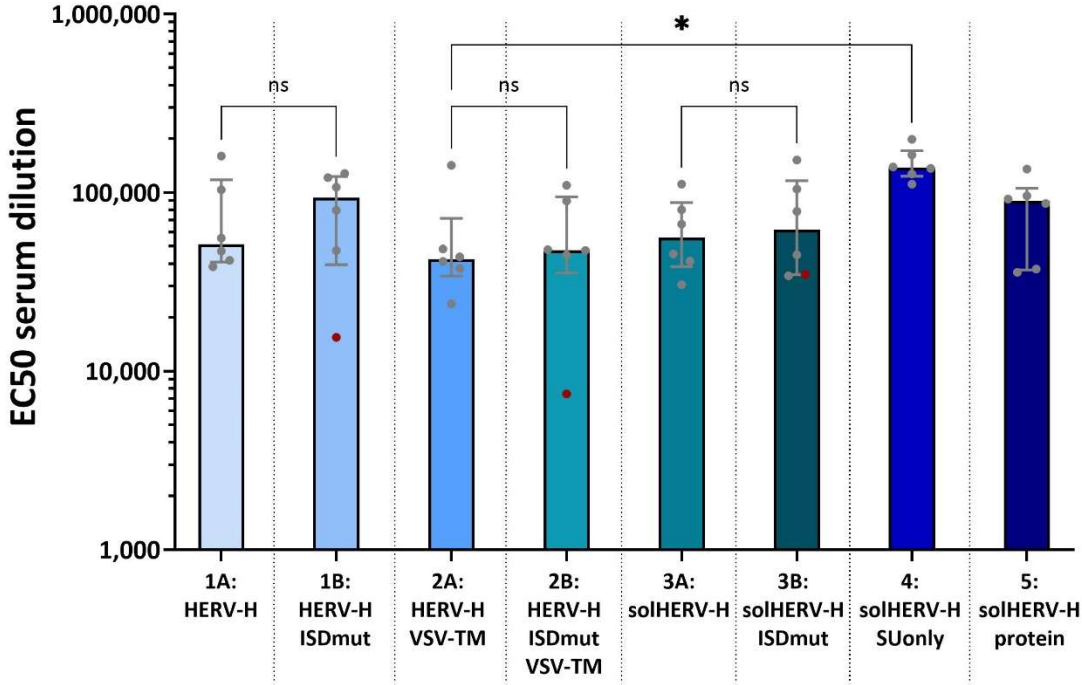


Figure 19: EC50 values of all 3rd bleed sera.

The EC50 values were calculated from a serum titration ELISA starting at a 1:50 dilution and following a 4-fold serial dilution (Figure S 11). The recombinant solHERV-H env protein was directly coated to the Nunc MaxiSorp plates and as a secondary antibody, an HRP-coupled anti-mouse Ig antibody was used. Each dot represents the mean value of technical duplicates from one mouse. The red dots indicate the three mice, that only received one protein boost immunization. The bars represent the median EC50 value within one group and the interquartile range is shown. The asterisks indicate significant differences between two groups determined using a Kruskal-Wallis test (ns: not significant; *: $p \leq 0.05$; **: $p \leq 0.01$; ***: $p \leq 0.001$).

Next, it was determined which broad regions of the HERV-H env protein were targeted by the antibodies. Therefore, a serum titration ELISA against the SUonly protein and the extracellular part of the TM subunit called the ectodomain was conducted and the EC50 values were analyzed (Figure S 12, Figure S 13). The ectodomain protein was kindly produced and provided by our cooperation partners at InProTher. All mice developed antibodies targeting the SUonly as well as the ectodomain. In all groups, the antibody titers against the SUonly region were higher than the titers of antibodies targeting the ectodomain with the differences being significant for all groups except 3A, 3B and 4 (Figure 20). Group 4, which was immunized with the SUonly DNA, also showed antibodies against the ectodomain since the protein boost was performed with the complete soluble env protein.

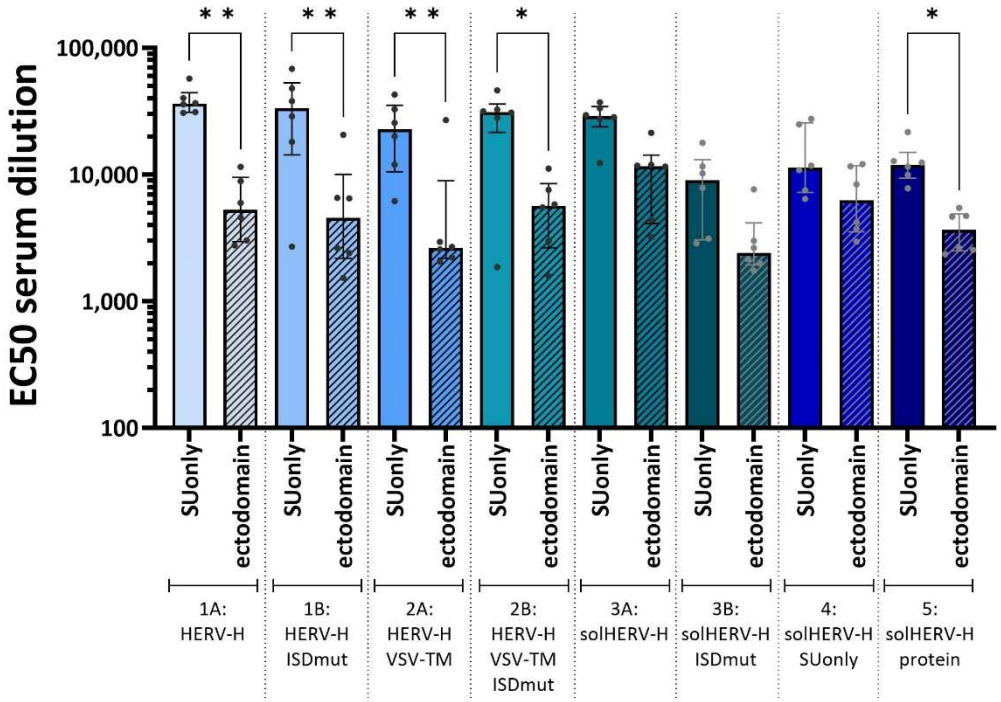


Figure 20: EC50 values for two different target-regions of the HERV-H env protein.

The 3rd bleed serum was titrated as described in Figure 19 against the SUonly and the ectodomain protein of HERV-H (Figure S 12, Figure S 13). As a secondary antibody, an HRP-coupled anti-mouse Ig antibody was used. Each dot represents the mean value of technical duplicates from one mouse. The bars indicate the median EC50 value within one group and the interquartile range is shown. The asterisks indicate significant differences between the mean EC50 values of the two target regions within one group determined using a Kruskal-Wallis test (*: $p \leq 0.05$; **: $p \leq 0.01$; ***: $p \leq 0.001$).

In conclusion, all HERV-H env variants tested elicited a specific antibody response against at least two regions of the env protein. Especially immunization with the recombinant soluble solHERV-H protein resulted in high antibody titers. On the other hand, the DNA vector immunization also showed to be effective, thereby giving an alternative to viral vector vaccines, if immunization with the full-length env is desired. An effect of the HERV-H env ISD on the humoral immune response in mice could not be observed in this study.

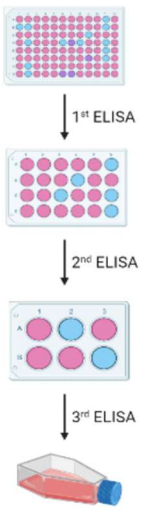
6.3 Generation of Monoclonal Antibodies

Monoclonal antibodies (mAbs) can be useful therapeutic tools to treat cancer by targeting specific tumor antigens expressed on the surface of tumor cells. Over the years, several therapeutic mAbs have been approved for cancer treatment (Zahavi and Weiner, 2020), the first one being the CD20-targeting mAb rituximab. This mAb showed to be effective in B-cell

lymphoma (Maloney *et al.*, 1997). MAbs can be administered as the recombinant antibody itself or be used as modules for CAR-T cells. Furthermore, mAbs can be used as a diagnostic tool for tumor stratification to determine which cancer types are suitable for treatment by targeting the respective antigen. The hybridoma method, which was developed by George Köhler and Cesar Milstein in 1975, is the most frequently used technology to generate mAbs. Here, splenocytes of mice immunized with the desired target antigen are fused with a myeloma cell line thus generating clones of immortal single B cells producing one specific type of antibody (Köhler and Milstein, 1975).

Here, this method was used for the generation of HERV-H env-specific mAbs. For this, 5 mice of the HERV-H immunization experiment that had high antibody titers were immunized a 5th time with the soluble recombinant solHERV-H env protein, and 3-4 days later the mice were sacrificed and the spleen extracted (Figure 17). Splenocytes were fused with the myeloma cell line Sp2/0-Ag14. The mice chosen for the procedure were mice 1 and 4 from group 1A, mice 4 and 5 from group 1B and mouse 6 from group 4. The resulting hybridoma clones were screened for HERV-H env-specific antibody secretion in three screening steps using an ELISA directed against the solHERV-H env protein (Figure 21A). Positive clones were expanded after each screening step and clones with an OD ≥ 0.1 in the last screening were defined as terminally positive. The first fusion of splenocytes from mouse 1 of group 1A failed and did not result in any HERV-H env antibody-positive clones. From the other 4 fusions, 2126 hybridoma clones were screened in total. Of these 2126 clones, 19 were defined as terminally positive for antibodies targeting HERV-H env after the three screening steps (Figure 21B). These clones were named according to the fusion and the position in the ELISA plate of the first ELISA screening (for example 2.7D12 is a clone from the 2nd fusion which was analyzed in the first ELISA on plate 7, position D12). Two of the 19 clones (3.5C11 and 3.6D9) died before they could have been cryopreserved, therefore only 17 clones were available for further analysis. Supernatants from cultures of most clones reached an OD between 0.1 and 0.5 in the final screening and only the 5 clones 2.7D12, 2.4C12, 4.3H10, 4.5H10 and 5.2E1 had an OD higher than 0.5 (Figure 21C).

A



B

mouse	Colonies harvested	1st ELISA screening	2nd ELISA screening	3rd ELISA screening
Group 1A m4	672	24/672	3/24	2/3
Group 4 m6	840	39/840	15/39	8/15
Group 1B m4	501	10/501	8/10	6/8
Group 1B m5	113	8/113	3/8	3/3
total	**2126**	**81/2126**	**29/81**	**19/29**

C

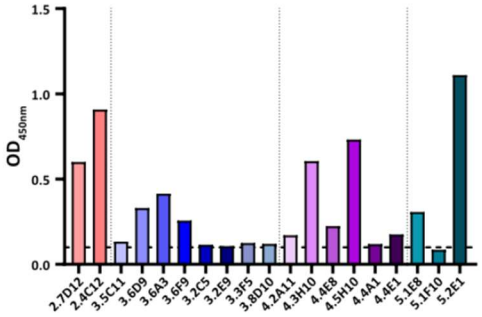


Figure 21: Overview of HERV-H env hybridoma generation.

Hybridoma clones were generated using the ClonaCell™-HY Hybridoma kit from StemCell. A: The clones that resulted from the fusion were screened for antibodies targeting HERV-H env through three ELISA screening steps. Positive clones from the first ELISA round were transferred from a 96-well plate into a 24-well plate and were expanded. After 3-4 days, a second ELISA was performed, and positive clones were transferred into a 6-well plate. After another expansion time, the clones were screened a third time. Clones that were still positive in the third screening were termed terminally positive, expanded further and finally cryopreserved. This figure was created with BioRender.com. B: Overview of all clones resulting from the 4 successful fusions and the ratio of positive clones per screening. C: OD of the finally defined positive hybridoma clones after three ELISA screenings. The vertical dotted lines separate clones from the different fusions while the horizontal dotted line indicates the cut-off for positive hybridomas at an OD of 0.1.

To further analyze the antibodies produced by the hybridoma clones, the Ig class of the antibodies was determined. For this, an ELISA was conducted using secondary antibodies specifically recognizing mouse-IgG1, -IgG2a, -IgG2b, -IgG3, or -IgM isotypes. All hybridoma clones showed a high signal for the IgM-specific secondary antibody, confirming IgM as the antibody class of the hybridoma-produced antibodies (Figure 22). As control, the same secondary antibody that was used in the screening to determine positive hybridoma clones was used. This antibody recognizes multiple mouse Ig classes. Comparing the results of the last screening ELISA (Figure 21C) with the ones from the multiple Ig-class control ELISA (Figure 22, Ig multiple) differences in the obtained OD from individual clones can be observed. These differences most likely result from variable cell densities at the time, the supernatants were analyzed.

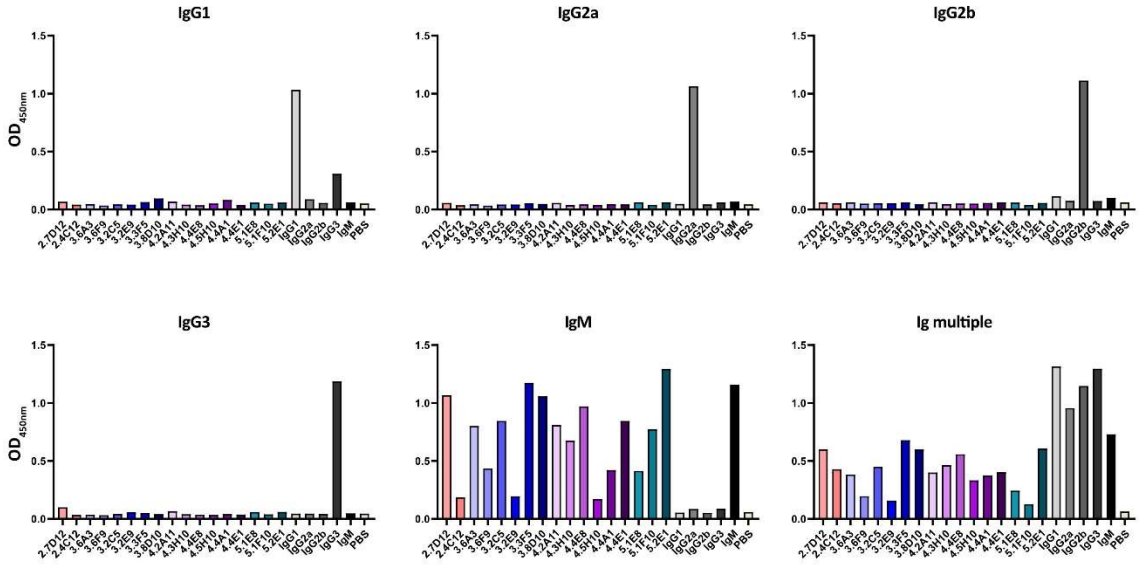


Figure 22: Ig class ELISA of hybridoma clones.

Hybridoma clones were analyzed for the Ig class of produced antibodies by ELISA using different HRP-coupled mouse Ig class-specific secondary antibodies. As positive controls, mouse Ig class-specific isotype control antibodies were used. The hybridoma supernatants and the isotype control antibodies were directly coated to the MaxiSorp plate.

Additionally, the antibodies of all clones were tested for their binding specificity to exclude cross-reactivity to other proteins. Therefore, an ELISA against different His6-tagged or non-tagged soluble proteins was performed. The proteins used in this assay included the His-tagged solHERV-H env protein as a positive control, a His6-tagged stabilized HA protein from the influenza strain A/England/195/2009 (solStabHA30_His6) and a His6-tagged (BG505_His6) as well as a non-tagged (BG505_noTag) soluble env protein from the HIV-1 clade BG505. These proteins were chosen due to their representation of another retroviral env (BG505) and a non-retroviral surface protein. The non-HERV-H env proteins were produced, purified, and kindly provided by Michael Schachtner (AG Wagner).

For all tested proteins, signals could be observed for all hybridoma clone supernatants, indicating strong cross-reactivity of the hybridoma antibodies. Only the serum from one of the mice used for the hybridoma fusion showed HERV-H env-specific binding and no binding to the non-HERV proteins (Figure 23).

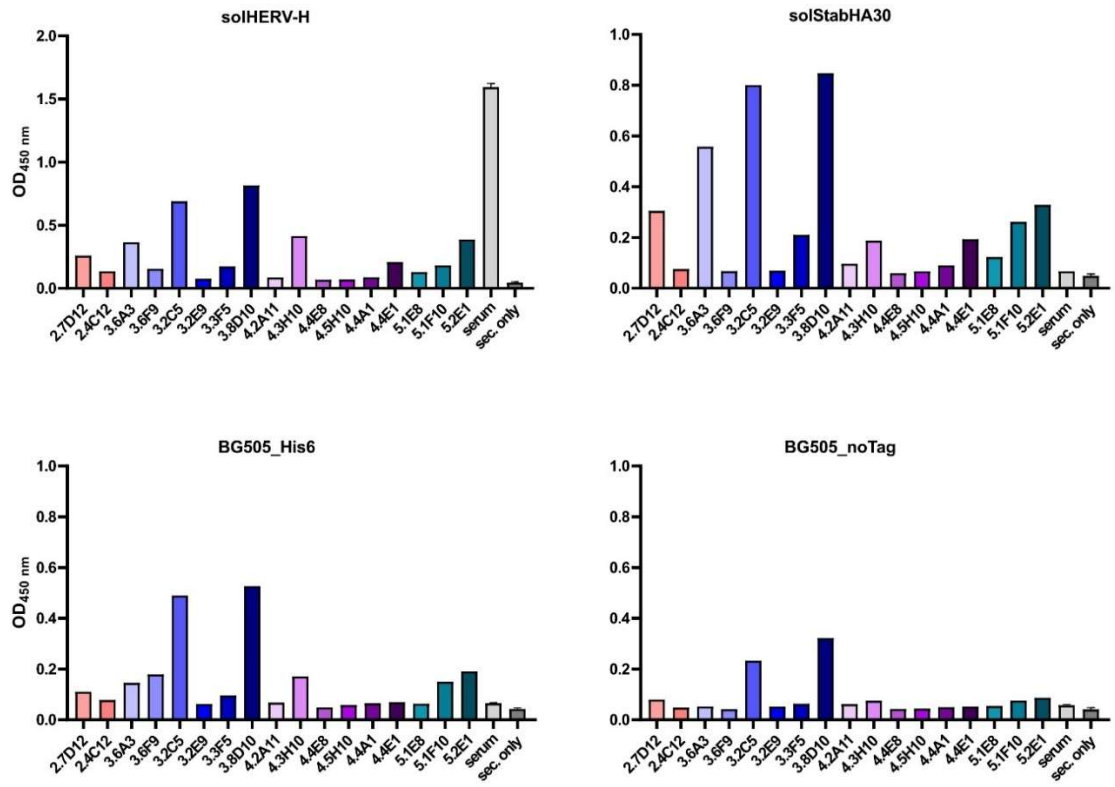


Figure 23: Cross-reactivity ELISA screening of hybridoma clones.

Hybridoma clones were analyzed for cross-reactive antibody binding by ELISA against three non-HERV recombinant proteins. As a secondary antibody, an HRP-coupled anti-mouse Ig antibody was used. As control a serum from one of the mice used for the hybridoma fusion was used.

As none of the clones from the analysis above exhibited specific HERV-H env reactivity, a second attempt to generate such mAbs was initiated. For this, mice were first immunized with HERV-H env RNA lipid nanoparticles (LNPs), which were produced and kindly provided by Dr. Martina Pfranger and Dominik Seidl from the Wagner lab. The immunizations and hybridoma fusions were performed by Davids Biotechnologie (Regensburg/Germany) as a contractor. The mice were immunized with two different immunization protocols in two groups, each comprising three mice. Group 1 followed the proposed protocol from the contractor, immunizing the mice every second week on 5 time points with 10 µg RNA-LNPs. Blood samples were taken one day before the first immunization, in week 5 and in week 9 (Figure 24A). The second group got immunized on three time points, each 4 weeks apart with 2 µg mRNA-LNPs, and blood samples were collected one day prior to the first immunization, in week 6 and in week 9 (Figure 24B).

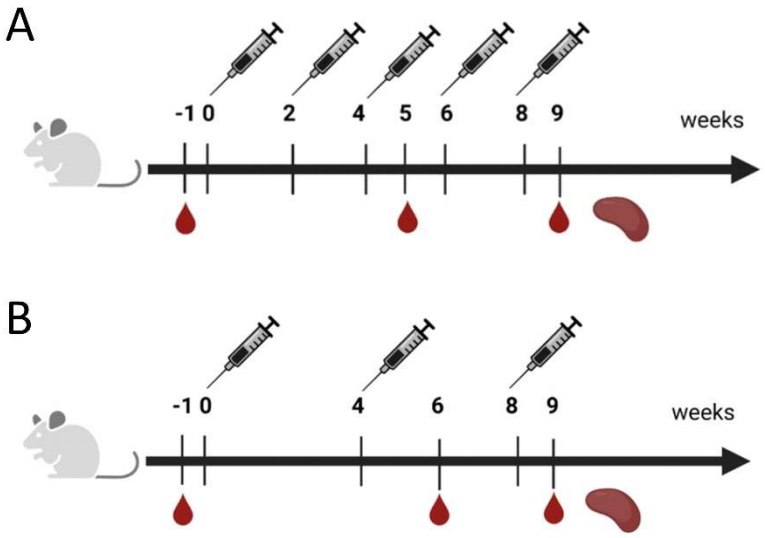


Figure 24: Immunization schedule for mRNA-LNP immunizations.

For the HERV-H env mRNA-LNP immunizations, two different immunization schedules were performed. A: Mice were immunized with 10 µg mRNA-LNPs in week 0, 2, 4, 6 and 8. Blood samples were collected one day before the first immunization, in week 5 and in week 9. B: Mice were immunized with 2 µg mRNA-LNPs in week 0, 4 and 8 and blood samples were collected before the first immunization, in week 4 and in week 9. All immunizations were performed by Davids Biotechnologie. This figure was created and kindly provided by Dr. Martina Pfranger using BioRender.com.

Since the RNA-LNP immunization did not lead to a detectable antibody titer in any of the mice (Figure S 14), further boost immunizations with the WT HERV-H env DNA vector were conducted, the first one in week 12, followed by another immunization in week 14 and two final immunization in weeks 20 and 24. After these additional DNA immunizations, some mice showed stable antibody titers, and those with the highest titer were chosen for the hybridoma generation. From three fusions, only one hybridoma clone termed clone 41 (C41) showed a stable antibody expression, the other two clones (clone 68 and clone 77, both IgG-class antibodies) lost their antibody expression over time (data not shown). The Ig-class analysis revealed that the mAbs produced by C41 belonged to the IgM class and the cross-reactivity screening showed specific binding to the HERV-H env protein (Figure 25A + B).

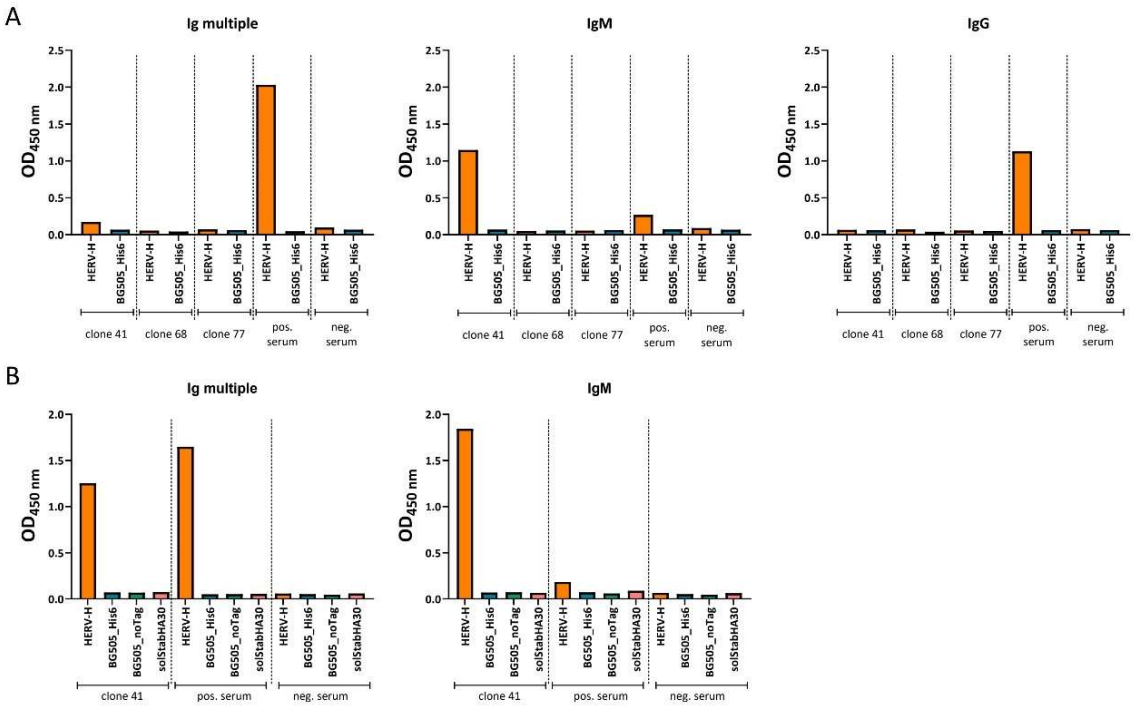


Figure 25: Antibody screening of hybridoma supernatants produced by Davids Biotechnologie.

Hybridoma supernatants were screened for their HERV-H env antibody binding via ELISA using different mouse Ig class-specific secondary antibodies. A: Hybridoma clones 41, 68 and 77 were tested for their mAb production against the soluble HERV-H env protein and the HIV BG505_His6 protein. B: Mabs from the supernatant of clone 41 were screened for cross-reactivity towards BG505_His6, BG505_noTag and the solStabHA30 proteins. In all assays, serum from HERV-H env antibody positive (pos. serum) and negative (neg. serum) mice were included as controls.

The clone 41 was further subcloned two times by the contractor, resulting in the three mAbs producing subclones 41.8.23, 41.18.6 and 41.18.85. The mAbs were purified from these clones via protein A purification. The mAb purification resulted in antibody yields ranging from 0.04-0.08 mg (Table 27).

Table 27: Yields of purified anti-HERV-H env IgM mAbs.

mAb name	Concentration/(mg/ml)	Volume/ml	Yield/mg
41.8.23	0.08	35	2.8
41.18.6	0.04	25	1
41.18.85	0.08	17	1.3

All three mAbs were tested in a western blot assay for their env-binding properties. Therefore, DF-1 cells were transfected with the three HERV-H env plasmids encoding for HERV-H/env62, HERV-H/env60 and HERV-H/env59. Cell lysates were separated by SDS PAGE and subjected to western blot analysis. The mAbs were applied to the western blots in two different dilutions (1:500 and 1:1000). As controls, one blot was stained with a polyclonal anti-HERV-H env

antibody (see section 6.2.3) and one blot was stained with an anti-tubulin antibody (Figure 26D + E). None of the mAbs showed binding to any of the env proteins, whereas binding to all three proteins could be detected in the control blot (Figure 26A-D). The tubulin control showed equal amounts of cell lysates for all variants tested (Figure 26E).

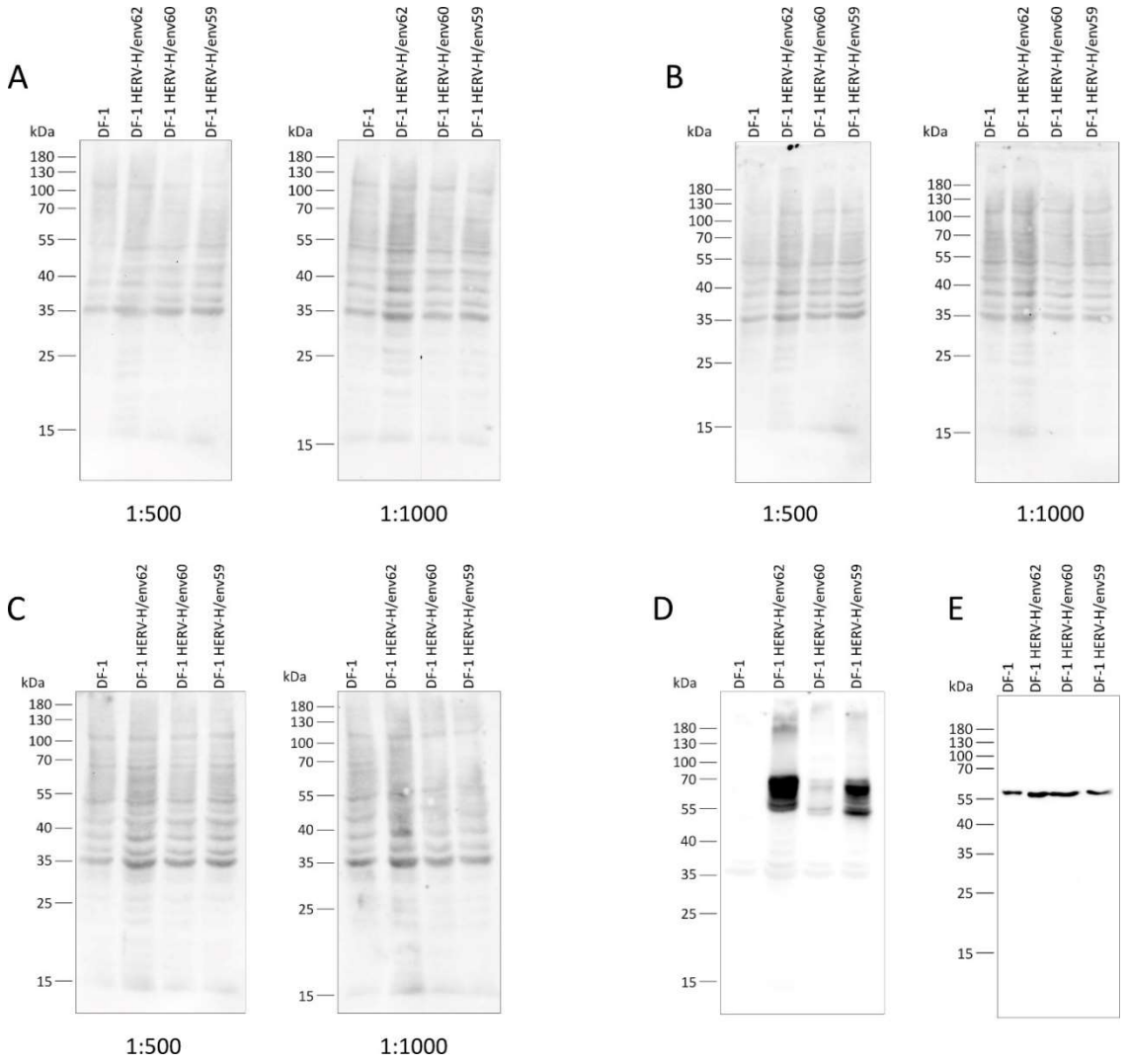


Figure 26: Western blot of purified mAbs produced by Davids Biotechnologie.

The purified mAbs were tested for their HERV-H env binding properties using DF-1 cells transfected with the three different HERV-H env proteins in a western blot, using a secondary anti-mouse IgM/HRP antibody. Western blots with the mAbs purified from clone 41.8.23 (A), 41.18.6 (B) and 41.18.85 (C) as primary antibody. All mAbs were tested in two different dilutions indicated below the respective blots. D: A protein-A-purified polyclonal anti-HERV-H env antibody was used as control. This antibody was detected using an anti-rabbit IgG/HRP secondary antibody. E: For the tubulin control, a primary anti-tubulin antibody was detected via a secondary anti-mouse Ig/HRP antibody. The PageRuler Prestained Protein ladder was used in all blots as a marker for the molecular weight.

In summary, it was not possible to generate mAbs targeting the HERV-H env protein. Even though, three mAbs were purified by a contractor, none of them were able to recognize the HERV-H env protein in a western blot assay. Furthermore, the results suggest, that

immunization with HERV-H env favorizes the production of IgM-class mAbs. Since antibodies are a vital reagent for both, diagnostic and therapeutic applications, a new approach was chosen by generating polyclonal antibodies (pAbs).

6.4 Generation of Polyclonal Antibodies

In contrast to mAbs which are derived from a single B cell clone and therefore recognize one single epitope of the antigen, pAbs are produced from multiple B cells resulting in multi-epitope binding properties. As reviewed by Ascoli and Aggeler, this can be an advantage for certain applications since polyclonality can enhance sensitivity by recognizing several epitopes of the target antigen. Especially in the diagnostic context, the multi-epitope binding properties of pAbs can have a great advantage since patients may show a diversity of naturally occurring antigen variations (Ascoli and Aggeler, 2018).

A suite of pAbs targeting HERV-H was generated by the company Davids Biotechnologie (Regensburg/Germany) using the soluble HERV-H env protein (described in 6.2.1a) as an immunogen. Two rabbits were immunized with the protein according to the contractor`s protocol. Half of the serum from one rabbit was used to purify the pAbs via protein A-purification resulting in the isolation of the total IgG fraction. The other half of the serum was affinity-purified using the solHERV-H env, catching only those pAbs binding to this protein. The serum from the second rabbit was left untreated. For the purifications, resin-filled columns with either protein A or the solHERV-H env protein covalently bound to the resin, were used and the purifications were performed following common protocols. In summary, after the serum had been loaded to the column, the bound antibodies were eluted under acetic conditions, typically at a pH of 2-3. Eluted antibodies were then directly neutralized with an alkaline buffer (pH 8.5-9) and the purified pAbs were buffer exchanged to a 0.1 M sodium-acetate buffer for protein-A-purified preparations, and a buffer containing 0.1 M phosphate and 10 mM sodium-acetate for the HERV-H-affinity purified pAbs.

The two pAb preparations and the serum were characterized for their binding abilities in ELISA, western blot and flow cytometry in the present work.

A titration ELISA was conducted using different recombinant proteins to test the pAb preparations for cross-reactivity and HERV-H env-specific binding. The pAbs were tested

against the solHERV-H protein, the solStabHA30_His6 and the two HIV-1 BG505 env proteins BG505_His6 and BG505_noTag. The protein-A-purified pAb preparation exhibited high-affinity binding to solHERV-H with an EC50 value of 0.035 µg/ml. While no binding could be detected for the non-tagged BG505, the pAb bound to both His6-tagged non-HERV proteins when applied at a concentration of approximately 5 µg/ml or higher (Figure 27A; Table 28). This effect was even more visible for the HERV-H affinity-purified pAb. Here, a signal for the His6-tagged non-HERV-H proteins was already detected at a pAb concentration between 0.3 and 1 µg/ml. This antibody had an EC50 value of 0.1 µg/ml for solHERV-H, 1.78 µg/ml for solStabHA30_His6 and 1.7 µg/ml for BG505_His6. Like the protein-A-purified pAb, no binding to the BG505_noTag protein was observed (Figure 27B).Therefore, it can be concluded that the anti-HERV-H env pAb preparations also included antibodies targeting the His6-tag. The untreated rabbit serum showed similar binding to the His6-tagged non-HERV-H proteins as the protein-A-purified pAb, with signals detected only for the highest pAb concentration. For solHERV-H, the serum bound to the protein with an EC50 value of a serum dilution factor of $1.69x10^{-6}$. Again, no binding to BG505_noTag was seen (Figure 27C).

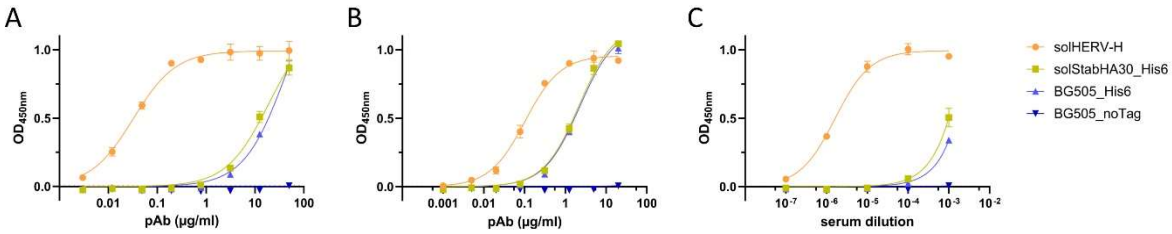


Figure 27: Specificity ELISA of anti-HERV-H env pAbs.

The pAb preparations were tested for binding towards the proteins solHERV-H (orange), solStabHA30_His6 (green), BG505_His6 (light blue) and BG505_noTag (dark blue) in a titration ELISA. A: The protein-A-purified pAb was titrated in a 4-fold serial dilution starting at a concentration of 50 µg/ml. B: The HERV-H affinity-purified pAb was titrated in a 4-fold serial dilution starting at 20 µg/ml. C: The rabbit serum was titrated in a 2-fold serial dilution starting at a 1:1000 dilution. For all pAb preparations, an HRP-coupled anti-rabbit Ig was used as secondary antibody.

Table 28: EC50 values of the pAbs for the different soluble proteins.

Antibody	solHERV-H	solStabHA30_His6	BG505_His6	BG505_noTag
Protein A/(µg/ml)	0.035	11.8	21.0	n. a.
Affinity/(µg/ml)	0.1	1.78	1.71	n. a.
Serum/dilution factor	$1.69x10^{-6}$	$1.57x10^{-4}$	$1.31x10^{-4}$	n. a.

To reduce the reactivity against the His6-tag, a depletion of the anti-His6 tag antibodies was performed on the HERV-H affinity-purified pAb by Davids Biotechnologie (Regensburg/Germany). Here, a negative selection was used, binding anti-His6 tag antibodies to the matrix of a column and collecting the flow through containing the anti-HERV-H antibody fraction. After the depletion, the pAb exhibited a His6-tag reactivity only at the highest pAb concentration tested (Figure 28B).

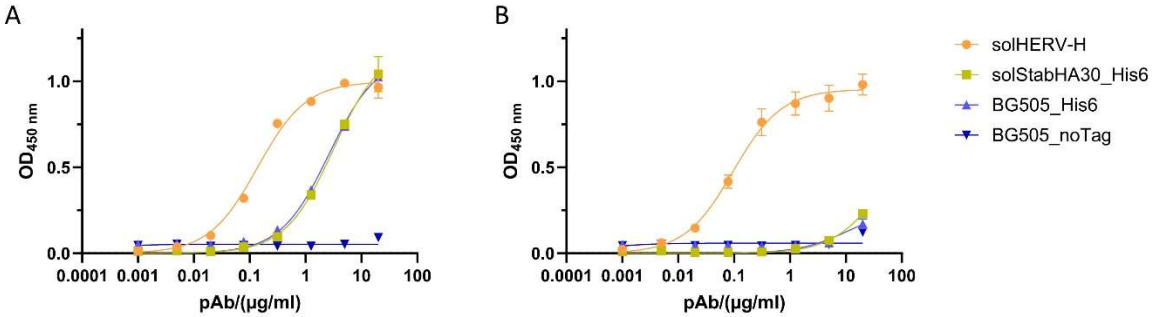


Figure 28: Specificity ELISA of the HERV-H affinity-purified pAb before and after anti-His6 antibody depletion.

The HERV-H affinity-purified pAb was depleted of anti-His6-tag antibodies and tested for reactivity towards the proteins solHERV-H (orange), solStabHA30_His6 (green), BG505_His6 (light blue) and BG505_noTag (dark blue) in a titration ELISA. A: Affinity-purified pAb before anti-His6 tag antibody depletion. B: Anti-His6 tag antibody depleted affinity-purified pAb.

Next, the pAb preparations were tested for application in a western blot assay. In a first assay, the optimal pAb concentration for the use in western blot was determined and simultaneously, the binding properties against the three HERV-H ORFs HERV-H/env62, HERV-H/env60 and HERV-H/env59 were assessed. Therefore, BHK-21 cells, which should be negative for HERV-H due to their non-human origin, were transiently transfected with the three HERV-H plasmids and stained with the pAb preparations in different antibody concentrations. The protein-A-purified pAb was used at concentrations of 200 µg/ml, 50 µg/ml and 1 µg/ml. For the HERV-H affinity-purified pAb, concentrations of 20 µg/ml, 5 µg/ml and 0.1 µg/ml were chosen, and the serum was applied at the dilutions 1:100, 1:1000 and 1:100000. The different pAb preparations all exhibited binding to the three HERV-H env proteins with strong bands for HERV-H/env62 and HERV-H/env59 and weaker bands for HERV-H/env60 (Figure S 15). Binding was seen for all antibody concentrations but with very faint bands for the lowest concentration of the HERV-H affinity-purified pAb and the serum, and high amounts of unspecific background bands for the highest concentrations of all pAbs. For further assays, the

protein-A-purified pAb was used at a concentration of 1 µg/ml, the HERV-H affinity-purified pAb at 5 µg/ml and the serum was applied at a 1:1000 dilution.

The pAb preparations were also tested in the western blot assay for cross-reactivity to other membrane-bound viral proteins. Here, a panel of in the lab available plasmids encoding viral surface proteins from different virus strains were chosen. These included the glycoproteins of the hemorrhagic fever viruses Ebolavirus (EBOV-GP), Marburgvirus (MARV-GP) and Lassa virus (LASV-GP), provided by Dr. Martina Pfranger from the AG Wagner, and the Spike protein of the SARS-CoV-2 Delta virus variant (SARS-CoV-2_S), provided by Sebastian Einhauser (AG Wagner). The plasmids were transiently transfected into BHK-21 cells and stained with the different pAb preparations. As a positive control, cells transfected with HERV-H/env62 were included. All pAbs showed specific binding to HERV-H transfected BHK-21 cells with a strong band at 70 kDa in all western blots and no bands at the expected sizes for the other proteins. Nevertheless, some additional unspecific background bands could be seen in all blots. Especially a band at 35 kDa was detected in all samples, including the untransfected BHK-21 cells (Figure 29A-C). The blot stained with the HERV-H affinity-purified pAb showed a higher variety of background bands, with bands being stronger and more distinct compared to the other two pAb preparations (Figure 29B). The tubulin control showed equal amounts of cell lysates for all variants tested (Figure 29D). The presence of the non-HERV-H proteins was validated by staining the blots with suitable antibodies recognizing these proteins specifically (data not shown).

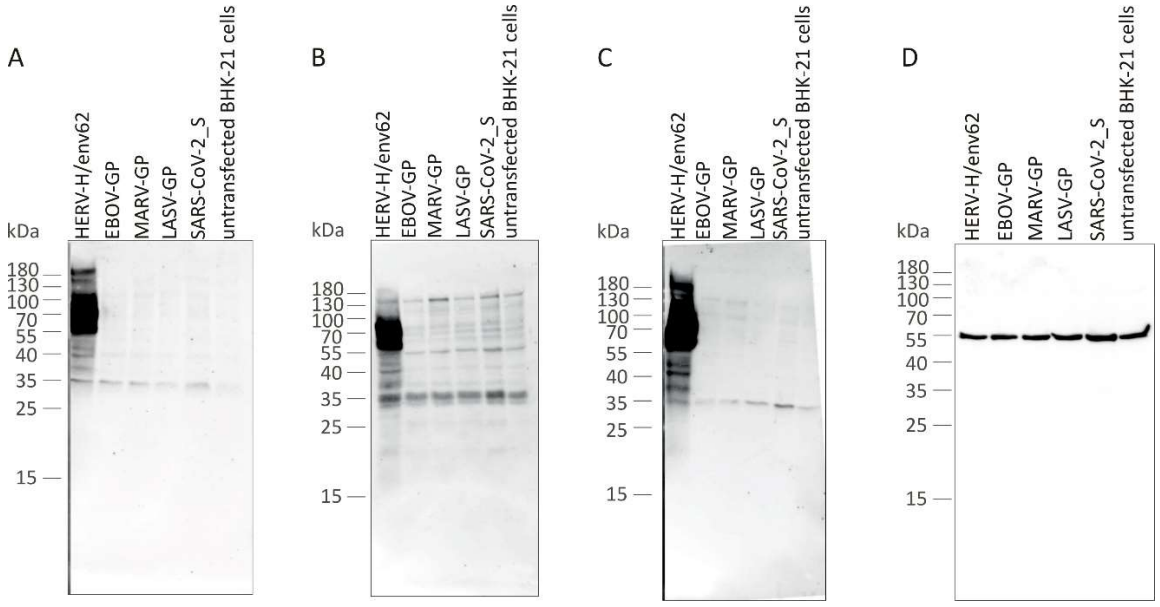


Figure 29: Specificity screening of anti-HERV-H env pAbs in a western blot assay.

BHK-21 cells were transiently transfected with plasmids encoding the HERV-H/env62 protein, the glycoproteins of the Ebolavirus (EBOV-GP), Marburgvirus (MARV-GP) and Lassa virus (LASV-GP), and the SARS-CoV-2 Spike protein (SARS-CoV-2_S). Untransfected BHK-21 cells were used as a negative control. A: Cells were stained with the 1 µg/ml protein-A-purified pAb. B: The staining was performed using the HERV-H affinity-purified pAb at a concentration of 5 µg/ml. C: Cells were stained with the rabbit serum diluted 1:1000. For the pAbs, an HRP-coupled anti-rabbit Ig secondary antibody was used. D: Anti-tubulin was used as control, stained with a secondary anti-mouse Ig/HRP antibody. The molecular weight was determined via the PageRuler Prestained Protein ladder.

Concluding, for the two pAb preparations and the rabbit serum good conditions for a specific HERV-H env staining with low background signals could be determined. Especially the protein-A-purified pAb preparation showed to be suitable for HERV-H env detection, due to its low background staining. Furthermore, a concentration of 1 µg/ml was sufficient for a strong env detection, making this pAb the preferred antibody for further analysis.

As a third application, the pAb preparations were tested in a flow cytometry assay. For this, HERV-H env transfected and untransfected BHK-21 cells were stained in a cell surface staining protocol with the pAbs. The pAbs were titrated in a 4-fold serial dilution for the protein-A-purified pAb, starting at 500 µg/ml, and the HERV-H affinity-purified pAb starting at 50 µg/ml, and a 5-fold serial dilution for the serum starting at a dilution of 1:20. All three pAb preparations exhibited high background signals in both, the percentage of HERV-H env positive cells and the MFI. Furthermore, the staining efficiency did not reach more than 60 % of positive cells (Figure 30A-C). As controls, the cells (transfected and untransfected) were either stained only with the secondary antibody (sec. only) or were left unstained. Here, the percentage of positive cells stayed below 2 %, and the MFI did not reach values above 40

(Figure 30D). This indicates that the high background seen for the pAb preparations is not caused by the secondary antibody.

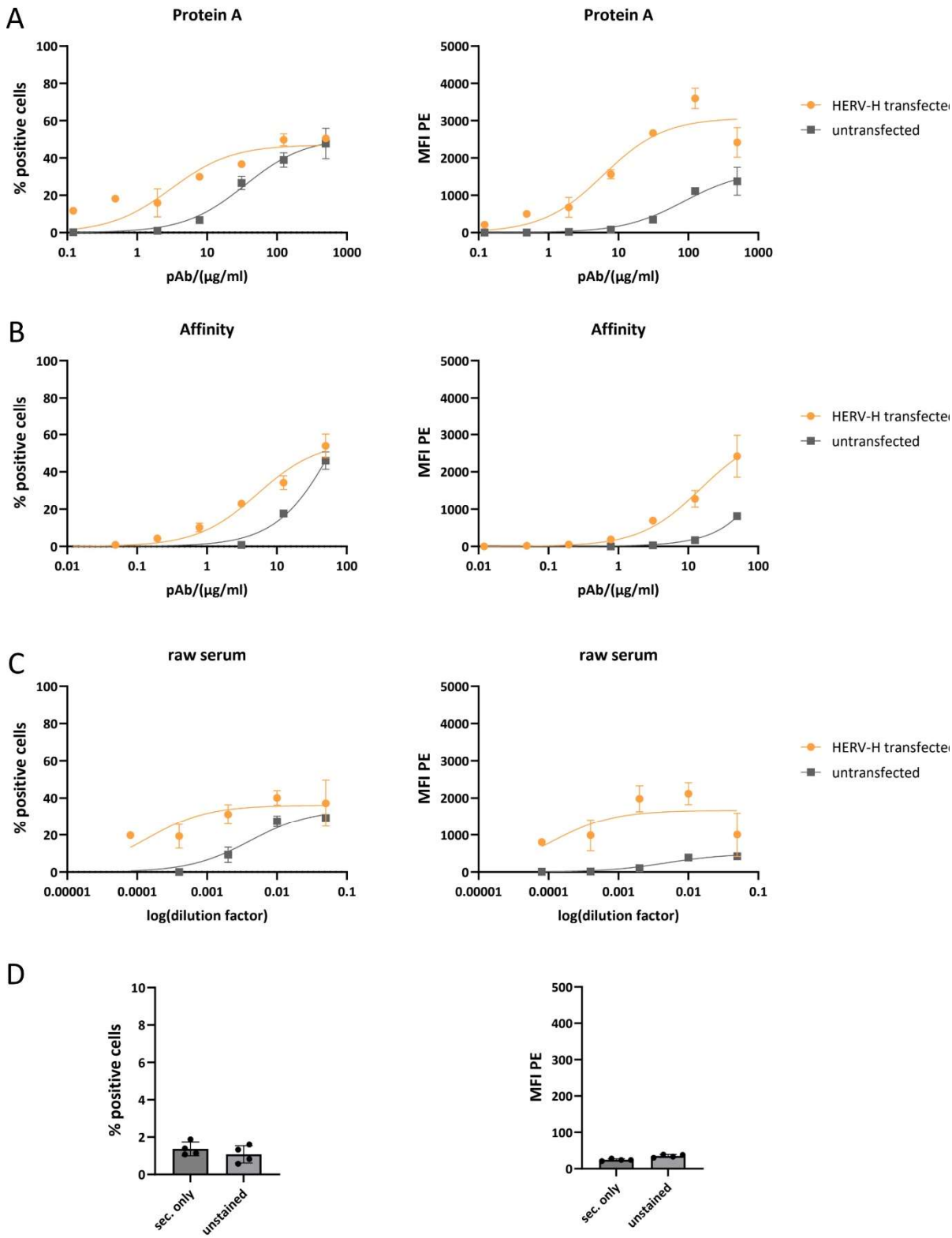


Figure 30: Flow cytometry assay with polyclonal anti-HERV-H env antibodies.

HERV-H env transfected BHK-21 cells were stained for env expression with the three different pAb preparations. As negative control untransfected BHK-21 cells were used. Antibody-binding was detected via a PE-labeled anti-rabbit IgG

secondary antibody. A: The percentage of HERV-H env positive cells (left) and the MFI of the PE-labeled secondary antibody (right) of transfected and untransfected cells stained with the protein-A-purified pAb. B: Cells were stained with the affinity-purified pAb. C: Raw serum of immunized mice was used to stain BHK-21 cells. D: As negative controls, the cells were stained only with the secondary antibody (sec. only) or were left unstained. For a better perception, the Y-axis of the negative controls was set to 1/10 of the axis from the measure samples. All samples were measured in technical duplicates, the controls were measured in 4 replicates and the symbols indicate the mean value at each pAb concentration including the standard error of the mean.

In conclusion, a panel of different anti-HERV-H env pAb preparations was produced and characterized for HERV-H env binding in different assays. Even though high background staining was seen in flow cytometry, all pAb preparations exhibited specific binding to the HERV-H env protein in ELISA and western blot. Therefore, it was possible to produce a suitable and specific reagent to detect HERV-H env expression. Since the protein-A-purified pAb performed the best, this preparation was used in all further assays.

6.5 Characterizing HERV-H Env Expression in Human Tumor Cell Lines

In order to screen tumor tissue for HERV-H env expression, a diagnostic platform needs to be developed. Therefore, different diagnostic tools for detecting HERV-H env on RNA and protein levels were tested and established using HERV-H env-positive human tumor cell lines.

6.5.1 Characterization on RNA and genomic DNA Level

Yi *et al.* conducted a study in 2006 screening a variety of human tissues and tumor cell lines for the transcription of HERV-H *env* on RNA level by RT-PCR. Therefore, the authors designed primers binding in the 3' part of the *env* sequence (Yi, Kim and Kim, 2005). In this thesis, this primer design was used for a first screening of human tumor cell lines comprising some of the cell lines tested by Yi e*t al.*, as well as new cell lines that have not been included in this study. These encompassed the human cell lines MDA-MB-231, HEK293, THP-1, A-673, as well as the primate cell line VeroE6 and the non-primate cell lines DF-1 (chicken) and BHK-21 (hamster). The cell lines were chosen in order to get the most comprehensive panel of human, primate and non-primate (tumor) cell lines possible.

RNA was extracted from all cell lines and a one-step RT-PCR was conducted. The PCR fragments were analyzed on an agarose gel and GAPDH was used as a reference gene. The agarose gel showed an amplification product for the HERV-H PCR between 500 and 600 bp, matching the expected size of 596 bp, for all human cell lines as well as the green monkey cell

line VeroE6, while no band was seen for the non-primate cell lines BHK-21 and DF-1 (Figure 31).

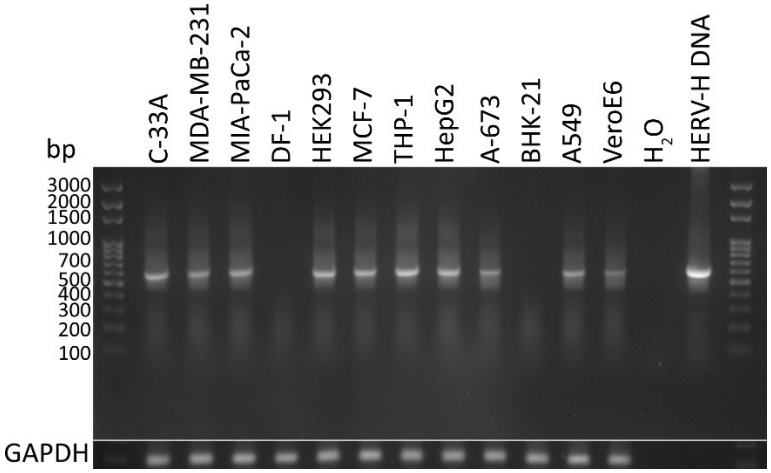


Figure 31: HERV-H env RT-PCR of human and non-human cell lines.

An RT-PCR using HERV-H primers previously published by Yi *et al.* was performed with 500 ng RNA extracted from various human, one primate and two non-primate cell lines. The PCR fragments were separated by agarose gel electrophoresis. For size determination, a 100 bp DNA ladder was used. GAPDH was used as a reference gene.

For a quantitative analysis of the HERV-H *env* expression, an RT-qPCR was conducted using the three tumor cell lines C-33A, MCF-7 and MIA-PaCa-2. These cell lines were of particular interest since - according to Yi *et al.* - they contain a complete approximately 600 bp *env* 3' part without any frameshift mutations or premature stop codons. The RT-qPCR was performed using the same primer pair as before, and GAPDH as reference gene. DF-1 and H_2O were used as negative control, while a plasmid encoding the HERV-H/env62 gene was used as positive control. All three cell lines showed comparable *env* amplification, with C_T values around 18, while for the negative controls a $C_T > 30$ was observed (Figure 32B), thereby clearly separating HERV-H env positive and negative cell lines.

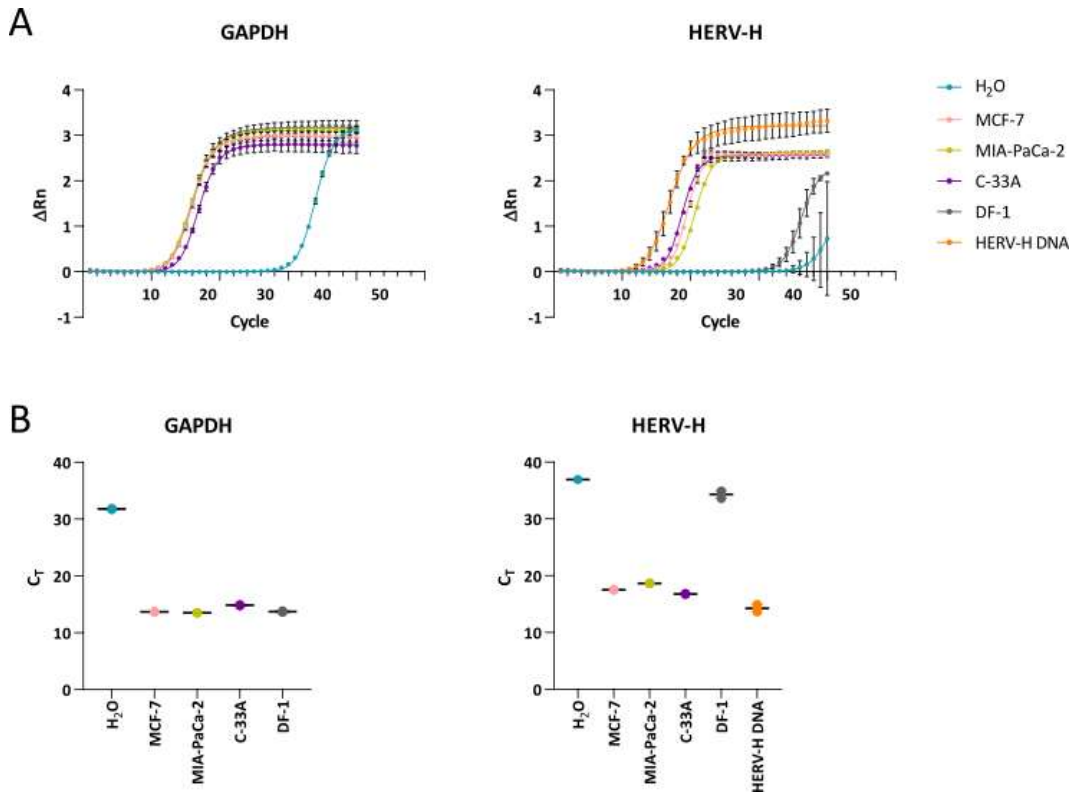


Figure 32: HERV-H env RT-qPCR with RNA isolated from three different human tumor cell lines.

An RT-qPCR with 500 ng RNA extracted from human and non-human tumor cell lines was conducted using a primer pair previously published by Yi *et al.* The amplification signal was detected with a StepOnePlus Real-Time PCR System using SYBR green. GAPDH was used as reference gene and HERV-H env DNA as positive control in the HERV-H PCR. A: The ΔRn values are plotted against the number of PCR cycles for each cell line. B: C_T values are depicted for each cell line. The graphs on the left side show the ΔRn and C_T values for the GAPDH PCR, while the graphs on the right depict the respective values for the HERV-H PCR. All samples were measured in technical triplicates. The amplification data were analyzed using the StepOnePlus Real-Time PCR software.

This first assessment showed that all primate cell lines tested transcribe a partial HERV-H *env* mRNA and therefore have the potential to contain a complete *env* ORF, confirming the previous data by Yi *et al*.

To test the performance of the RT-qPCR assay with samples that simulate conditions as they would occur in a diagnostic context, the three cell lines MCF-7, C-33A and MIA-PaCa-2 as well as the non-human cell line DF-1 were embedded in paraffin, and sections were prepared on microscope slides. These formalin-fixed paraffin-embedded (FFPE) samples were later also used for immunocytochemistry assays (see section 6.5.2). For the RT-qPCR assay, the RNA was extracted and pooled from two slides of each FFPE sample using the Pinpoint Slide RNA Isolation System II, and the RT-qPCR was performed using the same primer pair as described earlier. As reference, RNA extracted from fresh cells was included. Compared to the references, the FFPE samples showed only weak HERV-H *env* expression with ΔRn values not

reaching a plateau and C_T values above 20 for both, the HERV-H and GAPDH PCR (Figure 33A + B). Nevertheless, all FFPE samples except those obtained from MCF-7 cells showed lower C_T values than the HERV-H env negative cell line DF-1 ($C_T \approx 30$), although the difference was rather small (ΔC_T of 25.2 and 24.6 for C-33A and MIA-PaCa-2). The PCR products were additionally analyzed on an agarose gel. Here, only weak bands were observed for the C-33A and MIA-PaCa-2 FFPE samples. The MCF-7 FFPE sample exhibited no band for the HERV-H PCR product, but a weak band for GAPDH (Figure 33C).

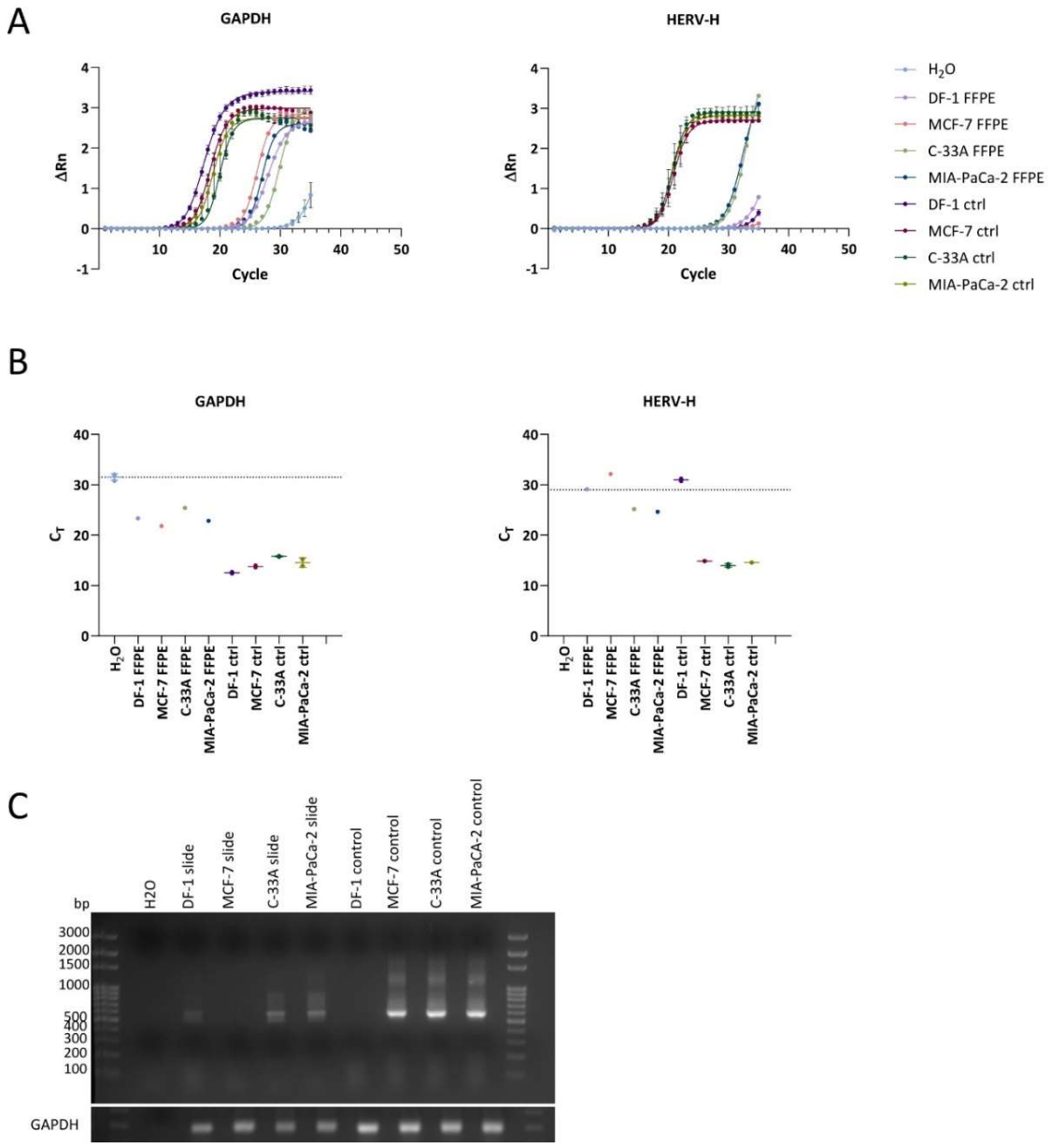


Figure 33: RT-qPCR with RNA extracted from FFPE cell samples.

RNA was extracted from FFPE cells, and the RT-qPCR was performed using HERV-H specific primers published by Yi *et al*. As reference gene, GPADH was used, and RNA extracted from fresh cells was used as internal assay control. A: ΔRn values are plotted against the cycle number of the PCR. B: The C_T values are shown for all samples. The horizontal dotted line indicates the lowest C_T of the negative controls. In A + B the graphs on the left hand side show the values

from the GAPDH PCR, while the graphs on the right hand side depict the results from the HERV-H PCR. C: The RT-qPCR products were separated on a 1 % agarose gel. As DNA standard, a 100 bp DNA ladder was used. The upper section shows the bands of the HERV-H PCR, the lower section bands from the GAPDH PCR. FFPE samples were measured in monoplicates, while the reference samples were measured in technical duplicates.

Since no HERV-H *env* transcription could be observed in the MCF-7 FFPE cells, the RT-qPCR was repeated with this sample. This time, the MCF-7 FFPE sample exhibited a lower C_T value than the DF-1 FFPE control (C_T of 22 for MCF-7 FFPE and $C_T \approx 30$ for DF-1 FFPE) (Figure 34).

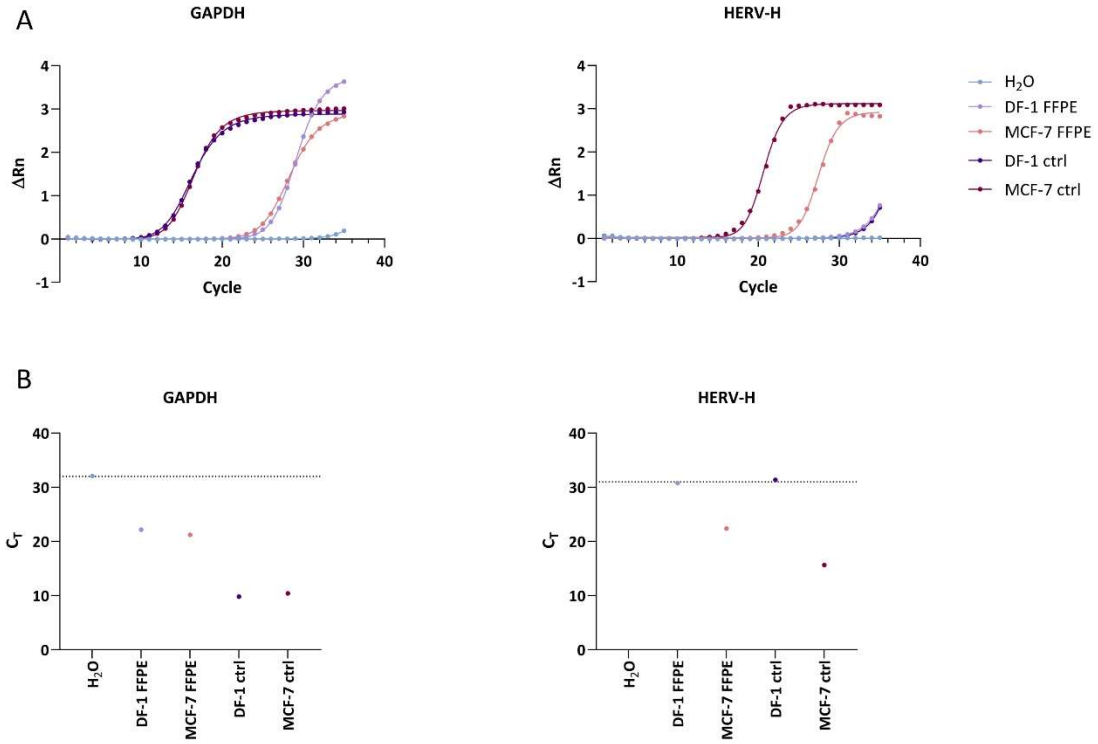


Figure 34: Repetition of the RT-qPCR with MCF-7 cells.

The RT-qPCR was performed as described in Figure 33. A: ΔRn values are plotted against the cycle number of the PCR. B: The C_T values are shown for all samples. The horizontal dotted line indicates the lowest C_T of the negative controls.

In conclusion, the results show that a RT-qPCR with RNA extracted from FFPE samples is possible, but the results should be interpreted with caution since a separation between positive and negative samples was not as clear as with RNA preparations from fresh cells. Also, the low cell numbers on the FFPE slides probably had an influence on the RT-qPCR results (Table S 2).

To further analyze the HERV-H *env* sequence encoded in these tumor cell lines and to assess, if these cells have the genomic prerequisite to express a full-length env protein, genomic DNA

(gDNA) was isolated from the three cell lines MCF-7, C-33A and MIA-PaCa-2. The complete *env* sequence was then amplified by PCR using primers binding up- and downstream of the *env* reading frame. In a first attempt, the purified PCR products were directly sent to Sanger sequencing. However, the obtained sequencing results were ambiguous with many double peaks in the sequencing histogram. Overall, the results of all three cell lines nevertheless showed a complete *env* sequence, resembling the HERV-H/env62 variant. Although several mutations leading to amino acid substitutions could be observed, most of them were conservative substitutions. However, one particular site at around 1.15 kb from the 1.9 kb total sequence was clearly different from the original HERV-H/env62 sequence in all three cell lines. Here, the sequence results were again highly ambiguous with one peak showing a thymine at that position and one indicating a cytosine (Figure 35A-C). While the cytosine would maintain the original sequence, the thymine would introduce a premature stop codon at this site, thereby ending the protein translation right at the beginning of the cleavage site at the C-terminal end of the SU.

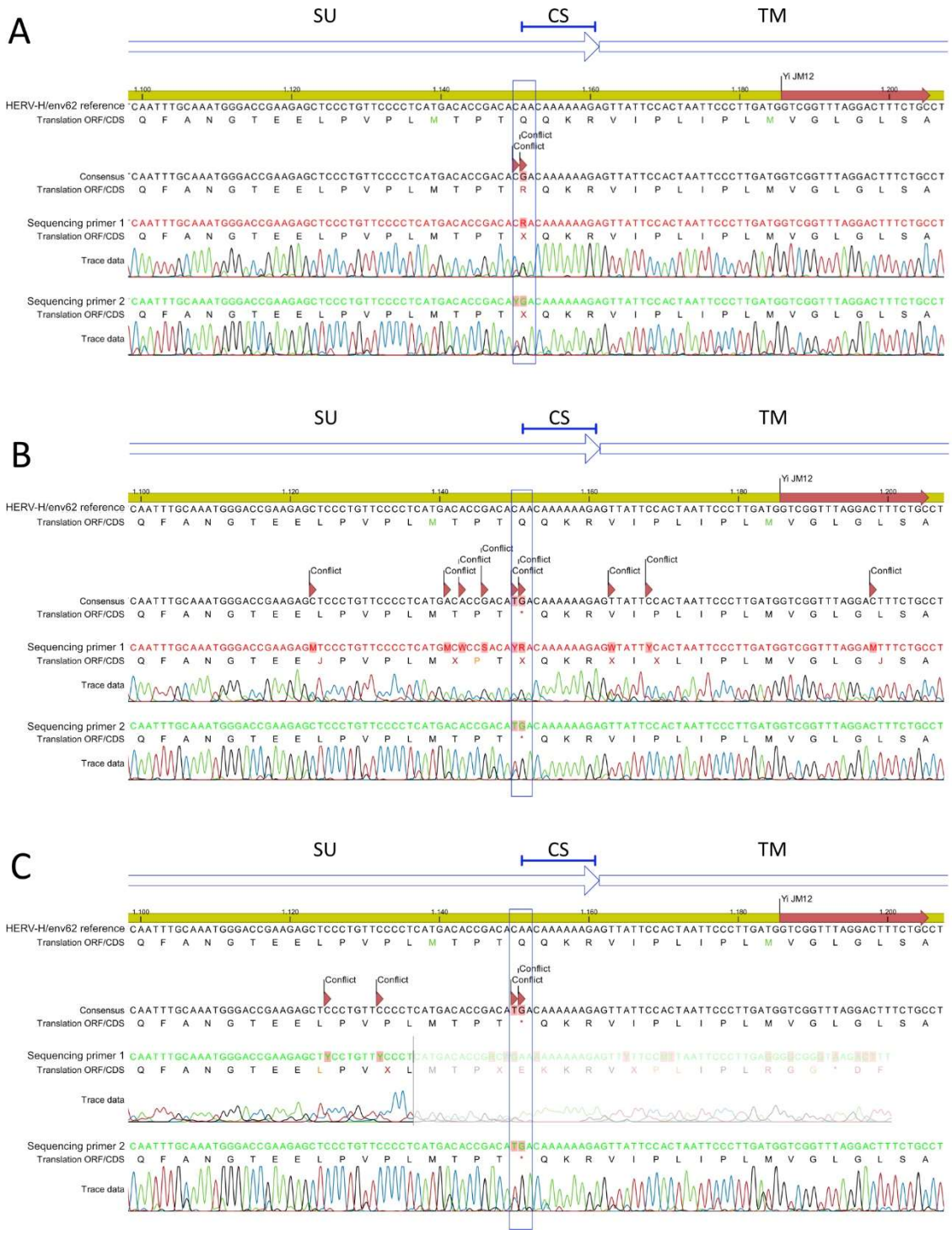


Figure 35: Sequencing results of the HERV-H env DNA locus of the three cell lines C-33A, MCF-7 and MIA-PaCa-2 from env nucleotide position 1099 to 1205.

Genomic DNA was isolated and the *env* sequence was amplified via PCR. The purified PCR product was sequenced via Sanger sequencing. The sequencing results were assembled to the HERV-H/env62 reference sequence. Depicted are the sequencing results from nucleotide position 1099-1205. Sequencing results from gDNA extracted from A: the cervical cancer cell line C-33A, B: the breast cancer cell line MCF-7, C: the pancreatic cancer cell line MIA-PaCa-2. On top of the reference sequence, arrows indicate the corresponding protein domain. SU: surface subunit; CS: cleavage site; TM: transmembrane subunit. The blue rectangles highlight the nucleotide substitution encoding for a potential premature stop codon. The sequence results were analyzed using the CLC Main Workbench 22. These data were generated by Lisa Messerer during her bachelor's thesis under my experimental supervision.

In order to unravel the ambiguity of this particular site and also all the other ambiguous sequencing results, the PCR fragments were blunt-end cloned into the pJET1.2 sequencing vector before sequencing. With that method, 41 clones were sequenced from all cell lines with 14 sequences obtained from C-33A cells, 19 sequences from MCF-7 cells, and 8 sequences from MIA-PaCa-2 cells. The sequences obtained in this attempt were clearer than from the directly sequenced PCR fragments. Overall, all but one sequence showed mutations altering the *env* sequence in a matter that would either truncate the env protein or turn it non-functional. The observed mutations included frameshifts introduced by insertions or deletions, as well as premature stop codons. More than 50 % of all sequences from C-33A and MCF-7 cells had premature stop codons, whereas for MIA-PaCa-2 cells, this genotype was observed in 75 % of all sequences. In C-33A cells more deletion (29 %) than insertion (14 %) mutations were observed. In MCF-7 cells, this effect was reverse with 11 % of deletions and 32 % insertions, and MIA-PaCa-2 cells had an equal proportion of deletion and insertion mutations with 13 % each. Only one sequence obtained from MCF-7 cells showed a complete *env* sequence (Table 29). The premature stop codon seen in the sequencing results of the purified PCR products (Figure 35) was observed in a total of 13 % of all cell line sequences after subcloning to the pJET1.2 vector (Table 30).

Table 29: Overview of sequencing results from gDNA of tumor cell lines.

Cell line	Seq. with pre stop	% pre stop	Seq. with del	% del	Seq. with in	% in	Complete seq.	% complete
C-33A	8/14	57	4/14	29	2/14	14	-	-
MCF-7	10/19	53	2/19	11	6/19	32	1	5
MIA-PaCa-2	6/8	75	1/8	13	1/8	13	-	-
total	24/41	59	7/41	17	9/41	22	1	2

Pre stop: premature stop codon; del: deletion; in: insertion. Data shown in this table were generated by Maximilian Fuß during his bachelor's thesis under my experimental supervision.

Table 30: Number of sequences with a premature stop codon at nucleotide position 1149-1150.

Cell line	Seq. with pre stop at position 1149-1150	% of pre stop at position 1149-1150 (relative to total number of pre stop)
C-33A	1	14
MCF-7	0	0
MIA-PaCa-2	2	33
total	3	13

Data shown in this table were generated by Maximilian Fuß during his bachelor's thesis under my experimental supervision.

The complete sequence found in MCF-7 cells showed an overall high sequence identity to the HERV-H/env62 reference sequence with only minor substitutions in the N-terminal part of the protein sequence.

All in all, the sequencing results show that the majority of HERV-H *env* sequences encoded in the cell lines are truncated and would therefore not get translated into a full-length env protein. However, the presence of the one full-length sequence found in MCF-7 cells indicates, that this cell line possesses the genomic prerequisite for the expression of an env protein. Moreover, the high ratio of incomplete-to-complete sequences suggest, that in order to detect full-length HERV-H *env* sequences, a large number of clones need to be analyzed.

6.5.2 Characterization on Protein Level

To use the HERV-H env protein as a target for cancer immunotherapy, a detection on mRNA level would not be sufficient to ensure protein expression in the tumor tissue. Therefore, an additional detection on protein level would be necessary.

To show HERV-H env expression on protein level in the cell lines, lysates of all cells tested in the RT-PCR assay were generated and analyzed in a western blot using the protein-A-purified anti-HERV-H pAb. As positive control, cell lysates of HERV-H transfected DF-1 cells were used, and anti-tubulin staining was performed to confirm equal cell lysate concentrations. Except for THP-1 cells, the anti-tubulin blot showed equal band strength for all cell lines, indicating that the lysate of THP-1 cells was not applied at the same concentration. All human cell lines and the green monkey cell line VeroE6 exhibited a band at approximately 70 kDa, which was not detectable for the non-primate cell lines BHK-21 and DF-1 (Figure 36A). Since the THP-1 cells were applied at a lower concentration than the other cells, only a weak HERV-H band was detected for this cell line. The protein band of the positive control (DF-1 HERV-H) was detected at a higher molecular weight than the bands from the cell lines. This could be caused by the induced protein overexpression or different post-translational modifications. An analysis of the HERV-H env band intensities normalized to the tubulin signal showed the highest signal for HEK293 cells with a relative intensity of around 0.7. The other human cell lines had a range of intensity values from approximately 0.05-0.3, while the VeroE6 cells showed the lowest detectable intensity value of 0.02 (Figure 36B). Since the positive control exhibited a very

strong band due to the artificial HERV-H env overexpression, it was not included in the analyses of the signal intensities.

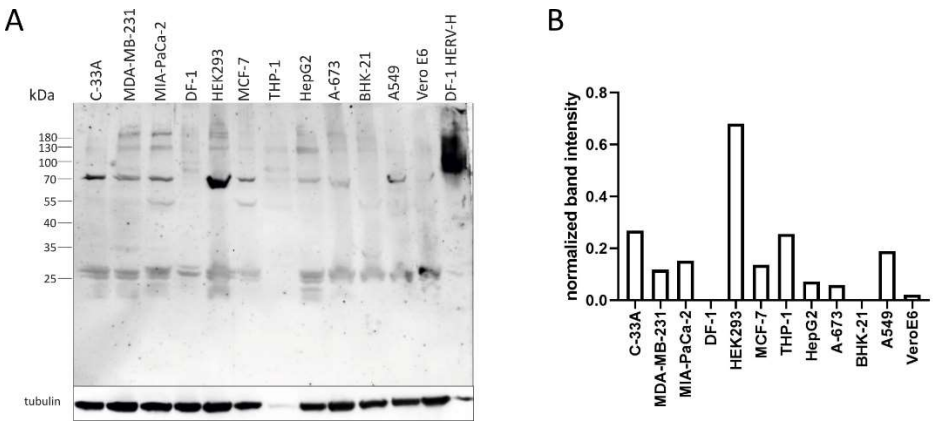
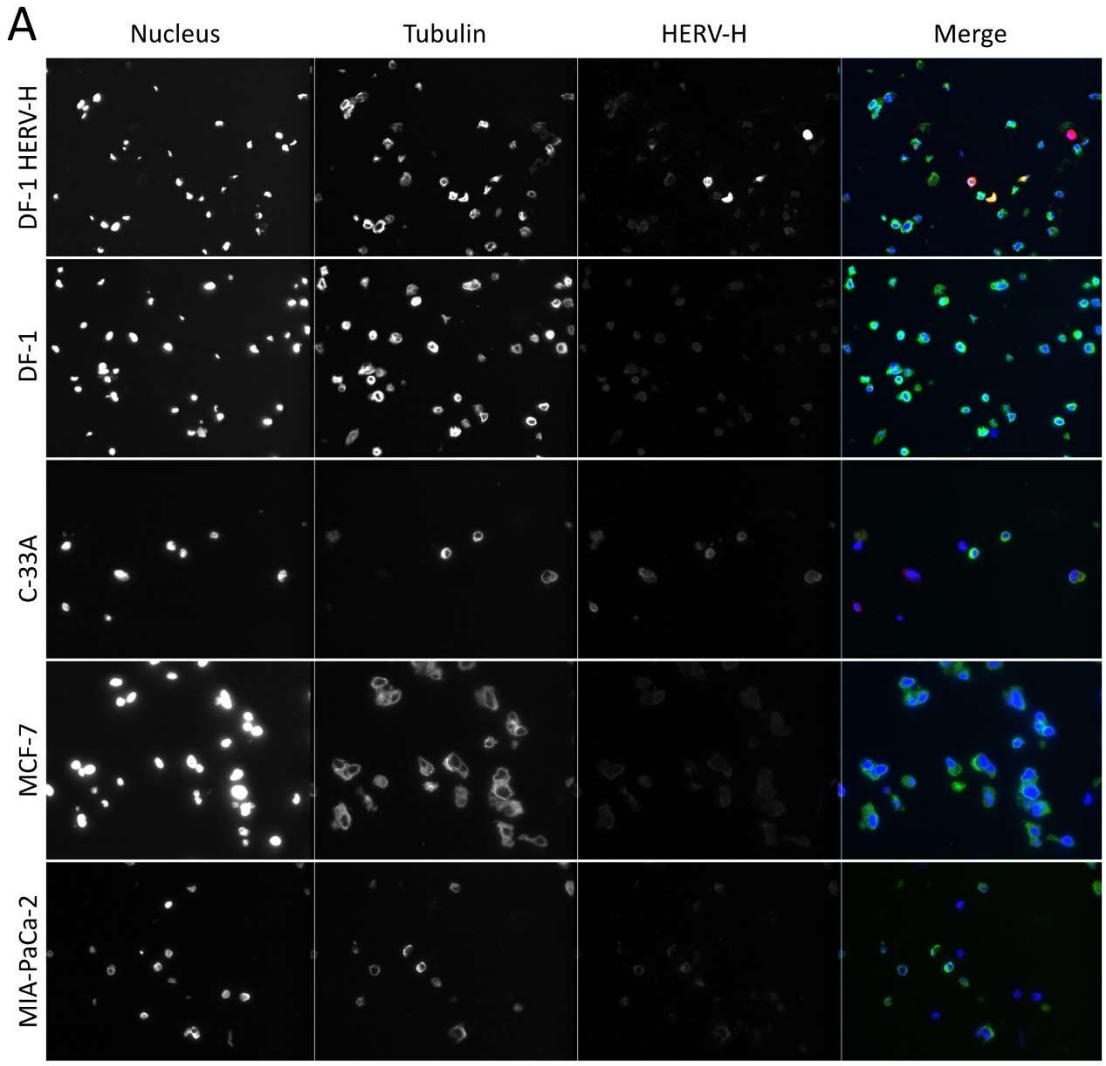


Figure 36: Anti-HERV-H env western blot with cell lysates of human and non-human cell lines.

Cell lysates of various human, one primate and two non-primate cell lines were analyzed on a western blot for HERV-H env protein expression. As primary antibody a protein-A-purified polyclonal anti-HERV-H antibody was used, which was captured by an anti-rabbit IgG/HRP secondary antibody. For the loading control, an anti-tubulin, and a secondary anti-mouse IgG/HRP antibody were used. The PageRuler Prestained Protein ladder was used as a molecular weight standard. A: Western blot of all cell lysates. B: HERV-H env band intensity values of all cell lines normalized to the respective anti-tubulin signal. The band intensity measurement was performed using the GelAnalyzer 23.1.1 software.

These results extend the RNA data shown earlier in this work and by Yi *et al.* (Yi, Kim and Kim, 2005), proving the expression of HERV-H env in the tested cell lines on protein level.

To further develop a diagnostic test suitable for routine screening of tumor tissues, an anti-HERV-H immunocytochemistry (ICC) assay was established by staining paraffin-embedded HERV-H transfected and untransfected DF-1 cells. For this, the cells were stained with the protein-A-purified anti-HERV-H pAb and an AlexaFluor647-coupled secondary antibody. Additionally, the cells were stained with an AlexaFluor488-coupled anti-tubulin antibody to visualize the cytoplasm, and a Hoechst33342 stain to detect the cell nucleus. The ICC assay was then performed on the three previously described HERV-H env positive cell lines MCF-7, C-33A and MIA-PaCa-2. All three cell lines showed a weak signal in the HERV-H staining, which was significantly higher compared to the anti-tubulin normalized signal intensity of untransfected HERV-H negative DF-1 cells (Figure 37).


A
Nucleus
Tubulin
HERV-H
Merge
DF-1 HERV-H
DF-1
C-33A
MCF-7
MIA-PaCa-2

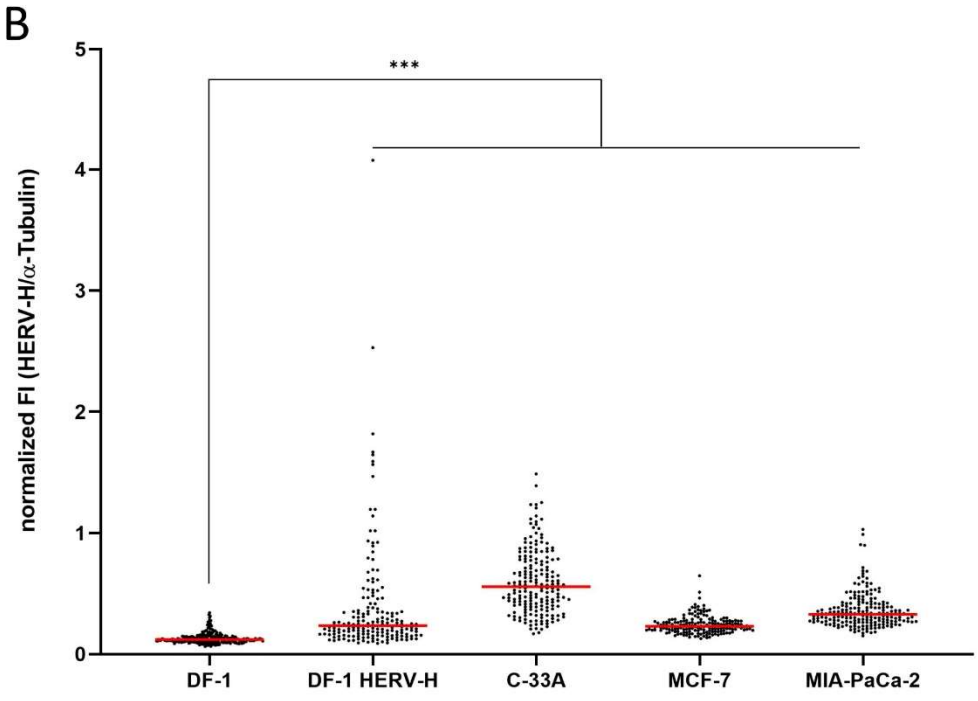

B
***
normalized FI (HERV-H/α-Tubulin)
0
1
2
3
4
5
DF-1
DF-1 HERV-H
C-33A
MCF-7
MIA-PaCa-2

Figure 37: ICC staining of FFPE tumor cell lines.

Sections of paraffin-embedded human (MCF-7, C-33A, MIA-PaCa-2) and non-human (DF-1) cell lines were stained with the protein-A-purified anti-HERV-H pAb as primary, and an anti-rabbit IgG/Alexa647 coupled secondary antibody (red). Additionally, the cells were co-stained with an anti-tubulin/Alexa488 coupled antibody (green) and a Hoechst33342 stain marking the nucleus (blue). HERV-H transfected DF-1 cells were used as positive control. A: Fluorescent microscopic images were taken with the Keyence BZ-9000 fluorescence microscope at 600x magnification. B: The signal intensities of the HERV-H signals were normalized to the anti-tubulin signals and blotted for every cell line. Each dot represents the normalized fluorescence intensity of one single cell. The red line indicates the median value of all cells within one cell line, taken from a total of 5 images per sample from one experiment. This experiment was replicated a few times, always showing comparable results (data not shown). The asterisks indicate significant differences between the median values of each cell line compared to the DF-1 cells determined using a Kruskal-Wallis test (*: $p \leq 0.05$; **: $p \leq 0.01$; ***: $p \leq 0.001$). The data displayed in this figure were collected during the bachelor's thesis of Lisa Messerer under my experimental supervision.

The ICC assay was further adapted from the fluorescence staining to a diagnostically more conventional DAB staining. Therefore, an HRP-coupled secondary antibody was used which was detected using DAB. The cytoplasm and nucleus were visualized through an H&E diagnostic routine staining. The DAB staining was established using the HERV-H transfected DF-1 cells and a stable HERV-H env expressing NIH-3T3 cell line, which was generated using lentiviral transduction and kindly provided by Sirion Biotech (Graefelfing, Germany). While no DAB signal could be detected in DF-1 cells, some cells in the HERV-H transfected DF-1 cells and all cells from the stable cell line were detected as HERV-H positive (Figure 38A). As control, the cells were stained without the secondary antibody. Here, only the blue H&E staining was visible, indicating that the brown precipitate seen in the complete staining resulted from the DAB staining (Figure 38B).

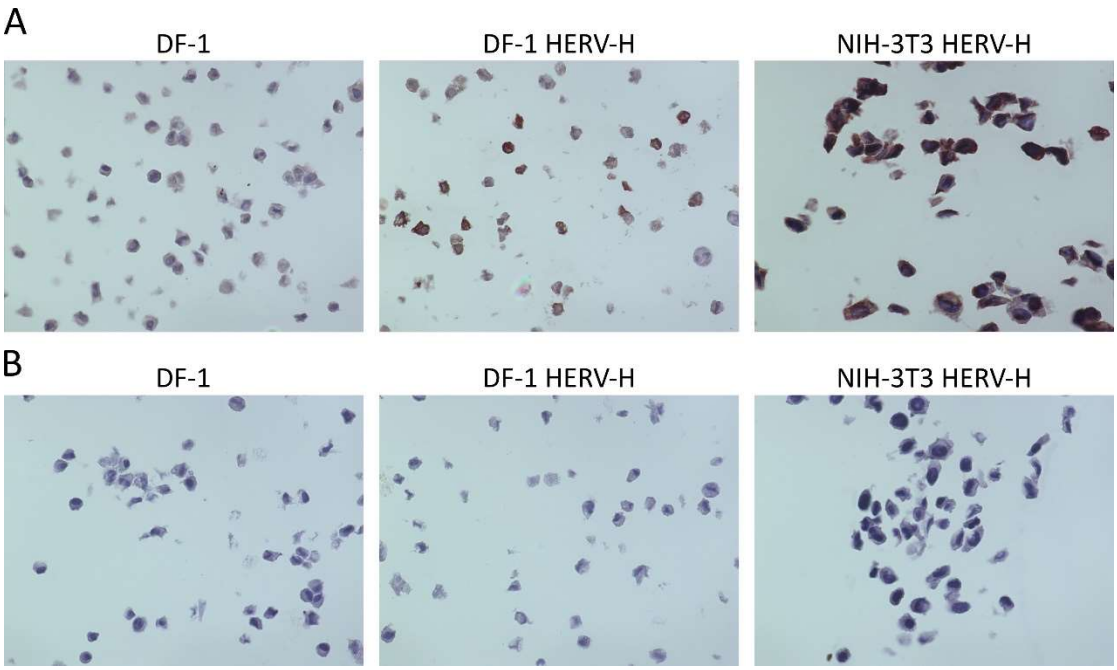


Figure 38: DAB staining with HERV-H env positive and negative FFPE cell lines.

Paraffin-embedded HERV-H positive (DF-1 HERV-H; NIH-3T3 HERV-H) and negative (DF-1) cell lines were stained with the protein-A-purified anti-HERV-H pAb as primary, and an anti-rabbit IgG/HRP coupled secondary antibody. The signal of the secondary antibody was detected by DAB and the cells were counterstained with an H&E staining. Brown precipitation indicates HERV-H env antibody binding. A: Complete staining. B: Control staining without a secondary antibody. The images were obtained by light microscopy using the Keyence BZ-9000 microscope at a 600x magnification.

The ICC data confirmed the protein expression seen in the western blot for the three cell lines C-33A, MCF-7 and MIA-PaCa-2. Furthermore, this experiment builds the foundation for a diagnostic platform, screening tumor tissues for HERV-H env expression.

6.5.3 Effect of Bortezomib Treatment on the HERV-H Env Protein Levels

Since cell stress often leads to changes in gene expression patterns, it was investigated if inducing cell stress in the HERV-H env-expressing human tumor cell lines would lead to an enhanced expression of the HERV-H env protein. Therefore, different cell stress inducers were tested and the HERV-H env protein expression levels were compared to untreated cells. From 4 tested compounds (staurosporine, thapsigargin, H_2O_2 and Bortezomib), only the proteasome inhibitor Bortezomib had an effect on env expression. Here, the three cell lines C-33A, MCF-7 and MIA-PaCa-2 were incubated with different amounts of Bortezomib (1000 nM, 500 nM, 100 nM, 50 nM, 10 nM and 1 nM) for 24 h and the cell lysates were analyzed for their HERV-H env expression in a western blot. In all blots strong bands at approximately 70 kDa could be

detected for the cells treated with high concentrations of Bortezomib, whereas for untreated cells or cells treated with low Bortezomib concentrations, the HERV-H env bands were barely visible (Figure 39A-C). These findings were further validated by analyzing the band intensities of each HERV-H env band normalized to the anti-tubulin signals (Figure 39D).

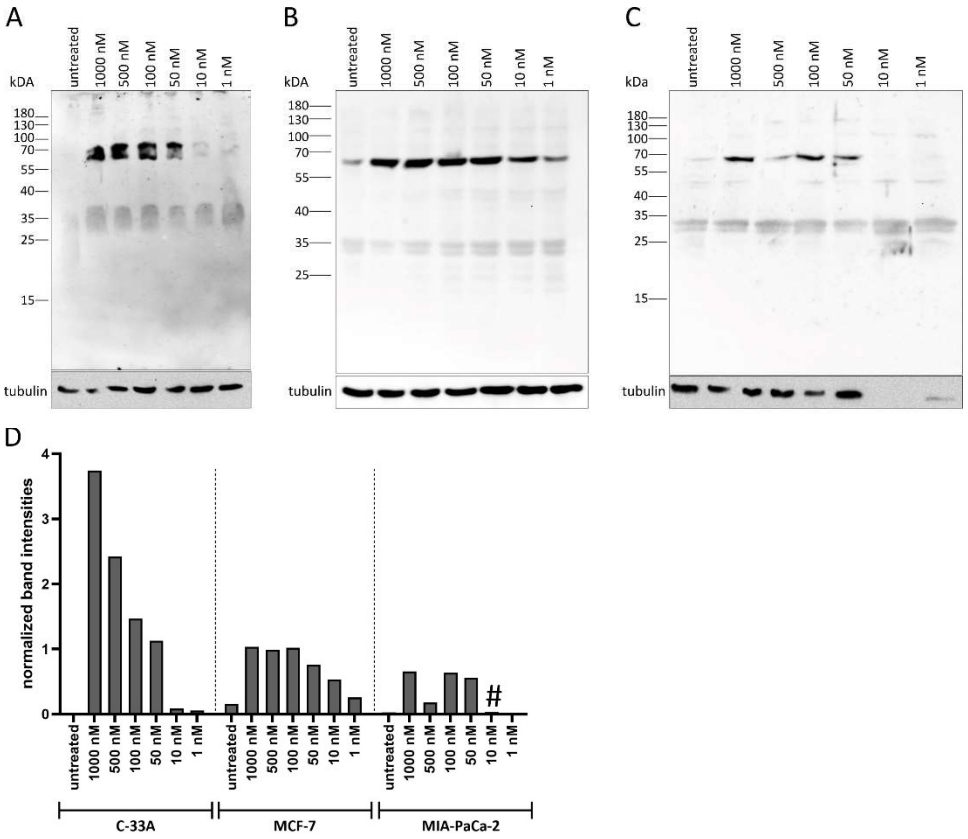


Figure 39: HERV-H env expression of human tumor cells upon treatment with Bortezomib.

Human tumor cell lines were treated with different concentrations of Bortezomib (1000, 500, 100, 50, 10 and 1 nM), and the HERV-H env protein expression was analyzed via western blot. The protein-A-purified polyclonal anti-HERV-H antibody was detected with an anti-rabbit IgG/HRP secondary antibody. As loading control, an anti-tubulin antibody was detected via an anti-mouse IgG/HRP secondary antibody. The PageRuler Prestained Protein ladder was used as molecular weight standard. Western blot of treated and untreated A: C-33A cells, B: MCF-7 cells and C: MIA-PaCa-2 cells. D: The band intensities were analyzed using the GelAnalyzer 23.1.1 software, and the HERV-H env band intensities were normalized to the anti-tubulin bands. #: No band was seen in the anti-tubulin blot for the MIA-PaCa-2 sample treated with 10 nM Bortezomib. Therefore, the signal intensity could not be determined.

Furthermore, it was tested whether the observed effect induced by the Bortezomib treatment could be ascribed to a regulatory effect on a protein or RNA level. Therefore, an RT-qPCR was conducted with RNA isolated from Bortezomib-treated and untreated cells. GAPDH was used in all samples as reference gene (Figure 40A). Analyzing the C_T values of treated and untreated cells revealed quite similar values, with only minor differences between the samples (Figure 40B). Using GAPDH as the reference gene and a factor of 1.29 for the HERV-H PCR efficiency

(Figure S 16B and Table S 3), $\Delta\Delta C_T$ values were calculated for all cell lines relative to the respective untreated control samples (see 5.2.4). The $\Delta\Delta C_T$ values did not show major differences between treated and non-treated cells, except for those treated with 1 nM Bortezomib. These cells showed lower $\Delta\Delta C_T$ values than their respective non-treated samples for all human tumor cell lines (Figure 40C).

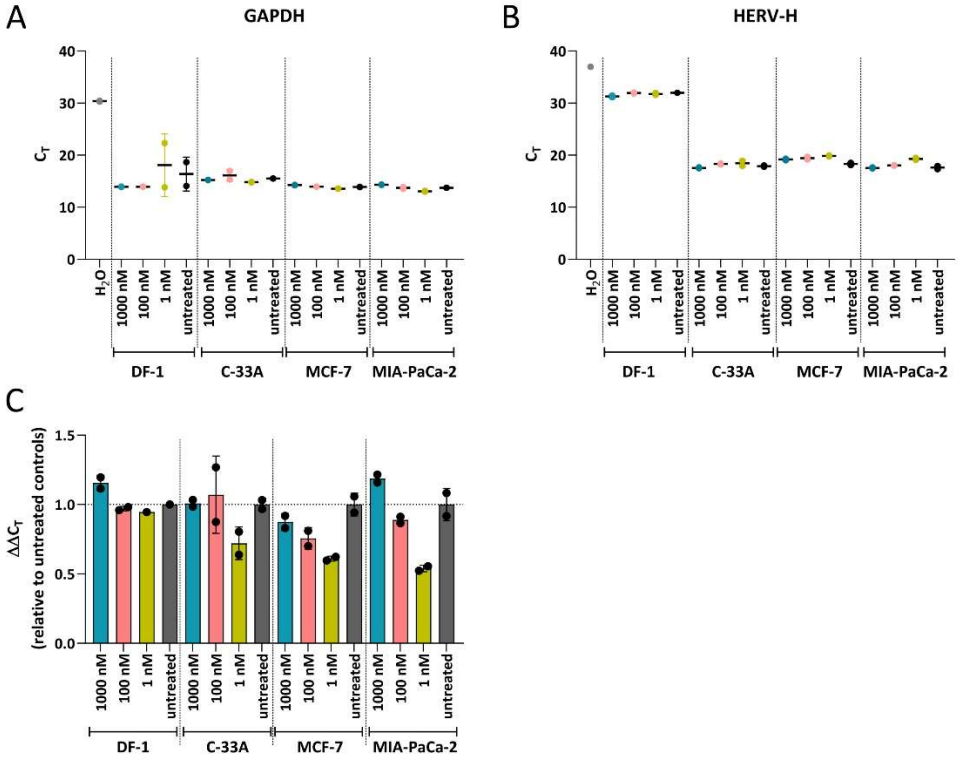


Figure 40: RT-qPCR with RNA extracted from tumor cells treated with Bortezomib.

An RT-qPCR was conducted with RNA extracted from human and non-human cell lines treated with different concentrations of Bortezomib (1000, 100 and 1 nM) using the primer pair published by Yi *et al*. All samples were measured in technical duplicates. A: Depicted are the C_T values of the reference gene GAPDH. B: The C_T values of the RT-qPCR performed with an HERV-H env-specific primer pair are shown. C: $\Delta\Delta C_T$ values were calculated to the respective untreated control sample. The amplification data were analyzed using the StepOnePlus Real-Time PCR software.

In conclusion, the Bortezomib treatment resulted in higher amounts of HERV-H env protein in human tumor cell lines compared to untreated cells. Furthermore, the results from the RT-qPCR indicate, that this effect was not caused on a regulatory RNA level. Therefore, these results hint at a potential degradation of HERV-H env proteins by the proteasome.

6.5.4 Characterization of the HERV-H Env Expression from the Culture Supernatant of Human Tumor Cell Lines

Since the sequencing results from the three HERV-H env positive cell lines indicated a possible expression of a truncated protein caused by premature stop codons (directly or following in/del mutations), the cells' supernatants were analyzed for the presence of a potentially secreted env protein in a western blot assay. To that end, the proteins were precipitated from the supernatant of cultured cells treated with or without Bortezomib and loaded next to the respective cell lysates. As references, the supernatant and cell lysates of DF-1 cells transfected either with the solHERV-H or with the solHERV-H-SUonly plasmids were analyzed. Additionally, cell lysates from DF-1 cells transfected with the membrane-bound HERV-H env were applied. The western blots showed a protein band in the precipitated supernatant for all cell lines, which became more prominent when Bortezomib was added to the cells (Figure 41A-C). The bands from the supernatants exhibited roughly the same size as the protein band from the respective cell lysates, but for MCF-7 and MIA-PaCa-2 cells, additional weaker bands could be detected that were not visible in the cell lysate samples. From these additional bands, the lowest one could be seen at approximately 55 kDa, showing the same size as the SUonly env protein observed in the cell lysates of the transfected cells (Figure 41B + C). The tubulin blot only showed bands for the cell lysates and supernatants of the cell lines, and only weak bands for proteins precipitated from the transfection control. Since tubulin is an intracellular protein, a detection from culture supernatants would not have been expected. However, processing of the cells before precipitation and especially the treatment with Bortezomib could trigger cell membrane damage, thereby releasing cellular proteins into the supernatant. Due to the low concentrations used, the cell lysates of the transfection controls exhibited no tubulin band (Figure 41).

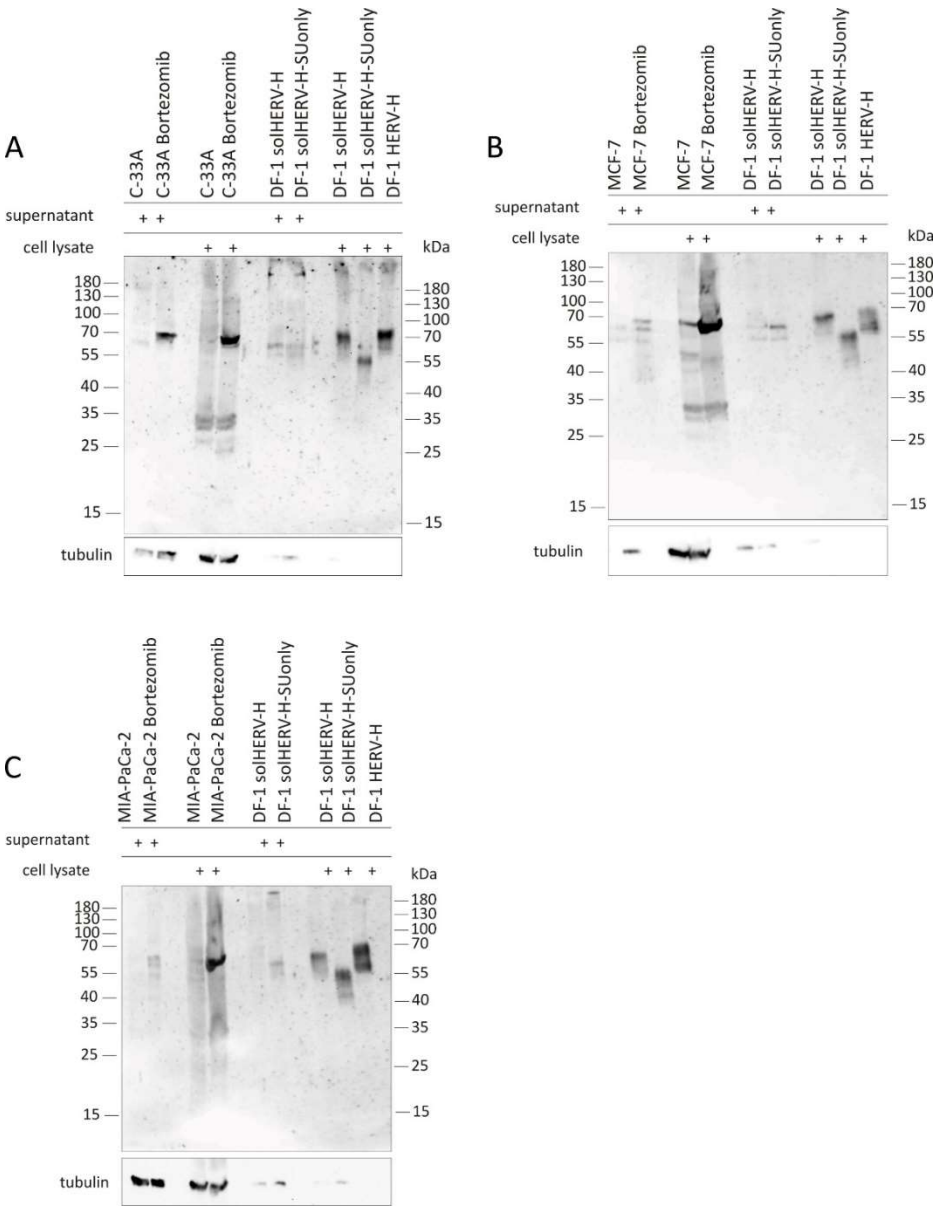


Figure 41: Western blot of HERV-H env protein precipitated from the supernatant of human tumor cell lines.

Human tumor cell lines were treated with 500 nM Bortezomib or left untreated. Cell lysates were prepared and secreted proteins were precipitated from the cell culture supernatant using TCA. As references, DF-1 cells were transfected with the membrane-bound HERV-H env or the soluble solHERV-H and solHERV-H-SUonly DNA constructs, and cell lysates as well as the supernatants were prepared. Western blots of supernatants and cell lysates of A: C-33A cells, B: MCF-7 cells and C: MIA-PaCa-2 cells. All blots were stained with the polyclonal protein-A-purified anti-HERV-H env antibody as primary and an anti-rabbit IgG/HRP as secondary antibody. As loading control, the blots were stained with an anti-tubulin antibody. As molecular weight standard the PageRuler Prestained Protein ladder was used on both sides of the SDS gel.

Concluding, it could be shown, that HERV-H env protein was present in both, cell lysates and culture supernatants of human tumor cell lines. Furthermore, some of these proteins exhibited comparable molecular weights, indicating that the secreted proteins have a quite similar form to the cellular proteins. However, the presence of tubulin in the samples

precipitated from the supernatants indicate, that cellular proteins could also be found in the supernatant to some extent.

6.6 Characterization of HERV-H Env Expression in Human Colorectal Cancer Tissue

The established ICC methods for detecting the HERV-H env protein in FFPE tissue were further evaluated in a diagnostic immunohistochemistry (IHC) approach. Tumor tissues of three different CRC cases (19172/23_B, 23442/23_B and 31786/23_B) were kindly prepared and provided by Prof. Dr. Daniela Hirsch and Doris Gaag from the UKR pathology institute. Additionally, sections from one healthy colorectal tissue (23422/23_D) and normal colon tissue prepared on a tissue microarray (Colon TMA) were provided as controls. Fluorescence and DAB IHC assays were conducted as described earlier for the FFPE cell lines and microscopic images were taken via fluorescence or light microscopy. For the fluorescence staining anti-HERV-H env signal intensities of different regions were assessed and normalized to the respective tubulin signals. As an internal assay control, DF-1 and DF-1 HERV-H FFPE slides were included in both approaches (Figure S 17).

The fluorescence staining showed no HERV-H env signal that was visible to the naked eye, neither for the normal nor the CRC tissues (Figure 42A + B). However, when comparing the tubulin-normalized anti-HERV-H fluorescence intensities, significant differences could be obtained between each normal tissue control and each CRC tissue sample, with p values ranging from 0.005 to <0.001 (Figure 42C; Table 31).

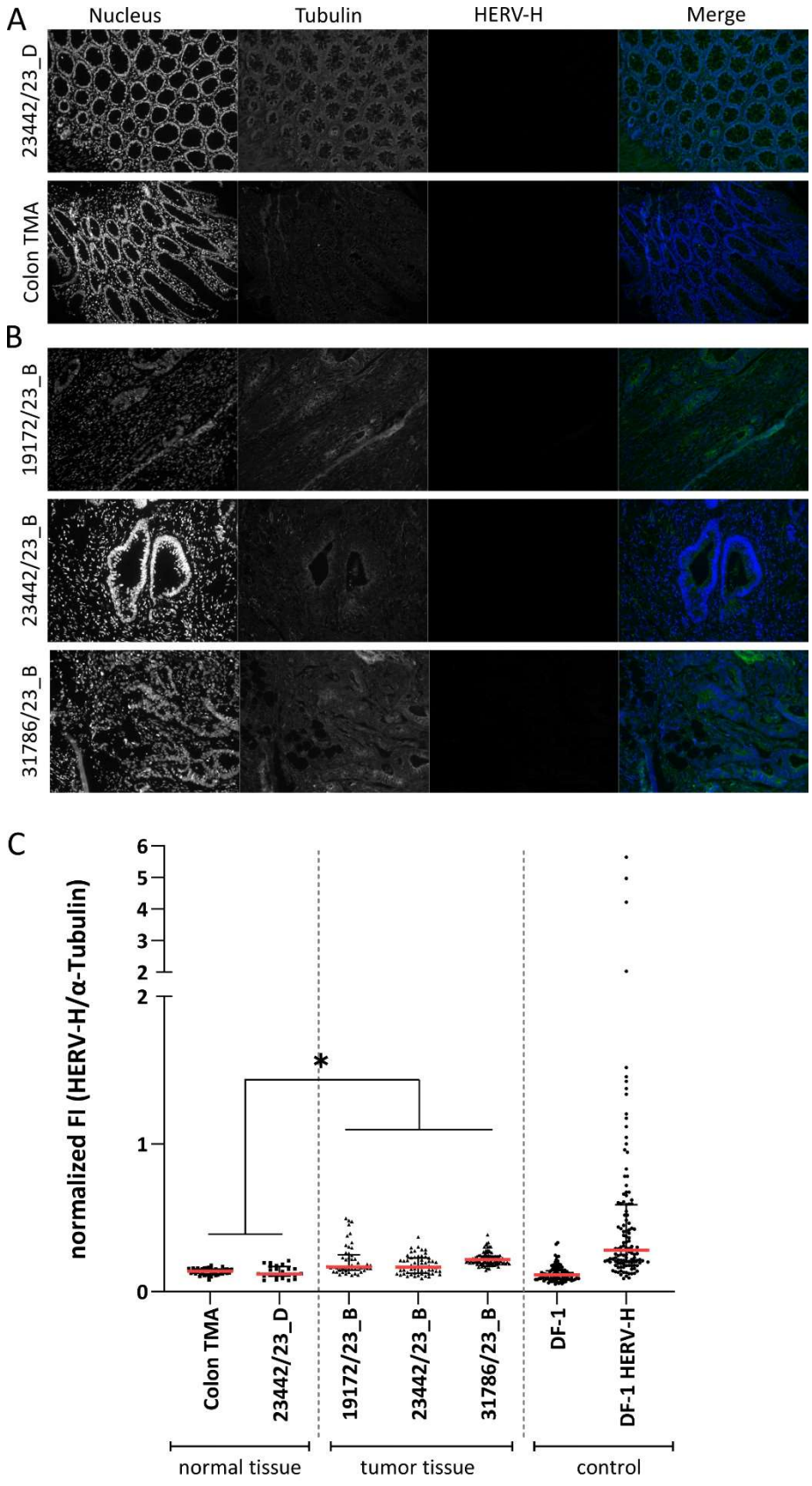


Figure 42: IHC staining of healthy and tumorous colon and colorectal tissue.

FFPE tissue samples were stained in an IHC assay against the HERV-H env protein using the protein-A-purified pAb and an Alexa647-labeled anti-rabbit secondary antibody (red). The samples were co-stained with an Alexa488-labeled anti-tubulin antibody (green) and Hoechst33342 staining the nucleus (blue). Images were taken with the

Keyence BZ-9000 fluorescence microscope with a 200x magnification. A: Tissue samples of healthy colon (Colon TMA) and colorectal (23442/23_D) tissue were stained. B: Staining of tissue samples from three different CRC patients. C: The signal intensities of the HERV-H signals were normalized to the anti-tubulin signals and blotted for every FFPE sample. Each dot represents the normalized fluorescence intensity of one defined tissue region. The red line indicates the median value of all regions within one tissue sample, measured from 5 images per sample. The asterisk indicates significant differences between the median values of each tumor tissue compared to both normal colon tissue samples determined using a Kruskal-Wallis test. A detailed list of all p values can be found in Table 31.

Table 31: p values of normalized signal intensities between normal and CRC tissues from the IHC fluorescence staining.

	Colon TMA	**23442/23_D**
19172/23_B	<0.001	<0.001
23442/23_B	0.002	0.005
31786/23_B	<0.001	<0.001

Tissue samples were also subjected to DAB staining. However, no differences in the signal intensity between normal and CRC tissues could be observed. Isolating the DAB signal through color deconvolution revealed no staining above the background for any sample (Figure 43). The internal assay control showed HERV-H env staining in the positive control (DF-1 HERV-H), indicating that the staining procedure itself was successful (Figure S 17B).

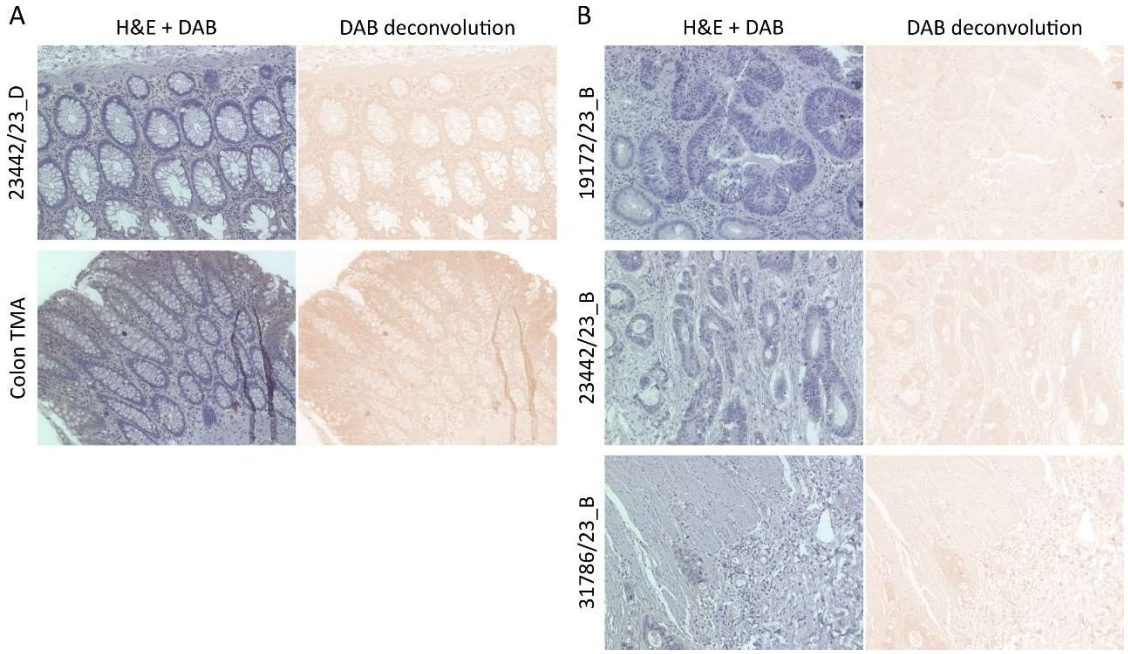


Figure 43: DAB staining of normal colon/colorectal and CRC tissue.

FFPE colon tissue samples were stained with the protein-A-purified anti-HERV-H env pAb as primary, and an anti-rabbit IgG/HRP coupled secondary antibody. The signal of the secondary antibody was detected by DAB staining and the cells were counterstained by H&E staining. A: Healthy colon (Colon TMA) and colorectal (23442/23_D) tissue samples were stained. B: CRC tissue samples were stained for HERV-H env expression. A+B: The images on the left hand side show the complete DAB and H&E counterstaining while the images on the right show the isolated DAB signal after color deconvolution. All images were taken at 200x magnification using the Keyence BZ-9000 microscope.

The results of both IHC staining methods showed no HERV-H env staining that was visible to the naked eye. However, in the fluorescence staining the normalized signal intensities of the CRC tissues were significantly higher than the ones of the normal tissue control, indicating a potentially weak HERV-H env expression. However, these results indicate, that the IHC staining with tumor and healthy tissues should be further optimized, and that a second diagnostic approach via for example RT-qPCR would be useful to underline the results.

7 Discussion

With nearly 10 million deaths in 2020, cancer remains one of the leading causes for death worldwide (*https://www.who.int/news-room/fact-sheets/detail/cancer*). Identifying new diagnostic markers and targets for immunotherapy could be a key player in the treatment of cancer. Human endogenous retroviruses, and especially their env proteins are promising candidates, due to their frequent association of enhanced transcription and expression in a variety of cancers. Within this work, detection reagents targeting the HERV-H env protein were generated and used to assess human tumor cell lines for their expression of the HERV-H env.

7.1 HERV-H Env is a Suitable Candidate for Cancer-Immunotherapy

Based on the occurrence of their transcription and expression in cancer, env proteins from the 5 HERV families HERV-FRD, HERV-K, HERV-W, HERV-H and ERV3.1 were chosen as potential candidates for anti-cancer immunotherapy (Bermejo *et al.*, 2020). A complete env sequence was available for each HERV family and it could be shown, that all the sequences indeed expressed an env protein upon transfection of HEK293 T. Regarding protein expression levels, HERV-W exhibited the highest expression, followed by HERV-FRD and HERV-H (Figure 4).

A fusogenic activity could only be observed in the env proteins of HERV-W and HERV-FRD, but not in HERV-H or the others (Figure 5). This result is in agreement with the current state of knowledge because so far, cell-cell fusion as a biological function is only known for HERV-W and HERV-FRD (Blaise *et al.*, 2003; Bonnaud *et al.*, 2004).

When using the env proteins as targets for immunotherapy through for example therapeutic vaccines, this activity could become problematic though. For instance, the production of recombinant viral vectors is highly dependent on cell culture work. A vector encoding a protein able to fuse the cells used for its development would most probably interfere with the production success. Also, regarding clinical studies, vaccination with a fusogenic antigen could induce unwanted fusion-related side effects. Therefore, a non-fusogenic protein like HERV-H env would be favored. Even though, the fusion activity was abrogated when exchanging the transmembrane-anchored domain and the cytoplasmic tail in the HERV-W and HERV-FRD env proteins, the modification also led to a highly reduced protein expression in HERV-W, indicating the functional importance of the region. Therefore, it was investigated if there are

further options to abrogate this function while keeping the impact on the protein structure and expression as limited as possible. The results of a N-glycosylation knock-out based on a study by Cui *et al.* led to a reduced or completely abrogated cell fusion for HERV-FRD (Cui *et al.*, 2016). However, the knock-out also led to a significantly reduced protein expression compared to the WT HERV-FRD env. For HERV-W, this method of cell-fusion abrogation was not successful (Figure 7).

Taken together, regarding the combination of protein expression levels and the lack of cell-cell fusion activity in its WT form, the HERV-H env showed to be a favorable candidate for a HERV env-based approach in cancer immunotherapy. Furthermore, HERV-H is one of the HERVs most frequently associated with different types of cancer, and transcription of the *env* sequence has been repeatedly observed especially in colorectal cancer (Wentzensen *et al.*, 2004, 2007; Alves *et al.*, 2008; Pérot *et al.*, 2015).

7.2 The HERV-H Env Protein Elicits a Specific Antibody Response in BALB/c Mice

Two main approaches were pursued to target HERV env proteins in the immunotherapeutic context: the development of therapeutic vaccines (main work package of the cooperation partners) and the production of therapeutic monoclonal antibodies. For both approaches, suitable immunogens eliciting a protein-specific antibody response are of great importance. To that matter, the immunogenicity of the HERV-H env protein was assessed by immunizing BALB/c mice with a variety of WT or modified env protein variants, either as DNA vaccination or by immunization with recombinant soluble protein.

First results showed that two DNA immunizations already resulted in high, but not uniform antibody titers (Figure 18). Nevertheless, these results indicate that the DNA vectors could serve as a valuable vaccination platform. DNA vaccines have several advantages compared to protein vaccines. First, a DNA vector vaccine enables immunization with the full-length, membrane-bound env protein which would not be possible in a platform using recombinant, soluble proteins. In addition to that, as reviewed by Suschak *et al.*, DNA vaccines elicit an *in vivo* production of the antigen, thereby enabling the presentation on both, MHC I and II, leading to a specific B-cell and cytotoxic T-cell (CTL) response (Suschak, Williams and Schmaljohn, 2017). Especially the induction of a strong CTL response could be advantageous

in the context of therapeutic anti-cancer vaccination. The use of viral vectors also provides the possibilities of full-length antigen presentation and CTL induction, but compared to viral vectors, DNA vaccines are easier in production, long-time storage, and overall stability (Suschak, Williams and Schmaljohn, 2017).

The high responses of the protein boost and the homologous protein prime-boost group (Figure 18) proved the potential of the recombinant protein to be a strong antigen for the generation of anti-HERV-H env antibodies in mice, despite its aggregated state. Aggregation of retroviral recombinant soluble env proteins is a known challenge. For example, findings of the early HIV-1 research and env purification trials showed, that without stabilizing modifications the soluble HIV-1 env protein tends towards aggregation (Binley *et al.*, 2000). A stabilization method using an artificial trimerization domain was not feasible for the HERV-H env protein, though (see Figure S 9).

Regarding the ISD, no significant differences were observed between groups immunized with the WT-ISD and their respective ISDmut counterparts (Figure 19). So far, the ISD activity of HERV-H has only been shown on a cellular level (Mangeney *et al.*, 2001). An activity on a humoral immune response has not been analyzed yet. However, since Mangeney *et al.* could prove an effect of the ISD mutation on an antibody level in mice for both, syncytin-1 and -2, it was a reasonable expectation to see a similar effect for HERV-H env (Mangeney *et al.*, 2007). It should be noted though, that Mangeney *et al.* did not use the complete soluble proteins for their immunization experiments, but only the trimerized ectodomain of the TM subunit containing the ISD. Furthermore, several studies on the isolated ISD of different HERV and HIV-1 env proteins showed, that a certain degree of oligomerization was necessary for the ISD peptides to be active (Denner, Norley and Kurth, 1994; Haraguchi, Good and Day-Good, 2008; Tolosa *et al.*, 2012). Considering the protein aggregation of the recombinant HERV-H env protein used, it can be assumed that the ISD will probably not assemble into an active form as described above, thereby potentially explaining the lack of differences in the antibody response between ISD-WT and ISDmut proteins.

Regarding therapeutic anti-cancer vaccination strategies targeting murine and human ERV env proteins, an inactivation of the ISD has been shown to be beneficial before. A study conducted by Daradoumis *et al.* with the ISD-mutated envelope protein of the endogenous murine melanoma-associated retrovirus (MelARV) showed a positive effect of the mutation regarding

T-cell and anti-tumor immunogenicity in mice (Daradoumis *et al.*, 2023). Another study that was conducted with our cooperation partners could show the same effect for a HERV-W env targeted vaccine. Here, the WT variant, which is known for its non-immunosuppressive activity, elicited a higher specific T cell response than a presumably immunosuppressive mutant (Skandorff *et al.*, 2023). However, in both studies no significant differences were observed regarding the antibody responses. Nevertheless, these findings underline the importance of the ISD in a (H)ERV-targeted vaccination strategy at least on the cellular level. Therefore, more research on the HERV-H ISD activity, and especially an inactivation thereof, should be conducted in the feature.

All in all, the HERV-H env was shown to be a strong antibody-inducing immunogen, both as DNA and soluble protein, thus being suitable for vaccine development and the generation of HERV-H env targeted antibodies. Nevertheless, it should be noted that these strong antibody responses were observed in mice, an organism not known to harbor HERV-H elements in its genome. Whether a HERV-H vaccine in humans could overcome the potential immunogenic self-tolerance remains an unanswered question for now. This limitation is one of the main disadvantages of therapeutic anti-cancer vaccination strategies (Makkouk and Weiner, 2015; Hollingsworth and Jansen, 2019). An alternative immunotherapeutic strategy that could overcome this limitation is the adaptive transfer of immune effectors. Here, tumor-specific monoclonal antibodies or T cells engineered to express chimeric antigen receptors (CAR) can be directly administered, thereby bypassing the need for *in vivo* antigen presentation (Makkouk and Weiner, 2015). Therapeutic mAbs targeting CD20 in Non-Hodgkin's lymphoma, and anti-HER-2 targeted mAbs that are used in breast cancer treatment, are just two examples of FDA-approved anti-cancer mAbs (Vacchelli *et al.*, 2013). CAR T cells are generated by fusing a single-chain variable fragment to a T-cell receptor signaling domain (Ellis, Sheppard and Riley, 2021). They can be used in a wide spectrum of tumors, ranging from solid tumors (for example CXCR2-targeting CAR T cells in pancreatic tumors) to hematologic malignancies (for example CD20- and CD19-trageting CAR T cells in B cell lymphoma) (Jogalekar *et al.*, 2022).

7.3 HERV-H Env Immunization favors the Generation of IgM-secreting Hybridoma Clones

As discussed above (see section 7.2), mAbs against the env protein would be a useful tool in the context of HERV-targeted immunotherapy. Therapeutic mAbs could be either administered as the purified mAb itself or could be used as a module in CAR T cells. In extension to the therapeutic application, mAbs could also be valuable in the diagnostic context. Here, mAbs could be useful by identifying env-expressing tumor tissue, thereby enabling tumor stratification.

In this work, the generation of HERV-H env-specific mAbs was approached using the hybridoma technology, which was first described by Köhler and Milstein in 1975 and has been the gold standard for the production of mAbs since then (Köhler and Milstein, 1975). However, even despite several attempts, it was not possible to generate specific, strongly binding mAbs targeting HERV-H env within this thesis.

So far, it is unknown why the hybridoma generation turned out to be so difficult. There are some reports on hybridomas being very unstable and losing mAb production (Harris and Koropatnick Pearson, 1990; Grigsby, Fairbairn and O'Neill, 1993; Barnes, Bentley and Dickson, 2003). However, it is unclear why the first batch of anti-HERV-H hybridomas exhibited binding to multiple, non-HERV-H proteins (Figure 23). The HERV-H env-specific binding of a control serum from one of the mice used for the fusion indicated, that the immunization itself elicited a specific antibody response. Therefore, the non-specificity of the hybridoma clones is most likely a result of the hybridoma fusion process. Furthermore, it could be excluded, that the cross-binding to other proteins was induced by mAbs binding to the C-terminal His6-tag instead of a HERV-H env epitope, since binding was also detected to a non-tagged HIV-1 env protein (Figure 23). Overall, literature on cross-reactive hybridoma clones is scarce. Nevertheless, there is one study showing genetic diversity in hybridomas containing one or more additional productive heavy or light antibody chains (Bradbury *et al.*, 2018). Bradbury *et al.* observed this diversity in approximately 32 % of examined hybridomas, whereas in the first HERV-H attempt, 100 % of hybridomas showed this particular behavior. Therefore, it is questionable, if the findings from Bradbury and colleagues can be transferred to the cross-reactivity observed in the HERV-H hybridomas, unless the occurrence of genetic diversity is a target-dependent effect.

Furthermore, the HERV-H env hybridoma generation seemed to favor IgM-secreting hybridomas (Figure 22 and Figure 25). The reason for this also remains unclear. A study by Lee *et al.* on Ebolavirus glycoprotein hybridomas showed, that DNA immunizations preferentially generated IgM-producing hybridomas. Yet, a booster immunization with protein shifted the preference to IgG-producing hybridomas (Lee *et al.*, 2017). Therefore, the DNA vaccination may explain the IgM-secreting hybridoma from the second attempt where only RNA and DNA vaccines were used, but not the high abundance of IgM hybridomas in the first attempt where the mice were boosted with not only one, but three protein immunizations.

In the second attempt to generate mAbs, two IgG-producing clones could be identified, but the clones lost their antibody production capacity over time (data not shown). Nevertheless, efforts were made to save the antibody sequences from the two clones by following a published protocol for the molecular cloning of mAbs from hybridomas (Meyer *et al.*, 2019). However, only sequences from an aberrant κ light chain could be obtained. This sequence has been identified before in hybridomas resulting from a non-functional rearrangement of an aberrant transcript found in the non-secreting myeloma fusion partner (Carroll, Mendel and Levy, 1988). Carroll *et al.* could also show that the expression of this sequence could be quite high and even exceed the expression levels of the productive light chain, which also seemed to be the case in the two HERV-H hybridomas where this sequence was found quite often (data not shown).

Lastly, even though it was possible to purify three IgM mAbs in the study performed by the contractor, antibody yields were low and the mAbs showed no env-binding (Figure 26). This phenomenon of purified IgM mAbs no longer showing binding to their antigen, that was previously bound by the hybridoma supernatant, was also observed by Lee *et al.* (Lee *et al.*, 2017). More importantly, they observed this effect not only for the Ebolavirus glycoprotein hybridomas of this study, but also for other surface antigens of infectious viruses.

Taken together, all attempts to generate HERV-H env-specific mAbs failed. It is impossible to pinpoint the exact reason why the hybridoma generation does not seem to work properly for HERV-H. So far, all evidence hints at the target being problematic for this kind of approach.

7.4 Tumor Stratification through HERV-H Env Screening

If targeting the HERV-H env protein should be implemented as potential cancer immunotherapy, it would be indispensable to develop a diagnostic platform proving HERV-H env expression, and more importantly surface expression, on tumor tissue. Two methods that are highly implemented in diagnostic tumor screening routines are RT-qPCR and IHC staining assays (Bernard and Wittwer, 2002; Duraiyan *et al.*, 2012). These assays represent two different approaches to detect on the one hand HERV-H *env* transcription (RT-qPCR), and on the other hand protein expression (IHC) in tumor tissue samples. Both approaches should be implemented to get the most reliable diagnostic outcome. Since tumor tissue samples are rare materials, the first steps in the development of these methods were conducted using human tumor cell lines presumably transcribing HERV-H env based on literature data (Yi, Kim and Kim, 2005) as replacement.

7.4.1 HERV-H Env Can Be Detected in Several Human Tumor Cell Lines on mRNA and gDNA Level

RT-PCR and RT-qPCR assays using a previously published primer pair binding the 600 bp 3' part of the HERV-H *env* sequence (Yi, Kim and Kim, 2005) could show HERV-H *env* transcription in a panel of human, primate and non-primate tumor cell lines (Figure 31). These results confirmed the findings of Yi *et al.* (cell lines C-33A, MIA-PaCa-2, MCF-7, HepG2 and A549), and extended them by also detecting HERV-H *env* transcription in the human (tumor) cell lines MDA-MB-231, HEK293, THP-1 and A673, as well as the green monkey cell line VeroE6. The finding of HERV-H *env* transcription in the green monkey cell line was not surprising, considering that the majority of HERV-H integrations happened around 30-40 million years ago, while the Old World monkeys, where the green monkeys are categorized, separated from the primate lineage approximately 30 million years ago, thereby after HERV-H integration happened (Magiorkinis, Blanco-Melo and Belshaw, 2015; Izsvák *et al.*, 2016).

Also, even though the HERV-H env primer efficiency was too low for a quantitative read-out (30-60 %, depending on the method used) (Figure S 16; Table S 3), a RT-qPCR was conducted with FFPE cell line samples, to get a first insight if this assay would be suitable in a diagnostic context. However, compared to fresh mRNA samples, where the C_T values of HERV-H env positive and negative cell lines could be clearly separated, the results obtained from the FFPE samples were not as clear. While it was still possible to separate the C_T values from the HERV-

H env positive cell lines from those of the negative controls, the values were generally much higher for both, HERV-H env and GAPDH, compared to the fresh sample controls. These results show that for a diagnostic RT-qPCR test optimization of the HERV-H env primer pair would be necessary.

Furthermore, even though it was possible to show HERV-H *env* transcription with the RT-(q)PCR, this method is highly limited regarding the proof of an actual intact HERV-H *env* open reading frame. The primer pair used here only detects a small fragment of the 3' part of the *env* sequence, thereby not giving any evidence regarding the integrity of the rest of the sequence. Importantly, it should also be mentioned that Yi *et al.* sequenced the obtained RT-PCR fragments and could only show an intact sequence in the three cell lines C-33A, MCF-7 and MIA-PaCa-2, while all other cell lines contained frameshift mutations and premature stop codons (Yi, Kim and Kim, 2005). Efforts were also made to conduct the RT-qPCR with different primer pairs amplifying the complete *env* sequence, but even though amplification could be detected in the positive cell lines, again only small fragments were amplified (data not shown). However, since RT-qPCR works best with small amplicons and is usually not used for long-fragment amplification these results were not surprising. Considering all these limitations, the RT-(q)PCR method only gives a first indication about the presence of a partial HERV-H *env* transcript in tumor cells, but not about the integrity of the sequence.

To tackle this issue and to assess the integrity of the complete *env* sequence encoded by the human tumor cell lines, the HERV-H *env* reading frame was amplified and sequenced from isolated gDNA. In 2001, De Parseval *et al.* already showed the existence of three HERV-H proviruses encoding a complete *env* sequence isolated from DNA of human peripheral blood leukocytes (De Parseval *et al.*, 2001). Furthermore, they could also show that these sequences get translated via a protein truncation test. Therefore, this sequencing experiment further served the purpose to confirm the existence of these ORFs in the human tumor cell lines. The results obtained showed a broad range of different *env* sequences with frameshift mutations and premature stop codons at various sites (Table 29). This variety of *env* sequences was not unexpected, considering that approximately 100 HERV-H sites contain a putative *env* sequence (Hirose, Takamatsu and Harada, 1993). The finding of one complete HERV-H *env* sequence, corresponding to 2.4 % of all analyzed sequences, furthermore confirmed the findings of De Parseval and colleagues. They identified the three complete *env* sequence out

of the approximately 100 sites, which corresponds to a quite similar ratio of complete-to-incomplete *env* sequences (De Parseval *et al.*, 2001).

Taken together, even though RT-(q)PCR assays are valuable tools in a diagnostic context, they can only give a first inside whether human tumor cells transcribe HERV-H *env* on an mRNA level. However, they do not give any indication regarding the sequence integrity. Also, the detection of a full-length HERV-H *env* DNA sequence only gives an indication, that these cells possess the genetic prerequisite to express a membrane-anchored env protein, but not if the protein will be translated and surface-expressed. However, this would be a prerequisite for a HERV-H env-targeted immunotherapeutic strategy. Therefore, the RT-(q)PCR can be used as a first screening method to identify HERV-H *env* transcribing tumors, but for positive samples a subsequent detection of actual translated HERV-H env protein would be inevitable.

7.4.2 HERV-H Env Can Be Detected on a Protein Level in Human Tumor Cell Lines

Using the generated protein-A-purified polyclonal anti-HERV-H env antibody (see section 6.4), the expression of HERV-H env on a protein level could be shown for all human tumor cell lines tested in both, a western blot and an ICC assay. These results extend the RNA data by Yi *et al*. (Yi, Kim and Kim, 2005) by detecting a translated HERV-H env product, which has not been shown in these cell lines before. However, due to the nature of both assays (cell lysis for western blot and permeabilization of the cell membrane through processing and fixating of the sample in ICC) they do not indicate whether the observed env proteins were intact and presented on the cell surface. For an env-targeted cancer immunotherapy this would be a crucial prerequisite though, because only proteins presented on the cell surface can be detected by mAbs or vaccine-induced antibodies. Nevertheless, these assays can be used for a preselection of tumors expressing the HERV-H env protein.

Treating the cells with Bortezomib elevated the expression levels seen in the western blot. Bortezomib, formerly known as PS341, is a proteasome inhibitor that hinders the cell to degrade mis- or unfolded proteins (Adams *et al.*, 1999). It is a dipeptide boric acid homolog that forms a complex with the threonine hydroxyl group in the chymotrypsin-like active site of the 26S proteasome, thus inhibiting its proteolysis activity (Adams *et al.*, 1998). Since the proteasome degrades misfolded proteins via a ubiquitination signal, inhibiting its function would lead to an enhanced level of wrongly folded proteins (Rock *et al.*, 1994; Coux, Tanaka

and Goldberg, 1996). The elevated HERV-H env levels observed in cells treated with Bortezomib indicate, that the major proportion of the expressed HERV-H proteins is somehow misfolded and thereby normally degraded. Interestingly, this phenomenon was not only seen in proteins from cell lysates, but also in secreted proteins that were precipitated from the cell culture supernatants. This further indicates that a substantial amount of expressed env proteins is structurally impaired, and simultaneously confirms the observations from the sequencing results, showing a presumably secreted form of the env protein caused by a premature stop codon. These data strongly support the findings by Mullins and Linnebacher from 2011. Via a reverse immunology approach, they could show, that retroviral peptides displaying HLA-A2.1-binding motifs based on a 273 amino acid long HERV-H env protein induced the proliferation of peripheral $CD8^+$ T cells. Furthermore, these peptide-specific CTLs were able to lyse CRC cancer lines (Mullins and Linnebacher, 2012). They suggest, that these env proteins are expressed, degraded by the proteasome, and subsequently presented on HLA molecules (Mullins and Linnebacher, 2012). Their findings combined with the results shown here indicate that aside from the antibody-inducing approach, HERV-H env could also be an interesting target towards a specific T cell-inducing therapeutic vaccination strategy.

In summary, even though HERV-H env protein expression was successfully detected in human tumor cell lines, the form in which the env proteins are presented remains highly speculative for the time being. A surface-staining flow cytometry assay would resolve the question if HERV-H env is presented on the surface of tumor cells, providing that a suitable HERV-H env-targeted antibody can be generated.

7.5 The ICC Staining Can Be Adapted to CRC Tissue Samples

In the last part of this thesis, the previously established ICC assay (see section 6.5.2) was applied to healthy colon and colorectal, as well as CRC tissue samples (see section 6.6). In general, both methods, the fluorescence and the DAB staining, could be transferred to the tissue samples as shown by the successful staining of the nucleus and cytoplasm in both. However, both staining methods did not result in a HERV-H env signal that was visible to the naked eye. Even though the relative signal intensities measured from the fluorescence staining showed a significant higher value for all CRC tissues compared to the healthy controls, these results should be considered with caution due to the limited sample size. To validate

these findings, it would be crucial to assess a larger panel of healthy and tumorous tissue samples.

Nevertheless, this work shows a first attempt at detecting HERV-H env expression in actual tumor tissue. So far, all data regarding the HERV-H env association with CRC are limited to the detection of *env* transcripts. Therefore, even if the first results remain inconclusive, the IHC assay shown here provides a valuable starting point for the detection of HERV-H env in tumor tissues.

8 Conclusion

In this thesis, HERV env proteins were assessed as potential targets for cancer immunotherapy. Therefore, env antigens from 5 HERV families were designed and tested for their protein expression capacities and protein activity regarding cell-cell fusion properties. Furthermore, the immunogenic properties of different HERV-H env variants were assessed in a mouse immunization experiment and monoclonal and polyclonal antibodies targeting HERV-H env were generated and characterized for their binding properties. Lastly, diagnostic platforms to screen for HERV-H env protein transcription and expression in tumors on gDNA, mRNA and protein levels were developed and assessed.

In conclusion, this work could show that:

I. All five evaluated HERV families possess protein-expressing env sequences that can be used as potential antigens for immunization experiments or vaccine development. In most proteins, modifications within the protein sequence did not enhance expression levels compared to the WT variants. However, further modifications of the HERV-W and HERV-FRD env proteins were necessary to abrogate their fusogenic activities.

II. HERV-H env DNA and protein immunizations induced specific immune responses in mice. These responses did not differ with regard to env modifications (including ISD and TM modifications). Overall, all immunization regimens were effective in inducing a humoral response.

III. Generating monoclonal Abs targeting HERV-H env failed in all attempts. However, a suite of polyclonal antibody preparations could be generated that exhibited HERV-H env-specific binding in different applications.

IV. All diagnostical methods established in this work could indicate some sort of HERV-H env transcription and expression in the tested tumor cell lines. However, due to the nature of these methods, it was not possible to distinctly prove protein expression on the cell surface. Sequencing results showed a high proportion of HERV-H *env* sequences within tumor cell lines are truncated and unfunctional. Yet, a full-length sequence that encodes a functional reading-frame could be detected in one cell line. Additionally, it could be shown that most of the actually expressed

env proteins get degraded by the proteasome, which hints at a possible misfolding of these proteins.

Taken together, even though it is tempting to propose HERV env proteins as new tumor antigens due to the strong evidence of their enhanced transcription in cancer, the case of HERV-H shown here highlights the challenges that can be faced for such an approach toward tumor therapy. Not only unexpected difficulties in the development of mAbs occurred, but also in the actual detection of HERV-H env in tumor samples, which is a crucial part of any clinical application. Moreover, the observed extensive protein degradation by the proteasome raises the question in which amounts HERV-H env is actually expressed on the surface of tumor cells and if this is enough for an env-targeted immunotherapy approach. Therefore, while this work provides an initial insight into the topic of HERV-H env targeted cancer immunotherapy and tumor stratification, more research to improve the diagnostic methods is needed.

9 Outlook

Regarding the insights gained about HERV-H env in this work, a lot of research and methodical improvement should be considered in the future. For example, a new method to develop monoclonal antibodies should be conducted, to overcome the problems that occurred with the hybridoma technology. Two of the most used alternatives for this are single B cell screening or phage-display screening (Schardt, Sivaneri and Tessier, 2024). In the single B cell screening, B cells of immunized mice are first purified by magnetic-activated cell sorting (MACS), followed by sorting antigen-binding cells using fluorescence-activated cell sorting (FACS). Afterwards, the transcripts of the Ig heavy (V_H) and light chain (V_L) variable region are amplified via molecular cloning technologies, and on that basis the mAbs can be produced recombinantly (Tiller, Busse and Wardemann, 2009). The phage-display screening takes advantage of the ability of bacteriophages to present peptides or protein fragments on their surface. Therefore, first, a library of all V_H and V_L genes from the B cells of immunized mice is generated by RT-PCR. These cDNA sequences are then fused to the sequence encoding for the gene III phage coat protein of the phage, thereby enabling the expression and presentation of the antibody fragments (Chan *et al.*, 2011). These phage libraries can then be screened for binding to the desired antigen. Importantly, with both methods, only the sequences of the variable region of the mAbs will be obtained. Therefore, it is necessary to further clone these sequences into expression vectors containing an appropriate antibody backbone. However, compared to the hybridoma technology where pure, but completely murine antibodies are generated, this enables the production of chimeric, largely humanized antibodies. This can be advantageous in a clinical setting, where a so-called human anti-mouse antibody response (HAMA) has been observed after treatment with mAbs purified from hybridoma clones (Schroff *et al.*, 1985; Shawler *et al.*, 1985).

Furthermore, the diagnostical approach should be improved regarding the detection of surface-expressed env protein. The ICC/IHC method used so far is unsuitable for specific surface detection, since processing and fixating the tissues permeabilizes the cell membranes. In that matter, methods like flow cytometry would be a better option. Viable single-cell suspensions can be prepared from tumor tissues in a non-permeabilizing way, thereby enabling the staining of surface-presented proteins (Leelatian *et al.*, 2017). Therefore, the only limiting factor for a flow cytometry-based diagnostical approach would be a suitable HERV-H-

env-recognizing antibody. Since the pAbs tested here showed high background staining, a mAb, generated as described above, could be a possible solution to this problem.

References

Adams, J. *et al.* (1998) 'Potent and selective inhibitors of the proteasome: Dipeptidyl boronic acids', *Bioorganic & Medicinal Chemistry Letters*, 8(4), pp. 333–338. Available at: https://doi.org/10.1016/S0960-894X(98)00029-8.

Adams, J. *et al.* (1999) 'Proteasome Inhibitors: A Novel Class of Potent and Effective Antitumor Agents', *CANCER RESEARCH*, 59, pp. 2615–2622. Available at: http://aacrjournals.org/cancerres/article-pdf/59/11/2615/3244742/ch119902615p.pdf (Accessed: 25 October 2024).

Ahn, K. and Kim, H.S. (2009) 'Structural and Quantitative Expression Analyses of HERV Gene Family in Human Tissues', *Molecules and Cells*, 28(2), pp. 99–104. Available at: https://doi.org/10.1007/S10059-009-0107-Y.

Alcazer, V., Bonaventura, P. and Depil, S. (2020) 'Human Endogenous Retroviruses (HERVs): Shaping the Innate Immune Response in Cancers', *Cancers 2020, Vol. 12, Page 610*, 12(3), p. 610. Available at: https://doi.org/10.3390/CANCERS12030610.

Almagro Armenteros, J.J. *et al.* (2019) 'SignalP 5.0 improves signal peptide predictions using deep neural networks', *Nature Biotechnology*, 37(4), pp. 420–423. Available at: https://doi.org/10.1038/s41587-019-0036-z.

Altschul, S.F. *et al.* (1990) 'Basic local alignment search tool', *Journal of molecular biology*, 215(3), pp. 403–410. Available at: https://doi.org/10.1016/S0022-2836(05)80360-2.

Alves, P.M.S. *et al.* (2008) 'Identification of tumor-associated antigensby large-scale analysis of genes expressed in human colorectal cancer', *Cancer Immunity : a Journal of the Academy of Cancer Immunology*, 8, p. 11. Available at: /pmc/articles/PMC2935784/ (Accessed: 22 May 2024).

Antony, J.M. *et al.* (2006) 'Quantitative analysis of human endogenous retrovirus-W env in neuroinflammatory diseases', *AIDS research and human retroviruses*, 22(12), pp. 1253–1259. Available at: https://doi.org/10.1089/AID.2006.22.1253.

Bannert, N. and Kurth, R. (2024) 'The Evolutionary Dynamics of Human Endogenous Retroviral Families', 18, p. 24. Available at: https://doi.org/10.1146/annurev.genom.7.080505.115700.

Barnes, L.M., Bentley, C.M. and Dickson, A.J. (2003) 'Stability of protein production from recombinant mammalian cells', *Biotechnology and Bioengineering*, 81(6), pp. 631–639. Available at: https://doi.org/10.1002/BIT.10517.

Bateman, A. *et al.* (2023) 'UniProt: the Universal Protein Knowledgebase in 2023', *Nucleic Acids Research*, 51(D1), pp. D523–D531. Available at: https://doi.org/10.1093/NAR/GKAC1052.

Bermejo, A.V. *et al.* (2020) 'Cancer Associated Endogenous Retroviruses: Ideal Immune Targets for Adenovirus-Based Immunotherapy', *International Journal of Molecular Sciences*, 21(14), pp. 1–21. Available at: https://doi.org/10.3390/IJMS21144843.

Bernard, P.S. and Wittwer, C.T. (2002) 'Real-Time PCR Technology for Cancer Diagnostics', *Clinical Chemistry*, 48(8), pp. 1178–1185. Available at: https://doi.org/10.1093/CLINCHEM/48.8.1178.

Binley, J.M. *et al.* (2000) 'A recombinant human immunodeficiency virus type 1 envelope glycoprotein complex stabilized by an intermolecular disulfide bond between the gp120 and gp41 subunits is an antigenic mimic of the trimeric virion-associated structure', *Journal of virology*, 74(2), pp. 627–643. Available at: https://doi.org/10.1128/JVI.74.2.627-643.2000.

Bjerregaard, B. *et al.* (2006) 'Syncytin is involved in breast cancer-endothelial cell fusions', *Cellular and Molecular Life Sciences*, 63(16), pp. 1906–1911. Available at: https://doi.org/10.1007/S00018-006-6201-9/METRICS.

Blaise, S. *et al.* (2003) 'Genomewide screening for fusogenic human endogenous retrovirus envelopes identifies syncytin 2, a gene conserved on primate evolution', *Proceedings of the National Academy of Sciences of the United States of America*, 100(22), pp. 13013–13018. Available at: https://doi.org/10.1073/pnas.2132646100.

Blaise, S., Mangeney, M. and Heidmann, T. (2001) 'The envelope of Mason-Pfizer monkey virus has immunosuppressive properties', *Journal of General Virology*, 82, pp. 1597–1600.

Blond, J.-L. *et al.* (2000) 'An Envelope Glycoprotein of the Human Endogenous Retrovirus HERV-W Is Expressed in the Human Placenta and Fuses Cells Expressing the Type D Mammalian Retrovirus Receptor', *Journal of Virology*, 74(7), pp. 3321–3329. Available at: https://doi.org/10.1128/JVI.74.7.3321-3329.2000/FORMAT/EPUB.

References

Boiko, A.D. *et al.* (2010) 'Human Melanoma Initiating Cells Express Neural Crest Nerve Growth Factor Receptor CD271', *Nature*, 466(7302), p. 133. Available at: https://doi.org/10.1038/NATURE09161.

Bonnaud, B. *et al.* (2004) 'Evidence of selection on the domesticated ERVWE1 env retroviral element involved in placentation', *Molecular Biology and Evolution*, 21(10), pp. 1895–1901. Available at: https://doi.org/10.1093/molbev/msh206.

Bradbury, A.R.M. *et al.* (2018) 'When monoclonal antibodies are not monospecific: Hybridomas frequently express additional functional variable regions', *mAbs*, 10(4), p. 539. Available at: https://doi.org/10.1080/19420862.2018.1445456.

Bresk, C.A. *et al.* (2019) 'Induction of Tier 1 HIV Neutralizing Antibodies by Envelope Trimers Incorporated into a Replication Competent Vesicular Stomatitis Virus Vector', *Viruses 2019, Vol. 11, Page 159*, 11(2), p. 159. Available at: https://doi.org/10.3390/V11020159.

Bühring, H.J. *et al.* (2007) 'Novel markers for the prospective isolation of human MSC', *Annals of the New York Academy of Sciences*, 1106, pp. 262–271. Available at: https://doi.org/10.1196/ANNALS.1392.000.

Cancer (no date). Available at: https://www.who.int/news-room/fact-sheets/detail/cancer (Accessed: 15 December 2024).

Carroll, W.L., Mendel, E. and Levy, S. (1988) 'Hybridoma fusion cell lines contain an aberrant kappa transcript', *Molecular Immunology*, 25(10), pp. 991–995. Available at: https://doi.org/10.1016/0161-5890(88)90005-3.

Chan, C.E.Z. *et al.* (2011) 'Comparison of the efficiency of antibody selection from semi-synthetic scFv and non-immune Fab phage display libraries against protein targets for rapid development of diagnostic immunoassays', *Journal of Immunological Methods*, 373(1–2), pp. 79–88. Available at: https://doi.org/10.1016/J.JIM.2011.08.005.

Cianciolo, G.J. *et al.* (1985) 'Inhibition of Lymphocyte Proliferation by a Synthetic Peptide Homologous to Retroviral Envelope Proteins', *Science*, 230(4724), pp. 453–455. Available at: https://doi.org/10.1126/SCIENCE.2996136.

Coux, O., Tanaka, K. and Goldberg, A.L. (1996) 'Structure and functions of the 20S and 26S proteasomes', *Annual Review of Biochemistry*, 65(Volume 65, 1996), pp. 801–847. Available

at: https://doi.org/10.1146/ANNUREV.BI.65.070196.004101/CITE/REFWORKS.

Cui, L. *et al.* (2016) 'Effects of individually silenced N-glycosylation sites and non-synonymous single-nucleotide polymorphisms on the fusogenic function of human syncytin-2', *Cell Adhesion and Migration*, 10(1–2), pp. 39–55. Available at: https://doi.org/10.1080/19336918.2015.1093720.

Curtin, F. *et al.* (2015) 'Preclinical and early clinical development of GNbAC1, a humanized IgG4 monoclonal antibody targeting endogenous retroviral MSRV-Env protein', *mAbs*, 7(1), pp. 265–275. Available at: https://doi.org/10.4161/19420862.2014.985021.

Daradoumis, J. *et al.* (2023) 'An Endogenous Retrovirus Vaccine Encoding an Envelope with a Mutated Immunosuppressive Domain in Combination with Anti-PD1 Treatment Eradicates Established Tumours in Mice', *Viruses*, 15(4). Available at: https://doi.org/10.3390/V15040926.

Denner, J., Norley, S. and Kurth, R. (1994) 'The immunosuppressive peptide of HIV-1: functional domains and immune response in AIDS patients', *AIDS (London, England)*, 8(8), pp. 1063–1072. Available at: https://doi.org/10.1097/00002030-199408000-00005.

Duraiyan, J. *et al.* (2012) 'Applications of immunohistochemistry', *Journal of Pharmacy & Bioallied Sciences*, 4(Suppl 2), p. S307. Available at: https://doi.org/10.4103/0975-7406.100281.

Ellis, G.I., Sheppard, N.C. and Riley, J.L. (2021) 'Genetic engineering of T cells for immunotherapy', *Nature reviews. Genetics*, 22(7), p. 427. Available at: https://doi.org/10.1038/S41576-021-00329-9.

Grandi, N. and Tramontano, E. (2018a) 'HERV envelope proteins: Physiological role and pathogenic potential in cancer and autoimmunity', *Frontiers in Microbiology*. Frontiers Media S.A. Available at: https://doi.org/10.3389/fmicb.2018.00462.

Grandi, N. and Tramontano, E. (2018b) 'Human endogenous retroviruses are ancient acquired elements still shaping innate immune responses', *Frontiers in Immunology*, 9(SEP), p. 400848. Available at: https://doi.org/10.3389/FIMMU.2018.02039/BIBTEX.

Griffiths, D.J. (2001) 'Endogenous retroviruses in the human genome sequence', *Genome Biology*, 2(6), pp. 1–5. Available at: https://doi.org/10.1186/gb-2001-2-6-reviews1017.

References

Grigsby, R. V., Fairbairn, D. and O'Neill, K.L. (1993) 'Differential DNA damage detected in hybridomas', *Hybridoma*, 12(6), pp. 755–761. Available at: https://doi.org/10.1089/HYB.1993.12.755.

Hanahan, D. (1983) 'Studies on transformation of Escherichia coli with plasmids', *Journal of Molecular Biology*, 166(4), pp. 557–580. Available at: https://doi.org/10.1016/S0022-2836(83)80284-8.

Haraguchi, S., Good, R.A. and Day-Good, N.K. (2008) 'A potent immunosuppressive retroviral peptide: Cytokine patterns and signaling pathways', *Immunologic Research*, 41(1), pp. 46–55. Available at: https://doi.org/10.1007/S12026-007-0039-6/FIGURES/2.

Harris, J.F. and Koropatnick Pearson, J.J. (1990) 'Spontaneous and radiation-induced genetic instability of heteromyeloma hybridoma cells.', *Molecular Biology & Medicine*, 7(6), pp. 485–493. Available at: https://europepmc.org/article/med/2077350 (Accessed: 29 October 2024).

Hartung, H.P. *et al.* (2022) 'Efficacy and safety of temelimab in multiple sclerosis: Results of a randomized phase 2b and extension study', *Multiple sclerosis (Houndmills, Basingstoke, England)*, 28(3), pp. 429–440. Available at: https://doi.org/10.1177/13524585211024997.

Henzy, J.E. and Coffin, J.M. (2013) 'Betaretroviral Envelope Subunits Are Noncovalently Associated and Restricted to the Mammalian Class', *Journal of Virology*, 87(4), p. 1937. Available at: https://doi.org/10.1128/JVI.01442-12.

Hirose, Y., Takamatsu, T. and Harada, H. (1993) 'Presence of env Genes in Members of the RTVL-H Family of Human Endogenous Retrovirus-like Elements', *Virology*, 192(1), pp. 52–61. Available at: https://doi.org/10.1006/VIRO.1993.1007.

Hogan, V. and Johnson, W.E. (2023) 'Unique Structure and Distinctive Properties of the Ancient and Ubiquitous Gamma-Type Envelope Glycoprotein', *Viruses*, 15(2), p. 274. Available at: https://doi.org/10.3390/V15020274.

Hollingsworth, R.E. and Jansen, K. (2019) 'Turning the corner on therapeutic cancer vaccines', *npj Vaccines 2019 4:1*, 4(1), pp. 1–10. Available at: https://doi.org/10.1038/s41541-019-0103-y.

Imai, T. *et al.* (2013) 'CD271 Defines a Stem Cell-Like Population in Hypopharyngeal Cancer', *PLOS ONE*, 8(4), p. e62002. Available at: https://doi.org/10.1371/JOURNAL.PONE.0062002.

Izsvák, Z. *et al.* (2016) 'Pluripotency and the endogenous retrovirus HERVH: Conflict or serendipity?', *BioEssays : news and reviews in molecular, cellular and developmental biology*, 38(1), pp. 109–117. Available at: https://doi.org/10.1002/BIES.201500096.

Jern, P. *et al.* (2005) 'Sequence Variability, Gene Structure, and Expression of Full-Length Human Endogenous Retrovirus H', *Journal of Virology*, 79(10), pp. 6325–6337. Available at: https://doi.org/10.1128/jvi.79.10.6325-6337.2005.

Jern, P., Sperber, G.O. and Blomberg, J. (2004) 'Definition and variation of human endogenous retrovirus H', *Virology*, 327(1), pp. 93–110. Available at: https://doi.org/10.1016/J.VIROL.2004.06.023.

Jern, P., Sperber, G.O. and Blomberg, J. (2005) 'Use of Endogenous Retroviral Sequences (ERVs) and structural markers for retroviral phylogenetic inference and taxonomy', *Retrovirology*, 2. Available at: https://doi.org/10.1186/1742-4690-2-50.

Jogalekar, M.P. *et al.* (2022) 'CAR T-Cell-Based gene therapy for cancers: new perspectives, challenges, and clinical developments', *Frontiers in Immunology*, 13, p. 925985. Available at: https://doi.org/10.3389/FIMMU.2022.925985.

Johanning, G.L. *et al.* (2017) 'Expression of human endogenous retrovirus-K is strongly associated with the basal-like breast cancer phenotype', *Scientific Reports 2017 7:1*, 7(1), pp. 1–11. Available at: https://doi.org/10.1038/srep41960.

Kassiotis, G. and Stoye, J.P. (2017) 'Making a virtue of necessity: The pleiotropic role of human endogenous retroviruses in cancer', *Philosophical Transactions of the Royal Society B: Biological Sciences*. Available at: https://doi.org/10.1098/rstb.2016.0277.

Katzourakis, A., Rambaut, A. and Pybus, O.G. (2005) 'The evolutionary dynamics of endogenous retroviruses', *Trends in Microbiology*, 13(10), pp. 463–468. Available at: https://doi.org/10.1016/j.tim.2005.08.004.

Köhler, G. and Milstein, C. (1975) 'Continuous cultures of fused cells secreting antibody of predefined specificity', *Nature 1975 256:5517*, 256(5517), pp. 495–497. Available at: https://doi.org/10.1038/256495a0.

Kolbe, A.R. *et al.* (2020) 'Human endogenous retrovirus expression is associated with head and neck cancer and differential survival', *Viruses*, 12(9). Available at:

https://doi.org/10.3390/v12090956.

Kong, Y. *et al.* (2019) 'Transposable element expression in tumors is associated with immune infiltration and increased antigenicity', *Nature Communications*, 10(1). Available at: https://doi.org/10.1038/S41467-019-13035-2.

Kraus, B. *et al.* (2013) 'Vaccination Directed against the Human Endogenous Retrovirus-K Envelope Protein Inhibits Tumor Growth in a Murine Model System', *PLoS ONE*, 8(8). Available at: https://doi.org/10.1371/JOURNAL.PONE.0072756.

Krishnamurthy, J. *et al.* (2015) 'Genetic engineering of T cells to target HERV-K, an ancient retrovirus on melanoma', *Clinical Cancer Research*, 21(14), pp. 3241–3251. Available at: https://doi.org/10.1158/1078-0432.CCR-14-3197/86537/AM/GENETIC-ENGINEERING-OF-T-CELLS-TO-TARGET-HERV-K-AN.

Kudo-Saito, C. *et al.* (2014) 'Induction of Immunoregulatory CD271 + Cells by Metastatic Tumor Cells That Express Human Endogenous Retrovirus H', *Cancer Research*, 74(5), pp. 1361–1370. Available at: https://doi.org/10.1158/0008-5472.CAN-13-1349.

Lander, S. *et al.* (2001) 'Initial sequencing and analysis of the human genome International Human Genome Sequencing Consortium* The Sanger Centre: Beijing Genomics Institute/Human Genome Center', *NATURE*, 409. Available at: www.nature.com (Accessed: 8 October 2024).

Lavialle, C. *et al.* (2013) 'Paleovirology of "syncytins", retroviral env genes exapted for a role in placentation', *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1626). Available at: https://doi.org/10.1098/rstb.2012.0507.

Lavie, L. *et al.* (2005) 'CpG Methylation Directly Regulates Transcriptional Activity of the Human Endogenous Retrovirus Family HERV-K(HML-2)', *Journal of Virology*, 79(2), p. 876. Available at: https://doi.org/10.1128/JVI.79.2.876-883.2005.

Lee, S.H. *et al.* (2017) 'Preferential production of IgM-secreting hybridomas by immunization with DNA vaccines coding for Ebola virus glycoprotein: use of protein boosting for IgG-secreting hybridoma production', *Clinical and Experimental Vaccine Research*, 6(2), p. 135. Available at: https://doi.org/10.7774/CEVR.2017.6.2.135.

Leelatian, N. *et al.* (2017) 'Preparing Viable Single Cells from Human Tissue and Tumors for

Cytomic Analysis Current Protocols in Molecular Biology UNIT 25C.1', *Current protocols in molecular biology*, 118, p. 25C.1.1. Available at: https://doi.org/10.1002/CPMB.37.

Lessard, J.C. (2013) 'Molecular Cloning', *Methods in Enzymology*, 529, pp. 85–98. Available at: https://doi.org/10.1016/B978-0-12-418687-3.00007-0.

Li, M. *et al.* (2017) 'Downregulation of human endogenous retrovirus type K (HERV-K) viral env RNA in pancreatic cancer cells decreases cell proliferation and tumor growth', *Clinical Cancer Research*, 23(19), pp. 5892–5911. Available at: https://doi.org/10.1158/1078-0432.CCR-17-0001/87047/AM/DOWN-REGULATION-OF-HUMAN-ENDOGENOUS-RETROVIRUS.

Lu, X. *et al.* (2014) 'The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity', *nature structural & molecular biology*, 21. Available at: https://doi.org/10.1038/nsmb.2799.

Lv, H. *et al.* (2014) 'ISDTool 2.0: A computational model for predicting immunosuppressive domain of retroviruses', *Journal of Theoretical Biology*, 360, pp. 78–82. Available at: https://doi.org/10.1016/j.jtbi.2014.06.033.

Mager, D.L. *et al.* (1999) 'Endogenous Retroviruses Provide the Primary Polyadenylation Signal for Two New Human Genes (HHLA2 and HHLA3)', *Genomics*, 59(3), pp. 255–263. Available at: https://doi.org/10.1006/GENO.1999.5877.

Mager, D.L. and Henthorn, P.S. (1984) 'Identification of a retrovirus-like repetitive element in human DNA.', *Proceedings of the National Academy of Sciences*, 81(23), pp. 7510–7514. Available at: https://doi.org/10.1073/PNAS.81.23.7510.

Magiorkinis, G., Blanco-Melo, D. and Belshaw, R. (2015) 'The decline of human endogenous retroviruses: Extinction and survival', *Retrovirology*, 12(1), pp. 1–12. Available at: https://doi.org/10.1186/S12977-015-0136-X/TABLES/3.

Makkouk, A. and Weiner, G.J. (2015) 'Cancer Immunotherapy and Breaking Immune Tolerance-New Approaches to an Old Challenge', *Cancer research*, 75(1), p. 5. Available at: https://doi.org/10.1158/0008-5472.CAN-14-2538.

Maloney, D.G. *et al.* (1997) 'IDEC-C2B8 (Rituximab) Anti-CD20 Monoclonal Antibody Therapy in Patients With Relapsed Low-Grade Non-Hodgkin's Lymphoma', *Blood*, 90(6), pp. 2188–2195. Available at: https://doi.org/10.1182/BLOOD.V90.6.2188.

References

Mameli, G. *et al.* (2007) 'Brains and peripheral blood mononuclear cells of multiple sclerosis (MS) patients hyperexpress MS-associated retrovirus/HERV-W endogenous retrovirus, but not human herpesvirus 6', *Journal of General Virology*, 88(1), pp. 264–274. Available at: https://doi.org/10.1099/VIR.0.81890-0/CITE/REFWORKS.

Manca, M.A. *et al.* (2022) 'HERV-K and HERV-H Env Proteins Induce a Humoral Response in Prostate Cancer Patients', *Pathogens*, 11(1). Available at: https://doi.org/10.3390/pathogens11010095.

Mangeney, M. *et al.* (2001) 'The full-length envelope of an HERV-H human endogenous retrovirus has immunosuppressive properties', *Journal of General Virology*, 82, pp. 2515–2518.

Mangeney, M. *et al.* (2007) 'Placental syncytins: Genetic disjunction between the fusogenic and immunosuppressive activity of retroviral envelope proteins', *Proceedings of the National Academy of Sciences of the United States of America*, 104(51), pp. 20534–20539. Available at: https://doi.org/10.1073/PNAS.0707873105/SUPPL_FILE/07873FIG5.PDF.

Mangeney, M. and Heidmann, T. (1998) 'Tumor cells expressing a retroviral envelope escape immune rejection in vivo', *Proceedings of the National Academy of Sciences of the United States of America*, 95(25), pp. 14920–14925. Available at: https://doi.org/10.1073/PNAS.95.25.14920/ASSET/F033505F-3790-47F2-8ED8-6AF13E4BF6C9/ASSETS/GRAPHIC/PQ2581277005.JPEG.

Matteucci, C. *et al.* (2018) 'Human endogenous retroviruses role in cancer cell stemness', *Seminars in Cancer Biology*, 53, pp. 17–30. Available at: https://doi.org/10.1016/J.SEMCANCER.2018.10.001.

Meyer, L. *et al.* (2019) 'A simplified workflow for monoclonal antibody sequencing'. Available at: https://doi.org/10.1371/journal.pone.0218717.

Mi, S. *et al.* (2000) 'Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis', *Nature*, 403(6771), pp. 785–789. Available at: https://doi.org/10.1038/35001608.

Mullins, C.S. and Linnebacher, M. (2012) 'Endogenous retrovirus sequences as a novel class of tumor-speciWc antigens: an example of HERV-H env encoding strong CTL epitopes', *Cancer*

Immunol Immunother, 61, pp. 1093–1100. Available at: https://doi.org/10.1007/s00262-011-1183-3.

Nevalainen, T., Autio-Kimura, A. and Hurme, M. (2024) 'Human endogenous retrovirus W in multiple sclerosis: transcriptional activity is associated with decline in oligodendrocyte proportions in the white matter of the brain', *Journal of NeuroVirology*, pp. 1–13. Available at: https://doi.org/10.1007/S13365-024-01208-9/FIGURES/3.

De Parseval, N. *et al.* (2001) 'Characterization of the Three HERV-H Proviruses with an Open Envelope Reading Frame Encompassing the Immunosuppressive Domain and Evolutionary History in Primates', *Virology*, 279(2), pp. 558–569. Available at: https://doi.org/10.1006/VIRO.2000.0737.

Pérot, P. *et al.* (2015) 'Expression of young HERV-H loci in the course of colorectal carcinoma and correlation with molecular subtypes', *Oncotarget*, 6(37), p. 40095. Available at: https://doi.org/10.18632/ONCOTARGET.5539.

Rabinovich, S. *et al.* (2014) 'A Novel, Live-Attenuated Vesicular Stomatitis Virus Vector Displaying Conformationally Intact, Functional HIV-1 Envelope Trimers That Elicits Potent Cellular and Humoral Responses in Mice', *PLOS ONE*, 9(9), p. e106597. Available at: https://doi.org/10.1371/JOURNAL.PONE.0106597.

Rhyu, D.W. *et al.* (2014) 'Expression of Human Endogenous Retrovirus env Genes in the Blood of Breast Cancer Patients', *International Journal of Molecular Sciences*, 15(6), p. 9173. Available at: https://doi.org/10.3390/IJMS15069173.

Rock, K.L. *et al.* (1994) 'Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules', *Cell*, 78(5), pp. 761–771. Available at: https://doi.org/10.1016/S0092-8674(94)90462-6.

Schardt, J.S., Sivaneri, N.S. and Tessier, P.M. (2024) 'Monoclonal Antibody Generation Using Single B Cell Screening for Treating Infectious Diseases', *BioDrugs*, 38(4), pp. 477–486. Available at: https://doi.org/10.1007/S40259-024-00667-0/FIGURES/2.

Schroff, R.W. *et al.* (1985) 'Human anti-murine immunoglobulin responses in patients receiving monoclonal antibody therapy', *Cancer research*, 45(2), pp. 879–885. Available at: https://pubmed.ncbi.nlm.nih.gov/3871353/ (Accessed: 29 October 2024).

Shawler, D.L. *et al.* (1985) 'Human immune response to multiple injections of murine monoclonal IgG', *Journal of immunology (Baltimore, Md. : 1950)*, 135(2), pp. 1530–1535. Available at: https://doi.org/10.4049/jimmunol.135.2.1530.

Skandorff, I. *et al.* (2023) 'Human Ad19a/64 HERV-W Vaccines Uncover Immunosuppression Domain-Dependent T-Cell Response Differences in Inbred Mice', *International journal of molecular sciences*, 24(12), p. 9972. Available at: https://doi.org/10.3390/IJMS24129972.

Sonnhammer, E.L., von Heijne, G. and Krogh, A. (1998) 'A hidden Markov model for predicting transmembrane helices in protein sequences.', *Proceedings / ... International Conference on Intelligent Systems for Molecular Biology ; ISMB. International Conference on Intelligent Systems for Molecular Biology*, 6(1), pp. 175–182. Available at: www.aaai.org (Accessed: 28 June 2021).

Sperber, G.O. *et al.* (2007) 'Automated recognition of retroviral sequences in genomic data—RetroTector©', *Nucleic Acids Research*, 35(15), p. 4964. Available at: https://doi.org/10.1093/NAR/GKM515.

Stoye, J.P. (2001) 'Endogenous retroviruses: Still active after all these years?', *Current Biology*, 11(22), pp. R914–R916. Available at: https://doi.org/10.1016/S0960-9822(01)00553-X.

Strick, R. *et al.* (2007) 'Proliferation and cell-cell fusion of endometrial carcinoma are induced by the human endogenous retroviral Syncytin-1 and regulated by TGF-β', *Journal of Molecular Medicine*, 85(1), pp. 23–38. Available at: https://doi.org/10.1007/S00109-006-0104-Y/FIGURES/9.

Strissel, P.L. *et al.* (2012) 'Reactivation of codogenic endogenous retroviral (ERV) envelope genes in human endometrial carcinoma and prestages: Emergence of new molecular targets.', *Oncotarget*, 3(10), pp. 1204–1219. Available at: https://doi.org/10.18632/ONCOTARGET.679.

Suschak, J.J., Williams, J.A. and Schmaljohn, C.S. (2017) 'Advancements in DNA vaccine vectors, non-mechanical delivery methods, and molecular adjuvants to increase immunogenicity', *Human Vaccines & Immunotherapeutics*, 13(12), p. 2837. Available at: https://doi.org/10.1080/21645515.2017.1330236.

Sverdlov, E.D. (1998) 'Perpetually mobile footprints of ancient infections in human genome', *FEBS Letters*, 428(1–2), pp. 1–6. Available at: https://doi.org/10.1016/S0014-5793(98)00478-

5.

Tiller, T., Busse, C.E. and Wardemann, H. (2009) 'Cloning and expression of murine Ig genes from single B cells', *Journal of Immunological Methods*, 350(1–2), pp. 183–193. Available at: https://doi.org/10.1016/J.JIM.2009.08.009.

Tolosa, J.M. *et al.* (2012) 'The endogenous retroviral envelope protein syncytin-1 inhibits LPS/PHA-stimulated cytokine responses in human blood and is sorted into placental exosomes', *Placenta*, 33(11), pp. 933–941. Available at: https://doi.org/10.1016/J.PLACENTA.2012.08.004.

Vacchelli, E. *et al.* (2013) 'Trial watch: Monoclonal antibodies in cancer therapy', *Oncoimmunology*, 2(1), p. e22789. Available at: https://doi.org/10.4161/ONCI.22789.

Vargiu, L. *et al.* (2016) 'Classification and characterization of human endogenous retroviruses mosaic forms are common', *Retrovirology*, 13(1), pp. 1–29. Available at: https://doi.org/10.1186/S12977-015-0232-Y/TABLES/5.

Wang-Johanning, F. *et al.* (2006) 'Expression of multiple human endogenous retrovirus surface envelope proteins in ovarian cancer', *International Journal of Cancer*, 120(1), pp. 81–90. Available at: https://doi.org/10.1002/IJC.22256.

Wang-Johanning, F. *et al.* (2011) 'Immunotherapeutic Potential of Anti-Human Endogenous Retrovirus-K Envelope Protein Antibodies in Targeting Breast Tumors', *JNCI: Journal of the National Cancer Institute*, 104(3), pp. 189–210. Available at: https://doi.org/10.1093/JNCI/DJR540.

Wang, J. *et al.* (2014) 'Primate-specific endogenous retrovirus-driven transcription defines naive-like stem cells'. Available at: https://doi.org/10.1038/nature13804.

Wang, T. *et al.* (2007) 'Species-specific endogenous retroviruses shape the transcriptional network of the human tumor suppressor protein p53', *Proceedings of the National Academy of Sciences of the United States of America*, 104(47), pp. 18613–18618. Available at: https://doi.org/10.1073/PNAS.0703637104/SUPPL_FILE/IMAGE1513.JPG.

Wentzensen, N. *et al.* (2004) 'Identification of differentially expressed genes in colorectal adenoma compared to normal tissue by suppression subtractive hybridization.', *International journal of oncology*, 24(4), pp. 987–994. Available at:

https://doi.org/10.3892/IJO.24.4.987/HTML.

Wentzensen, N. *et al.* (2007) 'Expression of an endogenous retroviral sequence from the HERV-H group in gastrointestinal cancers', *International Journal of Cancer*, 121(7), pp. 1417–1423. Available at: https://doi.org/10.1002/ijc.22826.

Wilkinson, D.A., Mager, D.L. and Leong, J.-A.C. (1994) 'Endogenous Human Retroviruses', *The Retroviridae*, pp. 465–535. Available at: https://doi.org/10.1007/978-1-4899-1730-0_9.

Wu, S. *et al.* (2016) 'A computational model for predicting fusion peptide of retroviruses', *Computational Biology and Chemistry*, 61, pp. 245–250. Available at: https://doi.org/10.1016/j.compbiolchem.2016.02.013.

Yang, X. *et al.* (2013) 'One cell, multiple roles: contribution of mesenchymal stem cells to tumor development in tumor microenvironment', *Cell & Bioscience 2013 3:1*, 3(1), pp. 1–10. Available at: https://doi.org/10.1186/2045-3701-3-5.

Yi, J.M., Kim, H.M. and Kim, H.S. (2005) 'Human endogenous retrovirus HERV-H family in human tissues and cancer cells: Expression, identification, and phylogeny', *Cancer Letters*, 231(2), pp. 228–239. Available at: https://doi.org/10.1016/j.canlet.2005.02.001.

Zahavi, D. and Weiner, L. (2020) 'Monoclonal Antibodies in Cancer Therapy', *Antibodies*, 9(3), p. 34. Available at: https://doi.org/10.3390/ANTIB9030034.

Zhou, F. *et al.* (2015) 'Chimeric antigen receptor T cells targeting HERV-K inhibit breast cancer and its metastasis through downregulation of Ras', *Oncoimmunology*, 4(11). Available at: https://doi.org/10.1080/2162402X.2015.1047582.

Zhou, F. *et al.* (2016) 'Activation of HERV-K Env protein is essential for tumorigenesis and metastasis of breast cancer cells', *Oncotarget*, 7(51), pp. 84093–84117. Available at: https://doi.org/10.18632/ONCOTARGET.11455.

Appendix

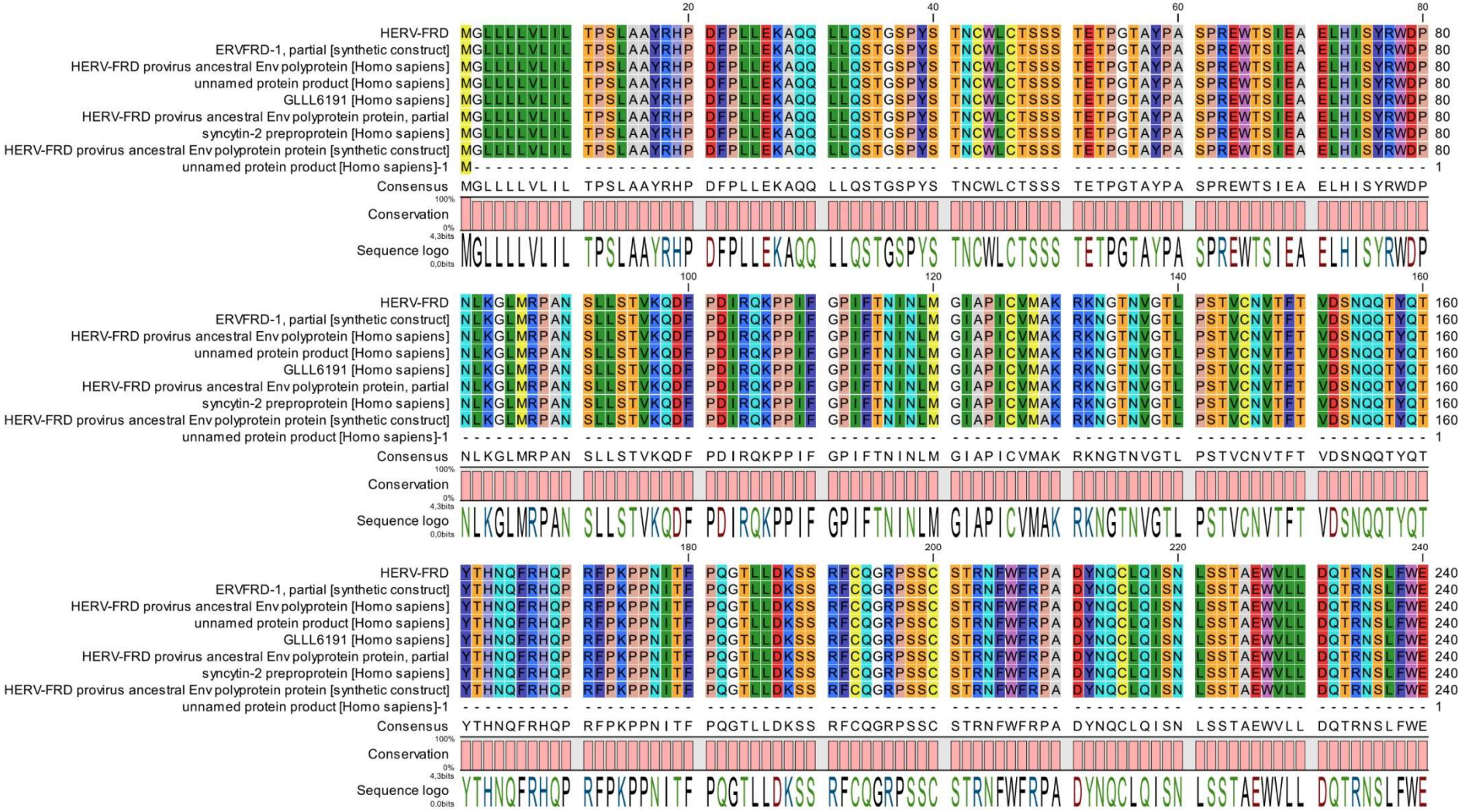
Table S 1: Oligonucleotides used for cloning and sequencing.

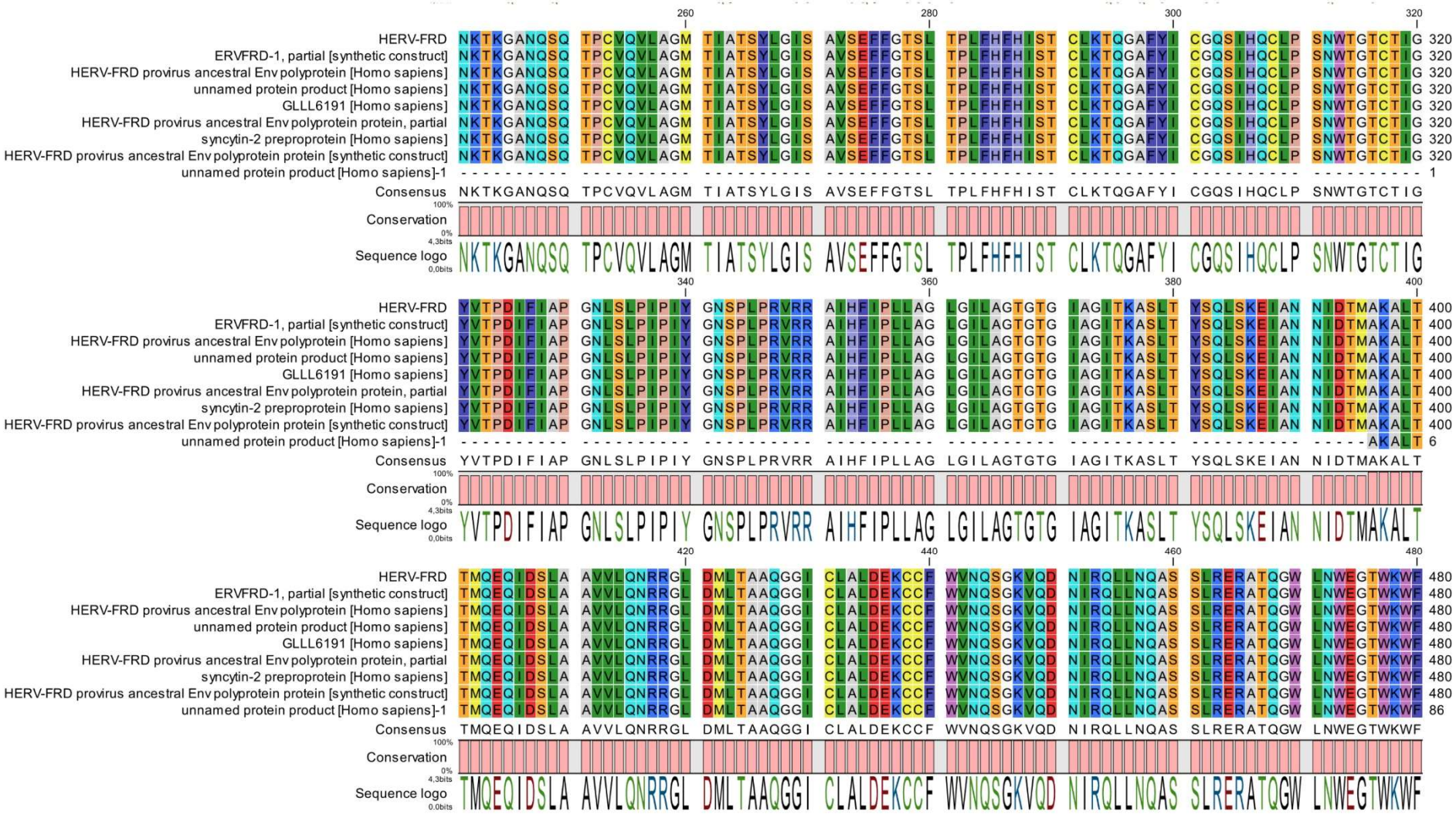
Name	Sequence
2 Con HERV-H fwd	CTGGCTTTGCATTTCCC
3 Con HERV-H rev	CTGGGAGGAGAGTCTG
3'ConHERV-H rev	CGATGTTTCTCAGGGCTG
4 Con HERV-H fwd	GCAGTAATTATTGCTTAGGAAGAC
5 HERV Con fwd	CTTTCAGCACTCCTTCTC
5 HERV con rev	GTCAACTTGGGCCTG
7 Con HERV-H fwd	CTTGCCAACCAAGCAAG
Con HERV rev	GTTAAAGGAAGTTCGGAGG
ConHERV-H-fwd	GATATCTCTTGGTGCTATCCC
ConHERV-H-fwd2	CTACTTATAAATGCCCTACTCTTG
ERV3.1-192C-fwd	GATCCCTCTGGAACCCGACTGCAAGACCAGCACCTGTAACAGCG
ERV3.1-192C-rev	CGCTGTTACAGGTGCTGGTCTTGCAGTCGGGTTCCAGAGGGATC
ERV3.1-BsaI-fwd	GATCGGTCTCGCAGCGCCGCCGAACCTTGGGAAGGCTGCCTG
ERV3.1-Env-VSV-fwd	GCCTCGGCGGCTTTAAGACAGGTGGCGGAGGATCTGGC
ERV3.1-Env-VSV-rev	GCCAGATCCTCCGCCACCTGTCTTAAAGCCGCCGAGGC
ERV3.1-KpnI-fwd	GATCGGTACCGCCACCATGCTGG
ERV3.1-KpnI-rev	GATCGGTACCTCAGGCGTAATCTGGCACGTC
ERV3.1-NotI-rev	GATCGCGGCCGCTCAGGCGTAATCTGGCACGTC
ERV-int-fwd	GTCCACCAACCAGCAAAG
ERV-int-rev	TAGGGGCCTGCCATGTGTT
FRD-BsaI-fwd	GATCGGTCTCGCAGCGCCGCCTACAGACACCCTGATTTCCCACTG
FRD-Env-VSV-fwd	GAATTGGGAAGGCACCTGGAAGGGTGGCGGAGGATCTGGC
FRD-Env-VSV-rev	GCCAGATCCTCCGCCACCCTTCCAGGTGCCTTCCCAATTC
FRD-int-fwd	AGAGAAAGAACGGCACCAAC
FRD-int-rev	CAGGTAGCTTGTGGCGATT
FRD-KpnI-fwd	GATCGGTACCGCCACCATGGGATTG
FRD-KpnI-rev	GATCGGTACCTCAGGCATAATCGGGCACATCG
FRD-N133Q-fwd	CTGCGTGATGGCCAAGAGAAAGcagGGCACCAACGTGGGCACCC
FRD-N133Q-rev	GGGTGCCCACGTTGGTGCCctgCTTTCTCTTGGCCATCACGCAG
FRD-N312Q-fwd	CCACCAGTGTCTGCCCTCTcagTGGACCGGCACCTGTACCATC
FRD-N312Q-rev	GATGGTACAGGTGCCGGTCCActgAGAGGGCAGACACTGGTGG
FRD-N443Q-fwd	GAGAAGTGCTGCTTCTGGGTCcagCAGTCTGGCAAGGTGCAGGAC
FRD-N443Q-rev	GTCCTGCACCTTGCCAGACTGctgGACCCAGAAGCAGCACTTCTC
FRD-NotI-rev	GATCGCGGCCGCTCAGGCATAATCGGGCACATCG
FRD-tpA-fwd	GTTCGTGTCCCCCAGCGCCGCCTACAGACACCCTGATTTCC
FRD-tpA-rev	ggtttaaacgggcccggtaccTCAGGCATAATCGGGCAC
FRD-VSV-HA-NotI-rev	AGATCTGCGGCCGCTCAGGCATAATCGGGCACATCGTAGGGGTAGCTGCCGC TGCCCTTGCCCAGCCGGTTCATCTC
GAPDH-fwd	CAAAGTTGTCATGGATGACC
GAPDH-rev	CCATGGAGAAGGCTGGGG

HA-Oligo-fwd	GATCCGGCAGCTACCCCTACGATGTGCCCGATTATGCCGGCAGCGGCAGCGGCAGCC
H-BamHI-rev	GATCGGATCCAGACATCCAGTTGCTCAGAGC
H-BsaI-fwd	GATCGGTCTCGCAGCGCCGCCAGCTACCTGCACCACACCATC
Henv59-HA-NotI-rev	GATCGCGGCCGCTCAAGCGTAGTCGGGCACATCGTAAGGGTAGCTGCCAGATCCAATGCTGTGGTTGGTGATGGCC
Henv59-int-rev	AGAGATGTCGGTGATGCTGG
Henv59-KpnI-fwd	GATCGGTACCGCCACCATGATTCTTGC
Henv60-HA-NotI-rev	GATCGCGGCCGCTCAAGCGTAGTCGGGCACATCGTAAGGGTAGCTGCCAGATCCAATGCTGTGATCGGTGATGGCC
Henv62-HA-NotI-rev	GATCGCGGCCGCTCAAGCGTAGTCGGGCACATCGTAAGGGTAGCTGCCAGATCCAGCAGATGGCAGGTCCTGTG
H-Env-VSV-fwd	GGGCTCTGAGCAACTGGATGTCTGGTGGCGGAGGATCTGGC
H-Env-VSV-rev	GCCAGATCCTCCGCCACCAGACATCCAGTTGCTCAGAGCCC
HERV-tpa-KpnI-fwd	GATCGGTACCGCCACCATGGACGCCATGAAG
H-HA-GS-XhoI-rev	GATCCTCGAGGCTGCCGCTGCCGCTGCCGGCATAATCGGGCACATCGTAGGGGTAGCTGCCGCTGCCAGACATCCAGTTGCTCAGAGC
H-HindIII-fwd	GATCAAGCTTGCCACCATGATCTTTGCCGG
H-HindIII-short-fwd	GATCAAGCTTGCCACCATGATCTTTG
H-int-fwd	ACCACACAGACAACAATCCC
H-int-rev	GCCACACAGAAAGAACAGG
H-ISDmut-fwd	GGATCTGCTGACCGCCGAGagaGGCGGCCTGTGTATCTTCC
H-ISDmut-rev	GGAAGATACACAGGCCGCCtctCTCGGCGGTCAGCAGATCC
H-KpnI-fwd	GATCGGTACCGCCACCATGATCTTTGC
H-KpnI-rev	GATCGGTACCTCAAGCGTAGTCGGGCACATC
H-NotI-rev	GATCGCGGCCGCTCAGGCATAATCGGGCACATCG
H-SU-BamHI-rev	GATCGGATCCTGTAGGGGTCATCAGAGGCAC
H-SU-XhoI-rev	GATCCTCGAGTGTAGGGGTCATCAGAGGCAC
H-VSV-HA-NotI-rev	AGATCTGCGGCCGCTCAAGCGTAGTCGGGCACATCGTAAGGGTAGCTGCCAGATCCCTTGCCCAGCCGGTTCATCTC
H-XhoI-rev	GATCCTCGAGAGACATCCAGTTGCTCAGAGCC
ISPCR	AAGCAGTGGTATCAACGCAGAG
JM02_mod_rev	TGTGGGAACCTAGAGCGGGAGGCGCTTGGCAGATCCTGTG
JM02_rev	TGTGGGAACCTAGAGCGGGA
JM12_fwd	GTCGGTTTAGGACTTTCTGC
JM12_mod_fwd	GTCGGACTGGGCCTGTCTGC
K-Env-BsaI-fwd	GATCGGTCTCGCAGCGCCGCCAATTACACATACTGGGCCTACG
K-Env-HA-NotI-rev	AGATCTGCGGCCGCTCAGGCATAATCGGGCACATCGTAGGGGTAGCTGCCGCTGCCCACGGACACGGTCACGATCTG
K-Env-KpnI-rev	GATCGGTACCTCACACGGACACGGTCACGATC
K-Env-VSV-fwd	CCTGTGACCTGGGTCAAGACCGGTGGCGGAGGATCTGGC
K-Env-VSV-rev	GCCAGATCCTCCGCCACCGGTCTTGACCCAGGTCACAGG
K-int-fwd	GAGGAAGAGGGCATGATGA
K-int-rev	TCTGTTCAGCACGCCCTT
K-tpA-fwd	GTTCGTGTCCCCCAGCGCCGCCAATTACACATACTGGGCCTAC
K-tpA-rev	ggtttaaacgggcccggtaccTCACACGGACACGGTC
mIGHG PCR	GGGATCCAGAGTTCCAGGTC

mIGHG RT	AGCTGGGAAGGTGTGCACAC
mIGK PCR	ACATTGATGTCTTTGGGGTAGAAG
mIGK RT	TTGTCGTTCACTGCCATCAATC
mIGL PCR	ATCGTACACACCAGTGTGGC
mIGL RT	GGGGTACCATCTACCTTCCAG
pJET-fwd	CGACTCACTATAGGGAGAGCGGC
pJET-rev	AAGAACATCGATTTTCCATGGCAG
Seq HERV-H 62	CTTTACGAAGTCACCCC
solH-HA-NotI-rev	GATCGCGGCCGCTCAGGCATAATCGGGCACATC
W-Env-BsaI-fwd	GATCGGTCTCGCAGCGCCGCCGCACCACCACCTTGCAGGTG
W-Env-HA-NotI-rev	AGATCTGCGGCCGCTCAGGCATAATCGGGCACATCGTAGGGGTAGCTGCCGC TGCCACTTGAGCCAGCAGAGTTTGGTC
W-Env-KpnI-fwd	GATCGGTACCGCCACCATGGCCCTGCCATATC
W-Env-KpnI-rev	GATCGGTACCTCAACTTGAGCCAGCAGAGTTTGG
W-Env-VSV-fwd	GATTCTGCCCTTCCTGGGACCTGGTGGCGGAGGATCTGGC
W-Env-VSV-rev	GCCAGATCCTCCGCCACCAGGTCCCAGGAAGGGCAGAATC
W-int-fwd	AGACACTGCGGACACACAC
W-int-rev	GGAGGCACCAGAAAAGACAGGA
W-N409Q-fwd	GCGAGGAGTGCTGTTACTATGTGcagCAGAGCGGCATCGTGACAG
W-N409Q-rev	CTGTCACGATGCCGCTCTGctgCACATAGTAACAGCACTCCTCGC
pC_CMV-fwd (1A3)	GTAGGCGTGTACGGTGGGAGG
pcDNA3.1seq BGHpA rev (10B4)	CCCCCAGAATAGAATGACACC

Appendix


HERV-FRD
ERVFRD-1, partial [synthetic construct]
HERV-FRD provirus ancestral Env polyprotein [Homo sapiens]
unnamed protein product [Homo sapiens]
GLLL6191 [Homo sapiens]
HERV-FRD provirus ancestral Env polyprotein protein, partial
syncytin-2 preproprotein [Homo sapiens]
HERV-FRD provirus ancestral Env polyprotein protein [synthetic construct]
unnamed protein product [Homo sapiens]-1
Consensus
Conservation
Sequence logo
MGLLLLVLIL TPSLAAYRHP DFPLLEKAQQ LLQSTGSPYS TNCWLCTSSS TETPGTAYPA SPREWTSIEA ELHISYRWDP
NLKGLMRPAN SLLSTVKQDF PDIRQKPPIF GPIFTNINLM GIAPICVMAK RKNGTNVGTL PSTVCNVTFT VDSNQQTYQT
YTHNQFRHQP RFPKPPNITF PQGTLLDKSS RFCQGRPSSC STRNFWFRPA DYNQCLQISN LSSTAEWVLL DQTRNSLFWE


HERV-FRD
ERVFRD-1, partial [synthetic construct]
HERV-FRD provirus ancestral Env polyprotein [Homo sapiens]
unnamed protein product [Homo sapiens]
GLLL6191 [Homo sapiens]
HERV-FRD provirus ancestral Env polyprotein protein, partial
syncytin-2 preproprotein [Homo sapiens]
HERV-FRD provirus ancestral Env polyprotein protein [synthetic construct]
unnamed protein product [Homo sapiens]-1
Consensus
Conservation
Sequence logo
NKTKGANQSQ TPCVQVLAGM TIATSYLGIS AVSEFFGTSL TPLFHFHIST CLKTQGAFYI CGQSIHQCLP SNWTGTCTIG
YVTPDIFIAP GNLSLPIPIY GNSPLPRVRR AIHFIPLLAG LGILAGTGTG IAGITKASLT YSQLSKEIAN NIDTMAKALT
TMQEQIDSLA AVVLQNRRGL DMLTAAQGGI CLALDEKCCF WVNQSGKVQD NIRQLLNQAS SLRERATQGW LNWEGTWKWF

Appendix

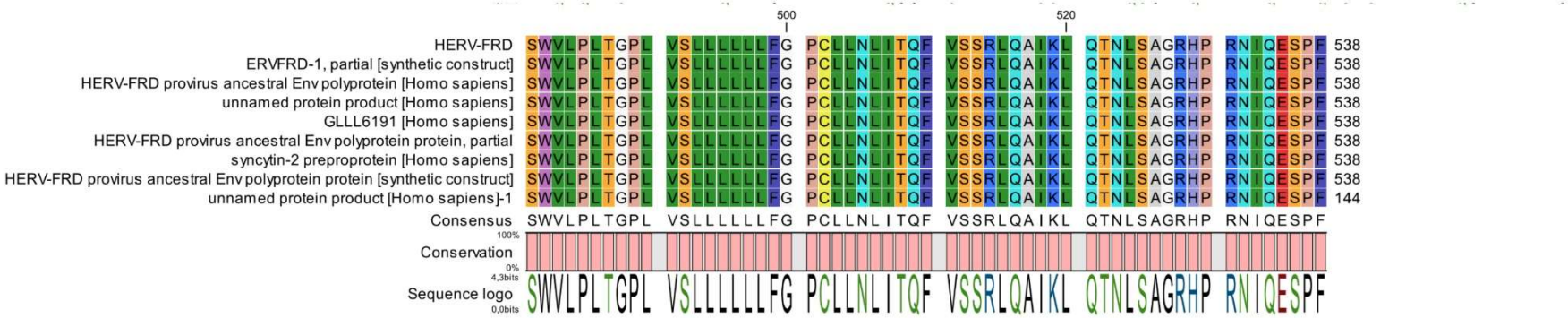
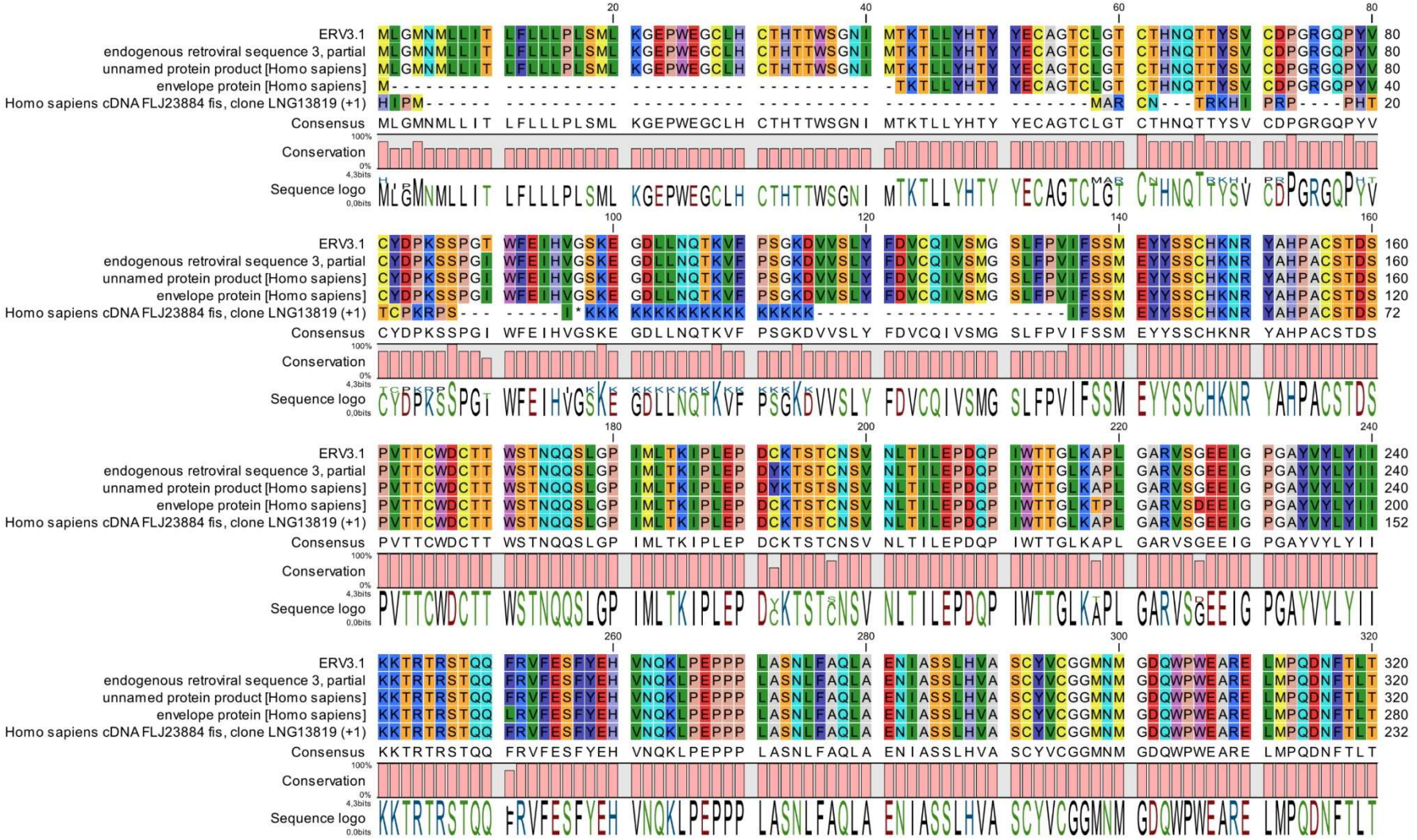
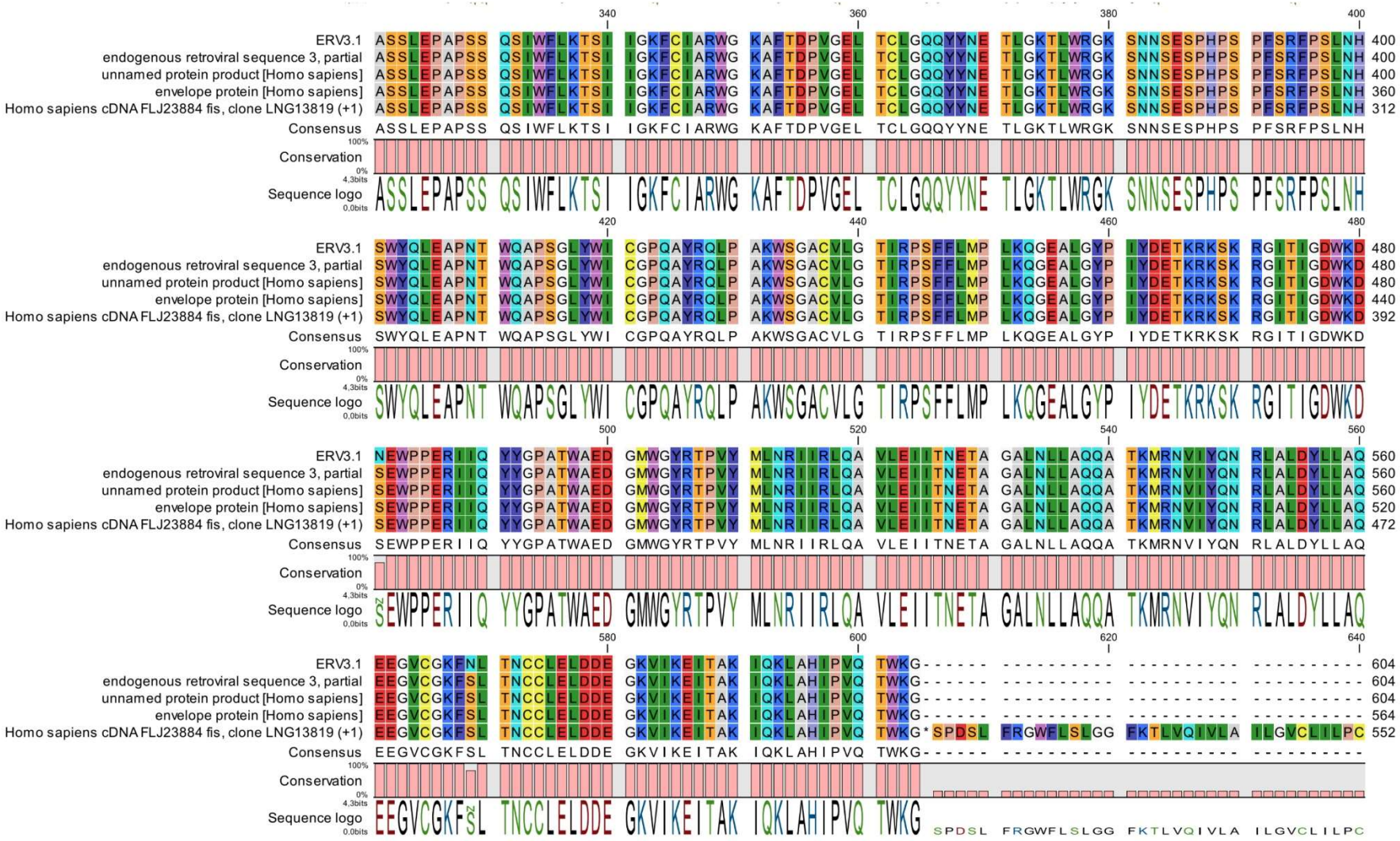


Figure S 1: Alignment of all HERV-FRD env sequences obtained from protein and nucleotides databases.

HERV-FRD env sequences were obtained from the UniProtKB database and a subsequent tblastn search. All resulting protein sequences were aligned using the CLC Main Workbench 22 alignment tool. The pink bars below the sequences indicate the level of sequence conservation.


ERV3.1
endogenous retroviral sequence 3, partial
unnamed protein product [Homo sapiens]
envelope protein [Homo sapiens]
Homo sapiens cDNA FLJ23884 fis, clone LNG13819 (+1)
Consensus
Conservation
Sequence logo

Appendix


ERV3.1
endogenous retroviral sequence 3, partial
unnamed protein product [Homo sapiens]
envelope protein [Homo sapiens]
Homo sapiens cDNA FLJ23884 fis, clone LNG13819 (+1)
Consensus
Conservation
Sequence logo

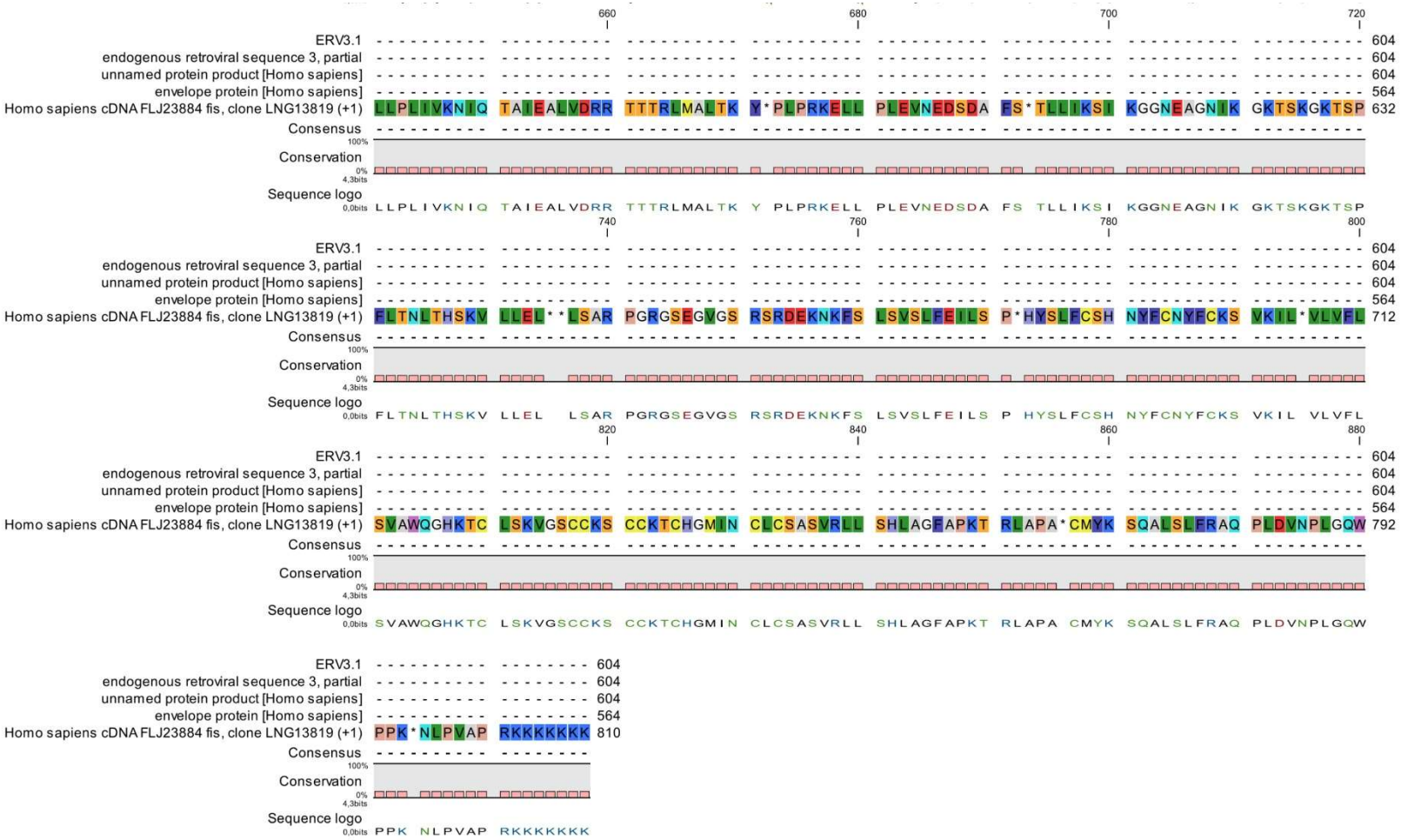
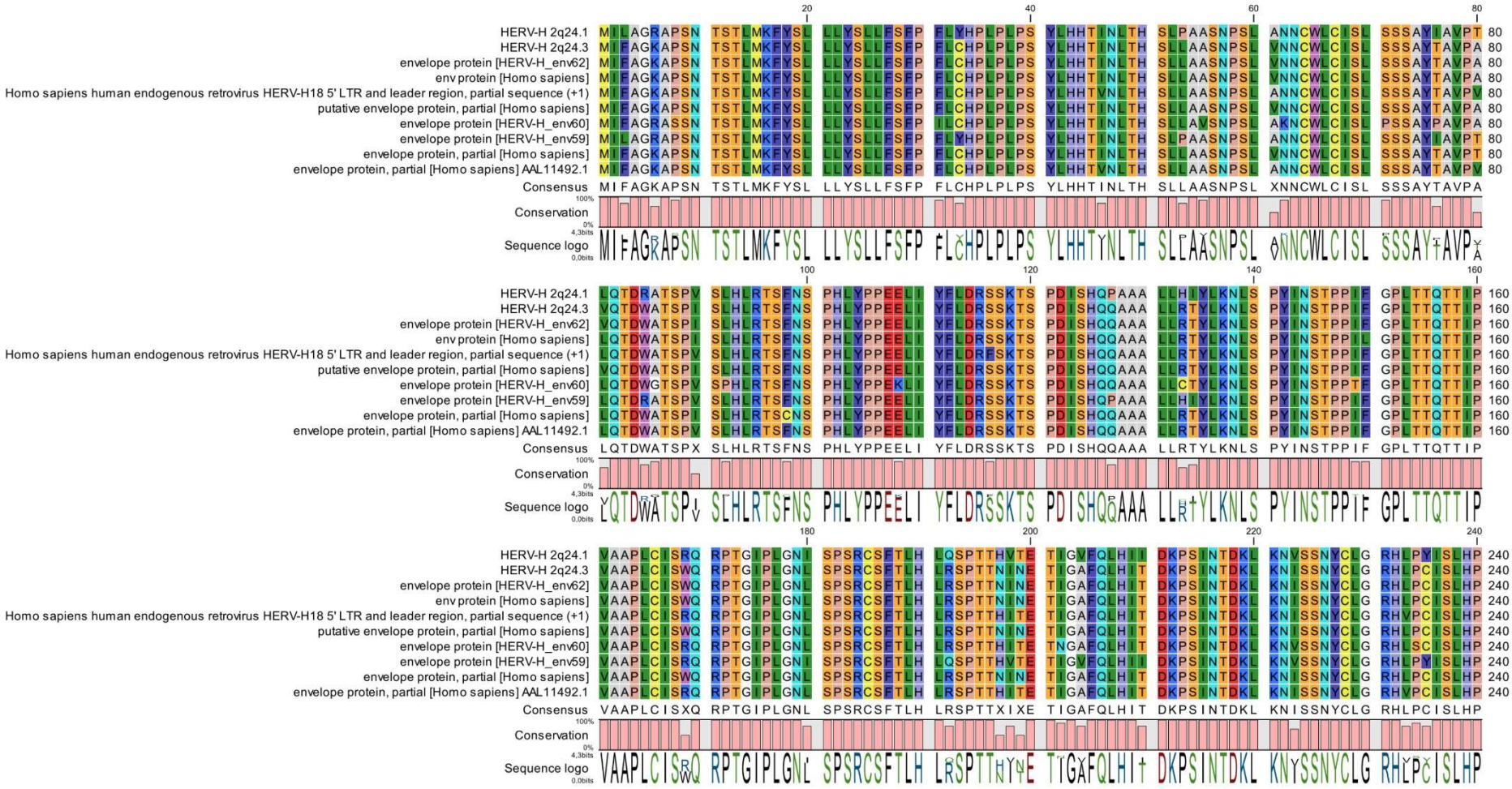
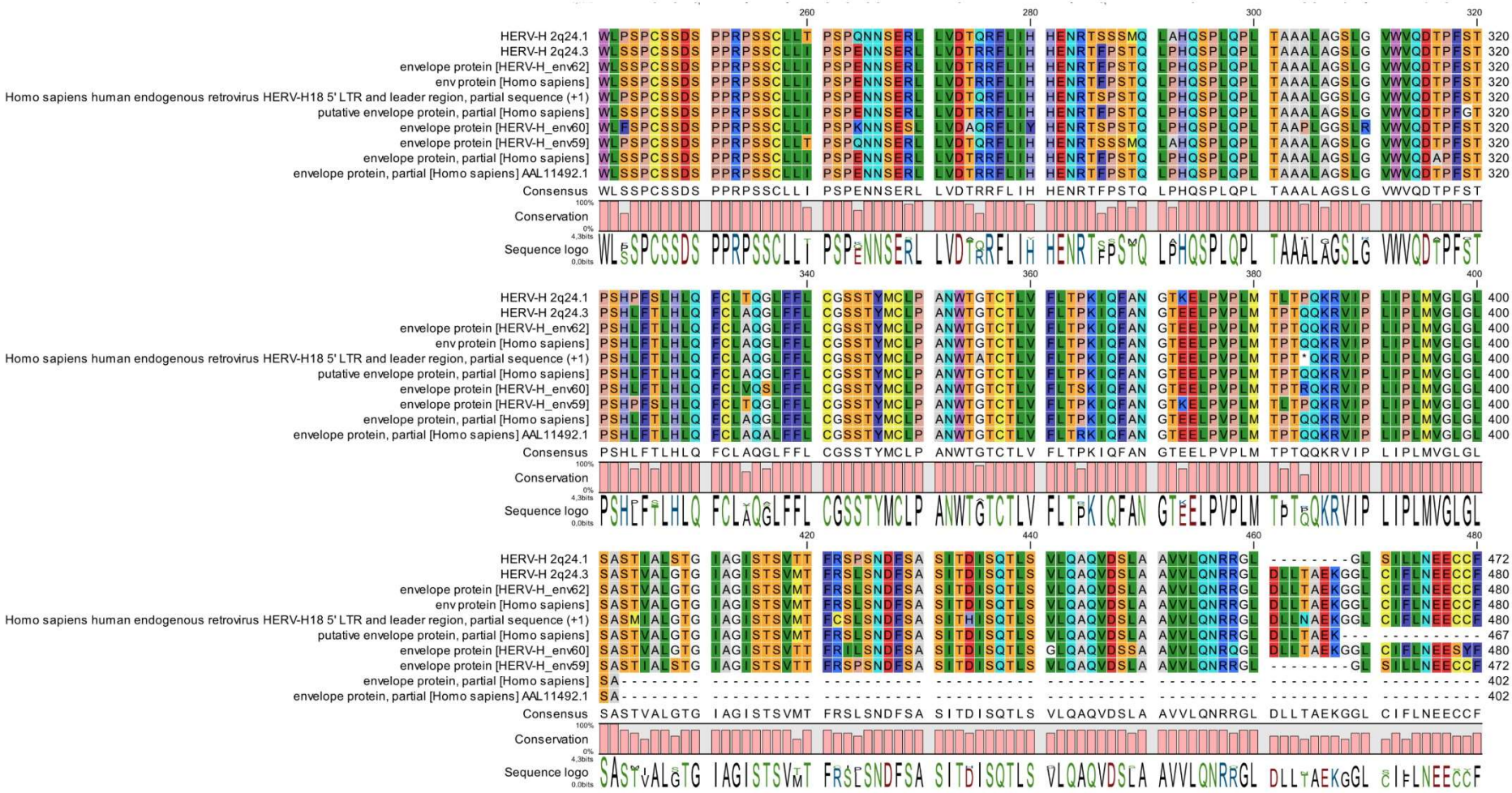


Figure S 2: Alignment of all ERV3.1 env sequences obtained from protein and nucleotides databases.

ERV3.1 env sequences were obtained from the UniProtKB database and a subsequent tblastn search. All resulting protein sequences were aligned using the CLC Main Workbench 22 alignment tool. The pink bars below the sequences indicate the level of sequence conservation.

Appendix


HERV-H 2q24.1
HERV-H 2q24.3
envelope protein [HERV-H_env62]
env protein [Homo sapiens]
Homo sapiens human endogenous retrovirus HERV-H18 5' LTR and leader region, partial sequence (+1)
putative envelope protein, partial [Homo sapiens]
envelope protein [HERV-H_env60]
envelope protein [HERV-H_env59]
envelope protein, partial [Homo sapiens]
envelope protein, partial [Homo sapiens] AAL11492.1
Consensus
Conservation
Sequence logo


260
280
300
320
HERV-H 2q24.1
HERV-H 2q24.3
envelope protein [HERV-H_env62]
env protein [Homo sapiens]
Homo sapiens human endogenous retrovirus HERV-H18 5' LTR and leader region, partial sequence (+1)
putative envelope protein, partial [Homo sapiens]
envelope protein [HERV-H_env60]
envelope protein [HERV-H_env59]
envelope protein, partial [Homo sapiens]
envelope protein, partial [Homo sapiens] AAL11492.1
Consensus
Conservation
Sequence logo
340
360
380
400
HERV-H 2q24.1
HERV-H 2q24.3
envelope protein [HERV-H_env62]
env protein [Homo sapiens]
Homo sapiens human endogenous retrovirus HERV-H18 5' LTR and leader region, partial sequence (+1)
putative envelope protein, partial [Homo sapiens]
envelope protein [HERV-H_env60]
envelope protein [HERV-H_env59]
envelope protein, partial [Homo sapiens]
envelope protein, partial [Homo sapiens] AAL11492.1
Consensus
Conservation
Sequence logo
420
440
460
480
HERV-H 2q24.1
HERV-H 2q24.3
envelope protein [HERV-H_env62]
env protein [Homo sapiens]
Homo sapiens human endogenous retrovirus HERV-H18 5' LTR and leader region, partial sequence (+1)
putative envelope protein, partial [Homo sapiens]
envelope protein [HERV-H_env60]
envelope protein [HERV-H_env59]
envelope protein, partial [Homo sapiens]
envelope protein, partial [Homo sapiens] AAL11492.1
Consensus
Conservation
Sequence logo

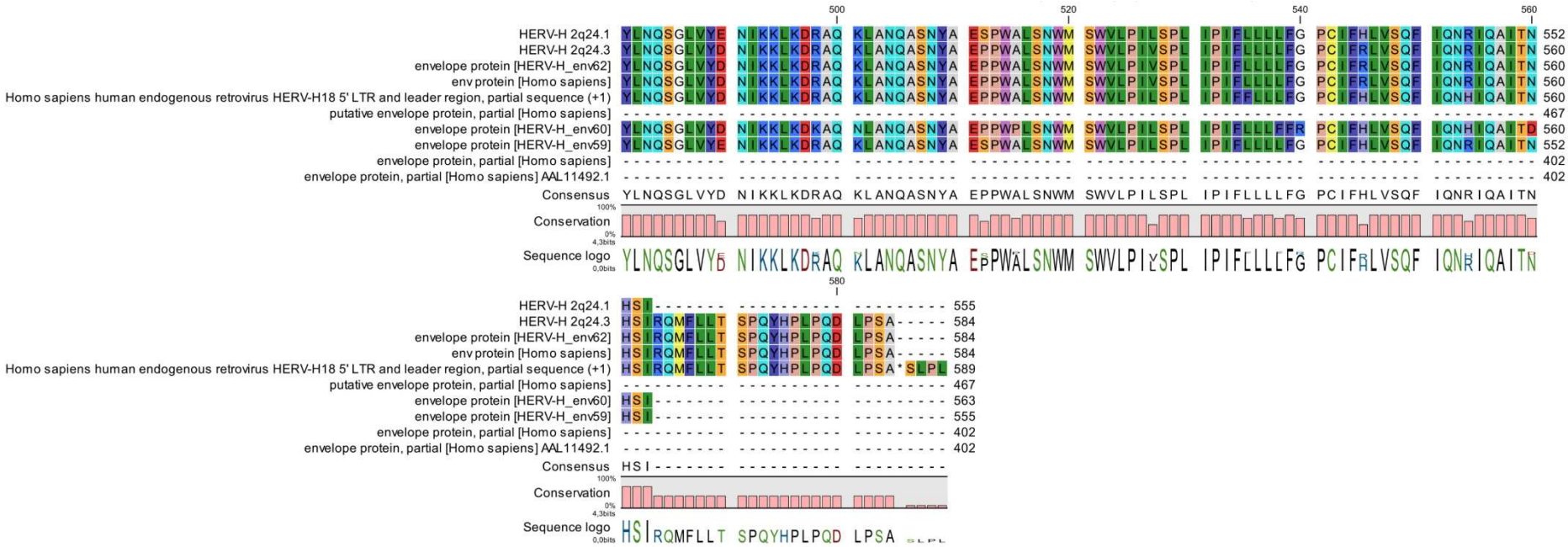


Figure S 3: Alignment of all HERV-H env sequences obtained from protein and nucleotides databases.

HERV-H env sequences were obtained from the UniProtKB databases and a subsequent tblastn search. All resulting protein sequences were aligned using the CLC Main Workbench 22 alignment tool. The pink bars below the sequences indicate the level of sequence conservation.

Figure S 4: Env protein sequence of HERV-FRD.

Amino acid sequence of the HERV-FRD env protein. SP: signal peptide; Surface SU: surface subunit; CS: cleavage site; ISD: immunosuppressive domain; Transmembrane SU: transmembrane subunit; TM: transmembrane domain; (GS)2: 2x glycine-serine linker; HA tag: hemagglutinin tag.

Figure S 5: Env protein sequence of HERV-H.

Amino acid sequence of the HERV-H env protein. SP: signal peptide; Surface SU: surface subunit; CS: cleavage site; ISD: immunosuppressive domain; Transmembrane SU: transmembrane subunit; TM: transmembrane domain; (GS)2: 2x glycine-serine linker; HA tag: hemagglutinin tag.

Figure S 6: Env protein sequence of ERV3.1_192C.

Amino acid sequence of the ERV3.1_192C env protein. SP: signal peptide; Surface SU: surface subunit; CS: cleavage site; ISD: immunosuppressive domain; Transmembrane SU: transmembrane subunit; TM: transmembrane domain; (GS)2: 2x glycine-serine linker; HA tag: hemagglutinin tag. The red triangle marks the amino acid position 192 where ERV31. 192C differs from ERV3.1_192Y (not shown) and the dark red triangle marks the alanine that was manually added instead of a stop codon.

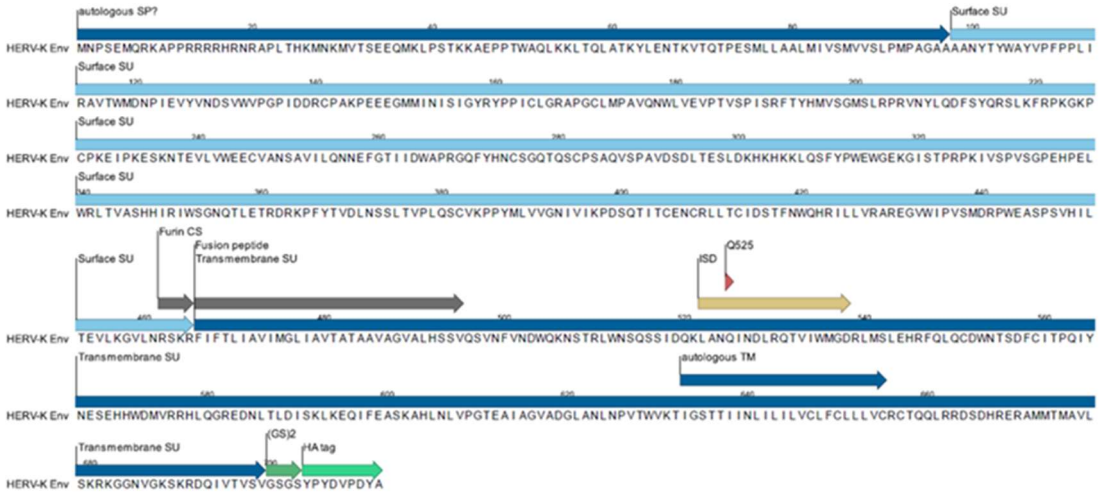


Figure S 7: Env protein sequence of HERV-K.

Amino acid sequence of the HERV-K env protein. SP: signal peptide; Surface SU: surface subunit; CS: cleavage site; ISD: immunosuppressive domain; Transmembrane SU: transmembrane subunit; TM: transmembrane domain; (GS)2: 2x glycine-serine linker; HA tag: hemagglutinin tag. The red triangle indicates the mutation site within the ISD.

Figure S 8: Env protein sequence of HERV-W.

Amino acid sequence of the HERV-W env protein. SP: signal peptide; Surface SU: surface subunit; CS: cleavage site; ISD: immunosuppressive domain; Transmembrane SU: transmembrane subunit; TM: transmembrane domain; (GS)2: 2x glycine-serine linker; HA tag: hemagglutinin tag.

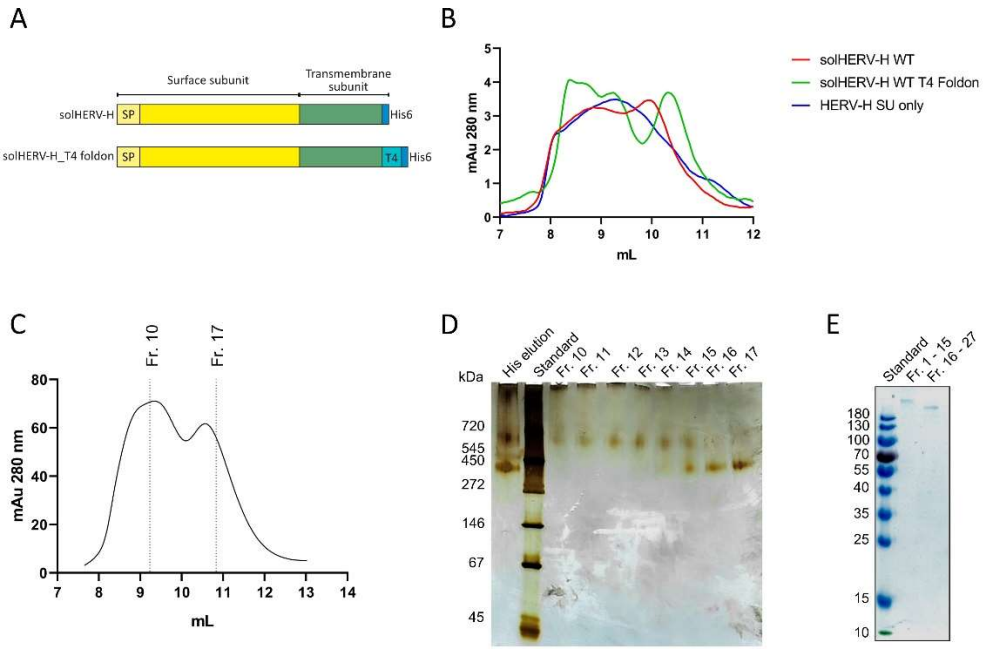


Figure S 9: Characterization of the solHERV-H env protein stabilized by a C-terminal T4 Foldon domain.

The solHERV-H env protein was stabilized in a potential trimeric conformation by adding the T4 bacteriophage fibritin foldon domain to its C-terminus. The recombinant protein was purified from transiently transfected Exppi293 cells via IMAC. A: Schematic overview of the solHERV-H_T4 foldon protein compared to the solHERV-H WT env. SP: signal peptide; T4: T4 bacteriophage fibritin foldon domain; His6: 6x Histidine-tag. B: The IMAC purified protein was analyzed in an analytical SEC run (green) and the SEC profile was compared to the solHERV-H WT (red) and the SUonly (blue) protein. C: Preparative SEC run of the IMAC purified stabilized protein. The marked range from fraction 10-17 was analyzed further in a Blue Native PAGE. D: Silver-stained Blue Native PAGE of fractions 10-17 from the preparative SEC run. As molecular weight marker the Serva Native Marker was used. E: Fractions 1-15 and 16-27 from the preparative SEC run were pooled and 3 µg of each fraction pool was loaded onto a non-reducing SDS gel. As molecular weight marker the PageRuler Prestained Protein Ladder was used. The data presented in this figure were collected by Jan Grebner during his masters' thesis under my experimental supervision.

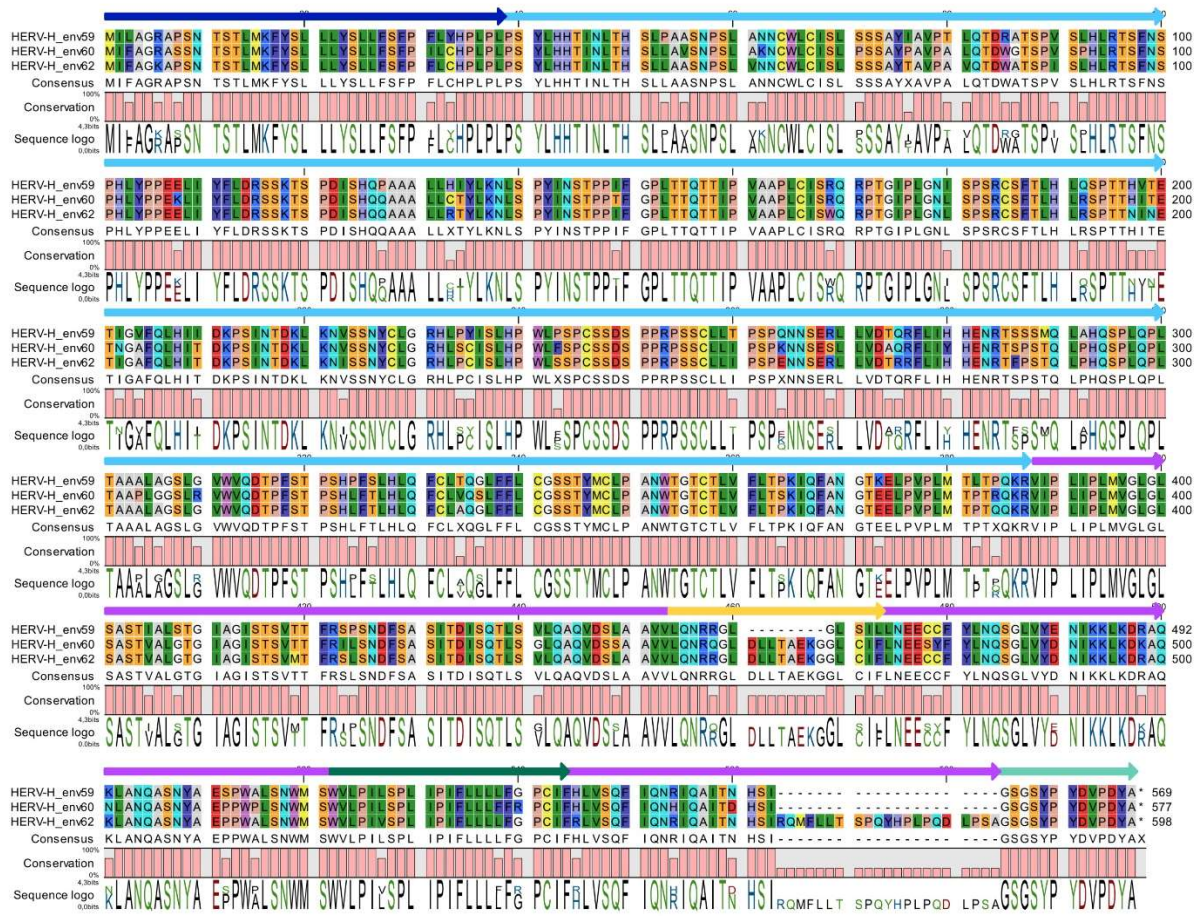


Figure S 10: Sequence alignment of the three HERV-H env protein sequences.

Alignment of the amino acid sequences of HERV-H/env62, HERV-H/env60 and HERV-H/env59 performed using the alignment tool of the CLC Main Workbench 22. The pink bars show the level of conservation between all three sequences. The arrows mark the protein subunits and regions (dark blue: signal peptide; light blue: surface subunit; lilac: transmembrane subunit; yellow: ISD; dark green: transmembrane-anchor domain; teal: glycine-serine linker and HA-tag).

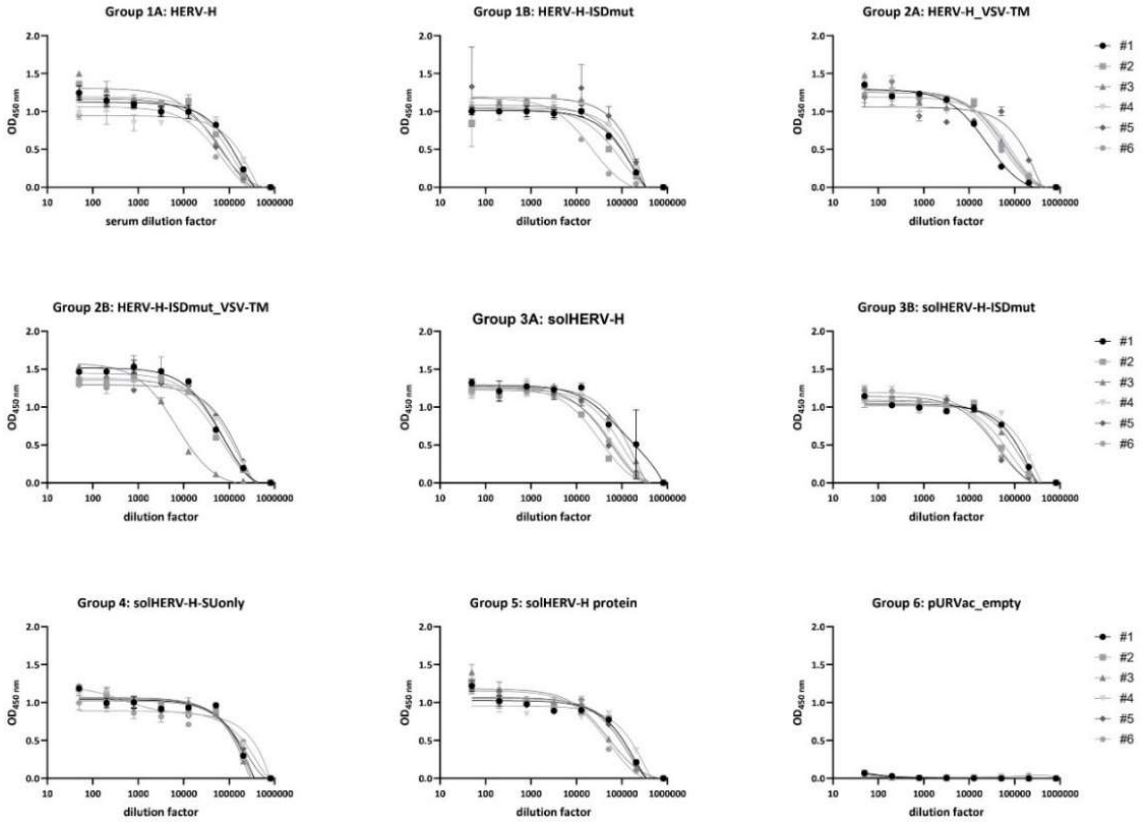


Figure S 11: Serum titration ELISA.

BALB/c mice were immunized, and serum samples were collected as described in Figure 19. The extracted sera were analysed for HERV-H env binding antibodies via ELISA by titrating the serum in a 4-fold serial dilution starting at 1:50 against the recombinant solHERV-H env protein. Binding was detected via an HRP-labeled secondary anti-mouse Ig antibody. Titration curves for all 6 individual mice per immunization group (#1-#6) are shown.

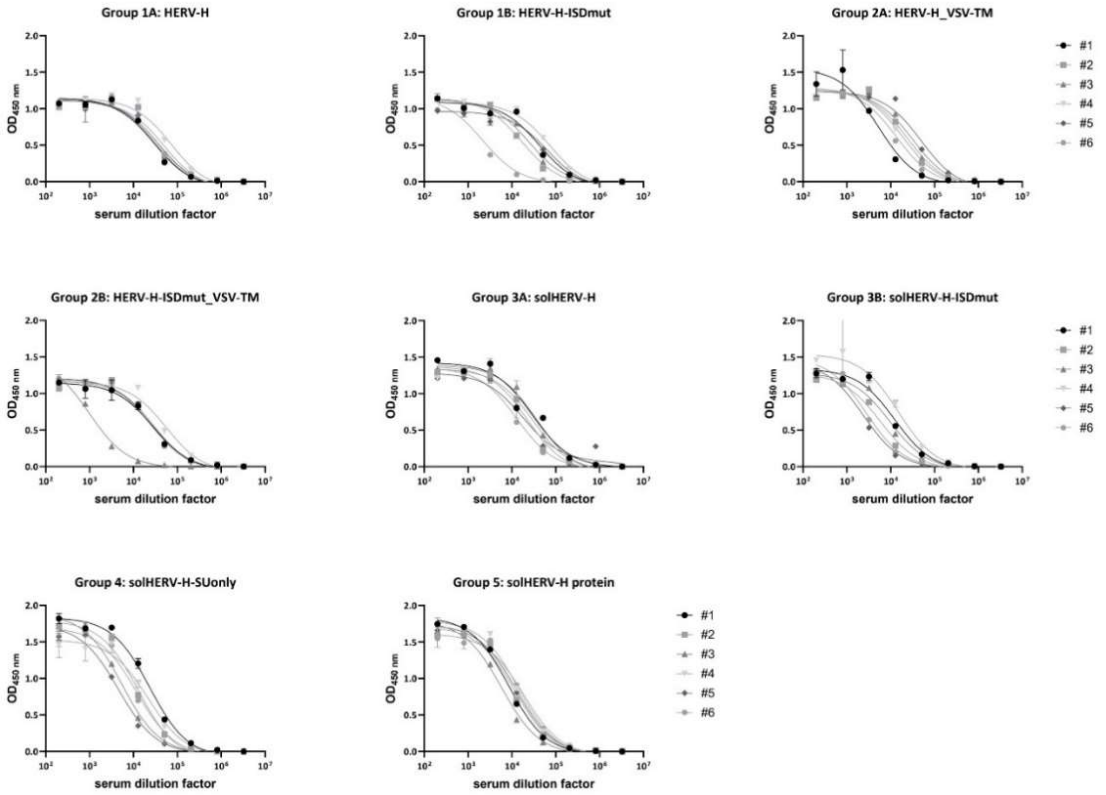


Figure S 12: Serum titration against the SUonly protein.

The 3rd bleed sera of BALB/c mice were titrated in a 4-fold serial dilution starting at 1:50 against the recombinant solHERV-H-SUonly protein. An anti-mouse Ig/HRP antibody was used as secondary antibody. Titration curves for all 6 individual mice per immunization group (#1-#6) are shown.

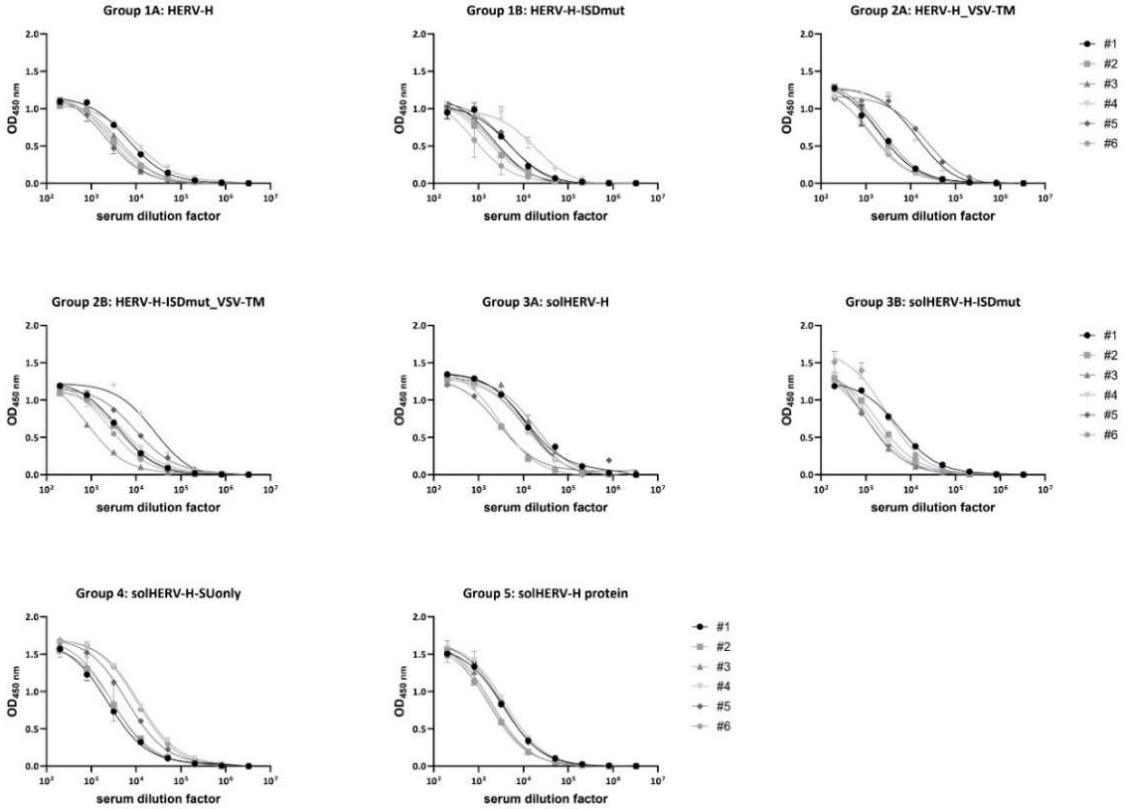


Figure S 13: Serum titration ELISA against the HERV-H env ectodomain.

The 3rd bleed sera of BALB/c mice were titrated in a 4-fold serial dilution starting at 1:50 against the recombinant solHERV-H ectodomain. An anti-mouse Ig/HRP antibody was used as secondary antibody. Titration curves for all 6 individual mice per immunization group (#1-#6) are shown.

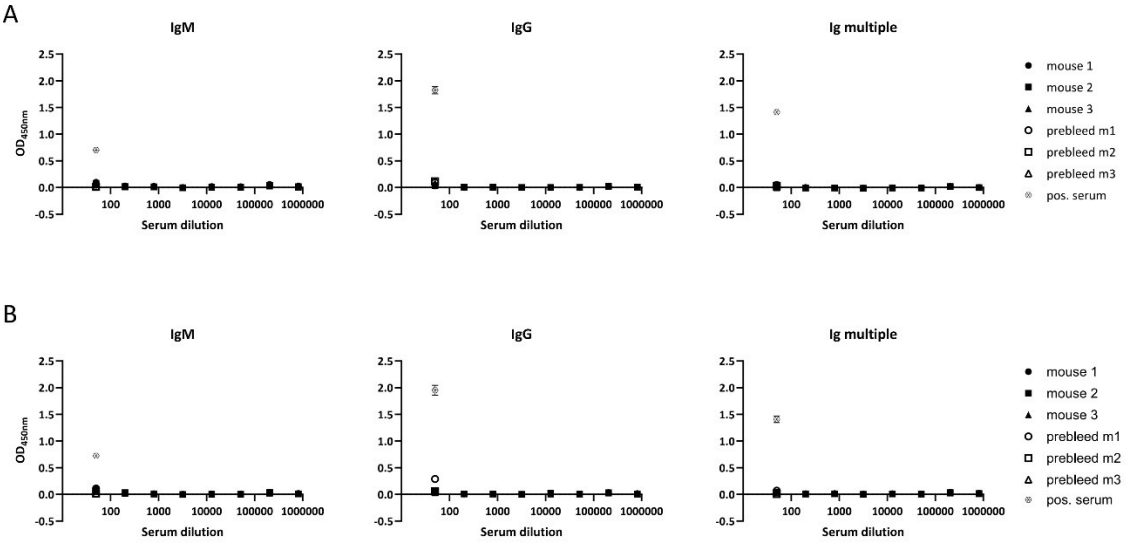


Figure S 14: ELISA with serum of mice immunized with HERV-H env RNA-LNPs.

Serum from mice immunized with different amounts of HERV-H env RNA-LNPs (mouse 1-3) and respective prebleed serum (prebleed m1-3) was titrated in a 4-fold serial titration starting at a 1:50 dilution against the HERV-H env protein. HRP-coupled secondary antibodies against mouse IgM, IgG and multiple Ig species were used for detection. As positive control, a HERV-H env positive mouse serum from 6.2.1d.was used (pos. serum). A: Mice were immunized 5 times every second week over a time course of 8 weeks with 10 µg RNA-LNPs. B: Mice were immunized three times

every 4 weeks over a time course of 8 weeks with 2 µg RNA-LNPs. Measurements were performed in technical duplicates, from which the mean value and standard deviation is depicted.

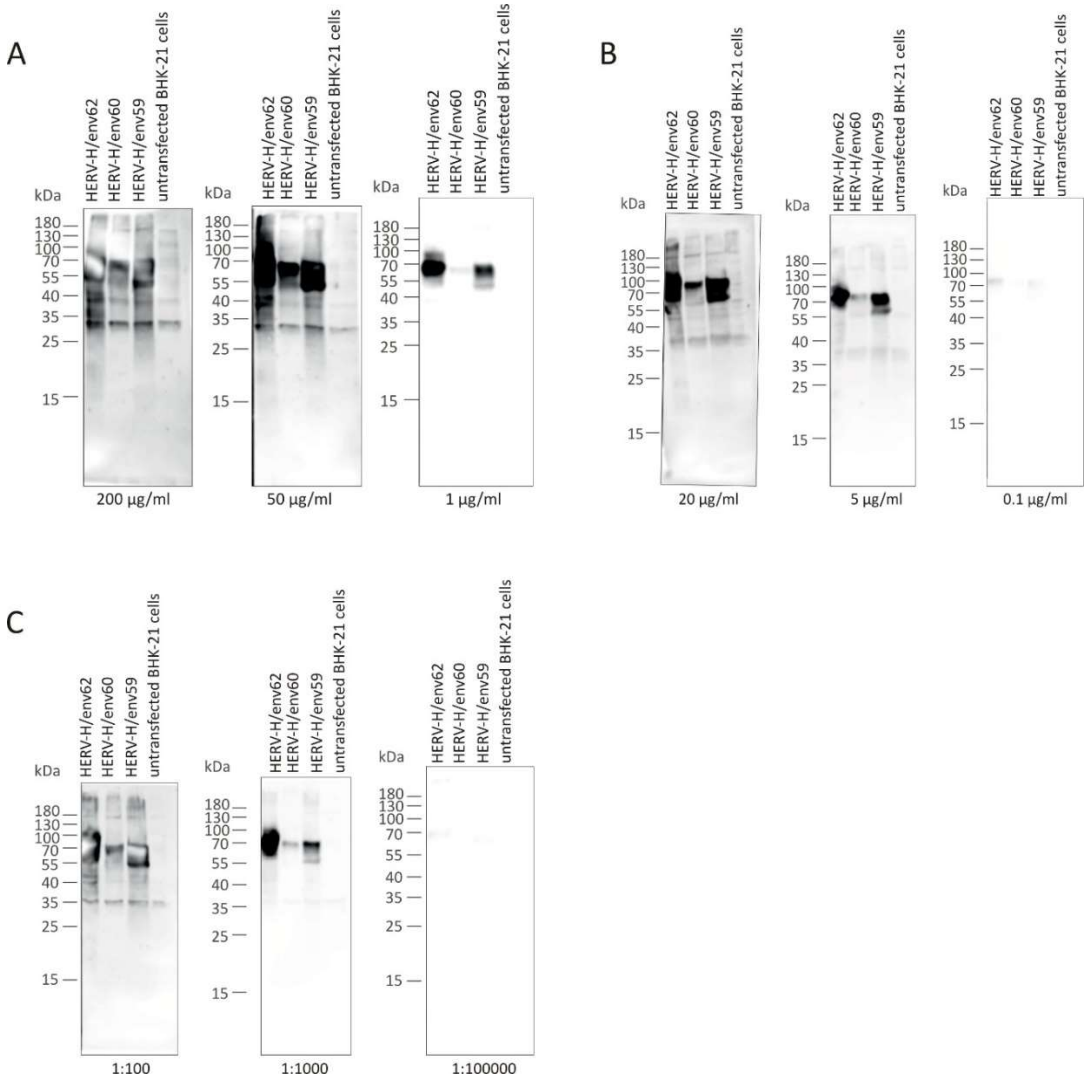


Figure S 15: Determining the optimal concentration of pAbs for western blot.

Different concentrations of the purified anti-HERV-H env pAb preparations were tested for their binding properties in a western blot against HERV-H transfected and untransfected BHK-21 cells. A: The Protein-A-purified pAb was applied at 200, 50 and 1 µg/ml. B: For the HERV-H affinity-purified pAb concentrations ranging from 0.1-20 µg/ml were tested. C: The unpurified rabbit serum was tested with the following dilutions: 1:100, 1:1000 and 1:100000. Binding was tested against the three HERV-H env variants HERV-H/env62, HERV-H/env60 and HERV-H/env59. For all blots an HRP-coupled anti-rabbit secondary antibody was used and the Prestained PageRuler Protein Ladder was applied as molecular weight marker.

Table S 2: Approximate number of cells per FFPE slide.

	DF-1	C-33A	MCF-7	MIA-PaCa-2
Approx. number of cells on one slide	78095	2626	50476	7222

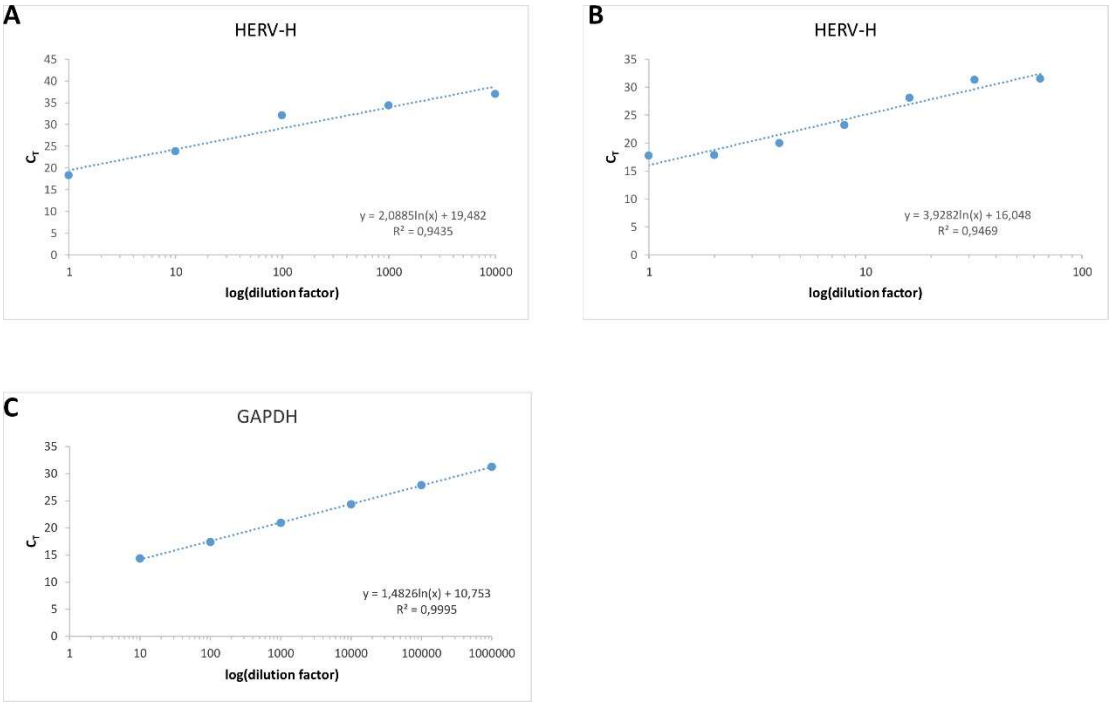


Figure S 16: Efficiency testing of HERV-H and GAPDH primers.

To determine the efficiency of primers used in the RT-qPCR mRNA from a HERV-H positive tumor cell line was titrated in either a 10-fold or a 2-fold serial dilution. The RT-qPCR was conducted using the HERV-H and GAPDH primer and detection via SYBR green. The C_T values were plotted against the logarithmic mRNA dilution factor and a logarithmic regression was performed. The curve formula and the R^2 value are depicted for each graph. A: mRNA extracted from a HERV-H positive cell line was titrated in a 10-fold serial dilution and the RT-qPCR was carried out using the HERV-H primer pair. B: HERV-H RT-qPCR with mRNA titrated in a 2-fold serial dilution. C: GAPDH RT-qPCR with mRNA titrated in a 10-fold serial dilution.

Table S 3: Calculated efficiency of primer pairs used for RT-qPCR.

Primer pair	mRNA dilution	Efficiency factor	PCR efficiency/%
HERV-H	1:10	1.61	61
HERV-H	1:2	1.29	29
GAPDH	1:10	1.97	97

The primer efficiency was calculated using the following equation: eff. = $e^{(-1/\text{-slope})}$. The values for the slope were taken from the regression formula shown in Figure S 16.

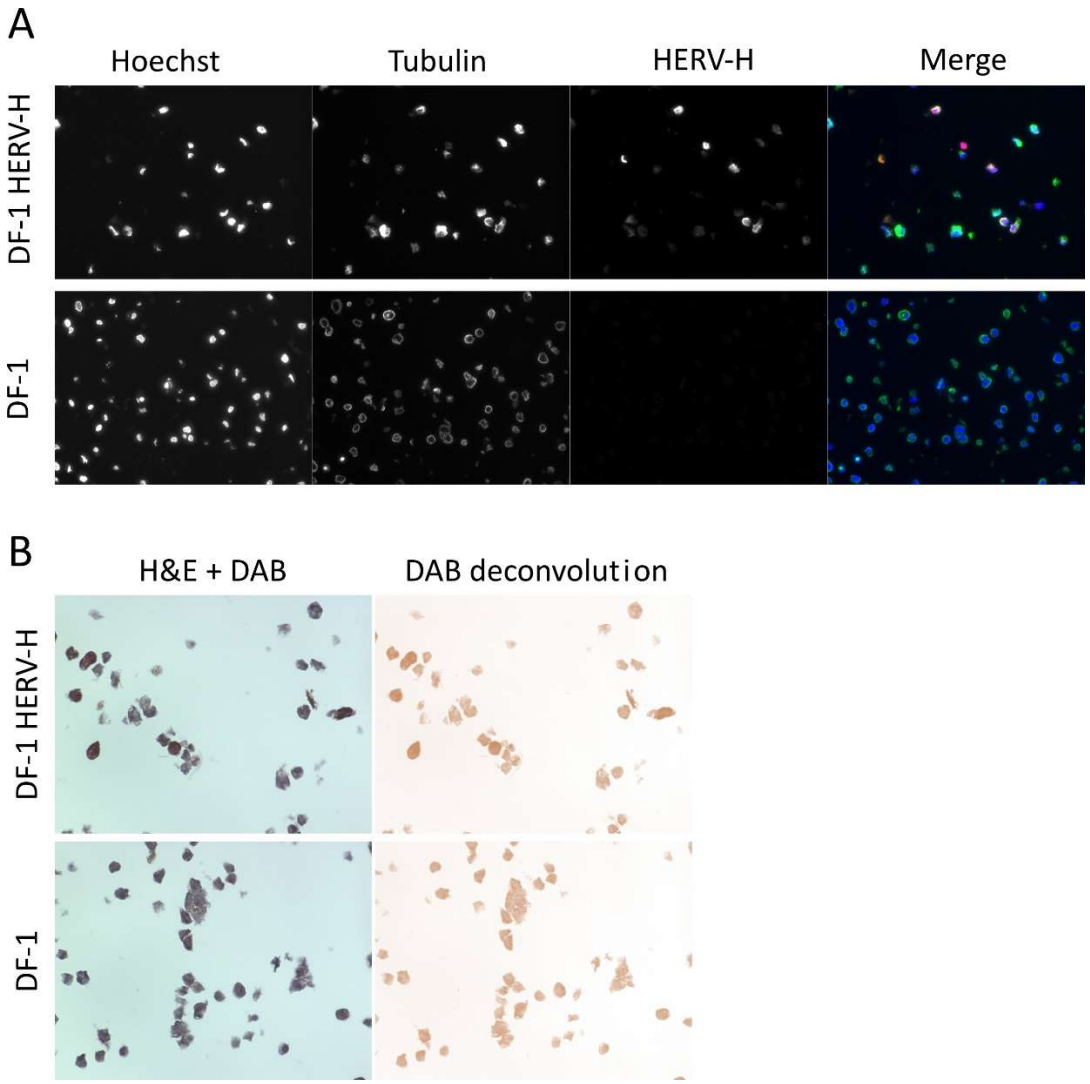


Figure S 17: ICC control staining with DF-1 and DF-1 HERV-H FFPE cell samples.

As control in the CRC IHC staining, paraffin-embedded DF-1 and HERV-H transfected DF-1 cells were stained in parallel to the CRC and normal tissues. A: Fluorescence control staining with an Alexa647-coupled anti-rabbit secondary antibody detecting HERV-H env (red), Hoechst33342 staining the nucleus (blue) and an Alexa488-coupled anti-tubulin antibody (green). B: DF-1 and DF-1 HERV-H FFPE samples were stained with the protein-A-purified anti-HERV-H env pAb and a HRP-coupled anti-rabbit secondary antibody. The HRP signal was detected via DAB staining and the cells were counterstained with a commercial H&E staining kit. The images on the left side show the complete staining, the images on the right side the DAB deconvoluted version. All images were taken with the Keyence BZ-9000 fluorescence microscope at 600x magnification.

Danksagung

Während der Zeit meiner Doktorarbeit habe ich viel Unterstützung von verschiedenen Menschen erhalten, denen ich hier danken möchte.

Zuallererst gilt mein Dank meinem Doktorvater Prof. Dr. Ralf Wagner. Vielen Dank, dass ich meine Doktorarbeit in deiner Arbeitsgruppe durchführen durfte. Du hast mich nicht nur jederzeit in meiner Arbeit unterstützt, sondern mich auch nach diversen Fehlschlägen wieder aufgebaut und immer an mich geglaubt. Außerdem bin ich dir sehr dankbar für die Erfahrungen, die du mir während meiner Doktorarbeitszeit ermöglicht hast. Der 4-monatige Forschungsaufenthalt in Kopenhagen und die Cold Spring Harbor Konferenz in New York waren absolute Highlights in meiner Zeit hier und haben mich auch persönlich stark geprägt.

Des Weiteren möchte ich meinen Mentoren Prof. Dr. Matthias Mack und Peter Holst für ihre vielen Ratschläge und Ideen danken, mit denen sie regelmäßig meine Arbeit weitergebracht habt. Peter Holst und dem gesamten ehemaligen InProTher Team möchte ich außerdem dafür danke, dass sie mich für 4 Monate bei sich aufgenommen haben und wesentlich dazu beigetragen haben, dass ich so eine großartige Zeit in Kopenhagen hatte. Hier möchte ich auch besonders Isabella Skandorff Petersen danken für die tolle Zusammenarbeit über die gesamte Zeit hinweg. Als Doktorandin auf der InPoTher Seite an diesem Projekt hast du mich nicht nur experimentell unterstützt, sondern auch mit mir zusammen gelitten und dich über Erfolge gemeinsam gefreut.

Bei Dr. Benedikt Asbach möchte ich mich für das Korrekturlesen sowie die vielen wertvollen Diskussionen und Anmerkungen bedanken.

Mein weiterer Dank gilt allen Mitgliedern der AG Wagner, sowie den Studierenden, die so tatkräftig an diesem Projekt mitgearbeitet haben. Euch allen ist es zu verdanken, dass ich immer gerne ins Labor gekommen bin. In einem Umfeld zu arbeiten, in dem man von seinen Arbeitskollegen nicht nur bei PhD-bezogenen, sondern auch bei privaten Krisen aufgefangen wird, ist unglaublich wertvoll und etwas ganz Besonderes. Und auch wenn mein Dank allen gilt, möchte ich doch ein paar Personen namentlich erwähnen. Da wäre zuerst einmal meine 65er Labor Truppe mit Antonia Senninger, Michael Schachtner und Sebastian Einhauser. Michi und Basti, ihr habt mich immer wieder aufgebaut und mir gut zugeredet, wenn mal wieder nichts zu funktionieren schien. Eure Unterstützung hat mich immer wieder weiter machen

lassen. Toni, meine Pommesfreundin, vielen Dank dass du mit mir den Testosteronspiegel in unserem Labor ausgeglichen hast. Mit dir habe ich einfach immer eine gute Zeit, egal ob bei der Arbeit, beim Klettern, bei Spieleabenden oder Burger essen.

Außerdem möchte ich mich bei Franz Attenkofer bedanken. Du hast mich nicht nur regelmäßig mit Schokolade versorgt, sondern warst mit deiner Erfahrung auch eine starke Unterstützung unter anderem bei der Etablierung und Auswertung der ICC Assays. Ein großer Dank geht auch raus an Dr. Martina Pfranger. Du warst von Anfang an dabei, hast mich immer wieder motiviert und immer an mich geglaubt. Danke auch, dass du zusammen mit Dominik Seidl so viel Zeit und Energie in die Generierung der RNA-LNPs gesteckt hast.

Ein Teil meiner Arbeit wäre nicht möglich gewesen ohne die Hilfe aus dem Institut für Pathologie. Daher vielen Dank an Prof. Dr. Daniela Hirsch und Doris Gaag für das Einbetten meiner Zelllinien und dem zur Verfügung stellen der Tumorschnitte.

Als letztes möchte ich mich bei den Menschen aus meinem privaten Umfeld bedanken. Vielen Dank an meine Familie für die konstante Unterstützung in allen Bereichen meines Lebens, ohne euch wäre dieser Weg für mich gar nicht erst möglich gewesen. Ein weiterer großer Dank geht außerdem an Lara Jeworowski. Danke für die gemeinsamen Urlaube, in denen ich einfach mal abschalten konnte, danke für die stundenlangen Skype-Anrufe, wenn ich mal wieder am Verzweifeln war und danke, dass du generell immer für mich da bist. Das aus einer 9-monatigen gemeinsamen Zeit in der Arbeitsgruppe so eine tolle Freundschaft entstanden ist, ist definitiv ein weiteres Highlight aus dieser Zeit für mich.