Ollscoil na Gaillimhe
University of Galway

Provided by the author(s) and University of Galway in accordance with publisher policies. Please cite the published version when available.

Title	Human endogenous retrovirus K (HERV-K) in prostate cancer
Author(s)	Downey, Rónán
Publication Date	2016-08-08
Item record	http://hdl.handle.net/10379/5928

Downloaded 2024-04-25T13:14:28Z

Some rights reserved. For more information, please see the item record link above.

Human endogenous retrovirus K (HERV-K) in prostate cancer

A thesis submitted to the National University of Ireland, Galway in fulfilment of the requirements for the degree of

Doctor of Philosophy

By

Rónán Downey

Prostate Cancer Institute,

School of Medicine,

National University of Ireland, Galway

Thesis supervisor: Dr. Sharon Glynn

March 2016

Table of Contents

Declaration

This thesis describes work that I undertook between 2012 and 2015 at the Prostate Cancer Institute, National University of Ireland, Galway. This work was supervised by Dr. Sharon Glynn (Lecturer in Pathology, School of Medicine, National University of Ireland, Galway).

I declare that the results presented herein are from original experimental work which has been carried out by me for the purpose of this thesis. The results from chapter 4 of this thesis represent work from a collaborative paper of which I was joint first author. Work conducted by myself towards this publication included assay optimization, cDNA synthesis, qPCR and initial data analysis. More complex statistical analysis was performed in conjunction with my supervisor Dr. Sharon Glynn. The work described within this thesis has not been submitted for degree, diploma or other qualification at any other university.

I have no conflict of interest pertaining to the subject matter of this work.

Funding Authorities

This work was funded by the Health Research Board of Ireland Clinical Research Facility, Galway and by the Irish Research Council.

Abstract

Millions of years of infection by viral agents have left their mark on the human genome in the form of human endogenous retroviruses (HERVs). These retroelements integrated into our germline, became subject to Mendelian inheritance, and now occupy 8% of the genome. The majority of these HERVs are replication incompetent due to the accumulation of defective mutations. The most conserved HERV group is that of HERV-K – more specifically HERV-K (HML-2). A number of proviral members of this sub-family contain intact ORFs for all the canonical retroviral genes (gag-pol-pro-env), including accessory proteins such as Rec.

HERVs have been implicated in the etiology of cancers such as breast, ovarian and melanoma due to the observation of high levels of HERV-K (HML-2) mRNA and protein. However, a paucity of data exists regarding the expression of HERVs in prostate cancer.

We set out to investigate HERV-K (HML-2) expression in prostate cancer. We analysed prostate cancer cell lines and clinical samples for HML-2 expression. Our results show that both HML-2 mRNA and protein is expressed in prostate cancer cell lines and in clinical samples. We also found evidence for the expression of spliced *Rec* and *Np9* mRNA and Rec protein in prostate cancer cell lines.

We completed a case-control study to investigate the potential of using HML-2 as a non invasive biomarker for the diagnosis of prostate cancer by analysing blood for HML-2 mRNA expression using qPCR. Univariate logistic regression analysis revealed that HML-2 gag mRNA expression in PBMCs was significantly associated with a prostate cancer diagnosis and that this association was strongest in older men and smokers. Further to this, we conducted immunohistochemistry on prostate tissue microarrays containing cores from prostate adenocarcinoma and benign prostatic hyperplasia (BPH) cases. We observed significantly greater epithelial

HERV-K gag protein expression in BPH cores versus prostate adenocarcinoma cores (Fisher exact test, p=0.001). Our data indicate that measuring HERV-K (HML-2) expression may improve the discriminatory ability of prostate specific antigen (PSA) in the diagnosis of prostate cancer versus BPH especially in older men.

We also investigated whether Rec plays a functional role in prostate cancer initiation or progression. To this end, we created a construct to overexpress Rec in prostate cancer cells and transfected this into non tumorigenic RWPE-1 prostate cells. We found Rec expression to be low to negative in the transfected cell line but the cloning process did reveal the expression of a polymorphic provirus from chromosome 5 in LNCap cells which most likely represents ERVK-9.

Overall, this thesis has demonstrated that HERV-K (HML-2) is expressed in prostate cancer and that this expression has the potential to be utilised as a biomarker for disease detection. Finally, the functional role of HML-2 expression in prostate cancer initiation and progression warrants further investigation.

List of figures

List of tables

Abbreviations

AA	African American
AML	Acute myeloid leukaemia
APOBEC3G	Apolipoprotein B mRNA-editing, enzyme catalytic, polypeptide-like 3G
AR	Androgen receptor
Bp	Base pair
BPH	Benign prostatic hyperplasia
cDNA	Complementary DNA
ChIP	Chromatin immunoprecipitation
CML	Chronic myeloid leukaemia
CTC	Circulating tumour cell
DMSO	Dimethyl sulfoxide
dNTP	deoxynucleotide triphosphate
EA	European American
EBV	Epstein-Barr virus
eMLV	Ectropic murine leukaemia virus
ERG1	ETS-related gene
ETV1	ETS translocation variant
FFPE	Formalin fixed paraffin embedded

FSC	Forward scatter
GFP	Green fluorescent protein
HERV	Human endogenous retrovirus
HIV	Human immunodeficiency virus
HML-2	Human mouse mammary tumour like
HPV	Human papillomavirus
hSGT	Human small glutamine-rich tetratricopeptide repeat protein
HTLV-1	Human T-lymphotropic virus
IHC	Immunohistochemistry
IL	Interleukin
INF	Interferon
IRB	Institutional review board
kDa	Kilodalton
LTR	Long terminal repeat
MHC	Major histocompatability locus
MMTV	Mouse mammary tumour virus
mRNA	Messenger RNA
NCI	National cancer Institute
NO	Nitric oxide
NOS2	Nitric oxide synthase 2
OR	Odds Ratio

ORF	Open reading frame
PBMC	peripheral blood mononuclear cell
PCR	Polymerase chain reaction
PIA	Proliferative inflammatory atrophy
PIN	Prostatic Intraepithelial Neoplasia
PLZF	Promyelocytic leukaemia zinc finger
PSA	Prostate specific antigen
qPCR	Quantitative PCR
ROC	Receiver operating characteristic
RQ	relative quantification
RT-PCR	Reverse transcription PCR
Sags	Superantigens
SEREX	Serological recombinant cDNA expression cloning
siRNA	Small interfering RNA
TE	Transposable element
TLR	Toll-like receptor
TMPRSS2	Transmembrane protease, serine 2
TZFP	Testicular zinc-finger protein

Chapter 1

Introduction

The following chapter is adapted from Downey et al*. Human endogenous retrovirus K and cancer: Innocent bystander or tumorigenic accomplice? International Journal of Cancer Volume 137, Issue 6, pages 1249–1257, 15 September 2015 (Downey* et al*., 2015). An appended PDF together with a copyright agreement is situated in the appendices of this thesis.*

1 Introduction

1.1 Viruses

Viruses represent unique infectious entities which are ubiquitous across planet earth (Fields et al., 2007). Not considered to be truly living, viruses cannot replicate independently, requiring another cell to do so (Fields et al., 2007). They successfully infect organisms from across all the domains of life (Fields et al., 2007). Evidence for viruses stretches back to at least 407 million years ago. Ancient coelacanth genomes found from this time were shown to have been infected by foamy virus like elements (Han and Worobey, 2012). Viral biology is extraordinarily diverse but three main types exist: DNA viruses, RNA viruses and reverse transcribing viruses or retroviruses (Fields et al., 2007).

1.2 Retroviruses

Retroviruses are defined by their innate ability to reverse transcribe an RNA template into DNA (Kurth and Bannert, 2010b) **(Figure 1).** This ability is endowed on retroviruses due to the presence of an inbuilt reverse transcriptase activity which is encoded in its genome (Kurth and Bannert, 2010a). The canonical structure of all retroviruses consists of gag, pro, pol and env genes bordered by long terminal repeat (LTR) units which regulate their expression (Kurth and Bannert, 2010b).

Retroviruses are responsible for a large amount of human pathologies including AIDs which is caused by infection with the human immunodeficiency virus 1 (HIV-1) (Evans and Kaslow, 1997).

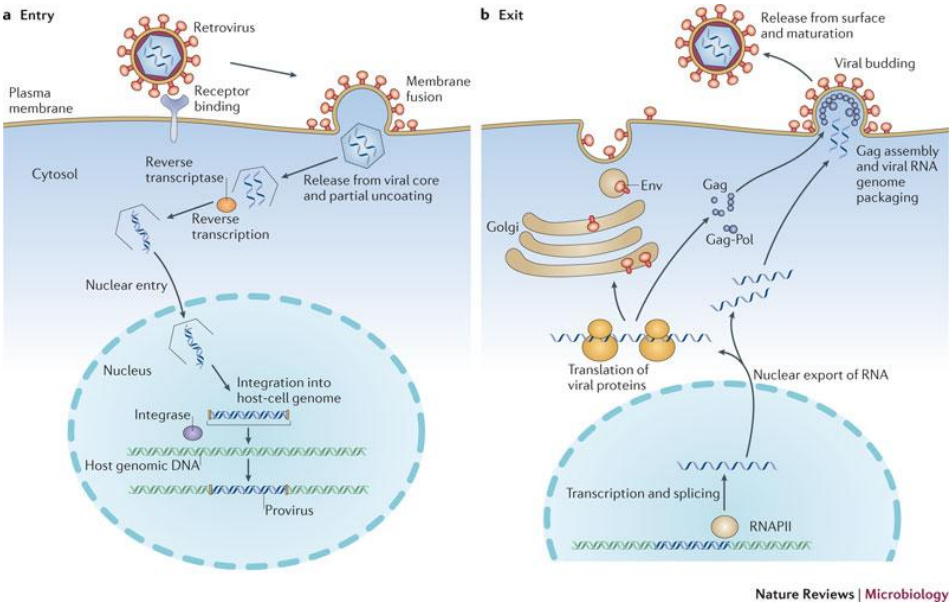


Figure 1 - Life cycle of a retrovirus

Retroviruses begin their infection of a cell by binding to cellular receptors via their envelope protein. After binding, the retroviral core is taken into the cytoplasm and the RNA genome gets reverse transcribed into a DNA copy which is then integrated into the genome. The next phase of the life cycle begins when the DNA genome is transcribed back into RNA. The retroviral genes are translated and the mature proteins gather at the cell membrane where they assemble to form the mature virus which eventually buds off from the surface. In the case of endogenous retroviruses (ERVs), this process occurs in a cell of the germline. The proviral DNA then gets transmitted in a Mendelian fashion. Adapted from (Stoye, 2012).

1.3 Endogenous Retroviruses

On rare occasions, retroviruses infect cells of the germline and become incorporated into the genome and transmitted in a Mendelian fashion (Stoye, 2012). Such retroviruses are from that point known as endogenous retroviruses (ERVs). ERVs have been found in the genomes of most of the jawed vertebrates (Belshaw et al., 2004). An interesting case study is that of the koala (*Phascolarctos cinereus*), in which Koala retrovirus (KoRV) is currently undergoing endogenisation (Stoye, 2006).

ERVs were first discovered in chickens in the early 1970s when an endogenised form of avian leukosis virus (ALV) was discovered in chickens

(Weiss, 2006). Thanks to new breakthroughs in molecular biology around this time, many more ERVs were discovered in many different kinds of organisms (Weiss, 2006). Human endogenous retroviruses (HERVs) were also discovered during the 1970s when scientists were exploring the possibility of a viral etiology for breast cancer (Ono et al., 1986). Using probes for mouse mammary tumour virus (MMTV) they accidently pulled out sequences which exhibited high homology to MMTV but were in fact ERVs. Today we know that 8% of the human genome consists of HERVs (Bannert and Kurth, 2004). HERVs make up a large part of what is termed 'junk DNA' (Lander et al., 2001). The notion that large swathes of the human genome are junk has changed over the last decade as researchers discover that much of this sequence contains function. The ENCODE study was the largest such study to date (Consortium, 2004). It concluded that 80% of the genome had some biochemical function. This conclusion has been somewhat controversial in that sequence with a biochemical function does not necessarily mean that it contributes to phenotype (Doolittle, 2013).

HERVs can also be classified as transposable elements **(Figure 2).** These make up about 45% of the human genome (Lander et al., 2001). Long Interspersed Nuclear Elements (LINEs) and Short Interspersed Nuclear Elements (SINEs) are other examples of transposable elements known specifically as retrotransposons (Lander et al., 2001).

HERVs are a potential source of new genetic material from which the generation of new genes can stem from (Kurth and Bannert, 2010a). The prototypical example is that of the gene syncytin. It originates from a HERV-W env gene (Mi et al., 2000). Syncytin is essential for the formation of the placenta, helping to form the syncytiotrophoblast by utilizing the fusogenic activity of env (Mi et al., 2000). Many LTRs of HERVs have been co-opted by protein coding genes so as provide them with different transcriptional characteristics (Kurth and Bannert, 2010a). Many different groups of HERVs exist. Each group is related to the original exogenous retrovirus which infected the cell (Bannert and Kurth, 2006). HERV groups are named according to the primer binding site of the t-RNA e.g HERV-K

uses lysine (Bannert and Kurth, 2004). The large majority of HERVs have become mutated over evolutionary time such that most have lost their coding capacity (Bannert and Kurth, 2006). These mutations include mainly frameshift mutations and premature stop codons. Other proviruses have had their genetic structure deleted due to homologous Recombination between LTRs. This has meant that the large majority of remaining HERV material in the human genome consists of solitary LTRs as opposed to actual proviruses.

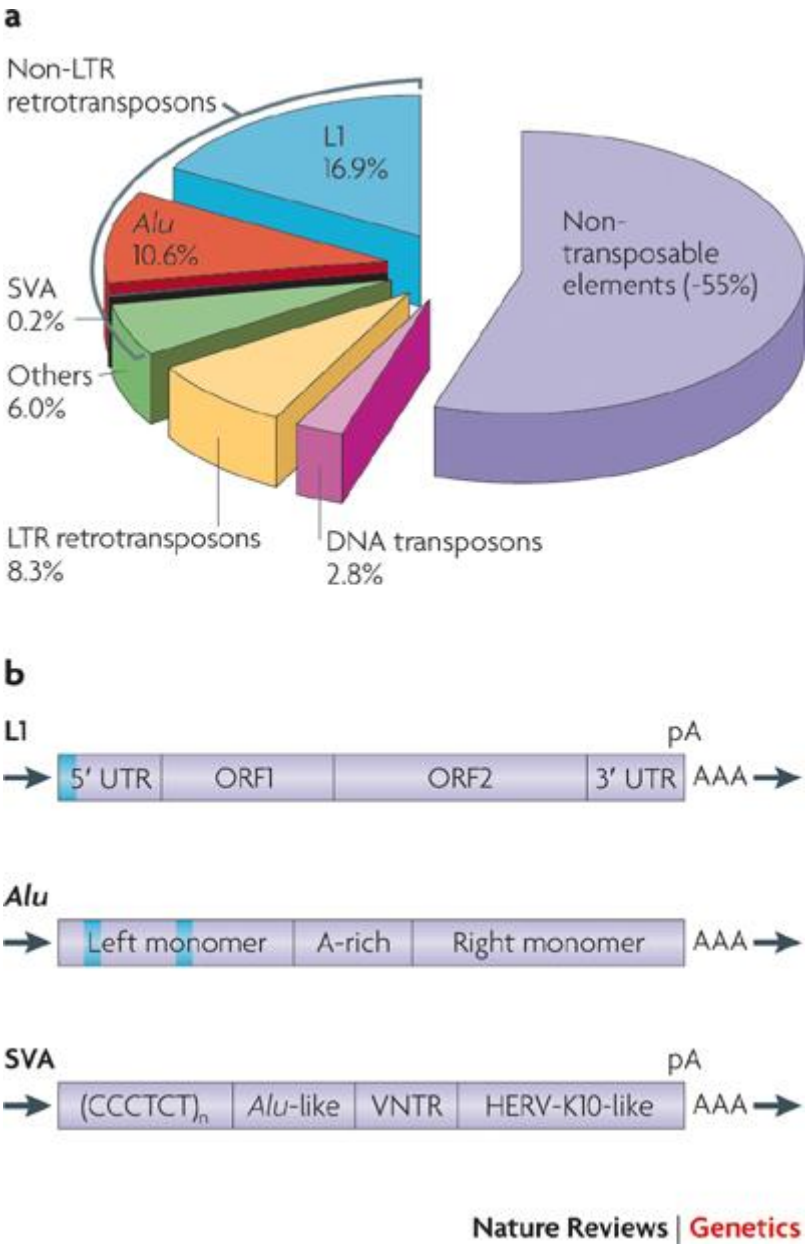


Figure 2 - The transposable element content of the human genome

Around 45% of the human genome consists of transposable elements. Non LTR retrotransposons such as Alu, SINEs and LINEs make up around 33% of the genome. HERVs are classed as LTR retrotransposons and make up 8.3% of the genome. Adapted from (Cordaux and Batzer, 2009)

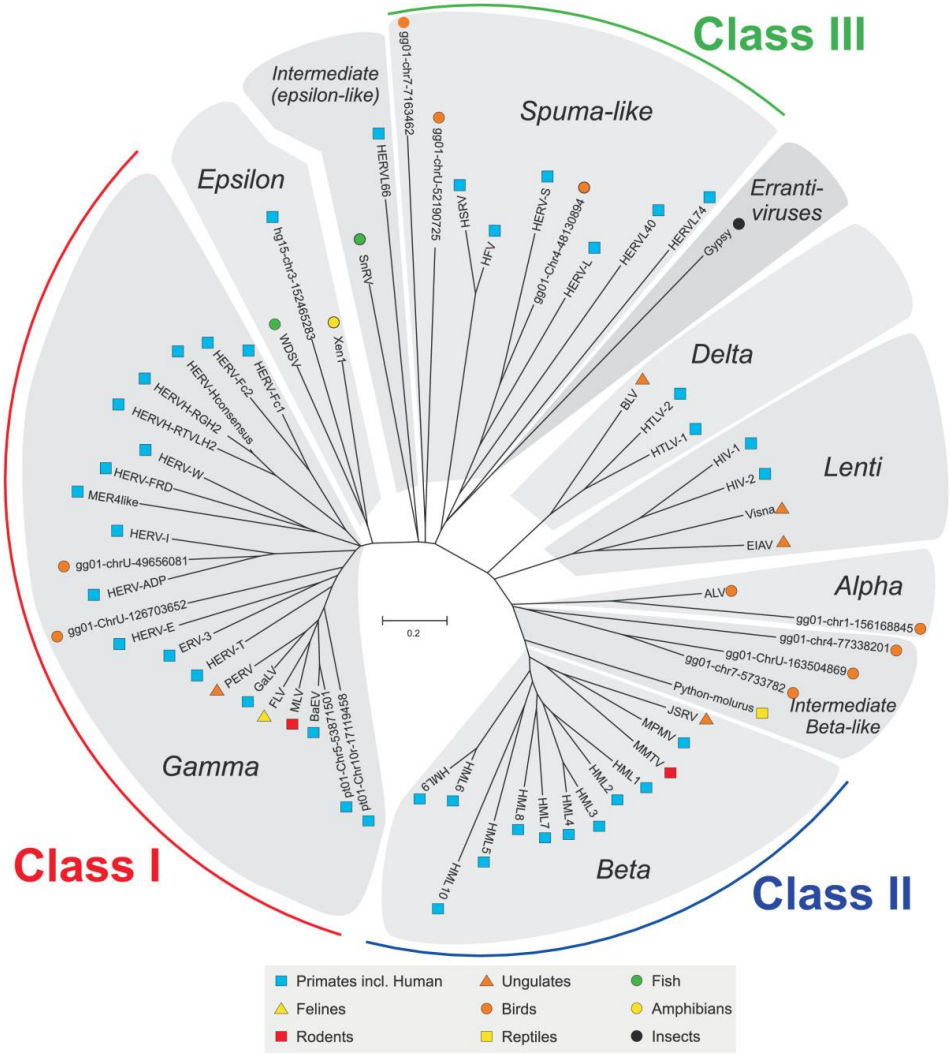


Figure 3 - HERV neighbour joining dendrogram

HERVs can be broadly grouped into three classes depending on their similarity to exogenous retroviruses. The HML HERVs exhibit closest homology to the Beta retroviruses such as MMTV and JSRV. Adapted from (Jern et al., 2005)

The HERV-K group are class II HERVs, and exhibit closest homology to betaretroviruses, which cluster as class II retroelements (Bannert and Kurth, 2006) **(Figure 3)**. It consists of 11 subgroups (HML-1 to HML-11), each the result of a separate germline infection (Subramanian et al., 2011). One of these subgroups, HML-2, has been subject to intensive research, due to the fact that it maintains an unrivalled coding competence, with many of its proviruses maintaining complete, or near-complete, open reading frames (ORFs) for all viral polyproteins (**Figure 4**) (Bannert and Kurth, 2006). Finally, it represents the most recently integrated HERV group into the human genome. Some HML-2 proviruses are both human specific and/or polymorphic indicating integration events subsequent to the human-chimpanzee split and within modern humans (Moyes et al., 2007). This likely contributes to the fact that HML-2 is the least defective and most active retroviral family. In this regard, HML-2 is considered the most interesting HERV group to study in terms of potential oncogenic activity.

Overall, HML-2 is represented in the genome by 91 proviruses and 944 solitary (solo) LTRs. Solo LTRs are the result of unequal crossing over due to highly homologous sequences (Subramanian et al., 2011). Two main types of HERV-K (HML-2) are found in humans: type I is characterized by a 292 base pair deletion at the boundary of the *pol* and *env* (envelope) genes (**Figure 4**), while type II lacks it (Downey et al., 2015). The deletion in type I proviruses leads to an alternative splicing event culminating in a protein known as Np9, while type II proviruses express a complete accessory protein known as Rec (Bannert and Kurth, 2006).

HERVs play an important role in normal physiological function. For example, the protein syncytin 1 mediates cellular fusion of the placental trophoblast and is encoded by an *env* gene from the HERV-W group (Mi et al., 2000). Another syncytin – known as syncytin 2 – plays a similar physiological role and is encoded by an env gene from the HERV-FRD group (Blaise et al., 2004). Finally, the presence of HERVs, in particular their LTR elements, has added an additional layer of complexity to our genome, in that many of these LTRs have been co-opted by protein coding

genes and serve as regulatory elements directing tissue specific expression (Cohen et al., 2009).

The association of HERVs with disease has garnered the most attention from researchers. HERVs have been implicated in autoimmune disorders (Balada et al., 2009; Urnovitz and Murphy, 1996), but with conflicting reports particularly involving multiple sclerosis (MS) (Curtin et al., 2012; Mameli et al., 2012; Perron and Lang, 2010). Recent research refutes a role for deregulated HERV-W *env* in MS lesions, including the high-level-transcribed *ERVWE1* locus encoding Syncytin-1 (Schmitt et al., 2013b). In this introduction, we discuss the most recent developments in the field of HML-2 and human tumour biology, in particular emerging evidence of a role for HML-2 in immunomodulation and the presence of HML-2 in tumour derived exosomes, further indicating the potentially important role of HML-2 in human carcinogenesis.

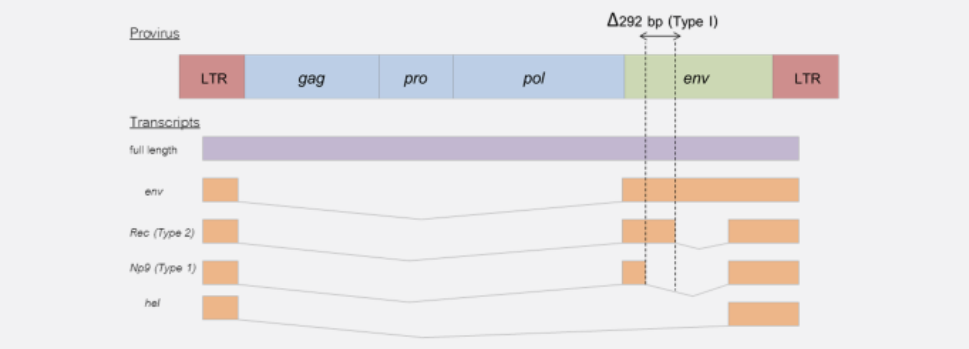


Figure 4 - Structure of HERV-K provirus

The full length (gag) HML-2 transcript encodes the gag, pro and pol polyproteins. A singly spliced transcript encodes the env polyprotein, while a doubly spliced transcript encodes either the Rec or Np9 accessory proteins depending on the presence or absence of a 292-bp deletion at the pol/env boundary—a characteristic that defines a HML-2 provirus as either Type 1 (deleted) or Type 2 (intact). HML-2 also expresses a 1.5-kb transcript of unknown function known as the hel transcript. Adapted from (Downey et al., 2015)

1.4 HML-2 and Solid Tumours

To date there is evidence for HML-2 activation in ovarian cancer (Iramaneerat et al., 2011; Wang-Johanning et al., 2007), melanoma (Buscher et al., 2005; Reiche et al., 2010; Serafino et al., 2009), breast (Burmeister et al., 2004; Golan et al., 2008; Wang-Johanning et al., 2003; Wang-Johanning et al., 2001; Wang-Johanning et al., 2008), prostate (Agoni et al., 2013a; Goering et al., 2011; Ishida et al., 2008; Tomlins et al., 2007) , lymphomas (Contreras-Galindo et al., 2008), leukaemia's (Depil et al., 2002) and sarcomas (Schiavetti et al., 2002). In the 1980s Ono et al. successfully cloned HML-2, thanks to its similarity to mouse mammary tumour virus (MMTV) (Ono et al., 1986). They also found that stimulation of human breast cancer cell lines with female steroid hormones led to an upregulation of HML-2 mRNA (Ono et al., 1987). Several groups followed with reports of HML-2 mRNA and viral particle expression in breast cancer (Contreras-Galindo et al., 2008; Etkind et al., 1997; Seifarth et al., 1995; Willer et al., 1997). Wang-Johanning and colleagues refined this work to produce data which accurately quantified HML-2 *env* transcripts and spliced transcripts in breast tumours demonstrating elevated levels compared to unaffected controls (Wang-Johanning et al., 2003; Wang-Johanning et al., 2001). They also demonstrated an association between HML-2 Env protein expression in breast tumours and increased risk of lymph node metastasis and poor outcome, in two separate US cohorts and a Chinese cohort of breast cancer patients (Wang-Johanning et al., 2012; Zhao et al., 2011), corroborating the findings of Golan et al.(Golan et al., 2008). Most recently Wang-Johanning et al. demonstrated that HML-2 serum mRNA and anti-Rec antibody titers are predictive of early stage breast cancer. Additionally HML-2-gag copy number tended to be higher in breast cancer patients with a primary tumour who later on developed the metastatic (Wang-Johanning et al., 2014).

High levels of expression of HML-2 *env*, *rec* and *np9* mRNA, and Env protein have been reported in ovarian cancer cell lines and tumours (Wang-

Johanning et al., 2007), while in another study Np9 mRNA was not detectible in two ovarian tumours tested (Armbruester et al., 2002). One possible mechanism of altered HML-2 expression in ovarian cancer may be due to alterations in its methylation status (Iramaneerat et al., 2011).

1.5 HML-2 and Haematological Malignancies

Brodsky et al. discovered a potential role for HML-2 in leukaemia. They showed that HML-2 pol mRNA was expressed in the blood of patients suffering from chronic myeloid leukaemia (CML) and acute myeloid leukaemia (AML) (Brodsky et al., 1993a; Brodsky et al., 1993b). Others also reported that HML-2 *gag* mRNA is present at higher levels in PBMCs of leukaemia patients compared to healthy controls (Depil et al., 2002). Similar studies have reported HML-2 viral particles in lymphomas (Contreras-Galindo et al., 2008), and HML-2 env expression in the H9 human T-cell lymphoma cell line (Vogetseder et al., 1995). Additionally the human lymphotropic herpesvirus Epstein-Barr virus (EBV), which has been implicated in the development of lymphoma, was shown to induce HERV-K18 *env* gene expression. The HERV-K18 env has been reported to have superantigen (SAg) activity by several groups (Hsiao et al., 2009; Sutkowski et al., 2001), while others have found no evidence of SAg activity (Azar and Thibodeau, 2002; Badenhoop et al., 1999). Indeed multiple HML-2 env proteins elicit antibody responses (Herve et al., 2002; Wang-Johanning et al., 2008). A direct association between HERV-K18 env SAg and carcinogenesis has yet to be shown. HML-2 expression has also been seen to decrease after lymphoma therapy, indicating that it may be useful for monitoring therapeutic response (Contreras-Galindo et al., 2008).

1.6 HML-2 and Melanoma

The prevalence of HML-2 *pol*, *gag*, and *env* mRNA, and Gag and Env proteins in melanoma is well established (Buscher et al., 2006; Buscher et al., 2005; Humer et al., 2006; Muster et al., 2003; Reiche et al., 2010; Serafino et al., 2009; Singh et al., 2009). In 2002, a sequence homologous to HERV-K (HML-6) was identified in melanoma patients (HERV-K-MEL)

(Schiavetti et al., 2002). HERV-K-MEL, which produces an antigen spliced from a defective non-coding env-like open reading frame, was reported in cutaneous and ocular melanomas, and nevi. Antibodies raised against the HERV-K-MEL antigen were detectable in melanoma patients (Schiavetti et al., 2002). Melanoma cell lines (SKMel-28, SKMel-1, 518A2, MelJuso, HS-Mel2 and JH-Mel6 and HV-Mel7), but not cultured melanocytes (NHEM neo 5935, NHEM neo 4528, and NHEM neo 6083) produce retrovirus-like particles that exhibit reverse transcriptase activity (Muster et al., 2003), which contain mature Gag and Env proteins. HML-2 Pol, Gag and Env (Muster et al., 2003), and accessory proteins Rec and Np9 have also been detected in melanoma (Buscher et al., 2006; Buscher et al., 2005). Further studies sought to predict the prognostic value of HML-2 in melanoma and found that HML-2 was a statistically significant marker of acrolentiginous, mucosal and uveal melanoma. Patients with serological response against HML-2 had a significantly decreased disease-specific overall survival (Hahn et al., 2008). Additionally HML-2 rec mRNA is expressed in melanoma cells but not in benign nevi or normal skin, indicating aberrant activation in melanoma. Furthermore rec mRNA positivity correlated with the vertical growth phase of melanoma, a step which increases the risk of metastatic melanoma (Singh et al., 2013). A recent study by Schmitt and co-workers defined the HML-2 transcriptome in melanoma, identifying 23 different HML-2 loci as transcribed to varying degrees in different patient specimens and melanoma cell lines (Schmitt et al., 2013a).

1.7 Polymorphic HML-2 group members

Of the ninety one known HML-2 proviruses, eleven are polymorphic(Subramanian et al., 2011). The most recent insertions (approximately 1million years ago) include HERV-K-19p12 (K113) (29% of individuals), and HERV-K-8p23.1 (K115) (16% of individuals) as measured using a pool of mixed backgrounds (Turner et al., 2001). Other polymorphic HML-2 proviruses include: HERV-K-11q22.1(K118), HERV-K-6q14.1 (K109), HERV-K-7p22.1a(K108R), HERV-K-8p23.1 (K115) and HERV-K-1p31.1(K116) (Belshaw et al., 2005; Hughes and Coffin, 2004),

in addition to HERV-K-3q13.2(K106), HERV-K-7p22.1b(K108L), HERV-K-10p12.1(K103), HERV-K-12q13.2, and finally HERV-K-U219(K105) located in the unassembled centromeric region (Un_g1000219) (Subramanian et al., 2011).

It is currently not known whether inheriting specific HML-2 polymorphisms increases the risk of cancer. Burmeister et al. investigated the frequency of the polymorphic full-length HERV-K115 and HERV-K113 in 102 female breast cancer cases and 102 controls, but did not find a significant association with breast cancer (HERV-K-K113, 16.7% vs. 12.7%; HERV-K-K115, 4.9% vs. 9.8%). (Note the lower prevalence than reported above (Turner et al., 2001) for both. This suggests ethnic differences in frequency of inheritance) (Burmeister et al., 2004).

1.8 Mechanisms of HERV Activation and Regulation

The abundance of inactive HERVs present in our genome suggests that active, integrating proviruses are largely deleterious to the host. Novel intrinsic restriction factors exist which impede retroviral infection and some of these have the ability to target both exogenous and endogenous infection. APOBEC proteins can inhibit viral RNA, thus blocking their translation (Mullins and Linnebacher, 2012). Additionally APOBEC3G can hypermutate and inactivate HERV DNA (Lee et al., 2008). Activation of these retroelements can therefore be an indication that cellular programs, crucial to a healthy phenotype, have gone awry.

A crucial question which needs to be addressed is whether activation of HERVs is simply an epiphenomenon or is necessary for disease progression? A large proportion of HERV loci have become silenced via DNA hypermethylation, an epigenetic phenomenon (Schulz et al., 2006). Many cancers display a globally hypomethylated state (Ehrlich, 2002), thus activation of HERVs during tumourigenesis may simply be a bystander effect of this epigenetic state. It has become increasingly clear that genomic instability, including deregulated transcription and genome plasticity, is enabled as a result of epigenetic changes that take place within tumours.

Demethylation of specific HERV families, including HERV-W, HERV-K, and HERV-H, has been reported in various cancers (Gimenez et al., 2010). Moreover, demethylation of transposable elements correlates with their transcriptional activation in prostate cancer (Goering et al., 2011). This indicates that where HERV transcription is increased in cancer cells, it is likely due in part to hypomethylation of their LTRs. HML-2 DNA hypomethylation has been reported in melanoma cell lines (Stengel et al., 2010), prostate tumours (Goering et al., 2011) and ovarian tumour (Iramaneerat et al., 2011). Interestingly, age was negatively associated with HML-2 methylation in PBMCs from healthy donors aged 20-88 years. The average onset of HML-2 methylation in PBMCs occurred at 40-63 years, implicating HML-2 DNA hypomethylation in aging (Jintaridth and Mutirangura, 2010). Another important epigenetic mechanism that influences transcriptional activity is histone modification, but the influence of histone methylation, acetylation or other modifications on HERV expression in malignancy is still unknown (Hurst et al., 2016).

Known inducers of HML-2 *in vitro* include ultraviolet radiation in melanoma (Reiche et al., 2010; Schanab et al., 2011), hormones, including progesterone, estrogen and androgen in breast (Ono et al., 1987; Wang-Johanning et al., 2003) and prostate (Tomlins et al., 2007) cancer cell lines, and bone morphogenetic proteins and retinoic acid in testicular germ cell tumour cell lines (Caricasole et al., 2000). Urine from smokers has also been shown to lead to an increase in HERV expression in normal human dermal fibroblasts and urothelium *in vitro* (Gabriel et al., 2010). Other factors which may activate or be activated by HML-2 include infectious viruses such as EBV (Sutkowski et al., 2004), and human immunodeficiency virus (HIV-1) (Jones et al., 2012), and transcription factors including NF-κB, NF-AT (Gonzalez-Hernandez et al., 2012), MITF-M (Katoh et al., 2011), Sp1, Sp3 (Fuchs et al., 2011), and YY1 (Knossl et al., 1999).

1.9 Possible Mechanisms of HML-2 Induced Oncogenesis

1.9.1 Insertional Mutagenesis

HERVs may be oncogenic via insertional mutagenesis. However, to date, no fully intact and infectious HERV derived retrovirus has been demonstrated in vivo. Retrovirus-like particles observed using electron microscopy in human placental trophoblasts (Kalter et al., 1973), and teratocarcinoma (Boller et al., 1993) and melanoma (Muster et al., 2003) were identified as HML-2 derived. Efforts to identify an infectious HML-2 are compounded by the fact that the large majority are partially defective and also that a somatic integration event would be a relatively rare occurrence (Bannert and Kurth, 2006). Two independent groups have succeeded in resurrecting full retroviral particles after constructing consensus sequences representing ancestors of now defunct proviruses (Dewannieux et al., 2006; Lee and Bieniasz, 2007). Although these viruses were found to be only weakly infectious, these studies will prove invaluable in formulating hypotheses regarding the potential oncogenic mechanisms of an infectious HML-2 **(Figure 5)**.

HERV-K113 and HERV-K115 are some of the most recently integrated HERVs in the human genome and represent obvious candidates for infectious proviruses. Boller and colleagues investigated this possibility and observed that HERV-K113 is able to produce fully intact retroviral particles *in vitro* (Boller et al., 2008). However, the authors concluded that an infectious HERV-K113 virus would be unlikely due to a lack of a functional reverse transcriptase.

1.9.2 HML-2 Rec and Np9 as Putative Oncogenes

Rec exhibits functional homology to the Rev protein of HIV-1, a nucleocytoplasmic shuttle protein (Bannert and Kurth, 2006). Np9 is spliced from an alternative splice donor site to Rec, and shares only 14aa with Rec and Env, with no homology to Rev (Ruggieri et al., 2009). Functional studies found that both proteins bind the promyelocytic leukaemia zinc finger (PLZF) protein, a transcriptional repressor of the *C-MYC* proto-

oncogene (Denne et al., 2007), leading to the de-repression of C-MYC. Rec also binds a related protein known as testicular zinc-finger protein (TZFP), a transcriptional repressor of the AR. Rec inhibits the ability of TZFP to repress AR transcription (Kaufmann et al., 2010). Hanke et al. identified an additional binding partner of Rec known as human small glutamine-rich tetratricopeptide repeat protein (hSGT) which also acts as a co-repressor of the AR (Hanke et al., 2013). Moreover, they proposed a 'vicious cycle' model, whereby increased cellular AR led to increased transcription at HML-2 loci and thus increased levels of Rec leading to further AR de-repression. The involvement of such hormonal regulators will be interesting to study in castration-resistant prostate cancer, in which disruption of the AR signalling axis is a key factor in development of resistance.

Reinforcing the possible importance of these proteins in tumourigenesis was the finding that mice transgenic for the *Rec* gene are prone to seminomas (Galli et al., 2005). Np9 has been shown to interact with LNX – an E3 ubiquitin ligase which targets members of the NUMB/NOTCH pathway (Armbruester et al., 2002). This pathway has been implicated in the regulation of proliferation of cancers of the breast and prostate (Roy et al., 2007). Finally, a recent study has shown that Np9 acts as a critical molecular switch for co-activating β-catenin, ERK, Akt and Notch1 and promoting the growth of human leukaemia stem/progenitor cells (Chen et al., 2013). Finally, a recent study by Heyne et al. found that Np9 can bind and inhibit MDM2 and that this inhibition supports the transactivation of genes by p53 (Heyne et al., 2015).

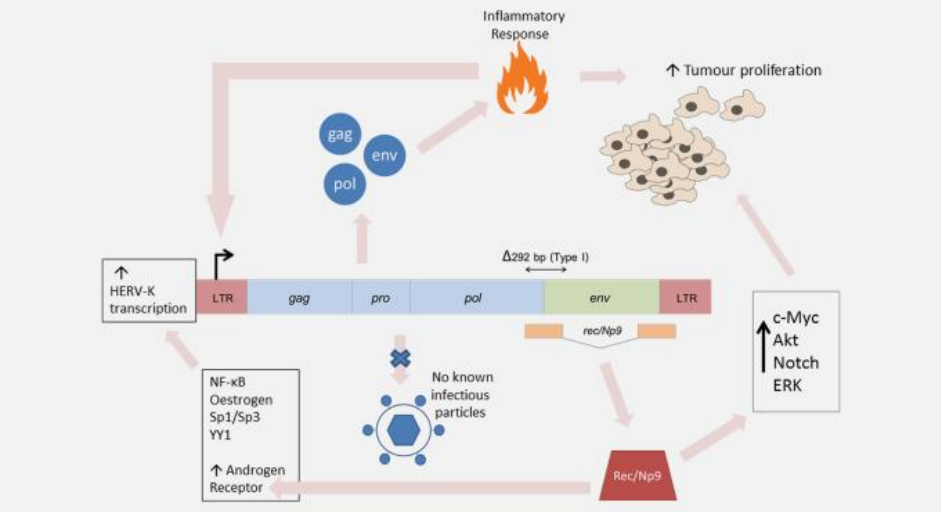


Figure 5 - Proposed model of HERV-K (HML-2)-driven cancer progression

Global DNA hypomethylation during early-stage cancer leads to activation of otherwise silenced TEs, including HERVs. A humoral response to HERV-K gag has been observed in some cancers. Such a response to high levels of HERV-K protein expression may culminate in chronic inflammation. Conversely, it has been hypothesised that HERV-K LTRs are responsive to inflammatory transcription factors—a phenomenon that may explain the high levels of HERV-K mRNA and protein seen in some inflammatory diseases. HERV-K (HML-2) accessory proteins Rec and Np9 have been shown to lead to the de-repression of the c-myc protooncogene, while Np9 has been shown to co-activate Akt, Notch and ERK pathways in leukaemia. Rec has also been observed to lead to the de-repression of the androgen receptor, which directly or indirectly causes a further increase in HERV-K transcription. Overall, the synergistic effects of chronic inflammation and dysregulated signalling/protooncogene activation caused by HERV-K protein expression may help to create a protumourigenic microenvironment culminating in further proliferation and metastasis. Finally, it is important to note that an active, infectious HML-2 provirus has not been isolated to date, but the existence of such a particle cannot be ruled out. It would potentially be oncogenic via mechanisms such as insertional mutagenesis **(Figure 5)** (Adapted from (Downey et al., 2015)).

1.9.3 HML-2 Induced Immunomodulation

In a Darwinian sense, cancerous tissue utilizes the inflammatory associated milieu of the tumour microenvironment to confer a selective advantage

(Coussens and Werb, 2002). The apparent immunogenicity of HERV proteins therefore represents a potential contributor to, or initiator of a chronic inflammatory state, beneficial to tumour survival (Manghera and Douville, 2013). HML-2 antibodies have been observed in patients with melanoma (Buscher et al., 2005), breast (Wang-Johanning et al., 2013) and ovarian cancers (Wang-Johanning et al., 2007). In breast cancer, studies have found that both humoral and cell-mediated immune responses to HERVs were enhanced in patients when compared to controls (Wang-Johanning et al., 2008). HERV-K18 Env protein has been shown to elicit T-cell responses and can be upregulated in response to EBV infection (Stauffer et al., 2001; Sutkowski et al., 2001), and may be a prerequisite of B cell lymphomas (Zur Hausen, 2009).

Similar to discoveries in HIV-1, HML-2 may encode env proteins with immunosuppressive transmembrane domains. A recent study by Morozov et al. identified an immunosuppressive HML-2 env protein which altered cytokine expression and suppressed immune cell proliferation *in vitro* (Morozov et al., 2013).

Nitric oxide (NO) is an endogenous free radical signalling molecule which has been intimately linked with inflammation, wound healing responses and cancer (Burke et al., 2013; Rapozzi et al., 2013). A significant association between nitric oxide synthase 2 (NOS2) and HML-2 Env expression has been demonstrated in breast cancer (Wang-Johanning et al., 2012). NOS2 is an independent predictor of poor outcome in estrogen receptor negative breast cancer, associated with macrophage infiltration, deregulated p53 signalling, increased proliferation and resistance to apoptosis (Ambs and Glynn, 2011; Burke et al., 2013; Glynn et al., 2010). Can HML-2 Env proteins mediate downstream inflammatory effects via their activation of NO signalling? Intriguingly, β-catenin, ERK and Akt, which are activated by Np9 (Chen et al., 2013), are also activated by NO signalling (Glynn et al., 2010; Switzer et al., 2012).

1.9.4 Exosomes

An evolving hypothesis in cancer research over the last few years has been the involvement of tumour exosomes in metastasis (Duijvesz et al., 2011; Yang and Robbins, 2011). Exosomes are nanoscale membrane vesicles which are secreted from cells and are thought to be important intercellular communicators, or, in a cancer setting, drivers of metastatic spread (Thery et al., 2002). A recent study has now implicated HERVs in this process, with the finding that HML-2 mRNA is selectively packaged into tumour exosomes and that this genetic material can be transferred to normal cells (Balaj et al., 2011). The authors also demonstrated that these exosomes were enriched for the *C-MYC* proto-oncogene, which has been shown to be regulated by PLZF, a target of Rec and Np9 (Denne et al., 2007). Therefore, it is possible that there is a link between the high levels of HML-2 mRNA and *C-MYC* in these exosomes, but further investigation will have to be done in this regard. Another important point is that HML-2 driven metastasis via exosomes would not require an envelope gene, as exosomes gain entry to target cells via a plasma membrane fusion event. In essence, exosomes could potentially empower the abundance of defective HERVs with a new-found infectivity (Gould et al., 2003).

1.9.5 HML-2 viral proteins as potential vaccines

While the direct oncogenic effects of HERVs in cancer remains to be fully elucidated, there is potential for their use as diagnostic or prognostic biomarkers and for immunotherapeutic purposes including vaccines. Independent groups have demonstrated a strong association between HML-2 antibodies and clinical manifestation of disease and therapeutic response (Contreras-Galindo et al., 2008; Golan et al., 2008). Antibodies recognizing synthetic HML-2 proteins were detected at a very low frequency in the sera of healthy donors (Wang-Johanning et al., 2007; Wang-Johanning et al., 2008). Humoral anti-HML-2 immune response may provide additional prognostic information to that of established melanoma markers (Hahn et al., 2008; Schiavetti et al., 2002). Data from these studies reveal a significant inverse correlation between serological anti-HML-2 reactivity

and patient survival probability in melanoma patients. Among the different classes of tumour antigens recognizable by the immune system, mutated self-antigens and viral antigens are unique because they are foreign to the host and not subjected to pre-existing antigen-specific tolerance (Foster and Rooney, 2006; Gottschalk et al., 2005; Smith et al., 2012). HML-2 exons coding for mature proteins are spread out over the genome and are a repository of immunogenic retroviral gene products that can be 'reawakened' when genetic damage occurs through chromosome breaks, frameshifts, and mutations, removing sequences normally silencing protein expression.

HERV-K MEL is an antigenic peptide which is encoded by a short ORF from a processed HERV-K (HML-6) pseudogene and has been shown to be recognised by cytotoxic T cells in human melanoma (Schiavetti et al., 2002). BCG, vaccinia and yellow fever vaccinations are associated with a reduced risk of developing melanoma (Krone et al., 2003; Mastrangelo et al., 2009; Pfahlberg et al., 2002), although conflicting data exists for yellow fever vaccines (Hodges-Vazquez et al., 2012). It is suggested that this effect is due to antigen sequence homology between these vaccines and HERV-K-MEL leading to cross-reaction between vaccine-elicited cytotoxic T cells and melanoma cells (Krone et al., 2005). Reintroduction of these vaccines has been suggested as a novel method of melanoma immunoprevention; otherwise, HERV-K MEL represents a legitimate target for cellular immunotherapy (Krone and Grange, 2010; Krone et al., 2005).

1.10 Future Perspectives

Over the course of evolution, our genome has been locked in a molecular 'war' with exogenous infectious agents. Ultimately it is this very battleground, together with viral endogenisation, which has bestowed upon us the diverse genetic repertoire we possess today. Constituting 8% of our genome, these HERVs have supplied us with an additional layer of plasticity and physiological functionality, yet scientists now believe that

hidden detrimental processes fuelled by HERVs may be present, which are inducing chronic diseases such as cancer and autoimmunity. As of yet, no truly infectious HERVs have been observed. However, as outlined in this review, a range of potential molecular mechanisms involving the retroviral proteins may be aiding and abetting both tumour formation and metastasis. Ultimately, it is likely that many of these mechanisms are working synergistically to produce these effects, and the heralding of a single molecular event induced by a HERV protein is improbable.

Ascribing a causative role for a particular agent to a disease has long been a challenging task. Criteria such as Hills criteria (Hill, 1965) and Koch's postulates (Evans, 1976) have been formulated to address this problem. These criteria have recently been refined and built upon in light of HERVs postulated role in human disease (Cegolon et al., 2013; Krone and Grange, 2010; Sarid and Gao, 2011). However, even if a direct link between HERVs and carcinogenesis is never established, their presence may be highly advantageous in terms of the implementation of novel biomarkers for cancer. Further work will need to effectively correlate their presence with various disease stages and also make the necessary comparisons against 'gold-standard' biomarkers. Equally promising, is the potential to take advantage of tumour specific HERV expression for the use of targeted immunotherapies. Wang-Johanning et al. have demonstrated the potential of anti-HML-2-Env antibodies in inhibiting tumour growth and inducing apoptosis, both *in vitro* and in *in vivo* mouse xenograft models (Wang-Johanning et al., 2012). This work represents a major milestone in research into HERVs and cancer and it is likely that targeting Env in a similar fashion in other cancers will be equally effective. However, it remains imperative that these studies are evaluated in a clinical setting. Additionally, it may also be possible to conjugate these antibodies to cytotoxic drugs for increased effect (Carter, 2006). Similarly, Kraus et al demonstrated that HML-2-Env targeted vaccine reduced renal tumour metastasis in a murine model (Kraus et al., 2013). Novel therapies, such as these, are key to making inroads towards a future cure for the increasingly complex and

multistep disease that is cancer. However their safety must be assessed given the newly established role of HML-2 in embryonic stem cells and iPS cells (Fuchs et al., 2013), which may have implications for pregnancy. Their role in adult stem cells is not currently known.

Several limitations exist in the field of cancer-related HML-2 research, including a lack of adequately powered patient population studies to determine the role of HML-2 in the etiology of cancer, and/or its association with metastasis, therapeutic response and overall patient survival. A gap exists in our knowledge as to which HML-2 loci are specifically activated in cancer. A recent study by Schmitt et al. has defined the HML-2 transcriptome in melanoma, identifying 23 loci as transcribed (Schmitt et al., 2013a), and it is an imperative that similar studies be initiated in other cancers. A causal role for HML-2 has yet to be established. Generally retroviruses induce tumours by insertional mutagenesis targeting specific oncogenes, as is the case with HBV (Guerrieri et al., 2013). This is an unlikely mechanism though in the case of HML-2. Evidence does suggest that Rec and Np9 may be putative oncogenes, but whether Gag or Env are also oncogenic is not known. In exceptional cases such as Jagsiekte sheep retrovirus (JSRV) the Env protein has been found to be causal (ovine pulmonary adenocarcinoma) (Hofacre and Fan, 2010). However, it is unlikely that HML-2 Gag or Env have a similar causal effect, potentially they may influence carcinogenesis by activating or perturbing inflammation responses against cancer.

It is our belief that within the next decade these genetic ‘squatters’ will have firmly established themselves within the modern multistep model of cancer progression and their expression will be viewed as an ‘enabling characteristic’ of cancer, giving new meaning to the famous words of Nobel laureate J. Michael Bishop when he stated that "the seeds of cancer are within us” (Conti, 2008).

1.11 Overview of prostate cancer

1.11.1 Prostate function

The prostate is a walnut sized glandular organ located in the pelvic cavity and is held together by a muscular stroma consisting of fibroblast cells (Standring and Gray, 2015). The main function of the prostate is to produce prostatic fluid which is an essential component of semen. The glandular epithelium consists of three main cell types: basal cells, secretory luminal cells and neuroendocrine cells (**Figure 6**) (Czyz et al., 2012). The existence of prostate cancer stem cells is a controversial topic but some markers such as CD133 have been suggested (Collins and Maitland, 2006). The luminal cells secrete prostatic fluid into the lumen of the gland helped via contractions of the smooth muscle stromal cells.

1.11.2 Prostate cancer incidence/mortality

Prostate cancer is the second most common cause of cancer in men in Ireland with over 3000 men being diagnosed each year. Mortality rates of prostate cancer in Ireland are currently very good with five year survival rates currently standing at 91% (National Cancer Registry Ireland).

1.11.3 Prostate cancer epidemiology

Increasing age is the main risk factor associated with prostate cancer (Gronberg, 2003). Other risk factors include, diet and obesity. It is thought that a link between increasing age and prostate cancer may be due to a buildup of inflammatory lesions as people age (De Marzo et al., 2007). Many inflammatory lesions of the prostate including PIN and PIA have been linked to prostate cancer. As one ages, the chances of these lesions developing further molecular changes increases and this may lead to prostate cancer. Other genetic changes which increase with age such as telomere shortening may also contribute to disease development (Meeker et al., 2002).

While cigarette smoking may not influence early disease development (Giovannucci et al., 1999; Hickey et al., 2001), it increases the risk of fatal prostate cancer (Giovannucci et al., 1999; Huncharek et al., 2010; Kenfield

et al., 2011)This observation was replicated in studies showing that current smokers develop distant metastasis more frequently than non-smokers (Kobrinsky et al., 2003; Roberts et al., 2003). Because smoking cessation reduces the risk of metastasis (Giovannucci et al., 1999; Kenfield et al., 2011) a tobacco-related factor appears to induce reversible molecular alterations in the cancerous prostate that facilitate metastatic spread.

Obesity is known to affect the regulation of hormones in the body (Kirschner et al., 1981). The main male hormone in the body is androgen and thus obesity leading to androgen dysregulation may cause an increase in the risk of prostate cancer (Hsing et al., 2007).

Certain diets have also been linked to prostate cancer (Gronberg, 2003). This link came from the observation that when east Asian men – who have low rates of prostate cancer – moved to the USA, it was found that their rates of prostate cancer increased to levels seen in American men. The mechanistic link in this case is thought to be diets high in red meats which are known to cause inflammation via the production of compounds such as polyaromatic hydrocarbons which are formed when red meats are cooked at high temperatures (De Marzo et al., 2007). Foods which are thought to prevent prostate cancer include garlic and tomatoes which include compounds which may prevent DNA damage in response to inflammation (Itsiopoulos et al., 2009).

1.11.4 Prostate pathobiology

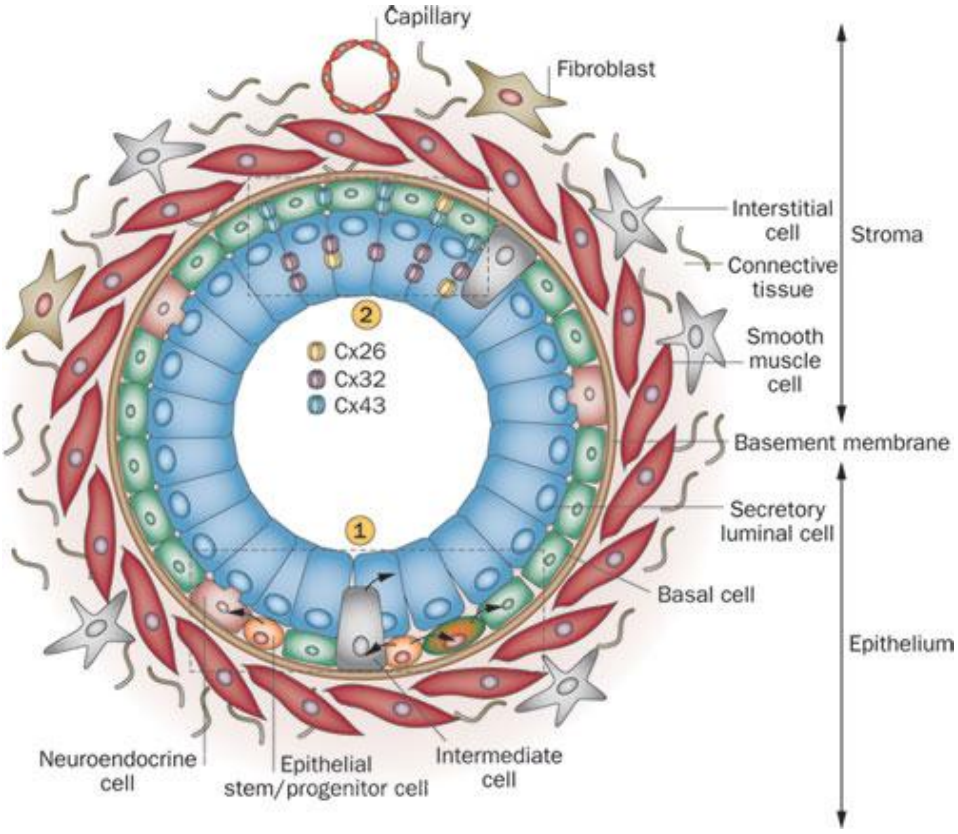


Figure 6 - Cell types of the prostate

The prostate is composed of a glandular epithelium surrounded by stroma. The epithelium is itself composed of secretory luminal cells, basal cells, neuroendocrine cells, intermediate cells and stem/progenitor cells. The stroma is composed of interstitial cells, smooth muscle cells, fibroblasts and connective tissue. (adapted from (Czyz et al., 2012))

1.11.4.1 BPH

Benign prostatic hyperplasia (BPH) is a common non malignant pathology of the prostate usually seen in older men and thought to be caused by androgen dysregulation (Thorpe and Neal, 2003). On average, 8% of 40 year old men will have BPH with this increasing to 90 % for men greater than 80 years old. Men with this disease clinically present with lower urinary tract symptoms (LUTS) such as urinary filling or voiding symptoms and this can progress to acute urinary retention. These are thought to be due to the obstruction of the bladder caused by an enlargement of the prostate.

Increasing age is the main risk factor for BPH probably because of age-related fibrosis of prostate tissue, more specifically, myofiber injury. The myofibers are a key component of the prostate stroma and aid in the movement of prostatic fluid into the lumen of the gland (Standring and Gray, 2015). These fibers have a low regeneration rate and thus are at increased risk of age-related fibrosis. This fibrosis is thought to lead to the enlargement of the prostate gland. Androgens are known to impact the development of both diseases. Evidence for the role of androgens in BPH development comes firstly from the observation that castrated boys do not develop BPH in later life (Guess, 1992) and secondly the positive effect that 5-alpha-reductase inhibitors have on the treatment of the disease (Roehrborn et al., 2002). These antiandrogens prevent the conversion of testosterone to the more potent androgen – DHT. The second common treatment for BPH is alpha blockers which are antagonistic drugs of adrenergic receptors. Other surgical interventions are common which attempt to deal with the LUTS associated with BPH. These include transurethral resection of the prostate (TURP) and transurethral microwave therapy (TUMT). Although, there is no known link between BPH and prostate cancer, BPH can cause PSA levels to rise which leads to false positives in the diagnosis of prostate adenocarcinoma (Thorpe and Neal, 2003).

1.11.4.2 PIA

Another inflammatory lesion which may lead to prostate adenocarcinoma is proliferative inflammatory atrophy (PIA) (De Marzo et al., 2007). These lesions are most commonly located in the peripheral zone of the prostate and their presence increases with advancing age. Cells in this lesion exhibit increased expression of genes such as *GSTP1* and *p16* and downregulation of *PTEN* and *CDKN1B*. Molecular changes such as these can lead to genetic instability which may be a precursor to prostate adenocarcinoma (Abate-Shen and Shen, 2000). Proliferation of prostate cells in the setting of chronic inflammation is thought to lead to PIA (De Marzo et al., 2007). Chronic inflammation in prostate cancer has been linked to infectious agents and dietary carcinogens amongst others.

1.11.4.3 PIN

Prostatic intraepithelial neoplasia (PIN) is an inflammatory lesion of the prostate. Men with prostate cancer are statistically more likely to present with PIN (De Marzo et al., 2007). Transitions from PIA to PIN have been observed in various zones of the prostate. Genes which exhibit increased expression in PIN include *Bcl-2* and *HER2/neu*. PIN can be defined as low-grade or high-grade PIN depending on the amount of cellular irregularity, for example, PIN cells have shortened telomeres compared to normal prostate cancer cells and this together with methylation of *GSTP1* is linked to the transition from low grade to high grade PIN (**Figure 7**) (Abate-Shen and Shen, 2000).

1.11.4.4 Prostate adenocarcinoma

Why should such a nondescript organ with a low rate of cellular turnover be so susceptible to cancer? The answer most likely lies in the dysregulation of hormones during older age (Feldman et al., 2002). However, the study of human cadavers has shown evidence for cancerous lesions in the prostate in men from the 3rd decade of life onwards (Sanchez-Chapado et al., 2003). Even so, increasing age remains the number one risk factor for prostate cancer (Gronberg, 2003). From an evolutionary perspective, prostate cancer is a post reproductive disease and thus the mechanisms behind its pathobiology have not been purified by evolutionary selection to a great extent (Greaves, 2007).

Prostate cancer is driven by androgens (Nelson et al., 2003). An androgen is defined as any molecule which can bind to the nuclear receptor known as the androgen receptor (AR). The activated AR itself binds to androgen response elements (AREs) which are nucleotide sequences present in the promoters of androgen inducible genes.

Over time, most prostate adenocarcinomas will eventually become resistant to androgen deprivation treatment, turning into castrate resistant prostate cancer (CRPC) (Damber and Aus, 2008). CRPC is the main cause of prostate cancer mortality. There are a number of molecular mechanisms by

which CRPC is thought to form (Feldman and Feldman, 2001). One of these is the hypersensitivity of the AR through gene amplification. Another - the 'outlaw pathway' hypothesis - states that the AR becomes activated through phosphorylation which is independent of ligand binding. Finally, it is thought that the spectrum of AR ligands can broaden to include non androgen molecules. This is known as the 'promiscuous pathway'.

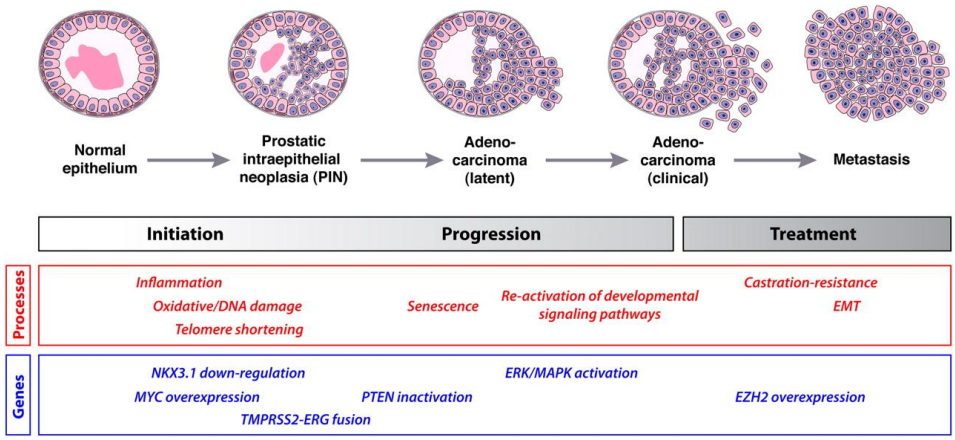


Figure 7 - Molecular progression of prostate cancer

Inflammatory lesions such as PIA and PIN are thought to be precursors to prostate adenocarcinoma. Most early prostate cancers are androgen dependent but the majority progress to an androgen independent state otherwise known as castrate resistant prostate cancer (CRPC) and metastasis. Various molecular pathways have been delineated for the different stages of the disease. (adapted from (Shen and Abate-Shen, 2010))

1.11.5 Prostate cancer biomarkers

1.11.5.1 PSA

Prostate specific antigen is the gold standard biomarker for prostate cancer (Makarov et al., 2009). It is a protein which is specifically expressed in the prostate. It remains an imperfect biomarker due mainly to the false positives which benign diseases such as BPH cause. A PSA score of greater than 4ng/ml is generally regarded as a cutoff point. Over this threshold and a patient will be asked to undergo a digital rectal examination (DRE) or a TRUS-guided biopsy. Histological investigation of this biopsy will confirm the presence or absence of prostate cancer.

1.11.5.2 Other biomarkers

Other promising biomarkers for prostate cancer include early prostate cancer antigen (EPCA) and EPCA2 (Makarov et al., 2009). These nuclear matrix proteins were first observed to change in conjunction with nuclear pleiomorphism seen in early prostate cancer. Since then they have successfully been used as histological markers during immunostaining experiments to help differentiate between organ and non organ confined prostate cancer. Finally, methylation levels of the glutathione s-transferase (*GSTP1*) gene have been linked to the presence or absence of prostate cancer, with hypermethylation seen in prostate cancer cases and hypomethylation observed in BPH and normal prostate tissue (Cairns et al., 2001).

1.11.5.3 The need for new biomarkers

There is an urgent need for new biomarkers in prostate cancer (Makarov et al., 2009). The low specificity of PSA is a major cause of the misdiagnosis of prostate cancer and this causes a great deal of stress for men, who in actual fact only have a benign disease but who will have to now undergo invasive biopsies in order to confirm this.

The blood is an ideal component of the body for non invasive disease testing. This is especially the case for prostate cancer whose diagnosis

usually entails the use of TRUS-guided biopsies or DREs (Nelson et al., 2003). Circulating tumour cells (CTCs) are cancer cells which circulate in the bloodstream and may offer a novel method of prostate cancer diagnosis (Moreno et al., 2005). They are, however, very difficult to detect, requiring special filters in which they can be separated from blood cells. They are also extremely rare, possibly with only 5-10 cells being present per ml of blood. If technological advances can improve how CTCs are detected then they may well represent a new generation of biomarkers for the disease.

The sentinel hypothesis postulates that the transcriptome of circulating normal blood cells can be an indication as to ones state of health or disease (Mohr and Liew, 2007). Blood cells circulate around the whole body and upon interacting with diseased tissue, can be affected by these tissues release of factors such as cytokines, which then influence the transcriptome of the blood cell. In the case of solid tumours such as prostate cancer, this would allow one to non invasively test the blood as a surrogate marker for the disease and would alleviate the need for invasive testing of the solid tumour.

1.11.6 Prostate cancer therapeutics

1.11.6.1 Provenge

Sipuleucel-T (Provenge) was the first immunotherapy approved by the FDA for any type of cancer (Higano et al., 2010). It is essentially an autologous T-cell transplant. Patient's T-cells are harvested and grown *ex vivo* in the presence of PMSA which is a marker specific to prostate cancer cells. The T-cells become educated and are transplanted back into the patient where they attack prostate cancer cells.

1.11.6.2 Hormone therapy

Since prostate cancer is essentially hormonally driven, hormone therapy is a widely used treatment for the disease (Damber and Aus, 2008). Surgery can be used as a form of physical castration to remove the testicles. This is known as an orchidectomy but it is a now a rarely performed procedure (Hellerstedt and Pienta, 2002). Chemical castration involves the use of

drugs to lower or ablate levels of androgens. The four main types are: luteinising hormone (LH) blockers such as histrelin, gonadotrophin releasing hormone (GnRH) blockers such as degarelix, anti- androgens such as enzalutimide and finally abiraterone. Anti- androgens are broadly broken down into first generation and second generation anti-androgens. Bicalutamide is an example of a first generation anti-androgen. Second generation versions of these drugs were created in order to circumvent the problems associated with partial agonism seen when using these drugs in engineered cells expressing high levels of the AR. An example is enzalutumide which not only acts as an AR antagonist but also prevents AR signalling by preventing the translocation of the AR to the nucleus (Jung et al., 2010). Finally abiraterone acetate is a drug which targets the cytochromes responsible for the critical enzymatic conversions needed to form androgen molecules (Logothetis et al., 2011). Without these catalysts androgen levels are drastically reduced.

1.11.6.3 Radiation

Two main types of radiotherapy for prostate cancer exist: external beam and brachytherapy (Duchesne, 2001). External beam radiation is the main type of radiotherapy used for prostate cancer and consists of using a machine known as a linear accelerator to fire x-rays at the cancerous tissue. The main disadvantage to external beam radiation is that it will affect normal non cancerous cells which are in close vicinity to the tumour. This results in high levels of side effects. Brachytherapy is a radiotherapy which involves the insertion of radioactive seeds into the prostate and is a more targeted form of radiation therapy when compared to external beam resulting in fewer side effects. As prostate cancer is usually a multi focal disease it thus represents an ideal platform in which to spatially target the various foci of adenocarcinoma (Nelson et al., 2003). The main advantage of brachytherapy over surgery is that the side effects are greatly reduced.

1.11.6.4 Chemotherapy

Chemotherapy for prostate cancer is normally only used when hormone treatment for metastatic prostate cancer has failed (Gilligan and Kantoff,

2002). Double chemotherapy or chemotherapy combined with radiation therapy is also sometimes used. Examples of prostate cancer chemotherapeutics include docetaxol and cabazitaxel.

1.11.6.5 Surgery

Radical prostatectomy involves the complete removal of the prostate (Damber and Aus, 2008). The procedure results in many major side effects including incontinence and impotence. The two major types of prostatectomy are robot-assisted radical prostatectomy (RARP) and open radical prostatectomy (ORP). Some debate remains as to which is the superior method, but a recent study has concluded that both techniques have similar outcomes with similar rates of complications (Gandaglia et al., 2014).

1.11.6.6 Watchful waiting

Many prostate cancers will remain indolent throughout the life of a patient (Nelson et al., 2003). Invasive treatments such as radical prostatectomy or radiation therapy lead to harmful side effects which may have never been needed in the first place. Watchful waiting entails keeping a patients treatment in reserve and closely monitoring the disease for indications of progression after which the treatment regime may be started (Bill-Axelson et al., 2011).

1.12 HML-2 and prostate cancer

Retrovirus-like particles and the expression of HML-2 mRNA and proteins are detectable in prostate cancer tissues. Ishida et al. (2008) isolated a HML-2 Gag protein in the serum of a prostate cancer patient using serological recombinant cDNA expression cloning (SEREX) technology (Ishida et al., 2008). They subsequently detected HML-2 *gag* mRNA in the serum of 6 of 9 prostate cancer patients, but failed to detect HML-2 *gag* mRNA in LNCap, DU145 or PC3 prostate cancer cells (Ishida et al., 2008). Gene fusions are a frequent occurrence in prostate cancer, the majority of which involve the fusion of the transcription factors ETS translocation variant (ETV1) or ETS related gene (ERG1), to the transmembrane protease, serine 2 (TMPRSS2). In these fusions the androgen responsive TMPRSS2 drives

expression of the ETV1 or ERG1 oncogenes. Recently ETV1-HML-2 fusions have been described, corresponding to the 5'-untranslated region (UTR) of HERV-K-22q11.23 (Tomlins et al., 2007) and HERV-K17 (Hermans et al., 2008). Additionally the ETV1-HERV-K-22q11.23 fusion is also inducible in LNCap in response to androgen (Tomlins et al., 2007), similar to HML-2 induction by estrogen and progesterone in breast cancer cell lines (Ono et al., 1987).

Goering et al., detected significant expression of HERV-K-22q11.23 and HERV-K17 in the androgen-responsive prostate cancer cell lines 22Rv1, LNCap and MDA-PCa-2b (Goering et al., 2011). Normal prostate cells and androgen insensitive prostate cancer cells (PC-3, DU-145 and BPH-1) exhibited expression near the limit of detection (Goering et al., 2011). Expression of two other proviruses HERV-K-11q23.3 and HERV-K-22q.11.21 were not detectable in prostate cancer cell lines. Assessing HERV-K-22q11.23 5'UTR-*gag, env* and *Np9* gene expression in prostate tumours (n=45) versus benign tissue (n=11), the expression of the 5'UTR-*gag* and *env* region was significantly elevated in tumours compared to benign tissues. *Np9* was detectable only in a subset of carcinomas (18/45). In contrast HERV-K17 was reduced in prostate tumours compared to benign. Where HERV-K-22q11.23 and HERV-K17 were expressed, they correlated with prostate specific antigen (PSA) levels, suggesting that HERV-K-22q11.23 and HERV-K17 retroelements are under androgen inducible control, while HERV-K-11q23.3 and HERV-K-22q.11.21 are not (Goering et al., 2011). Agoni et al. published two such papers describing HML-2 expression in prostate cancer. The first paper investigated the expression of HML-2 mRNA in prostate cancer cells lines (Agoni et al., 2013a). They found that four prostate cancer cell lines were positive for full length and spliced HML-2 mRNA. They also conducted an analysis into the exact loci which were giving rise to the HML-2 transcriptional signal. They concluded that the HML-2 signal was varied in that multiple loci were being expressed in each cell line and that no particular expression pattern was evident across all the lines. Interestingly, they also discovered the presence

of antisense HML-2 transcripts. The same group followed up this work by publishing a study which described the presence of alternatively spliced HML-2 mRNAs in prostate cancer, together with data which indicated that HML-2 expression in prostate cancer was upregulated by ionizing radiation (Agoni et al., 2013b). Another two publications originated from the Goering group. The first found evidence of DNA hypomethylation at the 22q11.23 HML-2 locus (Kreimer et al., 2013), while the second looked at the loci responsible for HML-2 signal in prostate cancer (Goering et al., 2015). Finally, Reis et al. published a study which found that prostate cancer progression correlates with an immune response to a HML-2 gag protein from the 22q11.23 locus (Reis et al., 2013).

1.13 Aims

The aims of this thesis are:

1. To define HML-2 expression in prostate cancer

   Little is known about HML-2 expression in prostate cancer. This thesis will investigate basal HML-2 levels in prostate cancer cell lines and in clinical samples using techniques including western blotting, qPCR and immunohistochemistry. Potential activators of HML-2 in prostate cancer will also be investigated using treatments with steroid hormones

2. To investigate the potential of HML-2 as a biomarker in prostate cancer

   HML-2 expression is observed in many types of cancer and may represent a novel biomarker. This thesis will explore the utility of using HML-2 expression as a biomarker for prostate cancer by conducting an epidemiological study which will look at differences in HML-2 mRNA expression between prostate cancer cases and healthy controls.

3. To investigate if HML-2 plays any functional role in prostate cancer initiation or progression

   This thesis will investigate whether HML-2 expression in prostate cancer is simply an epiphenomenon or if it actually promotes tumour development and proliferation. In order to do so, HML-2 Rec will be cloned and overexpressed in non tumourigenic RWPE-1.

Chapter 2

Materials and Methods

2 Materials and Methods

2.1 Reverse transcription and qPCR analysis of prostate cancer cell line HML-2 expression

Total RNA was isolated using the TRIzol reagent (Thermo Fisher Scientific) according to the manufacturer's instructions. RNA purity and concentration was determined using a NanoDrop ND-1000 Spectrophotometer (Nano-Drop Technologies). RNA was treated with TURBO DNAse (Ambion) for 30 minutes at 37 °C to eliminate genomic DNA. DNase was inactivated using 50 mM EDTA at 75 °C for 10 minutes. First strand cDNA synthesis was performed on 1 µg of RNA using a Tetro cDNA Synthesis Kit (Bioline) primed with oligo(dT) primers. qPCR took place in a StepOne Plus real-time PCR system (Applied Biosystems) together with Sensifast SYBR Hi-ROX (Bioline) and the relevant primers as shown in

Table 1.

Thermal cycling consisted of enzyme activation: 95 °C for 2 minutes, followed by 40 cycles of both denaturation and annealing/extension: 95 °C for 5 seconds and 60 °C for 15 seconds respectively. Subsequent melt curve analysis was carried out using the following conditions: 95 °C for 15 minutes, 60 °C for 1 minute and 95 °C for 15 minutes. Gene expression levels in all samples were normalised to an *ACTB* reference gene (RefSeq accession number NM_001101.3) using the delta Ct method and values were displayed as the mean of triplicate samples. Statistical significance of differences between cases and controls was calculated using Student's t-test in GraphPad Prism version 5 (GraphPad Software).

Error bars for qPCR experiments represent the RQ min and RQ max for each sample. Statistical differences were measured by calculating the standard deviation of ct values from three technical replicates.

Table 1 - HML-2 qPCR primers

Gag forward	5'-GTAAAAACAGTATGGACCCAACTC -3'
Gag reverse	5'-CTATTTCTTCGGACCTGTTCTTG -3'
Env1 forward	5'-GGAGATGGTAACACCAGTCACAT-3'
Env1 reverse	5'-GGATAACGATACCCAATGGAAAT-3'
Env2 forward	5'-CAAAATGGTGACGTCAGAAGAA-3'
Env2 reverse	5'-CAGGCATAGGGAGACTTACCAC-3'
Rec forward	5'-TGACACAAACCCCAGAGAGTATG-3'
Rec reverse	5'-ACACCTGCAGACACCATTGATACA-3'
K7	5'- TCTGTTTAACAAAGCACATCCTGC-3'
K21	5'- ATGAACCCATCAGAGATGCAAAG-3'
ACTB forward	5'-CATGTACGTTGCTATCCAGGC-3'
ACTB reverse	5'-CTCCTTAATGTCACGCACGAT-3'

2.2 RT-PCR of Rec and Np9 from prostate cancer cell lines

RNA was treated with TURBO DNAse (Ambion) for 30 minutes at 37 °C to eliminate genomic DNA. DNase was inactivated using 50 mM EDTA at 75 °C for 10 minutes. One step RT-PCR was performed on 1 µg of RNA using a Tetro one step RT-PCR kit (Qiagen) and primers K21 ATG AAC CCA TCA GAG ATG CAA AG and K7: TCT GTT TAA CAA AGC ACA TCC TGC. Thermal cycling consisted of: reverse transcription 50 °C for 30 minutes; enzyme activation: 95 °C for 15 minutes; 30 cycles of denaturation 94 °C, annealing 55 °C and extension 72 °C; followed by a final extension step of 10 minutes at 72 °C. Amplification of PGK1 was conducted as a loading control and was independently carried out using the same RNA with primers: PGK1 forward 5'-TGGACGTTAAAGGGAAGCGG-3', PGK1 reverse 5'-GCTCATAAGGACTACCGACTTGG-3'. Bands were visualised on 2 % agarose gel stained with GelRed (Biotium).

2.3 Subject recruitment and collection of PBMC

Selection criteria for cases (n = 270) and population-based controls (n = 91) that participated in the National Cancer Institute (NCI) study: prostate

cancer patients were eligible for the study when a diagnosis of prostate cancer has been made within 2 years prior to recruitment; resided in Maryland and adjacent counties in Pennsylvania, Delaware, Virginia, or District of Columbia, if they were born in the USA; were either African-American or European-American by self-report; had a working home phone number; were physically and mentally able of performing the interview; were not severely ill; spoke English fluently and were able to give informed consent and did not reside in an institution such as a prison, nursing home or shelter. Male population controls were frequency matched on age and race to cases and had the same eligibility criteria with the exception that they could not have a personal history of cancer, radiation therapy or chemotherapy. Controls resided in the greater Baltimore area and adjacent counties in Maryland. Selection criteria for cases (n = 23) and hospital-based controls (n = 44) that participated in the Georgetown study: Prostate cancer patients were eligible for the study when a diagnosis of prostate cancer has been made by the attending physician, and these patients resided in the District of Columbia or its adjacent states of Maryland or Virginia. Patients were self-reported as African-American and were physically and mentally capable of providing informed consent. Male hospital controls were frequency matched on age to cases and had the same eligibility criteria as cases with the exception that they could not have a personal history of cancer. No significant difference was found between the age at diagnosis/recruitment in Georgetown Cohort versus the NIH cohort (P = 0.29) overall or when comparing African-Americans alone (P = 0.87). Blood was drawn at time of recruitment. PBMC were collected from both prostate cancer patients (n = 294) and men without a diagnosis of the disease (n = 135). The cells were isolated from whole blood by standard ficoll–hypaque density gradient centrifugation and stored at −80°C. Men with prostate cancer were recruited between 2004 and 2008 under two Institutional Review Board (IRB)-approved protocols (NCI IRB #05-C-N021 and Georgetown University IRB #2003–013) and had a prostate cancer diagnosis within the last 2 years prior to recruitment (median time between diagnosis and recruitment = 206 days, range 0–714 days). These

patients had prostate cancer at time of recruitment and came to the hospital for consultation or to seek treatment including prostatectomy, radiation therapy, or androgen ablation therapy. The subjects were recruited at four hospitals: the Veterans Affairs Medical Center and the University of Maryland Medical Center in Baltimore City, the Department of Urology at the Georgetown University Hospital and the Washington DC Veterans Affairs Medical Center. All completed an informed consent. Controls were either population-based controls (n = 91) recruited under the NCI IRB approved protocol #05-C-N021 (NCT00342771) (Hudson et al., 2012), or they were men without a previous cancer diagnosis (by self-report) visiting the Georgetown University Hospital (n = 44), accompanying other people or coming for a routine checkup. The latter were recruited under the Georgetown University IRB-approved protocol #2003–013. All controls completed an informed consent. Prostate-specific antigen (PSA) test results were available for 287 of the 294 prostate cancer patients; they were not available for the controls. Both cases and controls completed interviewer-administered questionnaires but only the NCI-based study collected information on smoking history from study participants. Information on smoking was available for 359 subjects (270 cases, 89 controls) and was categorised into current, former, and current smokers, or into pack-years. A never smoker was defined as a subject who did not currently smoke and also smoked <100 cigarettes in his lifetime. A past smoker did not smoke cigarettes in the 6 months prior to enrolment. Race/ethnicity was self-reported.

2.4 RNA Isolation from PBMC and Detection of HML-2 *gag* mRNA.

Total RNA was isolated using the TRIZOL reagent according to the manufacturer's instructions. 500 ng of RNA was reverse transcribed and the cDNA was added to the qRT-PCR assays. Previously published primers were used to amplify HML-2 *gag* transcripts (F, 5'-AGC AGG TCA GGT GCC TGT AAC ATT-3'; R, 5'-TGG TGC CGT AGG ATT AAG TCT

CCT-3') (Contreras-Galindo et al., 2008). In addition, primers specific for 18s were used as an internal standard reference. Data were collected using the ABI PRISM® 7500 Sequence Detection System. Normalised expression was calculated using the comparative C_T method and fold changes were derived from the $2^{-\Delta\Delta Ct}$ values (Bookout and Mangelsdorf, 2003).

2.5 Detection of HML-2 Env type I and Env type II transcripts from PBMC in a subset of cases and controls

RNA was treated with TURBO DNAse (Ambion, Biosciences, Ireland) for 30 minutes at 37 °C to eliminate genomic DNA. DNase was inactivated using 50 mM EDTA at 75 °C for 10 minutes. First strand cDNA synthesis was performed on 10 ng of RNA using a Tetro cDNA Synthesis Kit (Bioline, MyBio, Ireland) primed with random hexamers. qPCR took place in a StepOne Plus real-time PCR system (Applied Biosystems, Biosciences, Ireland) together with Sensifast SYBR Hi-ROX (Bioline, MyBio, Ireland) and the following primers: *env1* forward, 5'-GGAGATGGTAACACCAGTCACAT-3'; *env1* reverse, 5'-GGATAACGATACCCAATGGAAAT-3'; *env2* forward, 5'-CAAAATGGTGACGTCAGAAGAA-3'; *env2* reverse, 5'-CAGGCATAGGGAGACTTACCAC-3'.Thermal cycling consisted of enzyme activation: 95 °C for 2 minutes, followed by 40 cycles of both denaturation and annealing/extension: 95 °C for 5 seconds and 60 °C for 15 seconds respectively. Subsequent melt curve analysis was carried out using the following conditions: 95 °C for 15 minutes, 60 °C for 1 minute and 95 °C for 15 minutes. Gene expression levels in all samples were normalised to an *18S* rRNA reference gene (RefSeq accession number NR_003286) using the delta C_t method. Values were displayed as the mean of duplicate samples.

2.6 Statistical analysis: PBMC study

Data analysis was performed using the Stata/SE 11 (Stata Corp, College Station, TX) and GraphPad Prism 5 (GraphPad Software, San Diego, CA) statistical software packages. All statistical tests were two-sided. $P < 0.05$ was considered statistically significant. The Fisher's exact test, univariate, and multivariable logistic regression were used to analyse dichotomised data and to calculate the associated odds ratios (ORs). The multivariable models adjusted for age at diagnosis and race/ethnicity. The Mann Whitney test was used to compare the differences of both plasma cytokine levels and HML-2 Ct-based expression values between groups.

2.7 Immunohistochemistry of prostate tissue microarrays and prostate tumour samples

Slides with sections from formalin fixed paraffin embedded prostate adenocarcinoma samples and prostate tissue microarrays were deparaffinised twice for 10 minutes at 60 °C for a minimum of 20 minutes each. Slides were hydrated in a series of decreasing ethanol concentrations, as follows: three times in 100% for 3 minutes each, once in 95% for 3 minutes and once in 70% for 3 minutes. Slides were then rinsed in distilled water for 10 minutes. Heat-induced epitope retrieval was performed using Tris EDTA buffer at pH 9 in a microwave at 95 °C for 20 minutes. Slides were subsequently washed three times in PBS for 5 minutes each, followed by immersion in H_20_2 (30%w/v) in the dark for 30 minutes to quench endogenous peroxidise. Three more PBS washes were completed before blocking with 3% BSA in PBS for 1 hour at room temperature. Slides were then incubated overnight with anti-HML-2 gag mouse monoclonal antibody (LifeSpan Biosciences Inc, Seattle, Washington, USA #LS-C65287) diluted in 1:50 in 3% BSA in PBS in a humidified chamber. Slides were washed a further three times with PBS and then incubated with the avidin/biotin based Vectastain ABC System (Vector Laboratories) with subsequent colorimetric development using diaminobenzidine (ImmPACT DAB; Vector laboratories). Slides were counterstained with haematoxylin, mounted with

DPX mounting media, and scanned using an Olympus VS120 Digital Scanner.

2.8 Western Blot analysis of HML-2 proteins from prostate cancer cell lines

Cells were seeded in 10 cm^3 dishes at a cell density of $1X10^6$ per dish, and grown for 3 days. Cells were rinsed twice with cold PBS and lysed directly on the dish with cold RIPA buffer (Thermo-Scientific Pierce, Ireland #89900) supplemented with protease inhibitors (Thermo-Scientific Pierce, Ireland, #78410), scraped, and spun at 14,000 g for 15 min at 4 °C. Supernatant was collected, quantified using a BCA kitand stored at –20 °C for Western blot analysis of protein expression. Proteins were separated by SDS-PAGE using 8% gels and transferred onto nitrocellulose membranes using the iBlot™ Dry Blotting System (Life Technologies) HML-2 Env and HML-2 Gag levels were detected through use of primary anti-HML-2 env mouse monoclonal antibody (clone 6H5, Dr. Feng Wang-Johannning) and anti-HML-2 gag mouse monoclonal antibody (LifeSpan Biosciences Inc, Seattle, Washington, USA #LS-C65287), respectively. Rec protein expression was detected using a rabbit anti-Rec polyclonal antibody (Mueller-Lantzsch group, Germany). Antibodies were diluted 1:1000 in 5% skimmed milk reconstituted in 1X tris-buffered saline (TBS) (pH8) 0.1% Tween. These dilutions were added to the transfer membrane, and shaken overnight at 4 °C, following a 1 hour room temperature blocking in 5% skimmed milk in TBS. Mouse monoclonal anti-β-actin antibody (Thermo-Scientific Pierce, Ireland #10624754) was used to confirm even protein loading. Secondary antibodies used were IRDye 800CW Goat Anti-Rabbit IgG (LI-COR Biosciences) and IRDye 680LT Goat Anti-Mouse IgG (both diluted 1:1000) (LI-COR Biosciences) with detection on a LI-COR Odyssey® Infrared Imaging System (LI-COR Biosciences).

2.9 Western Blot analysis of Rec-FLAG from transfected RWPE-1

Protein extraction and immunoblotting were conducted as described above. Rec-FLAG levels were detected using a mouse an anti-FLAG M1 primary antibody (Sigma-Aldrich) and a IRDye 680LT Goat Anti-Mouse IgG secondary antibody (LI-COR Biosciences) with detection on a LI-COR Odyssey® Infrared Imaging System (LI-COR Biosciences). Antibodies were diluted 1:1000 in 5% skimmed milk reconstituted in 1X tris-buffered saline (TBS) (pH8) 0.1% Tween. These dilutions were added to the transfer membrane, and shaken overnight at 4 °C, following a 1 hour room temperature blocking in 5% skimmed milk in TBS. Mouse monoclonal anti-β-actin antibody (Thermo-Scientific Pierce, Ireland #10624754) was used to confirm even protein loading.

2.10 Measurement of cytokines in human plasma samples

Heparinised plasma was collected from prostate cancer patients and population-based controls in the NCI study and stored at −80°C. Plasma interferon-γ (IFNγ), IP10, tumour necrosis factor-α, and interleukin-1β (IL-1β) concentrations were determined at a Leidos Biomedical Research Inc/NCI core facility using the human electrochemiluminescence immunoassays from Mesoscale Discovery (Gaithersburg, MD) under standardised conditions. Ultrasensitive multiplex electrochemiluminescence immunoassay plates were custom designed and were analysed on the MesoScale Discovery 6000 instrument, following manufacturer's assay and analysis protocols.

2.11 Cell lines

RWPE1, DU145, PC-3, 22Rv1 and CWR22 were sourced from the American Type Culture Collection (Manassas, VA) and cultured according to recommendations. In brief, CWR22 and 22Rv1 were cultured in RPMI 1640 medium with L-glutamine (Sigma #R8758) and supplemented with 10% fetal bovine serum Sigma #F7524). DU145 was cultured in minimum

essential medium (1×) with Earle's (Gibco #22561-021) supplemented with 10% fetal bovine serum. PC-3 was cultured in F12 nutrient mixture (HAM) medium, with L-glutamine (Gibco #21765-029) supplemented with 10% fetal bovine serum and 1 % antibiotic-antimycotic. RWPE1 was cultured in keratinocyte medium, supplemented with epidermal growth factor and Bovine Pituitary Extract (Gibco #17005–042). Cells were frozen when needed in 10% DMSO (Sigma-Aldrich) in FBS and stored in liquid nitrogen. Cells were thawed back by pelleting into fresh media and spinning at 1000 rpm for 5 minutes before resuspension into fresh media. All five cell lines were authenticated by LGC standards (United Kingdom) in May 2013 via short tandem repeat profiling and were found to be the correct cell lines **(Appendix B - Cell line authentication reports)**.

2.12 Construction of pcDNA-Rec

A gBlock (Integrated DNA Technologies) representing the coding sequence of *ERVK-6* (GenBank: X82271.1) together with FLAG-tag and restriction sites was ordered. 200 ng of this gBlock was dissolved in milli-Q water to a final concentration of 2 ng/µl. 1 µl of this solution was used in a PCR reaction using the primers Rec-FLAG forward 5'gaggtaccatggactacaaagacgatgacgacaagaacccatcagagatgcaaaga 3' and Rec-FLAG reverse 5'gtcagaattctcatggcccgttctcgatgg 3' (Sigma-Aldrich). Thermal cycling consisted of enzyme activation: 98 °C for 30 seconds, followed by 35 cycles of both denaturation and annealing/extension: 95 °C for 5 seconds and 60 °C for 15 seconds respectively and a final extension of 72°C for 2 minutes. The subsequent amplicon was gel extracted and cut with *Kpn*I and *Eco*RI restriction enzymes (NEB). The plasmid pcDNA3.1(+). (Thermo Fisher Scientific) was also cut using *Kpn*I and *Eco*RI restriction sites. The fragments were ligated using t4 DNA ligase (NEB) and transformed into Subcloning Efficiency™ DH5α™ Competent Cells (Invitrogen) using the heat shock method. Transformed cultures were plated on agar plates and grown overnight at 37 °C. The next day positive colonies were subcultured into LB broth supplemented with ampicillin

(Sigma Aldrich) and grown with shaking overnight at 37 °C. Plasmids were isolated from this culture using a QIAprep® Miniprep kit (Qiagen). Subsequent plasmid was then subjected to a restriction digest using *Kpn*I and *Eco*RI restriction enzymes to confirm the correct insert (Sambrook and Russell, 2001).

2.13 Optimization of RWPE-1 transfection using a GFP plasmid

400,000 RWPE-1 cells were plated in a 6-well plate and grown for 48 hours in keratinocyte medium, supplemented with epidermal growth factor and Bovine Pituitary Extract (Gibco #17005–042). 200 µl of Opti-Mem (Gibco) was mixed with 4 µl of FuGENE® HD Transfection Reagent (Promega) and 1 µg of pmaxFPTM-Green-N (Amaxa Biosystems) and incubated for 15 minutes. After 15 minutes this mixture was added to each well of the 6-well plate. Cells were then returned to the incubator for 48 hours. After this time cells were trypsinised and diluted to 1 million cells per ml in FACS buffer (PBS with 0.5% BSA, 1 mM EDTA, and 0.05% sodium azide) and analysed on the BD Accuri C6 flow cytometer (BD Biosciences) for expression of GFP using the FL1 channel. The gating strategy involved firstly gating on FSC and TO-PRO 3 (Thermo Fisher Scientific) in order to exclude dead cells. Secondly, they were gated against FSC-H and FSC-A in order to discriminate between single cells and doublets. Finally, they were gated on FSC-A and GFP in order to determine the final transfection efficiency. Gates for GFP positive cells were determined using non-transfected controls.

2.14 Transfection of RWPE-1

400,000 RWPE-1 cells were plated in a 6-well plate and grown for 48 hours in keratinocyte medium, supplemented with epidermal growth factor and Bovine Pituitary Extract (Gibco #17005–042). 200 µl of Opti-Mem (Gibco) was mixed with 4ul of FuGENE® HD Transfection Reagent (Promega) and 1 µg of *Bgl*II linearised pcDNA-Rec and incubated for 15 minutes. After 15 minutes this mixture was added to each well of the 6-well plate. Cells were then returned to the incubator for 48 hours after which time protein lysates were extracted and analysed for transgene expression. For stable

transfections, transfected cells were treated with 200 µg/ml of G418 (Thermo Fisher Scientific) for six weeks until polyclonal populations of RWPE-1-Rec emerged.

2.15 siRNA knockdown of HML-2

The siRNA specific to HML-2 as published by Orrichio et al. (Oricchio et al., 2007)TCCCAGTAACGTTAGAACC was ordered from Sigma-Aldrich. 170,000 cells were plated in a 6 well plate and simultaneously reverse transfected with either 10 nM or 50 nM of siRNA together with 3 µl of DharmFECT reagent (Dharmacon/GELife Sciences) and grown for 48 hours before RNA was extracted and qPCR conducted according to above protocols.

2.16 Androgen and Oestrogen treatments

200,000 cells were cultured in 6 well plates for 24 hours. Cells were serum starved for 24 hours with phenol red free media (Gibco) together with 10 % charcoal stripped FBS (Gibco). Cells were then treated with 0.1 nM or 1 nM of R1881 (Sigma-Aldrich) or β-Estradiol (Sigma-Aldrich) for 24 hours. 0.1 % ethanol (Sigma-Aldrich) in cell culture grade water (Sigma-Aldrich) was used as a negative vehicle control. RNA was extracted and qPCR conducted according to above protocols.

Chapter 3

HML-2 expression patterns and regulation in human prostate cancer cell lines

3 HML-2 expression patterns and regulation in human prostate cancer cell lines

Much data exists to date on the expression of HML-2 in many cancers. So far there is evidence for HML-2 mRNA and/or protein expression in ovarian (Wang-Johanning et al., 2007), melanoma (Schanab et al., 2011), breast (Wang-Johanning et al., 2012), prostate (Ishida et al., 2008), lymphoma (Contreras-Galindo et al., 2008) and leukaemia's (Brodsky et al., 1993a). HML-2 was originally discovered during a time when researchers were searching for a viral etiology of breast cancer (Ono et al., 1986). They screened human breast cancer tissue with riboprobes and pulled out sequences related to MMTV. These sequences turned out to be HML-2. Independently, researchers also discovered retroviral-like particles budding from teratocarcinoma cell lines (Bronson et al., 1978). They subsequently discovered that these particles were also encoded by HML-2 (Boller et al., 1993). Over the coming years, other types of cancers were eventually screened and found to also express these sequences.

Studies of HML-2 expression in these cancers have generally focused on a few different areas. First, is the study of the spliced and unspliced mRNA. HML-2 has a complex expression pattern whereby it expresses full length as well as spliced and doubly spliced mRNA. Furthermore, there are two main types of HML-2: type I which contains a 292 base pair deletion and type II which is the full length product. A second area of study is that of protein activation and autoantibody production. HML-2 is unique amongst the 11 members of the HML family in that it is the only member which produces proteins (Bannert and Kurth, 2006). All of the other HML groups are more ancient than HML-2 and have thus acquired mutations which render protein production an impossibility. An example of such a mutation is a premature stop codon or nonsense mutation. Such a mutation leads to a non functional peptide which is unable to fold and will be quickly degraded by the proteosome. Being a complex retrovirus, HML-2 produces two accessory proteins, Rec and Np9 (Downey et al., 2015). Rec is translated from type II

transcripts and Np9 is translated from type I. Antibodies against HML-2 proteins have also been discovered in cancers such as breast (Wang-Johanning et al., 2013). A third area of research is related to HML-2 activation. How are these proviruses - which are normally silenced – activated during carcinogenesis? Some studies have pointed towards hypomethylation in cancer as being responsible (Schulz et al., 2006) while others have suggested that certain steroid hormones or transcription factors are responsible (Manghera and Douville, 2013). Other studies have sought to define HML-2 proviral integrations. A seminal study by Subramanian et al. (Subramanian et al., 2011) discovered that the human reference genome harbours 91 distinct HML-2 proviruses together with 944 solitary LTRs. Other groups have built on this work by defining exactly which proviruses are responsible for the HML-2 signal seen in different types of cancer (Schmitt et al., 2013a).

At the beginning of this study in 2012, a paucity of data was present in the literature which described HML-2 expression in prostate cancer. We set out to define HML-2 expression in prostate cancer. In terms of mRNA expression, we were interested in discovering the splicing patterns of HML-2 as well as investigating if any of the main two types of HML-2, type I or type II, were being expressed preferentially. We were also interested in HML-2 protein expression in prostate cancer and its localization patterns in the disease. Finally, we also investigated possible mechanisms of HML-2 activation in the disease.

3.1 HML-2 mRNA is expressed in prostate cancer cell lines

We were interested in defining the mRNA expression of HML-2 in prostate cancer. To do so we utilised prostate cell lines 22Rv1, DU145, LNCap, PC3 and RWPE-1. These cell lines together represent a good *in vitro* model of the progression of prostate cancer from normal to primary to metastatic disease. RWPE-1 is a nontumourigenic immortalised prostate cell line which is an alternative to using prostate epithelial cells which have a finite

number of passages (Bello et al., 1997). 22Rv1 is an androgen independent parental cell line of CWR22 which was itself originally propagated from a mouse xenograft model of prostate cancer (Sramkoski et al., 1999; Wainstein et al., 1994). DU145 and PC3 are both androgen independent prostate cancer cell lines which represent metastatic disease originating from the brain and bone respectively (Kaighn et al., 1979; Stone et al., 1978). Finally, LNCap is an androgen dependent prostate cancer cell line which originated from a lymph node metastasis (Horoszewicz et al., 1980). We designed primers for four HML-2 genes: env type I (env1), env type II (env2), gag and Rec. These primers were designed to amplify as many of the 91 HML-2 loci as possible. Each primer set was optimised by creating a standard curve to measure their efficiency. RNA was then extracted from each of the cell lines and treated with DNAse to prevent genomic DNA contamination. This treated RNA was then converted to cDNA using reverse transcriptase. cDNA was subsequently used in a qPCR reaction in order to relatively quantify the levels of each HML-2 gene. A housekeeping gene – in this case ACTB – was used in order to normalize the data against. RWPE-1 was used as the calibrator sample, with all other genes expressed as fold change compared to it.

Our results revealed that all three HML-2 genes: env 1, Rec and gag were expressed in all 5 cell lines (**Figure 8**). The nontumourigenic prostate cell line RWPE-1 expressed the lowest levels of mRNA of all three genes. In PC3 we observed a 2.5 fold higher in expression of env1 and gag when compared to RWPE-1, while Rec levels were similar. In LNCap we found that levels of env1 were nearly 5 fold higher than in RWPE-1 while gag and Rec levels were similar to levels observed in PC3. Rec levels in DU145 did not seem to higher in comparison to RWPE-1 but env1 and gag levels were increased 2.5 fold and 4 fold respectively. Finally, in 22Rv1 we observed the highest levels of expression of HML-2 mRNA. Env1 was expressed 10 fold in comparison to RWPE-1. No increases were seen in gag levels but Rec was expressed over 10 fold albeit with a high degree of variability amongst technical replicates.

Furthermore to this data, we were interested in exploring the expression of the two main types of HML-2 env transcript – type I and type II. We hypothesised that one type may have functional significance over the other. An independent analysis of env1 and env2 expression was completed in 4 prostate cancer cell lines: CWR22, 22RV1, PC3 and DU145 and one nontumourigenic prostate cell line, RWPE-1 **(Figure 9).** RNA was extracted from each cell line, treated with DNase, converted to cDNA and levels analysed using qPCR. Expression of the two genes was observed in all the cell lines tested, even in the nontumourigenic RWPE-1. We conducted a statistical analysis for each cell line comparing env1 vs env2 expression. Env1 levels were found to be significantly higher than env2 levels in PC3 cells (p= 0.0025) and env2 levels were found to be significantly higher than env1 levels in 22RV1 cells (0.0172). All other differences were deemed not statistically significant.

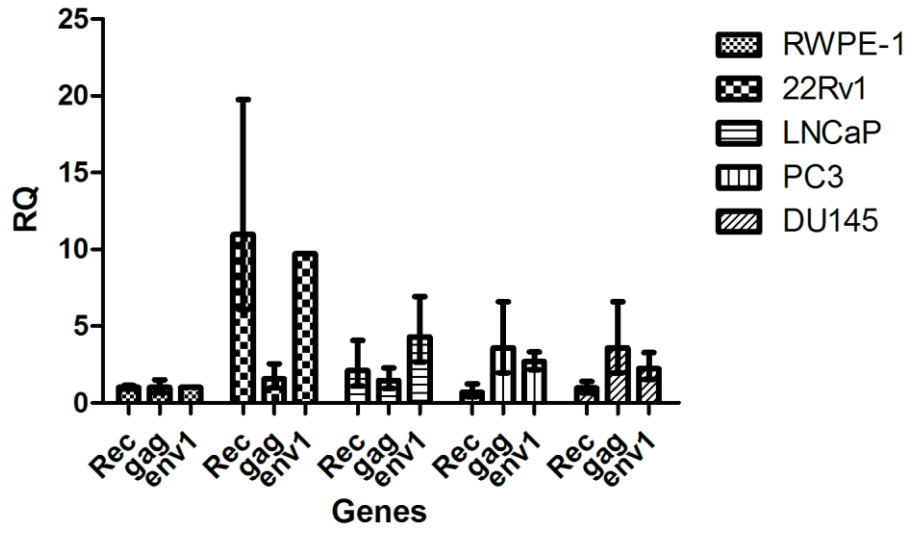


Figure 8 - HML-2 mRNA expression in prostate cancer cell lines

Expression of three HML-2 genes, env1, gag and Rec was observed in four prostate cancer cell lines: 22Rv1, DU145, LNCap and PC3 and one nontumourigenic prostate cell line, RWPE-1. Highest levels of expression were seen in the AR mutant cell line 22Rv1 for both env1 and Rec. DU145 and PC3 exhibited similar expression levels of all three of the genes, while LNCap

expressed four fold the amount of env1 in comparison to RWPE-1. Error bars represent RQ min and max from three technical replicates.

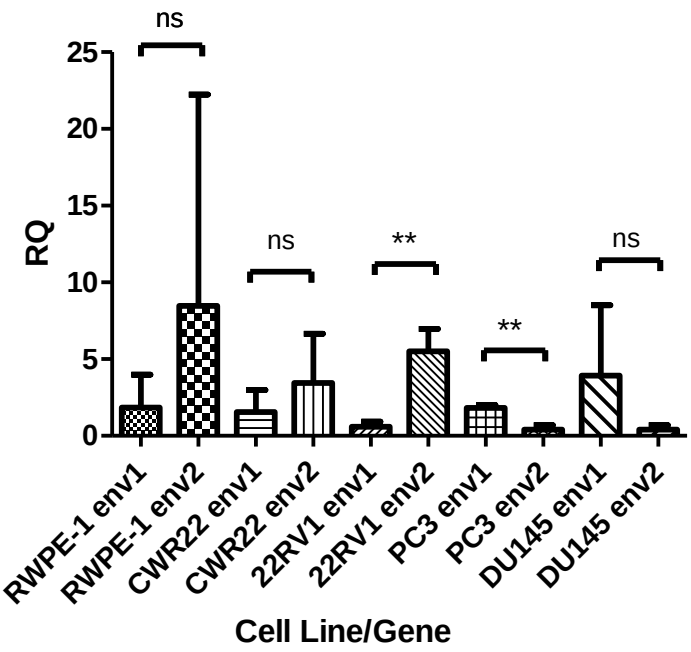


Figure 9 - env1/env2 expression in prostate cell lines

Expression of two HML-2 genes, env1 and env2 were observed in four prostate cancer cell lines: 22Rv1, DU145, CWR22 and PC3 and one nontumourigenic prostate cell line, RWPE-1. For each cell line a statistical analysis was completed comparing env1 and env2 levels. env1 levels were found to be significantly higher than env2 levels in PC3 cells (p=0.0019) and env2 levels were found to be significantly higher than env1 levels in 22RV1 cells (p=0.0048). All other comparisons were not significant. Results are expressed as fold change (RQ) in comparison to RWPE-1. Error bars represent standard deviation from three technical replicates (ns=not significant, *P ≤ 0.05, **P ≤ 0.01; Student's t-test)

Difficulty was encountered whilst attempting to analyse Np9 mRNA levels using qPCR so semi- quantitative RT-PCR was employed instead in order to investigate spliced Np9 expression in prostate cancer cell lines. RNA was extracted from prostate cancer cell lines using TRIzol and treated with DNAse to eliminate genomic DNA contamination. The treated RNA was

then converted to cDNA. Semi-quantitative RT-PCR was subsequently conducted using this cDNA together with primers designed against NP9. Bands representing Np9 mRNA were observed in all prostate cancer cell lines but not in immortalised non tumourigenic RWPE-1 or minus reverse transcriptase and no template negative controls **(Figure 10)**. These results suggest that HML-2 Np9 mRNA is expressed across most prostate cancer cell lines but not in non tumourigenic RWPE-1.

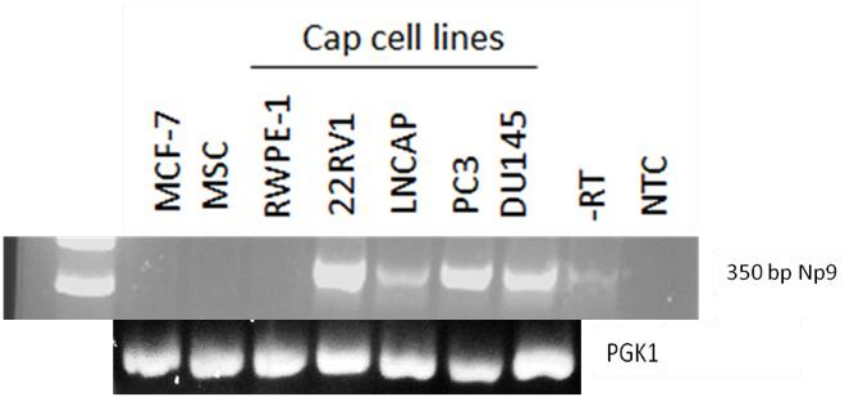


Figure 10 - HML-2 Np9 RT-PCR

Semi- quantitative RT-PCR representing Np9 mRNA expression in prostate cancer cell lines and in immortalised non tumourigenic RWPE-1. Negative controls: MSC, minus reverse transcriptase (-RT) and no template control (NTC). Positive MCF-7 breast cancer cell line control. PGK1 loading control.

3.2 HML-2 proteins are expressed in prostate cancer cell lines

Most of the hypotheses linking HML-2 with carcinogenesis concern the HML-2 protein products themselves. Even though RNA products such as long non coding RNA can have functional effects, the large majority of function is implicit in the protein products. HML-2 expresses the canonical retroviral proteins gag, pro, and pol alongside the accessory proteins Rec and Np9 (Downey et al., 2015). The 292 base pair deletion which is found in type I HML-2 proviruses leads to the fusion of the pol and env reading frames which eventually results in the translation of the Np9 protein. Rec is expressed from the non-deleted type II transcript **(Figure 4).**

Our study, utilised western blotting in order to detect the presence of HML-2 proteins in prostate cancer cell lines. Monoclonal antibodies against HML-2 gag and env were acquired as well as a polyclonal antibody against the Rec protein. The western blot probing for HML-2 gag revealed expression of the protein in both metastatic cell lines (DU145 and PC3) and also in primary lines (CWR22 and 22Rv1) **(Figure 11)**. Expression was also seen in the MDA-MB-231 positive breast cancer cell line control. A GAPDH loading control was used to confirm equal loading. Similarly, western blotting for HML-2 env using anti-env (6H5) antibody revealed the expression of the protein in all cell lines tested **(Figure 12).** Expression was also seen in this case in immortalised nontumourigenic RWPE-1. Highest expression was seen in RWPE-1 and PC3. The upper secondary bands represent a glycosylated form of HML-2 env, while the second largest band of ~155 kDa presumably represents trimers of HML-2 env surface units (55 kDa). However, there is also the possibility that these are non specific bands instead of trimers. A beta actin loading control was used to confirm equal loading. Finally, western blotting for HML-2 Rec revealed low expression of the protein in the four cancer cell lines as well as in nontumourigenic RWPE-1 **(Figure 13)**. A polyclonal antibody against the Np9 protein was also acquired but could not be optimised for use in western blot experiments.

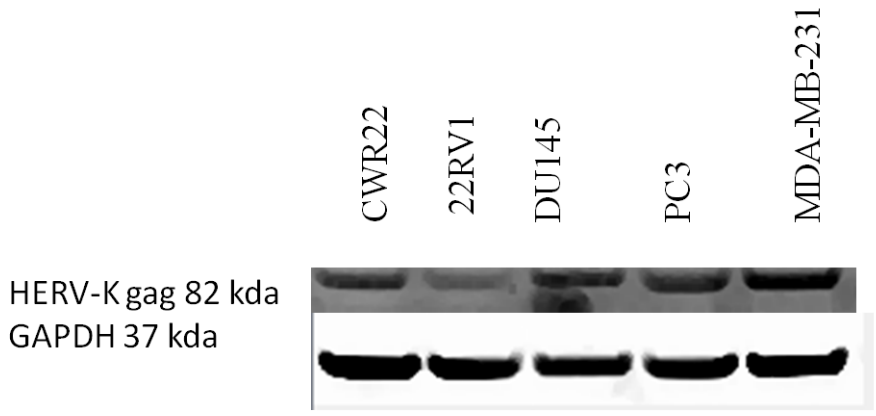


Figure 11 - Basal HML-2 gag protein expression in prostate cancer cell lines

HML-2 gag protein was observed to be expressed in four prostate cancer cell lines and MDA-MB-231 breast cancer positive control after western blotting. GAPDH loading control 37 kDa.

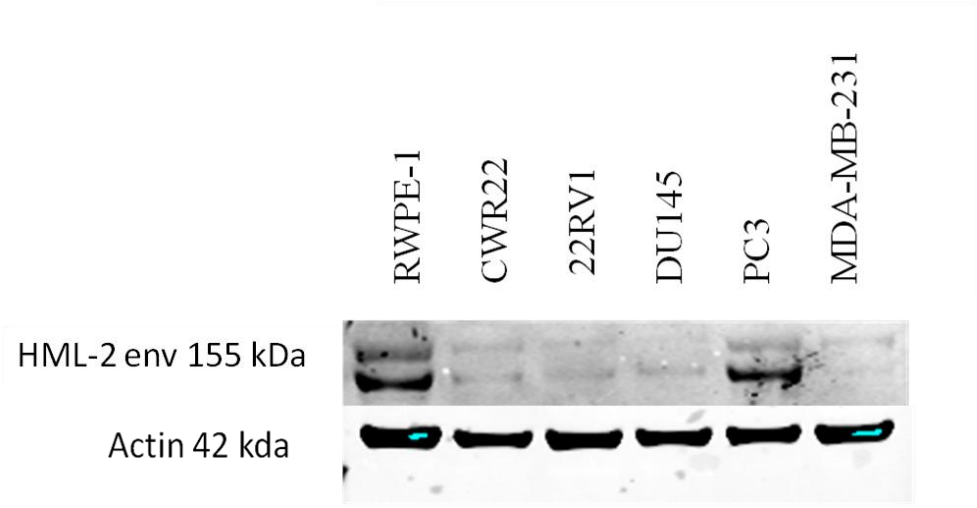


Figure 12 - Basal HML-2 env protein expression in prostate cancer cell lines

HML-2 env protein was observed to be expressed in four prostate cancer cell lines and in immortalised nontumourigenic RWPE-1 and MDA-MB-231 breast cancer positive control. Upper bands represent a glycosylated form of HML-2 env. Actin loading control 42 kDa.

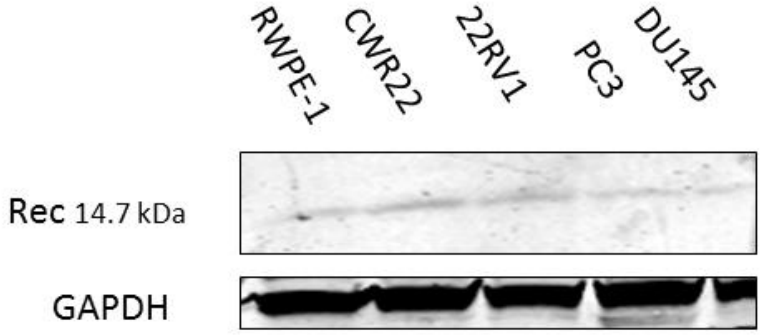


Figure 13 - Basal HML-2 Rec protein expression in prostate cancer cell lines

Low level basal expression of HML-2 Rec protein was observed in 4 prostate cancer cell lines as well as in immortalised non tumourigenic RWPE-1. GAPDH loading control 37 kDa.

3.3 HML-2 protein is expressed in clinical prostate cancer samples

We were also interested in investigating whether HML-2 proteins were expressed in clinical samples. For this, we utilised formalin-fixed paraffin-embedded (FFPE) prostate cancer samples. We subsequently conducted immunohistochemistry on microtome sections of these samples using an anti-HML-2 env antibody (6H5). Positive staining was confined mostly to the luminal epithelial cells but some low intensity staining was also visible in the stroma **(Figure 14)**.

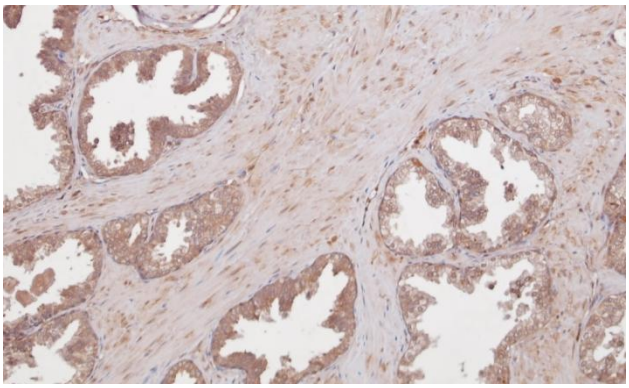


Figure 14 - HML-2 env protein expression in prostate cancer tissue

HML-2 env protein expression was observed after immunohistochemical staining in prostate tissue. Staining was confined to the luminal epithelium but with some present in the stroma. An isotype matched IgG control was used in place of primary antibody to confirm that staining was specific.

We also conducted IHC using an anti-HML-2 gag antibody on prostate cancer tissue microarrays which contained cores from prostate adenocarcinoma and benign prostatic hyperplasia (BPH) samples. These results make up part of chapter 2 of this thesis.

3.4 HML-2 activation and steroid hormone treatment in prostate cancer

Studies in various cancers have shown that HML-2 expression can be increased by treatment with steroid hormones. For example, in breast cancer; progesterone and estrogen have been shown to increase HML-2 levels (Wang-Johanning et al., 2001). Breast cancer is a disease which is essentially hormonally driven and thus it is interesting to speculate whether hormonal activation of HML-2 plays a role in carcinogenesis. Prostate cancer and breast cancer share many similarities at the molecular level (Risbridger et al., 2010). Because of this, we were interested in investigating the links between steroid hormones and HML-2 activation in prostate cancer. To this end, we treated androgen dependent prostate cancer cell lines with androgens and oestrogens. We subsequently used qPCR to analyse levels of HML-2 mRNA, looking for any changes in expression.

3.4.1 Androgen effects on HML-2 expression in prostate cancer cell lines

First we tested a synthetic form of androgen known as R1881. We extracted total RNA from AR positive LNCap prostate cancer cells. These cells had been treated with 0.1 nM and 1 nM of R1881. These concentrations were chosen as they represent physiologically relevant concentrations that would be present in the human body. LNCap cells are androgen dependent, expressing a wild type form of the AR which will bind free androgen molecules. Total RNA was extracted from the cells and treated with DNAse to prevent genomic DNA contamination and subsequently converted to cDNA which was used in a qPCR reaction together with primers for HML-2 gag, env1, env2 and Rec mRNA. A negative vehicle control was used in the form of cells treated with ethanol only. We simultaneously treated AR negative DU145 cells with the same treatments.

Our results indicated that treatment with R1881 leads to a substantial increase in expression of HML-2 gag mRNA (**Figure 15**). 0.1 nM concentration led to a 60 fold increase in gag expression while the 1 nm concentration led to a 110 fold increase compared to the non treated cells. We also tested the effect of R1881 treatment on spliced HML-2 expression. Our results indicated that all spliced HML-2 transcripts were upregulated by R1881 treatment (**Figure 16**). 0.1 nM concentration increased levels by approximately 25 fold for env1, env2 and Rec, while 1 nm led to an upregulation of 100 fold, 85 fold and 130 fold of env1, env2 and Rec respectively. Overall, our results show a dose dependent increase in HML-2 expression with increasing concentrations of R1881.

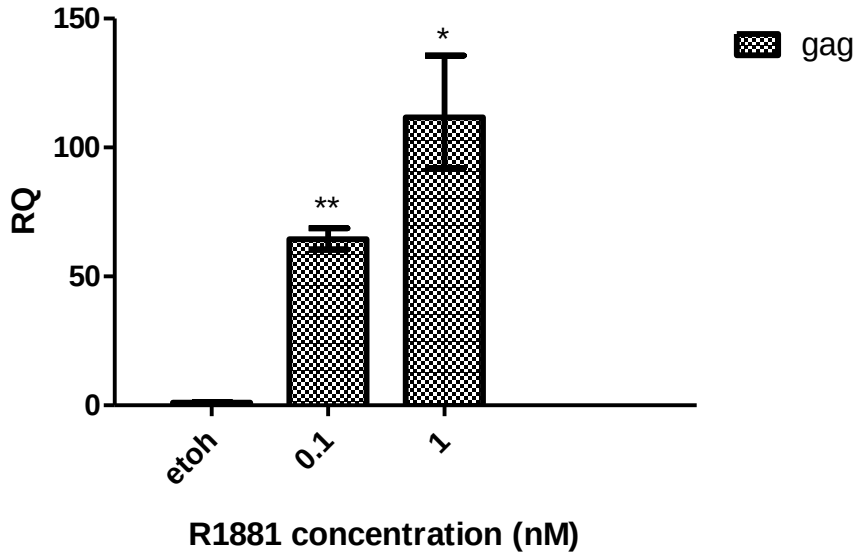


Figure 15 - Treatment of LNCap cells with R1881: full length HML-2 transcript expression

LNCap cells were treated with 0.1 nM or 1 nM R1881. A significant dose dependent increase was observed with increasing levels of R1881 leading to increasing levels of HML-2 gag transcription. Results are expressed as fold change (RQ) in comparison to a non treated vehicle control (ethanol (etoh)). Error bars represent standard deviation from three technical replicates (*P ≤ 0.05, ** P ≤ 0.01; Student's t-test)

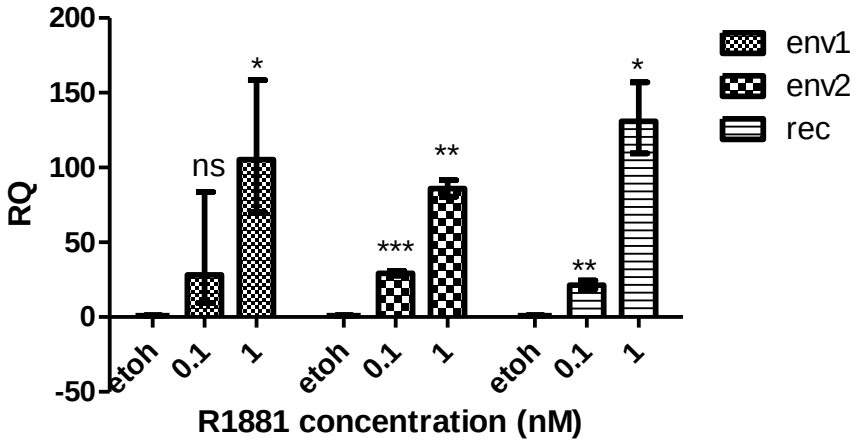


Figure 16 – Treatment of LNCap cells with R1881: spliced HML-2 transcript expression

LNCap cells were treated with 0.1 nM or 1 nM of R1881. A significant dose dependent increase was observed with increasing levels of R1881 leading to increasing levels of HML-2 env1, env2 and Rec transcription, with the exception of the 0.1 nM treatment env1 analysis which was deemed not to be a statistically significant increase (p=0.2169). Results are expressed as fold change (RQ) in comparison to a non treated vehicle control (ethanol (etoh)). Error bars represent standard deviation from three technical replicates (*P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001; Student's t-test)

Treatment of the AR negative DU145 cells with the same R1881 concentrations led to fold decreases of HML-2 gag in all cases compared to the untreated vehicle control (**Figure 13**). Interestingly, a large decrease of 333 fold was seen for 0.1 nM of R1881. The spliced transcripts, env1, env2 and Rec also exhibited fold decreases under the same conditions, however the 0.1 nM did not lead to such a pronounced decrease (**Figure 18**).

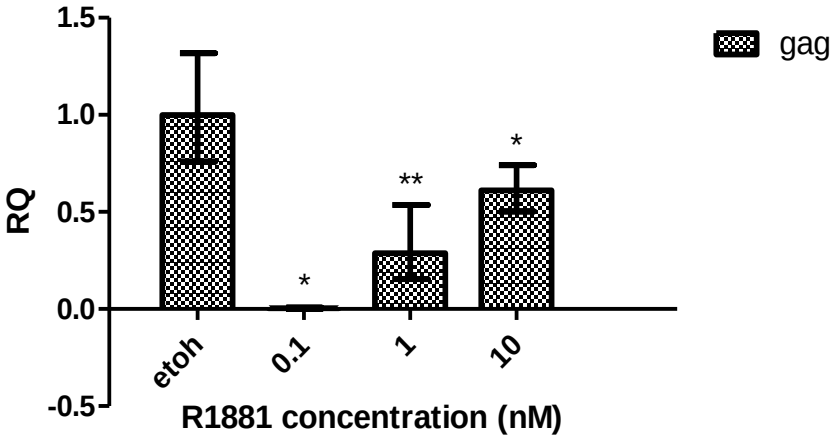


Figure 17 - Treatment of DU145 cells with R1881: full length HML-2 transcript expression

DU145 cells were treated with 0.1 nM, 1 nM or 10 nM concentrations of R1881. Significant fold decreases of HML-2 gag were observed after treatment with each concentration with a large decrease of 333 fold observed in the 0.1 nM treatment. Results are expressed as fold change (RQ) in comparison to a non treated vehicle control (ethanol (etoh)). Error bars represent standard deviation from three technical replicates (*P ≤ 0.05, ** P ≤ 0.01; Student's t-test)

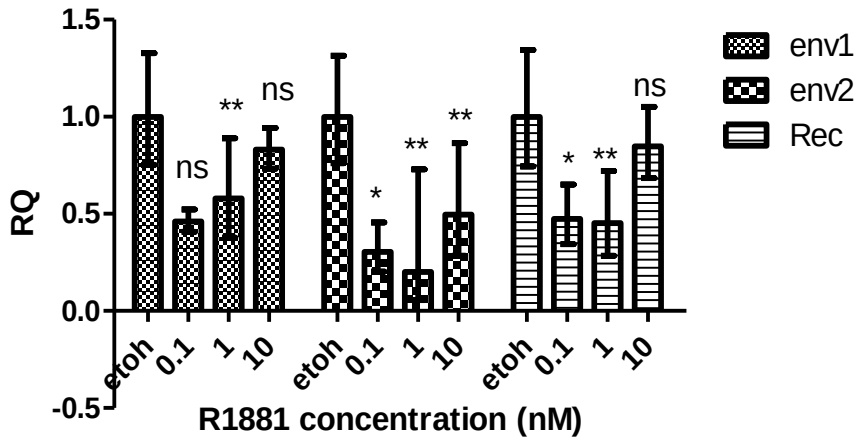


Figure 18 - Treatment of DU145 cells with R1881: spliced HML-2 transcript expression

DU145 cells were treated with 0.1 nM, 1 nM or 10 nM of R1881. Significant fold decreases were observed across all treatments for env2 in comparison to the untreated control. However, only 1 nM in the env1 analysis and 0.1 nM and 1 nM in the Rec analysis reached a statistically significant decrease. Results are expressed as fold change (RQ) in comparison to a non treated vehicle control (ethanol (etoh)). Error bars represent standard deviation from three technical replicates. (ns $P > 0.05$, *$P \leq 0.05$, ** $P \leq 0.01$; Student's t-test)

3.4.2 Estrogen effects on HML-2 expression in prostate cancer cell lines

Not long after their discovery of HML-2, Ono and colleagues completed experiments on breast cancer cell lines which induced HML-2 expression via treatment with oestrogens (Ono et al., 1987). These studies were replicated later by Wang-Johanning et al. who found that oestrogen treatment led to an increase in spliced HML-2 expression (Wang-Johanning et al., 2003).

Although it is the androgens which are the most important hormone in the formation of male characteristics, oestrogens are also expressed and play a lesser role (Hess et al., 1997). *In vivo* mouse models have implicated oestrogens in the formation of PIN and adenocarcinoma (Risbridger et al.,

2003). We were interested in investigating if the link between oestrogens and carcinogenesis was in any way associated with HML-2 expression. To this end, we treated AR positive LNCap cells and AR negative PC3 cells with varying concentrations of β-Estradiol (a form of oestrogen) and subsequently analysed HML-2 mRNA expression levels via qPCR.

In LNCap cells, we observed fold decreases in expression of HML-2 gag mRNA after treatment with β-Estradiol compared to the non treated vehicle control. These decreases were 6 fold and 11 fold for 0.1 nM and 1 nM of β-Estradiol respectively (**Figure 19**).

The same experiment was repeated except AR negative PC3 cells were used in place of LNCaps. Fold increases were seen for all concentrations of β-Estradiol, with a large increase of 49 fold seen for 0.1 nM (**Figure 20**).

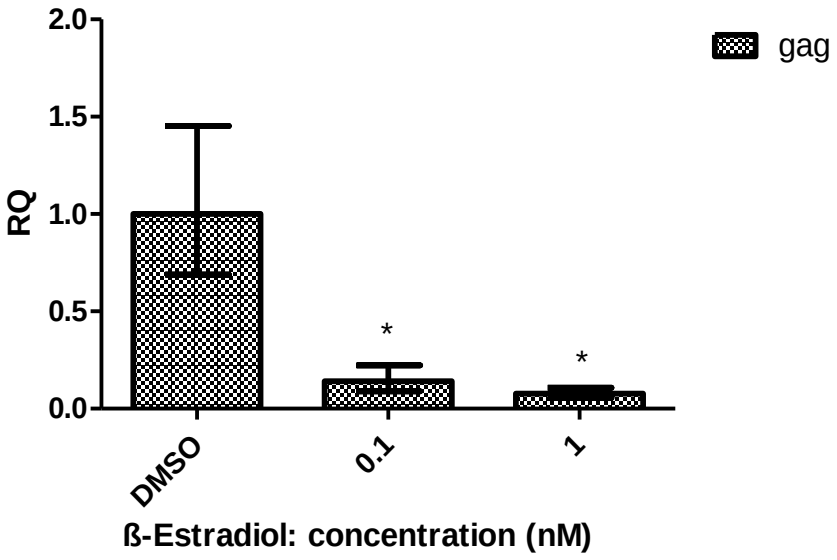


Figure 19 - Treatment of LNCap cells with β-Estradiol: full length HML-2 transcript expression

LNCap cells were treated with 0.1 nM or 1 nM of β-Estradiol. Significant fold decreases of HML-2 gag were observed in each of these treatments compared to the non treated control (DMSO). Results are expressed as fold change (RQ) in comparison to a non treated vehicle control (DMSO). Error bars represent standard deviation from three technical replicates. (*P ≤ 0.05, Student's t-test).

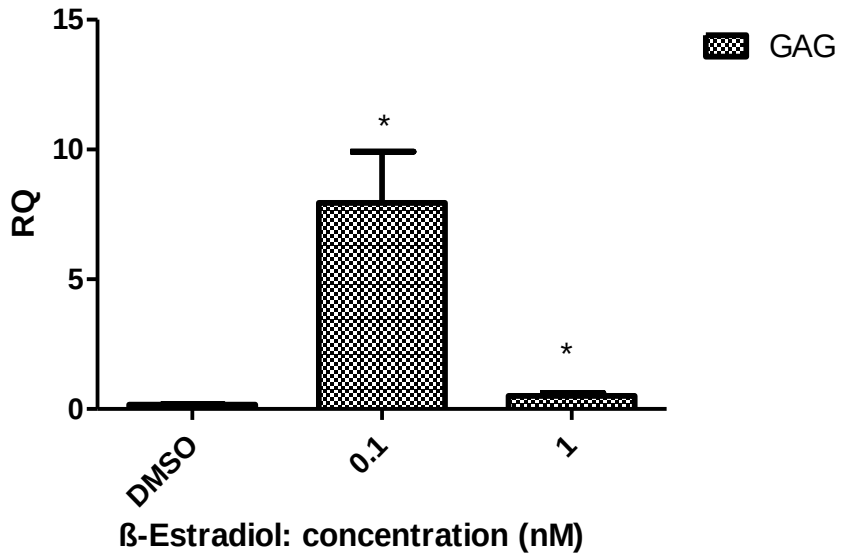


Figure 20 - Treatment of PC3 cells with β-Estradiol: full length HML-2 transcript expression

PC3 cells were treated with 0.1 nM or 1 nM of β-Estradiol. Significant fold increases were observed in each treatment compared to untreated DMSO control with a large increase of 49 fold observed in the 0.1 nM treatment (p=0.0161). Results are expressed as fold change (RQ) in comparison to a non treated vehicle control (DMSO). Error bars represent standard deviation from three technical replicates. (*P ≤ 0.05, Student's t-test).

3.5 Discussion

This study sought to define the expression of HML-2 in prostate cancer. The first aim of this study was to define the pattern of HML-2 transcription in prostate cancer cell lines. Each type of HML-2 expresses four types of transcripts. Type I HML-2 express full length (gag), env1, np9 and hel. Type II HML-2 express gag, env2, Rec and hel. Our qPCR analysis found that env1, env2, Rec and gag were expressed in 22Rv1, DU145, PC3

LNCap and RWPE-1. Highest levels overall were found in 22Rv1 while RWPE-1 expressed the lowest. Originally, RWPE-1 had been used as a negative control as it is a nontumourigenic prostate cell line. The finding that it expresses HML-2 at all is interesting and may be down to two reasons. Firstly, RWPE-1 is a cell line was originally produced via immortalization of prostate epithelium cells with human papillomavirus 18 (Bello et al., 1997). Epstein-Barr virus has been shown to transactivate the HML-2 LTR and thus the process of HPV viral immortalization may have inadvertently led to mutations of the RWPE-1 LTR which permit transcription of HML-2 (Hsiao et al., 2006). However, HPV and EBV are genetically very different viruses and so this mechanism seems unlikely. Agoni et al. found that all the cervical cancer cell lines they tested were HML-2 positive apart from C33A (Agoni et al., 2013b). Interestingly, C33A was the only cell line which was also negative for HPV. Therefore, there may remain an unknown link between HPV infection and HML-2 activation. Secondly, the signal from RWPE-1 may simply be a byproduct of pervasive transcription. Pervasive transcription is a term used to apply to large non functional areas of the genome – including intergenic regions - which are transcribed but from which the transcripts do not have any function (Lander, 2011). In this experiment RWPE-1 was used as the calibrator sample with all other cell lines expressed as fold change differences to it. Even though RWPE-1 was positive for transcription, the other cancer cell lines all exhibited fold change increases in all HML-2 genes tested. Interestingly, 22Rv1 expressed over 10 fold the levels of spliced HML-2 env1 and Rec that RWPE-1 did albeit with high levels of technical variability. 22Rv1 is a cell line that is derived from a primary prostate cancer and so this data point towards a possible function for spliced HML-2 transcripts in this cancer, although further work needs to be completed in clinical samples to investigate this link.

There may be a number of reasons for differences in expression between env1 and env2 in the cell lines tested. Firstly, the difference may simply be due to the fact that one cell line has a higher number of one type of HML-2

loci over another. Whole genome sequencing of cancer cell lines would enable us to answer this question in a conclusive manner. Differences in methylation between each cell line may also lead to the differences seen in expression. A more interesting possibility is that a cell line may express higher levels of one type over another because its survival as a cancer cell depends on it. The functional characterisation of the two env types would allow us to form hypotheses about their role in cancer. Also, the 292 bp deletion leads to the fusion of the pol and env reading frames in type I proviruses. Chen et al have shown that this fused ORF can produce a pol-env polyprotein in leukaemia (Chen et al., 2013). Therefore, it will be important for future work to define the protein expression which stems from env1 transcripts in prostate cancer.

A conflict in the data exists between the env1 levels observed in 22rv1 versus RWPE-1 in **Figure 8** compared to **Figure 9**. This could be potentially down to two reasons. Firstly, these experiments were conducted at different times and therefore passaging of the cell lines may have altered methylation profiles which may affect HML-2 expression levels. Secondly, this discrepancy may be related to how the analysis was run. The analysis could potentially be re-run by graphing delta cT instead of conducting a relative quantification.

A limitation of this study is that relative differences between env1 and env2 in the same sample were measured using RQ which is not a suitable quantification in this instance. Absolute quantification should be considered in order to accurately assess the true differences in these genes within these samples. This can be conducted by using an RNA standard and producing a standard curve.

Due to technical difficulty in designing qPCR primers specific for the Np9 transcript we instead attempted to investigate Np9 mRNA in prostate cancer via semi-quantitative RT-PCR. Expression of Np9 was observed in all four prostate cancer cell lines tested but not in nontumourigenic RWPE-1. MCF-7 breast cancer cell were used as a positive cell line control (Wang-

Johanning et al., 2012) and MSC cells as a second negative control (Fuchs et al., 2013) due to the previous finding of gag, env1 and Rec expression in RWPE-1. The MCF-7 breast cancer cells we tested did not express Np9 transcripts which is in contradiction to the literature which finds that spliced HML-2 is expressed in this cell line (Wang-Johanning et al., 2003; Wang et al., 1998) This discrepancy may be due to a difference in the source of the cell line or due to a high passage number which means that recombination mechanisms may have altered the LTR of some HML-2 loci. Highest levels of expression were observed in 22Rv1, DU145 and PC3 with lower levels seen in LNCap. This finding is an agreement with the levels of gag, env1 and Rec observed in the previous qPCR result. This is to be expected as Np9 is a spliced transcript of the full length gag mRNA. High levels in the metastatic lines point towards the fact that Np9 may be involved in the metastatic process but once again this will have to be investigated in clinical models. The lack of Np9 mRNA in RWPE-1 conflicts with the finding in the qPCR assay where expression was found to be present but low. This may be due to the technical differences between RT-PCR and qPCR or it may be a true biological reflection. This discrepancy could be fully understood by designing qPCR primers for Np9 – something which was attempted for this study but did not provide a single peak on the SYBR green melt curve and so was discarded. Also, semi quantitative RT-PCR is – as the name suggests – is only semi quantitative and so making conclusions on differences in expression between the cell lines in the case of this Np9 experiment is not ideal and should be followed up by fully quantitative qPCR.

After our investigation of HML-2 mRNA in prostate cancer we were keen to also delineate basal protein levels. We conducted western blotting using antibodies against HML-2 gag and env on four prostate cancer cell lines and a positive control breast cancer cell line known as MDA-MB-231 (Wang-Johanning et al., 2012). For the gag western, the results revealed positive expression of the protein in the four prostate cancer cell lines and also in

the positive control breast line. The highest levels of expression were seen in the two metastatic cell lines DU145 and PC3 and also in CWR22, while low levels were observed in 22Rv1. 22Rv1 is a parental cell line of CWR22 and thus would be expected to express similar levels of gag to it. However, 22Rv1 differs to CWR22 in that it is an androgen independent cell line even though it carries a functional but mutated AR. It is tempting to speculate that dysregulated androgen signalling is related to a drop in expression of HML-2 gag but further functional work needs to be completed to investigate this. The data is in concordance with the qPCR data of mRNA expression in cell lines where 22Rv1 exhibited low levels of gag mRNA.

Western blotting was also completed to investigate basal levels of HML-2 env. Low levels of expression were observed in CWR22, 22Rv1 and DU145, with high levels of expression seen in RWPE-1 and PC3. What appeared to be glycosylated forms of env were seen in all cell lines tested. This shows that prostate cancer cell lines can successfully modify env proteins. Many retroviruses are known to express glycosylated envelope proteins and it is thought that such modifications are necessary for successful target cell receptor binding (Hakomori, 2002). The receptor which HML-2 uses to infect cells is currently unknown. Glycosylated env proteins have been previously shown in breast cancer cell lines (Wang-Johanning et al., 2012). In other cancers, the lack of infectivity of HML-2 virions are thought to be due to immature env proteins , thus the presence of these modifications shows that these env proteins are mature in nature (Downey et al., 2015). To conclusively prove that the upper bands in this blot are indeed glycosylated bands, one would need to complete western blots using anti-glycosylation antibodies or mass spectrometry. Another method of doing this would be to use a glycolytic enzyme to treat the proteins before running the western blot and then check if the upper band disappears or decreases. The presence of high levels of protein in the RWPE-1 cells is conflicting with the env1 qPCR on cell lines. This may be because the proteins on this blot actually represent env2 proteins even though the antibody used is meant to have affinity for the two kinds. Due to

the relative nature of the qPCR quantification it is also possible that the RWPE-1 signal in the qPCR assay is truly reflective of the western blot levels and that the HML-2 mRNAs in the cancer cell lines are being degraded by host restriction factors or silenced by RNAi. This would be in disagreement with hypotheses which link HML-2 expression in a functional way to the progression of cancer. In order to clear up this discrepancy, a qPCR against env type II in RWPE-1 could be completed. Absolute quantification of HML-2 mRNA instead of relative quantification would also be of benefit.

A western blot was also completed against HML-2 Rec using a polyclonal anti-Rec antibody. Low levels of expression of Rec were observed in four prostate cancer cell lines and in nontumourigenic RWPE-1. This blot is in agreement with the expression of Rec mRNA seen through qPCR analysis. However, levels of Rec protein in 22Rv1 seem to be lower than that seen in the qPCR analysis. Once again, the disagreement between Rec protein levels compared to mRNA levels may be due to them being degraded by host restriction factors or silenced by RNAi. Experiments to look for restriction factor mutation signatures could be completed to try and investigate this. Being able to prove that a restriction factor was actively targeting HML-2 mRNAs in normal cell lines but is unable to do so in cancer cell lines would be evidence towards the fact that HML-2 is inherently carcinogenic and in a normal context needs to be kept under control by endogenous restriction machinery. Another reason for this discrepancy is that the antibody is non-specific for Rec. This may be the case since it is a polyclonal antibody.

Rec has been shown experimentally to lead to derepression of the AR (Kaufmann et al., 2010) and c-Myc (Boese et al., 2000). These are interesting proteins in the context of prostate cancer. The low basal levels of Rec seen in the prostate cancer cell lines would suggest that they would struggle to cause any significant effect on phenotype but it is possible that in clinical samples that some cancer cell clones acquire the ability to

upregulate Rec to high levels, which may lead to a dysregulation of androgen signalling which could lead to androgen independent growth.

We were interested in investigating HML-2 protein expression in clinical samples and so we conducted immunohistochemistry on prostate cancer FFPE tissue looking for HML-2 env. Our results showed strong env expression concentrated towards the lumen of the gland. The secretory luminal cells are thought to be the cell of origin for prostate adenocarcinoma (Wang et al., 2009) and so this data would seem to support this. Interestingly, we also observed low intensity staining throughout the stroma. One of the foremost hypotheses as to how HERVs are activated in cancer is that they become hypomethylated (Schulz et al., 2006). This would support the notion that the cancer cells of origin at the lumen are becoming hypomethylated at their HML-2 LTRs leading to its expression. However, stromal cells would not become hypomethylated and thus should not be expressing HML-2 at all in a cancerous context. There are a few possible reasons for the expression of HML-2 in the stroma. Firstly, prostate stroma may indeed get hypomethylated to a certain degree leading to HML-2 expression. Secondly, there may be some intercellular communication occurring between the epithelium and stroma which results in HML-2 expression. Thirdly, the proteins themselves or the transcripts may be being transferred intracellularly possible via exosome transport which was discussed in the introduction of this thesis (Downey et al., 2015). Fourthly, the positive signal in the stroma may stem from lymphocytes which are HML-2 positive. Finally, prostate cancer HML-2 expression may not require hypomethylation at all and the levels of expression seen in both cellular compartments are due to another activator. However, this would conflict with previous studies such as Goering et al. who suggest that HERV hypomethylation occurs in prostate cancer (Florl et al., 1999).

We also investigated the effect of different steroid hormones on HML-2 mRNA expression using qPCR. First we tested LNCap cells with the

synthetic androgen R1881. These cells were chosen to be tested as they are an androgen dependent line. Physiologically relevant levels of the synthetic androgen R1881 were chosen which were 0.1 nM and 1 nM respectively. For HML-2 gag we observed a dose dependent increase in expression with increasing concentrations of R1881. Levels of gag increased considerably with the 1 nM treatment leading to an over 100 fold increase in expression.

This experiment was repeated but instead primers were used for the spliced transcripts env1, env2 and Rec. A dose response was also evident for these genes with increasing concentrations of R1881 leading to increasing levels of the spliced transcript expression. Overall, it is clear that R1881 treatment leads to an increase in expression of all HML-2 genes in LNCap cells. The mechanism of this increase works via the binding of R1881 to the AR (Jenster et al., 1991). Further evidence for this could be gained by co-treating the cells with R1881 together with an AR inhibitor such as flutamide and looking for a reversal of the phenotype.

Even if R1881 was proven to be binding the AR, whether the increase in HML-2 was a direct or indirect result of activated AR would also have to be experimentally validated. To date, it has not been experimentally proven that the AR binds to the HML-2 LTR (Hanke et al., 2013). Some studies have suggested but not proven that the HML-2 LTR contains androgen response elements (AREs) (Hanke et al., 2013). In order to gain conclusive evidence of an AR/HML-2 LTR interaction, one would need to conduct chromatin immunoprecipitation (ChIP). This involves the shearing of the DNA and subsequent immunoprecipitation of bound proteins. ChIP followed by next-generation sequencing (ChIP-Seq) would enable one to conclude which exact LTR sequences were binding the AR. Hanke et al. proposed a vicious cycle model of HML-2 activation, whereby increasing levels of Rec protein led to increasing transcription of HML-2 (Hanke et al., 2013). It would be interesting to elucidate whether this feed-forward loop is present in LNCap cells. To do so would require a construct to overexpress the Rec protein. The effects of high levels of HML-2 Rec protein in prostate cancer could be an interesting phenomenon - due to the interactions Rec has

been shown to make with the AR and with c-myc – and will be discussed in greater detail in chapter 3 of this thesis.

Another way of gaining evidence towards R1881 treatment being directly responsible for the increase in HML-2 expression that we observed is to repeat the same experiment using an AR negative cell line. Theoretically, we should see no increase in expression of HML-2 if the effect is working through the AR. Our results of this experiment indicated that the R1881 treatment led to fold decreases for all HML-2 genes compared to the untreated control. There may be a few different reasons for this observed decrease. Firstly, R1881 may interact with specific DU145 proteins other than the AR which are not present in LNCap, these protein-protein interactions may directly or indirectly lead to a decrease in HML-2 expression. Secondly, the large fold decrease seen may actually just be due to high basal levels of HML-2 mRNA in DU145. However, this does not seem to be true as our previous data of basal mRNA levels across the prostate cancer cell lines suggested that levels were comparable between DU145 and LNCap. For the 0.1 nM concentration, the decreases observed in the spliced transcript levels were more modest compared to the over 300 fold decrease observed in the full length levels. This is a surprising result given that spliced transcripts have their origin in the full length transcript and so a large decrease in full length levels should lead to a concomitant decrease in spliced levels. The reason for this discrepancy may firstly be due to technical variation between the two experiments. Secondly, R1881 may be interfering with gag transcripts in a post transcriptional manner, either directly or indirectly at a time after all splicing has occurred. Rec is the molecule which shuttles unspliced HML-2 mRNA out of the nucleus (Downey et al., 2015). It does this so that full length transcripts survive any splicing machinery in order that critical proteins such as gag, pro, and pol can be successfully translated. It is possible that androgen increases the activity of the Rec protein and this leads to a drop in full length transcripts. For future work, it would be interesting to attempt to knockdown the Rec transcripts in this context. Such an interference should lead to less Rec

protein and thus a cessation of the movement of unspliced mRNA out of the nucleus.

Aromatase is an enzyme which is responsible for the conversion of androgens into oestrogens via a process called aromatization (Ellem and Risbridger, 2010). Aromatase levels have been studied previously in prostate cancer cell lines. Risbridger et al. concluded that aromatase was expressed in prostate cancer cell lines DU145, PC3 and LNCap but not in a normal prostate epithelium cell line PrEC (Risbridger et al., 2003). They found aromatase levels in the prostate cancer cell lines to be broadly comparable to levels seen in breast cancer cell lines. Because of this presence of aromatase, we were interested in testing the effects of oestrogen on HML-2 expression in prostate cancer cell lines. We tested two of the cell lines from the Risbridger et al. study: PC3 and LNCap. Interestingly, the two cell lines exhibited the opposite results. While the LNCap cells exhibited fold decreases of full length HML-2 mRNA across all concentrations of β-Estradiol, the PC3 cells all exhibited fold increases under the same conditions. The reason for these conflicting results are most likely not related to aromatase levels as they have been shown to be similar amongst the cell lines (Risbridger et al., 2003). One study by Castagnetta et al. has shown that oestrogen receptor levels differ between PC3 and LNCap cells with significantly lower levels found in PC3 (Castagnetta et al., 1997). They also found that treatment with β-Estradiol led to a decrease in cellular proliferation of PC3 cells but an increase in LNCap. Our results showed that β-Estradiol treatment led to an increase in HML-2 levels in PC3 but a decrease in LNCap. Thus our results do not fit into a model by which oestrogens would drive prostate cancer progression via stimulation of high levels of HML-2. Other studies have shown that there is an interacting effect between androgens and oestrogens in the formation of PIN (Risbridger et al., 2003). Future work, should therefore look at the combined effects of oestrogens and androgens in the stimulation of HML-2 in prostate cancer cell lines.

To summarise, treatment with steroid hormones yielded quite varied changes in HML-2 expression dependent on the cell line or hormone used. This is due, firstly, to the fact that some of the cell lines tested have a functional androgen receptor while others do not. Secondly, in terms of the oestrogen treatments, it has been shown that aromatase levels can differ between the cell lines tested.

Overall, from this study we can conclude that HML-2 mRNA and protein are expressed in prostate cancer cell lines and in clinical samples. This study marks the first detailed description of HML-2 HML-2 expression in prostate cancer and thus represents a milestone in this area of basic research. We can also conclude that androgen treatment leads to an increase in expression of full length and spliced HML-2 in androgen dependent LNCap cells. Chapter 2 of this thesis will include data which describes in greater detail the expression of HML-2 in clinical samples in the context of an epidemiological study.

The priority for future research in this area should be to define in greater detail the expression characteristics of HML-2 in prostate cancers using techniques such as immunocytochemistry to track HML-2 protein localization in cells and also to explore other activators like the androgens we have investigated here. Another interesting area to study would be that of the molecular predecessors of prostate adenocarcinoma such as PIA and PIN. Combined studies of methylation changes and HML-2 expression in these diseases may help to elucidate the molecular mechanisms which push these pre malignant diseases past the threshold into full blown adenocarcinoma. Finally, it is vitally important that the actual loci which lead to the HML-2 signal we see here are explored and defined in greater detail as this will be the key to discovering any functional link between HML-2 and cancer initiation or progression.

Chapter 4

HML-2 as a putative biomarker for prostate cancer detection

The following chapter is adapted from Wallace & Downey et al*. Elevated HERV-K mRNA expression in PBMC is associated with a prostate cancer diagnosis particularly in older men and smokers. Carcinogenesis (2014) 35 (9): 2074-2083 (Wallace* et al*., 2014). An appended PDF together with a copyright agreement is situated at the end of this thesis.*

4 HML-2 as a putative biomarker for prostate cancer detection

4.1 Introduction

It is estimated that prostate cancer will account for 28% of all new cancer diagnoses in US men in 2013 with approximately 30,000 deaths expected, ranking it as the second leading cause of cancer death in men in the country (Siegel et al., 2013). Epidemiological studies identified aging, disease family history, race/ethnicity, and obesity and diet as being significant risk factors for the development of the disease (Gronberg, 2003). Other studies suggest that inflammation and infections contribute to disease development (De Marzo et al., 2007; Sutcliffe and Platz, 2008). More recently, the reactivation of endogenous retroviruses in the HERV-K family has been associated with a prostate cancer diagnosis (Goering et al., 2011; Perot et al., 2013; Reis et al., 2013).

Human endogenous retroviruses (HERVs) represent the remnants of ancient germline infections by exogenous retroviruses (reviewed in Bannert and Kurth, 2004) (Bannert and Kurth, 2004) and today compose around 8 % of the human genome (Lander et al., 2001). For the most part, they have become defective over time via the accumulation of inactivating mutations and through silencing by epigenetic mechanisms such as DNA methylation (Schulz et al., 2006). The HERV-K (HML-2) subgroup (hereafter referred to as HML-2) is unique amongst HERVs in that a proportion of its constituent proviruses retain complete open reading frames (ORFs) for all retroviral genes (Bannert and Kurth, 2004). Furthermore, a selection of these proviruses are human specific and polymorphic (Moyes et al., 2007).

The genetic structure of an intact HML-2 provirus consists of open reading frames for the retroviral genes: *gag*, *pro*, *pol* and *env*, flanked by two long terminal repeats (LTRs) which regulate their expression (Bannert and Kurth, 2006). HML-2 consists of two major types defined by the presence or absence of a 292 base pair deletion at the junction of the *pol* and *env* genes which fuses their reading frames (Bannert and Kurth, 2004). Type 1 HML-

2 proviruses harbour the deletion while type 2 remains intact (Bannert and Kurth, 2004). Four HML-2 transcripts have been described to date: full length (*gag*) mRNA, singly spliced *env* mRNA, doubly spliced *rec*/*np9* mRNA and the '*hel'* transcript which lacks any known function (Hohn et al., 2013). High levels of HML-2 mRNA and protein have been observed in a variety of cancers, including germ cell, breast, ovarian, lymphoma and melanoma, but a causal link to any of these diseases remains to be identified (Downey et al., 2015; Ruprecht et al., 2008). HML-2 transcripts have also been detected in prostate cancer cell lines (Agoni et al., 2013a) and tissues (Goering et al., 2011) and a humoral response to the HML-2 gag protein has been observed in sera from prostate cancer patients (Reis et al., 2013). Furthermore, this immune response correlated with disease progression (Reis et al., 2013), indicating that an inflammatory immune response to HML-2 which does not eradicate the HML-2 expressing tumour may promote disease progression.

We hypothesised that HML-2 reactivation could serve as a non-invasive early disease detection marker for prostate cancer and therefore evaluated HML-2 expression in tumour and blood samples. Because peripheral blood mononuclear cells (PBMC) can provide a suitable surrogate to an individual's health status (Liew et al., 2006), and distinct PBMC gene expression profiles have been observed in a number of non-haematological cancers (Mohr and Liew, 2007), we analysed HML-2 mRNA expression in PBMC from prostate cancer patients and healthy volunteers using a case-control design. We also evaluated whether HML-2 reactivation may occur differently in patients of American and European ancestry. Using this approach, we observed differences in HML-2 reactivation between African-American and European-American men and found that blood-based HML-2 expression is a candidate early detection biomarker for prostate cancer.

4.2 Results

4.2.1 HML-2 gag mRNA is elevated in the PBMC of prostate cancer patients.

We compared HML-2 *gag* expression in total RNA isolated from PBMC collected from men without a cancer diagnosis (n=135) and prostate cancer patients (n=294) using qRT-PCR. The characteristics of the controls and cases are shown in **(Table 2)**.

Table 2 - Demographic and clinicopathological features of the male controls and patients with prostate cancer

		Controls (n=135)		Cases (n=294)		*P* value
		n	%	N	%	Controls vs.Cases
T Stage	T1a-c			95	35%	
	T2a-c			153	56%	
	T3a-b			15	6%	
	T4			8	3%	
	Missing			24		
Pathological Stage	p1			54	20%	
	p2			193	71%	
	p3			15	6%	

	p4			8	3%	
	Missing			24		
Gleason Score	≤6			167	59%	
	7			99	34%	
	8			16	5%	
	9			8	3%	
	10			3	1%	
	Missing			1		
Smoking status	Never	26	29 %	95	35%	0.014[1]
	Former	54	61 %	119	44%	
	Current	9	10 %	55	20%	
	Missing	27		47		
Race/ethnicity	AA	75	56 %	142	48%	0.163[1]
	EA	60	44 %	152	52%	
Age at Diagnosis/ Control Recruitment	(mean ± SD)	65.6 ± 8.7 (n=133)		63.7 ± 8.7 (n=294)		0.039[2]
Days from Diagnosis to Recruitment	(median (range))			206.5 (0-705) (n=270)		

PSA at Diagnosis	(median (range))		6.1 (0.2-535) (n=287)	
IFN γ	(median (range))	0.8 (0-29.5) (n=63)	1.0 (0.2-7.9) (n=86)	0.026[3]
IP10	(median (range))	203.8 (32-3922) (n=63)	234 (103-1261) (n=86)	0.035[3]
TNF-α	(median (range))	5.1 (2-161.1) (n=63)	5.15 (3.3-80.3) (n=86)	0.411[3]
IL-1β	(median (range))	1.5 (0.6-21.8) (n=63)	1.7 (0.3-15.5) (n=86)	0.359[3]
Pack years smoked	(median (range))	18.2 (0-67) (n=88)	16.3 (0-112) (n=267)	0.440[3]

[1]Chi square test. [2] Students' t-test. [3] Mann Whitney test. SD = standard deviation. AA = African American, EA = European American. A never smoker was defined as a subject who did not currently smoke and also smoked less than 100 cigarettes in his lifetime. A past smoker did not smoke cigarettes in the 6 months prior to enrolment.

Figure 21 shows that while HML-2 *gag* mRNA is detectable in the PBMC of both controls and prostate cancer patients, HML-2 *gag* levels were significantly elevated in patients compared to controls. Univariate and multivariable logistic regression analyses further showed that HML-2 *gag* mRNA expression is associated with a diagnosis of prostate cancer irrespective of whether the analysis was conducted using *gag* expression as a continuous variable, or comparing high vs. low values of HML-2 *gag* (when using either median or quartile values of HML-2 *gag* as cutoffs) **(Table 3).**

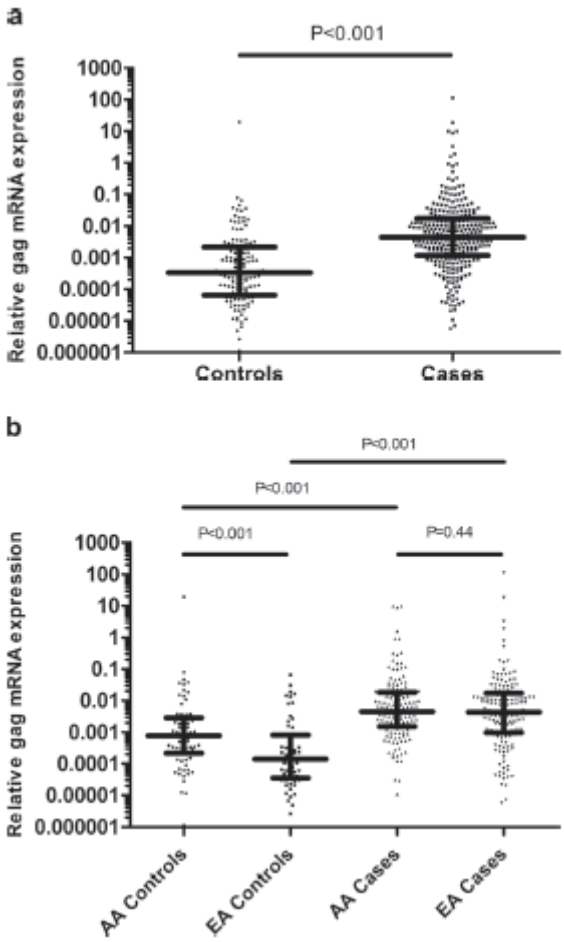


Figure 21 - HML-2 gag expression is elevated in the PBMC of prostate cancer patients versus healthy controls

Comparison of HML-2 gag mRNA expression in PBMC isolated from the blood of healthy controls (n = 135) and prostate cancer patients (n = 294) using previously published gag mRNA primer set. (a) HML-2 gag mRNA was significantly elevated in PBMC isolated from prostate cancer patients compared with the control population (Mann–Whitney test, $P < 0.001$). (b) HML-2 gag mRNA is significantly elevated in PBMC from African-American (AA) controls compared with European-American (EA) controls (Mann–Whitney test, $P < 0.001$), whereas there was no significant difference in the levels of HML-2 gag between African-American and European-American patients (Mann–Whitney test, $P = 0.44$). HML-2 gag was significantly elevated in patients compared with population controls in both African-Americans (Mann–Whitney test, $P < 0.001$) and European-Americans (Mann–Whitney test, $P < 0.001$).

Table 3 - Logistic regression analysis of the association between HML-2 gag mRNA levels in PBMC and a prostate cancer diagnosis

	Univariate Analysis[1]				Multivariable Analysis I[2]				Multivariable Analysis II[3]				Multivariable Analysis III[4]			
	OR	**95% CI**	***P* value**	**n**	**OR**	**95% CI**	***P* value**	**n**	**OR**	**95% CI**	***P* value**	**n**	**OR**	**95% CI**	***P* value**	**n**
Gag mRNA	1.33	1.24 - 1.43	<0.0001	429	1.36	1.26 - 1.47	<0.0001	427	1.49	1.35 - 1.65	<0.0001	358	1.67	1.42 - 1.96	<0.0001	148
Logistic regression comparing high vs. low HML-2 *gag* expression[5]																
Low gag	1				1				1				1			
High gag	5.85	3.67 - 9.36	<0.0001	429	6.02	3.73 - 9.72	<0.0001	427	10.35	5.43 - 19.8	<0.0001	358	24.52	8.61 - 69.8	<0.0001	148
Logistic regression to evaluate a dose response effect after stratification of HERV-K(HML-2) gag expression into quartiles[6]																
1st Quartile	1			48	1			48	1			42	1			30
2nd Quartile	1.38	0.61 - 3.11	0.433	57	1.58	0.69 - 3.62	0.280	57	1.53	0.61 - 3.81	0.363	43	2.25	0.50 - 10.1	0.288	20
3rd Quartile	4.14	1.98 - 8.67	<0.0001	98	5.90	2.69 - 13.0	<0.0001	96	7.15	3.01 - 16.9	<0.0001	80	7.80	2.05 - 29.7	0.003	32
4th Quartile	12.87	6.30 - 26.3	<0.0001	226	17.3	8.07 - 37.0	<0.0001	226	36.4	14.7 - 90.2	<0.0001	193	94.9	22.0 -408	<0.0001	66
		P_{trend}	<0.0001				P_{trend}	<0.0001			P_{trend}	<0.0001			P_{trend}	<0.0001

[1]294 cases and 135 controls. [2]Adjusted for age at diagnosis and race/ethnicity. [3]Adjusted for age at diagnosis, race/ethnicity and smoking status. [4]Adjusted for age at diagnosis, race/ethnicity, smoking status and plasma IFNγ and IP10. [5]HML-2 gag levels were dichotomized using the median expression in the control population. [6]HERV-K gag levels were divided into quartiles based on HML-2 gag quartile distribution in the control population.

These data highlight the robustness of the association between HERV-K marker expression and prostate cancer. For example, an above median *gag* expression in PBMC was associated with a 6-fold increased odds of having a prostate cancer diagnosis in the multivariable analysis [odds ratio (OR) = 6.02 (95% CI 3.73-9.72)], when compared with below median *gag* expression. In addition, our findings revealed a significant dose-relationship between expression levels of HML-2 in PBMC and the likelihood of being diagnosed with prostate cancer **(Table 3)**. Notable, those men with an HML-2 *gag* expression in the highest quartile had a more than 12-fold increased odds [OR =12.87 (95% CI 6.3-26.25)] of being diagnosed with prostate cancer when compared with men in the lowest quartile of PBMC HML-2 expression. After adjustment for age at diagnosis and race/ethnicity, this relationship remained, and men with HML-2 *gag* expression levels in the highest quartile had a 17.3-fold increased odds of being diagnosed with prostate cancer in the multivariable logistic regression analysis [OR =17.3 (95% CI 8.1-37.0)]. Next, we examined the effects of treatment on HML-2 expression in the PBMC of prostate cancer patients but found no significant difference in expression between men who had received prior treatment and those without it Figure 22 - Comparison of HML-2 gag mRNA expression in PBMC isolated from the blood of prostate cancer patients by treatment status **(Figure 22)**.

Additional analyses showed that HML-2 *gag* expression in the PBMC of prostate cancer patients did not correlate with PSA levels at diagnosis (Spearman's rho = -0.01, P = 0.92). Thus, they appear to be independent markers. We also examined if the time from diagnosis to blood draw had an effect on HML-2 *gag* expression in patients but did not find a significant relationship between these two variables (**Figure 23**).

Likewise, HML-2 gag expression did not differ significantly depending on Gleason score at diagnosis **(Figure 24).** In contrast, an analysis by stage showed that HML-2 *gag* expression was significantly higher in stage II patients than in stage I (P = 0.034) (**Figure 25**) while no difference was

found between patients with low stage disease (stage I or II) and the few patients with high stage disease (stage III or IV).

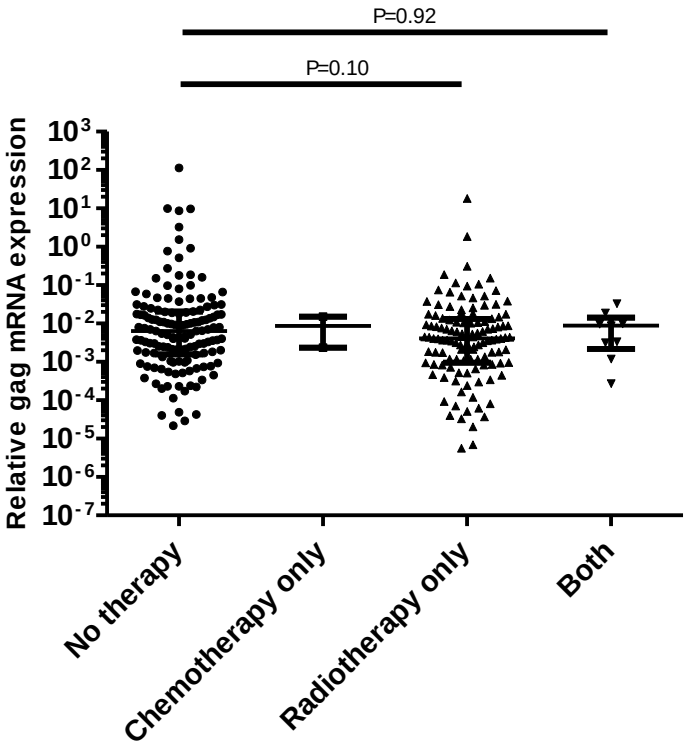


Figure 22 - Comparison of HML-2 gag mRNA expression in PBMC isolated from the blood of prostate cancer patients by treatment status

HML-2 gag mRNA expression did not differ significantly whether patients were either treatment naïve (n=140), received chemotherapy alone (n=2) the figure does not show hormone therapy, received radiotherapy alone (n=120) or had received both (n=9). (Mann Whitney test).

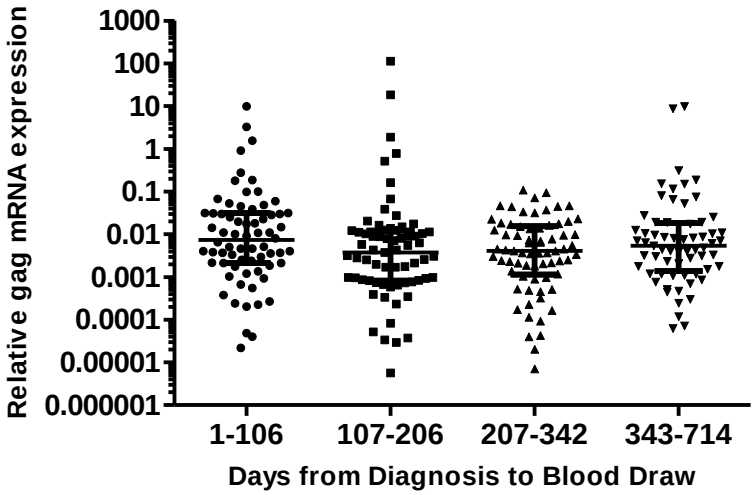


Figure 23 Comparison of HML-2 gag mRNA expression in prostate cancer patients according to days from diagnosis to blood draw.

No significant difference was detected in HML-2 gag mRNA expression depending on the length of time between diagnosis and blood draw for the HML-2 gag measurement (Kruskal Wallis test, P = 0.132; including Dunn's post-hoc test for multiple comparisons testing).

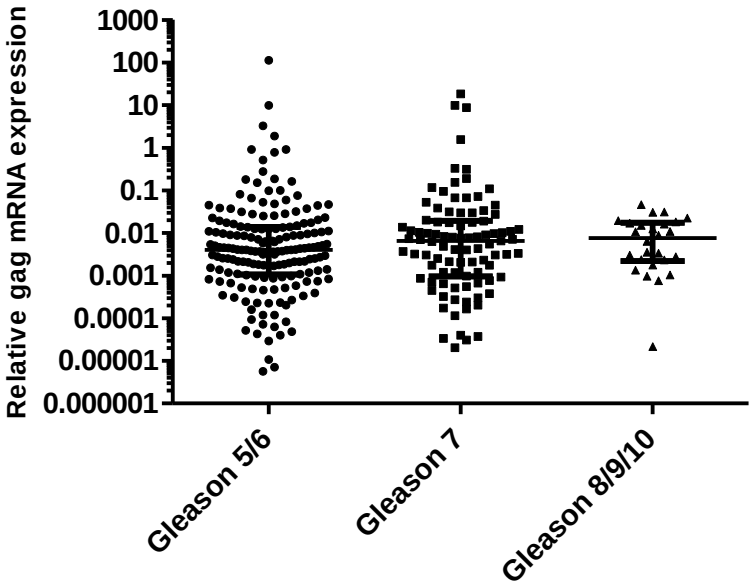


Figure 24 - Comparison of HML-2 gag mRNA expression in prostate cancer patients between different Gleason scores

No significant difference was detected in HML-2 gag mRNA expression depending on Gleason Score at diagnosis (Kruskal Wallis test, P = 0.532; including Dunn's post-hoc test for multiple comparisons testing).

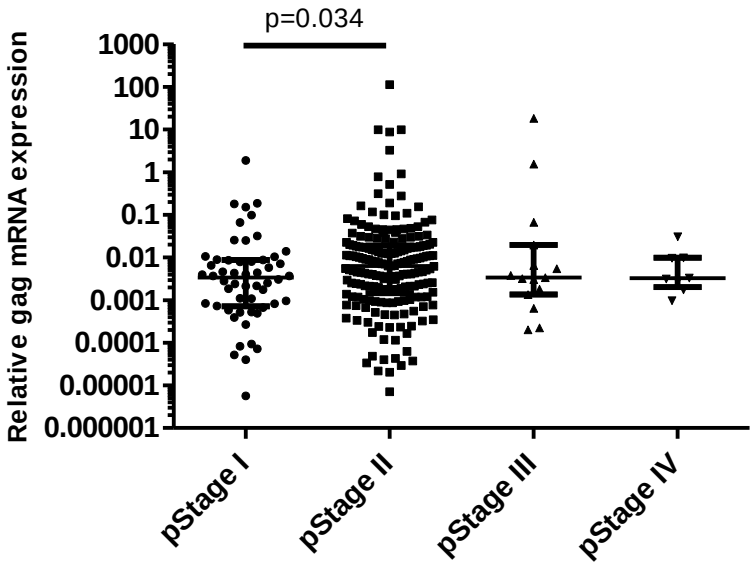


Figure 25 - Comparison of HML-2 gag mRNA expression in prostate cancer patients between different pathological stages (pStage).

A significant increase (1.9-fold) in HML-2 gag expression was detected in pStage II tumours when compared with pStage I tumours (Mann Whitney test, P = 0.034), No significant difference was detected in HML-2 gag mRNA expression when comparing pStage I and pStage II to pStage III or pStage IV tumours.

4.2.2 HML-2 gag mRNA is elevated in African-American controls compared to European-American controls

African-American men are at an increased risk of developing prostate cancer, and are also at an increased risk of diagnosis with advanced disease when compared with European-American men. We therefore examined whether race/ethnicity influences the association of HML-2 *gag* mRNA with prostate cancer. The analysis showed that *gag* expression is associated with a prostate cancer diagnosis in both population groups **(Table 4)** but as shown in **Figure 21**, African-American population controls expressed significantly more HML-2 *gag* mRNA in their PBMC than European-American population controls. Levels of HML-2 *gag* message were not different by race/ethnicity in the case population, suggesting that aberrant

HML-2 *gag* expression in PBMC of prostate cancer patients is similar irrespective of race/ethnicity.

Table 4 - Race/ethnicity-stratified logistic regression analysis of the association between HML-2 gag mRNA levels and prostate cancer

	Univariate Analysis				**Multivariable Analysis***			
	OR	**95% CI**	***P* value**	**N**	**OR**	**95% CI**	***P* value**	**n**
African-American men								
1st Quartile**	1			14	1			14
2nd Quartile	2.29	0.51-10.28	0.279	26	2.18	0.48-9.84	0.311	26
3rd Quartile	4.84	1.22-19.2	0.025	58	5.28	1.32-21.1	0.019	56
4th Quartile	15.30	3.95-59.3	<0.0001	119	15.61	4.01-60.8	<0.0001	119
		P_{trend}	<0.001	217		P_{trend}	<0.001	215
European-American men								
1st Quartile	1			34	1			34
2nd Quartile	1.16	0.42-3.17	0.776	31	1.37	0.48-3.87	0.558	31
3rd Quartile	6.31	2.27-17.55	<0.0001	40	7.54	2.60-21.8	<0.0001	40
4th Quartile	17.78	6.82-46.4	<0.0001	107	20.60	7.56-56.2	<0.0001	107
		P_{trend}	<0.001	212		P_{trend}	<0.001	212

**Adjusted for age at diagnosis. **HML-2 gag levels were divided into quartiles based on HML-2 gag quartile distribution in the control population.*

4.2.3 Association of HML-2 gag mRNA with a prostate cancer diagnosis increases with age and is most robust in older men

The sensitivity of PSA testing decreases with age, partially due to an increase in the prevalence of benign prostate hyperplasia (BPH) which also elevates PSA. To determine whether HML-2 *gag* expression is predictive of prostate cancer across all age group, we performed a stratified analysis comparing the association of HML-2 *gag* according to age groups. We stratified men into three groups similar sized groups, ages 41 to 59 (n=128), ages 60 to 69 (n=179) and ages 70 or older (n=120). The findings from the univariate logistic regression analysis in **Table 5** show that while HML-2 *gag* is predictive of prostate cancer across all ages, the strength of the association increased with age.

Table 5 - Age-stratified logistic regression analysis of the association between HML-2 gag mRNA levels and prostate cancer

	Univariate Analysis				Multivariable Analysis[1]			
	OR	**95% CI**	***P* value**	**N**	**OR**	**95% CI**	***P* value**	**n**
Men of 41-59 years of age								
1st Quartile[2]	1			12	1			12
2nd Quartile	1.12	0.25-4.91	0.880	18	1.22	0.27-5.60	0.795	18
3rd Quartile	3.15	0.76-13.00	0.113	26	4.67	1.02-21.38	0.047	26
4th Quartile	9.8	2.56-37.55	0.001	72	11.96	2.92-48.92	0.001	72
		P_{trend}	<0.0001			P_{trend}	<0.0001	
Men of 60-69 years of age								
1st Quartile	1			23	1			23
2nd Quartile	1.07	0.32-3.63	0.912	22	1.54	0.42-5.54	0.513	22
3rd Quartile	4.15	1.43-12.04	0.009	45	5.53	1.79-17.09	0.003	45
4th Quartile	13.29	4.58-38.57	<0.0001	89	19.52	6.13-62.12	0.001	89
		P_{trend}	<0.0001			P_{trend}	<0.0001	
Men of 70 years of age and older								
1st Quartile	1			13	1			13
2nd Quartile	3.0	0.40-18.24	0.233	17	2.96	0.48-18.06	0.239	17
3rd Quartile	8.25	1.49-45.42	0.015	25	9.07	1.59-51.71	0.013	25
4th Quartile	22.00	4.33-111.7	<0.0001	65	23.94	4.57-125.4	<0.0001	65
		P_{trend}	<0.0001			P_{trend}	<0.0001	

[1]Adjusted for race/ethnicity. [2]HML-2 gag levels were divided into quartiles based on HML-2 gag quartile distributions in the control population.

HML-2 *gag* was most predictive in men aged 70 or older, indicating that combining HML-2 testing with PSA testing may improve the efficacy of prostate cancer detection in these older men. This remained true in the multivariable analysis after adjusting for race/ethnicity. To determine if there is a modifying effect of age at diagnosis on the association between HML-2 *gag* expression and prostate cancer, we performed a statistical interaction test, first on the three individual age categories with HML-2 gag quartile expression, but found that the $P_{interaction}$ was not significant (P values ranged between 0.40-0.96). We then dichotomised age into high/low with the median as cutoff and assessed the modifying effect of the age variable on the association between HML-2 *gag* expression and prostate cancer within each HML-2 expression quartile (quartiles as described in **Table 5**). The test showed that age at diagnosis has a statistically significant modifying effect on the association of HML-2 *gag* with prostate cancer within each HML-2 gag quartile expression (Q2 $P_{interaction}$ = 0.049; Q3 $P_{interaction}$ = 0.019; Q4 $P_{interaction}$ = 0.017). These results are suggestive, but not definitive, of a stronger association of HML-2 with prostate cancer with increasing age.

4.2.4 Association of HML-2 gag mRNA expression with the risk of prostate cancer is modified by smoking status

The smoking status was available for 270 cases and 89 population-based controls in this cohort (total n = 359), categorised as current smoker, former smoker or never smoker. Moreover, pack years smoked information was available for 358 of them. While smoking status or pack years did not directly correlate with the level of HML-2 *gag* expression in the blood samples, smoking status was found to modify the association of HML-2 *gag* expression with a diagnosis of prostate cancer. As shown in **Table 6**, the strength of the association between HML-2 *gag* and disease was found to be highest in current smokers while lowest in never smokers.

Table 6 - Logistic regression analysis of the association between HML-2 gag mRNA levels and prostate cancer by smoking status

	Univariate Analysis				Multivariable Analysis[1]			
	OR	**95% CI**	***P* value**	**N**	**OR**	**95% CI**	***P* value**	**n**
Never Smokers								
Gag mRNA (continuous)	1.41	1.21-1.63	<0.001	121	1.43	1.22-1.67	<0.001	121
Gag mRNA[2] (median)	8.07	3.03-21.5	<0.001	121	9.38	3.33-26.4	<0.001	121
Former Smokers								
Gag mRNA (continuous)	1.44	1.27-1.63	<0.001	173	1.48	1.29-1.69	<0.001	173
Gag mRNA (median)	13.8	6.27-30.4	<0.001	173	17.4	7.29-41.8	<0.001	173
Current Smokers								
Gag mRNA (continuous)	1.93	1.31-2.87	0.001	64	2.35	1.33-4.16	0.003	64
Gag mRNA (median)	25.5	4.56-142	0.001	64	30.2	4.40-208	0.001	64

[1]Adjusted for age at diagnosis and race/ethnicity. [2]HML-2 gag levels were dichotomized using the median expression in the control population.

When stratified by pack years of smoking, there was an increase in the strength of the association between HML-2 *gag* expression and disease with an increase in pack years of tobacco exposure **(**

Table 7**)**. Because of these modifying effects of smoking on the association between HML-2 *gag* expression and disease, we performed an interaction analysis. An test for interaction between smoking status and HML-2 *gag* expression in prostate cancer development with HERV_K *gag* expression categorised as high/low (dichotomised at the median) indicated no statistically significant interaction between HML-2 *gag* and former smoker status ($P_{interaction}$ = 0.40) or HML-2 and current smoker status ($P_{interaction}$ = 0.26). We also performed an interaction test between pack years exposure (dichotomised at the median) and *gag* expression levels (dichotomised at the median). The detected interaction (or modifying effect of tobacco exposure) did not reach statistical significance ($P_{interaction}$ = 0.096). Therefore while HML-2 *gag* in PBMC is more closely associated with prostate cancer among heavy smokers than light or never smokers, there may not be a direct interaction between HML-2 expression and smoking exposure on prostate cancer risk. A larger and more appropriately powered study would be required to test this and show that such an interaction exists.

Table 7 - Logistic regression analysis of the association between HML-2 gag mRNA levels and prostate cancer by pack years

	Univariate Analysis				Multivariable Analysis*			
	OR	**95% CI**	***P* value**	**n**	**OR**	**95% CI**	***P* value**	**n**
0 Pack years								
Gag mRNA (continuous)	1.40	1.21-1.63	<0.0001	120	1.43	1.22-1.67	<0.0001	120
Gag mRNA** (median)	7.97	2.99-21.2	<0.0001	120	9.29	3.30-26.1	<0.0001	120
1-20 Pack years								
Gag mRNA (continuous)	1.45	1.23-1.71	<0.0001	112	1.46	1.23-1.73	<0.0001	112
Gag mRNA (median)	11.4	4.18-31.3	<0.0001	112	14.9	4.90-45.6	<0.0001	112
>21 Pack years								
Gag mRNA (continuous)	1.56	1.32-1.87	<0.0001	123	1.61	1.33-2.96	<0.0001	123
Gag mRNA (median)	24.3	8.63-68.4	<0.0001	123	26.5	8.44-83.1	<0.0001	123

**Adjusted for age at diagnosis and race/ethnicity. **HML-2 gag levels were dichotomised using the median expression in the control population.*

4.2.5 Prostate cancer patients with elevated HML-2 gag show increased serum expression of viral response IFNγ and IP10 expression.

To determine whether elevation of HML-2 *gag* mRNA in PBMC from prostate cancer results in the secretion of viral response interferons and inflammatory cytokines, we measured the expression of IFNγ, IP10, TNFα and IL1β in plasma of a subset of prostate cancer cases and controls. We found that the plasma levels of the viral response cytokines, IFNγ (**Figure 26a**) and its downstream effector IP10 (**Figure 26b**) were significantly elevated in the case population with high levels of HML-2 *gag* mRNA expression. We did not find the same elevated levels of these cytokines in cases with low *gag* expression, or among the control population. Neither TNFα (**Figure 26c**) nor IL1β (**Figure 26d**) were elevated in patients with high levels of HML-2 *gag* mRNA. Multivariable logistic regression (Model 3) in **Table 8** demonstrates that after adjusting for IFNγ and IP10 levels, HML-2 remains significantly associated with prostate cancer.

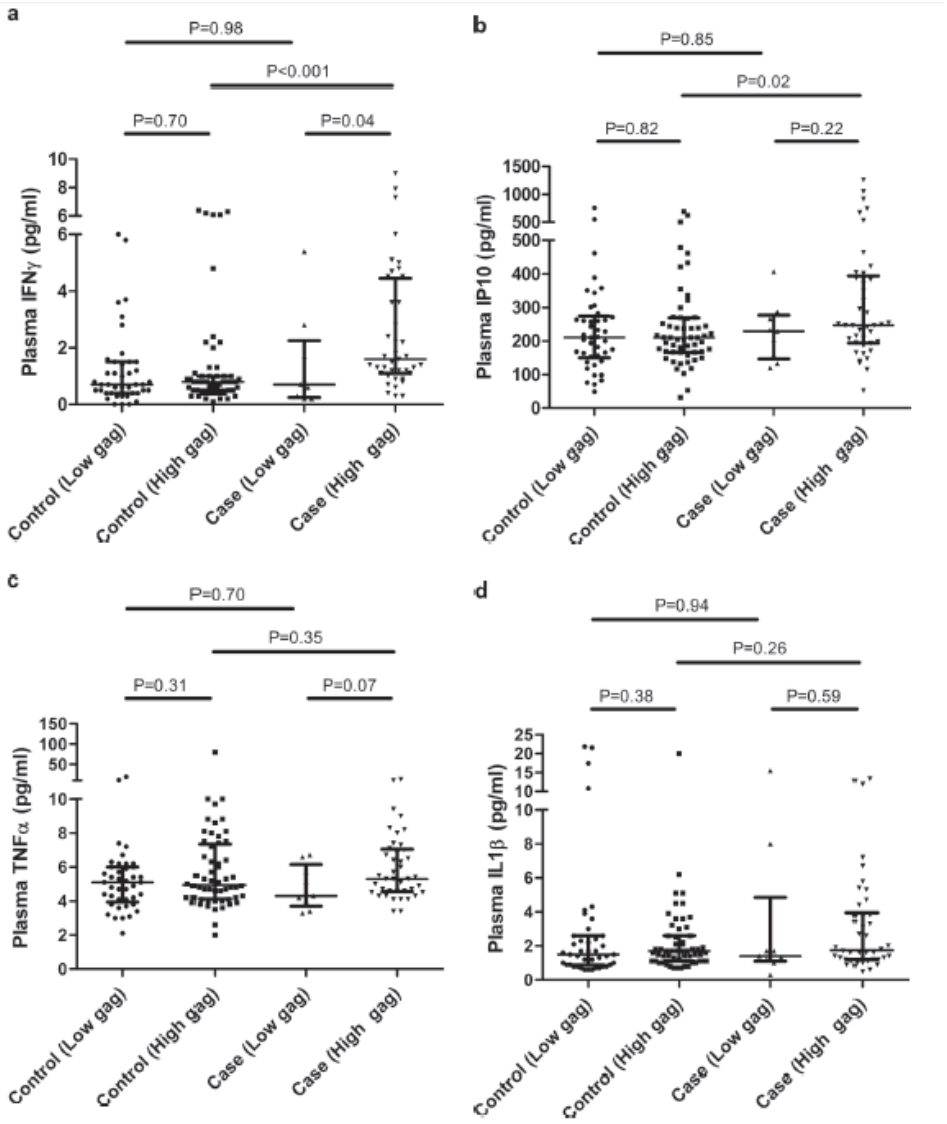


Figure 26 - Association between high HML-2 mRNA expression and elevated plasma IFNγ and IP-10 levels

Plasma IFNγ (a) and its downstream mediator IP10 (b) were significantly increased in patients with high gag expression compared with controls with high gag expression (Mann–Whitney test IFNγ, P < 0.001; Mann–Whitney test IP10, P = 0.02). Additionally, cases with high gag expression had significantly higher levels of plasma IFNγ compared with cases with low gag expression. No significant association of gag expression with tumour necrosis factor-α (c) or IL-1β (d) was observed.

Additionally we examined the impact of IFNγ or IP10, on the association of HML-2 gag mRNA with prostate cancer **(Table 8)**. The univariate logistic regression showed that high levels of HML-2 gag mRNA are more strongly

associated with prostate cancer in patients with high IFNγ or IP10, when compared to those with low IFNγ or IP10. Notable, IFNγ or IP10 levels above the median in the absence of high levels of HML-2 gag, were not found to be associated with prostate cancer. These associations upheld in the multivariable analysis after adjusting for age at diagnosis, race/ethnicity and smoking status.

Table 8 - Logistic regression analysis of the impact of IFN-γ and IP10 levels on the association of HML-2 gag mRNA with prostate cancer

	Univariate Analysis				Multivariable Analysis[1]			
	OR	**95% CI**	***P* value**	**N**	**OR**	**95% CI**	***P* value**	**n**
IFN-γ Effects[2]								
Gag Low/ IFNγ Low	1.0			31	1.0			30
Gag Low/ IFNγ High	0.74	0.16-3.35	0.691	20	0.77	0.17-3.62	0.746	20
Gag High/ IFNγ Low	9.52	3.2-28.3	<0.001	46	11.4	3.6-36.0	<0.001	46
Gag High/ IFNγ High	26.8	8.1-88.5	<0.001	52	31.2	8.9-109.8	<0.001	52
IP10 Effects								
Gag Low/ IP10 Low	1.0			27	1.0			26
Gag Low/ IP10 High	1.51	0.36-6.44	0.575	24	1.59	0.36-7.09	0.543	24
Gag High/ IP10 Low	15.5	4.49-53.4	<0.001	48	16.0	4.45-57.8	<0.001	48
Gag High/ IP10 High	30.2	8.20-111.2	<0.001	50	41.7	10.3-168.5	<0.001	50

[1]Adjusted for age at diagnosis, race/ethnicity and smoking status. [2]HML-2 *gag* levels and IFNγ were dichotomised into low/high with the median as cutoff.

4.2.6 Detection of HML-2 env type I and type II mRNA transcripts in a subset of the PBMC.

We designed primers specific to the env of HML-2 type I and HML-2 type II viruses. Type I and type II are distinguished by a 292bp deletion in the env gene (type I). **Figure 27** shows that both type I and type II env mRNA were elevated in the PBMC of the cancer patients, indicating that the HML-2 re-activation may arise from multiple HML-2 loci in prostate cancer patients.

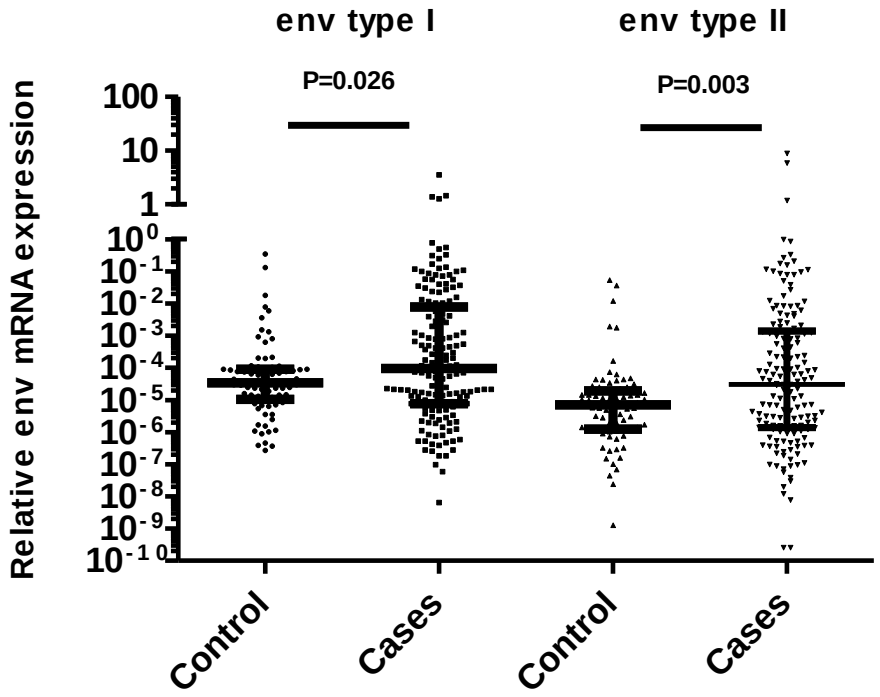


Figure 27 - Comparison of HML-2 env type I and env type II mRNA expression in a subset of PBMC isolated from the blood of healthy controls (n=79) and prostate cancer patients (n=159).

HML-2 env type I and env type II mRNA were significantly elevated in PBMC isolated from prostate cancer patients when compared with PBMC isolated from the control population (Mann Whitney test $_{\text{Type I env}}$, P = 0.026; Mann Whitney test $_{\text{Type II env}}$, P = 0.003).

4.2.7 HML-2 env protein is expressed in prostate tumours.

We used immunohistochemistry to assess whether prostate tumours express the HML-2 env protein using whole tissue sections to examine protein localization and a tissue microarray (TMA) for scoring. Consistent with our findings analyzing HML-2 expression levels in PBMC, expression of the HML-2 env protein in tumours varied considerably among patients. **Figure 28a-d** shows that HML-2 env protein expression is cytoplasmic and membrane located, and also localizing toward the lumen of the gland. Env protein expression levels were significantly higher in African-American patients than European-American patients (**Figure 28e**), and 61% of the

African-American patients presented with tumours that had high env expression (above median) while only 40% of the tumours from European-American patients fell into the same category ($P < 0.001$). In contrast, we did not detect aberrant HML-2 env expression in prostate tissues from patients with BPH (not shown). Western blot analysis of cell extracts further corroborated the presence of both HML-2 gag and HML-2 envelope protein in human prostate cancer cells, while expression was low to absent in the non-tumourigenic RWPE-1 cells (**Figure 28f**).

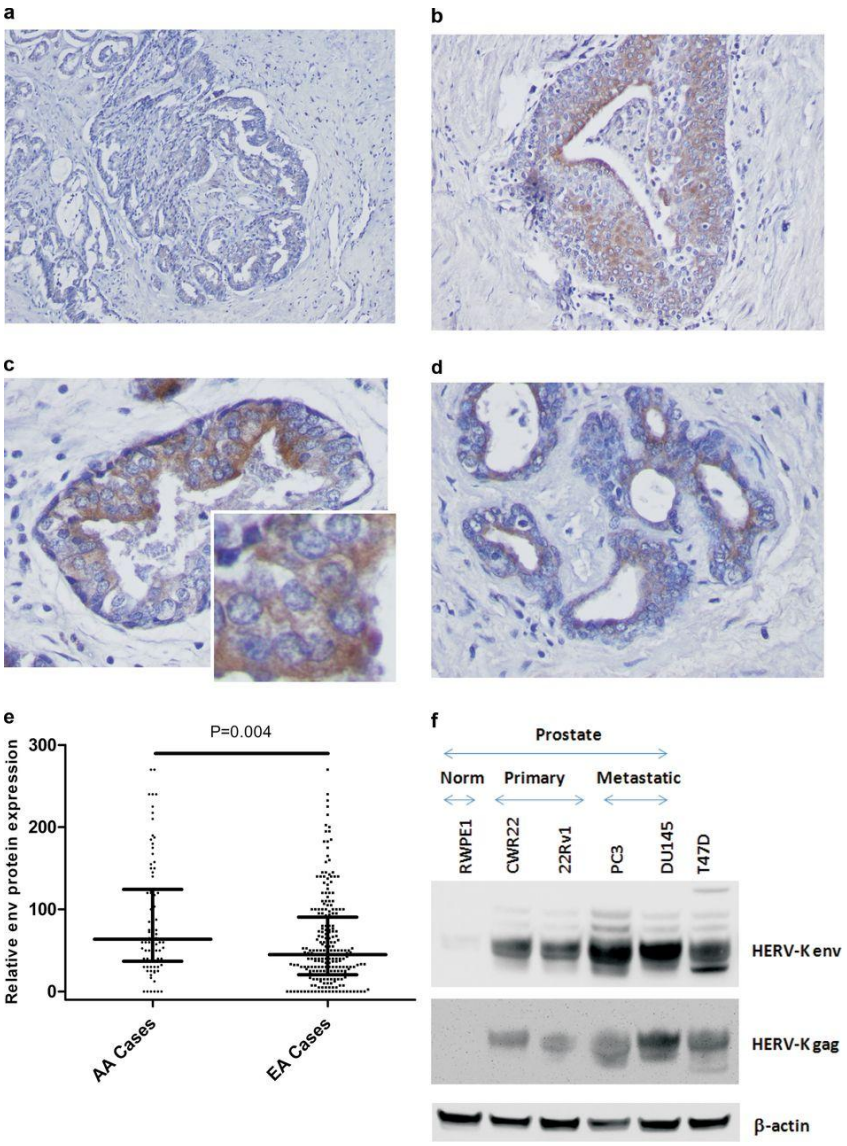


Figure 28 - HML-2 expression in human prostate adenocarcinomas.

HML-2 expression in human prostate adenocarcinomas. Shown is IHC analysis of four invasive adenocarcinomas for expression of HML-2 envelope (env) protein using the monoclonal anti-envelope antibody, 6H5. (a) Scattered positivity for HML-2 env expression in the tumour epithelium, showing a low to moderate antigen expression as indicated by the brown chromogen deposits. (b–d) Locally intensive staining for env expression in the tumour epithelium. Staining shows a cytosolic to membrane distribution with a more intensive staining of cancer cells toward the luminal side of the cancerous gland (c and d). a and b: Magnification: ×100. c and d: Magnification: ×400. Inset: higher resolution image for the env protein-positive tumour epithelium. Counterstain: Hematoxylin. (e) Immunostaining for env protein in human prostate tumours using a tissue microarray that included tumours from both African-American (n = 105) and European-American patients (n = 272). On

average, tumours from African-American patients showed a higher expression of the HML-2 env protein than tumours from European-Americans (Student's t-test, $P < 0.001$). (f) HML-2 gag and env protein expression were detected in human prostate cancer cell lines (CWR22, 22Rv1, PC-3 and DU145) but were not detected in the non-tumourigenic human prostate cell line, RWPE1. An extract from the HML-2-positive T47D human breast cancer cell line were included as a positive control.

4.3 Discussion

Our study makes the novel observation that HML-2 mRNA expression is aberrantly increased in PBMC from prostate cancer patients when compared with healthy male controls. Thus, the evaluation of blood-based HML-2 expression may serve an early disease detection biomarker in prostate cancer. Moreover, the expression as of HML-2 as an early disease biomarker may perform better in older than younger men, whereas the sensitivity of PSA testing decreases with age. Thus, combining HML-2 testing with PSA testing may improve the efficacy of prostate cancer detection in this age group. Other studies have observed the presence of autoantibodies to HML-2 gag in sera from prostate cancer patients and described the aberrant expression of HML-2 mRNA and proteins in prostate cancer tissues (Goering et al., 2011; Ishida et al., 2008; Reis et al., 2013). Hence, our observation of aberrantly increased blood-based expression of HML-2 in prostate cancer patients is consistent with these previous findings.

Prostate specific antigen (PSA) testing is the gold standard for prostate cancer screening. The traditional cut-off for an abnormal PSA level is 4.0 ng/ml. One of the major problems is that PSA has poor discriminating ability in men with symptomatic BPH versus those with prostate cancer (You et al., 2010). Additionally, PSA levels rise as men grow older, which can lead to increased false positive PSA tests and unnecessary biopsies (Gulati et al., 2011). Consistent with these data, PSA levels displayed a slight albeit non-significant increase with age at diagnosis in our patient cohort (Spearman's rho = 0.09, P=0.13), while HML-2 *gag* expression did not increase with either age at diagnosis for the cases (Spearman's rho = -0.01, P=0.91), or the age of recruitment for the controls (Spearman's rho = -0.04, P=0.61). Currently the American Urological Association only recommends the use of PSA screening in asymptomatic men between the ages of 55 to 69 (Carter, 2013). Our data suggest that blood-based HML-2 is a candidate biomarker for the detection of prostate cancer, potentially with a focus on older men. Future research must determine its predictive value in conjunction with PSA testing.

Incidence and mortality rates of prostate cancer are significantly higher in men of African ancestry when compared with men from other population groups in the United States, the Caribbean and the United Kingdom (Martin et al., 2013). This cancer health disparity may relate to unknown causative factors that influence disease pathology in men of African ancestry and induce a more aggressive disease among them. Therefore, we examined HML-2 expression in tumours and blood samples from both African-American and European-American patients. Indeed, we found that HML-2 expression in prostate tumours is significantly higher in the African-American patients, when compared with European-American patients, based on the immunohistochemical analysis of HML-2 envelope protein expression in these tissues. However, we did not find that HML-2 message levels are significantly different in PBMC from these two patient groups. Accordingly, HML-2 as biomarker should perform similarly among these patients. However, we found that baseline levels of HML-2 mRNA in PBMC of healthy controls were greater in African-Americans than European-Americans. Currently we do not know why African-American men may express higher baseline levels of HML-2, nor what effects, if any, there may be. However, it is known that retroviral reactivation is induced by both hormone exposure and stress signalling (Hohn et al., 2013; Serafino et al., 2009; Wang-Johanning et al., 2003), and these signalling pathways could become more commonly activated in African-American men. Additionally, Macfarlane and Simmonds reported that the frequency of allelic variation in the various HML-2 germline loci is greater in African populations than populations from Europe and Asia (Macfarlane and Simmonds, 2004). Furthermore, others reported significantly higher insertion frequencies of HERV-K113 (21%) and K115 (35%) in African-Americans compared with European-Americans (K113 9% and K115 6%) within the United States (Jha et al., 2009). Because African-Americans may have inherited a greater number of polymorphic HML-2 loci than European-Americans, this would increase the number of HML-2 loci that are potentially transcribed, leading to a higher baseline expression of HML-2 encoded genes from these loci.

Cigarette smoking is not associated with early disease development in prostate cancer (Giovannucci et al., 1999; Hickey et al., 2001), but it is associated with an increased risk of fatal prostate cancer (Giovannucci et al., 1999; Huncharek et al., 2010; Kenfield et al., 2011). It is also associated with increased risk of biochemical recurrence and metastasis (Moreira et al., 2013). Gabriel et al. previously showed that exposure of both normal human dermal fibroblasts and benign human uroepithelium to urine from current smokers increased the transcription of HERV-K (HML-6), indicating that HML-2 may be induced by tobacco metabolites (Gabriel et al., 2010), however the authors did not specifically test the HML-2 family. A recent study by Prueitt et al. found an antigenic response to nicotine from smoke in prostate cancer in the form of immunoglobulin deposits in tumours and an increased risk of metastasis (Prueitt et al., 2016). Because of these observations, we decided to evaluate the influence of smoking on the application of blood-based HML-2 expression as a disease biomarker. We had smoking status and pack-year information for a large subset of subjects in our case-control study, and therefore assessed the impact of smoking status on the association of HML-2 with prostate cancer. This analysis led to the finding that HML-2 was more predictive of prostate cancer in current smokers, and the association increased in strength with increased pack-years smoked. The underlying mechanism for this modifying effect of smoking on the relationship between HML-2 and prostate cancer remains unclear, and we could not demonstrate a statistically significant interaction between smoking, HML-2 and risk of prostate cancer in this study. Yet, we think our observation could be of particular significance because those patients who are current smokers tend to develop a more aggressive disease than other patients, as already pointed out by us, and improved disease detection at an early stage for this patient group may have a significant impact in reducing disease mortality.

The finding that radiotherapy did not affect HML-2 expression levels in the cohort is in conflict with previous findings from Agoni et al who discovered that gamma-irradiation led to increases in HML-2 expression in prostate cancer cell lines (Agoni et al., 2013b). This discrepancy is probably due to the fact that certain physiologically mechanisms exist *in vivo* which are not found in cancer cell lines which alter the way in which radiation affects HML-2 expression.

We also observed that elevated HML-2 mRNA in the PBMC of prostate cancer patients was associated with increased serum IFNγ and its effector IP10. It has been previously reported that patients with HML-2-positive tumours exhibit a humoral response to HML-2. Rakoff-Nahoum et al identified HML-2-specific T cells in seminoma patients, that display elevated IFN-γ secretion in response to HML-2 gag peptides (Rakoff-Nahoum et al., 2006). Wang-Johanning et al also showed that breast cancer patients mount anti-HML-2 responses, including production of anti-HERV-K(HML-2) env IgG, production of IFNγ and a T-helper cell cytokine response signature including increased production of IL-2, IL-6, IL-8 and IP10 during in vitro stimulation of BC PBMC with HML-2 antigen (Wang-Johanning et al., 2008). In this context, the finding of an interferon response in patients with an elevated blood-based HML-2 message is consistent with a humoral response to HML-2 that was observed by others, indicating aberrant HML-2 expression in prostate cancer patients not only affects tumour immunobiology but rather triggers a systemic antiviral response. The relationship of this antiviral response with disease outcome is still largely unknown, but Reis et al (2013) observed that 6.8% of patients with prostate cancer had serum antibodies to the gag-HML-2 protein encoded by the ch22q11.23 locus, and these patients tended to have a more aggressive disease and a higher disease mortality (Reis et al., 2013). We previously reported differences in the prostate tumour immunobiology between African-American and European-American men (Wallace et al., 2008). Prostate tumours from African-American men were characterised by the activation of immune response and host defence pathways. Johnston et al

2001 reported that activated macrophages display increased expression of HML-2, HERV-W and HERV-H mRNA, indicating that HERV reactivation may occur as a consequence of an elevated immune activity (Johnston et al., 2001). A frequent activation of immune response pathways in tumours from African-American men may therefore contribute to the increased expression of HML-2 env protein in these tumours that we observed in the current study.

The levels of HML-2 gag and env protein seen in Figure 11 and Figure 12 do not correlate with the levels shown by western blotting in figure Figure 28. This may be due to non specific binding of the antibody in Figure 11 and Figure 12 as RWPE-1 would not be expected to express HML-2 protein as it is a non tumourigenic prostate cell line. Also, the anti-env antibody used in Figure 28 was different to that used in the other figures and so this may be the reason behind the conflicting results.

One of the limitations of this study is our inability to tell where exactly the elevated HML-2 expression arises from in our PBMC from prostate cancer. HML-2 has previously been reported as being expressed at basal levels in healthy PBMC, with aberrant expression occurring in leukaemia cells (Brodsky et al., 1993a; Depil et al., 2002; Medstrand et al., 1992). However, circulating tumour cells (CTCs) are detectable in the blood of prostate cancer patients (approx. 2-8 per 7.5ml of blood), with higher levels of CTCs correlating with bone and visceral metastasis (Thalgott et al., 2013). Prostate CTCs reflect the biology of the cancer with the CTCs gene signatures switching from AR-on to AR-off within the first month of androgen deprivation therapy, indicating their sensitivity as indicators of tumour response to treatment (Miyamoto et al., 2012). When isolating PBMC from blood, CTCs may also be pulled down in the PBMC fraction (Kallergi et al., 2007). Therefore PBMC with high levels of HML-2 in prostate cancer patients may potentially reflect HML-2-positive CTCs. Another potential source of HML-2 is from exosomes secreted by tumours. Exosomes are nanoscale membrane vesicles which are secreted from cells and are thought to be important intercellular communicators, or, in a cancer

setting, drivers of metastatic spread (Thery et al., 2002). A recent study has now implicated HERVs in this process, with the finding that HML-2 mRNA is selectively packaged into tumour exosomes and that this genetic material can be transferred to normal cells (Balaj et al., 2011). Exosomes can be found abundantly in the blood of patients and may contribute to tumour dissemination (Zhang and Grizzle, 2014). Future work will focus on investigating these options and identifying the specific HML-2 loci responsible for elevated HML-2 in the PBMC of prostate cancer patients and will also evaluate the relative contribution of CTCs and exosomes to the HML-2 expression signal in PBMC from cancer patients.

In summary, we made the novel observation that blood-based HML-2 expression is a candidate early detection biomarker for prostate cancer that, may specifically improve disease detection among older men. Furthermore, we obtained evidence that this test would perform similarly in African-Americans and European-Americans. Lastly, aberrant HML-2 expression seems to occur more frequently in both healthy African-Americans and African-Americans with prostate cancer than in their European-American counterparts. While intriguing, these data need to be further studied.

Chapter 5

Cloning and overexpression of HML-2 in prostate cell lines

5 Functional analysis of HML-2 Rec

HML-2 is defined as a complex retrovirus because of its expression of the accessory proteins Rec and Np9 (Downey et al., 2015). These two proteins are transcribed from spliced transcripts of type II and type I HML-2 respectively. These proteins are of upmost interest due to their many links with diseases such as cancer and autoimmunity (Downey et al., 2015). Rec (formally known as c-ORF) is a 14 kDa protein which shares functionally homology with the HIV-1 rev protein in that its main function is to shuttle unspliced mRNA out of the nucleus (Bannert and Kurth, 2004). It does so by binding to Rec response elements (RcREs) in the 5' region of target mRNAs. Such a function is critical in the regulation of a retrovirus as it allows a balance between the translation of proteins from spliced transcripts and also from full length ones (Coffin et al., 1997). It is not known if Rec preferentially shuttles type I transcripts over type II.

A number of studies have identified Rec as a putative oncoprotein in cancer. Firstly, Boese et al. discovered that stable overexpression of Rec in Rat-1 Rat fibroblast cells led to cellular transformation and that Rec could support tumour formation in nude mice (Boese et al., 2000). These tumours were large and pleomorphic and infiltrated the subcutaneous fat. The same group built upon this work by finding that mice which were transgenic for Rec were prone to seminomas and developed lesions which resembled carcinoma *in situ* (Galli et al., 2005). They concluded that Rec may lead to carcinogenesis via its repression of the c-myc oncogene through its interactions with the transcriptional repressor PLZF. Furthermore, they found that Rec could lead to the derepression of the AR through its ability to form a trimeric complex with the AR and its repressor TZFP (Kaufmann et al., 2010). Similarly, Hanke et al. also discovered that Rec interacted with a repressor of the AR known as hSGT (Hanke et al., 2013). They found that Rec binding of hSGT led to a significant increase in AR activity.

These studies – particularly those concerning the AR – led us to believe that Rec may play a role in prostate cancer. We hypothesised that via its repression of the AR, Rec can contribute to androgen independent prostate

cancer by dysregulating androgen signalling. We set out to functionally characterize Rec in prostate cancer. To begin we would clone the Rec gene from prostate cancer cDNA derived from cell lines and construct a plasmid which would allow us to overexpress Rec in nontumourigenic RWPE-1 prostate cells. We would then complete assays which would look for malignant signs of transformation in cells overexpressing Rec. We also planned to conduct a gene expression microarray which would investigate changes in normal RWPE-1 compared to Rec overexpressed RWPE-1.

5.1 Construction of a Rec overexpressing plasmid

Our goal was to construct a plasmid via molecular cloning which would be used to overexpress Rec in RWPE-1 prostate cells. We had no access to a Rec antibody at this stage and so we planned to FLAG-tag the Rec gene. This would enable us to functionally characterise Rec through the use of an anti-FLAG antibody.

5.1.1 Cloning Strategy I – High fidelity polymerase

Our cloning strategy involved the use of primers which would contain start and stop codons, an N terminal FLAG-tagged Rec sequence as well as restriction sites for subsequent subcloning into an overexpression vector **(Figure 29)**. The final amplicon was expected to be 360 base pairs in length. The primers were based on the ERVK-6 HML-2 provirus. This provirus has been previously shown by Agoni et al. to be expressed in DU145 prostate cancer cells (Agoni et al., 2013a).

Figure 29 - Primer design for Rec cloning

The upstream primer contains the FLAG-tagged Rec N terminal sequence and a start codon. A *Kpn*I restriction site facilitates subcloning. The GA clamp allows for the binding of the restriction enzyme to the sequence. The downstream primer contains the Rec C terminal sequence as well as a stop codon and an *Eco*RI site to facilitate subcloning.

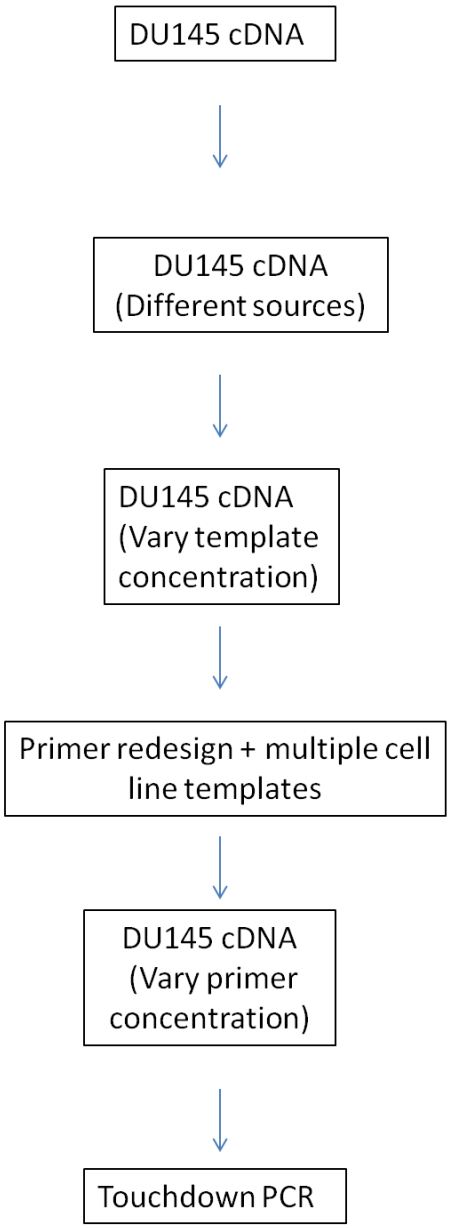


Figure 30 - Rec cloning optimization

Various optimization efforts were undertook to clone Rec including varying template source and concentration, redesigning primers, varying primer concentration and advanced methods such as touchdown PCR.

The optimisation workflow for Rec cloning is shown in **Figure 30**. DU145 was chosen first as the cell line in which to attempt to clone Rec from due to the previous report of its expression of ERVK-6 by Agoni et al (Agoni et al., 2013a). A high fidelity polymerase was chosen to ensure that no

mutations were introduced into the final protein. This PCR from DU145 cDNA led to no amplification of the correct band being observed **(Figure 31)**.

Different sources of DU145 cDNA template were used which originate from a different reverse transcriptase enzyme. However, this only led to the amplification of non-specific amplicons **(Figure 32)**. It was decided to redesign the downstream Rec-FLAG primer, specifically to shorten it in order to bring its melting temperature to a temperature closer to that of the upstream primer. It was hoped that this would improve the priming of Rec from the cDNA. The 'Superscript' source of DU145 cDNA was chosen. This new primer set resulted in the amplification of a non specific band of ~150 bp **(Figure 33)**. Primer dimers were also observed.

The next step was to vary the primer concentration in order to find a concentration which would alleviate the primer dimers and allow amplification of the Rec-Flag band. However, the results found the same amplification of the non specific band of ~150 bp together with the primer dimers **(Figure 34)**.

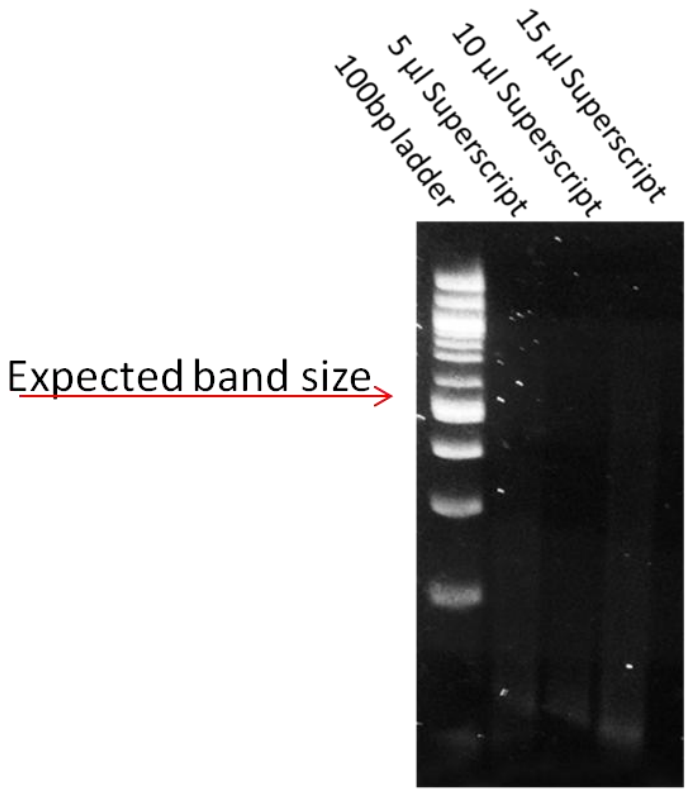


Figure 31 - Cloning Rec from DU145 cDNA

Lane 1 represents 100bp ladder. Lanes 2-4 represent increasing concentrations of DU145 'Superscript' cDNA template. No amplification was evident in any of the reactions. The faint bands near the bottom of the gel represent primer dimers.

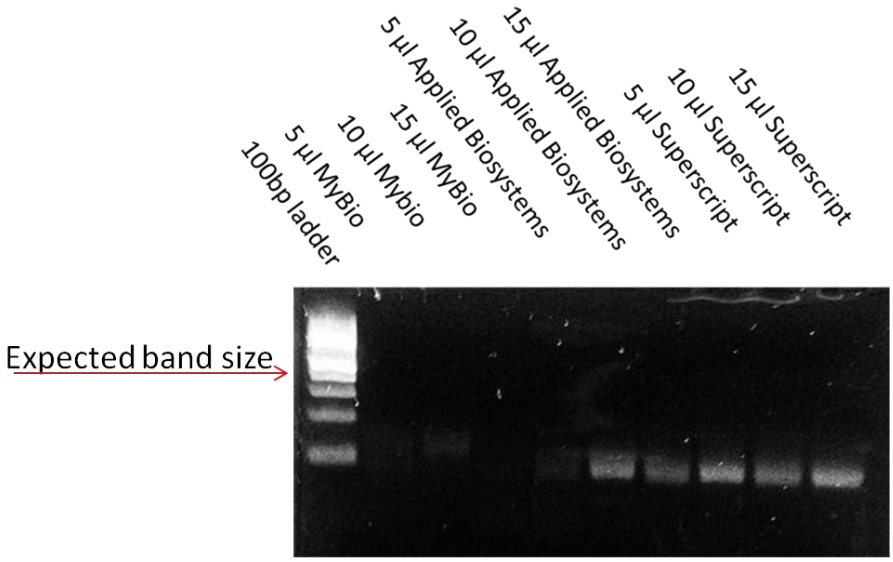


Figure 32 - Cloning Rec from DU145 cDNA: Varying template and sources

Varying amounts of template and different sources of DU145 template were attempted. Lane 1 represents 100 bp ladder. Lanes 2,3 and 4 represent 5, 10 and 15 µl respectively of 'MyBio' cDNA template. Lanes 5, 6 and 7 represent 5, 10 and 15 µl respectively of 'Applied biosystems template'. Lanes 8, 9 and 10 represent 5, 10, and 15 µl respectively of 'Superscript' template. Non specific amplification was observed in the 'Applied Biosystems' and 'Superscript' cDNA templates. However, the band of interest of 360 bp was not seen in any lane.

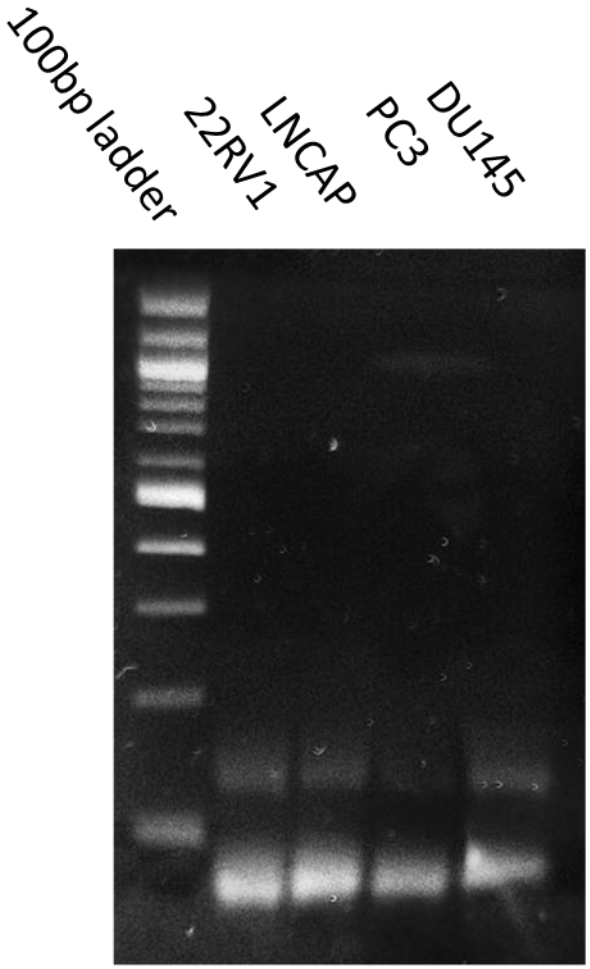


Figure 33 - Cloning Rec using rec-flagn-2 primer set

Lane 1 represents 100 bp ladder. Lanes 2- 5 represent, 22RV1, LNCap, PC3 and DU145 cDNA templates respectively. Primer dimers can be seen at the bottom of the gel but another band of ~150 bp is also present in all reactions.

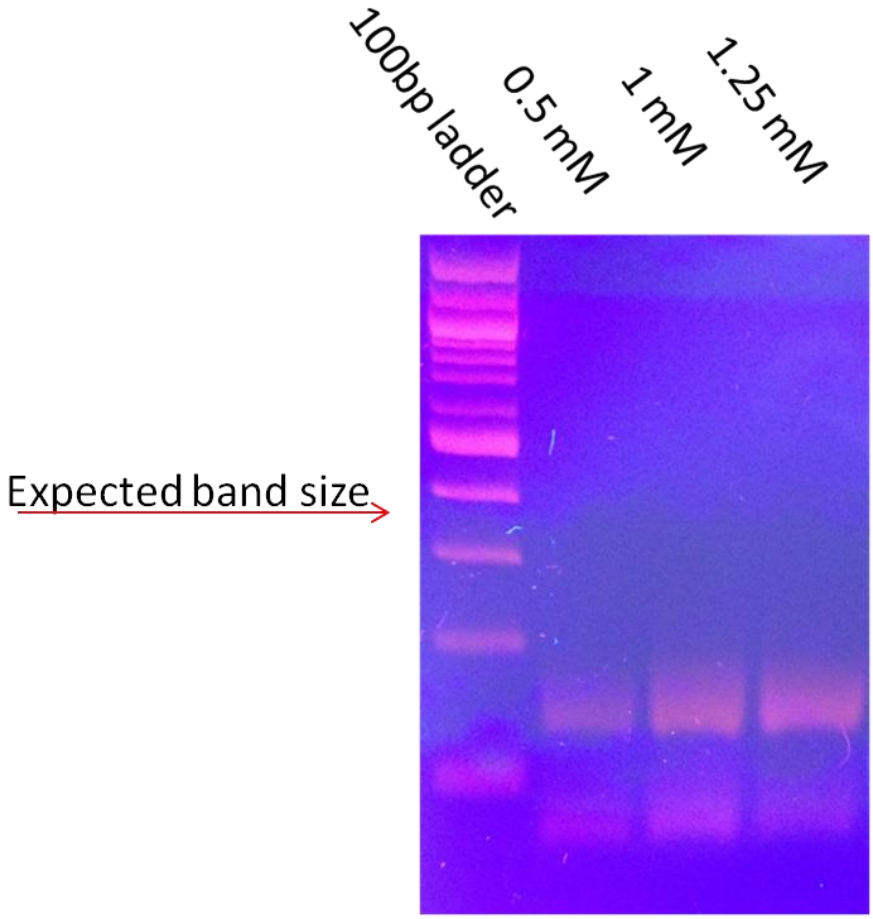


Figure 34 - Cloning Rec from DU145 cDNA – Varying primer concentration

Lane 1 represents 100 bp ladder. Lanes 2-4 represent 0.5,1 and 1.25 mM concentrations of primer respectively. Primer dimers can be seen at the bottom of the gel. The band of interest of 360bp was not amplified but a non specific band of ~150 bp was observed in each reaction.

5.1.2 Cloning Strategy II – Touchdown PCR

It was eventually decided to attempt a touchdown PCR protocol to try and alleviate the primer dimers that were appearing. Touchdown-PCR is a technique which starts off with an annealing temperature which is greater than the predicted tm of the primer sequences and gradually lowers this temperature throughout the cycles until the 'touchdown' temperature is reached. This process helps to increase the likelihood of specific amplicons

being produced in the early cycles of the reaction which will then tilt the balance in favour of the final desired amplicon being produced in large amounts at the end of the reaction at the expense of primer dimers and non specific amplicons. However, this approach also did not result in the correct amplicon being produced **(Figure 35)**.

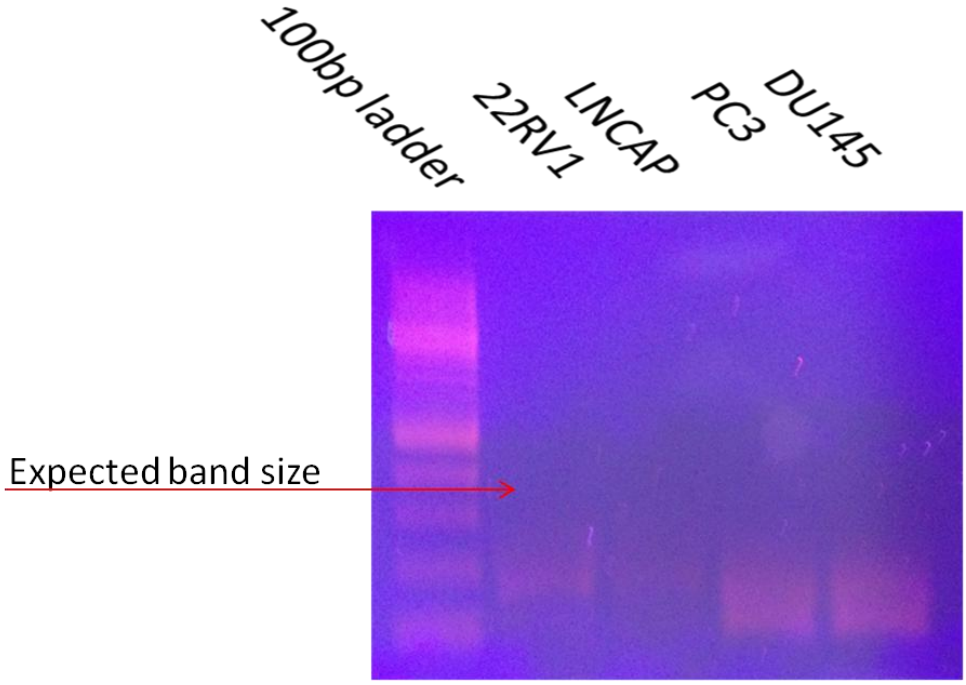


Figure 35 - Touchdown PCR – Rec cloning

A touchdown PCR using different prostate cancer cell line cDNA templates was attempted. Lane 1 represents 100 bp ladder. Lanes 2-5 represent 22RV1, LNCap, PC3 and DU145 template respectively. A non specific band of ~150 bp was observed in each reaction but was strongest in PC3 and DU145. The band of interest of 360 bp was not seen.

Previously, I had used primers k7/k21 in an RT-PCR reaction to amplify Rec and Np9 from prostate cancer cell line cDNA **(Section 3.1) (Figure 36).** Template from these completed reactions was chosen for a new round of PCR reactions with the Rec-FLAG primers as it should be enriched for the Rec and Np9 coding sequences. The Rec-FLAG primers are situated internally of the k7/k21 primers and so they should be able to amplify the

Rec coding sequence more easily from this template compared to total cDNA which would represent the whole genome **(Figure 37).** This process should also reduce the chances of non specific amplification occurring. The coding sequence together with the FLAG tag, start/stop codons, restriction sites and clamp make up an expected amplicon of 360 bp **(Figure 38).** This k7/k21 template was PCR cleaned in order to remove any primers that could be still present and interfere with the new primer set.

The results revealed the presence of an incorrect amplicon size of ~150 base pairs being visualised on the gel **(Figure 39).** A subsequent PCR using a no template control revealed that this band was not primer dimers and so it most likely represented Np9 or a non specific band.

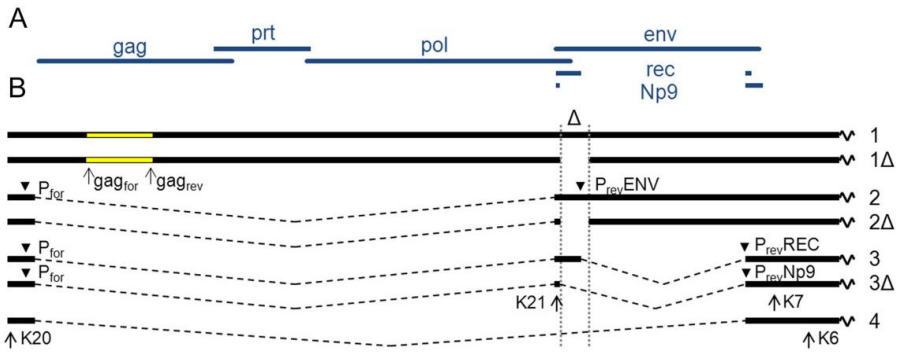


Figure 36 - Schematic of HML-2 splicing and K7/K21 primer location

The k7 and k21 primers are situated 5' and 3' respectively of the Rec transcript. This primer set will also amplify np9 transcripts. The Rec coding sequence is situated internally of this primer set and therefore template from a k7/k21 PCR should be enriched for the Rec coding sequence. (adapted from (Fuchs et al., 2013))

Figure 37 - Rec-FLAG primers

The Rec-FLAG primers are situated inside the predicted k7/k21 amplicon with the k21 primer partially overlapping that of the upstream Rec primer. Using the Rec-Flag primers on an enriched template of cleaned k7/k21 amplicon should theoretically increase the chances of the Rec-FLAG primers successfully amplifying and at the same time reducing the chances of non specific binding which may occur when the Rec-FLAG primers are used against the whole genome. The start of the k21 primer can be seen in green and the k7 primer can be seen in blue. The Rec-FLAG primers are seen in grey, with the forward primer overlapping with k21.

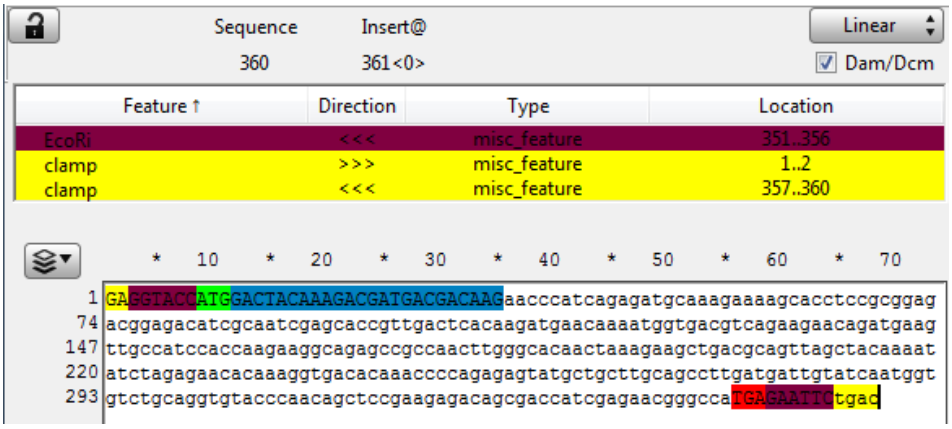


Figure 38 - Rec-FLAG amplicon

The Rec-FLAG amplicon is expected to be 360 bp in length. This consists of the Rec coding sequence (white), the FLAG tag (blue), start and stop codons (green/red), restriction sites (brown) and clamps (yellow).

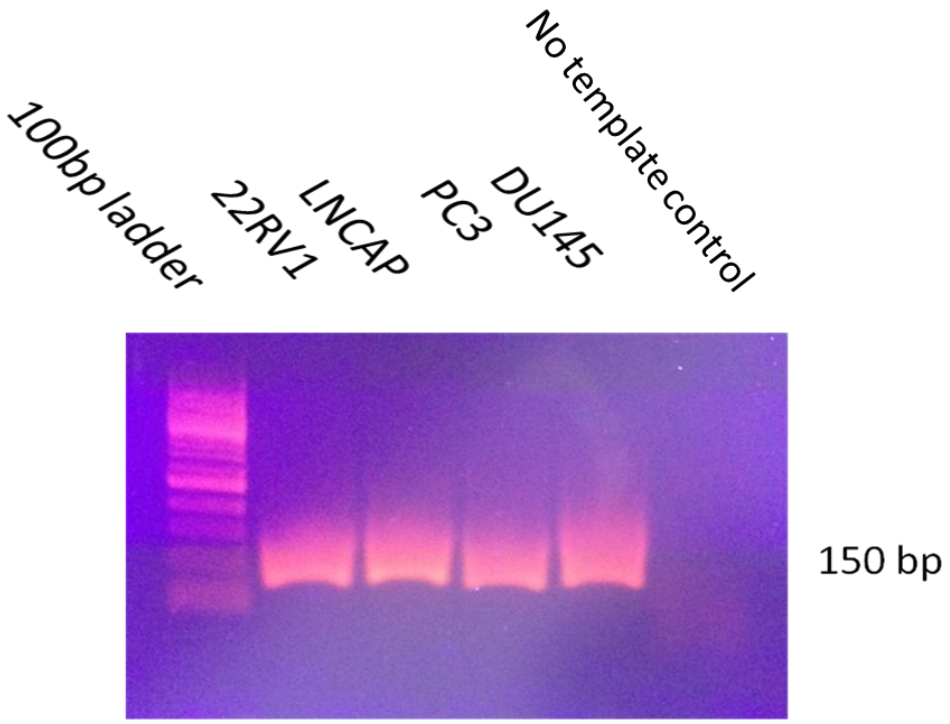


Figure 39 - Non specific amplicon

The Rec-FLAG primers on the k7/k21 template led to the amplification of a band of ~150bp. This was 110bp shorter than the expected Rec-FLAG amplicon and so it must represent either a non specific band or an Np9 amplicon. Lane 1 represents 100 bp ladder. Lanes 2-5 represent 22RV1, LNCap, PC3 and DU145 cDNA templates respectively. Lane 6 represents a no template control.

5.1.3 Cloning Strategy III - Non FLAG-tagged cloning attempt

To troubleshoot the cloning of Rec mRNA, primers were redesigned against Rec mRNA without any added FLAG tag or restriction sites. This would allow for a more specific reaction to take place which would hopefully discourage the amplification of any non specific bands. A gradient of melting temperatures was also tested in order to optimize the reaction. The template chosen was the PC3 template from the original k7/k21 RT-PCR **(Chapter 3)**. Results from these new primers revealed the amplicon of interest at 312 bp **(Figure 40).**

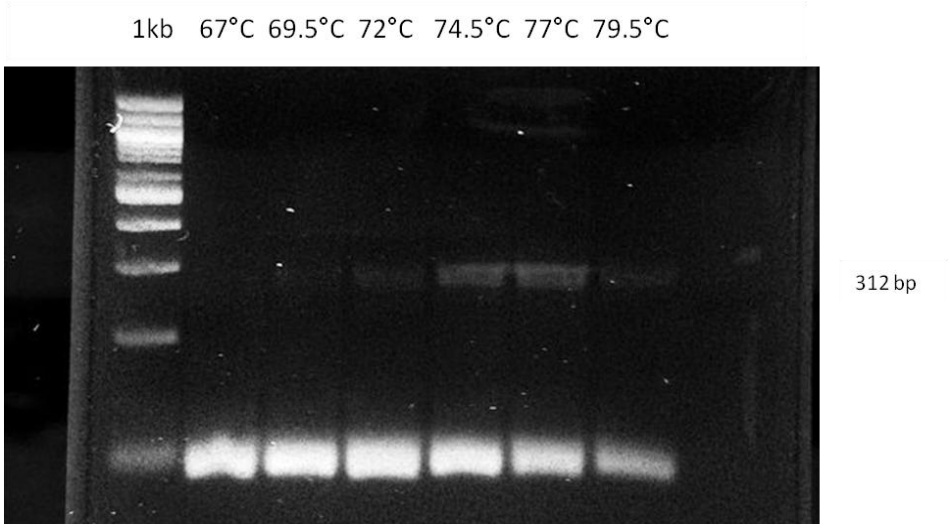


Figure 40 - Rec seq primers amplicon

The Rec seq primers do not contain any mismatched sequences such as restriction sites or FLAG tags. These primers were generated to attempt to prevent any non specific amplification which may have been due to the mismatches. The results once again revealed a band of ~100 bp which is either a non specific band or an Np9 amplicon. However, this time a band of the predicted size of 312bp was also seen which represents the Rec sequence. The different lanes represent identical reactions but with different melting temperatures tested. Highest levels of amplification were observed in the reactions which used 74.5°C and 77°C as melting temperatures. The 312 bp band from the 77°C lane was subsequently gel extracted and used as a template for the Rec-FLAG primers.

The non specific band that I had previously cloned was also present so the 312 bp band was gel extracted and cleaned and used as a template for a PCR reaction with the previous Rec-FLAG primers. To circumvent the amplification problems previously encountered with this primer set it was decided to split the annealing temperature. The reaction was begun with 5 cycles of a lower annealing temperature which would favour the binding of the homologous template. For the remaining 25 cycles this annealing temperature would be increased to match the tm of the homologous primer plus the mismatches. A gradient of melting temperatures was also introduced to try and further optimize the reaction. This reaction provided the predicted correct Rec amplicon of 360 bp **(Figure 41)**. This amplicon was gel extracted and subcloned into pcDNA3.1 (+) in order to be sequenced.

The sequencing results revealed that many mutations were present **(Figure 42)**. Different sequenced colonies had identical sequences and thus the mutations must have been incorporated during the amplification of Rec in the original RT-PCR reaction.

It was decided to re-clone Rec from LNCap prostate cancer cell line cDNA using the Rec sequence specific primers but with a high fidelity polymerase. As before, this reaction would then be used as template for a subsequent PCR reaction using the Rec-FLAG primers. Results from this secondary PCR revealed the same 360 bp band which was again subcloned into the sequencing vector. The sequencing results still revealed that the sequence contained mutations when compared to the ERVK-6 sequence. However, the coding sequence only differed by one mismatch to the last set of sequencing results and this was most likely due to replication slippage as it is situated at the end of a string of adenosine bases **(Figure 43)**.

A BLAST analysis of the first sequence - which lacks the inserted adenosine - revealed that it had 100 % identity to *Homo sapiens* chromosome 5 clone CTD-2022K1, complete sequence **(Figure 44)**.

The sequence was also run through NCBI ORF finder. This programme can identify the presence of ORFs from any given sequence. The Rec-FLAG ORF of 342 bp was not identified by the programme **(Figure 45).**

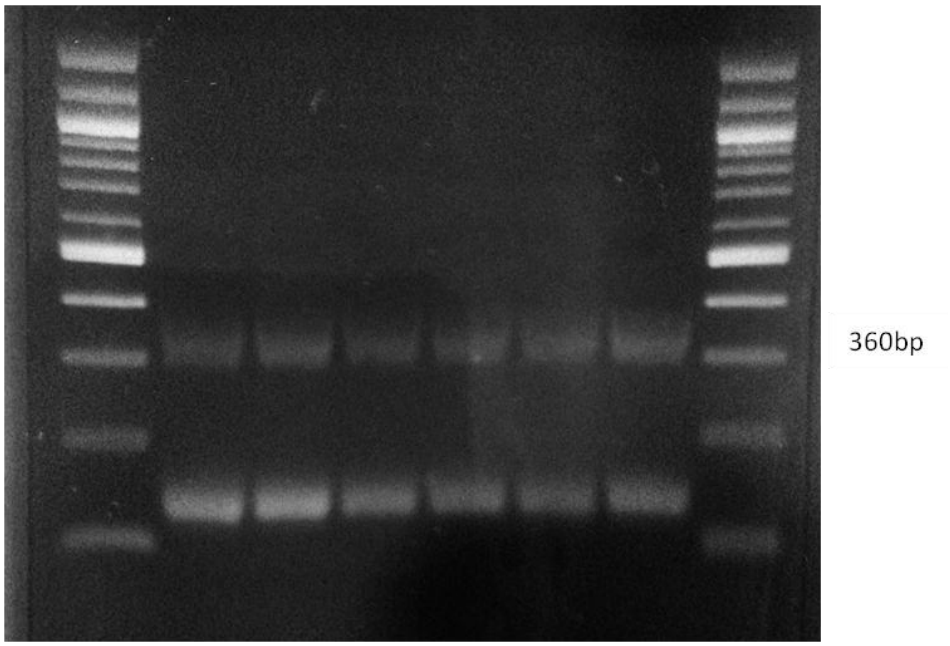


Figure 41 - Rec-FLAG 'split annealing'

The Rec enriched amplicon from the Rec 'seq' PCR was used as template in a further reaction with the Rec-FLAG primers. The annealing temperature was split and a gradient was also used in order to circumvent the original problems which were seen with this primer pair. The lanes on this gel represent identical reactions but with different melting temperatures used. The outer two lanes represent 1 kb ladder. The results revealed the presence of the original non specific band or Np9 amplicon but the correct band size of 360 bp was also seen which represents the Rec-FLAG amplicon.

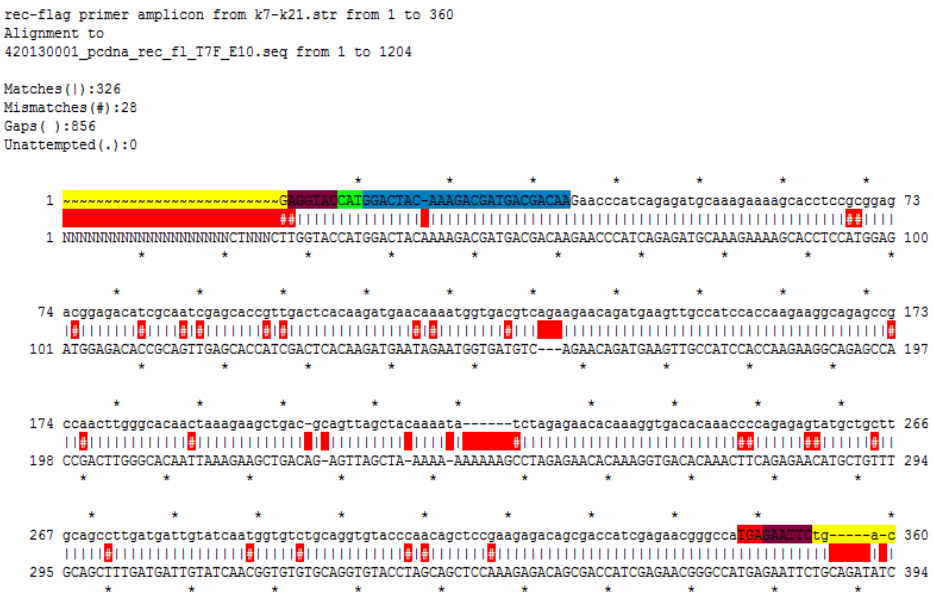


Figure 42 - Rec-FLAG 'split-annealing' sequencing results

The upper sequence represents the expected amplicon while the lower sequence represents the actual sequencing result. The sequencing revealed many mutations including 28 mismatches compared to the ERVK-6 coding sequence.

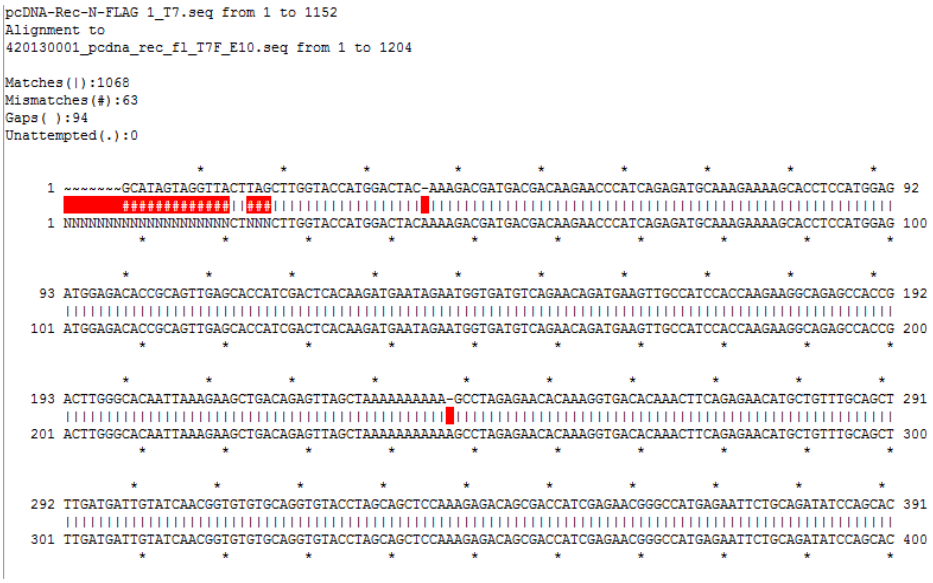


Figure 43 - Alignment of sequencing results

Only two mismatches were observed in the Rec-FLAG sequence. The first of these was observed in the FLAG-tag (top row) while the second was observed within the

coding sequence (third row) and is observed at the end of a string of adenosine bases.

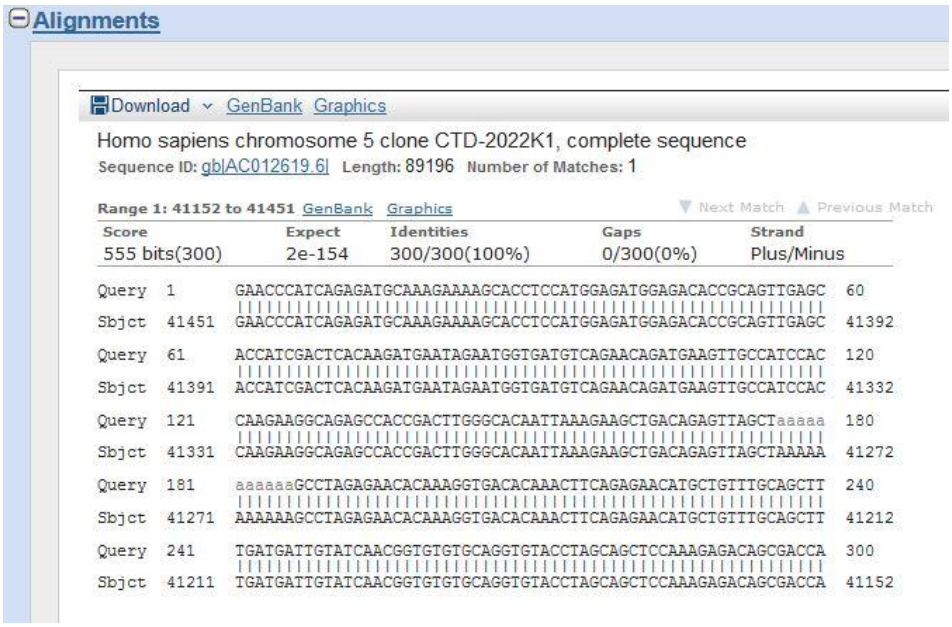


Figure 44 - BLAST analysis of Rec-FLAG amplicon

The amplicon from a Rec-FLAG PCR using a high fidelity polymerase was sequenced and was subsequently subjected to a BLAST analysis. This revealed that it had 100 % identity to *Homo sapiens* chromosome 5 clone CTD-2022K1.

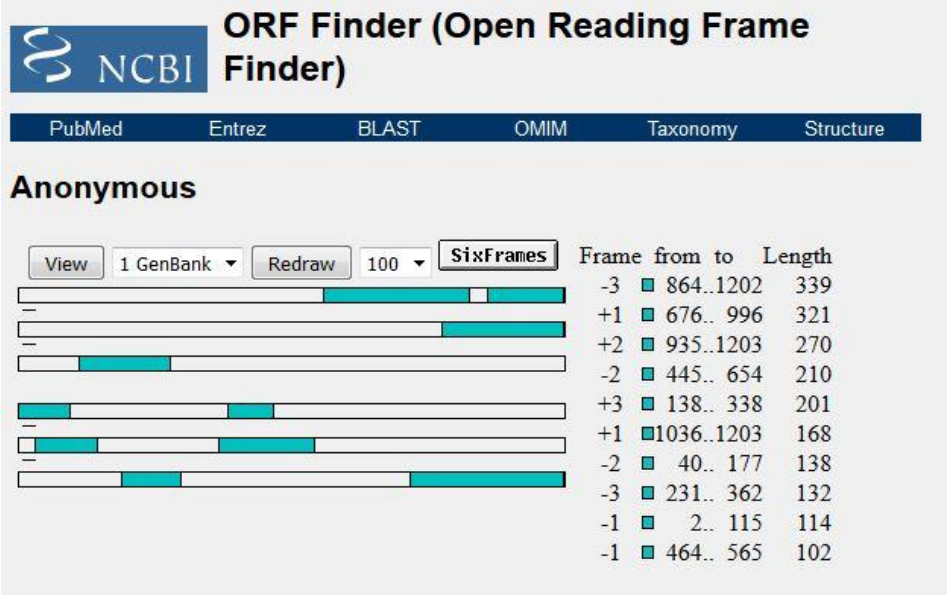


Figure 45 - ORF analysis of Rec-FLAG sequence

The sequence from the Rec-FLAG 'split annealing' experiment was analysed using the NCBI ORF finder programme. ORFs ranging in size from 339 bp to 102 bp were identified. However, the Rec-FLAG ORF of 342 bp was not identified by the programme.

5.1.4 gBlocks gene fragments

At this stage we became aware of a technology called gBlocks gene fragments. This technology allows the synthesis of gene fragments up to 2 kb and is commonly used in conjunction with Gibson assembly which is a novel cloning method which originated in the field of synthetic biology (Gibson et al., 2009). We abandoned our attempts to clone Rec from cDNA and instead ordered a gBlock which was identical to the Rec coding sequence of ERVK-6. The Rec-FLAG primers were used on this template in order to introduce the FLAG-tag and restriction sites for subcloning into the overexpression vector. We chose pcDNA3.1 (+) as our overexpression plasmid. This plasmid can constitutively overexpress proteins in mammalian cells and this expression is driven by a CMV promoter which is upstream of the gene of interest after subcloning **(Figure 46).** This new construct was sent for sequencing and was found to be free from any mutations (**Figure 47)**. ORF analysis confirmed the presence of the 342 bp Rec-FLAG ORF **(Figure 48)**.

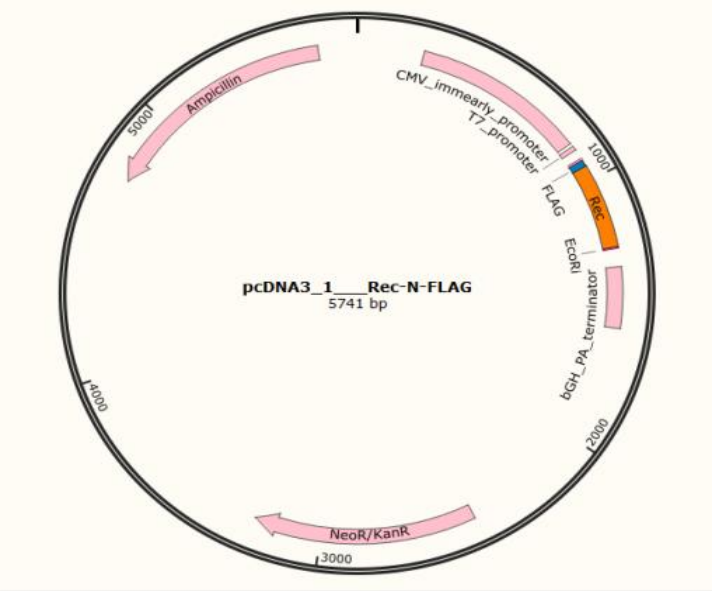


Figure 46 - Plasmid map of pcDNA3.1(+)-Rec-FLAG

The finished construct has the Rec sequence with an N-terminal FLAG tag being driven by a constitutive CMV promoter.

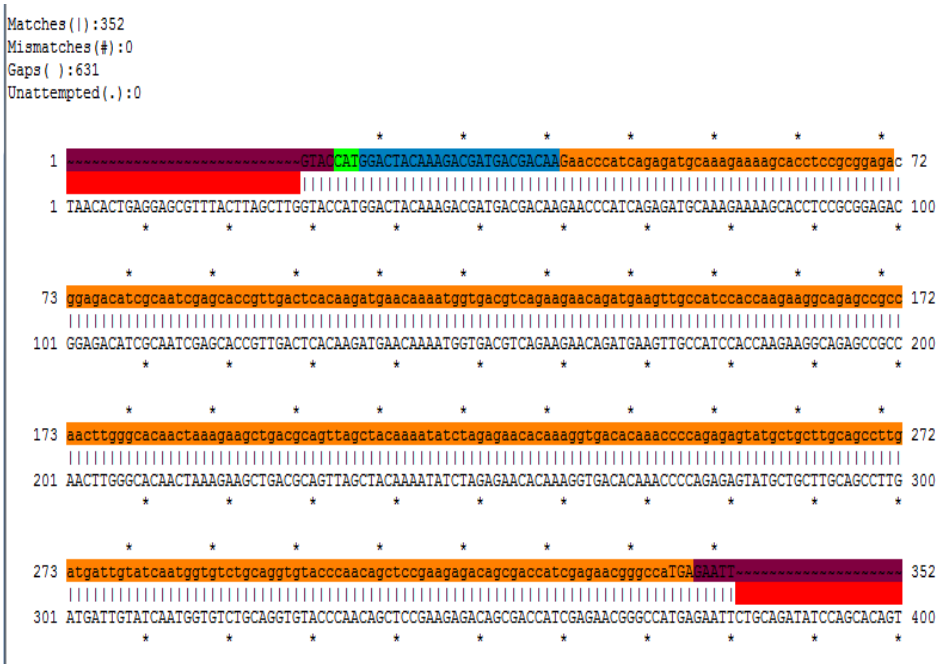


Figure 47 - Rec gBlock sequence

A gblock was ordered which contained a FLAG-tagged Rec sequence and restriction sites to facilitate subcloning into pcDNA 3.1 (+). This completed construct was then sent for sequencing to verify its integrity. The results revealed

that the sequence was free from mutations. The Rec sequence is highlighted in orange with the FLAG tag in blue.

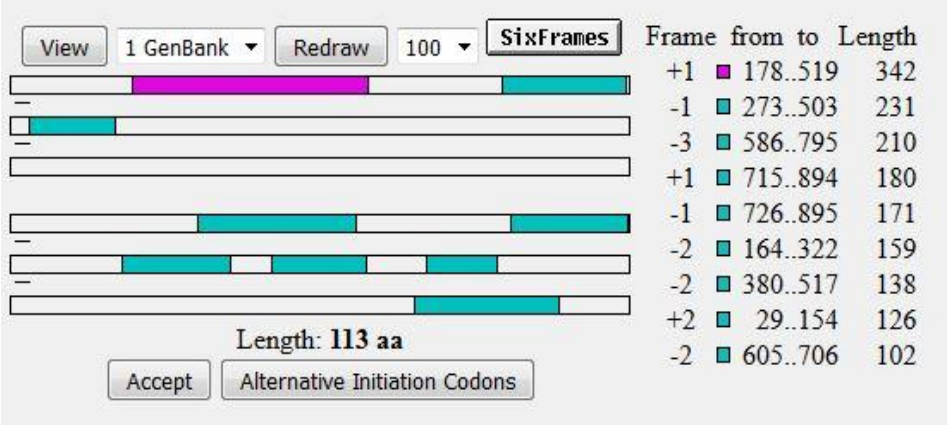


Figure 48 - ORF analysis of Rec-gBlock sequence

The Rec gBlock sequence was also analysed using NCBI ORF finder. The correct Rec-FLAG ORF of 342 bp was identified and is highlighted in pink. This predicts a final FLAG-tagged Rec protein of 113 amino acids.

5.2 Transfection of RWPE-1

The next step was to optimize transfection of RWPE-1 cells for uptake of the pcDNA-Rec plasmid. This proved to be a difficult task due to the fact that RWPE-1 is a nontumourigenic line. Optimization of transfection was conducted by testing different transfection reagents for their ability to transiently transfect a GFP plasmid (**Table 9**).

Various liposomal and non-liposomal reagents were tested. This analysis was completed using flow cytometry. This analysis led to the non-liposomal FuGene reagent being selected for the transfection of RWPE-1. This reagent was found to have transfection efficiencies of 73.8 % in the RWPE-1 cells **(Figure 49).**

Table 9 - Transfection reagents tested

Reagent	Company	Technology	% Efficiency
Lipofectamine 2000	Thermo Fisher Scientific	Liposomal	36.1
TransIT-Prostate	Mirus Bio	Polymer/Lipid mix	32.9
JetPrime	Polyplus Transfection	Polymer	8.7
Fugene HD	Promega	Lipid mix	73.8

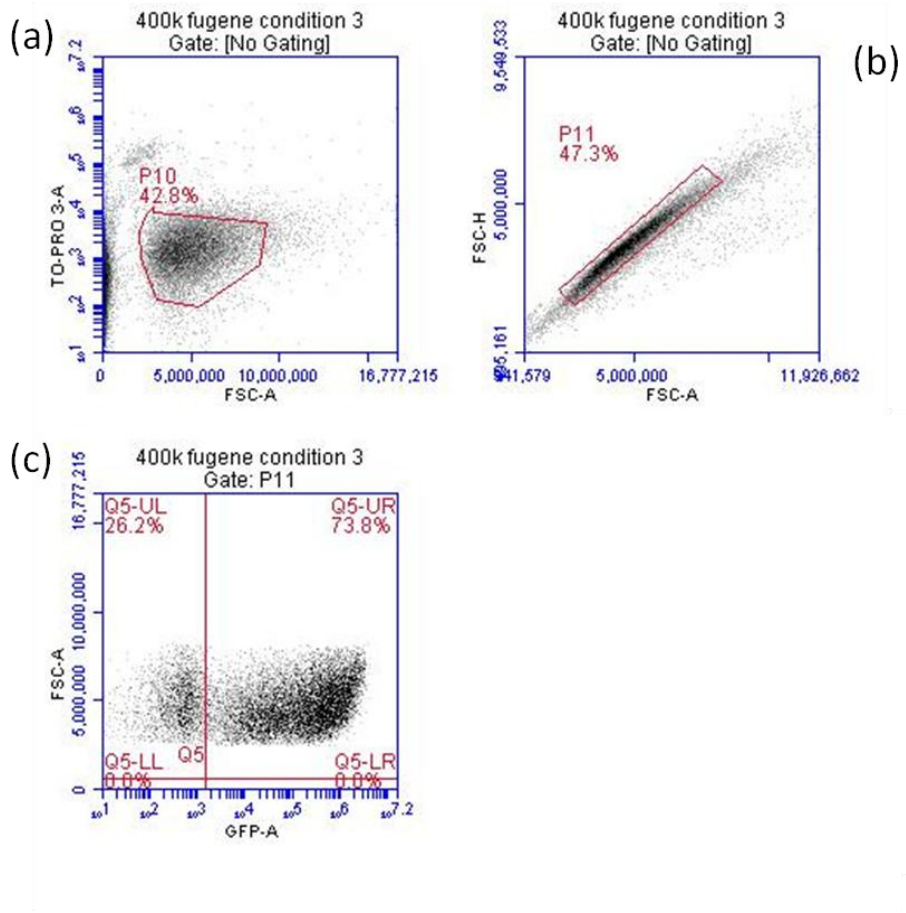


Figure 49 - Optimization of RWPE-1 transfection: FuGENE reagent

Flow cytometry in conjunction with transfection of a GFP plasmid was chosen as the technique to optimize the transfection of RWPE-1 cells. It was found that the non-liposomal FuGENE reagent could transfect RWPE-1 with the highest efficiency of 73.8%. The gating strategy involved firstly gating on FSC and TO-PRO 3 in order to exclude dead cells (a). Secondly, they were gated against FSC-H and FSC-A in order to discriminate between single cells and doublets (b). Finally, they were gated on FSC-A and GFP in order to determine the final transfection efficiency (c). Cells which are GFP positive shift towards the right and are seen in quadrant Q5-UR. These cells have successfully been transfected.

The pcDNA plasmid contains a neomycin cassette which confers resistance to G418. This would be used in order to select for stable transformants. First, a kill curve optimization experiment was conducted in order to find the optimal dose of G418 to be used. This was found to be 200 µg/ml. This

would represent the dose of G418 which would kill 100 % of RWPE-1 cells after one week. The construct was also linearised to attempt to increase the likelihood of stable transformants. Linearization presumably makes it easier for the DNA to integrate into the cell (Tymms, 2000).

After six weeks of selecting with G418 the polyclonal cell line RWPE-1-pcDNA-Rec was eventually created. Later on, a negative control stable line RWPE-1-pcDNA was also created by using the empty pcDNA vector to transfect the cells followed by the same G418 treatment as with the overexpression cell line. Protein lysates were extracted from RWPE-1-pcDNA-Rec and western blots were completed which probed for the presence of FLAG-tagged Rec using an anti-FLAG antibody M1. The results revealed the presence of a faint band at 15 kDa which represents the FLAG-tagged Rec protein (14 kDa Rec + 1 kDa FLAG-tag) **(Figure 50)**.

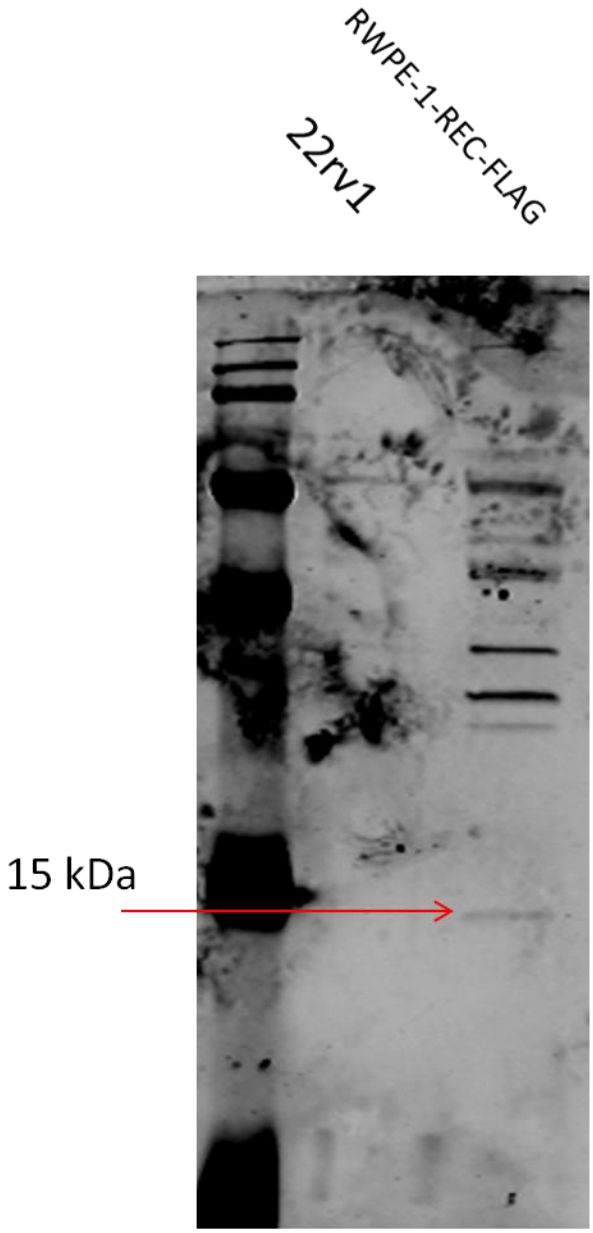


Figure 50 - Rec-FLAG expression in RWPE-1-Rec

Protein lysates from RWPE-1-Rec cells were extracted and probed for Rec-FLAG proteins in a western blot experiment using an anti-FLAG M1 antibody. A simultaneous western as a negative control was completed using the M1 antibody against FLAG in 22Rv1 protein extracts. The 22Rv1 cells revealed no expression of FLAG as expected. The RWPE1-Rec cells revealed a band of 15 kDa which represents FLAG-tagged Rec. The production of the RWPE-pcDNA negative control cell line was completed subsequent to this western blot and hence was not used instead of 22RV1 here. As expected RWPE-pcDNA did not express the 15 kDa band seen here.

This level of expression was lower than expected. Independent westerns were completed using new flasks of transformed cells but the faint band that had been seen earlier could not be replicated. Amounts of up to 100 µg were loaded on the gels but even with these large amounts no expression was seen. Further, transient assays were also completed but no Rec-FLAG protein expression was observed. A positive FLAG-tagged protein lysate of 250 kDa was sourced and westerns completed on this confirmed that the FLAG antibody was working **(Figure 51)**.

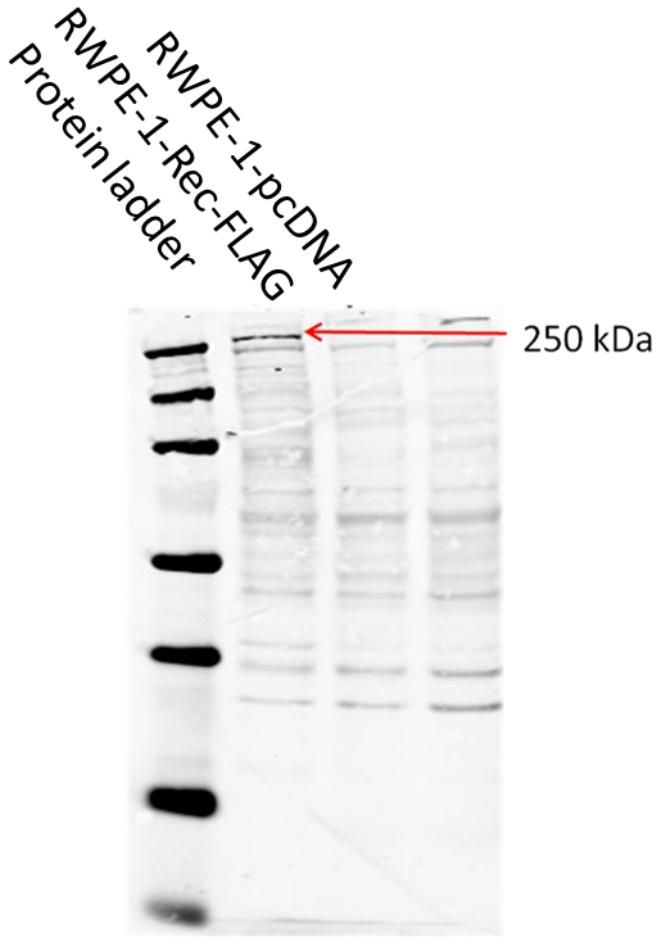


Figure 51 - Positive FLAG control experiment

A positive FLAG control was sourced and used to test that the FLAG antibody was working by western blot. Lane 1 represents protein ladder. Lane 2 represents RWPE-1-pcDNA. Lane 3 represents RWPE-1-Rec-FLAG. The FLAG control was only seen in the lane representing the positive control FLAG lysate. The other fainter bands seen represent non specific staining.

Passaging of polyclonal cell lines sometimes leads to the silencing of the transgene (Freshney, 2010) and so PCRs were completed to confirm that the FLAG-tagged Rec mRNA was still being produced. The results revealed that the RWPE-1-pcDNA cell line was still overexpressing the FLAG-tagged Rec mRNA and that the negative control cell line did not **(Figure 52)**.

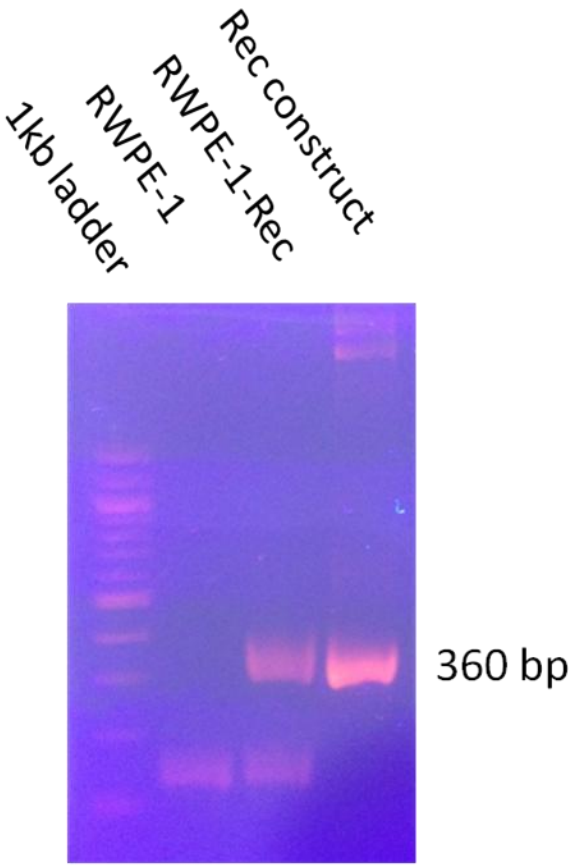


Figure 52 - PCR from RNA from RWPE-1 and RWPE-1-Rec

Lane one represents 100 bp ladder. Lane 2 represents RWPE-1 in which no Rec band is seen, only a non specific band. Lane 3 represents RWPE-1-Rec. Here the non specific band is still seen but the upper band is visible which represents the overexpression of FLAG-tagged Rec. Lane 4 represents a positive control which was a PCR conducted from the Rec construct itself.

It was decided to attempt to create a monoclonal cell line as the low protein expression may have been related to the polyclonality of the cell line (Freshney, 2010). The transfection was repeated and the same G418 selection was conducted except this time cloning rings were used to attempt to select single colonies. These colonies were then grown in independent flasks. Unfortunately, all attempts at growing up individual colonies were futile, probably due to the nature of RWPE-1 cells which are extremely hard to grow from a low density seed.

We investigated whether the low to negative protein expression in transfected RWPE-1 was due to proteosomal degradation. To this end, we treated both the transfected and non transfected RWPE-1 cells with the proteosome inhibitor MG132 (1 μM). We subsequently repeated the western blots we completed previously, probing for expression of FLAG-tagged Rec. The results did not reveal the presence of any bands **(Figure 53)**.

As a positive control experiment for the treatment, we completed an MCL-1 western blot. MCL-1 is a protein which accumulates in the presence of MG132 (Rahmani et al., 2005). This western blot revealed that levels of MCL-1 increased in a time-dependent manner following MG132 treatment **(Figure 54)**.

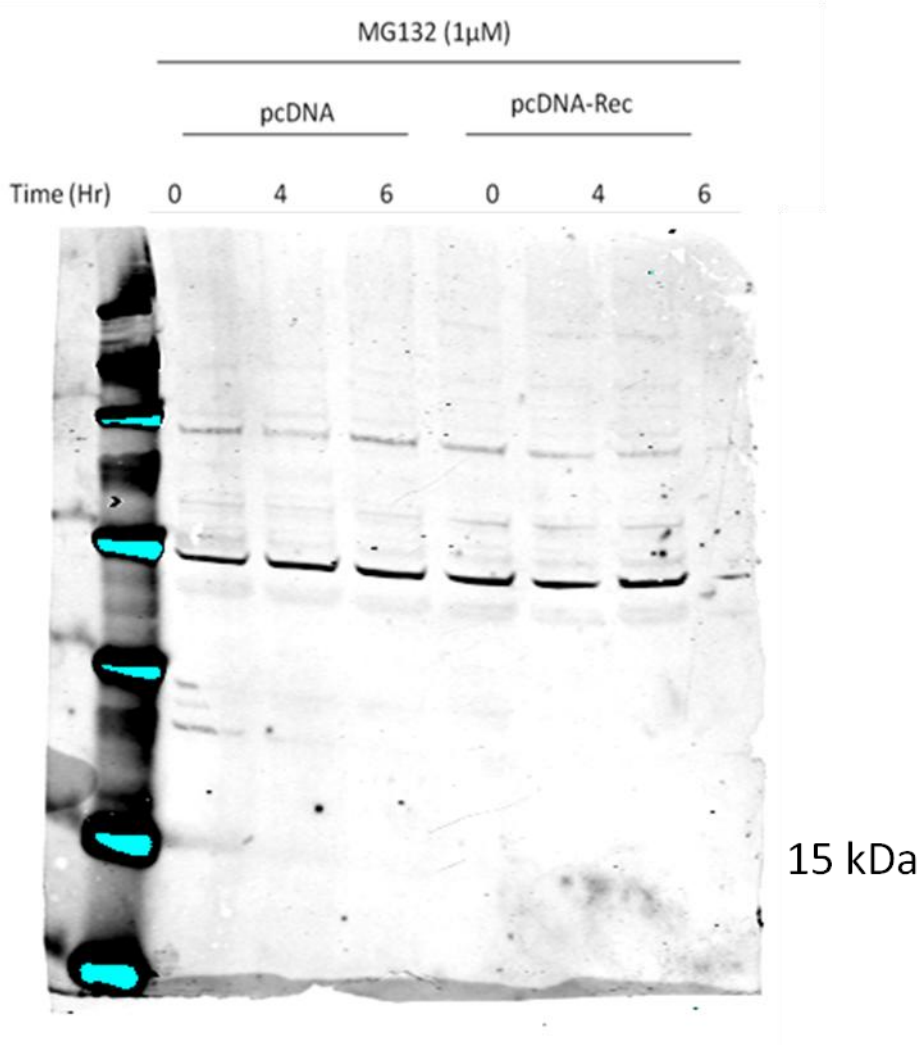


Figure 53 - MG132 treatment of transfected cell lines

RWPE-1-pcDNA and RWPE-1-Rec-FLAG were treated at 0, 4 and 6 hours with 1 µM of the proteosome inhibitor MG132 in an attempt to investigate if the lack of transgene expression was linked to ubiquitin-mediated proteosome degradation. No expression of Rec-FLAG was evident at 15 kDa in any of the treatments in either cell line. Other bands in each lane represent non specific binding.

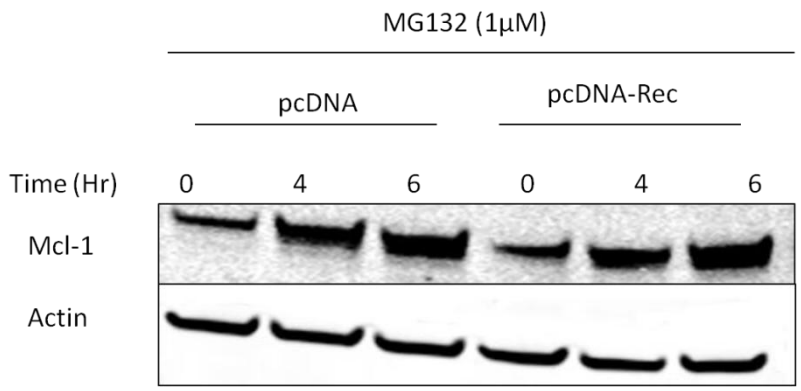


Figure 54 - MG132 treatment

Transfected and non-transfected RWPE-1 cells were subjected to 1μM MG132 treatment which was thought would lead to expression of Rec-FLAG in the transfected cell line if it was being degraded through ubiquitin-mediated proteosomal degradation. The rationale behind this is that MG132 is a proteosome inhibitor. This experiment did not yield any Rec-FLAG expression and thus a positive control experiment was conducted in which the same two treated cell lines were probed for MCL-1. This protein should be upregulated by MG132. The results revealed that the MG132 treatment led to an accumulation of MCL-1 in a time dependent manner in both cell lines. Actin was used as a loading control.

5.3 siRNA mediated knockdown of HML-2

Our attempts to overexpress Rec were intended to investigate the functional effects of Rec alone in prostate cancer. However, HML-2 expresses many other proteins such as env, gag and Np9 which may be playing a functional role in cancer. Their effects may be exerted in unison with one another. We were interested in investigating how these proteins may work together to initiate or progress the carcinogenic process. To this end, we decided to attempt to knockdown HML-2 in a prostate cancer cell line and to subsequently investigate the functional effects of doing so.

We ordered a siRNA which was identical to a study by Oricchio et al. *(Oricchio* et al*., 2007).* This siRNA targets a conserved region within the gag gene and should lead to the knock down of the majority of HML-2 loci. This was confirmed by BLAST analysis of the sequence. We optimised this siRNA for transfection into PC3 and DU145 cells. We completed a time

course experiment in PC3 cells. PC3 cells were transfected with siRNA and RNA was extracted every 24 hours for four days. The qPCR results revealed that HML-2 gag levels dropped after 24 hours and stayed at that level until 72 hours when they returned to normal levels **(Figure 55)**. However, a scrambled RNA control was also transfected and it caused gag levels to drop to levels similar to that of the targeted siRNA after 24 hours. We also attempted the same knockdown in DU145 cells but no differences in expression were seen between untreated cells, cells treated with scrambled RNA and cells treated with 10 nM or 50 nM of siRNA after 24 hours **(Figure 56)**. Western blots were also completed to look for knockdown of gag and env in PC3 and DU145 cells after siRNA transfection. However, non specific bands were observed in each western so expression levels of gag and env post transfection could not be deduced.

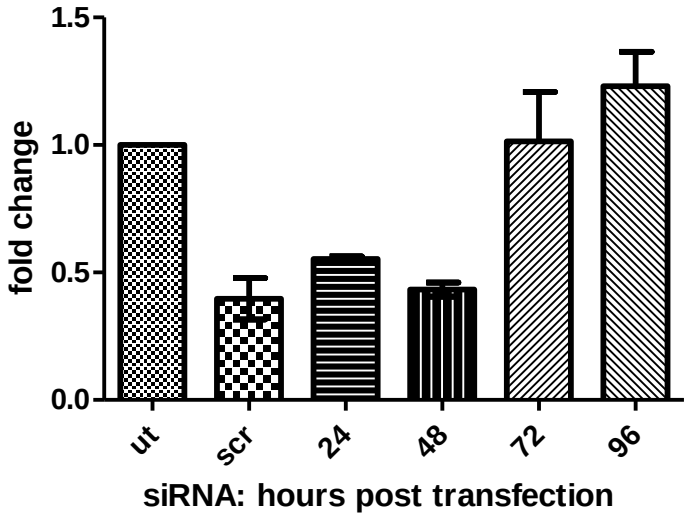


Figure 55 - siRNA transfection of PC3 cells

PC3 cells were transfected with 10 nM of siRNA targeting HML-2 gag. RNA was extracted from transfected cells at 24 hour time points. Valid conclusions could not be drawn from this experiment due to the effect which the scrambled RNA (scr) had on the cells. Results are expressed as fold change (RQ) in comparison to a non transfected control (ut). Error bars represent standard deviation from three technical replicates.

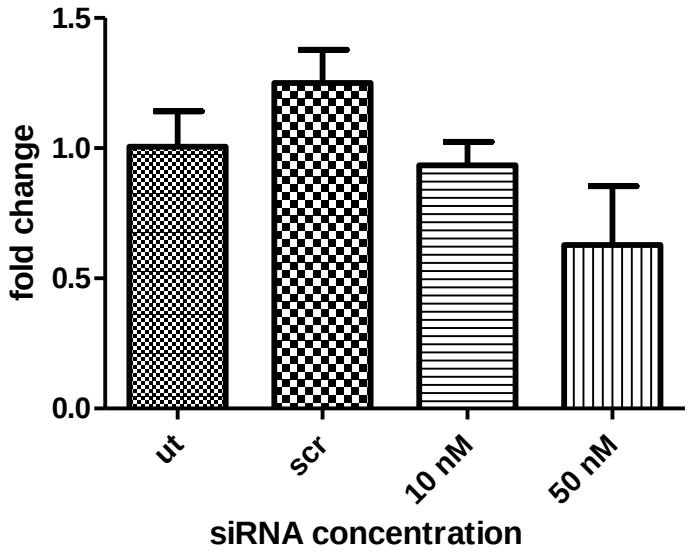


Figure 56 - siRNA knockdown of HML-2 in DU145 cells

siRNA knockdown of HML-2 in DU145 cells was attempted. RNA was extracted after 24 hours. 10nM and 50nM siRNA against HML-2 gag did not have any effect on gag levels compared to the untreated cells (ut) and to the scrambled RNA control (scr). Results are expressed as fold change (RQ) in comparison to a non transfected control (ut). Error bars represent standard deviation from three technical replicates

5.4 Discussion

Previous research which sought to find a link between HML-2 and cancer has concentrated on the HML-2 accessory proteins Rec and Np9 (Downey et al., 2015). These putative oncoproteins have been subject to a number of studies which have investigated their carcinogenic effects both *in vitro* and *in vivo*. Many interesting protein-protein interactions between previously known oncogenes and Rec and Np9 have been uncovered. Tentative links have been made with a number of cancers, especially germ cell tumours for which Rec seems like it may be playing a causative role (Herbst et al., 1998). In the context of prostate cancer, Rec has been shown to lead to the derepression of the AR through interactions with TZFP and hSGT (Hanke et al., 2013; Kaufmann et al., 2010). These studies spurred our interest into this

area and we set out to define the possible functional links between Rec and prostate cancer.

Our main tool for elucidating the function of Rec in prostate cancer was to be a Rec overexpression construct which we would transfect into immortalised non tumourigenic RWPE-1 prostate cells. The first stage of this process was to clone Rec from prostate cancer cell line cDNA and to subclone this into an overexpression vector. This process was complicated by the repetitive nature of HML-2 in the genome of which there are 91 proviral copies present (Subramanian et al., 2011). It was also complicated by the complex splicing pattern of HML-2 which can lead to either Rec or NP9 transcripts depending on what type of HML-2 is present (Bannert and Kurth, 2006).

Previously, RT-PCR had successfully been used to define Rec and Np9 expression in prostate cancer cell lines. This template is enriched for Rec and Np9 transcripts. The use of the Rec-FLAG primers on this template resulted in an incorrect smaller band size which represented either Np9 or a non specific band. Primer dimers were ruled out after the running of a no template control. The preferential amplification of this smaller band over the larger Rec band of interest is probably due to the nature of PCR itself which has been shown to amplify shorter fragments more efficiently than longer ones (Shagin et al., 1999). This greater efficiency effectively shores up all the available reagents during the early cycles and the larger amplicons cannot extend properly.

The presence of mismatches in our primers for restriction sites and FLAG tag was also a possible reason for the amplification of non specific bands. *In silico* analysis of the Rec coding sequence did not predict any binding sites for the introduced sequences. However, the 91 HML-2 loci harbour nucleotide differences from one locus to another which they have accumulated over millions of years of evolution (Flockerzi et al., 2008). This means that certain Rec coding sequences - or indeed sequences that have not yet been discovered – may have been binding to the introduced

sequences. Another possibility was that the original k7/k21 primer set was interacting with the Rec-FLAG primers to amplify spurious products. However, this was ruled out firstly by PCR cleaning the K7/K21 template and secondly by running a series of PCRs which mixed the primers. The latter experiments could not repeat the amplification of the ~150bp band and thus it was concluded that the band was a non specific band from the Rec-FLAG primers themselves or an Np9 band.

The sequencing of the 150 bp band revealed the presence of a 139 bp amplicon which presumably represents the ~150 bp seen on the gels. Excluding the FLAG-tag, this sequence has a total length of 101 bp but does not contain an ORF. According to Armbruester et al., Np9 has a protein sequence of 74 amino acids which would correspond to 222 bp (Armbruester et al., 2002). This sequence is therefore unlikely to represent Np9. Alignment of this sequence with the Rec coding sequence revealed that the sequence aligned with Rec at both the 5' and 3' ends but had a gap of 212 bp internally. No Np9 mRNA sequence was available from GenBank for alignment analysis but a BLAST analysis did reveal that the sequence had 93 % identity to Homo sapiens, clone IMAGE:5392784, mRNA GenBank: BC028044.1. This cDNA clone has no known protein product but does contain 2 ORFs.

To overcome the problem of non specific amplification, we designed new primers which targeted only the Rec sequence itself. PCRs with these primers led to the amplification of the band of interest but the smaller band was still present. It was clear that the mismatches had been the cause of the preferential amplification of the smaller band. The band of interest was gel extracted and used as template for a PCR with the Rec-FLAG primers. This time the annealing temperatures were split, with different melting temperatures used at different stages of the reaction. This method was successful in amplifying the band of interest in preference to the smaller spurious band. The correct amplicon was subcloned into a vector and sent for sequencing.

At this stage the sequence contained many mutations when aligned to the ERVK-6 prototype Rec sequence. Mutations would be expected at this stage as the original primers would be redundant amongst the HML-2 family, with each family member harbouring private nucleotide differences (Flockerzi et al., 2008). However, the large amount of mutations led us to suspect that they had been integrated at the original reverse transcriptase step of the k7/k21 reaction which utilised a low fidelity polymerase. The protocol was repeated with a high fidelity polymerase, this time cloning from LNCap cells. The final sequence once again had sequence differences to ERVK-6 Rec. Analysis confirmed that only one mismatch in the Rec coding sequence was apparent between LNCap and DU145 cells. This was an inserted adenosine base which was situated at the end of a string of adenosine base pairs. This insertion is most likely due to a molecular phenomenon known as replication slippage which is caused when a replicating DNA polymerase enzyme meets repetitive bases (Viguera et al., 2001).

The finding that two independent cloning procedures using cDNA originating from two different cell lines resulted in the same sequence being cloned is interesting and may be due to two reasons. Firstly, the exact primers and reaction conditions used may preferentially clone this sequence over others. Secondly, it may be possible that this sequence is expressed more abundantly then others in these particular cell lines. Goering et al found that 7 different HML-2 loci were expressed in PC3 cells, with 56 % of all transcripts originating from ERVK-5 (Goering et al., 2015).

The BLAST analysis of the sequence without the adenosine insertion revealed that the sequence had 100% identity to Homo sapiens chromosome 5 clone CTD-2022K1, complete sequence. This clone is located at 5q12-13. The only annotated HML-2 provirus in this region is that of ERVK-9 at 5q12.3 (Marchi et al., 2014). This provirus is not found in the human reference genome and Lee et al. found that it was present in only 12 % of human genomes they tested (Lee et al., 2012). Goering et al did not find that ERVK-9 was expressed in any of the tumours or cell lines tested including

PC3 (Goering et al., 2015). ORF analysis found that no intact Rec ORFs were present. Translation of a protein depends on the presence of an ORF with a start codon and a stop codon separated by triplicate codons (Crick, 1968). A single non synonymous mutation can disrupt an ORF by the creation of a premature stop codon. The result of this is a truncated protein. Other mutations such as frameshift mutations can result in spurious amino acid chains which lack any structure and will not form proteins (Watson, 2008). More exhaustive sequencing of LNCap and PC3 transcripts would be needed to gain conclusive evidence that the cloned Rec sequence is indeed from this locus. However, the lack of an ORF means that this provirus was not of interest for a functional study.

The correct construct overexpressing a Rec sequence from ERVK-6 was eventually created using gBLock technology. This construct was successfully transfected into RWPE-1 cells but we found that the expression of Rec protein in this cell line was low to negative.

We confirmed using PCR that high levels of Rec-FLAG mRNA were being produced **(Figure 52)** and so the problem was due to the protein expression. Other studies have also encountered difficulty in transiently overexpressing Rec including Galli et al. (Galli et al., 2005). One way of increasing the expression of Rec would be to codon optimize the sequence. The levels of the different tRNAs differ amongst various biological cell lines (Plotkin and Kudla, 2011). Codon optimization introduces synonymous mutations which match each codon with the most abundant tRNA in any given cell line. This increases the efficiency of translation and results in increased expression of the protein of interest. This method was utilised by Hanke et al. who created a codon optimised version of the HERV-K113 Rec sequence in order to overexpress it in HEK293T cells (Hanke et al., 2013). The main problem with codon optimizing a sequence in the case of a functional research study is that you are left with a different mRNA sequence to the wild type sequence which may affect certain viral restriction mechanisms, splicing or mRNA binding (Gustafsson et al., 2004).

It is possible that some prostate cancer clones pick up mutations which mirror those of a codon optimised Rec sequence, resulting in increased translation of Rec. These cells may have dysregulated androgen signalling due to the effects of high levels of Rec and this may eventually lead to androgen independent growth. Massively-parallel sequencing of Rec loci in prostate cancer clinical samples may be one way to investigate such a phenomenon.

We also confirmed that Rec was not being degraded through ubiquitin-mediated proteosomal degradation by treating the stable cell lines with the proteosome inhibitor MG132 and observing no increase in Rec expression.

It is also possible that viral restriction mechanisms are active in normal prostate cells in order to keep them silent. These mechanisms may break down in the context of a cancer cell which allows for increased expression of Rec. This could be investigated by looking at the activity of proteins such as TRIM and APOBEC3G which are known viral restriction enzymes (Huthoff and Towers, 2008). Lee et al have found that APOBEC3G can hypermutate DNA from a HERV virus that they resurrected (Lee et al., 2008) and Esnault et al. have shown that APOBEC3G can inhibit the retrotransposition of ERVs in mice (Esnault et al., 2005).

If Rec could be successfully overexpressed in RWPE-1 it would be interesting to complete a gene expression microarray study where the profile of the non transfected RWPE-1 cells is compared to that of the Rec overexpressing RWPE-1 cells. No such study has looked at the effects that the Rec protein induces across a genome wide level. It is possible that many novel proteins exist which Rec interacts with. Furthermore, high levels of Rec mRNA may induce viral restriction programmes which may alter immune cell signalling. The data from such a microarray could be subjected to gene set enrichment analysis which could elucidate the molecular pathways which Rec expression is affecting. A monoclonal stable cell line would be more suited than a polyclonal line in the context of a microarray

experiment because the expression would be higher and also it would be static which would be beneficial when it comes to reproducing experiments.

Transient assays could also be completed which would investigate the transforming potential of Rec expression in normal RWPE-1 cells. These assays could include cellular proliferation, invasion and wound healing assays, as well as expression analysis of stemness genes. The activity of the AR could also be monitored as a validation of previous studies which have shown that Rec expression can lead to derepression of the AR through interactions with TZFP and hSGT (Hanke et al., 2013; Kaufmann et al., 2010).

Another potential pitfall of trying to overexpress retroviral proteins in RWPE-1 is that RWPE-1 is a cell line which was created via immortalization with human papillomavirus (HPV) 18 (Bello et al., 1997). As previously mentioned Agoni et al. found a correlation between HML-2 positivity and HPV infection in cervical cancer cell lines (Agoni et al., 2013b), so it is possible that inserted HPV oncogenes in RWPE-1 can interact with HML-2 sequences such as Rec and control their expression.

We also attempted to knockdown HML-2 in prostate cancer cell lines DU145 and PC3 using siRNA. The siRNA used was identical to that from a publication by Oricchio et al. (2007) who used the sequence to knockdown HML-2 in melanoma cell lines (Oricchio et al., 2007). They found that knockdown of HML-2 in these cell lines reduced the tumourigenic potential of the cells when injected into nude mice. The same group also found that HML-2 knockdown prevented the transition from adherent to non adherent growth (Serafino et al., 2009). However, we found no changes in expression of full length HML-2 mRNA after treating prostate cancer cell lines with the same siRNA. The reason for this may be due to the differential expression of specific HML-2 loci in either cell line. The HML-2 transcriptome in both melanoma and prostate cancer has been investigated. It was found that the transcription of HML-2 in melanoma was complex with no obvious outlier of expression (Schmitt et al., 2013a). In prostate cancer the majority of

expression of HML-2 came from a few loci including the locus 22q11.23 (Goering et al., 2015). Redesigning a consensus siRNA against the majority loci in prostate cancer may help to knockdown HML-2 in this setting. Also, a large pool of siRNA could be utilised instead of a single sequence which may also improve the knockdown efficiency. Another possibility is to use the CRISPR-CAS9 system to 'delete' HML-2 from the genomes of these prostate cancer cell lines (Hsu et al., 2014). Either a consensus target sequence could be used which would target the majority of the HML-2 loci or a single locus could be targeted allowing its specific function to be investigated. In the melanoma study, Orrichio et al. had used a stable knockdown system while we only attempted a transient knockdown. It is possible that in prostate cancer cell lines a stable knockdown would be more effective than a transient knockdown.

The production of a Rec monoclonal antibody would be also beneficial for the future study of Rec in prostate cancer. This would alleviate the need to epitope tag the Rec protein, allowing the function of Rec to be investigated in a more physiologically relevant context. It would also allow Rec to be tracked more easily in immunofluorescence experiments allowing the subcellular localization of Rec in prostate cancer to be elucidated.

Np9 should not be overlooked in the context of prostate cancer. Similar to Rec, Np9 can relieve the transcriptional repression of the c-myc oncogene via its interaction with PLZF (Denne et al., 2007). Furthermore, Np9 can interact with LNX which is a key molecule in the LNX/NUMB/NOTCH pathway – a pathway which has previously been implicated in cancer (Armbruester et al., 2004). A global gene expression microarray together with pathway analysis may help further in elucidating the key molecular pathways which Np9 influences in human health and disease.

A potential limitation of this study is that the Rec sequence is chosen from one particular provirus, in this case that of ERVK-6. Agoni et al. found that ERVK-6 mRNA was expressed in DU145 cells (Agoni et al., 2013a). Goering et al. did not find that ERVK-6 was expressed in any of the prostate

cancer primary samples or cell lines that they tested (Goering et al., 2015). They did, however, identify that ERVK-4 was expressed and Mayer et al. confirmed that this provirus has an intact ORF for Rec (Mayer et al., 2004). The ERVK-4 Rec sequence has five amino acid differences to ERVK-6, two of which are contained within the nuclear localization signal (NLS) (Mayer et al., 2004). These amino acid differences may alter the function of each protein and so in the future it will be important to investigate these differences at the molecular level.

In summary, this study has successfully produced a Rec overexpressing construct. This construct will be of benefit to future studies which seek to explore the role of Rec in human health and disease. This study has also identified the expression of a Rec sequence in LNCap and PC3 cells, from a provirus originating on chromosome 5 which most likely represents the polymorphic ERVK-9.

Chapter 6

Overall discussion and future directions

6 Overall discussion and future directions

6.1 Overall discussion

From a medical viewpoint the onset of the 21st century has been notable due to the significant increase in chronic diseases (Eckel et al., 2010). This increase is in a large part due to the spread of westernised diets and lifestyles across the world (Kuh and Ben-Shlomo, 2004). With the increase in disease burden comes the need for novel medicines. HERVs represent 8% of the human genome yet we are only beginning to understand their impact on human health and disease (Downey et al., 2015). Basic research into HERVs including HML-2 will be the key to unlocking their eventual clinical utility. As the price of genome sequencing continues to tumble, clinical whole genome sequencing of patients will become routine (Hamburg and Collins, 2010). This deeper understanding of the human genome in the context of disease will help to uncover new HERV polymorphisms and novel loci. Landmark functional studies of the whole genome such as ENCODE (Consortium, 2012) and The Cancer Genome Atlas (TCGA) (Cancer Genome Atlas Research et al., 2013) are critical generators of new hypotheses into human disease. Unravelling the function of HERVs through these studies will be a catalyst for further studies into HERVs association with disease.

Progress in biomedical research is hugely influenced by the invention of new technologies and HERV research will be no exception. The ability to analyse HERV expression in the context of a single cell will benefit the search for an infectious particle, while CRISPR-CAS9 can be harnessed to investigate the function of individual loci.

The future clinical utility of HERVs can be broadly grouped into three different areas: vaccines, biomarkers and new therapeutic targets.

The last few years has seen considerable progress into HERV vaccines. Much of this progress has been made by the Wang-Johanning group. They

have successfully genetically engineered HML-2 chimeric antigen receptor T cells which can selectively kill melanoma (Krishnamurthy et al., 2015). Additionally, they have produced HML-2 T cells from dendritic cells and demonstrated their ability to kill breast (Zhou et al., 2015) and ovarian cancer cells (Rycaj et al., 2015). Theoretically, it should be possible to create these T cells to target any cancer which expresses HML-2 tumour associated antigens. The next step for these HERV vaccines will be clinical trials to test their safety and efficacy in humans.

The future of HML-2 as a biomarker in human disease also looks promising. More and more studies are being published which define HML-2 expression in greater detail in a wide array of diseases (Downey et al., 2015). Early diagnosis will be the key to lowering global cancer burden and new early detection biomarkers will be of immense benefit in reaching this goal (Etzioni et al., 2003). This thesis has demonstrated the utility of HML-2 expression in the diagnosis of prostate cancer and shown that it also may have the ability to improve PSA scores. Medicine has moved away from single use biomarkers and moved towards an age of multiple biomarker panels (Paik et al., 2004). These panels are more sensitive and specific than any one biomarker on its own. The fact that HERVs are expressed in so many diseases means that they stand ready to be introduced into many of these new panels.

Discovering new targets is one of the most important aspects of the drug discovery process (Overington et al., 2006). In this regard, HERVs represent a valuable untapped resource whose potential is waiting to be unlocked. Wang-Johanning et al. have demonstrated the immunotherapeutic potential of an anti-HML-2 env antibody in breast cancer (Wang-Johanning et al., 2012). They found that treatment with this antibody led to significant decreases in cellular proliferation and apoptosis both *in vitro* and *in vivo*. HERVs have also recently been implicated in the response of patients to immune checkpoint blockade with the finding from Chiappinelli et al. that DNA methyltransferase inhibitors can lead to the expression of dsRNA HERV products which upregulate immune signalling (Chiappinelli et al.,

2015). Another study which has reached Phase I clinical trials seeks to use antiretroviral therapy to treat patients with amyotrophic lateral sclerosis (ALS) who express high levels of HML-2 mRNA (Alfahad and Nath, 2013). More HERV drug targets are set to be discovered as we gain a greater understanding of HERV protein expression and function.

The search for an infectious HML-2 is one of the biggest questions still left unanswered in HERV research (Downey et al., 2015). There is evidence to suggest that some HML-2 proviruses have been active as recently as 200,000 years ago (Bannert and Kurth, 2006). It is possible that a rare pool of replicating HERVs is present in some human populations. Such pools of retroviruses are unlikely to reach high allelic frequencies due to the deleterious effects of insertional mutagenesis. Once again, routine clinical whole genome sequencing may uncover links between diseases and active infectious HERVs. Presumed ancestors of HERVs have been resurrected in the lab on two independent occasions (Dewannieux et al., 2006; Lee and Bieniasz, 2007). These studies are invaluable in gaining an insight into the true function of HERVs. However, the dangers of resurrecting infectious retroviruses should not be undermined and considerable ethical challenges remain.

Whether HML-2 expression is a cause or consequence of cancer remains to be seen (Downey et al., 2015). Although there is much evidence to suggest a link between the two, conclusive proof has yet to be found. In prostate cancer, the locus 22q11.23 was the focus of intensive research in this regard. Goering et al. found that this locus was responsible for the majority of HML-2 transcription in prostate cancer (Goering et al., 2015). Reis et al. had discovered that gag protein expression originating from this locus correlated with an immune response to prostate cancer (Reis et al., 2013). Finally, a fragment of this locus is responsible for gene fusions seen in prostate cancer. Goering et al. used siRNA to knockdown the expression of this locus in prostate cancer cells but did not find that this led to any changes in cellular proliferation or apoptosis and concluded that expression from this locus is most likely an epiphenomenon (Goering et al., 2015).

The beneficial aspect of HERVs to human physiology should also not be overlooked. Recent studies have highlighted the importance of HERVs to embryonic stem cell pluripotency (Lu et al., 2014; Wang et al., 2014) and it is possible that many other beneficial traits remain undiscovered.

Overall, this thesis has contributed to a greater understanding of HERVs in prostate cancer, both in terms of basic and clinically relevant research. Whether or not HML-2 plays a functional role in prostate cancer remains to be seen, but its potential as a clinical biomarker is obvious and warrants further development.

6.2 Future directions

Many questions remain unanswered in both clinical and basic HML-2 research in prostate cancer. Some of these are listed below together with directions for future work.

Basic research

Questions

- What induces HML-2 activation in hormone negative cell lines?
- Is HML-2 activation related to de-differentiation and stemness?
- Are there specific loci which contribute to prostate cancer progression or is it the collective HML-2 proteome which plays a role?
- Are anti-viral pathways activated by HML-2 expression?
- What exact cell PBMC cell types are responsible for the HML-2 signal?
- Are HML-2 virions expressed in prostate cancer?

Future work

- Define the role of methylation/demethylation in HML-2 silencing/activation

Clinical research

Questions

- Do hormone analogue drugs such as antiandrogens have an effect on HML-2 expression?
- What role does inflammation and infection play in HML-2 activation?
- Does HML-2 expression affect response to immune checkpoint blockade therapies?

- Is HML-2 activation related to the aging process?
- Is there a link between smoking and HML-2 activation?
- Is an infectious HML-2 particle present amongst humans?
- Can HML-2 represent a clinically useful vaccine?
- Do any HML-2 proteins represent potential new drug targets for the disease?

Future work

- Validate the HML-2 biomarker in a second independent cohort
- Gather data on HML-2 expression in relation to PSA scores
- Investigate the potential use of other HML-2 transcripts or proteins as biomarkers in the disease

Functional characterisation

Questions

- Does Rec expression contribute to androgen independent prostate cell growth through effects on androgen signalling?

Future work

- Attempt to overexpress Rec in prostate cancer cell line instead of non tumourigenic line
- Attempt to overexpress a provirus other than ERVK-6
- Attempt codon optimization of Rec coding sequence
- Attempt to overexpress Np9
- Attempt stable knockdown of HML-2 in prostate cancer
- Attempt deletion of HML-2 using CRISPR-CAS9
- Develop monoclonal antibodies against Rec and Np9
- Investigate novel interactors of Rec and Np9 in prostate cancer cells

Appendices

Appendix A - PhD output

Publications

First authored research papers

Wallace TA*, **Downey RF***, Seufert CJ, Schetter A, Dorsey TH, Johnson CA, Goldman R, Loffredo CA, Yan P, Sullivan FJ, Giles FJ, Wang-Johanning F, Ambs S, Glynn SA. Elevated HERV-K mRNA expression in PBMC is associated with a prostate cancer diagnosis particularly in older men and smokers. Carcinogenesis 2014;35:2074 – 2083. **(*joint first authors)**

First authored reviews

Downey RF, Sullivan FJ, Wang-Johanning F, Ambs S, Giles FJ, Glynn SA. Human endogenous retrovirus K and cancer: Innocent bystander or tumorigenic accomplice? Int J Cancer 2015;137:1249 – 1257.

Individual grants awarded

Government of Ireland Postgraduate Scholarship Award 2013:
http://www.research.ie/awards/postgraduate-scholarship-scheme-2013
Ronan Downey: Human Endogenous Retrovirus K Activation: A New Therapeutic Target for Prostate Cancer, Structured Doctoral Degree (48 months).
Project ID: GOIPG/2013/1061

Poster presentation prize awards

Biotechnology In Action - 3rd September 2012 – DCU - Human endogenous retrovirus activation in prostate cancer: association with disease progression *(1st prize – Best Poster Presentation)*

Irish Association of Cancer Research (IACR) 2014 27th-28th February – Galway Bay Hotel, Salthill, Galway - Human endogenous retrovirus k expression predicts prostate cancer *(1st prize – Best Poster Presentation)*

Other poster presentations

AACR 2015 – Philadelphia, USA - Human endogenous retrovirus k expression as a possible adjunct to PSA in the diagnosis of prostate cancer

RAMI 2012 - 14th June 2012 – NUI Galway - Human endogenous retrovirus activation in prostate cancer: association with disease progression

Roche Research Award in Life Sciences and Bioengineering – NUI Galway - Monday 19th November 2012 - Human endogenous retroviruses in prostate cancer

IACR 2013 – Santry, Dublin - Human endogenous retroviruses: a new etiology for prostate cancer?

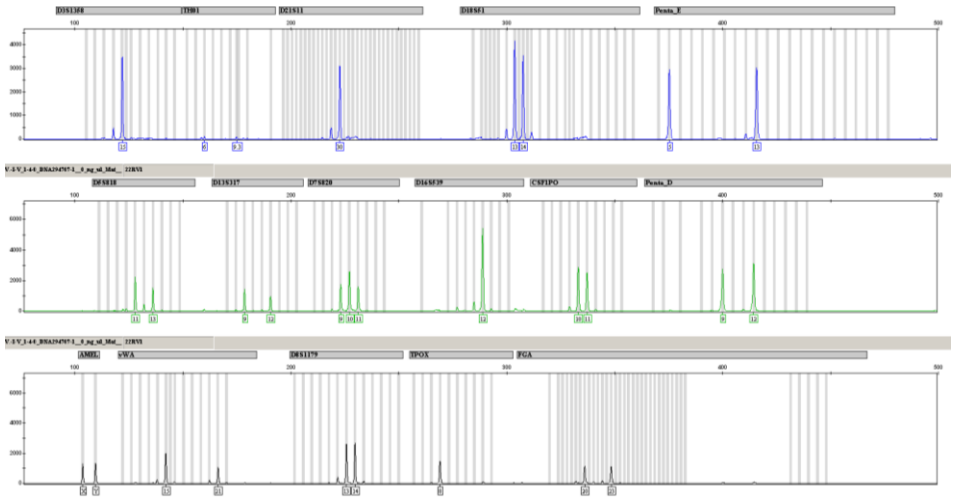
NUI Galway Cancer Research Day 2015 – NUI Galway – Human endogenous retroviruses as a biomarker in prostate cancer

Oral Presentations

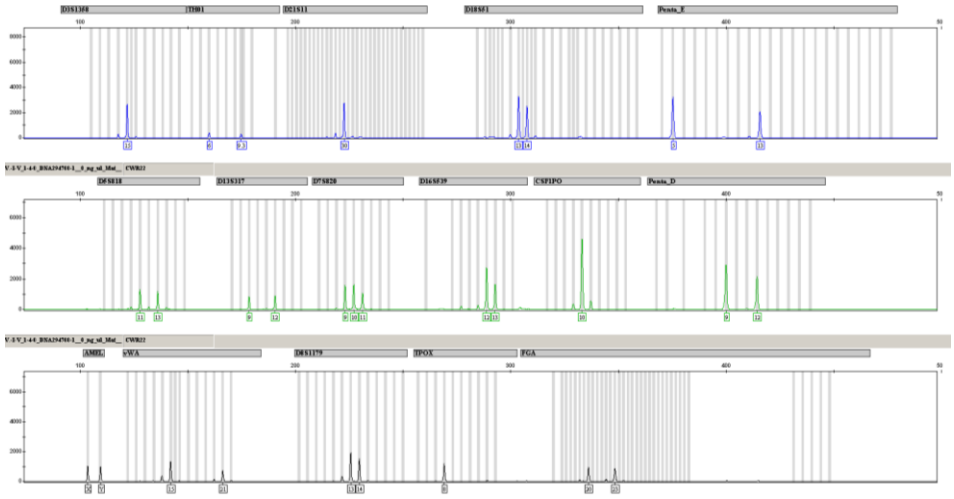
Irish Society of Urology (ISU) Annual Meeting 2014 – Killarney, Kerry – Human endogenous retrovirus k expression predicts a prostate cancer diagnosis

Appendix B - Cell line authentication reports

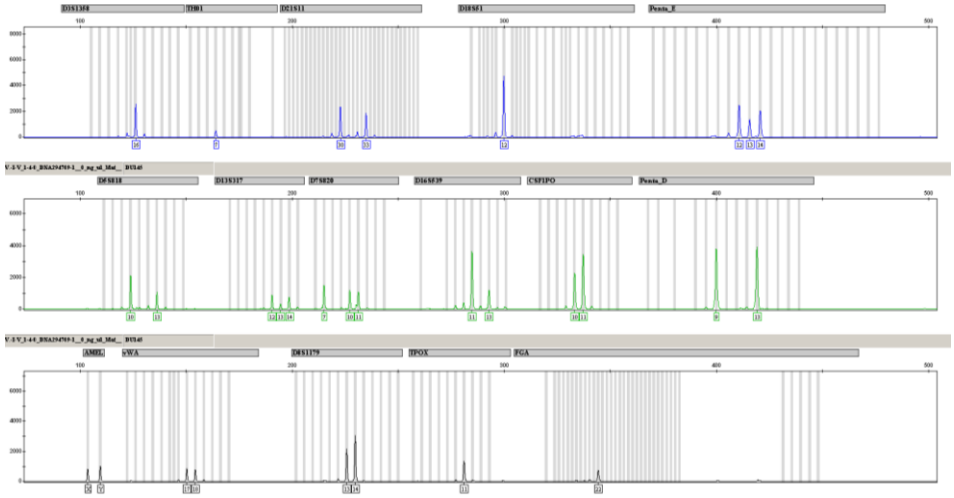
22RV1



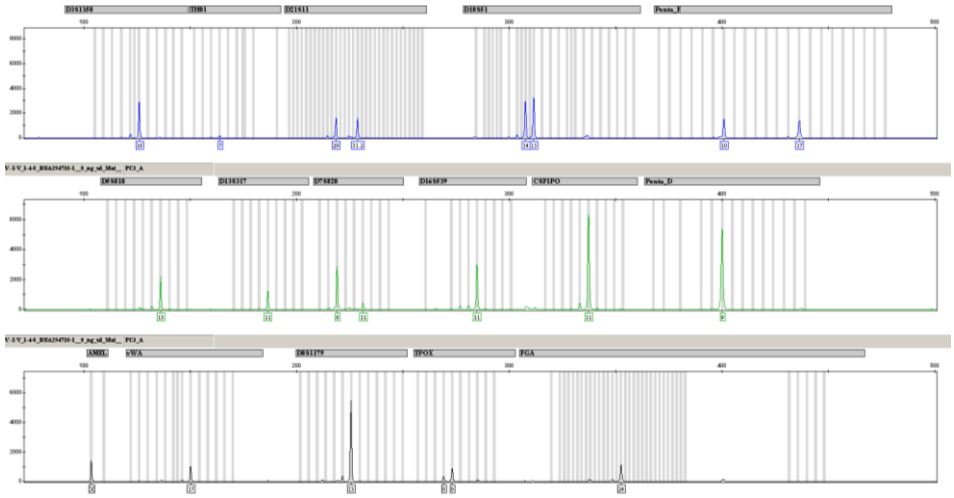
CWR22



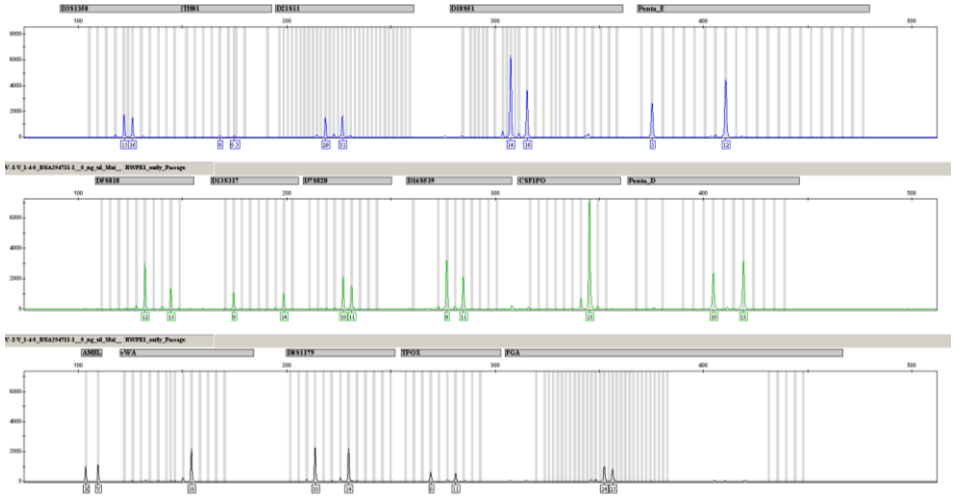
DU145



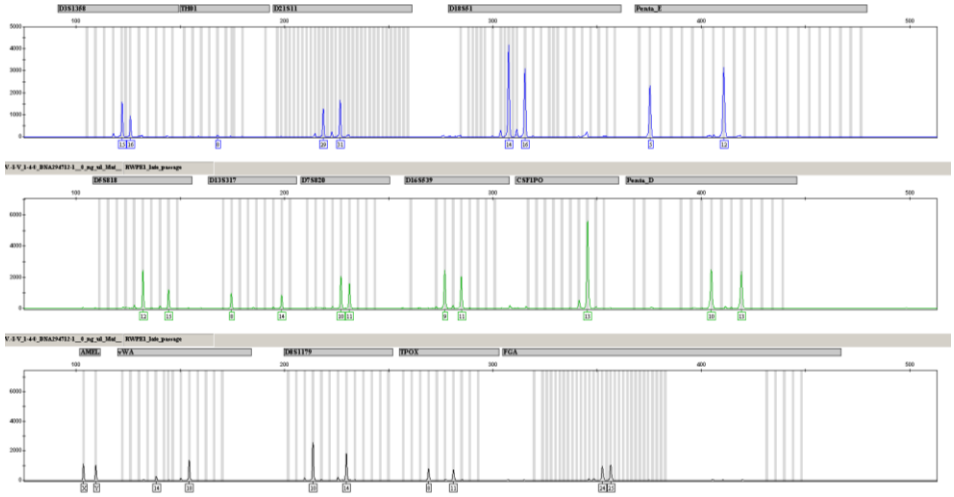
PC3_A



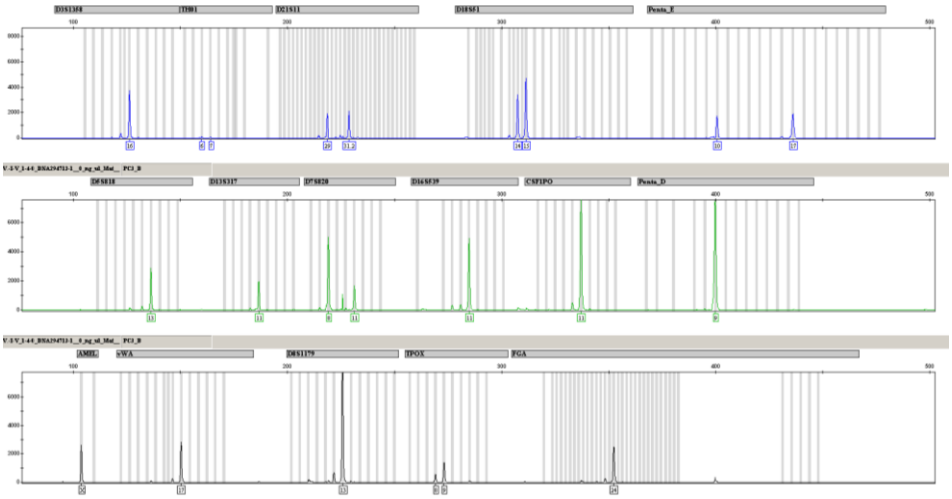
RWPE1_early_Passage



RWPE1_late_passage

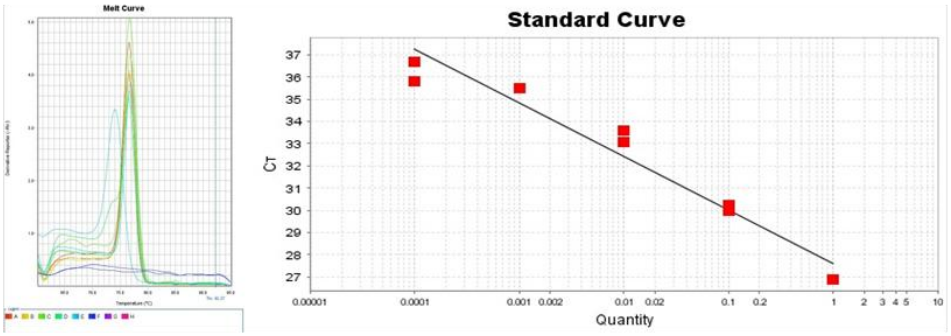


PC3_B

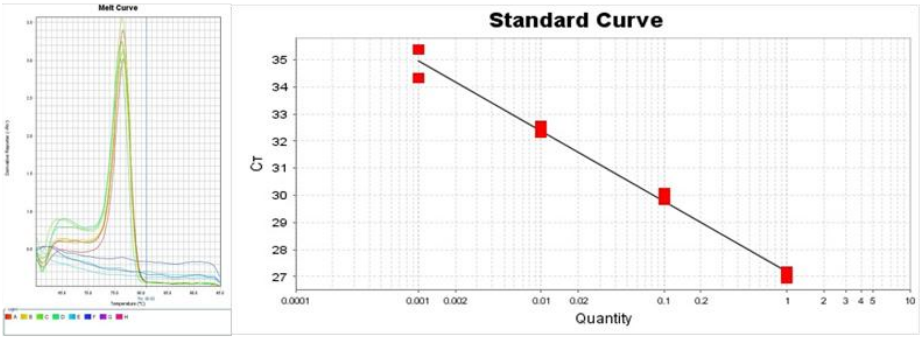


Appendix – Primer efficiency check

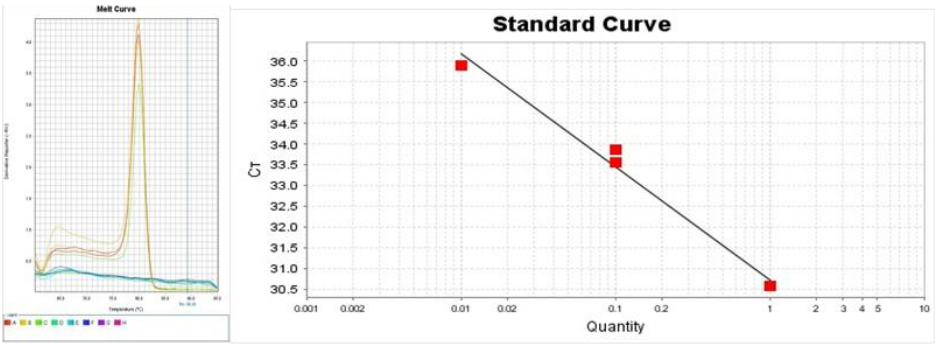
gag $r^2 = 0.95$



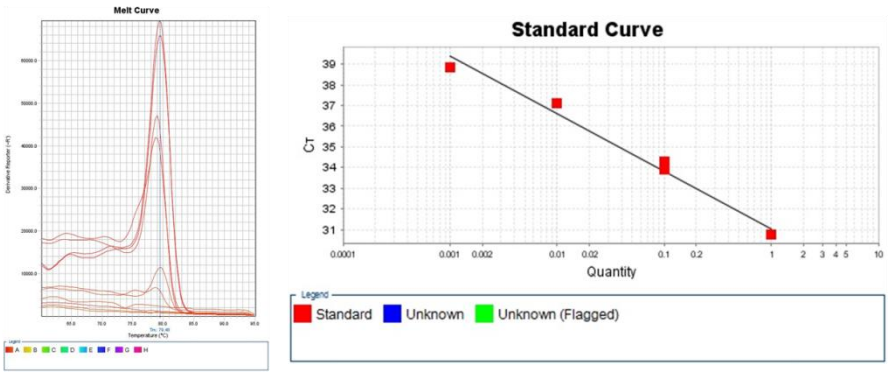
env1 $r^2 = 0.989$



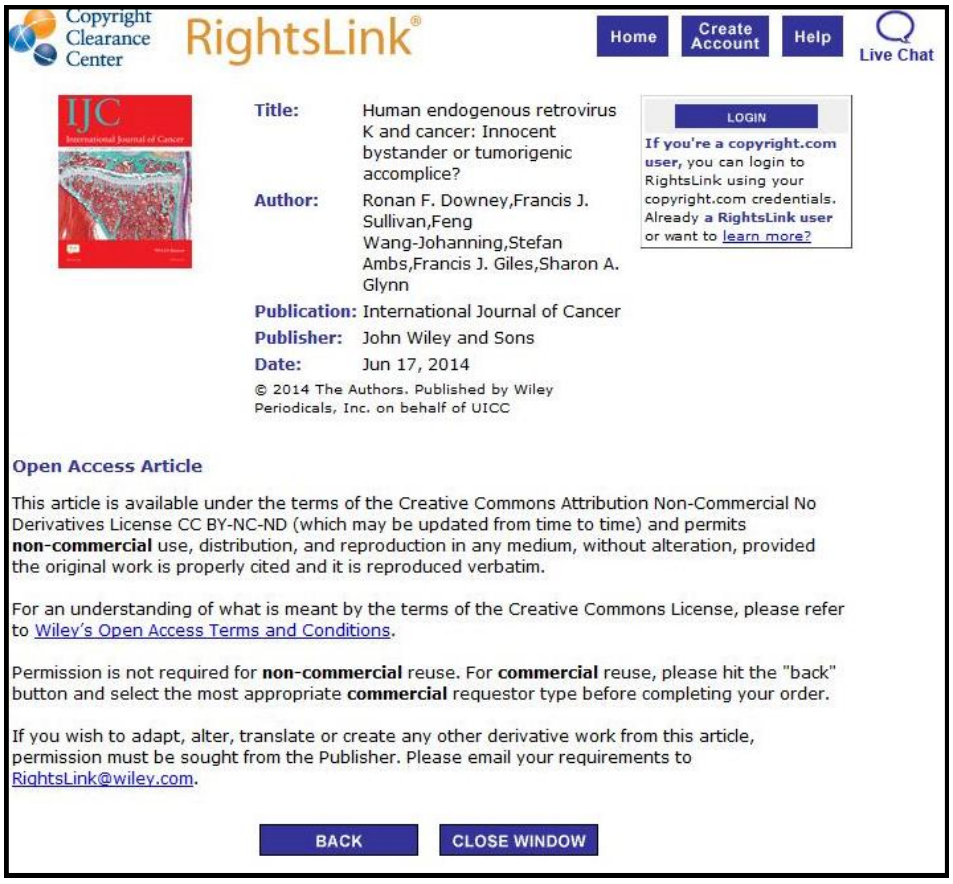
env2 $r^2 = 0.985$



Rec $r^2 = 0.983$



Appendix C - Copyright information and appended PDFs



Copyright Clearance Center RightsLink®

Home | Create Account | Help | Live Chat

Title: Human endogenous retrovirus K and cancer: Innocent bystander or tumorigenic accomplice?

Author: Ronan F. Downey,Francis J. Sullivan,Feng Wang-Johanning,Stefan Ambs,Francis J. Giles,Sharon A. Glynn

Publication: International Journal of Cancer

Publisher: John Wiley and Sons

Date: Jun 17, 2014

© 2014 The Authors. Published by Wiley Periodicals, Inc. on behalf of UICC

LOGIN

If you're a copyright.com user, you can login to RightsLink using your copyright.com credentials. Already a RightsLink user or want to learn more?

Open Access Article

This article is available under the terms of the Creative Commons Attribution Non-Commercial No Derivatives License CC BY-NC-ND (which may be updated from time to time) and permits **non-commercial** use, distribution, and reproduction in any medium, without alteration, provided the original work is properly cited and it is reproduced verbatim.

For an understanding of what is meant by the terms of the Creative Commons License, please refer to Wiley's Open Access Terms and Conditions.

Permission is not required for **non-commercial** reuse. For **commercial** reuse, please hit the "back" button and select the most appropriate **commercial** requestor type before completing your order.

If you wish to adapt, alter, translate or create any other derivative work from this article, permission must be sought from the Publisher. Please email your requirements to RightsLink@wiley.com.

BACK | CLOSE WINDOW

OXFORD UNIVERSITY PRESS LICENSE
TERMS AND CONDITIONS

Mar 01, 2016

This is a License Agreement between Ronan Downey ("You") and Oxford University Press ("Oxford University Press") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Oxford University Press, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3820300011595
License date	Mar 01, 2016
Licensed content publisher	Oxford University Press
Licensed content publication	Carcinogenesis
Licensed content title	Elevated HERV-K mRNA expression in PBMC is associated with a prostate cancer diagnosis particularly in older men and smokers:
Licensed content author	Tiffany A. Wallace, Ronan F. Downey, Caleb J. Seufert, Aaron Schetter, Tiffany H. Dorsey, Carol A. Johnson, Radoslav Goldman, Christopher A. Loffredo, Peisha Yan, Francis J. Sullivan, Francis J. Giles, Feng Wang-Johanning, Stefan Ambs, Sharon A. Glynn
Licensed content date	09/01/2014
Type of Use	Thesis/Dissertation
Institution name	None
Title of your work	Human endogenous retrovirus K (HERV-K) in prostate cancer
Publisher of your work	n/a
Expected publication date	Mar 2016
Permissions cost	0.00 EUR
Value added tax	0.00 EUR
Total	0.00 EUR
Total	0.00 EUR
Terms and Conditions	

IJC
International Journal of Cancer

Human endogenous retrovirus K and cancer: Innocent bystander or tumorigenic accomplice?

Ronan F. Downey[1], Francis J. Sullivan[1,2], Feng Wang-Johanning[3], Stefan Ambs[4], Francis J. Giles[1,5] and Sharon A. Glynn[1]

[1] Prostate Cancer Institute, National University of Ireland Galway, Galway, Ireland
[2] Department of Radiation Oncology, Galway University Hospitals, Galway, Ireland
[3] Center for Cancer and Metabolism, SRI International, Menlo Park, CA
[4] Laboratory of Human Carcinogenesis, National Cancer Institute, Bethesda, MD
[5] HRB Clinical Research Facilities Galway & Dublin, National University of Ireland Galway and Trinity College Dublin, Galway, Ireland

Mini Review

Harbored as relics of ancient germline infections, human endogenous retroviruses (HERVs) now constitute up to 8% of our genome. A proportion of this sequence has been co-opted for molecular and cellular processes, beneficial to human physiology, such as the fusogenic activity of the envelope protein, a vital component of placentogenesis. However, the discovery of high levels of HERV-K mRNA and protein and even virions in a wide array of cancers has revealed that HERV-K may be playing a more sinister role—a role as an etiological agent in cancer itself. Whether the presence of this retroviral material is simply an epiphenomenon, or an actual causative factor, is a hotly debated topic. This review will summarize the current state of knowledge regarding HERV-K and cancer and attempt to outline the potential mechanisms by which HERV-K could be involved in the onset and promotion of carcinogenesis.

One of the most striking findings that arose from the publication of the human genome sequence was the enormous swathe of transposable elements (TEs) it harbored.[1] Constituting ~45% of the entire sequence, they have co-evolved alongside the protein-coding component to contribute to modern-day phenotypes in ways which are still being deciphered. A subset of TEs, known as human endogenous retroviruses (HERVs), are ancestral relics of germline infections to which the genome succumbed over the course of evolution. The progenitors of these retroelements were exogenous retroviruses, which infected germline cells, subsequently became endogenized and subject to the laws of Mendelian inheritance.[2] HERVs share the genomic structure universal to all retroviruses: 5′LTR-*gag*, *pro*, *pol*, *env*−3′LTR. In retroviruses, these open reading frames (ORFs) encode viral polyproteins, which, after post-translational modification, become the critical structural and functional proteins, such as the reverse transcriptase or the transmembrane envelope, while the long

Key words: human endogenous retrovirus, HERV-K, carcinogenesis, melanoma, breast cancer, prostate cancer, HERV-K activation, oncogenesis, immunomodulation, Env, Gag, Np9, Rec

Abbreviations: AML: acute myeloid leukemia; APOBEC3G: apolipoprotein B mRNA-editing, enzyme-catalytic, polypeptide-like 3G; AR: androgen receptor; CML: chronic myeloid leukemia; EBV: Epstein–Barr virus; eMLV: ectropic murine leukemia virus; ERG1: ETS-related gene; ETV1: ETS translocation variant; HERV: human endogenous retrovirus; HIV-1: human immunodeficiency virus; HPV: human papillomavirus; hSGT: human small glutamine-rich tetratricopeptide repeat protein; HTLV-1: human T-lymphotropic virus; LTR: long terminal repeats; MHC: major histocompatability complex; MMTV: mouse mammary tumor virus; NO: nitric oxide; NOS2: nitric oxide synthase 2; ORF: open reading frames; PBMC: peripheral blood mononuclear cell; PLZF: promyelocytic leukemia zinc finger; Sags: superantigens; SEREX: serological recombinant cDNA expression cloning; TE: transposable element; TLR: toll-like receptor; TMPRSS2: transmembrane protease, serine 2; TZFP: testicular zinc-finger protein

Conflicts of interest: Nothing to report

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made

Grant sponsor: Galway University Foundation; **Grant number:** RNR1008; **Grant sponsor:** Health Research Board of Ireland Clinical Research Facility, Galway; **Grant number:** RSU004; **Grant sponsor:** Department of Defense Breast Cancer Research Program; **Grant number:** BC113114; **Grant sponsor:** Breast Cancer Campaign UK; **Grant number:** 2013MayPR019; **Grant sponsor:** Irish Cancer Society; **Grant number:** PCT13MCD

DOI: 10.1002/ijc.29003

History: Received 30 July 2013; Accepted 9 May 2014; Online 29 May 2014

Correspondence to: Sharon A. Glynn, Prostate Cancer Institute, National University of Ireland Galway, Galway, Ireland, Tel.: +353-91-494242, Fax: +353-91-494242, E-mail: sharon.glynn@nuigalway.ie

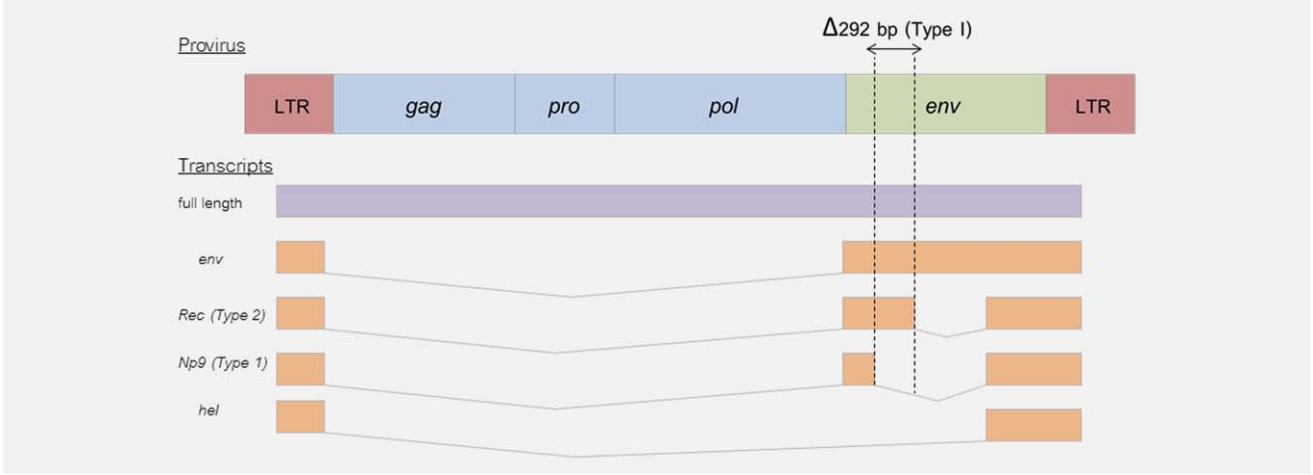


Figure 1. Structure of HERV-K provirus. The full length (gag) HML-2 transcript encodes the gag, pro and pol polyproteins. A singly spliced transcript encodes the env polyprotein, while a doubly spliced transcript encodes either the Rec or Np9 accessory proteins depending on the presence or absence of a 292-bp deletion at the pol/env boundary—a characteristic that defines a HML-2 provirus as either Type 1 (deleted) or Type 2 (intact). HML-2 also expresses a 1.5-kb transcript of unknown function known as the hel transcript.

terminal repeats (LTRs) specify promoter, enhancer and polyadenylation signals.[3] The vast majority of HERVs have acquired inactivating mutations such as stop codons or frameshifts, inhibiting the translation of functional proteins and thus making the production of a full, infectious retrovirus particle, from a single genetic locus, an impossibility.[4]

The HERV-K group is class II HERVs and exhibits closest homology to betaretroviruses, which cluster as class II retroelements. It consists of 11 subgroups (HML-1 to HML-11), each as the result of a separate germline infection.[5] One of these subgroups, HML-2, has been subject to intensive research because it maintains an unrivalled coding competence with many of its proviruses maintaining complete, or near-complete, ORFs for all viral polyproteins (Fig. 1).[3] Finally, it represents the most recently integrated HERV group into the human genome. Some HML-2 proviruses are both human specific and/or polymorphic indicating integration events subsequent to the human–chimpanzee split and within modern humans. This likely contributes to the fact that HML-2 is the least defective and most active retroviral family. In this regard, HML-2 is considered the most interesting HERV group to study in terms of potential oncogenic activity.

Overall, HML-2 is represented in the genome by 91 proviruses and 944 solitary (solo) LTRs. Solo LTRs are the result of unequal crossing over due to highly homologous sequences.[5] Two main types of HERV-K (HML-2) are found in humans: type I is characterized by a 292 base pair deletion at the boundary of the *pol* and *env* (envelope) genes (Fig. 1), whereas type II lacks it. The deletion in type I proviruses leads to an alternative splicing event culminating in a protein known as Np9, while type II proviruses express a complete accessory protein known as Rec.[3]

HERVs play an important role in normal physiological function. For example, the protein syncytin 1 mediates cellular fusion of the placental trophoblast and is encoded by an *env* gene from the HERV-W group.[6] Another syncytin—known as syncytin 2—plays a similar physiological role and is encoded by an env gene from the HERV-FRD group.[7] Finally, the presence of HERVs, in particular their LTR elements, has added an additional layer of complexity to our genome, in that many of these LTRs have been co-opted by protein-coding genes and serve as regulatory elements directing tissue-specific expression.[8]

The association of HERVs with disease has garnered the most attention from researchers. HERVs have been implicated in autoimmune disorders,[9,10] but with conflicting reports particularly involving multiple sclerosis (MS).[11–13] Recent research refutes a role for deregulated HERV-W *env* in MS lesions, including the high-level-transcribed *ERVWE1* locus encoding Syncytin-1.[14] In this review, we discuss the most recent developments in the field of HERV-K and human tumor biology, in particular emerging evidence of a role for HERV-K in immunomodulation and the presence of HERV-K in tumor-derived exosomes, further indicating the potentially important role of HERV-K in human carcinogenesis.

HERV-K and Solid Tumors

To date there is evidence for HML-2 activation in ovarian cancer,[15,16] melanoma,[17–19] breast,[20–24] prostate,[25–28] lymphomas,[29] leukemias[30] and sarcomas.[31] In the 1980s, Ono *et al.* successfully cloned HML-2, thanks to its similarity to mouse mammary tumor virus (MMTV).[32] They also found that stimulation of human breast cancer cell lines with female steroid hormones led to an upregulation of HML-2 mRNA.[33] Several groups followed with reports of HML-2 mRNA and viral particle expression in breast cancer.[29,34–36] Wang-Johanning *et al.* refined this work to produce data that accurately quantified HML-2 *env* transcripts and spliced transcripts in breast tumors demonstrating elevated levels

compared to unaffected controls.[20,21] They also demonstrated an association between HML-2 Env protein expression in breast tumors and increased risk of lymph node metastasis and poor outcome in two separate US cohorts and a Chinese cohort of breast cancer patients,[37,38] corroborating the findings of Golan *et al.*[23] Most recently, Wang-Johanning *et al.* demonstrated that HML-2 serum mRNA and anti-Rec antibody titers are predictive of early-stage breast cancer. Additionally, HERV-K-gag copy number tended to be higher in breast cancer patients with a primary tumor who later on developed the metastatic.[39]

High levels of expression of HML-2 *env*, *rec* and *np9* mRNA, and Env protein have been reported in ovarian cancer cell lines and tumors,[16] whereas in another study Np9 mRNA was not detectible in two ovarian tumors tested.[40] One possible mechanism of altered HML-2 expression in ovarian cancer may be due to alterations in its methylation status.[15]

Retrovirus-like particles and the expression of HML-2 mRNA and proteins are detectable in prostate cancer tissues. Ishida *et al.* isolated a HML-2 Gag protein in the serum of a prostate cancer patient using serological recombinant cDNA expression cloning (SEREX) technology.[25] They subsequently detected HML-2 *gag* mRNA in the serum of six of nine prostate cancer patients, but failed to detect HML-2 *gag* mRNA in LnCAP, DU145 or PC3 prostate cancer cells.[25] Gene fusions are a frequent occurrence in prostate cancer, the majority of which involve the fusion of the transcription factors ETS translocation variant (ETV1) or ETS-related gene (ERG1), to the transmembrane protease, serine 2 (TMPRSS2). In these fusions, the androgen-responsive TMPRSS2 drives expression of the ETV1 or ERG1 oncogenes. Recently, ETV1-HERV-K fusions have been described, corresponding to the 5′-untranslated region (UTR) of HERV-K-22q11.23[26] and HERV-K17.[41] Additionally, the ETV1-HERV-K-22q11.23 fusion is also inducible in LNCaP in response to androgen,[26] similar to HML-2 induction by estrogen and progesterone in breast cancer cell lines.[33]

Goering *et al.* detected significant expression of HERV-K-22q11.23 and HERV-K17 in the androgen-responsive prostate cancer cell lines 22Rv1, LNCaP and MDA-PCa-2b.[27] Normal prostate cells and androgen-insensitive prostate cancer cells (PC-3, DU-145 and BPH-1) exhibited expression near the limit of detection.[27] Expression of two other proviruses HERV-K-11q23.3 and HERV-K-22q.11.21 was not detectable in prostate cancer cell lines. Assessing HERV-K-22q11.23 5′UTR-*gag*, *env* and *Np9* gene expression in prostate tumors ($n = 45$) *versus* benign tissue ($n = 11$), the expression of the 5′UTR-*gag* and *env* region was significantly elevated in tumors compared to benign tissues. *Np9* was detectable only in a subset of carcinomas (18/45). In contrast, HERV-K17 was reduced in prostate tumors compared to benign. Where HERV-K-22q11.23 and HERV-K17 were expressed, they correlated with PSA levels, suggesting that HERV-K-22q11.23 and HERV-K17 retroelements are under androgen-inducible control, whereas HERV-K-11q23.3 and HERV-K-22q.11.21 are not.[27] Wallace *et al.* demonstrated that the HERV-K *gag* mRNA in peripheral blood mononuclear cells (PBMCs) is predictive of diagnosis with prostate cancer and correlates with elevated plasma interferon-γ and IP10.[42]

HERV-K and Hematological Malignancies

Brodsky *et al.* discovered a potential role for HERV-K in leukemia. They showed that HML-2 pol mRNA was expressed in the blood of patients suffering from chronic myeloid leukemia (CML) and acute myeloid leukemia (AML).[43,44] Others also reported that HML-2 *gag* mRNA is present at higher levels in PBMCs of leukemia patients compared to healthy controls.[30] Similar studies have reported HML-2 viral particles in lymphomas[29] and HML-2 env expression in the H9 human T-cell lymphoma cell line.[45] Additionally, the human lymphotropic herpesvirus Epstein–Barr virus (EBV), which has been implicated in the development of lymphoma, was shown to induce HERV-K18 *env* gene expression. The HERV-K18 env has been reported to have superantigen (SAg) activity by several groups,[46,47] whereas others have found no evidence of SAg activity.[48,49] Indeed, multiple HERV-K env proteins elicit antibody responses.[22,50] A direct association between HERV-K18 env SAg and carcinogenesis has yet to be shown. HML-2 expression has also been seen to decrease after lymphoma therapy, indicating that it may be useful for monitoring therapeutic response.[29]

HERV-K and Melanoma

The prevalence of HML-2 *pol, gag* and *env* mRNA, and Gag and Env proteins in melanoma is well established.[17–19,51–54] In 2002, a sequence homologous to HERV-K (HML-6) was identified in melanoma patients (HERV-K-MEL).[31] HERV-K-MEL, which produces an antigen spliced from a defective noncoding env-like ORF, was reported in cutaneous and ocular melanomas, and nevi. Antibodies raised against the HERV-K-MEL antigen were detectable in melanoma patients.[31] Melanoma cell lines (SKMel-28, SKMel-1, 518A2, MelJuso, HS-Mel2 and JH-Mel6 and HV-Mel7), but not cultured melanocytes (NHEM neo 5935, NHEM neo 4528 and NHEM neo 6083), produce retrovirus-like particles that exhibit reverse transcriptase activity,[52] which contain mature Gag and Env proteins. HML-2 Pol, Gag and Env,[52] and accessory proteins Rec and Np9 have also been detected in melanoma.[18,51] Further studies sought to predict the prognostic value of HERV-K in melanoma and found that HERV-K was a statistically significant marker of acrolentiginous, mucosal and uveal melanoma. Patients with serological response against HERV-K had a significantly decreased disease-specific overall survival.[55] Additionally, HML-2 rec mRNA is expressed in melanoma cells but not in benign nevi or normal skin, indicating aberrant activation in melanoma. Furthermore, rec mRNA positivity correlated with the vertical growth phase of melanoma, a step that increases the risk of metastatic melanoma.[56] A recent study by Schmitt *et al.*

defined the HML-2 transcriptome in melanoma, identifying 23 different HML-2 loci as transcribed to varying degrees in different patient specimens and melanoma cell lines.[57]

Polymorphic HML-2 Group Members

Of the 91 known HML-2 proviruses, 11 are polymorphic.[5] The most recent insertions (~1 million years ago) include HERV-K-19p12 (K113) (29% of individuals) and HERV-K-8p23.1 (K115) (16% of individuals) as measured using a pool of mixed backgrounds.[58] Other polymorphic HML-2 proviruses include: HERV-K-11q22.1 (K118), HERV-K-6q14.1 (K109), HERV-K-7p22.1a (K108R), HERV-K-8p23.1 (K115) and HERV-K-1p31.1(K116),[59,60] in addition to HERV-K-3q13.2 (K106), HERV-K-7p22.1b (K108L), HERV-K-10p12.1 (K103), HERV-K-12q13.2 and finally HERV-K-U219 (K105) located in the unassembled centromeric region (Un_g1000219).[5]

It is currently not known whether inheriting specific HML-2 polymorphisms increases the risk of cancer. Burmeister *et al.* investigated the frequency of the polymorphic full-length HERV-K115 and HERV-K113 in 102 female breast cancer cases and 102 controls, but did not find a significant association with breast cancer (HERV-K-K113, 16.7 *vs.* 12.7%; HERV-K-K115, 4.9 *vs.* 9.8%). (Note the lower prevalence than reported above[58] for both. This suggests ethnic differences in frequency of inheritance).[24]

Mechanisms of HERV Activation and Regulation

The abundance of inactive HERVs present in our genome suggests that active, integrating proviruses are largely deleterious to the host. Novel intrinsic restriction factors exist which impede retroviral infection and some of these have the ability to target both exogenous and endogenous infections. APOBEC proteins can inhibit viral RNA, thus blocking their translation.[61] Additionally, APOBEC3G can hypermutate and inactivate HERV DNA.[62] Activation of these retroelements can therefore be an indication that cellular programs, crucial to a healthy phenotype, have gone awry.

A crucial question that needs to be addressed is whether activation of HERVs is simply an epiphenomenon or is necessary for disease progression? A large proportion of HERV loci have become silenced *via* DNA hypermethylation, an epigenetic phenomenon.[63] Many cancers display a globally hypomethylated state[64]; thus, activation of HERVs during tumorigenesis may simply be a bystander effect of this epigenetic state. It has become increasingly clear that genomic instability, including deregulated transcription and genome plasticity, is enabled as a result of epigenetic changes that take place within tumors. Demethylation of specific HERV families, including HERV-W, HERV-K and HERV-H, has been reported in various cancers.[65] Moreover, demethylation of TEs correlates with their transcriptional activation in prostate cancer.[27] This indicates that where HERV transcription is increased in cancer cells, it is likely due in part to hypomethylation of their LTRs. HML-2 DNA hypomethylation has been reported in melanoma cell lines,[66] prostate tumors[27] and ovarian tumor.[15] Interestingly, age was negatively associated with HML-2 methylation in PBMCs from healthy donors aged 20–88 years. The average onset of HML-2 methylation in PBMCs occurred at 40–63 years, implicating HML-2 DNA hypomethylation in aging.[67] Another important epigenetic mechanism that influences transcriptional activity is histone modification, but the influence of histone methylation, acetylation or other modifications on HERV expression in malignancy is still unknown.

Known inducers of HML-2 *in vitro* include ultraviolet radiation in melanoma,[17,68] hormones, including progesterone, estrogen and androgen in breast[20,33] and prostate[26] cancer cell lines and bone morphogenetic proteins and retinoic acid in testicular germ cell tumor cell lines.[69] Urine from smokers has also been shown to lead to an increase in HERV expression in normal human dermal fibroblasts and urothelium *in vitro*.[70] Other factors that may activate or be activated by HERV-K include infectious viruses such as EBV[71] and human immunodeficiency virus (HIV-1),[72] and transcription factors including NF-κB, NF-AT,[73] MITF-M,[74] Sp1, Sp3[75] and YY1.[76]

Possible Mechanisms of HERV-K-Induced Oncogenesis

Insertional mutagenesis

HERVs may be oncogenic *via* insertional mutagenesis. However, to date, no fully intact and infectious HERV-derived retrovirus has been demonstrated *in vivo*. Retrovirus-like particles observed using electron microscopy in human placental trophoblasts,[77] and teratocarcinoma[78] and melanoma[52] were identified as HERV-K derived. Efforts to identify an infectious HERV-K are compounded by the fact that the large majority are partially defective and also that a somatic integration event would be a relatively rare occurrence.[3] Two independent groups have succeeded in resurrecting full retroviral particles after constructing consensus sequences representing ancestors of now defunct proviruses.[79,80] Although these viruses were found to be only weakly infectious, these studies will prove invaluable in formulating hypotheses regarding the potential oncogenic mechanisms of an infectious HERV-K (Fig. 2).[83,85,88,122,123]

HERV-K113 and HERV-K115 are some of the most recently integrated HERVs in the human genome and represent obvious candidates for infectious proviruses. Boller *et al.* investigated this possibility and observed that HERV-K113 is able to produce fully intact retroviral particles *in vitro*.[86] However, the authors concluded that an infectious HERV-K113 virus would be unlikely due to a lack of a functional reverse transcriptase.

HERV-K Rec and Np9 as putative oncogenes

Rec exhibits functional homology to the Rev protein of HIV-1, a nucleocytoplasmic shuttle protein.[3] Np9 is spliced from an alternative splice donor site to Rec, and shares only 14aa with Rec and Env, with no homology to Rev.[87] Functional

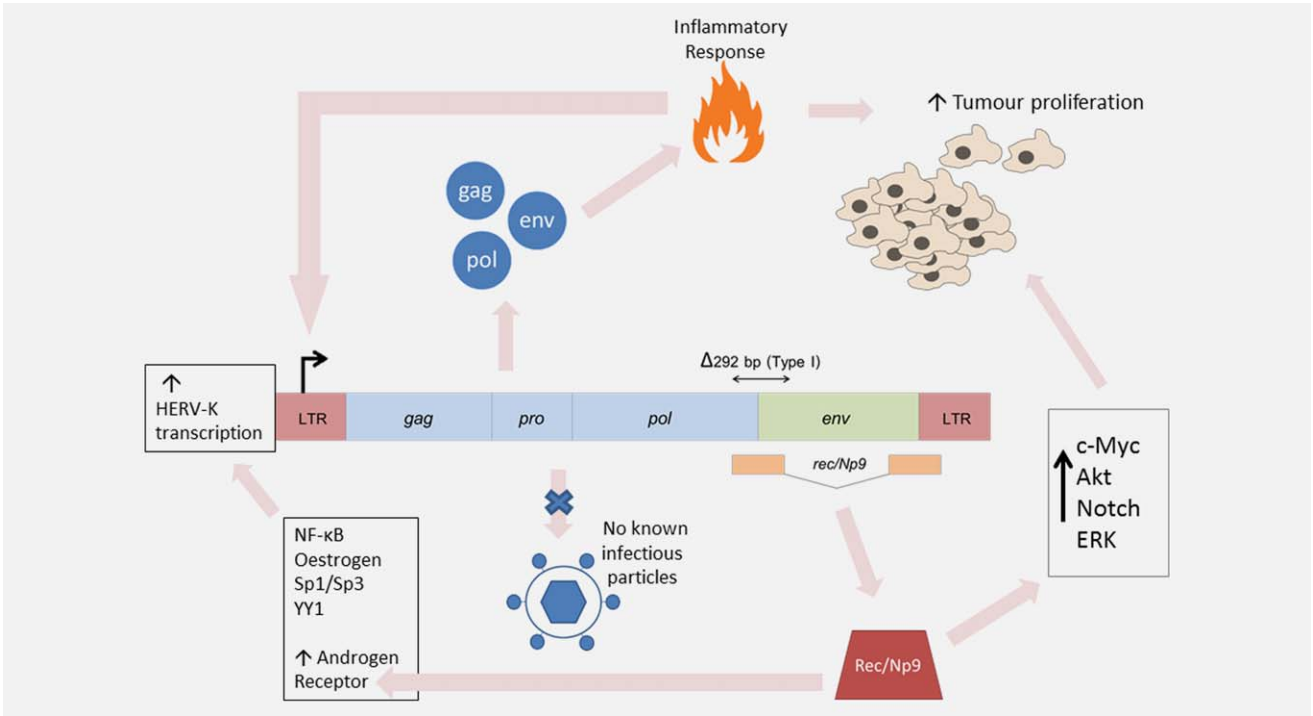


Figure 2. Proposed model of HERV-K (HML-2)-driven cancer progression. Global DNA hypomethylation during early-stage cancer leads to activation of otherwise silenced TEs, including HERVs. A humoral response to HERV-K gag has been observed in some cancers.[81] Such a response to high levels of HERV-K protein expression may culminate in chronic inflammation. Conversely, it has been hypothesized that HERV-K LTRs are responsive to inflammatory transcription factors—a phenomenon that may explain the high levels of HERV-K mRNA and protein seen in some inflammatory diseases.[82] HERV-K (HML-2) accessory proteins Rec and Np9 have been shown to lead to the derepression of the c-myc protooncogene,[83] while Np9 has been shown to co-activate Akt, Notch and ERK pathways in leukemia.[84] Rec has also been observed to lead to the derepression of the androgen receptor, which directly or undirectly causes a further increase in HERV-K transcription.[85] Overall, the synergistic effects of chronic inflammation and dysregulated signaling/protooncogene activation caused by HERV-K protein expression may help to create a protumorigenic microenvironment culminating in further proliferation and metastasis. Finally, it is important to note that an active, infectious HML-2 provirus has not been isolated to date, but the existence of such a particle cannot be ruled out. It would potentially be oncogenic *via* mechanisms such as insertional mutagenesis.

studies found that both proteins bind the promyelocytic leukemia zinc finger (PLZF) protein, a transcriptional repressor of the *C-MYC* proto-oncogene,[83] leading to the derepression of C-MYC. Rec also binds a related protein known as testicular zinc-finger protein (TZFP), a transcriptional repressor of the AR. Rec inhibits the ability of TZFP to repress AR transcription.[88] Hanke *et al.* identified an additional binding partner of Rec known as human small glutamine-rich tetratricopeptide repeat protein (hSGT), which also acts as a co-repressor of the AR.[85] Moreover, they proposed a "vicious cycle" model, whereby increased cellular AR led to increased transcription at HERV-K loci and thus increased levels of Rec leading to further AR derepression. The involvement of such hormonal regulators will be interesting to study in castration-resistant prostate cancer, in which disruption of the AR signaling axis is a key factor in development of resistance.

Reinforcing the possible importance of these proteins in tumorigenesis was the finding that mice transgenic for the *Rec* gene are prone to seminomas.[89] Np9 has been shown to interact with LNX—an E3 ubiquitin ligase that targets members of the NUMB/NOTCH pathway.[40] This pathway has been implicated in the regulation of proliferation of cancers of the breast and prostate.[90] Finally, a recent study has shown that Np9 acts as a critical molecular switch for co-activating β-catenin, ERK, Akt and Notch1 and promoting the growth of human leukemia stem/progenitor cells (Fig. 2).[84]

HERV-K-induced immunomodulation

In a Darwinian sense, cancerous tissue uses the inflammatory-associated milieu of the tumor microenvironment to confer a selective advantage.[91] The apparent immunogenicity of HERV proteins therefore represents a potential contributor to, or initiator of a chronic inflammatory state, beneficial to tumor survival (Fig. 2). HML-2 antibodies have been observed in patients with melanoma,[18] breast[39] and ovarian cancers.[16] In breast cancer, studies have found that both humoral and cell-mediated immune responses to HERVs were enhanced in patients when compared to controls.[22] HERV-K18 Env protein has been shown to elicit T-cell responses and can be upregulated in response to EBV infection,[46,92] and may be a prerequisite of B-cell lymphomas.[93]

Similar to discoveries in HIV-1, HERV-K may encode env proteins with immunosuppressive transmembrane domains.

A recent study by Morozov *et al.* identified an immunosuppressive HERV-K env protein, which altered cytokine expression and suppressed immune cell proliferation *in vitro*.[94]

Nitric oxide (NO) is an endogenous free radical signaling molecule that has been intimately linked with inflammation, wound healing responses and cancer.[95,96] A significant association between nitric oxide synthase 2 (NOS2) and HML-2 Env expression has been demonstrated in breast cancer.[37] NOS2 is an independent predictor of poor outcome in estrogen receptor-negative breast cancer, associated with macrophage infiltration, deregulated p53 signaling, increased proliferation and resistance to apoptosis.[95,97,98] Can HML-2 Env proteins mediate downstream inflammatory effects *via* their activation of NO signaling? Intriguingly, β-catenin, ERK and Akt, which are activated by Np9,[84] are also activated by NO signaling.[98,99]

Exosomes

An evolving hypothesis in cancer research over the last few years has been the involvement of tumor exosomes in metastasis.[100,101] Exosomes are nanoscale membrane vesicles that are secreted from cells and are thought to be important intercellular communicators, or, in a cancer setting, drivers of metastatic spread.[102] A recent study has now implicated HERVs in this process, with the finding that HML-2 mRNA is selectively packaged into tumor exosomes and that this genetic material can be transferred to normal cells.[103] The authors also demonstrated that these exosomes were enriched for the *C-MYC* protooncogene, which has been shown to be regulated by PLZF, a target of Rec and Np9.[83] Therefore, it is possible that there is a link between the high levels of HERV-K mRNA and *C-MYC* in these exosomes, but further investigation will have to be done in this regard. Another important point is that HML-2-driven metastasis *via* exosomes would not require an envelope gene, as exosomes gain entry to target cells *via* a plasma membrane fusion event. In essence, exosomes could potentially empower the abundance of defective HERVs with a new-found infectivity.[104]

HERV-K viral proteins as potential vaccines

Although the direct oncogenic effects of HERVs in cancer remain to be fully elucidated, there is potential for their use as diagnostic or prognostic biomarkers and for immunotherapeutic purposes including vaccines. Independent groups have demonstrated a strong association between HERV-K antibodies and clinical manifestation of disease and therapeutic response.[23,29] Antibodies recognizing synthetic HERV-K proteins were detected at a very low frequency in the sera of healthy donors.[16,22] Humoral anti-HERV-K immune response may provide additional prognostic information to that of established melanoma markers.[31,55] Data from these studies reveal a significant inverse correlation between serological anti-HERV-K reactivity and patient survival probability in melanoma patients. Among the different classes of tumor antigens recognizable by the immune system, mutated self-antigens and viral antigens are unique because they are foreign to the host and not subjected to preexisting antigen-specific tolerance.[105–107] HML-2 exons coding for mature proteins are spread out over the genome and are a repository of immunogenic retroviral gene products that can be "reawakened" when genetic damage occurs through chromosome breaks, frameshifts and mutations, removing sequences normally silencing protein expression.

HERV-K MEL is an antigenic peptide that is encoded by a short ORF from a processed HERV-K (HML-6) pseudogene and has been shown to be recognized by cytotoxic T cells in human melanoma.[31] BCG, vaccinia and yellow fever vaccinations are associated with a reduced risk of developing melanoma,[108–110] although conflicting data exist for yellow fever vaccines.[111] It is suggested that this effect is due to antigen sequence homology between these vaccines and HERV-K-MEL leading to cross-reaction between vaccine-elicited cytotoxic T cells and melanoma cells.[112] Reintroduction of these vaccines has been suggested as a novel method of melanoma immunoprevention; otherwise, HERV-K MEL represents a legitimate target for cellular immunotherapy.[112,113]

Future Perspectives

Over the course of evolution, our genome has been locked in a molecular "war" with exogenous infectious agents. Ultimately, it is this very battleground, together with viral endogenization, which has bestowed upon us the diverse genetic repertoire we possess today. Constituting 8% of our genome, these HERVs have supplied us with an additional layer of plasticity and physiological functionality, yet scientists now believe that hidden detrimental processes fueled by HERVs may be present, which are inducing chronic diseases such as cancer and autoimmunity. As of yet, no truly infectious HERVs have been observed. However, as outlined in this review, a range of potential molecular mechanisms involving the retroviral proteins may be aiding and abetting both tumor formation and metastasis. Ultimately, it is likely that many of these mechanisms are working synergistically to produce these effects, and the heralding of a single molecular event induced by a HERV protein is improbable.

Ascribing a causative role for a particular agent to a disease has long been a challenging task. Criteria such as Hills criteria[114] and Koch's postulates[115] have been formulated to address this problem. These criteria have recently been refined and built upon in light of HERVs-postulated role in human disease.[113,116,117] However, even if a direct link between HERVs and carcinogenesis is never established, their presence may be highly advantageous in terms of the implementation of novel biomarkers for cancer. Further work will need to effectively correlate their presence with various disease stages and also make the necessary comparisons against "gold standard" biomarkers. Equally promising is the potential to take advantage of tumor-specific HERV expression for the use of targeted immunotherapies. Wang-Johanning *et al.* have demonstrated the potential of anti-HML-2-Env

antibodies in inhibiting tumor growth and inducing apoptosis, both *in vitro* and in *in vivo* mouse xenograft models.[37] This work represents a major milestone in research into HERVs and cancer and it is likely that targeting Env in a similar fashion in other cancers will be equally effective. However, it remains imperative that these studies are evaluated in a clinical setting. Additionally, it may also be possible to conjugate these antibodies to cytotoxic drugs for increased effect.[118] Similarly, Kraus *et al.* demonstrated that HML-2-Env-targeted vaccine reduced renal tumor metastasis in a murine model.[119] Novel therapies, such as these, are key to making inroads toward a future cure for the increasingly complex and multistep disease that is cancer. However, their safety must be assessed given the newly established role of HERV-K in embryonic stem cells and iPS cells,[120] which may have implications for pregnancy. Their role in adult stem cells is not currently known.

Several limitations exist in the field of cancer-related HERV-K research, including a lack of adequately powered patient population studies to determine the role of HERV-K in the etiology of cancer, and/or its association with metastasis, therapeutic response and overall patient survival. A gap exists in our knowledge as to which HERV-K loci are specifically activated in cancer. A recent study by Schmitt *et al.* has defined the HML-2 transcriptome in melanoma, identifying 23 loci as transcribed,[57] and it is an imperative that similar studies be initiated in other cancers. A causal role for HML-2 has yet to be established. Generally, retroviruses induce tumours by insertional mutagenesis targeting specific oncogenes, as is the case with HBV.[121] This is an unlikely mechanism though in the case of HML-2. Evidence does suggest that Rec and Np9 may be putative oncogenes, but whether Gag or Env are also oncogenic is not known. In exceptional cases such as Jagsiekte sheep retrovirus (JSRV) the Env protein has been found to be causal (ovine pulmonary adenocarcinoma).[122] However, it is unlikely that HML-2 Gag or Env have a similar causal effect; potentially they may influence carcinogenesis by activating or perturbing inflammation responses against cancer.

It is our belief that within the next decade these genetic "squatters" will have firmly established themselves within the modern multistep model of cancer progression and their expression will be viewed as an "enabling characteristic" of cancer, giving new meaning to the famous words of Nobel laureate J. Michael Bishop when he stated that "the seeds of cancer are within us."[123]

Acknowledgements

This work was supported by the Galway University Foundation (RNR1008 to F.J.S. and S.A.G.); the Health Research Board of Ireland Clinical Research Facility, Galway (RSU004 to R.F.D. and F.J.G.); the Department of Defense Breast Cancer Research Program (BC113114 to F.J.W.); Breast Cancer Campaign UK (2013MayPR019 to S.A.G. and F.J.S.) and the Irish Cancer Society (PCT13MCD to S.A.G. and F.J.S.).

References

1. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature* 2001;409:860–921.
2. Lower R, Lower J, Kurth R. The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences. *Proc Natl Acad Sci USA* 1996;93:5177–84.
3. Bannert N, Kurth R. The evolutionary dynamics of human endogenous retroviral families. *Annu Rev Genomics Hum Genet* 2006;7:149–73.
4. Bannert N, Kurth R. Retroelements and the human genome: new perspectives on an old relation. *Proc Natl Acad Sci USA* 2004;101 (Suppl 2):14572–9.
5. Subramanian RP, Wildschutte JH, Russo C, et al. Identification, characterization, and comparative genomic distribution of the HERV-K (HML-2) group of human endogenous retroviruses. *Retrovirology* 2011;8:90.
6. Mi S, Lee X, Li X, et al. Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. *Nature* 2000;403:785–9.
7. Blaise S, Ruggieri A, Dewannieux M, et al. Identification of an envelope protein from the FRD family of human endogenous retroviruses (HERV-FRD) conferring infectivity and functional conservation among simians. *J Virol* 2004; 78:1050–4.
8. Cohen CJ, Lock WM, Mager DL. Endogenous retroviral LTRs as promoters for human genes: a critical assessment. *Gene* 2009;448: 105–14.
9. Balada E, Ordi-Ros J, Vilardell-Tarres M. Molecular mechanisms mediated by human endogenous retroviruses (HERVs) in autoimmunity. *Rev Med Virol* 2009;19:273–86.
10. Urnovitz HB, Murphy WH. Human endogenous retroviruses: nature, occurrence, and clinical implications in human disease. *Clin Microbiol Rev* 1996;9:72–99.
11. Perron H, Lang A. The human endogenous retrovirus link between genes and environment in multiple sclerosis and in multifactorial diseases associating neuroinflammation. *Clin Rev Allergy Immunol* 2010;39:51–61.
12. Curtin F, Lang AB, Perron H, et al. GNbAC1, a humanized monoclonal antibody against the envelope protein of multiple sclerosis-associated endogenous retrovirus: a first-in-humans randomized clinical study. *Clin Ther* 2012;34: 2268–78.
13. Mameli G, Poddighe L, Mei A, et al. Expression and activation by Epstein Barr virus of human endogenous retroviruses-W in blood cells and astrocytes: inference for multiple sclerosis. *PLoS One* 2012;7:e44991.
14. Schmitt K, Richter C, Backes C, et al. Comprehensive analysis of human endogenous retrovirus group HERV-W locus transcription in multiple sclerosis brain lesions by high-throughput amplicon sequencing. *J Virol* 2013; 87:13837–52.
15. Iramaneerat K, Rattanatunyong P, Khemapech N, et al. HERV-K hypomethylation in ovarian clear cell carcinoma is associated with a poor prognosis and platinum resistance. *Int J Gynecol Cancer* 2011;21:51–7.
16. Wang-Johanning F, Liu J, Rycaj K, et al. Expression of multiple human endogenous retrovirus surface envelope proteins in ovarian cancer. *Int J Cancer* 2007;120:81–90.
17. Reiche J, Pauli G, Ellerbrok H. Differential expression of human endogenous retrovirus K transcripts in primary human melanocytes and melanoma cell lines after UV irradiation. *Melanoma Res* 2010;20:435–40.
18. Buscher K, Trefzer U, Hofmann M, et al. Expression of human endogenous retrovirus K in melanomas and melanoma cell lines. *Cancer Res* 2005;65:4172–80.
19. Serafino A, Balestrieri E, Pierimarchi P, et al. The activation of human endogenous retrovirus K (HERV-K) is implicated in melanoma cell malignant transformation. *Exp Cell Res* 2009; 315:849–62.
20. Wang-Johanning F, Frost AR, Jian B, et al. Quantitation of HERV-K env gene expression and splicing in human breast cancer. *Oncogene* 2003;22:1528–35.
21. Wang-Johanning F, Frost AR, Johanning GL, et al. Expression of human endogenous retrovirus k envelope transcripts in human breast cancer. *Clin Cancer Res* 2001;7:1553–60.
22. Wang-Johanning F, Radvanyi L, Rycaj K, et al. Human endogenous retrovirus K triggers an antigen-specific immune response in breast cancer patients. *Cancer Res* 2008;68:5869–77.

23. Golan M, Hizi A, Resau JH, et al. Human endogenous retrovirus (HERV-K) reverse transcriptase as a breast cancer prognostic marker. *Neoplasia* 2008;10:521–33.
24. Burmeister T, Ebert AD, Pritze W, et al. Insertional polymorphisms of endogenous HERV-K113 and HERV-K115 retroviruses in breast cancer patients and age-matched controls. *AIDS Res Hum Retroviruses* 2004;20:1223–9.
25. Ishida T, Obata Y, Ohara N, et al. Identification of the HERV-K gag antigen in prostate cancer by SEREX using autologous patient serum and its immunogenicity. *Cancer Immun* 2008;8:15.
26. Tomlins SA, Laxman B, Dhanasekaran SM, et al. Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. *Nature* 2007;448:595–9.
27. Goering W, Ribarska T, Schulz WA. Selective changes of retroelement expression in human prostate cancer. *Carcinogenesis* 2011;32:1484–92.
28. Agoni L, Guha C, Lenz J. Detection of human endogenous retrovirus K (HERV-K) transcripts in human prostate cancer cell lines. *Front Oncol* 2013;3:180.
29. Contreras-Galindo R, Kaplan MH, Leissner P, et al. Human endogenous retrovirus K (HML-2) elements in the plasma of people with lymphoma and breast cancer. *J Virol* 2008;82:9329–36.
30. Depil S, Roche C, Dussart P, et al. Expression of a human endogenous retrovirus, HERV-K, in the blood cells of leukemia patients. *Leukemia* 2002;16:254–9.
31. Schiavetti F, Thonnard J, Colau D, et al. A human endogenous retroviral sequence encoding an antigen recognized on melanoma by cytolytic T lymphocytes. *Cancer Res* 2002;62:5510–16.
32. Ono M, Yasunaga T, Miyata T, et al. Nucleotide sequence of human endogenous retrovirus genome related to the mouse mammary tumor virus genome. *J Virol* 1986;60:589–98.
33. Ono M, Kawakami M, Ushikubo H. Stimulation of expression of the human endogenous retrovirus genome by female steroid hormones in human breast cancer cell line T47D. *J Virol* 1987;61:2059–62.
34. Seifarth W, Skladny H, Krieg-Schneider F, et al. Retrovirus-like particles released from the human breast cancer cell line T47-D display type B- and C-related endogenous retroviral sequences. *J Virol* 1995;69:6408–16.
35. Willer A, Saussele S, Gimbel W, et al. Two groups of endogenous MMTV related retroviral env transcripts expressed in human tissues. *Virus Genes* 1997;15:123–33.
36. Etkind PR, Lumb K, Du J, et al. Type 1 HERV-K genome is spliced into subgenomic transcripts in the human breast tumor cell line T47D. *Virology* 1997;234:304–8.
37. Wang-Johanning F, Rycaj K, Plummer JB, et al. Immunotherapeutic potential of anti-human endogenous retrovirus-K envelope protein antibodies in targeting breast tumors. *J Natl Cancer Inst* 2012;104:189–210.
38. Zhao J, Rycaj K, Geng S, et al. Expression of human endogenous retrovirus type K envelope protein is a novel candidate prognostic marker for human breast cancer. *Genes Cancer* 2011;2:914–22.
39. Wang-Johanning F, Li M, Esteva FJ, et al. Human endogenous retrovirus type K antibodies and mRNA as serum biomarkers of early-stage breast cancer. *Int J Cancer* 2014;134:587–95.
40. Armbruester V, Sauter M, Krautkraemer E, et al. A novel gene from the human endogenous retrovirus K expressed in transformed cells. *Clin Cancer Res* 2002;8:1800–7.
41. Hermans KG, van der Korput HA, van Marion R, et al. Truncated ETV1, fused to novel tissue-specific genes, and full-length ETV1 in prostate cancer. *Cancer Res* 2008;68:7541–9.
42. Wallace TA, Downey RF, Seufert CJ, et al. Elevated HERV-K mRNA expression in PBMC is associated with a prostate cancer diagnosis particularly in older men and smokers. *Carcinogenesis* 2014, May 23, pii: bgu114 (Epub ahead of print).
43. Brodsky I, Foley B, Gillespie D. Expression of human endogenous retrovirus (HERV-K) in chronic myeloid leukemia. *Leuk Lymphoma* 1993;11 (Suppl 1):119–23.
44. Brodsky I, Foley B, Haines D, et al. Expression of HERV-K proviruses in human leukocytes. *Blood* 1993;81:2369–74.
45. Vogetseder W, Feng J, Dierich MP. Reactivity of monoclonal antibodies established against a recombinant human endogenous retrovirus-K (HERV-K) envelope protein. *Immunol Lett* 1995;46:129–34.
46. Sutkowski N, Conrad B, Thorley-Lawson DA, et al. Epstein-Barr virus transactivates the human endogenous retrovirus HERV-K18 that encodes a superantigen. *Immunity* 2001;15:579–89.
47. Hsiao FC, Tai AK, Deglon A, et al. EBV LMP-2A employs a novel mechanism to transactivate the HERV-K18 superantigen through its ITAM. *Virology* 2009;385:261–6.
48. Azar GA, Thibodeau J. Human endogenous retrovirus IDDMK(1,2)22 and mouse mammary tumor virus superantigens differ in their ability to stimulate murine T cell hybridomas. *Immunol Lett* 2002;81:87–91.
49. Badenhoop K, Donner H, Neumann J, et al. IDDM patients neither show humoral reactivities against endogenous retroviral envelope protein nor do they differ in retroviral mRNA expression from healthy relatives or normal individuals. *Diabetes* 1999;48:215–18.
50. Herve CA, Lugli EB, Brand A, et al. Autoantibodies to human endogenous retrovirus-K are frequently detected in health and disease and react with multiple epitopes. *Clin Exp Immunol* 2002;128:75–82.
51. Buscher K, Hahn S, Hofmann M, et al. Expression of the human endogenous retrovirus-K transmembrane envelope, Rec and Np9 proteins in melanomas and melanoma cell lines. *Melanoma Res* 2006;16:223–34.
52. Muster T, Waltenberger A, Grassauer A, et al. An endogenous retrovirus derived from human melanoma cells. *Cancer Res* 2003;63:8735–41.
53. Humer J, Waltenberger A, Grassauer A, et al. Identification of a melanoma marker derived from melanoma-associated endogenous retroviruses. *Cancer Res* 2006;66:1658–63.
54. Singh S, Kaye S, Gore ME, et al. The role of human endogenous retroviruses in melanoma. *Br J Dermatol* 2009;161:1225–31.
55. Hahn S, Ugurel S, Hanschmann KM, et al. Serological response to human endogenous retrovirus K in melanoma patients correlates with survival probability. *AIDS Res Hum Retroviruses* 2008;24:717–23.
56. Singh S, Kaye S, Francis N, et al. Human endogenous retrovirus K (HERV-K) rec mRNA is expressed in primary melanoma but not in benign naevi or normal skin. *Pigment Cell Melanoma Res* 2013;26:426–8.
57. Schmitt K, Reichrath J, Roesch A, et al. Transcriptional profiling of human endogenous retrovirus group HERV-K(HML-2) loci in melanoma. *Genome Biol Evol* 2013;5:307–28.
58. Turner G, Barbulescu M, Su M, et al. Insertional polymorphisms of full-length endogenous retroviruses in humans. *Curr Biol* 2001;11:1531–5.
59. Hughes JF, Coffin JM. Human endogenous retrovirus K solo-LTR formation and insertional polymorphisms: implications for human and viral evolution. *Proc Natl Acad Sci USA* 2004;101:1668–72.
60. Belshaw R, Dawson AL, Woolven-Allen J, et al. Genomewide screening reveals high levels of insertional polymorphism in the human endogenous retrovirus family HERV-K(HML2): implications for present-day activity. *J Virol* 2005;79:12507–14.
61. Mullins CS, Linnebacher M. Human endogenous retroviruses and cancer: causality and therapeutic possibilities. *World J Gastroenterol* 2012;18:6027–35.
62. Lee YN, Malim MH, Bieniasz PD. Hypermutation of an ancient human retrovirus by APOBEC3G. *J Virol* 2008;82:8762–70.
63. Schulz WA, Steinhoff C, Florl AR. Methylation of endogenous human retroelements in health and disease. *Curr Top Microbiol Immunol* 2006;310:211–50.
64. Ehrlich M. DNA methylation in cancer: too much, but also too little. *Oncogene* 2002;21:5400–13.
65. Gimenez J, Montgiraud C, Pichon JP, et al. Custom human endogenous retroviruses dedicated microarray identifies self-induced HERV-W family elements reactivated in testicular cancer upon methylation control. *Nucleic Acids Res* 2010;38:2229–46.
66. Stengel S, Fiebig U, Kurth R, et al. Regulation of human endogenous retrovirus-K expression in melanomas by CpG methylation. *Genes Chromosomes Cancer* 2010;49:401–11.
67. Jintaridth P, Mutirangura A. Distinctive patterns of age-dependent hypomethylation in interspersed repetitive sequences. *Physiol Genomics* 2010;41:194–200.
68. Schanab O, Humer J, Gleiss A, et al. Expression of human endogenous retrovirus K is stimulated by ultraviolet radiation in melanoma. *Pigment Cell Melanoma Res* 2011;24:656–65.
69. Caricasole A, Ward-van Oostwaard D, Mummery C, et al. Bone morphogenetic proteins and retinoic acid induce human endogenous retrovirus HERV-K expression in NT2D1 human embryonal carcinoma cells. *Dev Growth Differ* 2000;42:407–11.
70. Gabriel U, Steidler A, Trojan L, et al. Smoking increases transcription of human endogenous retroviruses in a newly established in vitro cell model and in normal urothelium. *AIDS Res Hum Retroviruses* 2010;26:883–8.
71. Sutkowski N, Chen G, Calderon G, et al. Epstein-Barr virus latent membrane protein LMP-2A is sufficient for transactivation of the human endogenous retrovirus HERV-K18 superantigen. *J Virol* 2004;78:7852–60.
72. Jones RB, Garrison KE, Mujib S, et al. HERV-K-specific T cells eliminate diverse HIV-1/2 and

SIV primary isolates. *J Clin Invest* 2012;122: 4473–89.
73. Gonzalez-Hernandez MJ, Swanson MD, Contreras-Galindo R, et al. Expression of human endogenous retrovirus type K (HML-2) is activated by the Tat protein of HIV-1. *J Virol* 2012; 86:7790–805.
74. Katoh I, Mirova A, Kurata S, et al. Activation of the long terminal repeat of human endogenous retrovirus K by melanoma-specific transcription factor MITF-M. *Neoplasia* 2011;13:1081–92.
75. Fuchs NV, Kraft M, Tondera C, et al. Expression of the human endogenous retrovirus (HERV) group HML-2/HERV-K does not depend on canonical promoter elements but is regulated by transcription factors Sp1 and Sp3. *J Virol* 2011;85:3436–48.
76. Knossl M, Lower R, Lower J. Expression of the human endogenous retrovirus HTDV/HERV-K is enhanced by cellular transcription factor YY1. *J Virol* 1999;73:1254–61.
77. Kalter SS, Helmke RJ, Heberling RL, et al. Brief communication: C-type particles in normal human placentas. *J Natl Cancer Inst* 1973;50: 1081–4.
78. Boller K, Konig H, Sauter M, et al. Evidence that HERV-K is the endogenous retrovirus sequence that codes for the human teratocarcinoma-derived retrovirus HTDV. *Virology* 1993;196:349–53.
79. Dewannieux M, Harper F, Richaud A, et al. Identification of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements. *Genome Res* 2006;16:1548–56.
80. Lee YN, Bieniasz PD. Reconstitution of an infectious human endogenous retrovirus. *PLoS Pathog* 2007;3:e10.
81. Sauter M, Schommer S, Kremmer E, et al. Human endogenous retrovirus K10: expression of Gag protein and detection of antibodies in patients with seminomas. *J Virol* 1995;69:414–21.
82. Manghera M, Douville RN. Endogenous retrovirus-K promoter: a landing strip for inflammatory transcription factors? *Retrovirology* 2013;10:16.
83. Denne M, Sauter M, Armbruester V, et al. Physical and functional interactions of human endogenous retrovirus proteins Np9 and rec with the promyelocytic leukemia zinc finger protein. *J Virol* 2007;81:5607–16.
84. Chen T, Meng Z, Gan Y, et al. The viral oncogene Np9 acts as a critical molecular switch for co-activating beta-catenin, ERK, Akt and Notch1 and promoting the growth of human leukemia stem/progenitor cells. *Leukemia* 2013;27:1469–78.
85. Hanke K, Chudak C, Kurth R, et al. The Rec protein of HERV-K(HML-2) upregulates androgen receptor activity by binding to the human small glutamine-rich tetratricopeptide repeat protein (hSGT). *Int J Cancer* 2013;132: 556–67.
86. Boller K, Schonfeld K, Lischer S, et al. Human endogenous retrovirus HERV-K113 is capable of producing intact viral particles. *J Gen Virol* 2008;89:567–72.
87. Ruggieri A, Maldener E, Sauter M, et al. Human endogenous retrovirus HERV-K(HML-2) encodes a stable signal peptide with biological properties distinct from Rec. *Retrovirology* 2009;6:17.
88. Kaufmann S, Sauter M, Schmitt M, et al. Human endogenous retrovirus protein Rec interacts with the testicular zinc-finger protein and androgen receptor. *J Gen Virol* 2010;91: 1494–502.
89. Galli UM, Sauter M, Lecher B, et al. Human endogenous retrovirus rec interferes with germ cell development in mice and may cause carcinoma in situ, the predecessor lesion of germ cell tumors. *Oncogene* 2005;24:3223–8.
90. Roy M, Pear WS, Aster JC. The multifaceted role of Notch in cancer. *Curr Opin Genet Dev* 2007;17:52–9.
91. Coussens LM, Werb Z. Inflammation and cancer. *Nature* 2002;420:860–7.
92. Stauffer Y, Marguerat S, Meylan F, et al. Interferon-alpha-induced endogenous superantigen. a model linking environment and autoimmunity. *Immunity* 2001;15:591–601.
93. Zur Hausen H. The search for infectious causes of human cancers: where and why. *Virology* 2009;392:1–10.
94. Morozov VA, Dao Thi VL, Denner J. The transmembrane protein of the human endogenous retrovirus–K (HERV-K) modulates cytokine release and gene expression. *PLoS One* 2013;8: e70399.
95. Burke AJ, Sullivan FJ, Giles FJ, et al. The yin and yang of nitric oxide in cancer progression. *Carcinogenesis* 2013;34:503–12.
96. Rapozzi V, Della Pietra E, Zorzet S, et al. Nitric oxide-mediated activity in anti-cancer photodynamic therapy. *Nitric Oxide* 2013;30:26–35.
97. Ambs S, Glynn SA. Candidate pathways linking inducible nitric oxide synthase to a basal-like transcription pattern and tumor progression in human breast cancer. *Cell Cycle* 2011;10:619–24.
98. Glynn SA, Boersma BJ, Dorsey TH, et al. Increased NOS2 predicts poor survival in estrogen receptor-negative breast cancer patients. *J Clin Invest* 2010;120:3843–54.
99. Switzer CH, Glynn SA, Cheng RY, et al. S-Nitrosylation of EGFR and Src activates an oncogenic signaling network in human basal-like breast cancer. *Mol Cancer Res* 2012;10: 1203–15.
100. Yang C, Robbins PD. The roles of tumor-derived exosomes in cancer pathogenesis. *Clin Dev Immunol* 2011;2011:842849.
101. Duijvesz D, Luider T, Bangma CH, et al. Exosomes as biomarker treasure chests for prostate cancer. *Eur Urol* 2011;59:823–31.
102. Thery C, Zitvogel L, Amigorena S. Exosomes: composition, biogenesis and function. *Nat Rev Immunol* 2002;2:569–79.
103. Balaj L, Lessard R, Dai L, et al. Tumour microvesicles contain retrotransposon elements and amplified oncogene sequences. *Nat Commun* 2011;2:180.
104. Gould SJ, Booth AM, Hildreth JE. The Trojan exosome hypothesis. *Proc Natl Acad Sci USA* 2003;100:10592–7.
105. Smith C, Tsang J, Beagley L, et al. Effective treatment of metastatic forms of Epstein-Barr virus-associated nasopharyngeal carcinoma with a novel adenovirus-based adoptive immunotherapy. *Cancer Res* 2012;72:1116–25.
106. Foster AE, Rooney CM. Improving T cell therapy for cancer. *Expert Opin Biol Ther* 2006;6: 215–29.
107. Gottschalk S, Heslop HE, Rooney CM. Adoptive immunotherapy for EBV-associated malignancies. *Leuk Lymphoma* 2005;46:1–10.
108. Pfahlberg A, Kolmel KF, Grange JM, et al. Inverse association between melanoma and previous vaccinations against tuberculosis and smallpox: results of the FEBIM study. *J Invest Dermatol* 2002;119:570–5.
109. Krone B, Kolmel KF, Grange JM, et al. Impact of vaccinations and infectious diseases on the risk of melanoma—evaluation of an EORTC case-control study. *Eur J Cancer* 2003;39:2372–8.
110. Mastrangelo G, Krone B, Fadda E, et al. Does yellow fever 17D vaccine protect against melanoma? *Vaccine* 2009;27:588–91.
111. Hodges-Vazquez M, Wilson JP, Hughes H, et al. The yellow fever 17D vaccine and risk of malignant melanoma in the United States military. *Vaccine* 2012;30:4476–9.
112. Krone B, Kolmel KF, Henz BM, et al. Protection against melanoma by vaccination with Bacille Calmette-Guerin (BCG) and/or vaccinia: an epidemiology-based hypothesis on the nature of a melanoma risk factor and its immunological control. *Eur J Cancer* 2005;41:104–17.
113. Krone B, Grange JM. Melanoma, Darwinian medicine and the inner world. *J Cancer Res Clin Oncol* 2010;136:1787–94.
114. Hill AB. The environment and disease: association or causation? *Proc R Soc Med* 1965;58:295–300.
115. Evans AS. Causation and disease: the Henle-Koch postulates revisited. *Yale J Biol Med* 1976; 49:175–95.
116. Cegolon L, Salata C, Weiderpass E, et al. Human endogenous retroviruses and cancer prevention: evidence and prospects. *BMC Cancer* 2013;13:4.
117. Sarid R, Gao SJ. Viruses and human cancer: from detection to causality. *Cancer Lett* 2011; 305:218–27.
118. Carter PJ. Potent antibody therapeutics by design. *Nat Rev Immunol* 2006;6:343–57.
119. Kraus B, Fischer K, Buchner SM, et al. Vaccination directed against the human endogenous retrovirus-K envelope protein inhibits tumor growth in a murine model system. *PLoS One* 2013;8:e72756.
120. Fuchs NV, Loewer S, Daley GQ, et al. Human endogenous retrovirus K (HML-2) RNA and protein expression is a marker for human embryonic and induced pluripotent stem cells. *Retrovirology* 2013;10:115.
121. Guerrieri F, Belloni L, Pediconi N, et al. Molecular mechanisms of HBV-associated hepatocarcinogenesis. *Semin Liver Dis* 2013;33: 147–56.
122. Hofacre A, Fan H. Jaagsiekte sheep retrovirus biology and oncogenesis. *Viruses* 2010;2:2618–48.
123. Conti M. The selfish cell: an evolutionary defeat. Dordrecht, London: Springer, 2008.

Carcinogenesis vol.35 no.9 pp.2074–2083, 2014
doi:10.1093/carcin/bgu114
Advance Access publication May 23, 2014

Elevated HERV-K mRNA expression in PBMC is associated with a prostate cancer diagnosis particularly in older men and smokers

Tiffany A.Wallace[†], Ronan F.Downey[1,†], Caleb J.Seufert, Aaron Schetter, Tiffany H.Dorsey, Carol A.Johnson[1], Radoslav Goldman[2], Christopher A.Loffredo[2], Peisha Yan[3], Francis J.Sullivan[1,4], Francis J.Giles[1], Feng Wang-Johanning[5], Stefan Ambs[‡] and Sharon A.Glynn[1,*,‡]

Laboratory of Human Carcinogenesis, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA, [1]Prostate Cancer Institute, Biosciences Research Building, National University of Ireland Galway, Dangan, Corrib Village, Galway, Ireland, [2]Department of Oncology, Lombardi Comprehensive Cancer Center, Georgetown University, Washington, DC 20057-1465, USA, [3]Department of Pathology and Laboratory Medicine, The University of Texas Health Science Center at Houston, Houston, TX 77030, USA, [4]Department of Radiation Oncology, Galway University Hospital, Galway, Ireland and [5]Viral Oncology, Center for Cancer and Metabolism, Stanford Research Institute International, Menlo Park, CA 94025, USA

*To whom correspondence should be addressed. Tel: +353 91 49 4242; Fax: +353-91-494242;
Email: sharon.glynn@nuigalway.ie
Correspondence may also be addressed to Stefan Ambs. Tel: +1 301 496 4668; Fax: +1-301-496-0497;
Email: ambss@mail.nih.gov

Aberrant expression of subgroup k human endogenous retroviruses (HERV-K) has been observed in prostate cancer. This subgroup is unique because it encodes sequences in the human genome containing open reading frames for near intact retroviruses. We hypothesized that HERV-K reactivation could serve as a non-invasive early disease detection marker for prostate cancer. We evaluated HERV-K *gag* messenger RNA (mRNA) expression in blood samples of African-American and European-American men using a case–control design via quantitative real-time PCR. Additionally, we examined HERV-K envelope protein expression in prostate tumors by immunohistochemistry. HERV-K envelope protein was commonly upregulated in prostate tumors, but more so in tumors of African-American than European-American patients (61% versus 40%, $P < 0.01$). Examining HERV-K *gag* expression in peripheral blood mononuclear cells (PBMC) from 294 cases and 135 healthy men, we found that the abundance of HERV-K *gag* message was significantly higher in cases than controls and was associated with increased plasma interferon-γ. Men with *gag* expression in the highest quartile had >12-fold increased odds {odds ratio = 12.87 [95% confidence interval 6.3–26.25]} of being diagnosed with prostate cancer than those in the lowest quartile. Moreover, our results showed that HERV-K expression may perform better as a disease biomarker in older than younger men (whereas the sensitivity of prostate-specific antigen (PSA) testing decreases with age) and in men with a smoking history compared with never smokers. Combining non-invasive HERV-K testing with PSA testing may improve the efficacy of prostate cancer detection specifically among older men and smokers who tend to develop a more aggressive disease.

Introduction

It is estimated that prostate cancer will account for 28% of all new cancer diagnoses in USA men in 2013 with ~30 000 deaths expected, ranking it as the second leading cause of cancer death in men in the country (1). Epidemiological studies identified aging, disease family history, race/ethnicity and obesity and diet as being significant risk factors for the development of the disease (2). Other studies suggest that inflammation and infections contribute to disease development (3,4). More recently, the reactivation of endogenous retroviruses in the HERV-K family has been associated with a prostate cancer diagnosis (5–7).

Human endogenous retroviruses (HERVs) represent the remnants of ancient germline infections by exogenous retroviruses (reviewed in ref. 8) and today compose ~8% of the human genome (9). For the most part, they have become defective over time *via* the accumulation of inactivating mutations and through silencing by epigenetic mechanisms such as DNA methylation (10). The HERV-K (HML-2) subgroup (hereafter referred to as HERV-K) is unique among HERVs, in that a proportion of its constituent proviruses retain complete open reading frames for all retroviral genes (8). Furthermore, a selection of these proviruses are human specific and polymorphic (11).

The genetic structure of an intact HERV-K provirus consists of open reading frames for the retroviral genes: *gag*, *pro*, *pol* and *env*, flanked by two long terminal repeats that regulate their expression (12). HERV-K consists of two major types defined by the presence or absence of a 292 base pair deletion at the junction of the *pol* and *env* genes, which fuses their reading frames (8). Type 1 HERV-K proviruses harbor the deletion, whereas type 2 remains intact (8). Four HERV-K transcripts have been described to date: full length (*gag*) messenger RNA (mRNA), singly spliced *env* mRNA, doubly spliced *rec/np9* mRNA and the '*hel*' transcript that lacks any known function (13). High levels of HERV-K mRNA and protein have been observed in a variety of cancers, including germ cell, breast, ovarian, lymphoma and melanoma, but a causal link to any of these diseases remains to be identified (14,15). HERV-K transcripts have also been detected in prostate cancer cell lines (16) and tissues (6) and a humoral response to the HERV-K gag protein has been observed in sera from prostate cancer patients (5). Furthermore, this immune response correlated with disease progression (5), indicating that an inflammatory immune response to HERV-K, which does not eradicate the HERV-K expressing tumor, may promote disease progression.

We hypothesized that HERV-K reactivation could serve as a non-invasive early disease detection marker for prostate cancer and therefore evaluated HERV-K expression in tumor and blood samples. Because peripheral blood mononuclear cells (PBMC) can provide a suitable surrogate to an individual's health status (17) and distinct PBMC gene expression profiles have been observed in a number of non-hematological cancers (18), we analyzed HERV-K mRNA expression in PBMC from prostate cancer patients and healthy volunteers using a case–control design. We also evaluated whether HERV-K reactivation may occur differently in patients of American and European ancestry. Using this approach, we observed differences in HERV-K reactivation between African-American and European-American men and found that blood-based HERV-K expression is a candidate early detection biomarker for prostate cancer.

Materials and methods

Subject recruitment and collection of PBMC

Selection criteria for cases (n = 270) and population-based controls (n = 91) that participated in the National Cancer Institute (NCI) study: prostate cancer patients were eligible for the study when a diagnosis of prostate cancer has been made within 2 years prior to recruitment; resided in Maryland and adjacent counties in Pennsylvania, Delaware, Virginia, or District of Columbia, if they were born in the USA; were either African-American or European-American by self-report; had a working home phone number; were physically and mentally able of performing the interview; were not severely ill; spoke English fluently and were able to give informed consent and did not reside in an institution such as a prison, nursing home or shelter. Male population

Abbreviations: CTC, circulating tumor cell; INF, interferon; HERV, human endogenous retroviruses; IHC, immunohistochemistry;IL, interleukin; IRB, Institutional Review Board; mRNA, messenger RNA; NCI, National Cancer Institute; OR, odds ratio; PBMC, peripheral blood mononuclear cells; PSA, prostate-specific antigen.

[†]Joint first authors.
[‡]Joint last authors.

Published by Oxford University Press 2014.

Downloaded from http://carcin.oxfordjournals.org/ at National University of Ireland, Galway on September 19, 2014

controls were frequency matched on age and race to cases and had the same eligibility criteria with the exception that they could not have a personal history of cancer, radiation therapy or chemotherapy. Controls resided in the greater Baltimore area and adjacent counties in Maryland. Selection criteria for cases (n = 23) and hospital-based controls (n = 44) that participated in the Georgetown study: Prostate cancer patients were eligible for the study when a diagnosis of prostate cancer has been made by the attending physician, and these patients resided in the District of Columbia or its adjacent states of Maryland or Virginia. Patients were self-reported as African-American and were physically and mentally capable of providing informed consent. Male hospital controls were frequency matched on age to cases and had the same eligibility criteria as cases with the exception that they could not have a personal history of cancer. No significant difference was found between the age at diagnosis/recruitment in Georgetown Cohort versus the NIH cohort (P = 0.29) overall or when comparing African-Americans alone (P = 0.87).

Blood was drawn at time of recruitment. PBMC were collected from both prostate cancer patients (n = 294) and men without a diagnosis of the disease (n = 135). The cells were isolated from whole blood by standard ficoll–hypaque density gradient centrifugation and stored at −80°C. Men with prostate cancer were recruited between 2004 and 2008 under two Institutional Review Board (IRB)-approved protocols (NCI IRB #05-C-N021 and Georgetown University IRB #2003–013) and had a prostate cancer diagnosis within the last 2 years prior to recruitment (median time between diagnosis and recruitment = 206 days, range 0–714 days). These patients had prostate cancer at time of recruitment and came to the hospital for consultation or to seek treatment including prostatectomy, radiation therapy, or androgen ablation therapy. The subjects were recruited at four hospitals: the Veterans Affairs Medical Center and the University of Maryland Medical Center in Baltimore City, the Department of Urology at the Georgetown University Hospital and the Washington DC Veterans Affairs Medical Center. All completed an informed consent. Controls were either population-based controls (n = 91) recruited under the NCI IRB approved protocol #05-C-N021 (NCT00342771) (19), or they were men without a previous cancer diagnosis (by self-report) visiting the Georgetown University Hospital (n = 44), accompanying other people or coming for a routine checkup. The latter were recruited under the Georgetown University IRB-approved protocol #2003–013. All controls completed an informed consent. Prostate-specific antigen (PSA) test results were available for 287 of the 294 prostate cancer patients; they were not available for the controls. Both cases and controls completed interviewer-administered questionnaires but only the NCI-based study collected information on smoking history from study participants. Information on smoking was available for 359 subjects (270 cases, 89 controls) and was categorized into current, former, and current smokers, or into pack-years. A never smoker was defined as a subject who did not currently smoke and also smoked <100 cigarettes in his lifetime. A past smoker did not smoke cigarettes in the 6 months prior to enrolment. Race/ethnicity was self-reported.

RNA Isolation from PBMC and detection of HERV-K gag mRNA

Total RNA was isolated using the TRIZOL reagent according to the manufacturer's instructions. Five hundred nanograms of RNA was reverse transcribed and the complementary DNA was added to the quantitative real-time PCR assays. Previously published primers were used to amplify HERV-K *gag* transcripts (F, 5′-AGC AGG TCA GGT GCC TGT AAC ATT-3′; R, 5′-TGG TGC CGT AGG ATT AAG TCT CCT-3′) (20). In addition, primers specific for 18s were used as an internal standard reference. Data were collected using the ABI PRISM® 7500 Sequence Detection System. Normalized expression was calculated using the comparative C_T method and fold changes were derived from the $2^{-\Delta\Delta Ct}$ values (21).

Detection of HERV-K Env type I and Env type II transcripts in a subset of cases and controls

RNA was treated with TURBO DNAse (Ambion, Biosciences, Ireland) for 30 min at 37°C to eliminate genomic DNA. DNase was inactivated using 50 mM ethylene diamine tetra-acetic acid at 75°C for 10 min. First strand complementary DNA synthesis was performed on 10 ng of RNA using a Tetro cDNA Synthesis Kit (Bioline, MyBio, Ireland) primed with random hexamers. Quantitative PCR took place in a StepOne Plus real-time PCR system (Applied Biosystems, Biosciences, Ireland) together with Sensifast SYBR Hi-ROX (Bioline, MyBio, Ireland) and the following primers, 5′-CTAT TTCTTCGGACCTGTTCTTG-3′; env1 forward, 5′-GGAG ATGGTAA CACC AGTCACAT-3′; env1 reverse, 5′-GGATAACGATACCCAATGGAAAT-3′; env2 forward, 5′-CAAAATGGTGACGTCAGAAGAA-3′; env2 reverse, 5′-CAGG CATAG GGAGACTTACCAC-3′. Thermal cycling consisted of enzyme activation (95°C for 2 min), followed by 40 cycles of both denaturation and annealing/extension (95°C for 5 s and 60°C for 15 s, respectively). Subsequent melt curve analysis was carried out using the following conditions: 95°C for 15 min, 60°C for 1 min and 95°C for 15 min. Gene expression levels in all samples were normalized to an 18S rRNA reference gene (RefSeq accession number NR_003286) using the delta Ct method. Values were displayed as the mean of duplicate samples.

Measurement of cytokines in human plasma samples

Heparinized plasma was collected from prostate cancer patients and population-based controls in the NCI study and stored at −80°C. Plasma interferon-γ (INFγ), IP10, tumor necrosis factor-α, and interleukin-1β (IL-1β) concentrations were determined at a Leidos Biomedical Research Inc/NCI core facility using the human electrochemiluminescence immunoassays from Mesoscale Discovery (Gaithersburg, MD) under standardized conditions. Ultrasensitive multiplex electrochemiluminescence immunoassay plates were custom designed and were analyzed on the MesoScale Discovery 6000 instrument, following manufacturer's assay and analysis protocols.

Immunohistochemistry for HERV-K envelope expression

Immunohistochemistry (IHC) for HERV-K envelope protein expression was performed on formalin-fixed, paraffin-embedded tissue sections using standard protocols. We performed IHC on whole section tumors to examine protein localization. These tumors were obtained from patients recruited into the NCI study (IRB #05-C-N021). For IHC scoring, a tissue microarray (race/ethnicity) was obtained from the NCI Cooperative Prostate Cancer Tissue Resource. Two cores were scored per case on the tissue microarray and the average score between the two was calculated. Most tumors (304 out of 310) on the TMAs were of acinar adenocarcinoma histology, whereas the others had ductal carcinoma histology. Slides were deparaffinized, blocked with normal serum according to the VECTASTAIN® ABC protocol, and incubated with 1:200 diluted mouse monoclonal 6H5 antibody raised against the HERV-K envelope protein. This antibody was purified from a hybridoma cell supernatant and the specificity for the HERV-K envelope protein was demonstrated as described previously (22). After washing steps and incubation with a biotinylated secondary antibody, sections were incubated with VECTASTAIN® ABC reagent containing an antibiotin antibody labeled with peroxidase and stained with peroxidase substrate solution for desired stain intensity.

Cell lines

RWPE1, DU145, PC-3, 22Rv1 and CWR22 were sourced from the American Type Culture Collection (Manassas, VA) and cultured according to recommendations. In brief, CWR22 and 22Rv1 were cultured in RPMI 1640 medium with L-glutamine (Sigma #R8758) and supplemented with 10% fetal bovine serum Sigma #F7524). DU145 was cultured in minimum essential medium (1×) with Earle's (Gibco #22561-021) supplemented with 10% fetal bovine serum. PC-3 was cultured in F12 nutrient mixture (HAM) medium, with L-glutamine (Gibco #21765-029) supplemented with 10% fetal bovine serum. RWPE1 was cultured in keratinocyte medium, supplemented with epidermal growth factor and Bovine Pituitary Extract (Gibco #17005–042). All five cell lines were authenticated by LGC standards (United Kingdom) in May 2013 *via* short tandem repeat profiling and were found to be the correct cell lines.

Western blot analysis

Cells were seeded in 10 cm^3 dishes at a cell density of 1×10^6 per dish and grown for 3 days. Cells were rinsed twice with cold phosphate-buffered saline and lysed directly on the dish with cold radioimunoprecipitation assay buffer (Thermo-Scientific Pierce, Ireland #89900) supplemented with protease inhibitors (Thermo-Scientific Pierce, Ireland, #78410), scraped and spun at 14 000 *g* for 15 min at 4°C. Supernatant was collected and stored at −20°C for western blot analysis of protein expression. Extracted protein was quantified using a bicinchoninic acid assay kit. HERV-K env and HERV-K gag levels were detected through use of a primary anti-HERV-K env mouse monoclonal antibody (clone 6H5, Dr Feng Wang-Johanning) (22,23) and an anti-HERV-K gag mouse monoclonal antibody (LifeSpan Biosciences, Seattle, Washington, DC, #LS-C65287), respectively. Both antibodies were diluted 1:1000 in 5% skimmed milk reconstituted in 1× Tris-buffered saline (pH = 8) 0.1% Tween. These dilutions were added to the transfer membrane and shaken overnight at 4°C, following a 1 h room temperature blocking in 5% skimmed milk in Tris-buffered saline. Mouse monoclonal anti-β-actin antibody (Thermo-Scientific Pierce, Ireland #10624754) was used to confirm even protein loading. Secondary antibodies used were Goat Anti-Mouse horseradish peroxidase (Thermo-Scientific Pierce, Ireland #31430) and detection was imaged on the Alpha Imager imaging system.

Statistical analysis

Data analysis was performed using the Stata/SE 11 (Stata Corp, College Station, TX) and GraphPad Prism 5 (GraphPad Software, San Diego, CA) statistical software packages. All statistical tests were two sided. $P < 0.05$ was considered statistically significant. The χ^2 and Fisher's exact tests and univariate and multivariable logistic regression were used to analyze dichotomized data and to calculate odds ratios (ORs). The multivariable models were adjusted for

Downloaded from http://carcin.oxfordjournals.org/ at National University of Ireland, Galway on September 19, 2014

age at diagnosis and race/ethnicity. An interaction test was performed in the logistic regression model to assess statistical interactions between variables. The Mann–Whitney test was used to compare the differences of both plasma cytokine levels and HERV-K C_t-based expression values between groups.

Results

HERV-K gag mRNA is elevated in the PBMC of prostate cancer patients

We compared HERV-K *gag* expression in total RNA isolated from PBMC collected from men without a cancer diagnosis (n = 135) and prostate cancer patients (n = 294) using quantitative real-time PCR. The characteristics of the controls and cases are shown in Supplementary Table 1, available at *Carcinogenesis* Online. Figure 1A shows that

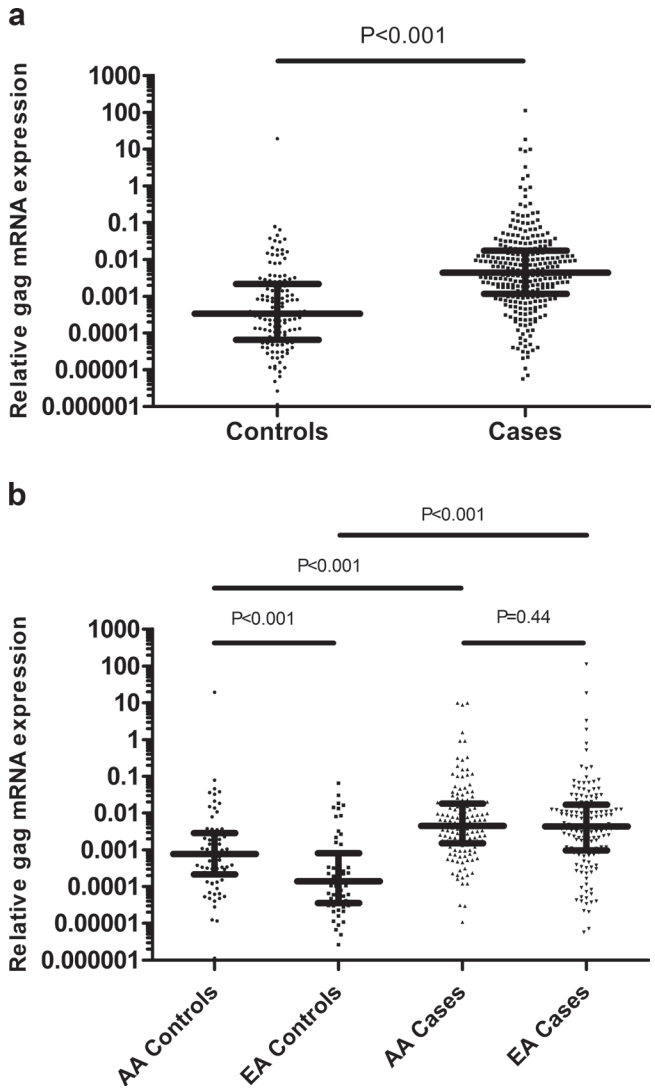


Fig. 1. Comparison of HERV-K *gag* mRNA expression in PBMC isolated from the blood of healthy controls (n = 135) and prostate cancer patients (n = 294) using previously published *gag* mRNA primer set. (**a**) HERV-K *gag* mRNA was significantly elevated in PBMC isolated from prostate cancer patients compared with the control population (Mann–Whitney test, $P < 0.001$). (**b**) HERV-K *gag* mRNA is significantly elevated in PBMC from African-American (AA) controls compared with European-American (EA) controls (Mann–Whitney test, $P < 0.001$), whereas there was no significant difference in the levels of HERV-K *gag* between African-American and European-American patients (Mann–Whitney test, $P = 0.44$). HERV-K *gag* was significantly elevated in patients compared with population controls in both African-Americans (Mann–Whitney test, $P < 0.001$) and European-Americans (Mann–Whitney test, $P < 0.001$).

although HERV-K *gag* mRNA is detectable in the PBMC of both controls and prostate cancer patients, HERV-K *gag* levels were significantly elevated in patients compared with controls. Univariate and multivariable logistic regression analyses further showed that HERV-K *gag* mRNA expression is associated with a diagnosis of prostate cancer irrespective of whether the analysis was conducted using *gag* expression as a continuous variable or comparing high versus low values of HERV-K *gag* (when using either median or quartile values of HERV-K *gag* as cutoffs; Table I). These data highlight the robustness of the association between HERV-K marker expression and prostate cancer. For example, an above median *gag* expression in PBMC was associated with a 6-fold increased odds of having a prostate cancer diagnosis in the multivariable analysis [OR = 6.02 (95% confidence interval 3.73–9.72)], compared with below median *gag* expression. In addition, our findings revealed a significant dose relationship between expression levels of HERV-K in PBMC and the likelihood of being diagnosed with prostate cancer (Table I). Notable, those men with an HERV-K *gag* expression in the highest quartile had >12-fold increased odds [OR =12.87 (95% confidence interval 6.3–26.25)] of being diagnosed with prostate cancer compared with men in the lowest quartile of PBMC HERV-K expression. After adjustment for age at diagnosis and race/ethnicity, this relationship remained, and men with HERV-K *gag* expression levels in the highest quartile had a 17.3-fold increased odds of being diagnosed with prostate cancer in the multivariable logistic regression analysis [OR =17.3 (95% confidence interval 8.1–37.0)]. Next, we examined the effects of treatment on HERV-K expression in the PBMC of prostate cancer patients but found no significant difference in expression between men who had received prior treatment and those without it (Supplementary Figure 1, available at *Carcinogenesis* Online). Additional analyses showed that HERV-K *gag* expression in the PBMC of prostate cancer patients did not correlate with PSA levels at diagnosis (Spearman's rho = –0.01, P = 0.92). Thus, they appear to be independent markers. We also examined if the time from diagnosis to blood draw had an effect on HERV-K *gag* expression in patients but did not find a significant relationship between these two variables (Supplementary Figure 2, available at *Carcinogenesis* Online). Likewise, HERV-K gag expression did not differ significantly depending on Gleason score at diagnosis (Supplementary Figure 3, available at *Carcinogenesis* Online). In contrast, an analysis by stage showed that HERV-K *gag* expression was significantly higher in stage II patients than in stage I (P = 0.034; Supplementary Figure 4, available at *Carcinogenesis* Online), whereas no difference was found between patients with low-stage disease (stages I or II) and the few patients with high-stage disease (stages III or IV).

HERV-K gag mRNA is elevated in African-American controls compared with European-American controls

African-American men are at an increased risk of developing prostate cancer and are also at an increased risk of diagnosis with advanced disease, compared with European-American men. We therefore examined whether race/ethnicity influences the association of HERV-K *gag* mRNA with prostate cancer. The analysis showed that *gag* expression is associated with a prostate cancer diagnosis in both population groups (Supplementary Table 2, available at *Carcinogenesis* Online), but as shown in Figure 1B, African-American population controls expressed significantly more HERV-K *gag* mRNA in their PBMC than European-American population controls. Levels of HERV-K *gag* message were not different by race/ethnicity in the case population, suggesting that aberrant HERV-K *gag* expression in PBMC of prostate cancer patients is similar irrespective of race/ethnicity.

Association of HERV-K gag mRNA with a prostate cancer diagnosis increases with age and is most robust in older men

The sensitivity of PSA testing decreases with age, partially due to an increase in the prevalence of benign prostate hyperplasia, which also elevates PSA. To determine whether HERV-K *gag* expression is predictive of prostate cancer across all age group, we performed a

Downloaded from http://carcin.oxfordjournals.org/ at National University of Ireland, Galway on September 19, 2014

Table I. Logistic regression analysis of the association between HERV-K *gag* mRNA levels in PBMC and a prostate cancer diagnosis

	Univariate analysis[a]			Multivariable analysis I[b]			Multivariable analysis II[c]			Multivariable analysis III[d]		
	OR (95% confidence interval)	*P* value	*n*	OR (95% confidence interval)	*P* value	*n*	OR (95% confidence interval)	*P* value	*n*	OR (95% confidence interval)	*P* value	w
Logistic regression with HERV-K *gag* coded as a continuous variable (per C_t value)												
Gag mRNA	1.33 (1.24–1.43)	<0.0001	429	1.36 (1.26–1.47)	<0.0001	427	1.49 (1.35–1.65)	<0.0001	358	1.67 (1.42–1.96)	<0.0001	148
Logistic regression comparing high versus low HERV-K *gag* expression[e]												
Low gag	1		1			1			1			
High gag	5.85 (3.67–9.36)	<0.0001	429	6.02 (3.73–9.72)	<0.0001	427	10.35 (5.43–19.8)	<0.0001	358	24.52 (8.61–69.8)	<0.0001	148
Logistic regression to evaluate a dose response effect after stratification of HERV-K (HML-2) gag expression into quartiles[f]												
1st quartile	1		48	1		48	1		42	1		30
2nd quartile	1.38 (0.61–3.11)	0.433	57	1.58 (0.69–3.62)	0.280	57	1.53 (0.61–3.81)	0.363	43	2.25 (0.50–10.1)	0.288	20
3rd quartile	4.14 (1.98–8.67)	<0.0001	98	5.90 (2.69–13.0)	<0.0001	96	7.15 (3.01–16.9)	<0.0001	80	7.80 (2.05–29.7)	0.003	32
4th quartile	12.87 (6.30–26.3)	<0.0001	226	17.3 (8.07–37.0)	<0.0001	226	36.4 (14.7–90.2)	<0.0001	193	94.9 (22.0–408)	<0.0001	66
	P_{trend}	<0.0001			P_{trend}	<0.0001		P_{trend}	<0.0001		P_{trend}	<0.0001

[a]294 cases and 135 controls.
[b]Adjusted for age at diagnosis and race/ethnicity.
[c]Adjusted for age at diagnosis, race/ethnicity and smoking status.
[d]Adjusted for age at diagnosis, race/ethnicity, smoking status and plasma IFNγ and IP10.
[e]HERV-K *gag* levels were dichotomized using the median expression in the control population.
[f]HERV-K *gag* levels were divided into quartiles based on HERV-K *gag* quartile distribution in the control population.

stratified analysis comparing the association of HERV-K *gag* according to age groups. We stratified men into three similar-sized groups, ages 41–59 (n = 128), ages 60–69 (n = 179) and ages ±70 (n = 120). The findings from the univariate logistic regression analysis in Table II show that although HERV-K *gag* is predictive of prostate cancer across all ages, the strength of the association increased with age. HERV-K *gag* was most predictive in men aged ±70, indicating that combining HERV-K testing with PSA testing may improve the efficacy of prostate cancer detection in these older men. This remained true in the multivariable analysis after adjusting for race/ethnicity. To determine if there is a modifying effect of age at diagnosis on the association between HERV-K *gag* expression and prostate cancer, we performed a statistical interaction test, first on the three individual age categories with HERV-K gag quartile expression but found that the $P_{interaction}$ was not significant (P values ranged between 0.40 and 0.96). We then dichotomized age into high/low with the median as cutoff and assessed the modifying effect of the age variable on the association between HERV-K *gag* expression and prostate cancer within each HERV-K expression quartile (quartiles as described in Table II). The test showed that age at diagnosis has a statistically significant modifying effect on the association of HERV-K *gag* with prostate cancer within each HERV-K gag quartile expression (Q2 $P_{interaction}$ = 0.049; Q3 $P_{interaction}$ = 0.019; Q4 $P_{interaction}$ = 0.017). These results are suggestive, but not definitive, of a stronger association of HERV-K with prostate cancer with increasing age.

Association of HERV-K gag mRNA expression with the risk of prostate cancer is modified by smoking status

The smoking status was available for 270 cases and 89 population-based controls in this cohort (total n = 359), categorized as current smoker, former smoker or never smoker. Moreover, pack-years smoked information was available for 358 of them. Although smoking status or pack-years did not directly correlate with the level of HERV-K *gag* expression in the blood samples, smoking status was found to modify the association of HERV-K *gag* expression with a diagnosis of prostate cancer. As shown in Table III, the strength of the association between HERV-K *gag* and disease was found to be highest in current smokers, whereas lowest in never smokers. When stratified by pack-years of smoking, there was an increase in the strength of the association between HERV-K *gag* expression and disease with an increase in pack-years of tobacco exposure (Supplementary Table 3, available at *Carcinogenesis* Online). Because of these modifying effects of smoking on the association between HERV-K *gag* expression and disease, we performed an interaction analysis. A test for interaction between smoking status and HERV-K *gag* expression in prostate cancer development with HERV_K *gag* expression categorized as high/low (dichotomized at the median) indicated no statistically significant interaction between HERV-K *gag* and former smoker status ($P_{interaction}$ = 0.40) or HERV-K and current smoker status ($P_{interaction}$ = 0.26). We also performed an interaction test between pack-years exposure (dichotomized at the median) and *gag* expression levels (dichotomized at the median). The detected interaction (or modifying effect of tobacco exposure) did not reach statistical significance ($P_{interaction}$ = 0.096). Therefore, although HERV-K *gag* in PBMC is more closely associated with prostate cancer among heavy smokers than light or never smokers, there may not be a direct interaction between HERV-K expression and smoking exposure on prostate cancer risk. A larger and more appropriately powered study would be required to test this and show that such an interaction exists.

Prostate cancer patients with elevated HERV-K gag show increased serum expression of viral response IFNγ and IP10 expressions

To determine whether elevation of HERV-K *gag* mRNA in PBMC from prostate cancer results in the secretion of viral response INFs and inflammatory cytokines, we measured the expression of IFNγ, IP10, tumor necrosis factor-α and IL-1β in plasma of a subset of prostate cancer cases and controls. We found that the plasma levels of the viral response cytokines, IFNγ (Figure 2A) and its downstream

Downloaded from http://carcin.oxfordjournals.org/ at National University of Ireland, Galway on September 19, 2014

Table II. Age-stratified logistic regression analysis of the association between HERV-K *gag* mRNA levels and prostate cancer

	Univariate analysis			Multivariable analysis[a]			
	OR (95% confidence interval)	*P* value	*N*	OR	95% confidence interval	*P* value	*n*
Men aged 41–59							
1st quartile[b]	1		12	1			12
2nd quartile	1.12 (0.25–4.91)	0.880	18	1.22 (0.27–5.60)		0.795	18
3rd quartile	3.15 (0.76–13.00)	0.113	26	4.67 (1.02–21.38)		0.047	26
4th quartile	9.8 (2.56–37.55)	0.001	72	11.96 (2.92–48.92)		0.001	72
	P_{trend}	<0.0001		P_{trend}		<0.0001	
Men aged 60–69							
1st quartile	1		23	1			23
2nd quartile	1.07 (0.32–3.63)	0.912	22	1.54 (0.42–5.54)		0.513	22
3rd quartile	4.15 (1.43–12.04)	0.009	45	5.53 (1.79–17.09)		0.003	45
4th quartile	13.29 (4.58–38.57)	<0.0001	89	19.52 (6.13–62.12)		0.001	89
	P_{trend}	<0.0001		P_{trend}		<0.0001	
Men aged ±70							
1st quartile	1		13	1			13
2nd quartile	3.0 (0.40–18.24)	0.233	17	2.96 (0.48–18.06)		0.239	17
3rd quartile	8.25 (1.49–45.42)	0.015	25	9.07 (1.59–51.71)		0.013	25
4th quartile	22.00 (4.33–111.7)	<0.0001	65	23.94 (4.57–125.4)		<0.0001	65
	P_{trend}	<0.0001		P_{trend}		<0.0001	

[a]Adjusted for race/ethnicity.
[b]HERV-K *gag* levels were divided into quartiles based on HERV-K *gag* quartile distributions in the control population.

Table III. Logistic regression analysis of the association between HERV-K *gag* mRNA levels and prostate cancer by smoking status

	Univariate analysis			Multivariable analysis[a]		
	OR (95% confidence interval)	*P* value	*N*	OR (95% confidence interval)	*P* value	*n*
Never smokers						
Gag mRNA (continuous)	1.41 (1.21–1.63)	<0.0001	121	1.43 (1.22–1.67)	<0.0001	121
Gag mRNA[b] (median)	8.07 (3.03–21.5)	<0.0001	121	9.38 (3.33–26.4)	<0.0001	121
Former smokers						
Gag mRNA (continuous)	1.44 (1.27–1.63)	<0.0001	173	1.48 (1.29–1.69)	<0.0001	173
Gag mRNA (median)	13.8 (6.27–30.4)	<0.0001	173	17.4 (7.29–41.8)	<0.0001	173
Current smokers						
Gag mRNA (continuous)	1.93 (1.31–2.87)	0.001	64	2.35 (1.33–4.16)	0.003	64
Gag mRNA (median)	25.5 (4.56–142)	0.001	64	30.2 (4.40–208)	0.001	64

[a]Adjusted for age at diagnosis and race/ethnicity.
[b]HERV-K *gag* levels were dichotomized using the median expression in the control population.

effector IP10 (Figure 2B) were significantly elevated in the case population with high levels of HERV-K *gag* mRNA expression. We did not find the same elevated levels of these cytokines in cases with low *gag* expression, or among the control population. Neither tumor necrosis factor-α (Figure 2C) or IL-1β (Figure 2D) was elevated in patients with high levels of HERV-K *gag* mRNA. Multivariable logistic regression (Model 3) in Table I demonstrates that after adjusting for IFNγ and IP10 levels, HERV-K remains significantly associated with prostate cancer. Additionally we examined the impact of IFNγ or IP10 on the association of HERV-K gag mRNA with prostate cancer (Supplementary Table 4, available at *Carcinogenesis* Online). The univariate logistic regression showed that high levels of HERV-K gag mRNA are more strongly associated with prostate cancer in patients with high IFNγ or IP10, compared with those with low IFNγ or IP10. Notable, IFNγ or IP10 level above the median in the absence of high levels of HERV-K gag was not found to be associated with prostate cancer. These associations upheld in the multivariable analysis after adjusting for age at diagnosis, race/ethnicity and smoking status.

Detection of HERV-K env type I and type II mRNA transcripts in a subset of the PBMC

We designed primers specific to the env of HERV-K type I and HERV-K type II viruses. Type I and type II are distinguished by a 292 bp deletion in the env gene (type I). Supplementary Figure 5, available at *Carcinogenesis* Online, shows that both type I and type II env mRNA were elevated in the PBMC of the cancer patients, indicating that the HERV-K reactivation may arise from multiple HERV-K loci in prostate cancer patients.

HERV-K env protein is expressed in prostate tumors

We used IHC to assess whether prostate tumors express the HERV-K env protein using whole tissue sections to examine protein localization and a tissue microarray for scoring. Consistent with our findings analyzing HERV-K expression levels in PBMC, expression of the HERV-K env protein in tumors varied considerably among patients. Figure 3A–D shows that HERV-K env protein expression is cytoplasmic and membrane located, and also localizing toward the lumen of the gland. Env protein expression levels were significantly higher in African-American than European-American patients (Figure 3E), and 61% of the African-American patients presented with tumors that had high env expression (above median), whereas only 40% of the tumors from European-American patients fell into the same category ($P < 0.001$). In contrast, we did not detect aberrant HERV-K env expression in prostate tissues from patients with benign prostate hyperplasia (data not shown). Western blot analysis of cell extracts further corroborated the presence of both HERV-K gag and HERV-K envelope protein in human prostate cancer cells, whereas expression was low to absent in the non-tumorigenic RWPE1 cells (Figure 3F).

Downloaded from http://carcin.oxfordjournals.org/ at National University of Ireland, Galway on September 19, 2014

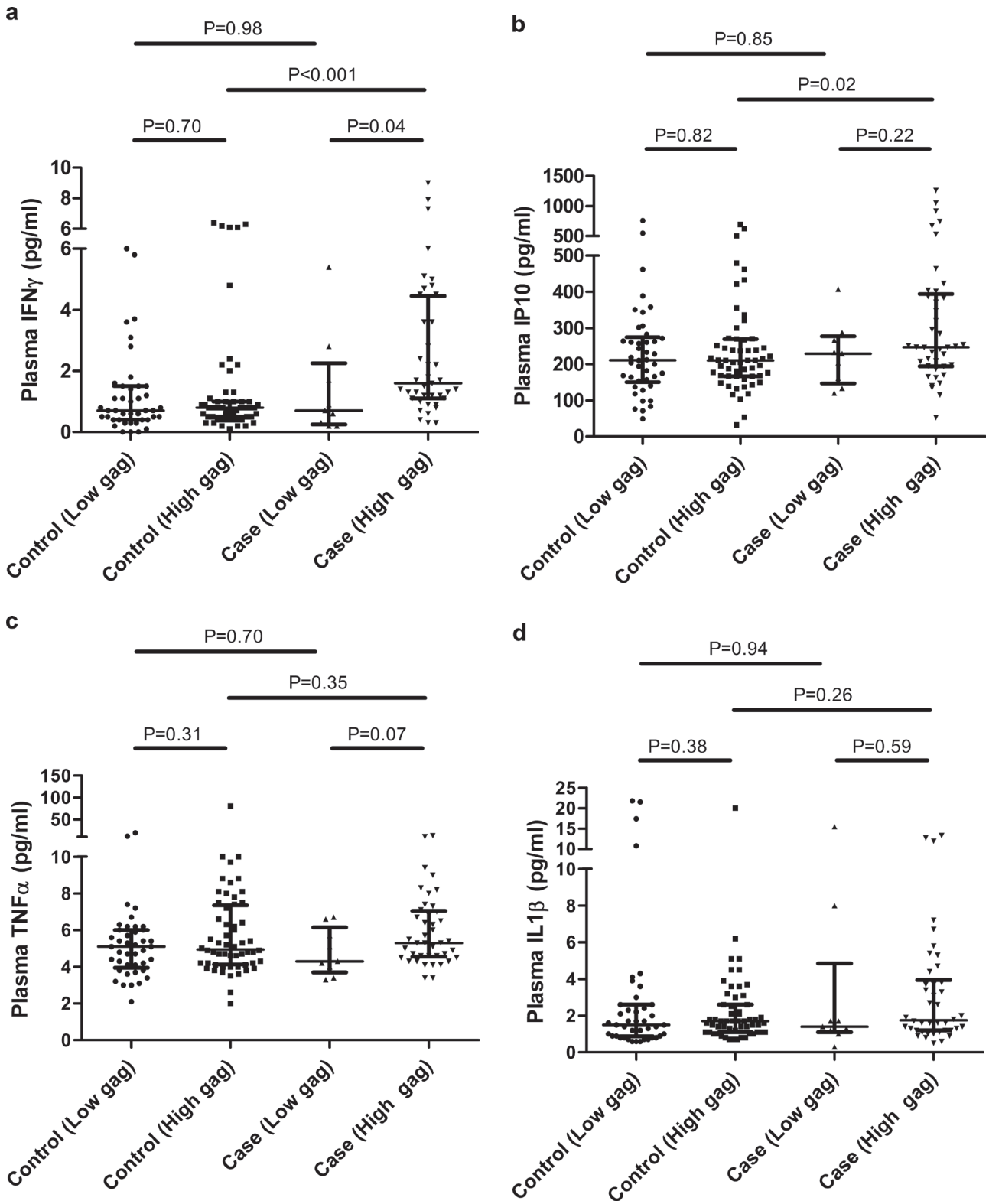


Fig. 2. Association between high HERV-K *gag* mRNA expression and elevated plasma IFNγ and IP10 levels. Plasma IFNγ (**a**) and its downstream mediator IP10 (**b**) were significantly increased in patients with high *gag* expression compared with controls with high *gag* expression (Mann–Whitney test IFNγ, $P < 0.001$; Mann–Whitney test IP10, $P = 0.02$). Additionally, cases with high *gag* expression had significantly higher levels of plasma IFNγ compared with cases with low *gag* expression. No significant association of *gag* expression with tumor necrosis factor-α (**c**) or IL-1β (**d**) was observed.

Discussion

Our study makes the novel observation that HERV-K mRNA expression is aberrantly increased in PBMC from prostate cancer patients compared with healthy male controls. Thus, the evaluation of blood-based HERV-K expression may serve an early disease detection biomarker in prostate cancer. Moreover, the expression as of HERV-K as an early disease biomarker may perform better in older than younger

Downloaded from http://carcin.oxfordjournals.org/ at National University of Ireland, Galway on September 19, 2014

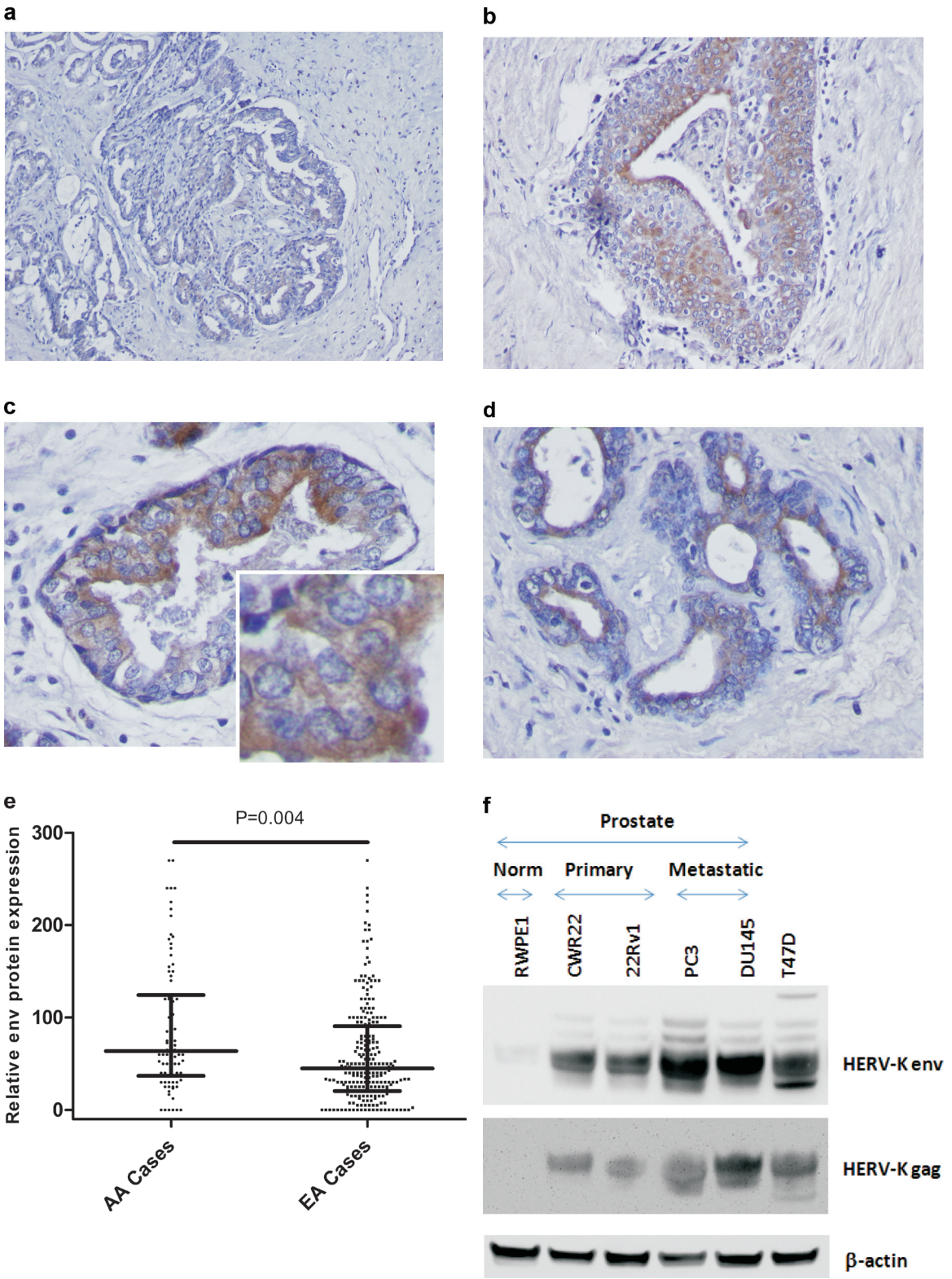


Fig. 3. HERV-K expression in human prostate adenocarcinomas. Shown is IHC analysis of four invasive adenocarcinomas for expression of HERV-K envelope (env) protein using the monoclonal anti-envelope antibody, 6H5. (**a**) Scattered positivity for HERV-K env expression in the tumor epithelium, showing a low to moderate antigen expression as indicated by the brown chromogen deposits. (**b**–**d**) Locally intensive staining for env expression in the tumor epithelium. Staining shows a cytosolic to membrane distribution with a more intensive staining of cancer cells toward the luminal side of the cancerous gland (c and d). a and b: Magnification: ×100. c and d: Magnification: ×400. Inset: higher resolution image for the env protein-positive tumor epithelium. Counterstain: Hematoxylin. (**e**) Immunostaining for env protein in human prostate tumors using a tissue microarray that included tumors from both African-American (n = 105) and European-American patients (n = 272). On average, tumors from African-American patients showed a higher expression of the HERV-K env protein than tumors from European-Americans (Student's t-test, $P < 0.001$). (**f**) HERV-K *gag* and env protein expression were detected in human prostate cancer cell lines (CWR22, 22Rv1, PC-3 and DU145) but were not detected in the non-tumorigenic human prostate cell line, RWPE1. An extract from the HERV-K-positive T47D human breast cancer cell line were included as a positive control.

Downloaded from http://carcin.oxfordjournals.org/ at National University of Ireland, Galway on September 19, 2014

men, whereas the sensitivity of PSA testing decreases with age. Thus, combining HERV-K testing with PSA testing may improve the efficacy of prostate cancer detection in this age group. Other studies have observed the presence of autoantibodies to HERV-K gag in sera from prostate cancer patients and described the aberrant expression of HERV-K mRNA and proteins in prostate cancer tissues (5,6,24). Hence, our observation of aberrantly increased blood-based expression of HERV-K in prostate cancer patients is consistent with these previous findings.

PSA testing is the gold standard for prostate cancer screening. The traditional cutoff for an abnormal PSA level is 4.0 ng/ml. One of the major problems is that PSA has poor discriminating ability in men with symptomatic benign prostate hyperplasia versus those with prostate cancer (25). Additionally, PSA levels rise as men grow older, which can lead to increased false positive PSA tests and unnecessary biopsies (26). Consistent with these data, PSA levels displayed a slight albeit non-significant increase with age at diagnosis in our patient cohort (Spearman's rho = 0.09, $P = 0.13$), whereas HERV-K *gag* expression did not increase with either the age at diagnosis for the cases (Spearman's rho = –0.01, $P = 0.91$) or the age of recruitment for the controls (Spearman's rho = –0.04, $P = 0.61$). Currently, the American Urological Association only recommends the use of PSA screening in asymptomatic men aged between 55 and 69 (27). Our data suggest that blood-based HERV-K is a candidate biomarker for the detection of prostate cancer, potentially with a focus on older men. Future research must determine its predictive value in conjunction with PSA testing.

Incidence and mortality rates of prostate cancer are significantly higher in men of African ancestry compared with men from other population groups in the USA, the Caribbean and the United Kingdom (28). This cancer health disparity may relate to unknown causative factors that influence disease pathology in men of African ancestry and induce a more aggressive disease among them. Therefore, we examined HERV-K expression in tumors and blood samples from both African-American and European-American patients. Indeed, we found that HERV-K expression in prostate tumors is significantly higher in the African-American patients, compared with European-American patients, based on the immunohistochemical analysis of HERV-K envelope protein expression in these tissues. However, we did not find that HERV-K message levels are significantly different in PBMC from these two patient groups. Accordingly, HERV-K as biomarker should perform similarly among these patients. However, we found that baseline levels of HERV-K mRNA in PBMC of healthy controls were greater in African-Americans than European-Americans. Currently, we do not know why African-American men may express higher baseline levels of HERV-K, nor what effects, if any, there may be. However, it is known that retroviral reactivation is induced by both hormone exposure and stress signaling (13,29,30), and these signaling pathways could become more commonly activated in African-American men. Additionally, Macfarlane and Simmonds reported that the frequency of allelic variation in the various HERV-K germline loci is greater in African populations than populations from Europe and Asia (31). Furthermore, others reported significantly higher insertion frequencies of HERV-K113 (21%) and K115 (35%) in African-Americans compared with European-Americans (K113 9% and K115 6%) within the USA (32). Because African-Americans may have inherited a greater number of polymorphic HERV-K loci than European-Americans, this would increase the number of HERV-K loci that are potentially transcribed, leading to a higher baseline expression of HERV-K encoded genes from these loci.

Cigarette smoking is not associated with early disease development in prostate cancer (33,34), but it is associated with an increased risk of fatal prostate cancer (33,35,36). It is also associated with increased risk of biochemical recurrence and metastasis (37). Gabriel *et al*. (38) showed previously that exposure of both normal human dermal fibroblasts and benign human uroepithelium to urine from current smokers increased the transcription of HERV-K (HML-6), indicating that HERV-K may be induced by tobacco metabolites; however, the authors did not specifically test the HERV-K family. Because of these observations, we decided to evaluate the influence of smoking on the application of blood-based HERV-K expression as a disease biomarker. We had smoking status and pack-year information for a large subset of subjects in our case–control study and therefore assessed the impact of smoking status on the association of HERV-K with prostate cancer. This analysis led to the finding that HERV-K was more predictive of prostate cancer in current smokers, and the association increased in strength with increased pack-years smoked. The underlying mechanism for this modifying effect of smoking on the relationship between HERV-K and prostate cancer remains unclear, and we could not demonstrate a statistically significant interaction between smoking, HERV-K and risk of prostate cancer in this study. Yet, we think our observation could be of particular significance because those patients who are current smokers tend to develop a more aggressive disease than other patients, as already pointed out by us, and improved disease detection at an early stage for this patient group may have a significant impact in reducing disease mortality.

We also observed that elevated HERV-K mRNA in the PBMC of prostate cancer patients was associated with increased serum IFNγ and its effector IP10. It has been previously reported that patients with HERV-K-positive tumors exhibit a humoral response to HERV-K. Rakoff-Nahoum *et al*. (39) identified HERV-K-specific T cells in seminoma patients, which display elevated IFNγ secretion in response to HERV-K gag peptides. Wang-Johanning *et al*. (40) also showed that breast cancer patients mount anti-HERV-K responses, including production of anti-HERV-K(HML-2) env immunoglobulin G, production of IFNγ and a T-helper cell cytokine response signature including increased production of IL-2, IL-6, IL-8 and IP10 during *in vitro* stimulation of breast cancer PBMC with HERV-K antigen. In this context, the finding of an INF response in patients with an elevated blood-based HERV-K message is consistent with a humoral response to HERV-K that was observed by others, indicating aberrant HERV-K expression in prostate cancer patients not only affects tumor immunobiology but also triggers a systemic antiviral response. The relationship of this antiviral response with disease outcome is still largely unknown, but Reis *et al*. (5) observed that 6.8% of patients with prostate cancer had serum antibodies to the gag-HERV-K protein encoded by the ch22q11.23 locus, and these patients tended to have a more aggressive disease and a higher disease mortality. We previously reported differences in the prostate tumor immunobiology between African-American and European-American men (41). Prostate tumors from African-American men were characterized by the activation of immune response and host defense pathways. Johnston *et al*. (42) reported that activated macrophages display increased expression of HERV-K, HERV-W and HERV-H mRNA, indicating that HERV reactivation may occur as a consequence of an elevated immune activity. A frequent activation of immune response pathways in tumors from African-American men may therefore contribute to the increased expression of HERV-K env protein in these tumors that we observed in this study.

One of the limitations of this study is our inability to tell where exactly the elevated HERV-K expression arises from in our PBMC from prostate cancer. HERV-K has previously been reported as being expressed at basal levels in healthy PBMC, with aberrant expression occurring in leukemia cells (43–45). However, circulating tumor cells (CTCs) are detectable in the blood of prostate cancer patients (~2–8 per 7.5 ml of blood), with higher levels of CTCs correlating with bone and visceral metastasis (46). Prostate CTCs reflect the biology of the cancer with the CTCs gene signatures switching from AR-on to AR-off within the first month of androgen deprivation therapy, indicating their sensitivity as indicators of tumor response to treatment (47). When isolating PBMC from blood, CTCs may also be pulled down in the PBMC fraction (48). Therefore, PBMC with high levels of HERV-K in prostate cancer patients may potentially reflect HERV-K-positive CTCs. Another potential source of HERV-K is from exosomes secreted by tumors. Exosomes are nanoscale membrane vesicles that are secreted from cells and are thought to be important intercellular communicators or, in a cancer setting, drivers of metastatic spread (49). A recent study has now implicated HERVs in this

Downloaded from http://carcin.oxfordjournals.org/ at National University of Ireland, Galway on September 19, 2014

process, with the finding that HERV-K mRNA is selectively packaged into tumor exosomes and that this genetic material can be transferred to normal cells (50). Exosomes can be found abundantly in the blood of patients and may contribute to tumor dissemination (51). Future work will focus on investigating these options and identifying the specific HERV-K loci responsible for elevated HERV-K in the PBMC of prostate cancer patients and will also evaluate the relative contribution of CTCs and exosomes to the HERV-K expression signal in PBMC from cancer patients.

In summary, we made the novel observation that blood-based HERV-K expression is a candidate early detection biomarker for prostate cancer that may specifically improve disease detection among older men. Furthermore, we obtained evidence that this test would perform similarly in African-Americans and European-Americans. Lastly, aberrant HERV-K expression seems to occur more frequently in both healthy African-Americans and African-Americans with prostate cancer than in their European-American counterparts. Although intriguing, these data need to be further studied.

Supplementary material

Supplementary Tables 1–4 and Figures 1–5 can be found at http://carcin.oxfordjournals.org/

Funding

The Intramural Research Program of the National Institutes of Health, National Cancer Institute, Center for Cancer Research, USA; the Galway University Foundation, Ireland (RNR1008 to F.J.S. and S.A.G.); the Health Research Board of Ireland Clinical Research Facility, Galway, Ireland (RSU004 to R.F.D. and F.J.G.); the Department of Defense Breast Cancer Research Program, USA (BC113114 to F.W.J.); the National Cancer Institute, National Institutes of Health, USA (U01CA171146 to R.G.); the Department of Defense Prostate Cancer Research Program, USA (PC081609 to R.G.); the Breast Cancer Campaign, UK (2013MayPR019 to S.A.G. and F.J.S.); the Irish Cancer Society, Ireland (PCT13MCD to S.A.G. and F.J.S.).

Conflict of Interest Statement: None declared.

Acknowledgement

We would like to thank Dr Aideen Ryan (National University of Ireland Galway) for her help with preparation of images for publication.

References

1. Siegel,R. *et al.* (2013) Cancer statistics, 2013. *CA. Cancer J. Clin*., **63**, 11–30.
2. Grönberg,H. (2003) Prostate cancer epidemiology. *Lancet*, **361**, 859–864.
3. De Marzo,A.M. *et al.* (2007) Inflammation in prostate carcinogenesis. *Nat. Rev. Cancer*, **7**, 256–269.
4. Sutcliffe,S. *et al.* (2008) Inflammation and prostate cancer: a focus on infections. *Curr. Urol. Rep*., **9**, 243–249.
5. Reis,B.S. *et al.* (2013) Prostate cancer progression correlates with increased humoral immune response to a human endogenous retrovirus GAG protein. *Clin. Cancer Res*., **19**, 6112–6125.
6. Goering,W. *et al.* (2011) Selective changes of retroelement expression in human prostate cancer. *Carcinogenesis*, **32**, 1484–1492.
7. Perot,P., *et al.* (2013) Microarray-based Identification of Individual HERV Loci Expression: Application to Biomarker Discovery in Prostate Cancer. *J Vis Exp*, e50713.
8. Bannert,N. *et al.* (2004) Retroelements and the human genome: new perspectives on an old relation. *Proc. Natl Acad. Sci. USA*, **101**(suppl 2), 14572–14579.
9. Lander,E.S. *et al.*; International Human Genome Sequencing Consortium. (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860–921.
10. Schulz,W.A. *et al.* (2006) Methylation of endogenous human retroelements in health and disease. *Curr. Top. Microbiol. Immunol*., **310**, 211–250.
11. Moyes,D. *et al.* (2007) Insertional polymorphisms: a new lease of life for endogenous retroviruses in human disease. *Trends Genet*., **23**, 326–333.
12. Bannert,N. *et al.* (2006) The evolutionary dynamics of human endogenous retroviral families. *Annu. Rev. Genomics Hum. Genet*., **7**, 149–173.
13. Hohn,O. *et al.* (2013) HERV-K(HML-2), the Best Preserved Family of HERVs: Endogenization, Expression, and Implications in Health and Disease. *Front. Oncol*., **3**, 246.
14. Ruprecht,K. *et al.* (2008) Endogenous retroviruses and cancer. *Cell. Mol. Life Sci*., **65**, 3366–3382.
15. Downey,R.F., *et al.* (2014) Human endogenous retrovirus K and cancer: innocent bystander or tumorigenic accomplice? *International Journal of Cancer*. doi:10.1002/ijc.29003. [Epub ahead of print].
16. Agoni,L. *et al.* (2013) Detection of Human Endogenous Retrovirus K (HERV-K) Transcripts in Human Prostate Cancer Cell Lines. *Front. Oncol*., **3**, 180.
17. Liew,C.C. *et al.* (2006) The peripheral blood transcriptome dynamically reflects system wide biology: a potential diagnostic tool. *J. Lab. Clin. Med*., **147**, 126–132.
18. Mohr,S. *et al.* (2007) The peripheral-blood transcriptome: new insights into disease and risk assessment. *Trends Mol. Med*., **13**, 422–432.
19. Hudson,R.S. *et al.* (2012) MicroRNA-1 is a candidate tumor suppressor and prognostic marker in human prostate cancer. *Nucleic Acids Res*., **40**, 3689–3703.
20. Contreras-Galindo,R. *et al.* (2008) Human endogenous retrovirus K (HML-2) elements in the plasma of people with lymphoma and breast cancer. *J. Virol*., **82**, 9329–9336.
21. Bookout,A.L. *et al.* (2003) Quantitative real-time PCR protocol for analysis of nuclear receptor signaling pathways. *Nucl. Recept. Signal*., **1**, e012.
22. Wang-Johanning,F. *et al.* (2012) Immunotherapeutic potential of anti-human endogenous retrovirus-K envelope protein antibodies in targeting breast tumors. *J. Natl Cancer Inst*., **104**, 189–210.
23. Wang-Johanning,F. *et al.* (2014) Human endogenous retrovirus type K antibodies and mRNA as serum biomarkers of early-stage breast cancer. *Int. J. Cancer*, **134**, 587–595.
24. Ishida,T. *et al.* (2008) Identification of the HERV-K gag antigen in prostate cancer by SEREX using autologous patient serum and its immunogenicity. *Cancer Immun*., **8**, 15.
25. You,J. *et al.* (2010) Innovative biomarkers for prostate cancer early diagnosis and progression. *Crit. Rev. Oncol. Hematol*., **73**, 10–22.
26. Gulati,R. *et al.* (2011) What if i don't treat my PSA-detected prostate cancer? Answers from three natural history models. *Cancer Epidemiol. Biomarkers Prev*., **20**, 740–750.
27. Carter,H.B. (2013) American Urological Association (AUA) guideline on prostate cancer detection: process and rationale. *BJU Int*., **112**, 543–547.
28. Martin,D.N. *et al.* (2013) Biological determinants of health disparities in prostate cancer. *Curr. Opin. Oncol*., **25**, 235–241.
29. Wang-Johanning,F. *et al.* (2003) Quantitation of HERV-K env gene expression and splicing in human breast cancer. *Oncogene*, **22**, 1528–1535.
30. Serafino,A. *et al.* (2009) The activation of human endogenous retrovirus K (HERV-K) is implicated in melanoma cell malignant transformation. *Exp. Cell Res*., **315**, 849–862.
31. Macfarlane,C. *et al.* (2004) Allelic variation of HERV-K(HML-2) endogenous retroviral elements in human populations. *J. Mol. Evol*., **59**, 642–656.
32. Jha,A.R. *et al.* (2009) Cross-sectional dating of novel haplotypes of HERV-K 113 and HERV-K 115 indicate these proviruses originated in Africa before Homo sapiens. *Mol. Biol. Evol*., **26**, 2617–2626.
33. Giovannucci,E. *et al.* (1999) Smoking and risk of total and fatal prostate cancer in United States health professionals. *Cancer Epidemiol. Biomarkers Prev*., **8**(Pt 1), 277–282.
34. Hickey,K. *et al.* (2001) Smoking and prostate cancer. *Epidemiol. Rev*., **23**, 115–125.
35. Huncharek,M. *et al.* (2010) Smoking as a risk factor for prostate cancer: a meta-analysis of 24 prospective cohort studies. *Am. J. Public Health*, **100**, 693–701.
36. Kenfield,S.A. *et al.* (2011) Smoking and prostate cancer survival and recurrence. *JAMA*, **305**, 2548–2555.
37. Moreira,D.M., *et al.* (2013) Cigarette smoking is associated with an increased risk of biochemical disease recurrence, metastasis, castration-resistant prostate cancer, and mortality after radical prostatectomy: Results from the SEARCH database. *Cancer*, **120**, 197–204.
38. Gabriel,U. *et al.* (2010) Smoking increases transcription of human endogenous retroviruses in a newly established *in vitro* cell model and in normal urothelium. *AIDS Res. Hum. Retroviruses*, **26**, 883–888.

Downloaded from http://carcin.oxfordjournals.org/ at National University of Ireland, Galway on September 19, 2014

39. Rakoff-Nahoum,S. *et al.* (2006) Detection of T lymphocytes specific for human endogenous retrovirus K (HERV-K) in patients with seminoma. *AIDS Res. Hum. Retroviruses*, **22**, 52–56.
40. Wang-Johanning,F. *et al.* (2008) Human endogenous retrovirus K triggers an antigen-specific immune response in breast cancer patients. *Cancer Res.*, **68**, 5869–5877.
41. Wallace,T.A. *et al.* (2008) Tumor immunobiological differences in prostate cancer between African-American and European-American men. *Cancer Res.*, **68**, 927–936.
42. Johnston,J.B. *et al.* (2001) Monocyte activation and differentiation augment human endogenous retrovirus expression: implications for inflammatory brain diseases. *Ann. Neurol.*, **50**, 434–442.
43. Medstrand,P. *et al.* (1992) Expression of human endogenous retroviral sequences in peripheral blood mononuclear cells of healthy individuals. *J. Gen. Virol.*, **73**(Pt 9), 2463–2466.
44. Brodsky,I. *et al.* (1993) Expression of human endogenous retrovirus (HERV-K) in chronic myeloid leukemia. *Leuk. Lymphoma*, **11**(suppl 1), 119–123.
45. Depil,S. *et al.* (2002) Expression of a human endogenous retrovirus, HERV-K, in the blood cells of leukemia patients. *Leukemia*, **16**, 254–259.
46. Thalgott,M. *et al.* (2013) Detection of circulating tumor cells in different stages of prostate cancer. *J. Cancer Res. Clin. Oncol.*, **139**, 755–763.
47. Miyamoto,D.T. *et al.* (2012) Androgen receptor signaling in circulating tumor cells as a marker of hormonally responsive prostate cancer. *Cancer Discov.*, **2**, 995–1003.
48. Kallergi,G. *et al.* (2007) Phosphorylation of FAK, PI-3K, and impaired actin organization in CK-positive micrometastatic breast cancer cells. *Mol. Med.*, **13**, 79–88.
49. Théry,C. *et al.* (2002) Exosomes: composition, biogenesis and function. *Nat. Rev. Immunol.*, **2**, 569–579.
50. Balaj,L. *et al.* (2011) Tumour microvesicles contain retrotransposon elements and amplified oncogene sequences. *Nat. Commun.*, **2**, 180.
51. Zhang,H.G. *et al.* (2014) Exosomes: a novel pathway of local and distant intercellular communication that facilitates the growth and metastasis of neoplastic lesions. *Am. J. Pathol.*, **184**, 28–41.

Received February 3, 2014; revised May 9, 2014; accepted May 12, 2014

Downloaded from http://carcin.oxfordjournals.org/ at National University of Ireland, Galway on September 19, 2014

References

Abate-Shen, C., and Shen, M.M. (2000). Molecular genetics of prostate cancer. Genes & development *14*, 2410-2434.

Agoni, L., Guha, C., and Lenz, J. (2013a). Detection of Human Endogenous Retrovirus K (HERV-K) Transcripts in Human Prostate Cancer Cell Lines. Front Oncol *3*, 180.

Agoni, L., Lenz, J., and Guha, C. (2013b). Variant splicing and influence of ionizing radiation on human endogenous retrovirus K (HERV-K) transcripts in cancer cell lines. PLoS One *8*, e76472.

Alfahad, T., and Nath, A. (2013). Retroviruses and amyotrophic lateral sclerosis. Antiviral research *99*, 180-187.

Ambs, S., and Glynn, S.A. (2011). Candidate pathways linking inducible nitric oxide synthase to a basal-like transcription pattern and tumor progression in human breast cancer. Cell Cycle *10*, 619-624.

Armbruester, V., Sauter, M., Krautkraemer, E., Meese, E., Kleiman, A., Best, B., Roemer, K., and Mueller-Lantzsch, N. (2002). A novel gene from the human endogenous retrovirus K expressed in transformed cells. Clin Cancer Res *8*, 1800-1807.

Armbruester, V., Sauter, M., Roemer, K., Best, B., Hahn, S., Nty, A., Schmid, A., Philipp, S., Mueller, A., and Mueller-Lantzsch, N. (2004). Np9 protein of human endogenous retrovirus K interacts with ligand of numb protein X. J Virol *78*, 10310-10319.

Azar, G.A., and Thibodeau, J. (2002). Human endogenous retrovirus IDDMK(1,2)22 and mouse mammary tumor virus superantigens differ in their ability to stimulate murine T cell hybridomas. Immunol Lett *81*, 87-91.

Badenhoop, K., Donner, H., Neumann, J., Herwig, J., Kurth, R., Usadel, K.H., and Tonjes, R.R. (1999). IDDM patients neither show humoral reactivities against endogenous retroviral envelope protein nor do they differ in retroviral mRNA expression from healthy relatives or normal individuals. Diabetes *48*, 215-218.

Balada, E., Ordi-Ros, J., and Vilardell-Tarres, M. (2009). Molecular mechanisms mediated by human endogenous retroviruses (HERVs) in autoimmunity. Rev Med Virol *19*, 273-286.

Balaj, L., Lessard, R., Dai, L., Cho, Y.J., Pomeroy, S.L., Breakefield, X.O., and Skog, J. (2011). Tumour microvesicles contain retrotransposon elements and amplified oncogene sequences. Nat Commun *2*, 180.

Bannert, N., and Kurth, R. (2004). Retroelements and the human genome: new perspectives on an old relation. Proc Natl Acad Sci U S A *101 Suppl 2*, 14572-14579.

Bannert, N., and Kurth, R. (2006). The evolutionary dynamics of human endogenous retroviral families. Annu Rev Genomics Hum Genet *7*, 149-173.

Bello, D., Webber, M.M., Kleinman, H.K., Wartinger, D.D., and Rhim, J.S. (1997). Androgen responsive adult human prostatic epithelial cell lines immortalized by human papillomavirus 18. Carcinogenesis *18*, 1215-1223.

Belshaw, R., Dawson, A.L., Woolven-Allen, J., Redding, J., Burt, A., and Tristem, M. (2005). Genomewide screening reveals high levels of insertional polymorphism in the human endogenous retrovirus family HERV-K(HML2): implications for present-day activity. J Virol *79*, 12507-12514.

Belshaw, R., Pereira, V., Katzourakis, A., Talbot, G., Paces, J., Burt, A., and Tristem, M. (2004). Long-term reinfection of the human genome by endogenous retroviruses. Proc Natl Acad Sci U S A *101*, 4894-4899.
Bill-Axelson, A., Holmberg, L., Ruutu, M., Garmo, H., Stark, J.R., Busch, C., Nordling, S., Haggman, M., Andersson, S.O., Bratell, S*., et al.* (2011). Radical prostatectomy versus watchful waiting in early prostate cancer. N Engl J Med *364*, 1708-1717.
Blaise, S., Ruggieri, A., Dewannieux, M., Cosset, F.L., and Heidmann, T. (2004). Identification of an envelope protein from the FRD family of human endogenous retroviruses (HERV-FRD) conferring infectivity and functional conservation among simians. J Virol *78*, 1050-1054.
Boese, A., Sauter, M., Galli, U., Best, B., Herbst, H., Mayer, J., Kremmer, E., Roemer, K., and Mueller-Lantzsch, N. (2000). Human endogenous retrovirus protein cORF supports cell transformation and associates with the promyelocytic leukemia zinc finger protein. Oncogene *19*, 4328-4336.
Boller, K., Konig, H., Sauter, M., Mueller-Lantzsch, N., Lower, R., Lower, J., and Kurth, R. (1993). Evidence that HERV-K is the endogenous retrovirus sequence that codes for the human teratocarcinoma-derived retrovirus HTDV. Virology *196*, 349-353.
Boller, K., Schonfeld, K., Lischer, S., Fischer, N., Hoffmann, A., Kurth, R., and Tonjes, R.R. (2008). Human endogenous retrovirus HERV-K113 is capable of producing intact viral particles. J Gen Virol *89*, 567-572.
Bookout, A.L., and Mangelsdorf, D.J. (2003). Quantitative real-time PCR protocol for analysis of nuclear receptor signaling pathways. Nucl Recept Signal *1*, e012.
Brodsky, I., Foley, B., and Gillespie, D. (1993a). Expression of human endogenous retrovirus (HERV-K) in chronic myeloid leukemia. Leuk Lymphoma *11 Suppl 1*, 119-123.
Brodsky, I., Foley, B., Haines, D., Johnston, J., Cuddy, K., and Gillespie, D. (1993b). Expression of HERV-K proviruses in human leukocytes. Blood *81*, 2369-2374.
Bronson, D.L., Ritzi, D.M., Fraley, E.E., and Dalton, A.J. (1978). Morphologic evidence for retrovirus production by epithelial cells derived from a human testicular tumor metastasis. J Natl Cancer Inst *60*, 1305-1308.
Burke, A.J., Sullivan, F.J., Giles, F.J., and Glynn, S.A. (2013). The yin and yang of nitric oxide in cancer progression. Carcinogenesis *34*, 503-512.
Burmeister, T., Ebert, A.D., Pritze, W., Loddenkemper, C., Schwartz, S., and Thiel, E. (2004). Insertional polymorphisms of endogenous HERV-K113 and HERV-K115 retroviruses in breast cancer patients and age-matched controls. AIDS Res Hum Retroviruses *20*, 1223-1229.
Buscher, K., Hahn, S., Hofmann, M., Trefzer, U., Ozel, M., Sterry, W., Lower, J., Lower, R., Kurth, R., and Denner, J. (2006). Expression of the human endogenous retrovirus-K transmembrane envelope, Rec and Np9 proteins in melanomas and melanoma cell lines. Melanoma Res *16*, 223-234.
Buscher, K., Trefzer, U., Hofmann, M., Sterry, W., Kurth, R., and Denner, J. (2005). Expression of human endogenous retrovirus K in melanomas and melanoma cell lines. Cancer Res *65*, 4172-4180.
Cairns, P., Esteller, M., Herman, J.G., Schoenberg, M., Jeronimo, C., Sanchez-Cespedes, M., Chow, N.H., Grasso, M., Wu, L., Westra, W.B*., et al.* (2001). Molecular detection of prostate cancer in urine by GSTP1 hypermethylation. Clin Cancer Res *7*, 2727-2730.
Cancer Genome Atlas Research, N., Weinstein, J.N., Collisson, E.A., Mills, G.B., Shaw, K.R., Ozenberger, B.A., Ellrott, K., Shmulevich, I., Sander, C., and Stuart, J.M.

(2013). The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet *45*, 1113-1120.
Caricasole, A., Ward-van Oostwaard, D., Mummery, C., and van den Eijnden-van Raaij, A. (2000). Bone morphogenetic proteins and retinoic acid induce human endogenous retrovirus HERV-K expression in NT2D1 human embryonal carcinoma cells. Dev Growth Differ *42*, 407-411.
Carter, H.B. (2013). American Urological Association (AUA) guideline on prostate cancer detection: process and rationale. BJU Int *112*, 543-547.
Carter, P.J. (2006). Potent antibody therapeutics by design. Nat Rev Immunol *6*, 343-357.
Castagnetta, L.A., Granata, O.M., Bellavia, V., Amodio, R., Scaccianoce, E., Notarbartolo, M., Follari, M.R., Miceli, M.D., and Carruba, G. (1997). Product of aromatase activity in intact LNCaP and MCF-7 human cancer cells. J Steroid Biochem Mol Biol *61*, 287-292.
Cegolon, L., Salata, C., Weiderpass, E., Vineis, P., Palu, G., and Mastrangelo, G. (2013). Human endogenous retroviruses and cancer prevention: evidence and prospects. BMC Cancer *13*, 4.
Chen, T., Meng, Z., Gan, Y., Wang, X., Gu, Y., Xu, X., Tang, J., Zhou, H., Zhang, X., Gan, X*., et al.* (2013). The viral oncogene Np9 acts as a critical molecular switch for co-activating beta-catenin, ERK, Akt and Notch1 and promoting the growth of human leukemia stem/progenitor cells. Leukemia.
Chiappinelli, K.B., Strissel, P.L., Desrichard, A., Li, H., Henke, C., Akman, B., Hein, A., Rote, N.S., Cope, L.M., Snyder, A*., et al.* (2015). Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses. Cell *162*, 974-986.
Coffin, J.M., Hughes, S.H., and Varmus, H. (1997). Retroviruses (New York: Cold Spring Harbor Laboratory Press).
Cohen, C.J., Lock, W.M., and Mager, D.L. (2009). Endogenous retroviral LTRs as promoters for human genes: a critical assessment. Gene *448*, 105-114.
Collins, A.T., and Maitland, N.J. (2006). Prostate cancer stem cells. Eur J Cancer *42*, 1213-1218.
Consortium, E.P. (2004). The ENCODE (ENCyclopedia Of DNA Elements) Project. Science *306*, 636-640.
Consortium, E.P. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57-74.
Conti, M. (2008). The selfish cell : an evolutionary defeat ([Dordrecht ; London]: Springer).
Contreras-Galindo, R., Kaplan, M.H., Leissner, P., Verjat, T., Ferlenghi, I., Bagnoli, F., Giusti, F., Dosik, M.H., Hayes, D.F., Gitlin, S.D*., et al.* (2008). Human endogenous retrovirus K (HML-2) elements in the plasma of people with lymphoma and breast cancer. J Virol *82*, 9329-9336.
Cordaux, R., and Batzer, M.A. (2009). The impact of retrotransposons on human genome evolution. Nat Rev Genet *10*, 691-703.
Coussens, L.M., and Werb, Z. (2002). Inflammation and cancer. Nature *420*, 860-867.
Crick, F.H. (1968). The origin of the genetic code. Journal of molecular biology *38*, 367-379.
Curtin, F., Lang, A.B., Perron, H., Laumonier, M., Vidal, V., Porchet, H.C., and Hartung, H.P. (2012). GNbAC1, a Humanized Monoclonal Antibody Against the Envelope Protein of Multiple Sclerosis-Associated Endogenous Retrovirus: A First-in-Humans Randomized Clinical Study. Clin Ther *34*, 2268-2278.

Czyz, J., Szpak, K., and Madeja, Z. (2012). The role of connexins in prostate cancer promotion and progression. Nature reviews Urology *9*, 274-282.
Damber, J.E., and Aus, G. (2008). Prostate cancer. Lancet *371*, 1710-1721.
De Marzo, A.M., Platz, E.A., Sutcliffe, S., Xu, J., Gronberg, H., Drake, C.G., Nakai, Y., Isaacs, W.B., and Nelson, W.G. (2007). Inflammation in prostate carcinogenesis. Nat Rev Cancer *7*, 256-269.
Denne, M., Sauter, M., Armbruester, V., Licht, J.D., Roemer, K., and Mueller-Lantzsch, N. (2007). Physical and functional interactions of human endogenous retrovirus proteins Np9 and rec with the promyelocytic leukemia zinc finger protein. J Virol *81*, 5607-5616.
Depil, S., Roche, C., Dussart, P., and Prin, L. (2002). Expression of a human endogenous retrovirus, HERV-K, in the blood cells of leukemia patients. Leukemia *16*, 254-259.
Dewannieux, M., Harper, F., Richaud, A., Letzelter, C., Ribet, D., Pierron, G., and Heidmann, T. (2006). Identification of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements. Genome Res *16*, 1548-1556.
Doolittle, W.F. (2013). Is junk DNA bunk? A critique of ENCODE. Proc Natl Acad Sci U S A *110*, 5294-5300.
Downey, R.F., Sullivan, F.J., Wang-Johanning, F., Ambs, S., Giles, F.J., and Glynn, S.A. (2015). Human endogenous retrovirus K and cancer: Innocent bystander or tumorigenic accomplice? Int J Cancer *137*, 1249-1257.
Duchesne, G.M. (2001). Radiation for prostate cancer. Lancet Oncol *2*, 73-81.
Duijvesz, D., Luider, T., Bangma, C.H., and Jenster, G. (2011). Exosomes as biomarker treasure chests for prostate cancer. Eur Urol *59*, 823-831.
Eckel, R.H., Alberti, K.G., Grundy, S.M., and Zimmet, P.Z. (2010). The metabolic syndrome. Lancet *375*, 181-183.
Ehrlich, M. (2002). DNA methylation in cancer: too much, but also too little. Oncogene *21*, 5400-5413.
Ellem, S.J., and Risbridger, G.P. (2010). Aromatase and regulating the estrogen:androgen ratio in the prostate gland. J Steroid Biochem Mol Biol *118*, 246-251.
Esnault, C., Heidmann, O., Delebecque, F., Dewannieux, M., Ribet, D., Hance, A., Heidmann, T., and Schwartz, O. (2005). APOBEC3G cytidine deaminase inhibits retrotransposition of endogenous retroviruses. Nature *433*, 430-433.
Etkind, P.R., Lumb, K., Du, J., and Racevskis, J. (1997). Type 1 HERV-K genome is spliced into subgenomic transcripts in the human breast tumor cell line T47D. Virology *234*, 304-308.
Etzioni, R., Urban, N., Ramsey, S., McIntosh, M., Schwartz, S., Reid, B., Radich, J., Anderson, G., and Hartwell, L. (2003). The case for early detection. Nat Rev Cancer *3*, 243-252.
Evans, A.S. (1976). Causation and disease: the Henle-Koch postulates revisited. The Yale journal of biology and medicine *49*, 175-195.
Evans, A.S., and Kaslow, R.A. (1997). Viral infections of humans : epidemiology and control, 4th ed., completely rev. and expanded / edited by Afred S. Evans and Richard A. Kaslow. edn (New York ; London: Plenum Medical Book Co).
Feldman, B.J., and Feldman, D. (2001). The development of androgen-independent prostate cancer. Nat Rev Cancer *1*, 34-45.
Feldman, H.A., Longcope, C., Derby, C.A., Johannes, C.B., Araujo, A.B., Coviello, A.D., Bremner, W.J., and McKinlay, J.B. (2002). Age trends in the level of serum testosterone and other hormones in middle-aged men: longitudinal results from

the Massachusetts male aging study. The Journal of clinical endocrinology and metabolism *87*, 589-598.
Fields, B.N., Knipe, D.M., and Howley, P.M. (2007). Fields' virology, 5th ed. / editors-in-chief, David M. Knipe, Peter M. Howley ; associate editors, Diane E. Griffin ... [et al.]. edn (Philadelphia, Pa. ; London: Wolters Kluwer/Lippincott Williams & Wilkins).
Flockerzi, A., Ruggieri, A., Frank, O., Sauter, M., Maldener, E., Kopper, B., Wullich, B., Seifarth, W., Muller-Lantzsch, N., Leib-Mosch, C*., et al.* (2008). Expression patterns of transcribed human endogenous retrovirus HERV-K(HML-2) loci in human tissues and the need for a HERV Transcriptome Project. BMC Genomics *9*, 354.
Florl, A.R., Lower, R., Schmitz-Drager, B.J., and Schulz, W.A. (1999). DNA methylation and expression of LINE-1 and HERV-K provirus sequences in urothelial and renal cell carcinomas. Br J Cancer *80*, 1312-1321.
Foster, A.E., and Rooney, C.M. (2006). Improving T cell therapy for cancer. Expert Opin Biol Ther *6*, 215-229.
Freshney, R.I. (2010). Culture of animal cells : a manual of basic technique and specialized applications, 6th ed. edn (Hoboken, N.J.: Wiley-Blackwell ; Chichester : John Wiley [distributor]).
Fuchs, N.V., Kraft, M., Tondera, C., Hanschmann, K.M., Lower, J., and Lower, R. (2011). Expression of the human endogenous retrovirus (HERV) group HML-2/HERV-K does not depend on canonical promoter elements but is regulated by transcription factors Sp1 and Sp3. J Virol *85*, 3436-3448.
Fuchs, N.V., Loewer, S., Daley, G.Q., Izsvak, Z., Lower, J., and Lower, R. (2013). Human endogenous retrovirus K (HML-2) RNA and protein expression is a marker for human embryonic and induced pluripotent stem cells. Retrovirology *10*, 115.
Gabriel, U., Steidler, A., Trojan, L., Michel, M.S., Seifarth, W., and Fabarius, A. (2010). Smoking increases transcription of human endogenous retroviruses in a newly established in vitro cell model and in normal urothelium. AIDS Res Hum Retroviruses *26*, 883-888.
Galli, U.M., Sauter, M., Lecher, B., Maurer, S., Herbst, H., Roemer, K., and Mueller-Lantzsch, N. (2005). Human endogenous retrovirus rec interferes with germ cell development in mice and may cause carcinoma in situ, the predecessor lesion of germ cell tumors. Oncogene *24*, 3223-3228.
Gandaglia, G., Sammon, J.D., Chang, S.L., Choueiri, T.K., Hu, J.C., Karakiewicz, P.I., Kibel, A.S., Kim, S.P., Konijeti, R., Montorsi, F*., et al.* (2014). Comparative effectiveness of robot-assisted and open radical prostatectomy in the postdissemination era. J Clin Oncol *32*, 1419-1426.
Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., 3rd, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nature methods *6*, 343-345.
Gilligan, T., and Kantoff, P.W. (2002). Chemotherapy for prostate cancer. Urology *60*, 94-100; discussion 100.
Gimenez, J., Montgiraud, C., Pichon, J.P., Bonnaud, B., Arsac, M., Ruel, K., Bouton, O., and Mallet, F. (2010). Custom human endogenous retroviruses dedicated microarray identifies self-induced HERV-W family elements reactivated in testicular cancer upon methylation control. Nucleic Acids Res *38*, 2229-2246.
Giovannucci, E., Rimm, E.B., Ascherio, A., Colditz, G.A., Spiegelman, D., Stampfer, M.J., and Willett, W.C. (1999). Smoking and risk of total and fatal prostate cancer in United States health professionals. Cancer Epidemiol Biomarkers Prev *8*, 277-282.

Glynn, S.A., Boersma, B.J., Dorsey, T.H., Yi, M., Yfantis, H.G., Ridnour, L.A., Martin, D.N., Switzer, C.H., Hudson, R.S., Wink, D.A*., et al.* (2010). Increased NOS2 predicts poor survival in estrogen receptor-negative breast cancer patients. J Clin Invest *120*, 3843-3854.
Goering, W., Ribarska, T., and Schulz, W.A. (2011). Selective changes of retroelement expression in human prostate cancer. Carcinogenesis *32*, 1484-1492.
Goering, W., Schmitt, K., Dostert, M., Schaal, H., Deenen, R., Mayer, J., and Schulz, W.A. (2015). Human endogenous retrovirus HERV-K(HML-2) activity in prostate cancer is dominated by a few loci. Prostate *75*, 1958-1971.
Golan, M., Hizi, A., Resau, J.H., Yaal-Hahoshen, N., Reichman, H., Keydar, I., and Tsarfaty, I. (2008). Human endogenous retrovirus (HERV-K) reverse transcriptase as a breast cancer prognostic marker. Neoplasia *10*, 521-533.
Gonzalez-Hernandez, M.J., Swanson, M.D., Contreras-Galindo, R., Cookinham, S., King, S.R., Noel, R.J., Jr., Kaplan, M.H., and Markovitz, D.M. (2012). Expression of human endogenous retrovirus type K (HML-2) is activated by the Tat protein of HIV-1. J Virol *86*, 7790-7805.
Gottschalk, S., Heslop, H.E., and Rooney, C.M. (2005). Adoptive immunotherapy for EBV-associated malignancies. Leuk Lymphoma *46*, 1-10.
Gould, S.J., Booth, A.M., and Hildreth, J.E. (2003). The Trojan exosome hypothesis. Proc Natl Acad Sci U S A *100*, 10592-10597.
Greaves, M. (2007). Darwinian medicine: a case for cancer. Nat Rev Cancer *7*, 213-221.
Gronberg, H. (2003). Prostate cancer epidemiology. Lancet *361*, 859-864.
Guerrieri, F., Belloni, L., Pediconi, N., and Levrero, M. (2013). Molecular mechanisms of HBV-associated hepatocarcinogenesis. Seminars in liver disease *33*, 147-156.
Guess, H.A. (1992). Benign prostatic hyperplasia: antecedents and natural history. Epidemiologic reviews *14*, 131-153.
Gulati, R., Wever, E.M., Tsodikov, A., Penson, D.F., Inoue, L.Y., Katcher, J., Lee, S.Y., Heijnsdijk, E.A., Draisma, G., de Koning, H.J*., et al.* (2011). What if i don't treat my PSA-detected prostate cancer? Answers from three natural history models. Cancer Epidemiol Biomarkers Prev *20*, 740-750.
Gustafsson, C., Govindarajan, S., and Minshull, J. (2004). Codon bias and heterologous protein expression. Trends in biotechnology *22*, 346-353.
Hahn, S., Ugurel, S., Hanschmann, K.M., Strobel, H., Tondera, C., Schadendorf, D., Lower, J., and Lower, R. (2008). Serological response to human endogenous retrovirus K in melanoma patients correlates with survival probability. AIDS Res Hum Retroviruses *24*, 717-723.
Hakomori, S. (2002). Glycosylation defining cancer malignancy: new wine in an old bottle. Proc Natl Acad Sci U S A *99*, 10231-10233.
Hamburg, M.A., and Collins, F.S. (2010). The path to personalized medicine. N Engl J Med *363*, 301-304.
Han, G.Z., and Worobey, M. (2012). An endogenous foamy-like viral element in the coelacanth genome. PLoS Pathog *8*, e1002790.
Hanke, K., Chudak, C., Kurth, R., and Bannert, N. (2013). The Rec protein of HERV-K(HML-2) upregulates androgen receptor activity by binding to the human small glutamine-rich tetratricopeptide repeat protein (hSGT). Int J Cancer *132*, 556-567.
Hellerstedt, B.A., and Pienta, K.J. (2002). The current state of hormonal therapy for prostate cancer. CA Cancer J Clin *52*, 154-179.

Herbst, H., Sauter, M., Kuhler-Obbarius, C., Loning, T., and Mueller-Lantzsch, N. (1998). Human endogenous retrovirus (HERV)-K transcripts in germ cell and trophoblastic tumours. APMIS *106*, 216-220.
Hermans, K.G., van der Korput, H.A., van Marion, R., van de Wijngaart, D.J., Ziel-van der Made, A., Dits, N.F., Boormans, J.L., van der Kwast, T.H., van Dekken, H., Bangma, C.H.*, et al.* (2008). Truncated ETV1, fused to novel tissue-specific genes, and full-length ETV1 in prostate cancer. Cancer Res *68*, 7541-7549.
Herve, C.A., Lugli, E.B., Brand, A., Griffiths, D.J., and Venables, P.J. (2002). Autoantibodies to human endogenous retrovirus-K are frequently detected in health and disease and react with multiple epitopes. Clin Exp Immunol *128*, 75-82.
Hess, R.A., Bunick, D., Lee, K.H., Bahr, J., Taylor, J.A., Korach, K.S., and Lubahn, D.B. (1997). A role for oestrogens in the male reproductive system. Nature *390*, 509-512.
Heyne, K., Kolsch, K., Bruand, M., Kremmer, E., Grasser, F.A., Mayer, J., and Roemer, K. (2015). Np9, a cellular protein of retroviral ancestry restricted to human, chimpanzee and gorilla, binds and regulates ubiquitin ligase MDM2. Cell Cycle *14*, 2619-2633.
Hickey, K., Do, K.A., and Green, A. (2001). Smoking and prostate cancer. Epidemiologic reviews *23*, 115-125.
Higano, C.S., Small, E.J., Schellhammer, P., Yasothan, U., Gubernick, S., Kirkpatrick, P., and Kantoff, P.W. (2010). Sipuleucel-T. Nature reviews Drug discovery *9*, 513-514.
Hill, A.B. (1965). The Environment and Disease: Association or Causation? Proceedings of the Royal Society of Medicine *58*, 295-300.
Hodges-Vazquez, M., Wilson, J.P., Hughes, H., and Garman, P. (2012). The yellow fever 17D vaccine and risk of malignant melanoma in the United States military. Vaccine *30*, 4476-4479.
Hofacre, A., and Fan, H. (2010). Jaagsiekte sheep retrovirus biology and oncogenesis. Viruses *2*, 2618-2648.
Hohn, O., Hanke, K., and Bannert, N. (2013). HERV-K(HML-2), the Best Preserved Family of HERVs: Endogenization, Expression, and Implications in Health and Disease. Front Oncol *3*, 246.
Horoszewicz, J.S., Leong, S.S., Chu, T.M., Wajsman, Z.L., Friedman, M., Papsidero, L., Kim, U., Chai, L.S., Kakati, S., Arya, S.K.*, et al.* (1980). The LNCaP cell line--a new model for studies on human prostatic carcinoma. Progress in clinical and biological research *37*, 115-132.
Hsiao, F.C., Lin, M., Tai, A., Chen, G., and Huber, B.T. (2006). Cutting edge: Epstein-Barr virus transactivates the HERV-K18 superantigen by docking to the human complement receptor 2 (CD21) on primary B cells. J Immunol *177*, 2056-2060.
Hsiao, F.C., Tai, A.K., Deglon, A., Sutkowski, N., Longnecker, R., and Huber, B.T. (2009). EBV LMP-2A employs a novel mechanism to transactivate the HERV-K18 superantigen through its ITAM. Virology *385*, 261-266.
Hsing, A.W., Sakoda, L.C., and Chua, S., Jr. (2007). Obesity, metabolic syndrome, and prostate cancer. The American journal of clinical nutrition *86*, s843-857.
Hsu, P.D., Lander, E.S., and Zhang, F. (2014). Development and applications of CRISPR-Cas9 for genome engineering. Cell *157*, 1262-1278.
Hudson, R.S., Yi, M., Esposito, D., Watkins, S.K., Hurwitz, A.A., Yfantis, H.G., Lee, D.H., Borin, J.F., Naslund, M.J., Alexander, R.B.*, et al.* (2012). MicroRNA-1 is a candidate tumor suppressor and prognostic marker in human prostate cancer. Nucleic Acids Res *40*, 3689-3703.

Hughes, J.F., and Coffin, J.M. (2004). Human endogenous retrovirus K solo-LTR formation and insertional polymorphisms: implications for human and viral evolution. Proc Natl Acad Sci U S A *101*, 1668-1672.
Humer, J., Waltenberger, A., Grassauer, A., Kurz, M., Valencak, J., Rapberger, R., Hahn, S., Lower, R., Wolff, K., Bergmann, M.*, et al.* (2006). Identification of a melanoma marker derived from melanoma-associated endogenous retroviruses. Cancer Res *66*, 1658-1663.
Huncharek, M., Haddock, K.S., Reid, R., and Kupelnick, B. (2010). Smoking as a risk factor for prostate cancer: a meta-analysis of 24 prospective cohort studies. American journal of public health *100*, 693-701.
Hurst, T., Pace, M., Katzourakis, A., Phillips, R., Klenerman, P., Frater, J., and Magiorkinis, G. (2016). Human endogenous retrovirus (HERV) expression is not induced by treatment with the histone deacetylase (HDAC) inhibitors in cellular models of HIV-1 latency. Retrovirology *13*, 10.
Huthoff, H., and Towers, G.J. (2008). Restriction of retroviral replication by APOBEC3G/F and TRIM5alpha. Trends Microbiol *16*, 612-619.
Iramaneerat, K., Rattanatunyong, P., Khemapech, N., Triratanachat, S., and Mutirangura, A. (2011). HERV-K hypomethylation in ovarian clear cell carcinoma is associated with a poor prognosis and platinum resistance. Int J Gynecol Cancer *21*, 51-57.
Ishida, T., Obata, Y., Ohara, N., Matsushita, H., Sato, S., Uenaka, A., Saika, T., Miyamura, T., Chayama, K., Nakamura, Y.*, et al.* (2008). Identification of the HERV-K gag antigen in prostate cancer by SEREX using autologous patient serum and its immunogenicity. Cancer Immun *8*, 15.
Itsiopoulos, C., Hodge, A., and Kaimakamis, M. (2009). Can the Mediterranean diet prevent prostate cancer? Molecular nutrition & food research *53*, 227-239.
Jenster, G., van der Korput, H.A., van Vroonhoven, C., van der Kwast, T.H., Trapman, J., and Brinkmann, A.O. (1991). Domains of the human androgen receptor involved in steroid binding, transcriptional activation, and subcellular localization. Mol Endocrinol *5*, 1396-1404.
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*, 50.
Jha, A.R., Pillai, S.K., York, V.A., Sharp, E.R., Storm, E.C., Wachter, D.J., Martin, J.N., Deeks, S.G., Rosenberg, M.G., Nixon, D.F.*, et al.* (2009). Cross-sectional dating of novel haplotypes of HERV-K 113 and HERV-K 115 indicate these proviruses originated in Africa before Homo sapiens. Mol Biol Evol *26*, 2617-2626.
Jintaridth, P., and Mutirangura, A. (2010). Distinctive patterns of age-dependent hypomethylation in interspersed repetitive sequences. Physiol Genomics *41*, 194-200.
Johnston, J.B., Silva, C., Holden, J., Warren, K.G., Clark, A.W., and Power, C. (2001). Monocyte activation and differentiation augment human endogenous retrovirus expression: implications for inflammatory brain diseases. Ann Neurol *50*, 434-442.
Jones, R.B., Garrison, K.E., Mujib, S., Mihajlovic, V., Aidarus, N., Hunter, D.V., Martin, E., John, V.M., Zhan, W., Faruk, N.F.*, et al.* (2012). HERV-K-specific T cells eliminate diverse HIV-1/2 and SIV primary isolates. J Clin Invest *122*, 4473-4489.
Jung, M.E., Ouk, S., Yoo, D., Sawyers, C.L., Chen, C., Tran, C., and Wongvipat, J. (2010). Structure-activity relationship for thiohydantoin androgen receptor antagonists for castration-resistant prostate cancer (CRPC). J Med Chem *53*, 2779-2796.

Kaighn, M.E., Narayan, K.S., Ohnuki, Y., Lechner, J.F., and Jones, L.W. (1979). Establishment and characterization of a human prostatic carcinoma cell line (PC-3). Investigative urology *17*, 16-23.
Kallergi, G., Mavroudis, D., Georgoulias, V., and Stournaras, C. (2007). Phosphorylation of FAK, PI-3K, and impaired actin organization in CK-positive micrometastatic breast cancer cells. Mol Med *13*, 79-88.
Kalter, S.S., Helmke, R.J., Heberling, R.L., Panigel, M., Fowler, A.K., Strickland, J.E., and Hellman, A. (1973). Brief communication: C-type particles in normal human placentas. J Natl Cancer Inst *50*, 1081-1084.
Katoh, I., Mirova, A., Kurata, S., Murakami, Y., Horikawa, K., Nakakuki, N., Sakai, T., Hashimoto, K., Maruyama, A., Yonaga, T.*, et al.* (2011). Activation of the long terminal repeat of human endogenous retrovirus K by melanoma-specific transcription factor MITF-M. Neoplasia *13*, 1081-1092.
Kaufmann, S., Sauter, M., Schmitt, M., Baumert, B., Best, B., Boese, A., Roemer, K., and Mueller-Lantzsch, N. (2010). Human endogenous retrovirus protein Rec interacts with the testicular zinc-finger protein and androgen receptor. J Gen Virol *91*, 1494-1502.
Kenfield, S.A., Stampfer, M.J., Chan, J.M., and Giovannucci, E. (2011). Smoking and prostate cancer survival and recurrence. JAMA *305*, 2548-2555.
Kirschner, M.A., Ertel, N., and Schneider, G. (1981). Obesity, hormones, and cancer. Cancer Res *41*, 3711-3717.
Knossl, M., Lower, R., and Lower, J. (1999). Expression of the human endogenous retrovirus HTDV/HERV-K is enhanced by cellular transcription factor YY1. J Virol *73*, 1254-1261.
Kobrinsky, N.L., Klug, M.G., Hokanson, P.J., Sjolander, D.E., and Burd, L. (2003). Impact of smoking on cancer stage at diagnosis. J Clin Oncol *21*, 907-913.
Kraus, B., Fischer, K., Buchner, S.M., Wels, W.S., Lower, R., Sliva, K., and Schnierle, B.S. (2013). Vaccination directed against the human endogenous retrovirus-K envelope protein inhibits tumor growth in a murine model system. PLoS One *8*, e72756.
Kreimer, U., Schulz, W.A., Koch, A., Niegisch, G., and Goering, W. (2013). HERV-K and LINE-1 DNA Methylation and Reexpression in Urothelial Carcinoma. Front Oncol *3*, 255.
Krishnamurthy, J., Rabinovich, B.A., Mi, T., Switzer, K.C., Olivares, S., Maiti, S.N., Plummer, J.B., Singh, H., Kumaresan, P.R., Huls, H.M.*, et al.* (2015). Genetic Engineering of T Cells to Target HERV-K, an Ancient Retrovirus on Melanoma. Clin Cancer Res *21*, 3241-3251.
Krone, B., and Grange, J.M. (2010). Melanoma, Darwinian medicine and the inner world. J Cancer Res Clin Oncol *136*, 1787-1794.
Krone, B., Kolmel, K.F., Grange, J.M., Mastrangelo, G., Henz, B.M., Botev, I.N., Niin, M., Seebacher, C., Lambert, D., Shafir, R.*, et al.* (2003). Impact of vaccinations and infectious diseases on the risk of melanoma--evaluation of an EORTC case-control study. Eur J Cancer *39*, 2372-2378.
Krone, B., Kolmel, K.F., Henz, B.M., and Grange, J.M. (2005). Protection against melanoma by vaccination with Bacille Calmette-Guerin (BCG) and/or vaccinia: an epidemiology-based hypothesis on the nature of a melanoma risk factor and its immunological control. Eur J Cancer *41*, 104-117.
Kuh, D., and Ben-Shlomo, Y. (2004). A life course approach to chronic disease epidemiology, 2nd ed. edn (Oxford: Oxford University Press).
Kurth, R., and Bannert, N. (2010a). Beneficial and detrimental effects of human endogenous retroviruses. Int J Cancer *126*, 306-314.

Kurth, R., and Bannert, N. (2010b). Retroviruses : molecular biology, genomics and pathogenesis ([Wymondham]: Caister Academic).
Lander, E.S. (2011). Initial impact of the sequencing of the human genome. Nature *470*, 187-197.
Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W*., et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.
Lee, E., Iskow, R., Yang, L., Gokcumen, O., Haseley, P., Luquette, L.J., 3rd, Lohr, J.G., Harris, C.C., Ding, L., Wilson, R.K*., et al.* (2012). Landscape of somatic retrotransposition in human cancers. Science *337*, 967-971.
Lee, Y.N., and Bieniasz, P.D. (2007). Reconstitution of an infectious human endogenous retrovirus. PLoS Pathog *3*, e10.
Lee, Y.N., Malim, M.H., and Bieniasz, P.D. (2008). Hypermutation of an ancient human retrovirus by APOBEC3G. J Virol *82*, 8762-8770.
Liew, C.C., Ma, J., Tang, H.C., Zheng, R., and Dempsey, A.A. (2006). The peripheral blood transcriptome dynamically reflects system wide biology: a potential diagnostic tool. J Lab Clin Med *147*, 126-132.
Logothetis, C.J., Efstathiou, E., Manuguid, F., and Kirkpatrick, P. (2011). Abiraterone acetate. Nature reviews Drug discovery *10*, 573-574.
Lu, X., Sachs, F., Ramsay, L., Jacques, P.E., Goke, J., Bourque, G., and Ng, H.H. (2014). The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity. Nature structural & molecular biology *21*, 423-425.
Macfarlane, C., and Simmonds, P. (2004). Allelic variation of HERV-K(HML-2) endogenous retroviral elements in human populations. J Mol Evol *59*, 642-656.
Makarov, D.V., Loeb, S., Getzenberg, R.H., and Partin, A.W. (2009). Biomarkers for prostate cancer. Annual review of medicine *60*, 139-151.
Mameli, G., Poddighe, L., Mei, A., Uleri, E., Sotgiu, S., Serra, C., Manetti, R., and Dolei, A. (2012). Expression and activation by Epstein Barr virus of human endogenous retroviruses-W in blood cells and astrocytes: inference for multiple sclerosis. PLoS One *7*, e44991.
Manghera, M., and Douville, R.N. (2013). Endogenous retrovirus-K promoter: a landing strip for inflammatory transcription factors? Retrovirology *10*, 16.
Marchi, E., Kanapin, A., Magiorkinis, G., and Belshaw, R. (2014). Unfixed endogenous retroviral insertions in the human population. J Virol *88*, 9529-9537.
Martin, D.N., Starks, A.M., and Ambs, S. (2013). Biological determinants of health disparities in prostate cancer. Curr Opin Oncol *25*, 235-241.
Mastrangelo, G., Krone, B., Fadda, E., Buja, A., Grange, J.M., Rausa, G., de Vries, E., and Koelmel, K.F. (2009). Does yellow fever 17D vaccine protect against melanoma? Vaccine *27*, 588-591.
Mayer, J., Ehlhardt, S., Seifert, M., Sauter, M., Muller-Lantzsch, N., Mehraein, Y., Zang, K.D., and Meese, E. (2004). Human endogenous retrovirus HERV-K(HML-2) proviruses with Rec protein coding capacity and transcriptional activity. Virology *322*, 190-198.
Medstrand, P., Lindeskog, M., and Blomberg, J. (1992). Expression of human endogenous retroviral sequences in peripheral blood mononuclear cells of healthy individuals. J Gen Virol *73 (Pt 9)*, 2463-2466.
Meeker, A.K., Hicks, J.L., Platz, E.A., March, G.E., Bennett, C.J., Delannoy, M.J., and De Marzo, A.M. (2002). Telomere shortening is an early somatic DNA alteration in human prostate tumorigenesis. Cancer Res *62*, 6405-6409.

Mi, S., Lee, X., Li, X., Veldman, G.M., Finnerty, H., Racie, L., LaVallie, E., Tang, X.Y., Edouard, P., Howes, S.*, et al.* (2000). Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis. Nature *403*, 785-789.
Miyamoto, D.T., Lee, R.J., Stott, S.L., Ting, D.T., Wittner, B.S., Ulman, M., Smas, M.E., Lord, J.B., Brannigan, B.W., Trautwein, J.*, et al.* (2012). Androgen receptor signaling in circulating tumor cells as a marker of hormonally responsive prostate cancer. Cancer discovery *2*, 995-1003.
Mohr, S., and Liew, C.C. (2007). The peripheral-blood transcriptome: new insights into disease and risk assessment. Trends Mol Med *13*, 422-432.
Moreira, D.M., Aronson, W.J., Terris, M.K., Kane, C.J., Amling, C.L., Cooperberg, M.R., Boffetta, P., and Freedland, S.J. (2013). Cigarette smoking is associated with an increased risk of biochemical disease recurrence, metastasis, castration-resistant prostate cancer, and mortality after radical prostatectomy: Results from the SEARCH database. Cancer.
Moreno, J.G., Miller, M.C., Gross, S., Allard, W.J., Gomella, L.G., and Terstappen, L.W. (2005). Circulating tumor cells predict survival in patients with metastatic prostate cancer. Urology *65*, 713-718.
Morozov, V.A., Dao Thi, V.L., and Denner, J. (2013). The transmembrane protein of the human endogenous retrovirus--K (HERV-K) modulates cytokine release and gene expression. PLoS One *8*, e70399.
Moyes, D., Griffiths, D.J., and Venables, P.J. (2007). Insertional polymorphisms: a new lease of life for endogenous retroviruses in human disease. Trends Genet *23*, 326-333.
Mullins, C.S., and Linnebacher, M. (2012). Human endogenous retroviruses and cancer: causality and therapeutic possibilities. World J Gastroenterol *18*, 6027-6035.
Muster, T., Waltenberger, A., Grassauer, A., Hirschl, S., Caucig, P., Romirer, I., Fodinger, D., Seppele, H., Schanab, O., Magin-Lachmann, C.*, et al.* (2003). An endogenous retrovirus derived from human melanoma cells. Cancer Res *63*, 8735-8741.
Nelson, W.G., De Marzo, A.M., and Isaacs, W.B. (2003). Prostate cancer. N Engl J Med *349*, 366-381.
Ono, M., Kawakami, M., and Ushikubo, H. (1987). Stimulation of expression of the human endogenous retrovirus genome by female steroid hormones in human breast cancer cell line T47D. J Virol *61*, 2059-2062.
Ono, M., Yasunaga, T., Miyata, T., and Ushikubo, H. (1986). Nucleotide sequence of human endogenous retrovirus genome related to the mouse mammary tumor virus genome. J Virol *60*, 589-598.
Oricchio, E., Sciamanna, I., Beraldi, R., Tolstonog, G.V., Schumann, G.G., and Spadafora, C. (2007). Distinct roles for LINE-1 and HERV-K retroelements in cell proliferation, differentiation and tumor progression. Oncogene *26*, 4226-4233.
Overington, J.P., Al-Lazikani, B., and Hopkins, A.L. (2006). How many drug targets are there? Nature reviews Drug discovery *5*, 993-996.
Paik, S., Shak, S., Tang, G., Kim, C., Baker, J., Cronin, M., Baehner, F.L., Walker, M.G., Watson, D., Park, T.*, et al.* (2004). A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med *351*, 2817-2826.
Perot, P., Cheynet, V., Decaussin-Petrucci, M., Oriol, G., Mugnier, N., Rodriguez-Lafrasse, C., Ruffion, A., and Mallet, F. (2013). Microarray-based Identification of Individual HERV Loci Expression: Application to Biomarker Discovery in Prostate Cancer. Journal of visualized experiments : JoVE.

Perron, H., and Lang, A. (2010). The human endogenous retrovirus link between genes and environment in multiple sclerosis and in multifactorial diseases associating neuroinflammation. Clin Rev Allergy Immunol *39*, 51-61.
Pfahlberg, A., Kolmel, K.F., Grange, J.M., Mastrangelo, G., Krone, B., Botev, I.N., Niin, M., Seebacher, C., Lambert, D., Shafir, R.*, et al.* (2002). Inverse association between melanoma and previous vaccinations against tuberculosis and smallpox: results of the FEBIM study. J Invest Dermatol *119*, 570-575.
Plotkin, J.B., and Kudla, G. (2011). Synonymous but not the same: the causes and consequences of codon bias. Nat Rev Genet *12*, 32-42.
Prueitt, R.L., Wallace, T.A., Glynn, S.A., Yi, M., Tang, W., Luo, J., Dorsey, T.H., Stagliano, K.E., Gillespie, J.W., Hudson, R.S.*, et al.* (2016). An Immune-Inflammation Gene Expression Signature in Prostate Tumors of Smokers. Cancer Res *76*, 1055-1065.
Rahmani, M., Davis, E.M., Bauer, C., Dent, P., and Grant, S. (2005). Apoptosis induced by the kinase inhibitor BAY 43-9006 in human leukemia cells involves down-regulation of Mcl-1 through inhibition of translation. J Biol Chem *280*, 35217-35227.
Rakoff-Nahoum, S., Kuebler, P.J., Heymann, J.J., M, E.S., G, M.O., G, S.O., Barbour, J.D., Lenz, J., Steinfeld, A.D., and Nixon, D.F. (2006). Detection of T lymphocytes specific for human endogenous retrovirus K (HERV-K) in patients with seminoma. AIDS Res Hum Retroviruses *22*, 52-56.
Rapozzi, V., Della Pietra, E., Zorzet, S., Zacchigna, M., Bonavida, B., and Xodo, L.E. (2013). Nitric oxide-mediated activity in anti-cancer photodynamic therapy. Nitric Oxide *30*, 26-35.
Reiche, J., Pauli, G., and Ellerbrok, H. (2010). Differential expression of human endogenous retrovirus K transcripts in primary human melanocytes and melanoma cell lines after UV irradiation. Melanoma Res *20*, 435-440.
Reis, B.S., Jungbluth, A.A., Frosina, D., Holz, M., Ritter, E., Nakayama, E., Ishida, T., Obata, Y., Carver, B., Scher, H.*, et al.* (2013). Prostate Cancer Progression Correlates with Increased Humoral Immune Response to a Human Endogenous Retrovirus GAG Protein. Clin Cancer Res.
Risbridger, G.P., Bianco, J.J., Ellem, S.J., and McPherson, S.J. (2003). Oestrogens and prostate cancer. Endocr Relat Cancer *10*, 187-191.
Risbridger, G.P., Davis, I.D., Birrell, S.N., and Tilley, W.D. (2010). Breast and prostate cancer: more similar than different. Nat Rev Cancer *10*, 205-212.
Roberts, W.W., Platz, E.A., and Walsh, P.C. (2003). Association of cigarette smoking with extraprostatic prostate cancer in young men. J Urol *169*, 512-516; discussion 516.
Roehrborn, C.G., Boyle, P., Nickel, J.C., Hoefner, K., Andriole, G., Aria, A., and Investigators, A.S. (2002). Efficacy and safety of a dual inhibitor of 5-alpha-reductase types 1 and 2 (dutasteride) in men with benign prostatic hyperplasia. Urology *60*, 434-441.
Roy, M., Pear, W.S., and Aster, J.C. (2007). The multifaceted role of Notch in cancer. Curr Opin Genet Dev *17*, 52-59.
Ruggieri, A., Maldener, E., Sauter, M., Mueller-Lantzsch, N., Meese, E., Fackler, O.T., and Mayer, J. (2009). Human endogenous retrovirus HERV-K(HML-2) encodes a stable signal peptide with biological properties distinct from Rec. Retrovirology *6*, 17.
Ruprecht, K., Mayer, J., Sauter, M., Roemer, K., and Mueller-Lantzsch, N. (2008). Endogenous retroviruses and cancer. Cell Mol Life Sci *65*, 3366-3382.

Rycaj, K., Plummer, J.B., Yin, B., Li, M., Garza, J., Radvanyi, L., Ramondetta, L.M., Lin, K., Johanning, G.L., Tang, D.G.*, et al.* (2015). Cytotoxicity of human endogenous retrovirus K-specific T cells toward autologous ovarian cancer cells. Clin Cancer Res *21*, 471-483.
Sambrook, J., and Russell, D.W. (2001). Molecular cloning : a laboratory manual, 3rd ed. edn (Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press).
Sanchez-Chapado, M., Olmedilla, G., Cabeza, M., Donat, E., and Ruiz, A. (2003). Prevalence of prostate cancer and prostatic intraepithelial neoplasia in Caucasian Mediterranean males: an autopsy study. Prostate *54*, 238-247.
Sarid, R., and Gao, S.J. (2011). Viruses and human cancer: from detection to causality. Cancer Lett *305*, 218-227.
Schanab, O., Humer, J., Gleiss, A., Mikula, M., Sturlan, S., Grunt, S., Okamoto, I., Muster, T., Pehamberger, H., and Waltenberger, A. (2011). Expression of human endogenous retrovirus K is stimulated by ultraviolet radiation in melanoma. Pigment Cell Melanoma Res *24*, 656-665.
Schiavetti, F., Thonnard, J., Colau, D., Boon, T., and Coulie, P.G. (2002). A human endogenous retroviral sequence encoding an antigen recognized on melanoma by cytolytic T lymphocytes. Cancer Res *62*, 5510-5516.
Schmitt, K., Reichrath, J., Roesch, A., Meese, E., and Mayer, J. (2013a). Transcriptional profiling of human endogenous retrovirus group HERV-K(HML-2) loci in melanoma. Genome Biol Evol.
Schmitt, K., Richter, C., Backes, C., Meese, E., Ruprecht, K., and Mayer, J. (2013b). Comprehensive analysis of human endogenous retrovirus group HERV-W locus transcription in multiple sclerosis brain lesions by high-throughput amplicon sequencing. J Virol *87*, 13837-13852.
Schulz, W.A., Steinhoff, C., and Florl, A.R. (2006). Methylation of endogenous human retroelements in health and disease. Curr Top Microbiol Immunol *310*, 211-250.
Seifarth, W., Skladny, H., Krieg-Schneider, F., Reichert, A., Hehlmann, R., and Leib-Mosch, C. (1995). Retrovirus-like particles released from the human breast cancer cell line T47-D display type B- and C-related endogenous retroviral sequences. J Virol *69*, 6408-6416.
Serafino, A., Balestrieri, E., Pierimarchi, P., Matteucci, C., Moroni, G., Oricchio, E., Rasi, G., Mastino, A., Spadafora, C., Garaci, E.*, et al.* (2009). The activation of human endogenous retrovirus K (HERV-K) is implicated in melanoma cell malignant transformation. Exp Cell Res *315*, 849-862.
Shagin, D.A., Lukyanov, K.A., Vagner, L.L., and Matz, M.V. (1999). Regulation of average length of complex PCR product. Nucleic Acids Res *27*, e23.
Shen, M.M., and Abate-Shen, C. (2010). Molecular genetics of prostate cancer: new prospects for old challenges. Genes & development *24*, 1967-2000.
Siegel, R., Naishadham, D., and Jemal, A. (2013). Cancer statistics, 2013. CA Cancer J Clin *63*, 11-30.
Singh, S., Kaye, S., Francis, N., Peston, D., Gore, M., McClure, M., and Bunker, C. (2013). Human endogenous retrovirus K (HERV-K) rec mRNA is expressed in primary melanoma but not in benign naevi or normal skin. Pigment Cell Melanoma Res.
Singh, S., Kaye, S., Gore, M.E., McClure, M.O., and Bunker, C.B. (2009). The role of human endogenous retroviruses in melanoma. Br J Dermatol *161*, 1225-1231.
Smith, C., Tsang, J., Beagley, L., Chua, D., Lee, V., Li, V., Moss, D.J., Coman, W., Chan, K.H., Nicholls, J.*, et al.* (2012). Effective treatment of metastatic forms of

Epstein-Barr virus-associated nasopharyngeal carcinoma with a novel adenovirus-based adoptive immunotherapy. Cancer Res *72*, 1116-1125.
Sramkoski, R.M., Pretlow, T.G., 2nd, Giaconia, J.M., Pretlow, T.P., Schwartz, S., Sy, M.S., Marengo, S.R., Rhim, J.S., Zhang, D., and Jacobberger, J.W. (1999). A new human prostate carcinoma cell line, 22Rv1. In Vitro Cell Dev Biol Anim *35*, 403-409.
Standring, S.e., and Gray, H.A. (2015). Gray's anatomy : the anatomical basis of clinical practice, Forty-First Edition. edn.
Stauffer, Y., Marguerat, S., Meylan, F., Ucla, C., Sutkowski, N., Huber, B., Pelet, T., and Conrad, B. (2001). Interferon-alpha-induced endogenous superantigen. a model linking environment and autoimmunity. Immunity *15*, 591-601.
Stengel, S., Fiebig, U., Kurth, R., and Denner, J. (2010). Regulation of human endogenous retrovirus-K expression in melanomas by CpG methylation. Genes Chromosomes Cancer *49*, 401-411.
Stone, K.R., Mickey, D.D., Wunderli, H., Mickey, G.H., and Paulson, D.F. (1978). Isolation of a human prostate carcinoma cell line (DU 145). Int J Cancer *21*, 274-281.
Stoye, J.P. (2006). Koala retrovirus: a genome invasion in real time. Genome Biol *7*, 241.
Stoye, J.P. (2012). Studies of endogenous retroviruses reveal a continuing evolutionary saga. Nature reviews Microbiology *10*, 395-406.
Subramanian, R.P., Wildschutte, J.H., Russo, C., and Coffin, J.M. (2011). Identification, characterization, and comparative genomic distribution of the HERV-K (HML-2) group of human endogenous retroviruses. Retrovirology *8*, 90.
Sutcliffe, S., and Platz, E.A. (2008). Inflammation and prostate cancer: a focus on infections. Current urology reports *9*, 243-249.
Sutkowski, N., Chen, G., Calderon, G., and Huber, B.T. (2004). Epstein-Barr virus latent membrane protein LMP-2A is sufficient for transactivation of the human endogenous retrovirus HERV-K18 superantigen. J Virol *78*, 7852-7860.
Sutkowski, N., Conrad, B., Thorley-Lawson, D.A., and Huber, B.T. (2001). Epstein-Barr virus transactivates the human endogenous retrovirus HERV-K18 that encodes a superantigen. Immunity *15*, 579-589.
Switzer, C.H., Glynn, S.A., Cheng, R.Y., Ridnour, L.A., Green, J.E., Ambs, S., and Wink, D.A. (2012). S-Nitrosylation of EGFR and Src Activates an Oncogenic Signaling Network in Human Basal-Like Breast Cancer. Mol Cancer Res.
Thalgott, M., Rack, B., Maurer, T., Souvatzoglou, M., Eiber, M., Kress, V., Heck, M.M., Andergassen, U., Nawroth, R., Gschwend, J.E*., et al.* (2013). Detection of circulating tumor cells in different stages of prostate cancer. J Cancer Res Clin Oncol *139*, 755-763.
Thery, C., Zitvogel, L., and Amigorena, S. (2002). Exosomes: composition, biogenesis and function. Nat Rev Immunol *2*, 569-579.
Thorpe, A., and Neal, D. (2003). Benign prostatic hyperplasia. Lancet *361*, 1359-1367.
Tomlins, S.A., Laxman, B., Dhanasekaran, S.M., Helgeson, B.E., Cao, X., Morris, D.S., Menon, A., Jing, X., Cao, Q., Han, B*., et al.* (2007). Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. Nature *448*, 595-599.
Turner, G., Barbulescu, M., Su, M., Jensen-Seaman, M.I., Kidd, K.K., and Lenz, J. (2001). Insertional polymorphisms of full-length endogenous retroviruses in humans. Curr Biol *11*, 1531-1535.
Tymms, M.J. (2000). Transcription factor protocols (Totowa, N.J.: Humana).

Urnovitz, H.B., and Murphy, W.H. (1996). Human endogenous retroviruses: nature, occurrence, and clinical implications in human disease. Clin Microbiol Rev *9*, 72-99.
Viguera, E., Canceill, D., and Ehrlich, S.D. (2001). Replication slippage involves DNA polymerase pausing and dissociation. The EMBO journal *20*, 2587-2595.
Vogetseder, W., Feng, J., and Dierich, M.P. (1995). Reactivity of monoclonal antibodies established against a recombinant human endogenous retrovirus-K (HERV-K) envelope protein. Immunol Lett *46*, 129-134.
Wainstein, M.A., He, F., Robinson, D., Kung, H.J., Schwartz, S., Giaconia, J.M., Edgehouse, N.L., Pretlow, T.P., Bodner, D.R., Kursh, E.D*., et al.* (1994). CWR22: androgen-dependent xenograft model derived from a primary human prostatic carcinoma. Cancer Res *54*, 6049-6052.
Wallace, T.A., Downey, R.F., Seufert, C.J., Schetter, A., Dorsey, T.H., Johnson, C.A., Goldman, R., Loffredo, C.A., Yan, P., Sullivan, F.J*., et al.* (2014). Elevated HERV-K mRNA expression in PBMC is associated with a prostate cancer diagnosis particularly in older men and smokers. Carcinogenesis *35*, 2074-2083.
Wallace, T.A., Prueitt, R.L., Yi, M., Howe, T.M., Gillespie, J.W., Yfantis, H.G., Stephens, R.M., Caporaso, N.E., Loffredo, C.A., and Ambs, S. (2008). Tumor immunobiological differences in prostate cancer between African-American and European-American men. Cancer Res *68*, 927-936.
Wang-Johanning, F., Frost, A.R., Jian, B., Epp, L., Lu, D.W., and Johanning, G.L. (2003). Quantitation of HERV-K env gene expression and splicing in human breast cancer. Oncogene *22*, 1528-1535.
Wang-Johanning, F., Frost, A.R., Johanning, G.L., Khazaeli, M.B., LoBuglio, A.F., Shaw, D.R., and Strong, T.V. (2001). Expression of human endogenous retrovirus k envelope transcripts in human breast cancer. Clin Cancer Res *7*, 1553-1560.
Wang-Johanning, F., Li, M., Esteva, F.J., Hess, K.R., Yin, B., Rycaj, K., Plummer, J.B., Garza, J.G., Ambs, S., and Johanning, G.L. (2013). Human endogenous retrovirus type K antibodies and mRNA as serum biomarkers of early-stage breast cancer. Int J Cancer.
Wang-Johanning, F., Li, M., Esteva, F.J., Hess, K.R., Yin, B., Rycaj, K., Plummer, J.B., Garza, J.G., Ambs, S., and Johanning, G.L. (2014). Human endogenous retrovirus type K antibodies and mRNA as serum biomarkers of early-stage breast cancer. Int J Cancer *134*, 587-595.
Wang-Johanning, F., Liu, J., Rycaj, K., Huang, M., Tsai, K., Rosen, D.G., Chen, D.T., Lu, D.W., Barnhart, K.F., and Johanning, G.L. (2007). Expression of multiple human endogenous retrovirus surface envelope proteins in ovarian cancer. Int J Cancer *120*, 81-90.
Wang-Johanning, F., Radvanyi, L., Rycaj, K., Plummer, J.B., Yan, P., Sastry, K.J., Piyathilake, C.J., Hunt, K.K., and Johanning, G.L. (2008). Human endogenous retrovirus K triggers an antigen-specific immune response in breast cancer patients. Cancer Res *68*, 5869-5877.
Wang-Johanning, F., Rycaj, K., Plummer, J.B., Li, M., Yin, B., Frerich, K., Garza, J.G., Shen, J., Lin, K., Yan, P*., et al.* (2012). Immunotherapeutic potential of anti-human endogenous retrovirus-K envelope protein antibodies in targeting breast tumors. J Natl Cancer Inst *104*, 189-210.
Wang, J., Xie, G., Singh, M., Ghanbarian, A.T., Rasko, T., Szvetnik, A., Cai, H., Besser, D., Prigione, A., Fuchs, N.V*., et al.* (2014). Primate-specific endogenous retrovirus-driven transcription defines naive-like stem cells. Nature *516*, 405-409.

Wang, X., Kruithof-de Julio, M., Economides, K.D., Walker, D., Yu, H., Halili, M.V., Hu, Y.P., Price, S.M., Abate-Shen, C., and Shen, M.M. (2009). A luminal epithelial stem cell that is a cell of origin for prostate cancer. Nature *461*, 495-500.
Wang, Y., Go, V., Holland, J.F., Melana, S.M., and Pogo, B.G. (1998). Expression of mouse mammary tumor virus-like env gene sequences in human breast cancer. Clin Cancer Res *4*, 2565-2568.
Watson, J.D. (2008). Molecular biology of the gene, 6th international ed. edn (San Francisco: Pearson/Benjamin Cummings).
Weiss, R.A. (2006). The discovery of endogenous retroviruses. Retrovirology *3*, 67.
Willer, A., Saussele, S., Gimbel, W., Seifarth, W., Kister, P., Leib-Mosch, C., and Hehlmann, R. (1997). Two groups of endogenous MMTV related retroviral env transcripts expressed in human tissues. Virus Genes *15*, 123-133.
Yang, C., and Robbins, P.D. (2011). The roles of tumor-derived exosomes in cancer pathogenesis. Clin Dev Immunol *2011*, 842849.
You, J., Cozzi, P., Walsh, B., Willcox, M., Kearsley, J., Russell, P., and Li, Y. (2010). Innovative biomarkers for prostate cancer early diagnosis and progression. Crit Rev Oncol Hematol *73*, 10-22.
Zhang, H.G., and Grizzle, W.E. (2014). Exosomes: A Novel Pathway of Local and Distant Intercellular Communication that Facilitates the Growth and Metastasis of Neoplastic Lesions. Am J Pathol *184*, 28-41.
Zhao, J., Rycaj, K., Geng, S., Li, M., Plummer, J.B., Yin, B., Liu, H., Xu, X., Zhang, Y., Yan, Y.*, et al.* (2011). Expression of Human Endogenous Retrovirus Type K Envelope Protein is a Novel Candidate Prognostic Marker for Human Breast Cancer. Genes Cancer *2*, 914-922.
Zhou, F., Krishnamurthy, J., Wei, Y., Li, M., Hunt, K., Johanning, G.L., Cooper, L.J., and Wang-Johanning, F. (2015). Chimeric antigen receptor T cells targeting HERV-K inhibit breast cancer and its metastasis through downregulation of Ras. Oncoimmunology *4*, e1047582.
Zur Hausen, H. (2009). The search for infectious causes of human cancers: where and why. Virology *392*, 1-10.