UNIVERSITÀ DEGLI STUDI
DI MILANO

DOTTORATO DI MEDICINA MOLECOLARE E TRASLAZIONALE

CICLO XXXII

Anno Accademico 2018/2019

TESI DI DOTTORATO DI RICERCA

MED/07

UNRAVELING THE ROLE OF THE HUMAN ENDOGENOUS RETROVIRUSES IN THE PATHOGENESIS OF COLON CANCER

Dottorando: Maria **DOLCI**

Matricola N°: **R11790**

TUTORE: Prof. Pasquale **FERRANTE**
CO-TUTORE: Prof. ssa Serena **DELBUE**

COORDINATORE DEL DOTTORATO: Prof. Michele **SAMAJA**

This thesis respects the four principles of the good research integrity: reliability, honesty, respect, and accountability of The European Code of Conduct for Research Integrity (ALLEA, Berlin, 2018) [1].

“Reliability in ensuring the quality of research reflected in the design, the methodology, the analysis and the use of resources.

Honesty in developing, undertaking, reviewing, reporting and communicating research in a transparent, fair, full and unbiased way.

Respect for colleagues, research participants, society, ecosystems, cultural heritage and the environment.

Accountability for the research from the idea to publication, for its management and organization, for training, supervision and mentoring, and for its wider impacts.”

SOMMARIO

Introduzione: *I retrovirus endogeni umani (HERV) sono elementi originari da infezioni da parte di retrovirus esogeni e rappresentano l'8% del genoma umano. La loro regolazione è basata sulla metilazione a livello dei promotori, le Long Terminal Repeats (LTRs). I trascritti di HERV sono stati associati ai tumori, ma gli studi riguardanti l'espressione degli HERV nel tumore del colon retto rimangono sporadici.*

Metodi: *Sono stati arruolati 63 pazienti italiani e 58 pazienti tunisini affetti da tumore al colon retto ad uno stadio avanzato. Sono stati analizzati, mediante pyrosequenziamento, la metilazione delle LTRs di HERV-H, -K, -P, -R, Alu, LINE-1 e, mediante RT-qPCR, l'espressione del gene env e pol degli HERV nei tessuti tumorali e adiacenti sani e, quando possibile, nel sangue e nelle vescicole extracellulari (EVs) plasmatiche. Sono state inoltre valutate le associazioni tra le caratteristiche cliniche e i livelli di espressione e di metilazione degli HERV. L'espressione delle proteine Pol e Env sono state investigate mediante saggi di Western Blot.*

Risultati: *Nella popolazione italiana, le sequenze Alu, LINE-1 e le LTRs di HERV-H e -K sono ipometilate nel tessuto tumorale rispetto al tessuto adiacente sano ($p<0,05$) mentre non sono state osservate differenze tra le espressioni del gene env degli HERV nei campioni biologici. Il gene env è espresso nelle vescicole extracellulari ($p<0,01$) del 54% (-H), 38% (-K) e 31% (-R) dei pazienti. Sono risultate significative le associazioni tra l'espressione di HERV e la posizione destra del tumore e tra la metilazione degli HERV e l'invasione vascolare ($p<0,05$). La proteina Pol di HERV-K è più espressa nel tessuto adiacente sano rispetto al tessuto tumorale ($p<0,01$). La proteina Env di HERV-K è invece espressa solamente nel tessuto tumorale. Nella popolazione tunisina, le sequenze LINE-1 sono ipometilate nel tessuto tumorale rispetto al tessuto adiacente sano ($p<0,05$), mentre Alu e le LTRs di HERV-H e -K presentano lo stesso andamento di metilazione della popolazione italiana ma la differenza tra i tessuti non è significativa. L'espressione dei geni env e pol sono simili nei campioni biologici. Non sono state rilevate differenze significative tra l'espressione/metilazione e le caratteristiche cliniche dei pazienti tunisini.*

Conclusioni: *Nel tumore al colon retto i cambiamenti nei profili di metilazione degli elementi retrovirali sono specifici ma non correlano con l'overespressione virale. L'espressione della proteina Pol nelle cellule sane potrebbe indurre la retrotrascrizione e il seguente trasferimento delle sequenze virali in altre cellule, forse attraverso le EVs. L'integrazione degli HERV nel genoma potrebbe causare la trasformazione cellulare. Nelle cellule tumorali, la proteina Env potrebbe contribuire alla progressione tumorale mediante la fusione cellula-cellula.*

ABSTRACT

Background: *Human endogenous retroviruses (HERV) are remnants of exogenous retroviral infections, representing 8% of the human genome. Their regulation is based on the DNA methylation of promoters, the long terminal repeats (LTRs). Transcripts from HERV have been associated with cancers, but reports concerning HERV expression in colorectal cancer remain sporadic.*

Methods: *63 Italian patients and 58 Tunisian patients with advanced stages of colorectal cancer were enrolled in this study. HERV-H, -K, -R, -P LTRs, and Alu, LINE-1 methylation levels, and the expressions of HERV env and pol gene were investigated by pyrosequencing and by RT qPCR, respectively, in the tumor, normal adjacent tissues, and, when possible, in the blood and plasmatic extracellular vesicles (EVs). Associations among clinical characteristics and HERV expression and methylation levels were also evaluated. The expression of the HERV-K Pol/Env proteins was also evaluated by Western Blot.*

Results: *As for the Italians patients, Alu, LINE-1, HERV-H and -K LTRs were demethylated in the tumor tissues compared to the normal adjacent tissues ($p<0.05$), while no differences were observed in HERV env gene expression levels, among the clinical specimens. The env gene was expressed in the EVs ($p<0.01$) of 54% (-H), 38% (-K), 31% (-R) patients. Associations were found between HERV expression and right tumor colon location and between HERV methylation and vascular invasion ($p<0.05$). HERV K Pol protein was more expressed ($p<0.01$) in the adjacent normal tissues compared to the tumor tissues. The Env protein was only expressed in the tumor tissue. As for the Tunisian population, LINE-1 was less methylated in the tumor tissue compared to the adjacent normal tissue ($p<0.05$), while Alu, HERV-K and -H LTRs showed the same trends, but the difference was not significant. The expression of the HERV env and pol genes were similar in the biological samples. No association was found between the HERV expression/methylation and the clinical characteristics of the Tunisian patients.*

Conclusions: *The changes in DNA methylation of retroelements are specific in colorectal cancer but do not correlate with viral overexpression. The Pol protein expression in the normal cells may induce the retrotrascription and the subsequent transfer of HERV sequences into other cells, possibly through EVs. HERV genome insertion might cause cells transformation. In the cancer cells, the Env protein may contribute to the cancer progression through cell to cell fusion.*

INDEX

LIST OF SYMBOLS

5-FU	5-Fluorouracil
AA	Amino Acids
ABs	Apoptotic Bodies
APC	Antigen Presenting Cells
BIO	Biotin
CA 19-9	Carbohydrate Antigen 19-9
CEA	Carcino-Embryonal Antigen
CIMP	CpG island methylator phenotype
CIN	Chromosomal Instability
CNS	Central Nervous System
CSF	Cerebro Spinal Fluid
Ct	Cycle Threshold
dsDNA	Double stranded DNA
EBV	Epstein Barr Virus
EDTA	EthyleneDiamineTetraacetic Acid
EGFRvIII	Epithelial Growth Factor Receptor vIII
Env	Envelope
ERV	Endogenous Retroviruses
EtOH	Ethanol
EVs	Extracellular Vesicles
Gag	Group-specific antigen
GAGE6	G Antigen 6
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
HDAC	Hystone Deacetylase
HERV	Human Endogenous Retroviruses
HNPCC	Hereditary NonPolyposis Colorectal Cancer
HSV-1	Herpes Simplex Virus 1
HTLV-1	Human T-Lymphotropic Virus-1

IFN	Interferon
IFN-γ	Interferon-γ
IL-1α	Interleukin 1 alpha
IL-1β	Interleukin 1 beta
IPSCs	Induced Pluripotent Stem Cells
KLF4	Kruppel-like factor 4
LS	Left-sided
LTRs	Long Terminal Repeats
MAP	Mitogen activated protein
mRNA	Messenger RNA
MS	Multiple Scelerosis
MSI	Microsatellite Instability
MSRV	Multiple Sclerosis Associated Retrovirus
MVBs	MultiVescicular Bodies
ncRNA	NonCoding RNA
Oct3/4	Octamer-binding transcription factor 3/4
ORFs	Open Reading Frames
PBMCs	Peripheral Blood Mononuclear Cells
PBS	Phosphate-buffered saline
PCR	Polymerase Chain Reaction
PMA	Phorbol-12-Myristate-13-Acetate
Pol	Polymerase
Pro	Protease
PS	Phosphatidylserine
PSF	Protein-associated Splicing Factor
PTM	Post Translational Modification
PSQ plate	Pyromark Q96 HS Plate
RNA Seq	RNA sequencing
RS	Right-sided

RT	Reverse Transcription
RT-PCR	Reverse Transcription-Polymerase Chain Reaction
SCZ	Schizophrenia
SOX-2	sex determining region Y-box 2
SU subunit	Surface subunit
T-reg	Regulatory T cells
TAG-72	Tumor-Associated Glycoprotein-72
TE	Transposable Elements
TM subunit	Transmembrane subunit
TNM	Tumors/Nodes/Metastases
TPS	Tissue Polypeptide Specific Antigen
U.S.	Unites States
VHL	Von Hippel-Lindau
VPT	Vacum Prep Tool

LIST OF FIGURES

LIST OF TABLES

Table B1: *Composition of bisulfite PCR mix.*

Table B2: *Thermal profile of HERV-H, -K. -P, and -W LTRs PCR.*

Table B3: *Thermal profile of Alu and LINE-1 PCR.*

1. INTRODUCTION

1.1 The Human Endogenous Retroviruses (HERV)

1.1.1 The transposable elements (TE)

The transposable elements (TE) represent the 45% of the human genome: they are DNA fragments that can transpose in the genome and at least the 90% of these sequences is repetitive elements [2]. TE include the non-Long Terminal Repeats (LTRs) retrotransposons and the LTRs retroelements. The long-interspersed elements (LINEs) and short interspersed elements (SINEs) belong to the non-LTRs family. These elements act as an insertional mutagen, they can affect the gene function and genome integrity by recombination-mediated gene rearrangements, genetic instability, transcriptional interference, alternative splicing, gene breaking, epigenetic effects [3]. Instead the LTRs retroelements are repeated sequences that flank internal coding region; they can be further divided in different groups: LTRs bounded elements, endogenous retrovirus (ERV), and LTRs retrotransposons [4]. ERV can be divided into three different classes (class I, II and III) according to the classification of retroviruses [5].

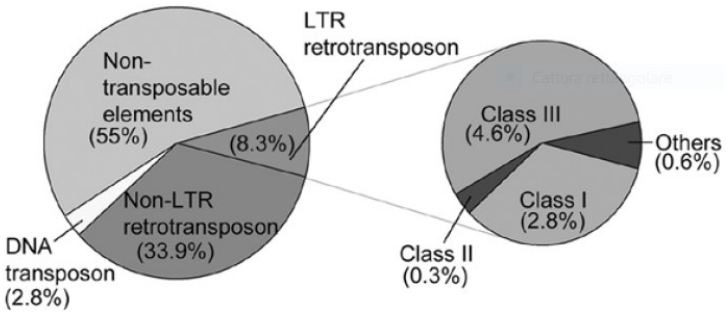


Figure 1: *The classification of TE in the human genome* [5].

1.1.2 Origin of HERV

The human endogenous retroviruses (HERV) are remnants of ancient exogenous retroviral infections [6] occurring millions of years ago. They were

described in the late 1960s [7]. All the human exogenous retroviruses present an horizontal transmission from an infected subject to a new host, integrating into the host genome, while the retroviruses that generate HERV infected also the germ cell lines [8]. The integrated HERV sequences into the germ cell line were inherited in a Mendelian way, being vertically transmitted to the offspring [8]. During the evolution, germ-line integrated proviruses can either expand in number within the germ-line and in the population, achieving fixation, or become extinct by random events or selection pressure against them [4]. HERV are integrated in the human genome as a proviruses [9] and represent the 8% of our genome [10]. The majority of HERV accumulated mutations inducing the silencing of the HERV expression. In other cases, the homologous LTRs recombination leads to the elimination of the internal HERV sequences, leaving solitary LTRs [8] (Figure 2).

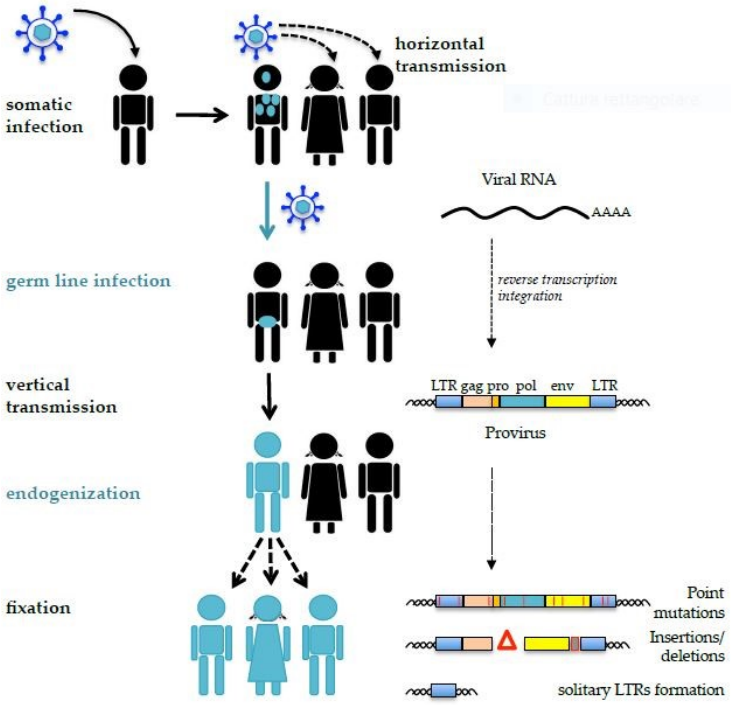


Figure 2: *HERV fixation and transmission into the human genome* [8].

1.1.3 HERV families

There is no standard nomenclature for HERV: a classification in three different classes (Class I, Class II and Class III) based on sequences similarity to different genera of exogenous retroviruses has been used [11] (Table 1).

Table 1: *The HERV families' classification* [12].

HERV classes	HERV families
I	HERV-H, -F, -W, -R, -P, -E, -I, -T, ERV-FTD, ERV-FRD
II	HERV-K, HML-1, -2, -3, -4, -5, -6, -7, -8, -9, -10, -11
III	HERV-L

HERV of the Class I are similar to the gamma and epsilon retroviruses, the Class II are similar to beta retroviruses, and HERV with the homology to spumaviruses are classified in the Class III [4]. In the three classes, HERV can be divided in different families, at least 40 families [13] [14], named after the amino acid carried by the tRNA complementary to the primer binding site of the HERV genome, for examples HERV-H (Histidine), -K (Lysine), -P (Proline), -R (Arginine), and -W (Tryptophan) [2]. The HERV-K family is the most studied, because it is the only known HERV that contains intact open reading frames (ORFs) coding for the retroviral proteins. HERV-K family are characterized by 11 subgroups, named from HML-1 to HML-11, and the most famous is HML-2 that seems to be involved in several types of cancer [15]. Furthermore, HML-2 can be divided into two different proviruses, the type I and the type II, and the difference is due to the presence (in HML-2 type I) or absence (in HML-2 type II) of a 292 base pairs deletion in the *pol* and *env* gene sequence.

1.1.4 HERV genome structure

The HERV genome is like the other retroviruses' genome and their length is approximately of 9000 base pair (bp) [6] (Figure 3): the genomic structure is 5'LTRs-*gag-pro-pol-env*-3'LTRs [16].

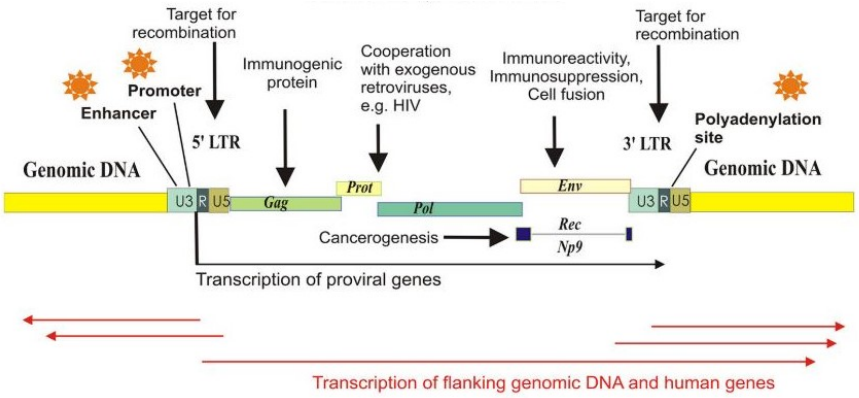


Figure 3: *HERV genomic and functional structure* [17]. *The genome is composed of four encoding genes: gag, pro, pol and env genes. These genes are flanked at 5' and 3' by two regulatory regions, named Long Terminal Repeats.*

The group specific antigen *(gag)* gene contains the sequences to form the structural protein matrix, the capsid and nucleocapsid of HERV virion [18].

The polymerase *(pol)* gene forms the reverse transcriptase (RT), with RNase H and polymerase domains, and the integrase. The RT protein leads to the reverse transcription of the RNA into single strand DNA, while the RNA in the RNA–DNA heteroduplex is degraded by the RNase H. Instead, the polymerase domain synthetizes the complementary strand of DNA, that is integrated into the host genome by the integrase [18].

The envelope *(env)* gene codes for the Env protein that mediates the fusion and the anchoring of the virions to the host cells; it is composed of three domains: the signal peptide, the surface unit and finally the transmembrane domain [18].

The others HERV's transcripts are *Rec*, that is produce only by the HML-2 type II provirus and the Rec protein has 87 amino acids (aa) in common with

the Env protein, and *Np9*, that is coded only by the type-I of HML-2 [19]. Rec protein interacts with several cellular proteins, for example with the promyelocytic zinc finger protein (PLZF), testicular zinc finger protein (TZFP). Np9 protein can interact with PLZF. These interactions may play an important role in the cells and they may contribute to tumor development [20].

These genes are flanked at 5' and at 3' by two regulatory regions, named Long Terminal Repeats (LTRs) (with a length of 1000 bp), that are composed of the upstream U3 region, a repeat region(R), and the downstream U5 region. U3 is characterized by enhancers and promoter sequences, driving the viral transcription. R region encodes the 5′ capping sequences and the polyadenylation signal. The HERV LTRs can affect the adjacent genes: LTRs act as a promoter [21], enhancer [22], polyadenylation signals [23] and binding sites for many nuclear proteins [24]. Although the majority of HERV genes are highly defective with large deletions, stop codons and frameshift in the ORFs, HERV LTRs still retain their function and potentially regulate their neighbouring viral and cellular genes [25].

1.1.5 The exogenous and endogenous retroviruses life cycle

The exogenous retroviruses entry into the cells using specific receptors in the surface that mediate the fusion with the target cells. The specificity of the retrovirus for the cells is determined by the envelope proteins. In the cytosol, the internal core is released, and the single-stranded positive-sense linear RNA is reverse transcripted into single strand DNA, using the RT; the RNA is degraded by RNaseH and the polymerase synthetizes the complementary DNA strand, producing dsDNA. Finally, the dsDNA is integrated into the cellular genome through the integrase. The expression of mRNA from the provirus produces both new genomic RNAs both the synthesis of viral structural proteins, which are assembled into new virions. The new virions are released from the membrane of the host cell [4]. HERV are not able to

create complete virions and to infect the cells due to the accumulation of mutations and deletions in their genes [26]. The majority of the HERV families can only transcribe the DNA into RNA, producing some transcripts and proteins. HML-2 presents functional ORFs and it can produce viral particle, even if its replication is limited [6].

1.1.6 HERV regulation and activation: host and external factors

The HERV expression depends on host factor, such as the regulation of the LTRs [27], but it also depends on external factors, such as environmental factors (Figure 4).

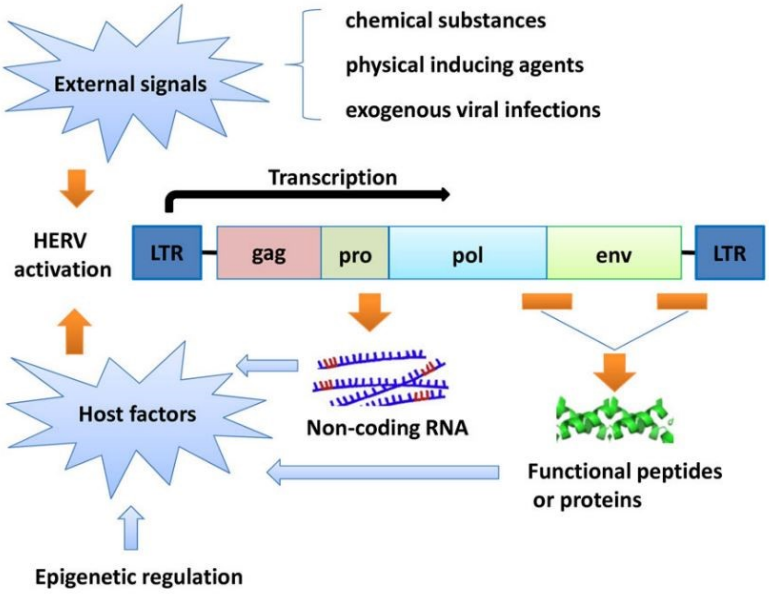


Figure 4: *Mechanisms of HERV regulation and activation* [16]*: host factors and external signals.*

The possible external factors involved in HERV regulation are external chemical substances, for example the treatment with phorbol-12-myristate-13-acetate (PMA) [16] and physical agents, such as UVB irradiation, and X-ray that can act as a promoter of the HERV expression [16]. In fact, stimulation with PMA of the U937 monocytoid cells showed increase of the

HERV-W, -K, and -H RNA levels [28]. Lee et al., by means of quantitative reverse transcription-polymerase chain reaction (RT-PCR), analysed the HERV-R *env* gene RNA in the cells irradiated with the γ-ray compared to the normal cells: they observed that there was an alteration in the epigenetic pattern that regulated the LTRs, inducing the HERV over-expression [29].
Moreover, the exogenous viral infections, such as HIV, can influence the HERV expression. In a study conducted *in vitro* and *in vivo* by Yamamura and colleague demonstrated that HERV-K transcription was related to the HIV. In fact, the HERV-K RNA and protein expression were higher in the HIV-infected-H9 cells compared to the H9 cells [30]. Moreover, in HIV patients, the HERV-K RNA expression was higher in the CD4 positive T cell compared to the healthy subjects [30]. Also, the immediate early protein of Herpes Simplex Virus 1 (HSV-1), ICP0, transactivated the HERV-K transcription, through the AP-1 binding site located in the HERV-K LTRs: in fact, the ICP0 can upregulate the activity of this site, increasing the HERV-K expression [31]. In the same study, it was observed that the deletion of AP-1 site did not allow the binding of the HSV-1. Also the Human T-Lymphotropic Virus-1 (HTLV-1), and Epstein Barr Virus (EBV) seem to be involved in the activation of HERV expression, showing similarities with the LTRs [32].
Although the majority of HERV are not able to produce complete virions, HERV LTRs maintain their functions and the HERV expression is regulated by epigenetic mechanisms, such as DNA methylation and histone modification, but also by proteins that act as a transcription factors, like cytokines, and small RNA [16]. The methylation is an epigenetic modification that allows altering the gene expression, the maturation of transcripts, and the regulation of protein functions, without altering the DNA sequences. Szpakowski and colleagues demonstrated that, analysing the different HERV families, HERV were more methylated in the healthy tissues [33] compared to the tumor tissues. Moreover, the HERV's age are correlated with their

methylation status: in the older families, HERV are less methylated respect with the younger families [33]. Loss of the methylation leads the alteration in the HERV expression, especially inducing the HERV overexpression, that can play an important role in the cancer pathogenesis. To this regard, Szpakowski and colleagues, analysed the methylation pattern in 33 tumor and 17 adjacent normal tissue, collected from patients affected by head and neck carcinoma. Using genome-wide microarray assay, HERV-H, HERV-E, HERV-K, and LINE-1 and Alu families showed a loss of methylation in the tumor tissue compared to the normal adjacent tissue [33].

The HERV expression may be also regulated by the histone modification, but there are very few studies regarding these findings and the obtained data are opposed. The histone modification is a covalent post-translational modification (PTM) of the histone proteins that alters chromatin structure, inducing the alteration of gene expression. Hurst and colleagues demonstrated that in HIV-1 latency T cell treated with Histone Deacetylase (HDAC) inhibitors, the HML-2 *env* and *pol* gene, and the *env* gene of HERV-W expression were not influenced by the treatment [34]. In contrast, White et al. analysed the HERV transcripts by means of RNA-Seq in the CD4 positive T cells after vorinostat treatment. They assessed that the use of the HDAC inhibitor-vorinostat in the HIV treatment may affects the HERV expression [35]. According to this study, Daskalakis observed that the use of epigenetic therapies, such as HDAC inhibitor, both *in vitro* and *in vivo,* could activate the HERV expression [36].

The HERV transcription can be also regulated by transcriptional factors [16]. In an *in vitro* study conducted by Katsumata and colleagues, it was reported that HERV-R *env* gene expression was controlled by some cytokines: the data were obtained by the stimulation of the human vascular endothelial and smooth muscle cells with cytokines: the *env* gene was over-expressed after

Tumor Necrosis factor (TNF-α), Interleukine-1alpha (IL-1α) and IL-1 beta (IL-β), while the treatment with IFN-γ induced downregulation [37].

1.1.7 The HERV expression in the healthy subjects

There are very few studies regarding the HERV expression in the healthy subjects. In 2015, Sinibaldi-Vallebona and colleagues investigated the changes in HERV expression related to age in the peripheral blood mononuclear cells (PBMCs) of 261 healthy subject [9]. The enrolled population was aged from 1 to 80 years old. They analysed the expression of HERV-H, -K, and -W by means of Real Time PCR: the HERV-H and -W *env* gene showed highest level in the young subjects, a decrease of expression until 40 years old and finally an increase in the population over 40 years old [9]. As for HERV-K, the *env* gene was less expressed in the younger individuals meanwhile the expression was higher in the older ones [9].

The difference of HERV expression related to the age in healthy subjects was also investigated by Nevalainen et al. They analysed the expression of 91 proviruses of HERV-K and 213 proviruses of HERV-W in 7 young adults and 7 nonagenarians by means of RNA-sequencing (RNA-seq): no significant differences were observed in the two groups, even if the expression level seems to follow an aging-related pattern [38].

Other study investigated for the first time the polymorphisms of HERV-K 6 and HERV-K 11 in the blood samples collected from 18 healthy subjects (10-79 years old, 8 males and 10 females) to analysed the HERV-K movements in individuals, by means of inter-retrotransposon amplified polymorphism PCR: as for HERV-K 6, the percentage of polymorphisms was between the 0-70%, while for HERV-K 11 it was between the 0-38%. No association was observed between the polymorphism and the age [39].

The expression of the HERV-K *env* mRNA, especially the *np9* and *rec* transcripts, and the capability of these two mRNAs to code for the proteins were analysed in various normal tissues collected from 16 healthy subjects by means of PCR. Schmitt and colleagues observed that transcripts were expressed in all the normal tissues and that some alternative splicing events of the two mRNAs produced proteins. These results suggested the presence of Rec and Np9 proteins as "a part of the proteome of normal human tissues" [20].

1.1.8 The positive side of HERV

There is a positive side of the HERV expression, especially the HERV *env* gene is involved in the development of the syncytiotrophoblast: the syncytiotrophoblast is the outer layer of the human placenta, bound to the endometrium, that allows the exchange of many components from the maternal blood to the fetal circulation, contributing to the feto-maternal immunotolerance [40]. The HERV Env protein seems to play an important role in the syncytiotrophoblast formation. In fact, the *env* gene of HERV-W, located into the locus 7q21.2, encodes for the protein syncytin-1 [41] that mediated the fusion and the formation of the syncytium, occurring in the placenta morphogenesis [42]. Moreover, the syncytin-1 regulates the trophoblast homeostasis: especially, it is involved in the differentiation, proliferation and survival of the cytotrophoblast, the inner part of the trophoblast [43]. These findings were observed by Frendo and colleagues, that evaluated the role of the HERV-W *env* gene in the differentiation of primary cultures of human villous cytotrophoblasts.

In addition, HERV-FRD, integrated into the locus 6p24.1, encodes for other protein involved in the placenta development, expressed only in the villous cytotrophoblasts [44], named syncytin-2 [45]; it induces the fusion of single cytotrophoblasts into the syncytiotrophoblast.

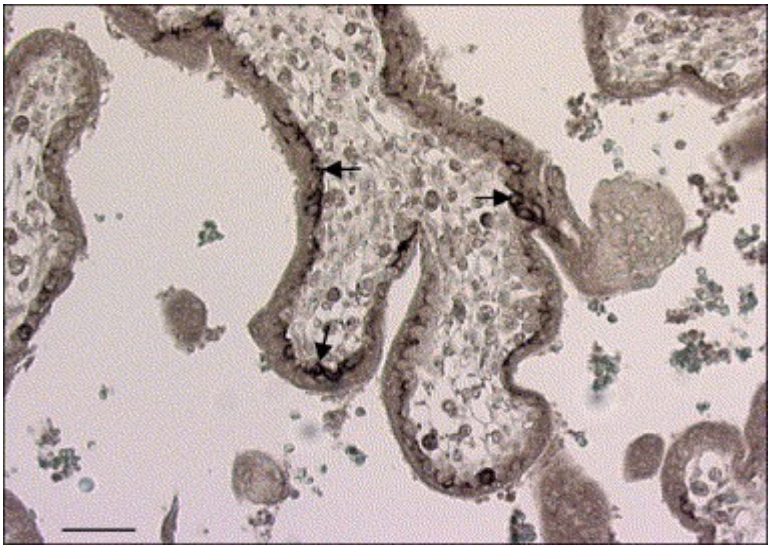


Figure 5: *Syncytin 2 protein expression in the villous trophoblast of placenta, analysed by means of immunohistochemical (16 weeks of gestation). The black arrows show the presence of the protein* [44].

The two proteins present a similar structure: they are about 73 kDalton glycosylated protein with two subunits, a surface (SU) and transmembrane subunit (TM). SU leads cell-cell fusion and the syncytia formation, while the TM, with immunosuppressive activity, can induce the escape of the tumor cells from the immune system. TM inhibits the production of the maternal T_H1 cytokines, such as TNF-α, IFN-γ (inflammatory pathway), switching to the T_H2 cytokines production, such as IL-4, IL-5, and IL-10 (anti-inflammatory pathway) [45].

Moreover, HERV-K plays an important role in the physiological activities of the cells, as reported in the study conducted by Suntsova et al.: HERV-K acts as an enhancer of the PRODH gene, involved in the catabolism of the proline, with some important function in the central nervous system (CNS), in particular in the hippocampus. The regulation of PRODH gene expression is due to the methylation status both of the PRODH promoter both of HERV-K, that act synergistically [46].

Ohnuk and colleagues demonstrated that HERV-H was hyperactivated only in the reprogramming phase of the induced pluripotent stem cells (iPSCs) and then its activities was decreased. They reported that the important transcription factors in the establish of the pluripotence in the stem cells, the POU transcription factor Oct-3/4 (Oct-3/4), Sex-determining region Y (SRY)-Box2 (SOX-2), and Krüppel-like factor 4 (KLF4) were able to bind and activate the HERV-H LTRs. These results suggest the role of HERV-H in the formation of the iPSCs [47].

1.1.9 The negative side of HERV: their role in neurologic diseases

HERV play an important role in the pathogenesis of the neurologic disease, such as multiple sclerosis (MS), schizophrenia (SCZ), and HIV associated dementia. The retrotrascription of HERV into DNA or the expression of HERV proteins produce viral signal, like the pathogen-associated molecular pattern of exogenous retroviruses, stimulating the innate immune response. But the pattern recognition receptor (PRR) activation shows implications for several inflammatory and immune-mediated diseases, such as MS. In fact, the immune responses to HERV, both humoral and cell-mediated, are increased in MS, and the target, the HERV antigen, are more expressed on peripheral blood mononuclear cells from patients with MS [48].

As for MS, HERV-H/F and HERV-W/Multiple Sclerosis Associated Retrovirus (MSRV) seem to be a risk factors and they seem to be involved in the pathogenesis of this disease [49]. In a study conducted on Sardinia population, the presence of MSRV transcripts in the cerebrospinal fluid (CSF) of MS patients presented a strong association with a severe disease course: MSRV transcript can be considered as a prognostic marker for the MS pathogenesis [50]. Moreover, the single-nucleotide polymorphism of HERV-Fc1 on the chromosome X were associated to this disease [51]. The HERV-W Env protein were also detected in the brain of MS patients, while the

expression of HERV-H, -K, and -W were increased [28]. Moreover, the syncytin 1 protein, expressed in the CNS of this patients, can regulate the neuroinflammation [16]: it mediates the oligodendrocyte cell death [52] and the reduction of the oligodendroglial differentiation capacity [53]. Also antibodies direct against viral antigens, such as HERV and some herpesvirus in the CSF of MS patients were found, in particular against HERV-W Env and HERV-H Env epitopes [49].

1.2 HERV and cancer

The existence of a strong relationship between HERV expression and cancer, based on the mRNA and proteins expression profile of HERV in normal and tumor tissues has been suggested [54], but conclusive evidence is still lacking. In particular abnormal expression of HERV is reported in literature for different types of cancer, such as melanoma, breast cancer, leukemia/ lymphoma, astrocytoma, prostate cancer, lung cancer, pancreatic cancer, gastro-intestinal cancer, ovarian cancer, endometrial and testicular cancer [2], as reported in the Table 2.

Table 2: *mRNA and proteins' HERV expression in different type of tumor* [2].

Tumors	HERV	RNA and proteins	References
Melanoma	-K -K -K -K -K	Protein (gag, env, rec) Protein (gag, env) RNA, Protein (gag, env, rec) RNA, Protein (env, rec, np9) RNA, Protein (env)	[55] [56] [57] [58] [59]
Breast	-K -K, E, F, W, T, FRD -K -K	Protein (gag) RNA (pol) RNA (env) RNA, Protein (gag, pol, env) RNA (gag)	[60] [61] [62] [63] [64] [65]
Leukemia/ lymphoma	-K -K,-H -K -K -K -E -H	RNA, Protein (gag, env) Protein (gag) RNA (gag) RNA (pol, env) RNA (LTRs) RNA (gag, pol, env) RNA (gag, env)	[62] [60] [56] [66] [67] [68] [69]
Astrocytoma	-K	Protein (env)	[70]
Prostate	-K -E, -R	RNA, Protein (gag) RNA (env)	[56] [63]
Lung	-K -E -R	Protein (gag) RNA (LTRs) RNA (env)	[56] [71] [72]
Pancreatic	-K -H	RNA (env) RNA (gag)	[73] [74]
Gastro- intestinal	-K -K -H	Protein (gag) RNA (env) RNA (gag)	[56] [75] [10]
Ovarian	-K -K, -E, -R -E -K -H	RNA, Protein (gag) RNA, Protein (gag) RNA (negative strain) Protein (gag) RNA (LTR)	[56] [76] [77] [78] [79]
Endometrial	-W	Protein (env)	[80]
Testicula r, germ cells	-K -K -K, -H -K	Protein (gag, env) RNA (gag) RNA (LTRs) Protein (gag)	[81] [82] [83] [60]

1.2.1 HERV: mechanisms of oncogenesis

HERV can be involved in oncogenesis with many mechanisms, as reported in the Figure 6.

- HERV can produce Long non-coding RNA, inducing gene silencing;
- HERV LTRs cause genomic instability and can regulate and modify the transcription of the adjacent genes, included the oncogenes or tumor-suppressor genes;
- HERV mediate insertional mutagenesis due to the non-allelic-recombination;
- HERV encode for two oncogenic proteins, Rec and Np9;
- HERV can induce the cell fusion, promoting the cell growth;
- HERV sequences are also the cargo of the tumor-derived retroviral-like microvesicles, involved in the tumour progression;
- finally, HERV can regulate the innate immunity.

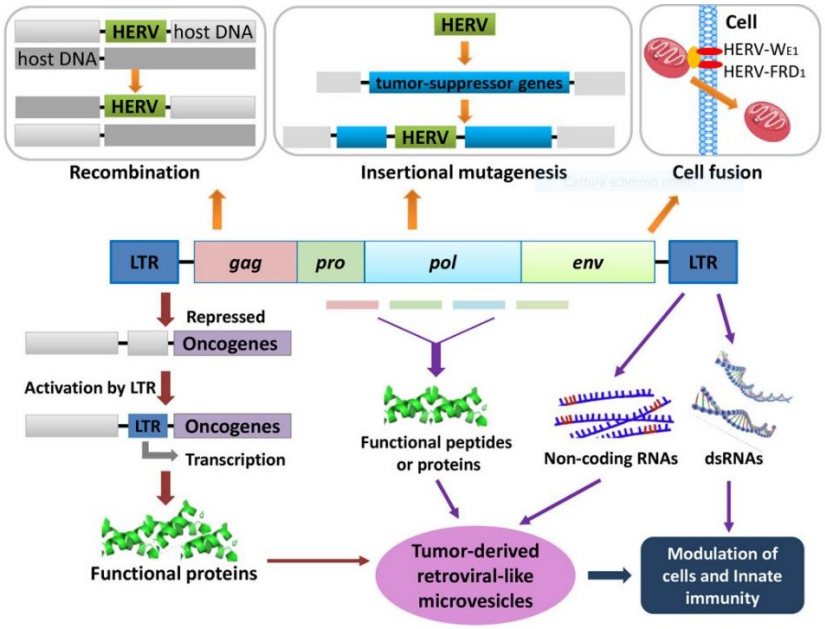


Figure 6: *The different mechanisms of HERV involved in the oncogenesis* [16].

HERV accumulate several mutations in the ORFs and most of them are transcriptionally silenced or they can produce non-coding RNAs (ncRNAs). The ncRNAs are small RNA sequences that are not translated into proteins. The oligonucleotide may bind and repress the transcription of the target genes. For example, MER11C (belonging to the HERV-K 11 group) codes for a ncRNA that binds the pre-mRNA of the Polypyrimidine tract-binding protein-associated splicing factor (PSF) factor. In normal condition, the PSF factor inhibits of the proto-oncogene G Antigen 6 (GAGE6) transcription[84]; but when PSF is bounded to the ncRNAs, GAGE6 promotes the oncogenic activities, especially it induces the hepatocellular carcinoma cells proliferation [85].

The LTRs and their regulation can be implicated in the malignant transformation. HERV LTRs may cause genome instability, for example through the insertional mutagenesis. In fact, the LTRs elements can be inserted into genomic sequences with an important biological and physiological functions, contributing to the tumor growth [5].

Alteration in the HERV LTRs' methylation is also an important mechanism in the development and progression of the tumor: it can be a peculiarity of the tumor or can be a result of the wave global demethylation, occurring in several tumors; the loss of methylation in the LTRs induces an increase of HERV expression in the pathologic tissue compared to the normal tissue [86]. In addition, the hypomethylated LTRs, acting as a promoters, can alter also the expression of host cellular genes [87]. The activation of the HERV LTRs may affect adjacent oncogenic sequences that in normal condition are subjected to silencing. For instance, the colony-stimulating factor 1 receptor (CSF1R) oncogene is restricted to the myeloid lineage, but this restriction is changed in Hodgkin lymphoma through the transcriptional activation of a normally silenced HERV LTR as an alternative promoter [88].

LTRs can also donate binding site for the cellular transcription factors or other regulatory proteins [16]. For example, some HERV LTRs (especially the LTR10 and MER61 families) are enriched of the p53 DNA binding sites. These sequences can alter the p53 expression, a tumor suppressor gene, leading the transcription of the p53 gene, contributing to carcinogenesis [89]. The non-allelic recombination between HERV sequences produces duplications, deletions, and other chromosomal rearrangements [90] that induce the activation of oncogenes and/ or the silencing of the tumor-suppressor genes. For example, in some prostate cancer cases, the HERV-K integrated into the chromosome 22q11.23 with 5′-LTR-UTR is translocated upstream of the transcription factor ETS translocation variant 1 (ETV1), a mitogen-activated protein (MAP) kinase; the translocation leads high level of expression of the ETV1 oncogene, especially in prostate cancer [91].

The two oncogenic proteins Rec and Np9, encoded by HML-2, interact with important cellular proteins and may contribute to tumor development and progression [20]. Rec is a functional homologue of the Rev and Rex protein of HIV and HTLV1 respectively, while Np9 is a nuclear protein. Both Rec and Np9 protein interact with several cellular proteins, for example with the promyelocytic zinc finger protein (PLZF), a transcriptional repressor involved in cell cycle control [92]; their binding causes the blocking of the repression of the proto-oncogene c-myc by PLZF, inducing the cell proliferation [93]. The *in vitro* coimmunoprecipitation of the HERV proteins and PLZF factor confirmed the binding between the proteins. The Rec expression was observed in the metastatic melanoma biopsies, and in the melanoma cell lines but not in normal melanocytes [58], instead the Np9 protein was detected in a cells derived from a Burkitt's lymphoma, and in an EBV-transformed human lymphoblastoid cell line [94].

Other mechanism involves in the development of the cancer is the HERV-mediated cell fusion. HERV-W encodes for the syncytin-1 protein that is

involved in the fusion of cells in placentas, an important step in the formation of the syncytiotrophoblast [40]. This protein is also involved in the promotion of the malignant transformation. In a study conducted by Strick et al., it was observed in the endometrial carcinomas that, after hormone treatments, the syncytin-1 protein expression was up-regulated and it mediated the fusion of the tumor cells, in a similar manner as in the syncytiotrophoblast formation; the role of syncytin-1 was also confirmed by its silencing [80]. Other protein, homologous to the syncytin-1, mediates the mitochondria transfer cell to cell, promoting the survival of the tumor cells and the cancer growth, as reported in a study conducted by Huang and colleagues: HERV-K Env protein leads the cell to cell fusion in melanoma. The HERV-K Env protein, with a fusogenic activity, could mediates the fusion of the melanoma tumor cells and the fusion of the melanoma tumor cells with the target cells, promoting also the metastasis process [95].

HERV mRNA can be transferred in the extracellular vesicle (EVs) mediating the cells transformation. EVs are part of an important mechanism of intercellular communication: EVs are composed of exosomes and microvesicles, secreted in plasma, containing also RNA. EVs play important roles in many diseases, such as cancer: several studies have implicated EVs in driving the formation of a pre-metastatic tumour niche [96] and in stimulating tumour progression [96]. In a study conducted by Skong and colleagues, it has been demonstrated that tumor microvesicles contain also retrotransposon elements [97] but the role of HERV in these vesicles is still unclear. In several types of cancer, such as breast cancer, melanoma and teratoma the use of therapeutical strategies reduces the HERV RT proteins and retroviral like particles, according to the role of HERV in promoting the cancer [16].

HERV are also able to regulate the immune system, producing two different effects on cancer: they are involved both in the oncogenic processes in

anticancer defences [16]. Some HERV Env proteins have immunosuppressive activities due to a conserved domain in the TM subunit. Especially, the human syncytin-2 can regulated the immune system but the role of syncytin-1 is not clear; in fact, syncytin-1 presents a lack in the aminoacids sequence of the TM [40]. Moreover, it has been recently demonstrated by Kudo-Saito et al. that HERV-H produces H17 peptide with immunosuppressive properties: it caused the epithelial-mesenchymal transition and stimulated the expression of the chemokine CCL19 in the cancer cells. The expression of the CCL19 induced the recruitment and the expansion of the CD271 positive cells, a multipotential mesenchymal stromal cells [98].

1.2.2 HERV and anti-oncogenic properties: immunologic target of the tumor

The immunotherapy is the one of the most successful cancer therapy: it is characterized by the reinforce of the adaptive immunity against specific antigens expressed by the tumor [99]. In cancers such as melanomas, the HERV genes (especially HERV-K) may be released from epigenetic regulation and expression of HERV sequences are upregulated. In the cancer cells, the expression of HERV proteins can act as tumor antigens and induce the T-cell and the B-cell responses against HERV-encoded antigens [100]. In melanoma [101], seminoma [81], renal cell carcinoma [102], and colorectal carcinoma [103] there are examples of proteins/antigens encoded by HERV and targeted by the T cells, although their anti-oncogenic properties are still unclear [104]. In patients affected by Von Hippel-Lindau (VHL) disease, Harashima and collaborators observed the presence of the Env peptide on the surface of tumor renal cells, offering a potential target for T-cell immunotherapy in the renal cancer [105]. In addition, HERV-specific antibodies are frequently elevated in cancer patients compared to the healthy

subjects [104]. Antibodies direct against HERV-K was detected in the sera of patients affected by germ cell cancers [60], prostate cancer [56], melanoma [57] and breast cancer [106]. Wang-Johanning et al. investigated the HERV-K Env protein expression in malignant and normal human breast cell lines and they targeted the protein with monoclonal antibodies to evaluate the role of the Env protein in the tumor growth both *in vivo* and *in vitro.* The use of the monoclonal antibodies stimulated the apoptosis and the reduction of the tumor growth, both *in vivo* and *in vitro* [107]. They supposed a similar effect of some antibodies in breast cancer patients.

1.2.3 HERV as a diagnostic and prognostic marker of the tumor

HERV could be used as potential biological diagnostic and prognostic markers of the tumor [108]. Kleiman and colleagues examined the serum titres of antibodies anti HML-2 Gag and Env proteins in germ cells tumor. They observed a link between the clinical manifestations of the tumor and chemotherapy success with the antibodies titre, showing an important diagnostic role [109]. For instance, a study conducted by Golan et al. analysed the HERV RT protein expression in 100 breast cancer patients with 5 to 10 years of follow-up: it has been reported a negative significant correlation between HERV transcription and disease-free survival. Moreover, the RT protein overexpression was linked to the presence of metastases, showing a possible prognostic role [64]. The prognostic role of HERV-K mRNA and protein was also reported by Zhao and colleagues: they analysed the *env* gene and Env protein expression in three cohorts of breast cancer patients, from the United States (U.S.) and China, by means of RT-PCR and immunoistochemistry. They observed in the U.S. cohort that the HERV-K *env* gene and Env protein expression were increased in breast cancer tissue compared to normal adjacent one, while in the healthy controls the levels of HERV expression was undetectable. It was also reported that the HERV-K

presence in the tumor tissue was associated with stage of the pathology and with the lymph node metastasis. In the Chinese cohort, they observed a strong relationship between the HERV-K expression and the tumor size, the tumor stage and the presence of metastasis in the lymph node [110]. The prognostic relevance of HERV-K was also investigated in melanoma patients: Hahn and colleagues analysed the presence of the antibodies directed against HERV-K Gag and Env proteins in 312 sera collected from melanoma patients compared to 70 sera of healthy subjects. 51 patients presented the HERV-K antibodies while no one healthy subject was positive. The presence of HERV-K antibodies correlated statistically with the stage and the location of the tumor; moreover, the HERV-K antibodies were associated with a reduced survival probability in the patients with the tumor at the early stages [111]. The correlation between the HERV mRNA expression, in particular of the *rec* gene, and the melanoma progression was also confirmed by Singh et al. [112].

1.3 HERV and Colon cancer

1.3.1 Epidemiology of colon cancer

Colorectal cancer (CRC) is the fourth malignancy in incidence and mortality in developed countries, with more than 1.8 million new cases in 2018 worldwide [113]. The National Cancer Institute estimates 145,600 new cases and 51,020 deaths in 2019 only in the United States [114]. The incidence and the mortality worldwide until 2012 are represented in the Figure 7 [115].

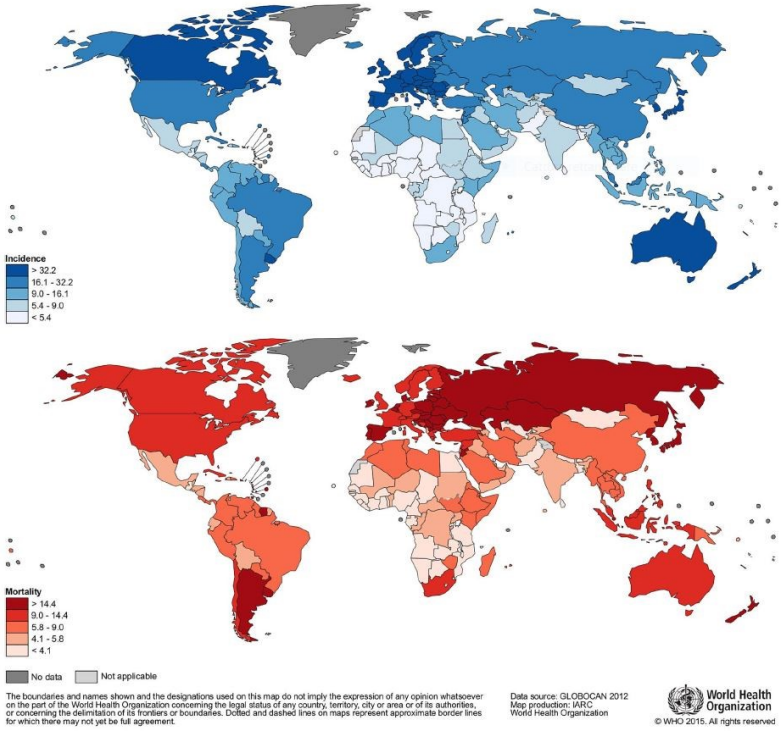


Figure 7: *Incidence and mortality of CRC in the world. The high incidence is in the dark blue countries, while the high mortality is indicated by the red countries* [116].

The difference of CRC incidence in different countries is due to the difference in CRC risk factors such as obesity, physical inactivity, smoking, high alcohol consumption, a diet high in fat and low in fruits and vegetables [117]. Generally, the risk factors of this cancer are associated with the economic development of the countries, which probably explains the highest cases of

CRC in western nations [118], in Canada, Australia, New Zealand and Europe, especially in Czech Republic and Slovakia [117]. In developing countries, such as in Tunisia, the increased of the CRC incidence may be explained by the adoption of a westernized life style, in particular an incorrect diet that induce obesity [119][120]. In fact, in the last years, the obesity increased by 67% in the Tunisian population, multiply by 2.5 times among the adult [117]. The undeveloped countries, such as India, Egypt and central west Africa, present the lowest rates of CRC incidence. CRC is most frequently diagnosed among people aged 65-74 and it affects men more frequently than women (44.2 new cases in men versus 33.9 new cases in woman per 100,000 person) [114]. CRC is the second cause of cancer death in the United States and the number of deaths was 14.2 per 100,000 men and women per year based on 2012-2016 deaths [114], and death rates increase with the age, while the better survival depends on the early diagnosis of the tumor.

1.3.2 Colon cancer types, risk factors, and genetic/epigenetic alterations

Most CRCs are caused by benign pre-neoplastic lesions, the adenomatous polyps or adenomas that progress into a malignant carcinoma through histological, morphological, and genomic mutations [121]: the tumor progression is known as the adenoma-carcinoma sequence model. According to etiological and genetic characteristics of the pathology, CRC is classified into three different groups: sporadic, familial, and hereditary [122]. The sporadic tumors represent about the 70% of CRC and they do not present any familial predisposition. The subjects affected by sporadic carcinoma are older than 50 years old: the main risks for CRC are related to the lifestyle of the population and the aging [122]. For example, some type of diet: several studies reported that a diet high in calories, rich in animal fat

and low in fiber is associated with an increase in intestinal tumors. Other important factors involve in the development of this cancer are also the overweight, the lack of physical inactivity and the smoke [114]. The transition adenoma-carcinoma-sequence in the sporadic CRC is characterized by an accumulation of genetic alteration [123] that induces malignant transformation of the normal tissue. About the 80% of sporadic CRC presents this sequence [124]. The early stages of the CRC are characterized by chromosomal instability (CIN) [124], with the loss-of-function mutations of tumor-suppressor genes such as the *adenomatous polyposis coli* (APC) gene: APC gene is located at the fifth chromosome and it encodes for the APC protein that is associates with β-catenin in the normal cells [125]. APC plays an essential role in suppressing the Wnt signalling pathway, with the binding to β-catenin, that controls the cell growth, the proliferation and the differentiation of the cells in the intestine [126]. Mutation of APC induces the accumulation of β-catenin, that translocates into the nucleus and activates the transcription of c-*MYC* and cyclin D1 gene [124]. In the progression of the tumor, there may be an activation of the K-RAS oncogene due to gain-of-function mutation [127]. K-RAS is involved in the regulation of cell division and its mutations switches K-RAS protein from inactive to active state, inducing the cell transformation and the resistance to chemotherapy [128]. Moreover, a mutation of p53 gene, that encodes for a protein involved in the cell cycle arrest, and apoptosis in case of cellular stress, leads to the final onset of CRC. The accumulation of these mutations in a specific order is essential for CRC progression [124].

The inherited predisposition to CRC is in fewer than 10% of patients: the major manifestation of this disease is the colonic polyps compared to the absence of polyps. The Familial Adenomatous Polyposis (FAP) is an inherited disorder characterized by the growth of multiple benign polyposis in the colon and in the rectum of a younger patients, that may develop

malignancies [129]. In the hereditary non polyposis colon cancer (HNPCC), an autosomal dominant condition, about 2-4% of patients inherit gene mutation, in particular in the MLH1 gene, inducing the onset of the Lynch syndrome [130]. This mutation is due to an alteration in DNA mismatch repair genes, characterized by DNA replication and repair defects [125], that lead to microsatellites instability (MSI) and accumulation of mutations in growth-regulating genes, involved in the CRC pathogenesis [124].

The last group of CRCs is the familial CRC. CRC develops very frequently in these susceptible families to be considered sporadic but does not present the characteristics of the inherited syndrome. About 25% of all cases of CRC may be grouped into this type of cancer [131].

There is also other group of CRCs characterized by the epigenetic alteration, alteration in the methylation status of gene promoters (gene specific methylation) and transposable elements (global methylation) [123]. In the normal cells, CpG islands are usually maintained in an unmethylated state with a gene expression at normal levels or in the methylated promoters with a downregulation of gene expression. Aberrant methylation of gene promoters is common in CRC, with a significant activation or inactivation of the oncogenes and tumor suppressor genes, respectively [132]. The CpG island methylator phenotype (CIMP) pathway is also involved in pathogenesis of CRC: CIMP is characterized by several hypermethylated genes and the tumorigenesis is induce by progressive genetic silencing, also without genetic mutations [124]. This type of CRC are not only induced by genomic instability but also by epigenetic instability [132]. For example, the promoter of MLH1 gene, show an aberrant profile of methylation, that occurs in 80% of sporadic colorectal cancers. Other gene, such as NEUROG1, IGF2, and SOCS1 can be considered as CIMP markers [127].

There are also numerous previous studies that report an epigenetic alteration of global hypomethylation in tumor including CRC, contributing to the

genomic instability [133]. The methylation of Alu and LINE-1 elements is affected in the tumor tissues of CRC patients [134]. The methylation status of LINE-1 sequences is used as marker to the global methylation, representing about the 17% of the human genome [133]. LINE-1 can contribute genome instability through the retrotransposition in the genome due to the overexpression as a consequences of the epigenetic alteration in the cancer [135]; moreover, studies conducted by Ongino and by Sunami reported the relationship between the hypomethylation and worse prognosis of CRC [136] and the correlation between the decrease of LINE-1 methylation and the tumor stages [137].

1.3.3 Pathogenesis, diagnosis and therapy of colorectal cancer

CRC is a disease that affects the last part of the digestive system. Colon or the large intestine is responsible for absorption of water from the indigestible residue of food and it is long about 1.5 metres. The ascending colon is the first part (right side): the material passes through this part and then through the transverse, descending and sigmoid part of the colon (left side), and finally into the rectum, where the waste is expelled out through the anal canal. There are several stages of CRC according to the TNM system (Tumors/Nodes/Metastases): from the stage 0 to the stage IV, as reported in the Figure 8.

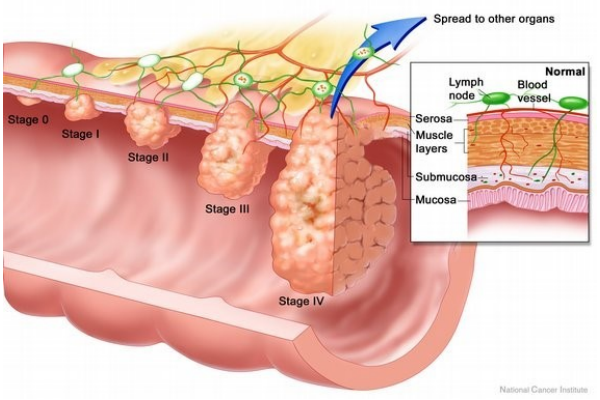


Figure 8: *Colorectal cancer stages. Stage 0: the tumor involves superficially the mucosa. Stage I: the tumor invades the mucosa and submucosa. Stage II: the tumor invades the muscle layers. Stage III: the tumor involves the lymph nodes. Stage IV: the tumor induces metastasis.*

The stage 0 is characterized by the presence of the abnormal cells in the mucosa of the colon. At this stage the tumor is also called carcinoma *in situ*. In the stage I, the cancer invades the mucosa and the submucosa it may also affect the deeper muscle layer of the colon wall, with a survival of the 90–100% of cases [122]. In the stage II, the malignant cancer cells spread outer of the colon wall, but they do not invade the lymph nodes, and the survival rates is about of the 75-85%. When the cancer cells affect the lymph nodes, but not yet other organs, is considered the stage III of colon cancer, with a survival rate of 30-40%. The late stage of the cancer is the IV stages: the malignant cells invade to other organs though the blood and the lymph nodes: the patients present metastasis and less than 5% of patients survive [122].

The diagnosis of CRC can be done with invasive and non-invasive methods [138]. Sigmoidoscopy and colonoscopy are invasive diagnostic methods, that allow to localize the tumor and to collect a part of the large intestine for the histological examination. The sensitivity and specificity of these techniques to identify the polyps and the tumor are very high. Fecal occult blood test is an easy non-invasive method that reveals hemoglobin in feces,

indicating the bleeding of the intestinal tract. The analysis of tumor markers, substances produced by tumor cells or healthy cells as a response to the tumor can be used for the CRC diagnosis. These substances can be reach in the biological fluids of the patients, for example in the blood, plasma, and urine and their expression levels are increased from the first stages of the tumor. The markers for CRC are Carcino-Embryonic Antigen (CEA), Carbohydrate Antigen (CA 19-9), Tissue Polypeptide Specific Antigen (TPS), and Tumor-Associated Glycoprotein-72 (TAG-72) [138].

The best therapy of CRC depends on the staging of the tumor. The treatment for the stages 0 and I, is the surgical ablation of the polyps or part of the colon. Adjuvant therapies, using chemotherapy, are recommended for the stage II and if the cancer presents a higher risk of coming back, when the tumor affects the lymph nodes or when there is metastasis. Patients affect by tumor at stage III and IV, the chemotherapy is the standard treatment: the adjuvants that can be administrated are 5-fluorouracil (5-FU) with leucovorin or with capecitabine [114]. For patients with metastasis or genetic alterations in their cancer cells, the treatment with an immunotherapy drug can be eligible. The use of immunotherapy improves the patient's immune system to eliminate cancer cells, augmenting the immune responses against the tumor through different strategies, for example vaccines and checkpoint inhibitors [139].

1.3.4 Colon cancer: role of HERV

Several studies reported that the modified expression of HERV is related to a pathological contest, such as the tumor, where there is a correlation between the prognosis of the disease and the increase level of transcripts or proteins expression in patients [140]. In literature there are few studies regarding the association between the HERV expression and CRC, especially the studies are focused on the HERV-H, as reported in the Table 3.

***Table 3**: Expression of HERV transcripts (RNA and proteins) in colorectal cancer.*

HERV	RNA and proteins	References
HERV-H	RNA (gag, pol, env, LTR)	[103]
HERV-H	RNA (pol)	[141]
HERV-H	RNA (EST)	[142]
HERV-H	RNA (pol)	[143]
HERV-H	RNA (env)	[144]
HERV-H	RNA (env)	[10]
HERV-H	RNA (env)	[145]
HERV-H	RNA (gag)	[146]
HERV-H	Pr (env)	[147]
HERV-H	RNA (env)	[142]
HERV-H	RNA (env)	[148]
HERV-H	RNA (env)	[10]
HERV-H	RNA (gag, pol, env)	[149]
HERV-H	RNA (gag, pol, env, rec)	[87]
HERV-H	RNA (gag, pol, env)	[74]
HERV-R	Pr (env)	[150]
HERV-W	RNA (env)	[151]
HERV-K	RNA (env)	[152]

A possible association between HERV and colon cancer was demonstrated in the 1986, when Moshier and colleagues assessed the expression of ERV-A *pol* gene in the tumor tissue of patients affected by CRC and the adjacent tissue [143]. After that, the study conducted by Stauffer et al. was focused on the HERV expression in the tissue of cancer patients, included the CRC patients: they analysed the digital expression patterns of the HERV-K, -W, -H, and -E families in several types of cancer [142]. In the CRC, HERV-H

presented high expression levels compared to the other tumor and to the normal tissues.

Then, Liang and colleagues investigated the association between HERV-H and colon cancer, studying the deletion of a part of HERV-H provirus *env* gene, HERV-H-X, located on chromosome Xp22, compared to other three *env* intact ORFs of HERV-H, *env* 59, *env* 60, and *env* 62. The authors examined the expressions in the normal and tumor tissue of 8 patients affected by CRC by means of PCR and Real Time PCR. The complete *env* HERV-H ORFs was found both in the normal and in the tumor tissue whereas the HERV-H-X was detected only in the tumor tissue and the HERV-H-X expression was up-regulated in CRC tissues compared to the normal tissues in a statistically significant way [144]. The HERV-H-X full-length transcript sequence was also investigated by Liang et al., demonstrating that HERV-H-X was upregulated in colon cancer tissue, while HERV-H with complete *env* gene were differently expressed in colon tumor and normal samples, suggesting the different role of HERV-H-X and HERV-H in colon cancer [10].

The HERV expression were also evaluated by means of Reverse Transcriptase-Polymerase Chain Reaction (RT-PCR) in twenty-five primary colon cancer tissues, six liver metastatic CRCs, and one sample of tubule-villous colon adenoma: the aim of this study was to identify a CRC specific antigen to be targeted by immunotherapy [146]. They tested several genes, included the HERV-H *gag* gene, located on the chromosome Xp22: the HERV-H *gag* transcript was differently expressed in most of the primary and liver metastatic CRC samples, while no expression was detected in the normal tissues. The *gag* transcripts' cDNAs were also sequenced and the same deletion of 280 bp downstream from the start codon was observed in the all the tumor samples, suggesting a role of the coding protein as target for the therapy, even if the role of this mutation was still unclear.

The relationship between the expression of the HERV-H transcripts in the tissue and the clinical characteristics of the patients enrolled was reported in a study conducted by Peròt and colleagues: they detected a strong correlation between HERV-H expression and MSI of the tumor, and the lymph node invasion of the tumor cells. These findings suggested the role of HERV-H expression as a diagnostic, prognostic and therapeutic target of the CRC [103].

HERV-H produces also noncoding RNAs in normal and in tumor tissues of CRC patients but also in the colon cell lines, as reported in different studies, even if their the expression pattern is still unclear [10] [148] [149]. In a study conducted by Liang et al. in 2012, they investigated for the first time the non-coding spliced transcripts of HERV-H and they reported that the expression of these noncoding RNAs were significantly different in the normal tissue compared to the tumor tissue [149].

The HERV over expression in colon cancer can be regulated by the liberation of HERV LTRs from epigenetic controls, especially through the hypomethylation, as reported by Liang et al. [149] and Wentzensen et al. [74]. By means of RT-PCR, they assessed that the expression of HERV-H located on the chromosome Xp22.3 was higher in the CRC tissue compared to the normal tissue in the 40% of the samples and the over-expression was associated to the demethylation at the 5′ LTR in the tumor tissue; this association supported that alterations in the methylation status were correlated to the tumor.

The HERV-H expression in CRC are correlated to the genetic and clinical characteristics of the patients, and to the tumor stages. A study of Peròt et al. provided more information regarding the associations between the HERV-H overexpression and the genetic and clinical characteristics of the colorectal cancer patients. For instance, the HERV-H was overexpressed in MSI CRC; its overexpression was also related to the lymph node infiltration and to the

aggressiveness of the tumor, but no correlation with the presence of metastasis was demonstrated [103]. HERV-H expression has been also associated with the epithelial to mesenchymal transition only in the early phases of cancer transformation; thus, HERV may not play a role in subsequent phases of tumor progression. During the epithelial to mesenchymal transition HERV-H overexpression may be also involved in the tumor progression through immunoevasion. In fact, the H17 peptide coded by HERV-H *env* gene were more expressed, inducing the production of CCL19, a chemokine that was associated to immune dysfunction and was involved in the recruitment and in the expansion the CD271positive cells, a population of pluripotent immunoregulatory cells; these groups of cells included mesenchymal stem cells and myeloid-derived suppressor cells. Changes in the expression of HERV-H and/or CCL19, and depletion of the CD271 positive cells enhanced immune responses *in vitro* and *in vivo* inducing the tumor regression [98].

Only one study reported the expression of HERV-R Env protein in the CRC: using microarray and western blot, 55 CRCs, normal colon, and metastatic CRC tissues were analysed [150]. The expression of HERV-R Env protein was significantly increased in the CRC tissues, compared to the adjacent normal tissues of the same patients, especially in the primary and metastasized CRCs.

HERV-W seems to be not associated to CRC, but involved only in the neurological diseases. Kim and colleagues analysed the HERV-W expression in CRC by means of Real Time PCR: no significant differences among the tumor and the normal adjacent tissues were reported [151]. The same author found high levels of expression of HERV-P in CRC tissues, compared to the adjacent normal tissues [152].

1.4 HERV and the extracellular vesicles

1.4.1 Extracellular vesicles: origins and roles

The vesicle are spherical structures delimited by a lipid bilayer (with a similar structure to the cell membranes), which contain soluble hydrophilic components in their cargo. They can be divided into intracellular vesicles, important for the transport of components between intracellular compartments, and in extracellular vesicles (EVs).

Based on their biogenesis, EVs can be divided into three different categories:

1. Microvesicles (or ectosomes) are large extracellular vesicles with a diameter of 50-1000 nm[153]. They express phosphatidylserine (PS) on the cell surface. They are released from many cells, including tumor cells;
2. Exosomes are small membranous vesicles with an endocytic origin, (diameter of 30-100 nm) [153]. The biogenesis of exosomes begins with the internalization of molecules through endocytosis [154]. Then they are recycled into the plasmatic membrane or directed towards multivesicular bodies (MVBs);
3. Apoptotic bodies (ABs) are released from the apoptotic cells with a diameter of about 50-5000 nm [153]. ABs are composed during the process of apoptosis and they represent dead cell fragments. As for ectosomes, the expression of PS on the cell surface is a characteristic of ABs [155].

In the last years, many studies asses that these vesicles, containing proteins, lipids and genetic material, play an important role as a vector. These components are released into the vesicles out of the cells and they bind the receptors on other cells, inducing intracellular signals that modify the physiological state of the receiving cells. For these reasons, they seem to play a key role in cell to cell communication mediated by extracellular vesicles [156]. EVs perform their roles with several mechanisms: they can

act as signaling complexes, using ligands expressed on the cell surface, stimulating the target cells; they can transfer the receptors between the cells; they can transport functional proteins or infectious particles to the target cells, by which can modify the function and physiological state of the recipient cells; finally, they can transfer genetic information by mRNA, microRNA (miRNA), transcription factors, retrotransposons and other DNA elements from one cell to another [157].

1.4.2 Extracellular vesicles and cancer

The cancer cells release more EVs compared to normal cells, which makes the EVs a potential targets for cancer therapy [158]. The increased levels of EVs can be due to the cellular intrinsic and environmental factors. For instance, the activation of oncogenic pathways, such as Epithelial Growth Factor Receptor vIII (EGFRvIII) pathway, induce the production of higher amount of EVs [159]. Moreover, the environmental conditions, like hypoxia, regulate the release of EVs from the cells [160].

EVs modify the behaviour of the stromal cells, promoting the tumor-niche formation, that support tumor angiogenesis and immunosuppression [158]. EVs are also involved in the extracellular matrix remodelling, and in the pre-metastatic niche formation [161].

Although tumor-derived EVs present on the surface tumor-antigens that could be involved in antitumor immunity [162], some studies demonstrated that they suppress the immune response, inhibiting the proliferation and activation of CD8+ cells while promoting the expansion of regulatory T cells (T-reg) [163].

The circulating EVs isolated from cancer patients have been also associated with metastasis or relapse: they could act as diagnostic/prognostic markers and as therapeutic targets [164].

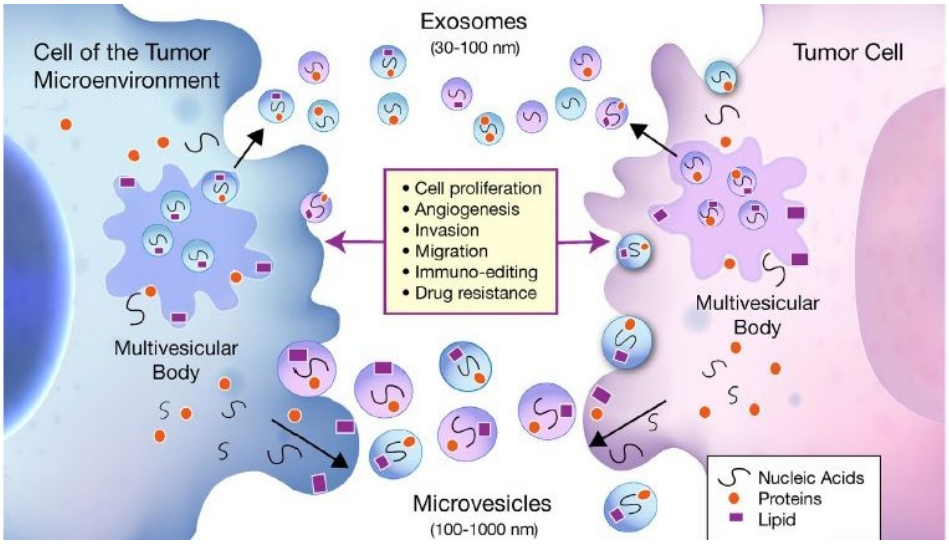


Figure 9: *The intercellular communication between the tumor cells and the cells of the tumor microenvironment via exosomes and microvesicles* [161].

1.4.3 HERV, extracellular vesicles and colon cancer

In a study conducted by Skog and colleagues, it has been demonstrated that tumour microvesicles contain also retrotransposon elements [97]: in the cancer, the hypomethylation of the genome seems to induce the increased of expression of the retrotransposable elements, as RNA and as proteins. These overexpressed elements can be a cargo for the EVs derived from the tumor and induce transformation in other normal cells. Especially, Skog analysed the cargo of the EVs of patients affected by glioblastoma, melanoma, medulloblastoma: the EVs were enriched for the transcritpts of HERV-K, HERV-C, and HERV-W [97]. The role of HERV in these vesicles is still unclear and further investigation is necessary to understand if HERV transcripts in the EVs induce the tumor progression or they can be considered like a biomarker of the disease [19].

2. AIM OF THE STUDY

Colorectal cancer (CRC) is the fourth malignancy in incidence and mortality in developed countries, with more than 1.8 million new cases in 2018 in the worldwide. There are many risk factors involved in the cancer development: environmental and genetics factors, but also infectious agents.

The Human Endogenous Retroviruses (HERV) are remnants of ancients' exogenous retrovirus infection, integrated in the human genome, and their regulation is based on the DNA methylation of promoters, the long terminal repeats (LTRs). They are involved in the pathogenesis of some diseases, such as neurological diseases, but also in the cancer. Several studies reported the association between the alteration in the HERV transcripts/proteins expression and cancers, especially melanoma and breast cancer. Instead, a few reports have shown the existence of a relationship between the HERV expression and colon cancer, based on HERV expression profile in normal and cancer tissues, but conclusive evidences are still lacking.

The main objective of this study was to investigate the involvement of HERV in colorectal cancer pathogenesis. HERV expression (RNA and protein) and LTRs methylation were examined in clinical specimens (tumor tissues, adjacent normal tissues and, when possible, blood) collected from two different populations, Italians and Tunisian. Association between viral characteristics and clinical data were also made. The patients were affected with operable, mostly locally advanced colorectal cancer. Moreover, the extracellular vesicles isolated from the plasma of the patients were tested for the presence and for the expression of HERV transcripts.

3. MATERIALS AND METHODS

3.1 Patients enrolment: two different populations

The study was conducted on two different populations. The first population was composed of 63 Italian colon cancer patients, who underwent surgery to remove the tumor at Istituto Clinico Città Studi (Milan) and at Ospedale di Circolo Busto Arsizio (Varese), Italy. The patients enrolled in this study signed the inform consent after Institutional Ethics Committee approval was obtained (Ethical Committee Istituto Clinico Città Studi, Ospedale Maggiore Policlinico Milano, protocol number 683_2017bis). The clinical and demographic characteristics are reported in the Table 4.

Table 4: *Demographic and clinical data of the Italian enrolled patients.*

Sex	**N (%)**
Female	*25 (39.7)*
Male	*31(49.2)*
n.a.	*7 (11.1)*
Age	**mean ± SD**
Years	*76.3 ± 10.6*
n.a.	*7 (11.1)*
Tumor stage	**N°**
pT1	*4 (6.4)*
pT2	*5 (7.9)*
pT3	*37 (58.7)*
pT4	*4 (6.4)*
n.a.	*13 (20.6)*
Lymph nodes +/-	**N (%)**
Positive	*22 (35.0)*
Negative	*28 (44.4)*
n.a.	*13 (20.6)*
Tumor location	**N (%)**
Right colon	*40 (63.5)*
Left colon	*15 (23.8)*
n.a.	*8 (12.7)*

Lymphocyte infiltration	N (%)
Present	*20 (31.8)*
Absent	*14 (22.2)*
n.a.	*29 (46.0)*
Vascular invasion	**N (%)**
Present	*20 (31.7)*
Absent	*26 (41.3)*
n.a.	*17 (27.0)*
Perineural invasion	**N (%)**
Present	*8 (13.0)*
Absent	*38 (60.0)*
n.a.	*17 (27.0)*

For each patient tumor and adjacent normal tissues and, when possible (49.2% of participants), 7 ml of blood and 2 ml of plasma were collected.

A second group of 58 Tunisian colorectal cancer patients were also enrolled in this study and the demographic data are reported in Table 5. The patients underwent surgery at the Habib Thameur Hospital (Tunis, Tunisia). For each patient adjacent normal and tumor tissues were collected. The patients enrolled in this study signed the inform consent after Institutional Ethics Committee approval was obtain (Ethical Committee Istituto Clinico Città Studi, Ospedale Maggiore Policlinico Milano, protocol number 683_2017bis).

Table 5: *Demographic and clinical data of the Tunisian enrolled patients.*

Sex	N (%)
Female	*30 (51.7)*
Male	*25 (43.1)*
n.a.	*3 (5.2)*
Age	**mean ± SD**
Years	*60.9 ± 14.35*
n.a.	*3 (5.2)*

Smoke	**N (%)**
Yes	*20 (34.5)*
No	*35 (60.3)*
n.a.	*3 (5.2)*
Use of alcohol	**N (%)**
Yes	*6 (10.3)*
No	*49 (84.5)*
n.a.	*3 (5.2)*
Tumor stage	**N° (%)**
II	*17 (29.3)*
III	*2 (3.4)*
IV	*36 (62.1)*
n.a.	*3 (5.2)*
Tumor location	**N (%)**
Right colon	*11 (19)*
Left colon	*44 (75.8)*
n.a.	*3 (5.2)*
Family history	**N (%)**
Yes	*8 (13.8)*
No	*47 (81.0)*
n.a.	*3 (5.2)*
Polyps presence	**N (%)**
Present	*21 (36.2)*
Absent	*34 (58.6)*
n.a.	*3 (5.2)*
Chemotherapy	**N (%)**
Yes	*20 (34.5)*
No	*35 (60.3)*
n.a.	*3 (5.2)*

3.2 Preparation of the biological samples

The tumor and adjacent normal tissue and blood samples were delivered to the laboratory after collection. The tissues were cut in pieces of approximately 20 mg and stored at -80°C in RNAlater medium until use.

The blood was collected in a sterile violet vacutainer tubes containing the anticoagulant ethylenediaminetetraacetic acid (EDTA). Immediately after blood collection and after dilution 1:1 with Phosphate Buffer Saline (PBS) (EuroClone, Milano, Italy), plasma was separated following the manufacturer's instruction of the Ficoll-Paque™ Plus method (GE Healthcare Bio-Sciences AB, Uppsala, Sweden) (Appendix A). 2ml of blood, and 2ml of plasma were stored at -80°C.

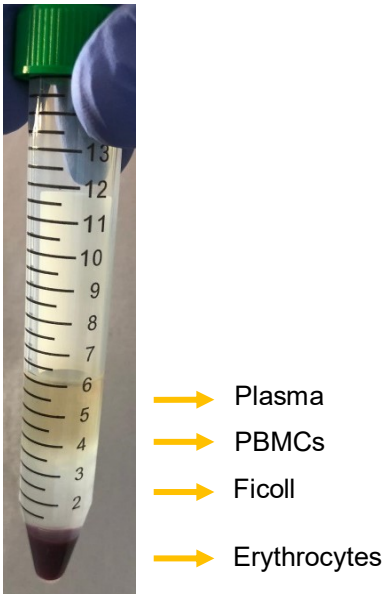


Figure 10: *Plasma isolation using Ficoll method.*

To isolate the EVs, one ml of plasma was centrifuged for three times at 1000 g, 2000 g and 3000 g for 15 minutes at 4°C, and the EVs were concentrated by ultracentrifugation (Beckman Coulter Optima-MAX-XP, Indianapolis, USA) at 110,000 g for 75 minutes at 4°C. The pellets of EVs were stored at -20°C for the following analysis.

3.3 DNA isolation

DNA was isolated from 200 µl of blood, and from 20 mg of tumor and the adjacent normal tissues using QIAmp DNA Blood Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's guidelines (Appendix B). For DNA isolation from tissue, 180 µl of ATL buffer and 20 µl of proteinase K (Qiagen, Hilden, Germany) were added to the sample and incubated overnight at 56°C for the tissue disruption. 200 µl of lysis buffer (AL buffer) and 20µl of proteinase K (Qiagen, Hilden, Germany) were added to blood and to lysed tissue. The sample was incubated for 10 minutes at 56°C in the thermo block. After the ethanol (EtOH) addition, the solution was put into a spin column with a positive charge membrane. Two different washes with AW1 and AW2 buffers were done and finally the DNA was eluted in RNase free water and the concentration was evaluated using Nanophotometer NP80 (Implen, Munich, Germany).

3.4 Bisulfite conversion, PCR and Pyrosequencing

500 ng of DNA was treated with bisulfite, a method to precisely measure DNA methylation. Bisulfite transforms the cytosine of the CpG dinucleotide that are unmethylated in thymine.

The analysis was performed using the EZ DNA Methylation-Gold Kit (Zymo Research, California, USA). 130 µl of CT conversion reagent was added to 20 µl of DNA (500 ng) and then incubated in the thermal cycler, following these steps: 10 minutes at 98°C, 2.5 hours at 64°C and up to 20 hours at 4°C. The treated DNA was stored at -20°C.

A bisulfite-PCR was conducted (Appendix B) using specific primers for the HERV-H, -K, -P, and -W LTRs and *psen,* a cellular control gene: a 50µl PCR was carried out in 25µl GoTaq Green Master Mix (Promega, Wisconsis, USA), 1 pmol forward primers and 1 pmol reverse primers, 50ng bisulfite-treated genomic DNA and water. One of the two primers was biotinylated.

The sequences of the primers for HERV-H, -K, -P and HERV-W LTRs and *psen* were draw using specific software (Table 6):

Table 6: *HERV primers sequences of bisulfited-PCR.*

HERV	Forward Primers	Reverse Primers
-H	5'-BIO-AAAAGGAGGAAAAGTAAAGAAAGA-3'	5'-CCAAAAAAAAAAATTTCACAAAA-3'
-K	5'-GTAAAGGGTTTGTGTTGAGGAG-3'	5'-BIO-ACTTATCCCACACCTCCAAC-3'
-P	5'-BIO-TGTGGAGAAAGAAGTTTGATGTTA-3'	5'-CCTTTTAATCTCTTCACTAATT-3'
-W	5'-AAGTTATAGTTGAAGGAAGA-3'	5'-BIO-CAATCCCCCATCTCAACAA-3'
psen	5'-BIO-AGGGTTTTATTTTTTGTATGATGA-3'	5'-ACTTCAACAACCATTTTACTCTTC-3'

The primers were synthesized by Eurofins Genomics. HERV-R LTRs methylation was not studied because it has not been possible to draw the specific primers.

Moreover, the global methylation was evaluated in the blood, adjacent and tumor tissues: the methylation levels of the repetitive elements, Alu and LINE-1, were analyzed. The sequences of their primers for the bisulfite PCR are reported in the Table 7 [165].

Table 7: *Alu and LINE-1 primers sequences of bisulfited-PCR.*

	Forward Primers	Reverse Primers
Alu	5'-BIO-TTTTTATTAAAAATATAAAAATT-3'	5'-CCCAAACTAAAATACAATAA-3'
LINE-1	5'-TTTTGAGTTAGGTGTGGGATATA-3'	5'-BIO-AAAATCAAAAAATTCCCTTTC-3'

The thermal cycling was 95°C for 5 minutes, followed by 45 cycles of 95°C for 30 seconds, annealing for 30 seconds, 72°C for 30 seconds and finally 72°C for 5 minutes. The annealing temperatures were the following: 43°C for Alu, 50°C for LINE-1 and HERV-H and 52°C for HERV-K and -P, 54°C for *psen*, and 55°C for HERV-W.

The biotin-labelled primer was used to purify the final PCR product using Sepharose beads (Appendix B). The PCR product was bound to Streptavidin Sepharose HP (Amersham-Biosciences, Little Chalfont, United Kingdom) and the Sepharose beads containing the immobilized PCR product were purified, washed and denaturized using a NaOH solution and washed again

using pyrosequencing Vacuum Prep Tool (Qiagen, Hilden, Germany). Pyrosequencing primer annealed to the purified single stranded PCR product, and classical pyrosequencing analysis using Pyro Mark (Qiagen, Hilden, Germany) was performed [165].

The sequences of the primers for pyrosequencing for HERV-H, HERV-K, HERV-P, and HERV-W LTRs and *psen* were draw using specific software, while the sequences of Alu and LINE-1 primers were previously published [165] (Table 8):

Table 8: *Sequences of HERV LTRs, psen and Alu and LINE-1 used for pyrosequencing.*

	Pyrosequencing Primers	
HERV -H	5'-CAATTACTTCAAACCATCTA-3'	Reverse
HERV -K	5'-TTTTGGGTAATGGAATG-3'	Forward
HERV -P	5'-CCCTTTAAATCACAACC-3'	Reverse
HERV -W	5'-AGTTTAAGATTAGATTTAT-3'	Forward
psen	5'-ATGAAAAGAATGTGTTATGA-3'	Reverse
Alu	5'-AATAACTAAAATTACAAAC-3'	Reverse
LINE-1	5'-AGTTAGGTGTGGGATATAGT-3'	Forward

The primers were synthesized by Eurofins Genomics. The degree of methylation was expressed as a percentage of methylated cytosines divided by the sum of methylated and unmethylated cytosines (%5mC) [166].

3.5 RNA isolation

RNA was isolated from 150 µl of blood and from 20 mg of the tumor and adjacent normal tissue using QIAmp RNA Blood Mini Kit (Qiagen, Hilden, Germany), according to the manufacturer's guidelines (Appendix C). The RNA isolation from tissue presented the following modification: 20mg of tissues were crumbled with the scalpel and then 600µl of Lysis Buffer (RLT) and 20 µl of Proteinase K (Qiagen, Hilden, Germany) were added to the sample. The solution was incubated overnight at 56°C in the thermomixer (Eppendorf, Milano, Italy). After the precipitation of RNA using ethanol, the

solution was added to the QIAamp spin column and the same protocol of the RNA isolation from blood was performed (Figure 11).

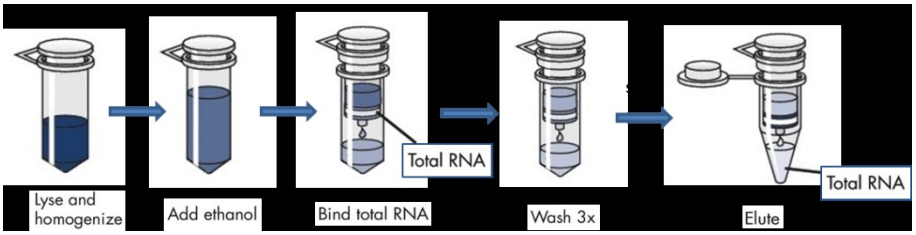


Figure 11: *RNA extraction from blood and tissue using QIAmp RNA Blood Mini Kit.*

RNA concentration was evaluated using Nanophotometer NP80 (Implen, Munich, Germany).

The RNA was also isolated from the EVs using QIAzol lysis reagent (Qiagen, Hilden, Germany) and RNeasy MinElute Cleanup kit (Qiagen, Hilden, Germany). The pellet of EVs was resuspended using 700µl of QIAzol lysis reagent and then 140µl of clorophormio. The mixed solution was centrifugated for 15 minutes at 12,000 g and the upper phase, the aqueous phase, that contained RNA was transferred to a new tube. After the addition of a volume of 70% EtOH, RNA was extracted according to the manufacturer's instruction.

3.6 Reverse transcription and Real Time PCR assay

RNA was reverse transcripted using QuantiTec Reverse Transcription Kit (Qiagen, Hilden, Germany) following the manufacturer's instruction. 1 µg of RNA isolated from the tissue and the blood, and 10ng of RNA extracted from the EVs pellet were reverse transcripted. The composition of the reaction mix is reported in the Table 9.

Table 9: *Composition mix for reverse transcription.*

Reverse-transcription enzyme	1 µl
Quantiscript RT buffer, 5x	4 µl
RT primer Mix	1 µl
Template RNA (1µg or 10ng)	up to14 µl
Final volume	20 µl

The mix reaction with RNA was incubated for 30 minutes at 42°C and then at 95°C for 3 minutes to inactivate the reverse transcription enzyme. The reaction took place in the thermal cycler (Applied Biosystem, California, USA).

Quantitative Sybr Green Real Time PCR was performed to evaluate the expression of the HERV-H, -K, -R and -P *env* gene, and Alu and LINE-1, using the 7500 Real Time PCR system (Applied Biosystem, California, USA). All the *env* genes and Alu and LINE-1 expressions were related to the expression of a housekeeping gene, the Glyceraldehyde 3-phosphate dehydrogenase (GAPDH). The reaction was conducted in a final volume of 25 µl, containing 0.3 µM forward primer, 0.3 µM reverse primer and 2 µl of cDNA (Table 10). For the Tunisian population, the expression of HERV-R, Alu, and LINE-1 were not evaluated but the expression of HERV-K *pol* was investigated. The final concentration of the HERV-K *pol* gene primers was 0.2 µM.

Table 10: *Composition of Real Time PCR mix.*

Reagents	**Initial concentration**	**Final concentration**	**Volume**
Master Mix Sybr Green	2X		12.5 µl
Forward Primer	70 µM	0.3 µM	0.11 µl
Reverse Primer	70 µM	0.3 µM	0.11 µl
cDNA			2 µl
H_2O			10.3 µl

The sequences of each primers were previously published [54] (Table 11).

Table 11: *Real Time PCR primers sequences.*

	SEQUENCES	**SEQUENCES**
	Forward	**Reverse**
GAPDH [54]	5'-GCCCAGGATGCCCTTGA-3'	5'-GTGTCCCCACTGCCAAC-3'
HERV-H *env* gene [54]	5'- TTCACTCCATCCTTGGCTAT-3'	5'- CGTCGAGTATCTACGAGCAAT-3'
HERV-K *env* gene [54]	5'-CACAACTAAAGAAGCTGACG-3'	5'-CATAGGCCCAGTTGGTATAG-3'
HERV-P *env* gene [54]	5- CAAGATTGGGTCCCCTCAC-3'	5'-CCTATGGGGTCTTTCCCTC-3'
HERV-R *env* gene [54]	5'-CATGGGAAGCAAGGGAACT-3'	5'-CTTTCCCCAGCGAGCAATAC-3'
HERV-K *pol* gene [167]	5'-ATCCCAAAAGATTGGCCTTTA-3'	5'-TTAAGCATTCCCTGAGGTAACA-3'
Alu	5'-CAAGGTGAAACCCCGTCTCTAC-3'	5'-CCATTCTCCTGCCTCAGCC-3'
LINE-1	5'-CTCAAGTGGGTCCCTGACTCC-3'	5'-TCTGTTGGAATACCCTGCCG-3'

The reaction steps were: 2 minutes at 50°C, 15 minutes at 95°C and 45 cycle, each of 95°C for 15 seconds, 54°C for 30 seconds and 72 °C for 30 seconds. The quantification of the HERV-H, -K, -P and -R *env* gene, HERV-K *pol* gene, and Alu and LINE-1 expression was performed using the comparative Ct method, and the differences between the levels of *env* and *pol* gene and Alu and LINE-1 expression in the biological samples were calculated by relative quantification (RQ).

$$\Delta Ct\ HERV = Ct\ HERV\ (or\ Alu\ and\ LINE-1) - Ct\ GAPDH$$

$$\Delta Ct\ mean\ HERV = \frac{\sum Ct\ HERV\ (or\ Alu\ and\ LINE-1)}{number\ of\ samples}$$

$$RQ\ expression = 2^{-(\Delta Ct\ HERV(or\ Alu\ and\ LINE-1) - \Delta Ct\ mean\ HERV\ (or\ Alu\ and\ LINE-1))}$$

3.7 Protein isolation

Proteins were isolated from 20 mg of adjacent normal and tumor tissue of 7 patients, who presented the HERV-K *env* gene expression in the EVs.

The tissue was homogenised using lysis buffer, gentleMACS C Tubes (Miltenyi Biotec, Bergisch Gladbach, Germany) and gentleMACS™ Dissociators (Miltenyi Biotec, Bergisch Gladbach, Germany) that allow the disruption of the tissues (Figure 12).

Figure 12: *GentleMACS™ Dissociators and gentleMACS C Tubes used to the protein extraction from the adjacent normal and tumor tissues.*

The specific pre-cool lysis buffer for the protein isolation from 20 mg of tissue was composed by 2.5 ml of RIPA Buffer (Thermo Fisher Scientific Waltham, Massachusetts, USA), one Pierce Protease Inhibitor Tablets to a final concentration of 3X (Thermo Fisher Scientific Waltham, Massachusetts, USA) and 0.5M EDTA to a final concentration of 5 mM. The homogenised tissue was transferred in a tube and it was centrifuged at 4,000g for 5 minutes at 4°C. The protein isolated were in the supernatant and they were stored at -80°C for the next experiments. The proteins were quantified used Bradford assay (Thermo Fisher Scientific Waltham, Massachusetts, USA) using the Nanophotometer NP80 (Implen, Munich, Germany).

3.8 Western Blot HERV-K Pol and Env assay

3.8.1 Preparation of the gels

The proteins, isolated from the adjacent normal and tumor tissues of 7 patients, were separated using gel electrophoresis, in particular SDS-PAGE separation was performed. SDS-PAGE uses two different gels to separate

the proteins: the stacking gel and the running gel. The stacking gel is used to concentrate the proteins in one band, allowing their migration in the running gel at the same time, while the running gel allows to separate the proteins based on their molecular weight.

The composition of the stacking and of the running gels were the followings:

Table 12: *Composition of stacking gel mix.*

MIX stacking gel	Final volume: 5ml
H_2O distilled	3.4 ml
30% acrylamide	0.83 ml
1.0 M TRIS (pH=6.8)	0.63 ml
10% SDS	0.05 ml
10% APS	0.05 ml
Temed	0.005 ml

Table 13: *Composition of running gel mix (10%).*

MIX running gel	Final volume: 15ml
H_2O distilled	5.9 ml
30% acrylamide	5.0 ml
1.5 m tris (pH=8.8)	3.8 ml
10% SDS	0.15 ml
10% APS	0.15 ml
Temed	0.006ml

Based on the molecular weight of the target protein, the gel composition is variable: in our study, the running gel was at 10% (for proteins with molecular weights between 20 kDalton and 300 kDalton).

3.8.2 Preparation of the samples and run of the gel

30 µg of proteins were heated for 5 minutes at 95°C in the thermoblock and they were loaded into the wells of the gel with 4 µl of loading dye solution 6x and RIPA buffer (Thermo Fisher Scientific, Waltham, USA) in a final volume of 24 µL. 6 µl of Prestained Protein SHARPMASS VI (5-245 kDalton) Protein MW Marker (EuroClone, Milano, Italy) was also added to the gel.

The gel ran for 30 minutes at 50 Volt constant and then for 3 hours at 90 Volt in the running buffer 1x, composed of Tris Glycine-SDS (Tris-Glycine-SDS, TGS) (Bio-Rad, Hercules, USA).

3.8.3 Blotting and incubation with antibodies

The proteins were blotted to a nitrocellulose using the iBlot 2 Dry Blotting System (Thermo Fisher Scientific, Waltham, USA). The nitrocellulose was stained with Ponceau solution (Sigma-Aldrich, Missouri, USA) to verify the presence of the proteins, and then in 5% Not Fat Dry Milk (Euroclone, Milano, Italy) in TBS-T 1x was blocked for one hour at room temperature. TBS-T 1x was composed of 900 mL of H_2O distilled, 100 mL of TBS 10X, and 1 mL of Tween 20 (Thermo Fisher Scientific, Waltham, USA). Overnight incubation with the primary antibodies (diluted in 5% Not Fat Dried Milk TBS-T 1x) in a cold room was performed: GAPDH (Bio-Techne, Minneapolis, USA) diluted 1:7500; ERVK Pol (Novus Biological, Centennial, USA) diluted 1:1000; ERVK-7 Env Polyclonal Antibody (Thermo Fisher Scientific, Waltham, USA) diluted 1:1000. The blot was rinsed 3 times for 5 minutes with TBS-T 1x and it was incubated for one hour at room temperature with the secondary antibodies, diluted in 5% Not Fat Dried Milk: the secondary antibodies used were Goat Anti-Mouse IgG Peroxidase Conjugated (Thermo Fisher Scientific, Waltham, USA) diluted 1:5000 and Goat anti-Rabbit IgG HRP linked (Cell Signalling, Massachusetts, USA) diluted 1:1000. Finally, the blot was rinsed 3 times for 5 minutes at room temperature with TBS-T 1x.

3.8.4 Western blot development method and analysis

In the dark room, the chemiluminescent substrate Pierce ECL plus Western Blot Substrate (Thermo Fisher Scientific, Waltham, USA) was added to the nitrocellulose membrane and it was incubated for 5 minutes. The development of the Western Blot was performed using photographic paper and the signal was fixed incubating in the Developer and Fixer solution

(Kodak, New York, USA). The GAPDH protein presents a molecular weight of 39 KDalton, while the ERV-K Pol can be divided in different complexes with different molecular weights, as shown in the Figure 13. The molecular weight of the ERVK-7 Env protein is 67 KDalton.

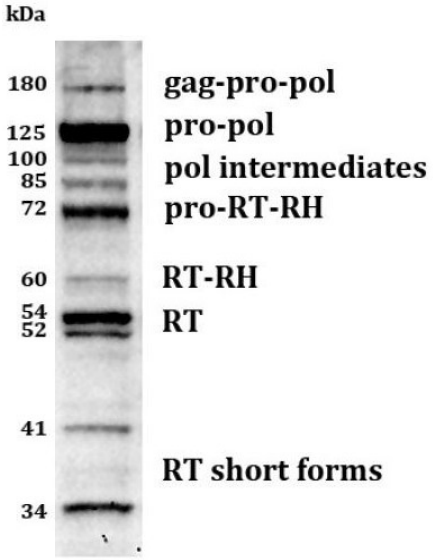


Figure 13: *The ERV-K Pol protein bands with different molecular weights* [168]*. 180 KDalton: complex gag-protease-polymerase; 125 KDalton: protease-polymerase complex; between 100-85 KDalton: polymerase intermediates; 72 KDalton: protease-retrotranscriptase-RNase H complex; 60 KDalton: retrotrotranscriptase-RNase H complex; two isoforms of retrotranscriptase (52 and 54 KDalton); finally, the short forms of retrotranscriptase, from 41 to 34 KDalton.*

The densitometric analysis of Western Blots can be assessed by making comparisons between the bands in different lanes, using ImageJ software. The intensity of the lane was converted in a densitometric signal and it was normalized using the expression of the GAPDH in the same sample.

3.9 Statistical analysis

A descriptive analysis was performed on the demographic and clinical characteristics. For the categorical variables, we reported frequencies and percentages, and the quantitative data are expressed as the mean ±

standard deviation. The graphical analysis with boxplots and spaghetti plots was utilized to describe the expression and methylation levels for each type of sample. We applied multivariable linear mixed models for paired data to investigate the relationship between the HERV *env* expression levels and the types of sample (blood, normal and tumor tissue). Logarithm transformation was performed to satisfy the normality assumption. Adjusted geometric means and 95% CI were calculated and compared. Multivariable linear mixed-effect models for paired data were fitted to evaluate the differences between mean levels of DNA methylation measured in the blood, normal and tumor tissues. CpG dinucleotide position was considered a random effect. Adjusted means and 95% CI were calculated and compared as well.

We also evaluated the association between clinical characteristics and methylation and expression levels in tumor tissue samples with multivariable linear mixed models.

All the linear mixed models were adjusted for age and gender, and an unstructured covariance structure was used to model within-subject errors. The Kenward-Roger approximation was used to estimate the degrees of freedom in the denominator. Where it was necessary to take multiple comparison into account, we selected the Tukey *post hoc* test with Kramer approximation for unbalanced data.

Correlations between the methylation and expression levels were evaluated using the Spearman correlation index. The statistical analysis was performed using SAS software (version 9.4, SAS Institute, Milan, Italy). A two-sided *p*-value of less than 0.05 was considered statistically significant.

GraphPad Prism 7.0 and 8.0 were used to compare the expression of HERV proteins in the adjacent normal and in the tumor tissues. The paired T test was used to evaluate the differences between the protein expression in the tissues.

4. RESULTS

4.1 Case study

Patients with different origin, Italians and Tunisian, who underwent a colorectal surgery, were enrolled in this study. As for the Italian patients, the mean age was 76.3 ± 10.6 years (range 37-91 years); the 49.2% of the patients were male; the majority of patients were affected by the tumor at the III stage (58.7%); the lymph-nodes were positive in 35% of the patients; the tumor location was in the right colon in 63.5% of the patients; the lymphocyte infiltration was present in 31.8% of the patients, and vascular and perineural invasions were present in 31.7% and 13.0% of the patients, respectively (Table 4).

As for the Tunisian patients, the mean age was 60.9 ± 14.4 years (range 28-86 years) the 43.1% of the patients were male; 65.5% of the patients were affected by the tumor at the advance stages (III and IV); the tumor location was in the left colon in 75.8% of the patients; most of the patients both were not a smoker (60.3%) and the 84.5% did not use alcohol (Table 5).

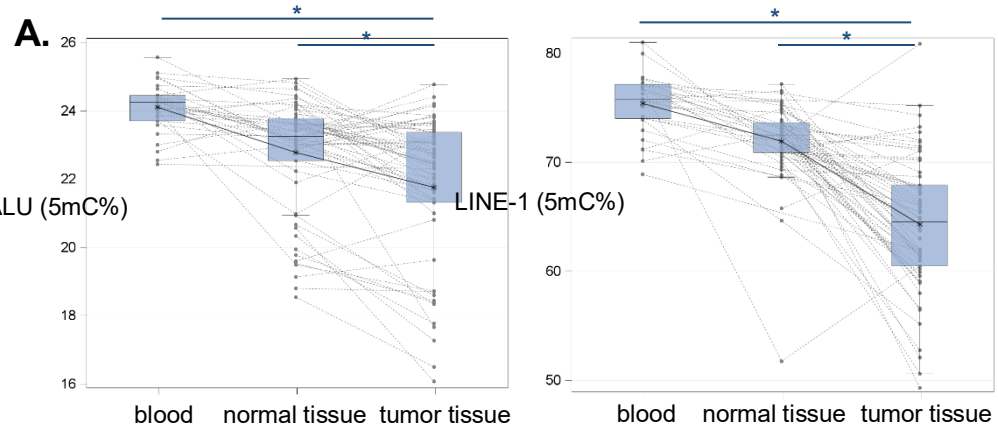
4.2 Methylation status of HERV LTRs, and Alu and LINE-1 in the Italians population

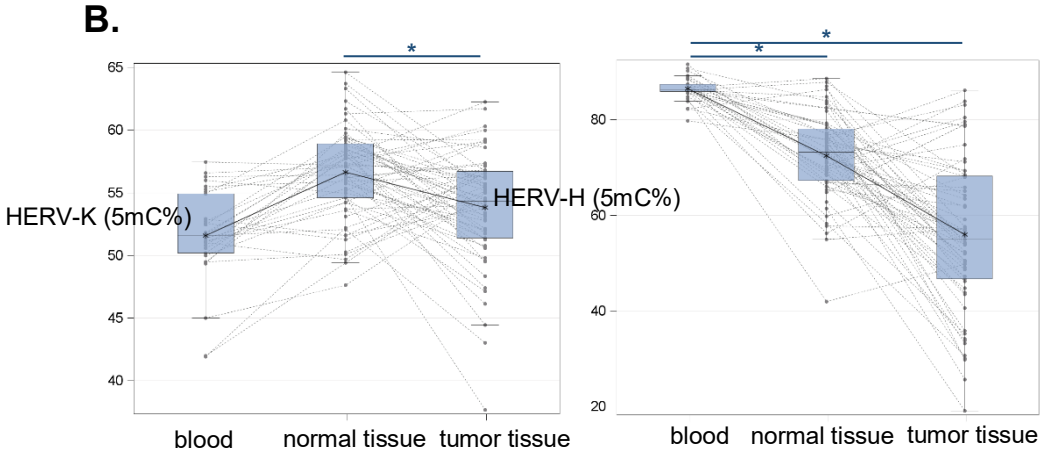
The adjusted mean levels of Alu and LINE-1 promoter DNA methylation were significantly ($p<0.0001$) decreased in the tumor tissues (mean: 21.64%, 95% CI: 21.23-22.05, and 64.38%, 95% CI: 63.20-65.56, respectively) compared to the adjacent tissues (mean: 22.59%, 95% CI: 22.59, 95% CI: 22.29-22.89 and 71.91%, 95% CI: 71.23-72.60, respectively) or the blood (mean: 23.52%, 95% CI: 23.18-23.85 and 75.31%, 95% CI: 74.50-76.13, respectively) (Figure 14 panel A, and Table 14).

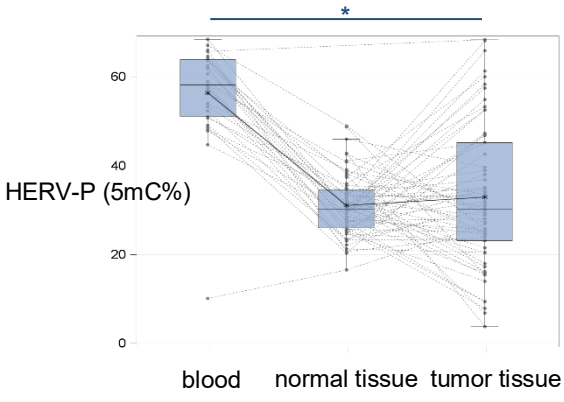
HERV-H LTRs were significantly ($p<0.0001$) demethylated in the tumor tissues (mean: 56.64%, 95% CI: 52.60-60.67) compared to the normal adjacent tissues (mean: 72.49%, 95% CI: 70.12-74.85), or the blood (mean: 86.54%, 95% CI: 85.63-87.44); HERV-K LTRs were significantly ($p<0.0001$) demethylated in the tumor tissues (mean: 53.8%, 95% CI: 52.52-55.09) compared to the adjacent normal tissues (mean: 56.64%, 95% CI: 55.63-57.64). HERV-P LTRs were significantly demethylated ($p<0.0001$) in the tumor tissues (mean: 31.58%, 95% CI: 27.36-35.80) compared to the blood (mean: 56.83%, 95% CI: 53.26-60.40). HERV-W LTRs were similarly methylated ($p>0.05$) in the three specimens: mean 93.80%, (95% CI: 92.76-93.70) in the tumor tissues; mean 94.34%, (95% CI: 94.10-94.58) in the adjacent normal tissues; and mean 94.74%, (95% CI: 94.46-95.02) in the blood (Figure 14, panel B and C, and Table 14).

Psen methylation status was assessed only in tumor and adjacent tissue, because of the lack of blood sample. The levels of its promoter DNA methylation were similar ($p>0.05$) in the tumor tissues (mean: 94.40%, 95% CI: 93.69-95.11) and in the adjacent tissues (mean: 93.50%, 95% CI: 92.58-94.42) (Figure 14, panel C, and Table 14).

*P-value< 0.0001







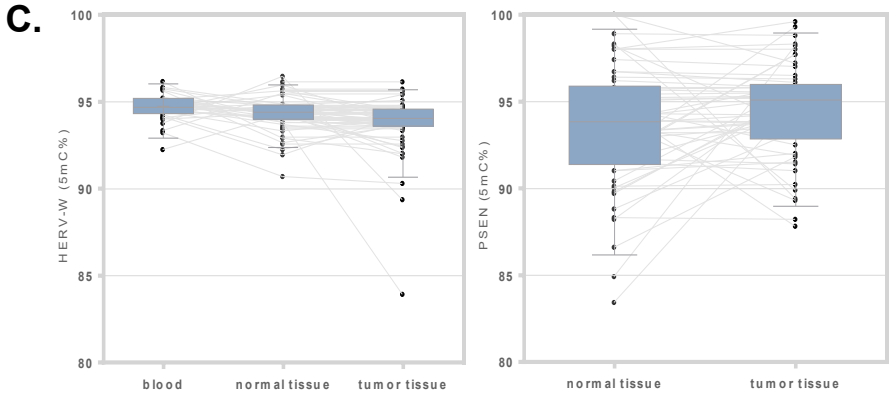


Figure 14: *Alu and LINE-1 (panel A), HERV-K, -H and -R LTRs (panel B), and HERV-W and psen gene (as controls, panel C) methylation mean levels measured in the blood (when possible), normal adjacent and tumor tissue samples in the Italian population. The spaghetti plots show changes in individual DNA methylation and the mean change among the samples. The box plots summarize the DNA methylation levels for each sample.*

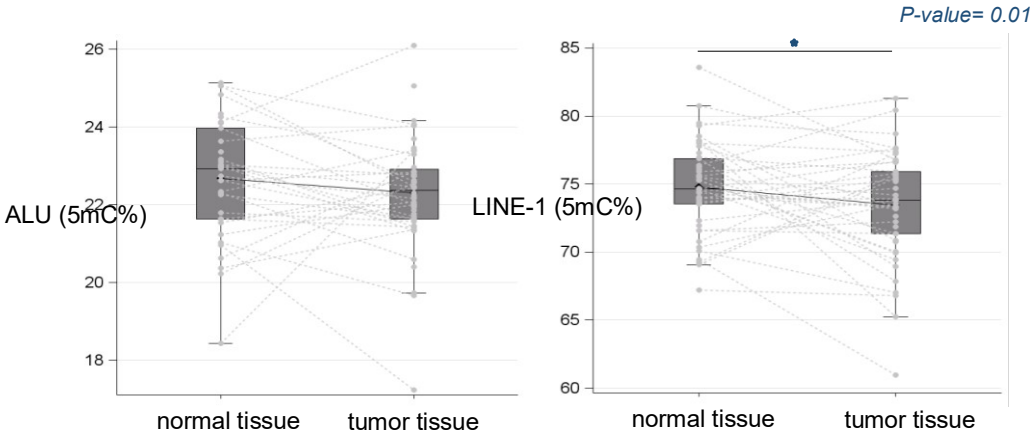
Table 14: *Difference in DNA methylation among the samples (blood, normal and tumor tissues) in the Italian population. Mean*: geometric mean*

		Mean *	95% CI		Overall P-value	Comparison	Mean difference *	P-value	Adjusted Tukey-Kramer P-value
ALU	*Blood*	23.52	23.18	23.85	**<0.0001**	*Blood vs Normal tissue*	0.93	**<0.0001**	**<0.0001**
	Normal tissue	22.59	22.29	22.89		*Blood vs Tumor tissue*	1.88	**<0.0001**	**<0.0001**
	Tumor tissue	21.64	21.23	22.05		*Normal vs Tumor tissue*	0.95	**<0.0001**	**<0.0001**
LINE-1	*Blood*	75.31	74.50	76.13	**<0.0001**	*Blood vs Normal tissue*	3.40	**<0.0001**	**<0.0001**
	Normal tissue	71.91	71.23	72.60		*Blood vs Tumor tissue*	10.94	**<0.0001**	**<0.0001**
	Tumor tissue	64.38	63.20	65.56		*Normal vs Tumor tissue*	7.54	**<0.0001**	**<0.0001**
HERV-H	*Blood*	86.54	85.63	87.44	**<0.0001**	*Blood vs Normal tissue*	14.05	**<0.0001**	**<0.0001**
	Normal tissue	72.49	70.12	74.85		*Blood vs Tumor tissue*	29.90	**<0.0001**	**<0.0001**
	Tumor tissue	56.64	52.60	60.67		*Normal vs Tumor tissue*	15.85	**<0.0001**	**<0.0001**
HERV-K	*Blood*	51.77	50.46	53.09	**<0.0001**	*Blood vs Normal tissue*	-4.86	**<0.0001**	**<0.0001**
	Normal tissue	56.64	55.63	57.64		*Blood vs Tumor tissue*	-2.03	0.0223	0.0573
	Tumor tissue	53.80	52.52	55.09		*Normal vs Tumor tissue*	2.83	0.0004	0.0010
HERV-P	*Blood*	56.83	53.26	60.40	**<0.0001**	*Blood vs Normal tissue*	25.63	**<0.0001**	**<0.0001**
	Normal tissue	31.21	29.41	33.00		*Blood vs Tumor tissue*	25.25	**<0.0001**	**<0.0001**
	Tumor tissue	31.58	27.36	35.80		*Normal vs Tumor tissue*	-0.38	0.8650	0.9841
HERV-W	*Blood*	94.74	94.46	95.02	0.1074	*Blood vs Normal tissue*	0.4	0.3394	0.1383
	Normal tissue	94.34	94.10	94.58		*Blood vs Tumor tissue*	1.51	0.0466	0.0696
	Tumor tissue	93.80	92.76	93.70		*Normal vs Tumor tissue*	1.11	0.0578	0.2213
psen	*Normal tissue*	93.50	92.58	94.42					
	Tumor tissue	94.40	93.69	95.11	0.1357	*Normal vs Tumor tissue*	-0.9	0.1357	0.2078

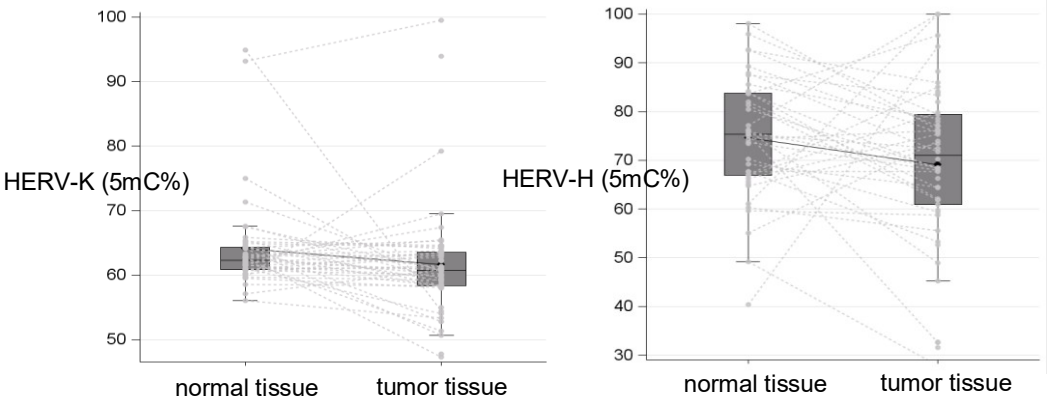
4.1 Methylation status of HERV LTRs, and Alu and LINE-1 in the Tunisian population

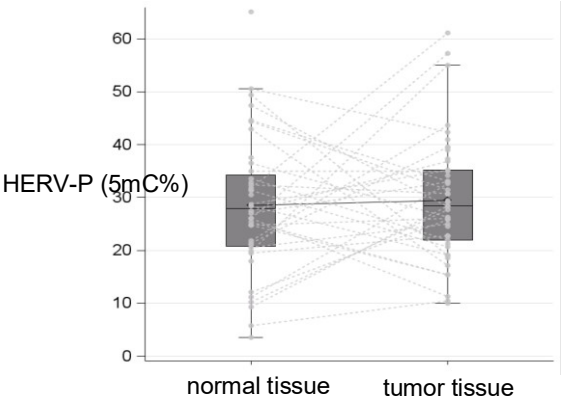
The promoter of LINE-1 sequences was hypomethylated (p=0.0109) in the tumor tissues (mean: 73.47%, 95% CI: 72.62-74.32) compared to the adjacent normal tissues (mean: 74.80%, 95% CI: 74.11-75.49) (Figure 15, panel A, Table 15). ALU was less methylated in the tumor tissues compared to the adjacent normal tissues (mean: 22.34%, 95% CI: 22.05-22.63 and mean: 22.70%, 95% CI: 22.35-23.04, respectively), but the difference was not statistically significant (Figure 15, panel A, Table 15). HERV-H LTRs were demethylated in the tumor tissues (mean: 69.11%, 95% CI: 64.32-73.90) compared to the normal adjacent tissues (mean: 74.37%, 95% CI: 70.41-78.32); also the HERV-K LTRs were demethylated in the tumor tissues (mean: 61.56%, 95% CI: 59.15-63.98) compared to the adjacent normal tissues (mean: 64.20%, 95% CI: 62.01-66.39); HERV-P and HERV-W LTRs were similarly methylated (p>0.05) in the two tissues: mean 29.42%, (95% CI: 26.06-32.77) and mean 94.56%, (95% CI: 93.96-95.15) in the tumor tissues respectively, and mean 28.43%, (95% CI: 24.24-32.62) and mean 94.81% (95% CI: 93.76-95.86) (Figure 15, panel B and C, Table 15).

A.


P-value= 0.01
ALU (5mC%)
LINE-1 (5mC%)
normal tissue
tumor tissue
normal tissue
tumor tissue

B.


HERV-K (5mC%)
HERV-H (5mC%)
normal tissue
tumor tissue
normal tissue
tumor tissue


HERV-P (5mC%)
normal tissue
tumor tissue

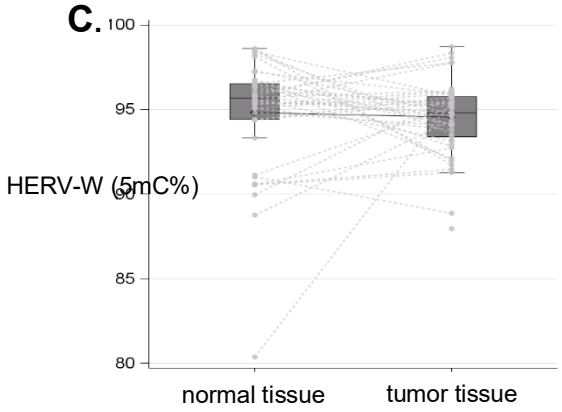


Figure 15: *Alu and LINE-1 (panel A), HERV-K, -H and -R LTRs (panel B), and HERV-W and (as controls, panel C) methylation mean levels measured in the normal adjacent and tumor tissue samples of the Tunisian population. The spaghetti plots show changes in individual DNA methylation and the mean change among the samples. The box plots summarize the DNA methylation levels for each sample.*

Table 15: *Difference in DNA methylation among the tissue of the Tunisian patients.*

		Mean	95% CI		P-value
LINE-1	*Normal tissue*	74.80	74.11	75.49	**0.01**
	Tumor tissue	73.47	72.62	74.32	
ALU	*Normal tissue*	22.70	22.35	23.04	0.08
	Tumor tissue	22.34	22.05	22.63	
HERV-H	*Normal tissue*	74.37	70.41	78.32	0.07
	Tumor tissue	69.11	64.32	73.90	
HERV-K	*Normal tissue*	64.20	62.01	66.39	0.06
	Tumor tissue	61.56	59.15	63.98	
HERV-P	*Normal tissue*	28.43	26.06	32.77	0.70
	Tumor tissue	29.42	24.24	32.62	
HERV-W	*Normal tissue*	94.81	93.76	95.86	0.67
	Tumor tissue	94.56	93.96	95.17	

4.2 Expression of HERV env, and LINE-1 and Alu in the clinical specimens of the Italians populations

The differences in the HERV *env* and in the LINE-1 and Alu expression between the samples are summarized in Figure 16. Expression levels of LINE-1 and Alu were assessed only in tumor and normal adjacent tissues, because of the lack of blood sample. The mean values of the expression levels for the four analyzed HERV *env* genes were similar in the blood, tumor and normal adjacent tissues (Figure 16). HERV-H and HERV-P were more expressed in the tumor tissue (geometric mean: 0.97 and 1.03, respectively; 95% CI: 0.67-1.41, and 0.64-1.64, respectively) than in the blood (geometric mean: 0.85 and 0.87; 95% CI: 0.50-1.43 and 0.53-1.44) or in the normal adjacent tissue (geometric mean: 0.84 and 0.79; 95% CI: 0.46-1.55 and 0.51-1.23). In contrast, HERV-K was more expressed in the blood (geometric mean 1.00; 95% CI: 0.69-1.45) than in the tumor tissue (geometric mean: 0.90; 95% CI: 0.66-1.23) or in the normal adjacent tissue (geometric mean: 0.77; 95% CI: 0.54-1.09).

Both LINE-1 and Alu were similarly expressed in the normal adjacent tissue (geometric mean: 0.93; 95% CI: 0.5-2.27 and geometric mean: 0.93; 95% CI: 0.33-1.9, respectively; p=0.144), then in the tumor tissue (geometric mean: 0.97; 95% CI: 0.46-1.48 and geometric mean: 0.96; 95% CI:0.52-1.44, respectively; p=0.157) (Figure 16).

HERV *env* sequences were detected in the EVs of 14/26 (54%, -H), 10/26 (38%, -K), 1/26 (4%, -P) and 8/26 (31%, -R) tested samples. The HERV *env* was more expressed in the EVs than in the other biological samples from the same patients, but with no significant differences, as described in Table 16.

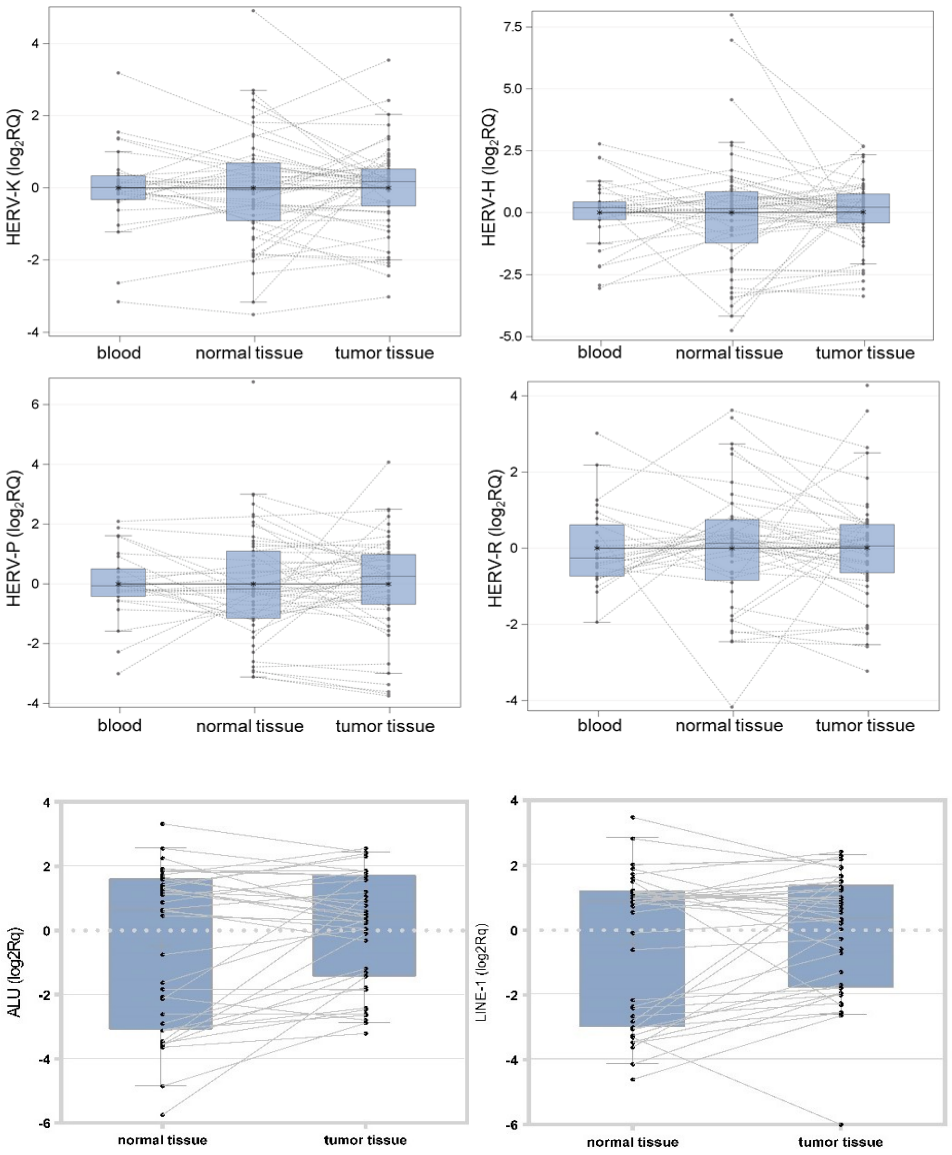


Figure 16: *HERV-K, -H, -P and -R, and LINE-1, Alu expression levels measured in the blood (when possible), normal tissue and tumor tissue samples of the Italian population. The spaghetti plots show changes in individual expression levels and the mean change among the samples. The box plots summarize the expression levels for each sample.*

Table 16: *HERV env expression levels in the EVs compared to the blood, adjacent normal and tumor tissues in the Italian population.*

		Mean Rq	Comparison	Mean Differences Rq	95% CI		Adjusted P-value
HERV-H	*EVs*	2.82					
	Blood	1.12	*EVs vs blood*	1.73	-0.98	4.44	0.25
	Normal tissue	1.37	*EVs vs normal tissue*	1.49	-1.25	4.22	0.37
	Tumor tissue	1.69	*EVs vs tumor tissue*	1.17	-1.14	3.76	0.51
HERV-K	*EVs*	1.24					
	Blood	1.06	*EVs vs blood*	0.18	-0.85	1.21	0.92
	Normal tissue	0.75	*EVs vs normal tissue*	0.49	-0.46	1.44	0.37
	Tumor tissue	1.08	*EVs vs tumor tissue*	0.16	-0.89	1.21	0.95
HERV-R	*EVs*	2.06					
	Blood	0.78	*EVs vs blood*	1.28	-0.76	3.32	0.23
	Normal tissue	1.04	*EVs vs normal tissue*	1.02	-1.27	3.30	0.46
	Tumor tissue	1.13	*EVs vs tumor tissue*	0.93	-1.07	3.01	0.74

4.3 Expression of HERV env and pol genes in the clinical specimens of the Tunisian populations

The mean values of the expression levels for the three analyzed HERV *env* genes and HERV-K *pol* gene were similar in the adjacent normal tissue (HERV-H: geometric mean 1.00, 95% CI 0.66-1.53; HERV-K *env*: geometric mean 1.00, 95%CI 0.76-1.32; HERV-K *pol*: geometric mean 1.03, 95%CI 0.64-1.65; HERV-P: geometric mean 1.01, 95%CI 0.64-1.61) and in the tumor tissue (HERV-H: geometric mean 1.00, 95% CI 0.82-1.22; HERV-K *env*: geometric mean 1.00, 95%CI 0.81-1.24; HERV-K *pol*: geometric mean 0.96, 95%CI 0.66-1.40; HERV-P: geometric mean 1.00, 95%CI 0.72-1.40) (Figure 17, Table 17).

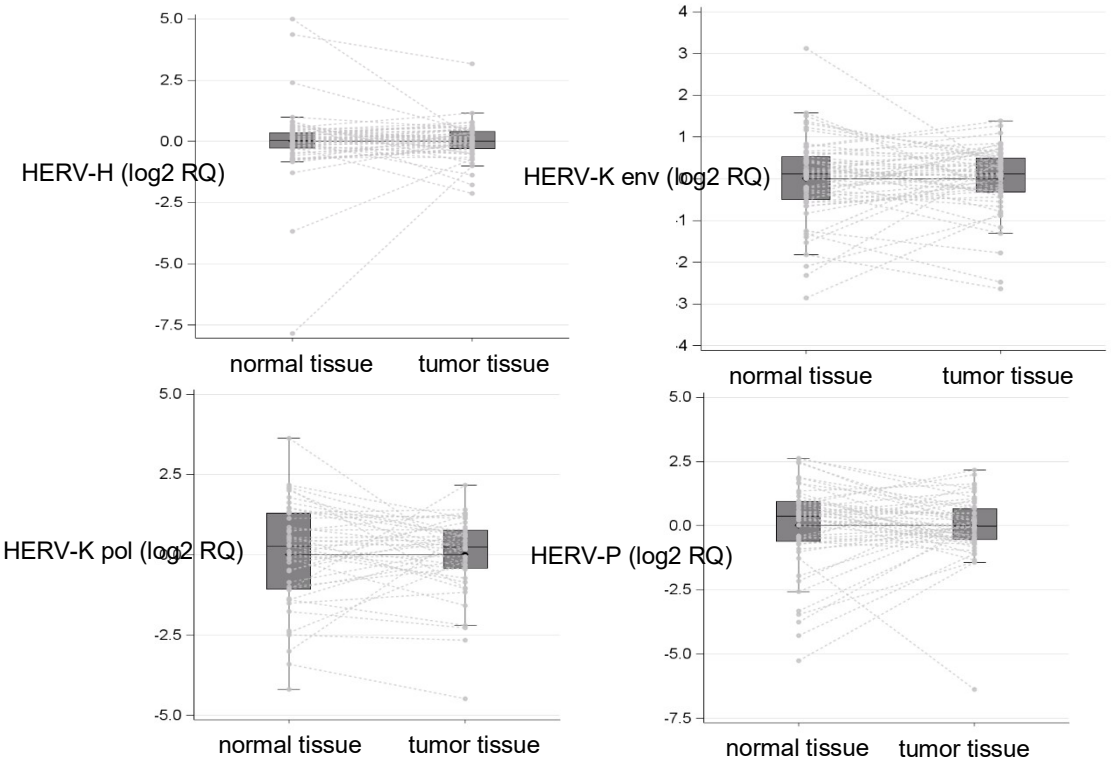


Figure 17: *HERV-H, -K env, -K pol and -P expression levels measured in normal tissue and tumor tissue samples of the Tunisian population. The spaghetti plots show changes in individual expression levels and the mean change among the samples. The box plots summarize the expression levels for each sample.*

Table 17: *Difference in HERV expression levels among the tissue of the Tunisian patients.*

		Mean	95% CI		P-value
HERV-H *env*	*Normal tissue*	1.00	0.66	1.53	0.1
	Tumor tissue	1.00	0.82	1.22	
HERV-K *env*	*Normal tissue*	1.00	0.76	1.32	0.1
	Tumor tissue	1.00	0.81	1.24	
HERV-K *pol*	*Normal tissue*	1.03	0.64	1.65	0.76
	Tumor tissue	0.96	0.66	1.40	
HERV-P *env*	*Normal tissue*	1.01	0.64	1.61	0.96
	Tumor tissue	1.00	0.72	1.40	

4.4 Correlation analysis in the Italian population

The Table 18, 19, and 20 reported the correlation between the HERV expressions, between the LTRs methylation levels, and the correlation between the expression and methylation in the three biological specimens, respectively. There is a significant correlation between the HERV-H and HERV-K expression in the three biological specimens; a correlation between HERV-H and HERV-P expression in the adjacent normal tissue, and between HERV-H and HERV-R expression in the tumor tissue ($p<0.05$). No significant correlation was observed between methylation status and the expression of any of the tested HERV.

Regarding the association between HERV *env* expression and clinical characteristics, no association was observed between HERV *env* expression and clinical stages of the diseases. Interestingly, the four tested HERV were more expressed in the tumors tissue arising in the right colon than in left colon, reaching a statistical difference in the expression of HERV-P and HERV-R in the right colon ($p=0.049$ and 0.045, respectively) (Table 21).

Table 18: *Correlation analysis between the HERV expression and methylation in the blood.*

	Met HERV_P	**rqExpr HERV_P**	**Met HERV_H**	**rqExpr HERV_H**	**Met HERV_K**	**rqExpr HERV_K**	**rqExpr HERV_R**
Met_HERV_P	1	-0.22	0.16	-0.34	-0.12	-0.004	-0.26
p-value		0.30	0.40	0.06	0.51	0.98	0.19
Number oservation	31	25	31	30	31	30	27
rqExpr_HERV_P	-0.22	1	0.19	-0.13	-0.32	-0.19	0.22
p-value	0.30		0.36	0.54	0.12	0.37	0.31
Number oservation	25	25	25	25	25	25	24
Met_HERV_H	0.16	0.19	1	-0.32	-0.29	-0.004	-0.25
p-value	0.40	0.36		0.09	0.11	0.98	0.21
Number oservation	31	25	31	30	31	30	27
rqExpr_HERV_H	-0.34	-0.13	-0.32	1	0.03	0.45	0.27
p-value	0.06	0.54	0.09		0.87	**0.01**	0.18
Number oservation	30	25	30	30	30	30	27
Met_herv_K	-0.12	-0.32	-0.29	0.03	1	0.05	0.35
p-value	0.51	0.12	0.11	0.87		0.80	0.08
Number oservation	31	25	31	30	31	30	27
rqExpr_HERV_K	-0.004	-0.19	-0.004	0.45	0.05	1	0.30
p-value	0.98	0.37	0.98	**0.01**	0.80		0.13
Number oservation	30	25	30	30	30	30	27
rqExpr_HERV_R	-0.26	0.22	-0.25	0.27	0.35	0.30	1
p-value	0.19	0.31	0.21	0.18	0.08	0.13	
Number oservation	27	24	27	27	27	27	27

Table 19: *Correlation analysis between the HERV expression and methylation in the adjacent normal tissue.*

	Met HERV_P	**rqExpr HERV_P**	**Met HERV_H**	**rqExpr HERV_H**	**Met HERV_K**	**rqExpr HERV_K**	**rqExpr HERV_R**
Met_HERV_P	1	-0.11	0.13	-0.09	-0.01	-0.11	-0.17
p-value		0.45	0.32	0.54	0.91	0.41	0.27
Number oservation	62	54	62	51	62	54	43
rqExpr_HERV_P	-0.11	1	0.06	0.50	-0.03	0.21	0.64
p-value	0.45		0.67	**0.0002**	0.85	0.13	**<.0001**
Number oservation	54	55	54	52	54	52	44
Met_HERV_H	0.13	0.06	1	-0.20	-0.10	-0.06	0.06
p-value	0.32	0.67		0.16	0.44	0.65	0.69
Number oservation	62	54	62	51	62	54	43
rqExpr_HERV_H	-0.09	0.50	-0.20	1	0.08	0.36	0.48
p-value	0.54	**0.0002**	0.16		0.57	**0.01**	**0.001**
Number oservation	51	52	51	52	51	49	44
Met_herv_K	-0.01	-0.03	-0.10	0.08	1	-0.25	0.06
p-value	0.91	0.85	0.44	0.57		0.06	0.72
Number oservation	62	54	62	51	62	54	43
rqExpr_HERV_K	-0.11	0.21	-0.06	0.36	-0.25	1	-0.16
p-value	0.41	0.13	0.65	**0.01**	0.06		0.30
Number oservation	54	52	54	49	54	55	42
rqExpr_HERV_R	-0.17	0.64	0.06	0.48	0.06	-0.16	1
p-value	0.27	**<.0001**	0.69	**0.001**	0.72	0.31	
Number oservation	43	44	43	44	43	42	44

Table 20: *Correlation analysis between the HERV expression and methylation in the tumor tissue.*

	Met HERV_P	**rqExpr HERV_P**	**Met HERV_H**	**rqExpr HERV_H**	**Met HERV_K**	**rqExpr HERV_K**	**rqExpr_ HERV_**
Met_HERV_P	1	0.22	0.43	0.21	0.10	0.23	0.21
p-value		0.10	**0.0005**	0.11	0.45	0.08	0.12
Number oservation	63	54	63	58	62	59	56
rqExpr_HERV_P	0.22	1	0.15	0.32	0.20	0.43	0.62
p-value	0.10		0.29	**0.02**	0,15	**0.001**	**<.0001**
Number oservation	54	54	54	53	53	54	53
Met_HERV_H	0.43	0.15	1	0.25	0.01	0.08	0.15
p-value	**0.0005**	0.30		0.06	0.97	0.55	0.27
Number oservation	63	54	63	58	62	59	56
rqExpr_HERV_H	0.21	0.32	0.25	1	-0.02	0.58	0.31
p-value	0.11	**0.02**	0.06		0.88	**<.0001**	**0.02**
Number oservation	58	53	58	58	57	58	55
Met_herv_K	0.10	0.20	0.005	-0.02	1	0.16	0.15
p-value	0.45	0.15	0.97	0.88		0.22	0.27
Number oservation	62	53	62	57	62	58	55
rqExpr_HERV_K	0.23	0.43	0.08	0.58	0.16	1	0.43
p-value	0.08	**0.001**	0.55	**<.0001**	0.22		**0.001**
Number oservation	59	54	59	58	58	59	56
rqExpr_HERV_R	0.21	0.62	0.15	0.31	0.15	0.43	1
p-value	0.12	**<.0001**	0.27	**0.02**	0.27	**0.001**	
Number oservation	56	53	56	55	55	56	56

Table 21: *Correlation analysis between the HERV expression in the tumor tissue and the clinical characteristics of the patients.*

		HERV-H					HERV-K					HERV-P					HERV-R				
Tumor stage	**N**	**Mean***	**95% CI**		**P-value**		**Mean***	**95% CI**		**P-value**		**Mean***	**95% CI**		**P-value**		**Mean***	**95% CI**		**P-value**	
pT1	4	0.32	0.06	1.64	ref		0.61	0.15	2.42	ref		0.36	0.04	3.03	ref		0.16	0.03	0.82	ref	
pT2	6	0.81	0.23	2.90	0.38	0.26	0.57	0.19	1.65	0.94	0.51	0.64	0.10	3.96	0.69	0.49	1.20	0.34	4.24	0.06	0.13
pT3	37	1.13	0.72	1.79	0.13		0.97	0.66	1.41	0.53		1.22	0.67	2.19	0.27		1.16	0.74	1.82	0.02	
pT4	10	0.41	0.11	1.55	0.81		0.48	0.16	1.49	0.79		0.47	0.08	2.68	0.84		0.67	0.18	2.53	0.18	
Lymph nodes +/-																					
+	22	1.03	0.53	1.98	-	0.79	0.82	0.48	1.40	-	0.63	0.77	0.33	1.78	-	0.31	0.74	0.38	1.43	-	0.15
-	28	0.92	0.53	1.58	ref		0.97	0.63	1.50	ref		1.34	0.67	2.70	ref		1.38	0.81	2.34	ref	
Tumor location																					
Right colon	40	1.00	0.63	1.56	-	0.61	0.95	0.66	1.38	-	0.51	1.39	0.79	2.46	-	**0.05**	1.32	0.85	2.06	-	**0.045**
Left colon	15	0.87	0.42	1.79	ref		0.76	0.42	1.37	ref		0.48	0.20	1.17	ref		0.56	0.28	1.14	ref	
Lymphocyte infiltration																					
Yes	20	0.75	0.32	1.74	-	**0.05**	0.93	0.48	1.78	-	0.85	0.97	0.31	3.03	-	0.92	0.88	0.38	2.01	-	0.37
No	14	1.06	0.35	3.22	ref		1.02	0.45	2.31	ref		0.89	0.24	3.29	ref		1.58	0.57	4.34	ref	
Vascular invasion																					
Yes	20	1.57	0.77	3.20	-	0.67	1.02	0.56	1.86	-	0.55	1.05	0.38	2.89	-	0.87	0.91	0.42	1.99	-	0.67
No	26	0.62	0.35	1.09	ref		0.81	0.50	1.29	ref		0.94	0.43	2.03	ref		1.13	0.62	2.06	ref	
Perineural infiltration																					
Yes	8	1.09	0.38	3.14	-	0.51	0.94	0.39	2.22	-	0.82	1.24	0.32	4.86	-	0.74	0.66	0.23	1.91	-	0.29
No	38	0.85	0.51	1.41	ref		0.84	0.56	1.27	ref		0.96	0.49	1.89	ref		1.23	0.74	2.05	ref	

Mean*: geometric mean

4.5 Correlation analysis in the Tunisian population

The correlations between the methylation and expression levels in the Tunisian population were evaluated using the Spearman correlation index (Table 22 and Table 23). There was a significant correlation between the HERV-K *env* and HERV-K *pol*, HERV-H, and HERV-P expression in the adjacent normal tissue and in the tumor tissue ($P<0.05$).

In the adjacent normal tissue, a significant correlation was observed between the expression of HERV-K *pol* and methylation status of HERV-K, and -W, and the expression of HERV-P ($P<0.05$). Moreover, the methylation of HERV-H was associated to the methylation of HERV-K and HERV-W ($P<0.05$). In addition, in the tumor tissue the methylation of HERV-P LTRs was correlated to the expression levels of HERV-P and of HERV-K *env* and to the methylation of HERV-H ($P<0.05$).

Regarding the association between HERV *env* and *pol* expression and clinical characteristics, no association was observed between HERV *env* expression and clinical characteristics of the enrolled patients (Table 24).

Table 22: *Correlation analysis between the HERV expression and methylation in the adjacent normal tissue in Tunisian population.*

	rqExpr HERV_P	**rqExpr HERV_H**	**rqExpr HERV_K_env**	**rqExpr HERV_k_pol**	**Met HERV_P**	**Met HERV_H**	**Met HERV_k**	**Met HERV_W**
rqExpr_HERV_P	1	0.01	0.60	0.34	-0.22	-0.22	-0.11	-0.21
p-value		0.96	**<.0001**	**0.02**	0.15	0.13	0.45	0.15
Number oservation	54	54	54	43	45	49	50	48
rqExpr_HERV_H	0.01	1	0.34	0.23	0.07	0.26	-0.16	0.18
p-value	0.96		**0.01**	0.14	0.63	0.07	0.27	0.22
Number oservation	54	56	56	44	46	49	52	50
rqExpr HERV_K_env	0.60	0.34	1	0.59	-0.16	-0.17	-0.23	-0.18
p-value	**<.0001**	**0.01**		**<.0001**	0.27	0.24	0.10	0.22
Number oservation	54	56	56	44	46	49	52	50
rqExpr HERV_K_pol	0.34	0.23	0.59	1	-0.11	-0.03	-0.32	-0.36
p-value	**0.02**	0.14	**<.0001**		0.50	0.87	**0.04**	**0.02**
Number oservation	43	44	44	44	37	39	41	40
Met_HERV_P	-0.22	0.07	-0.16	-0.11	1	0.09	0.09	-0.005
p-value	0.15	0.63	0.27	0.50		0.54	0.54	0.97
Number oservation	45	46	46	37	47	46	47	46
Met_HERV_H	-0.22	0.26	-0.17	-0.03	0.09	1	0.37	0.36
p-value	0.13	0.07	0.24	0.87	0.54		**0.008**	**0.01**
Number oservation	49	49	49	39	46	50	50	48
Met_HERV_K	-0.11	-0.16	-0.23	-0.32	0.09	0.37	1	0.31
p-value	0.45	0.27	0.10	**0.04**	0.54	**0.008**		**0.03**
Number oservation	50	52	52	41	47	50	53	51
Met_HERV_W	-0.21	0.18	-0.18	-0.36	-0.005	0.36	0.30	1
p-value	0.15	0.22	0.22	**0.02**	0.97	**0.01**	**0.03**	
Number oservation	48	50	50	40	46	48	51	51

Table 23: *Correlation analysis between the HERV expression and methylation in the tumor tissue in Tunisian population.*

	rqExpr HERV_P	**rqExpr HERV_H**	**rqExpr HERV_K_env**	**rqExpr HERV_k_pol**	**Met HERV_P**	**Met HERV_H**	**Met HERV_K**	**Met HERV_W**
rqExpr_HERV_P	1	0.10	0.52	0.47	0.40	-0.07	-0.15	-0.12
p-value		0.45	**<.0001**	**0.0008**	**0.01**	0.68	0.33	0.44
Number oservation	55	55	55	48	38	39	43	40
rqExpr_HERV_H	0.10	1	0.37	0.48	0,05	-0.11	0.16	-0.005
p-value	0.45		**0.005**	**0.0005**	0.78	0.50	0.31	0.98
Number oservation	55	56	56	48	38	39	43	40
rqExpr HERV_K_env	0.52	0.37	1	0.53	0.33	-0.03	0.29	-0.23
p-value	**<.0001**	**0.005**		**0.0001**	**0.04**	0.88	0.06	0.14
Number oservation	55	56	56	48	38	39	43	40
rqExpr HERV_K_pol	0.47	0,48	0.53	1	0.23	-0.02	-0,03	-0.45
p-value	**0.0008**	**0.0005**	**0.0001**		0.19	0.92	0.85	**0.006**
Number oservation	48	48	48	48	34	35	39	36
Met_HERV_P	0.40	0.05	0.33	0.23	1	-0.41	0.1	-0.16
p-value	**0.01**	0.78	**0.04**	0.19		**0.01**	0.58	0.35
Number oservation	38	38	38	34	40	38	39	37
Met_HERV_H	-0.07	-0.11	-0.03	-0.02	-0.41	1	0.09	-0.08
p-value	0.68	0.50	0.88	0.92	**0.01**		0.57	0.63
Number oservation	39	39	39	35	38	41	41	36
Met_HERV_K	-0.15	0.16	0.29	-0.03	0.1	0.09	1	0.10
p-value	0.33	0.31	0.06	0.85	0.56	0.57		0.53
Number oservation	43	43	43	39	39	41	45	40
Met_HERV_W	-0.12	-0.005	-0.23	-0.45	-0.16	-0.08	0.10	1
p-value	0.44	0.98	0.14	**0.006**	0.35	0.63	0.53	
Number oservation	40	40	40	36	37	36	40	42

Table 24: *Correlation analysis between the HERV expression in the tumor tissue and the clinical characteristics of the Tunisian patients.*

		HERV-H				HERV-K env				HERV-K pol				HERV-P			
		Mean*	**95% CI**		**P-value**	**Mean***	**95% CI**		**P-value**	**Mean***	**95% CI**		**P-value**	**Mean***	**95% CI**		**P-value**
Tumor location	left	0.95	0.73	1.24	0.86	0.92	0.70	1.21	0.44	0.95	0.56	1.62	0.70	1.02	0.74	1.40	0.36
	left/right	1.13	0.38	3.41		1.46	0.47	4.52		2.99	0.21	42.85		1.40	0.39	4.96	
	right	1.09	0.66	1.78		1.26	0.76	2.09		0.98	0.36	2.63		1.61	0.91	2.84	
Tumor stage	III, IV	0.92	0.70	1.22	0.41	1,00	0.75	1.34	0.94	1.18	0.66	2.09	0.38	1.20	0.86	1.68	0.66
	I, II	1.12	0.77	1.64		1.02	0.68	1.51		0.80	0.40	1.59		1.06	0.67	1.68	
Polyps	+	0.87	0.60	1.27	0.42	1.17	0.79	1.73	0.33	0.89	0.41	1.96	0.71	1.13	0.72	1.76	0.91
	-	1.05	0.80	1.38		0.93	0.70	1.23		1.05	0.62	1.78		1.16	0.83	1.63	
Chemiotherapy	+	0.78	0.55	1.12	0.11	0.93	0.63	1.36	0.60	1.10	0.49	2.44	0.78	1.16	0.74	1.82	0.96
	-	1.13	0.86	1.49		1.05	0.78	1.41		0.97	0.57	1.64		1.14	0.81	1.61	
Histologic analysis	Tubular/ villous	0.98	0.59	1.64	0.98	0.96	0.56	1.64	0.85	0.54	0.22	1.32	0.11	1.02	0.53	1.94	0.68
	Tubular	0.99	0.77	1.27		1.02	0.78	1.32		1.23	0.74	2.05		1.18	0.87	1.59	

Mean*: geometric mean

4.8 Expression of HERV-K Pol gene and proteins in the adjacent and normal tissue

The HERV-K *pol* gene and Pol protein expression was evaluated in the adjacent normal and in the tumor tissue of 7 Italians population's patients, who present the HERV-K *env* mRNA in the EVs: the analysis was performed to assess the association between the mRNA and the protein expression.
The demographic data of the seven patients were reported in the Table 25.

Table 25: *Demographic and clinical data of the patients who showed the HERV-K env transcript in the EVs*

Sex	**N°**
Female/Male	*1/6*
Age	**mean ± SD**
Years	*75.1 ± 18.7*
Tumor stage	**N°**
pT3	5/7
pT4	2/7
Tumor location	**N°**
Right colon	4/7
Left colon	3/7

As for the mRNA expression, the HERV-K *pol* gene was less expressed in the adjacent normal tissue compared to the tumor tissue, but the difference was not statistically significant (*p-value*=0.8) (Figure 18).

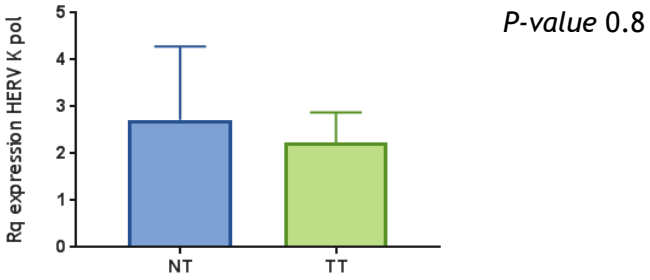


Figure 18: *Rq expression of the HERV-K pol gene in the adjacent normal (NT) and in the tumor tissue (TT).*

But, if the *pol* gene expression was similar in the adjacent normal and tumor tissue, the Pol protein presented a significant difference in the adjacent normal tissue compared to the tumor tissue. The majority of patients expressed the Pol protein in three different complexes: Pol intermediates with a molecular weight of 100 KDalton, the complex RT-RH with a molecular weight of 60 KDalton, and RT with a molecular weight of 54 KDalton. For each patient the GAPDH lane was reported. One patient did not express the HERV-K Pol protein.

The patient number #1 presented all the three different lanes in the adjacent normal tissue, corresponding to the Pol intermediates, the complex RT-RH, and the RT. Instead, in the tumor tissue, only the Pol intermediates of HERV-K were expressed at very low level (Figure 19). The results were also graphically reported.

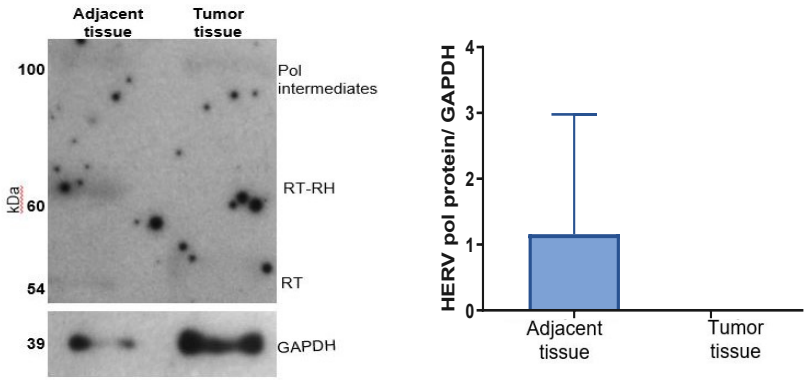


Figure 19: *HERV-K Pol and GAPDH expression in the adjacent normal and in the tumor tissue of patient #1 (left). HERV-K Pol expression normalised on the GAPDH expression (right).*

In the adjacent normal tissue of the patient #2, the Pol intermediates and the RT lanes were very intense compared to the lane at 60 KDalton. In the tumor tissue there was not expression of the Pol protein (Figure 20).

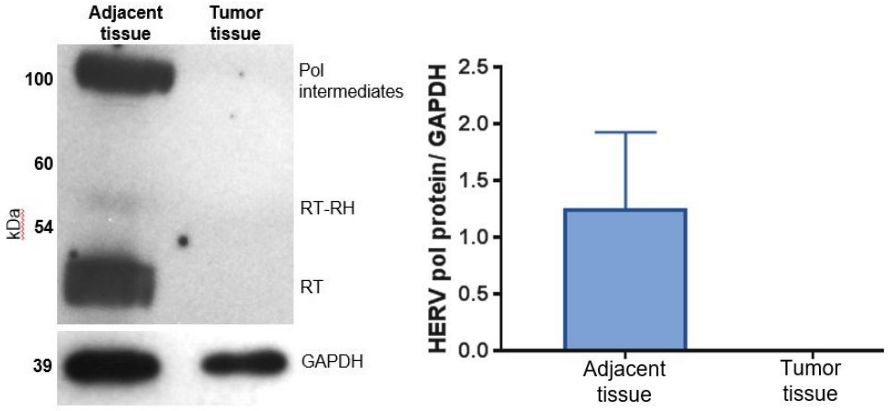


Figure 20: *HERV-K Pol and GAPDH expression in the adjacent normal and in the tumor tissue of patient #2 (left). HERV-K Pol expression normalised on the GAPDH expression (right).*

The patient #3 presented only one lane corresponding to the RT in the adjacent normal tissue. No HERV-K Pol expression was observed in the tumor tissue (Figure 21).

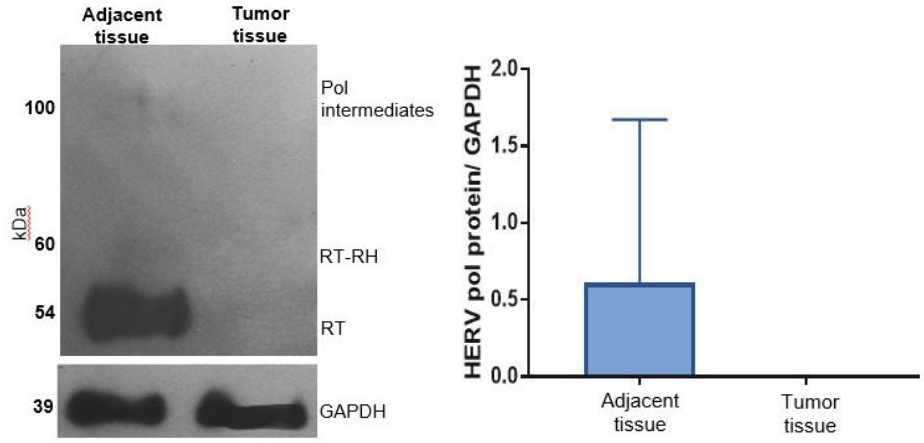


Figure 21: *HERV-K Pol and GAPDH expression in the adjacent normal and in the tumor tissue of the patients #3 (left). HERV-K Pol expression normalised on the GAPDH expression (right).*

The patient #4 expressed the Pol protein in the three different complexes, while in the tumor tissue the Pol intermediates were the only complex expressed (Figure 22).

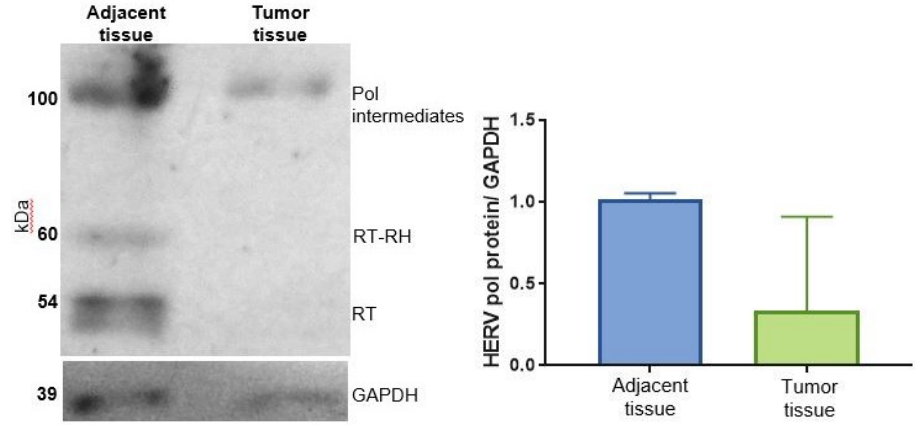


Figure 22: *HERV-K Pol and GAPDH expression in the adjacent normal and in the tumor tissue of the patients #4 (left). HERV-K Pol expression normalised on the GAPDH expression (right).*

The Pol intermediates, the RT-RH complex and the RT were only expressed in the adjacent normal tissue of the patient #5 and no expression of HERV-K Pol was observed in the tumor tissue (Figure 23).

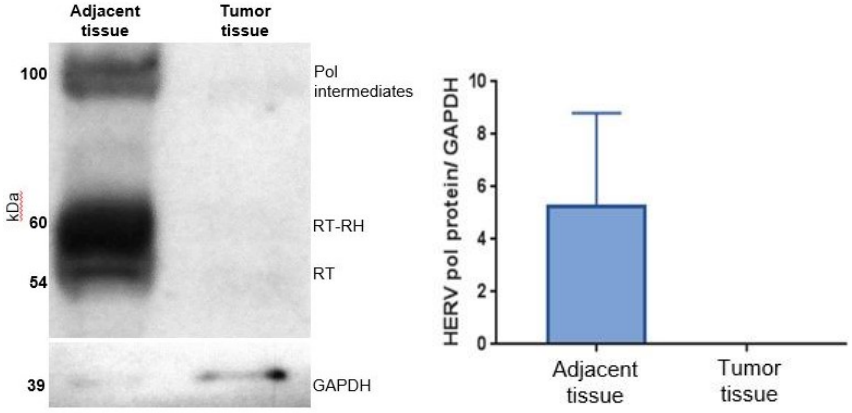


Figure 23: *HERV-K Pol and GAPDH expression in the adjacent normal and in the tumor tissue of the patients #5 (left). HERV-K Pol expression normalised on the GAPDH expression (right).*

Only the patient #6 expressed the HERV-K Pol protein in the same complexes both in the adjacent normal tissue and in the tumor tissue, but it was more expressed in the normal tissue compared to the tumor tissue (Figure 24).

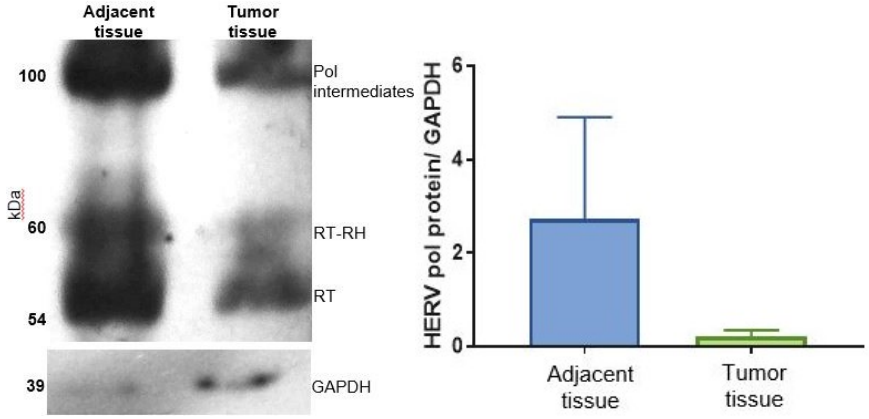


Figure 24: *HERV-K Pol and GAPDH expression in the adjacent normal and in the tumor tissue of the patients #6 (left). HERV-K Pol expression normalised on the GAPDH expression (right).*

The patient number#7 did not present the HERV-K Pol protein both in the tumor tissue both in the normal adjacent tissue.

Taken together our results, in the analysed patients, the HERV-K Pol protein was more expressed in the adjacent normal tissue compared to the tumor tissue and the difference was statistically significant, with a *p-value* of 0.0013, as reported in the Figure 25.

P value 0.0013

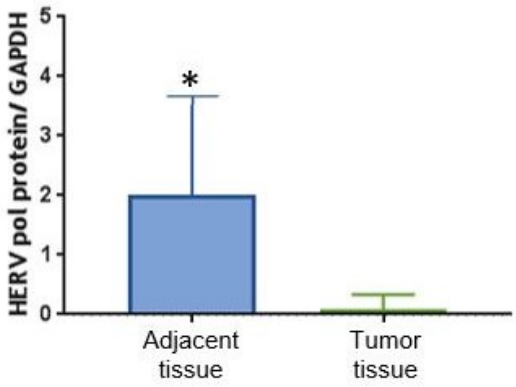


Figure 25: *Mean expression values of HERV-K Pol protein in the adjacent normal and in the tumor tissue of the enrolled patients.*

4.9 Expression of HERV-K Env gene and proteins in the adjacent and normal tissue

The HERV-K *env* gene and Env protein expression was evaluated in the adjacent normal and in the tumor tissue of 7 Italians population's patients, tested also for the presence and for the expression of the HERV-K Pol protein: the analysis was performed to assess the association between the mRNA and the protein expression.

The demographic data of the 7 patients were reported in the Table 25.

In the normal and in the tumor tissue of the patient number #1, the Env protein was not express (Figure 26).

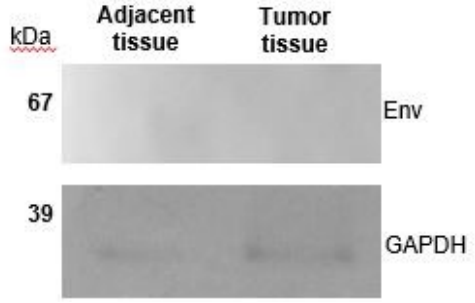


Figure 26: *HERV-K Env and GAPDH expression in the adjacent normal and in the tumor tissue of patient #1.*

The patient number #2 presented the Env protein only in the tumor tissue, even if the expression was very low, as represented in the Figure 27.

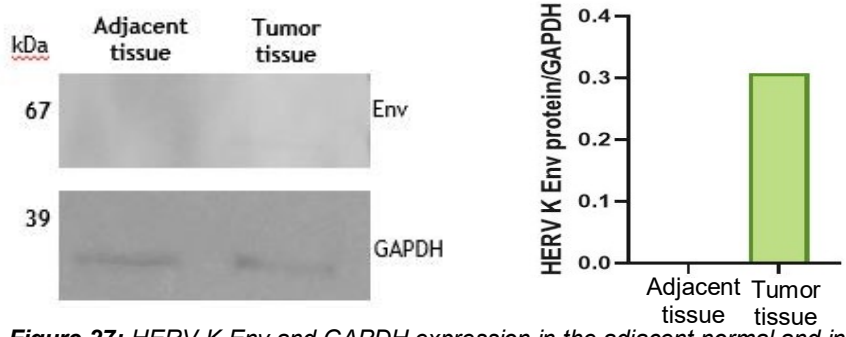


Figure 27: *HERV-K Env and GAPDH expression in the adjacent normal and in the tumor tissue of patient #2 (left). HERV-K Env expression normalised on the GAPDH (right).*

The Env protein was not expressed in the adjacent normal tissue and in the tumor tissue of the patient number #3 (Figure 28).

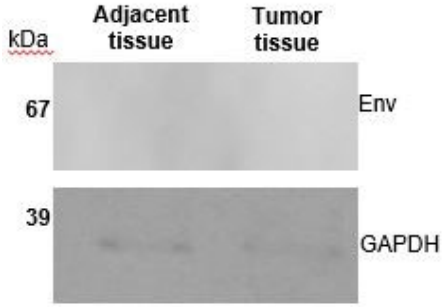


Figure 28: *HERV-K Env and GAPDH expression in the adjacent normal and in the tumor tissue of patient #3.*

The patient number #4 presents the HERV-K Env protein expression in the tumor tissue, while in the adjacent normal tissue there was no expression of the protein (Figure 29).

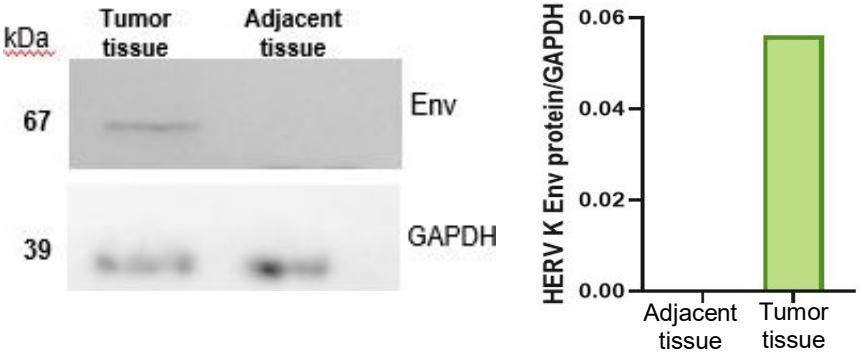


Figure 29: *HERV-K Pol and GAPDH expression in the adjacent normal and in the tumor tissue of the patients #4 (left). HERV-K Env expression normalised on the GAPDH expression (right).*

The Env protein was expressed also in the tumor tissue of the patient number #5. No expression of the HERV-K Env protein was observed in the adjacent normal tissue of the same patient (Figure 30).

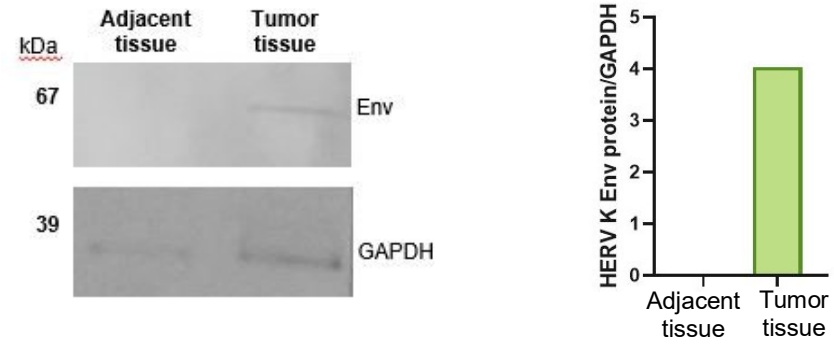


Figure 30: *HERV-K Env and GAPDH expression in the adjacent normal and in the tumor tissue of patient #5 (left). HERV-K Env expression normalised on the GAPDH expression (right).*

No Env protein expression was observed in the adjacent normal and tumor tissue of the patients #6, as reported in the Figure 31.

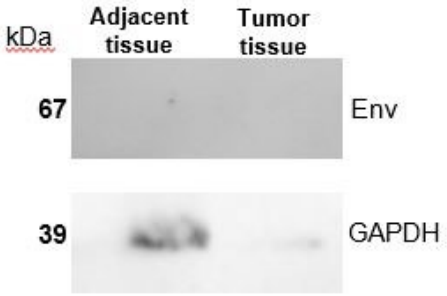


Figure 31: *HERV-K Env and GAPDH expression in the adjacent normal and in the tumor tissue of patient #6.*

In the patient #7 the HERV-K Pol protein was expressed only in the tumor tissue (Figure 32).

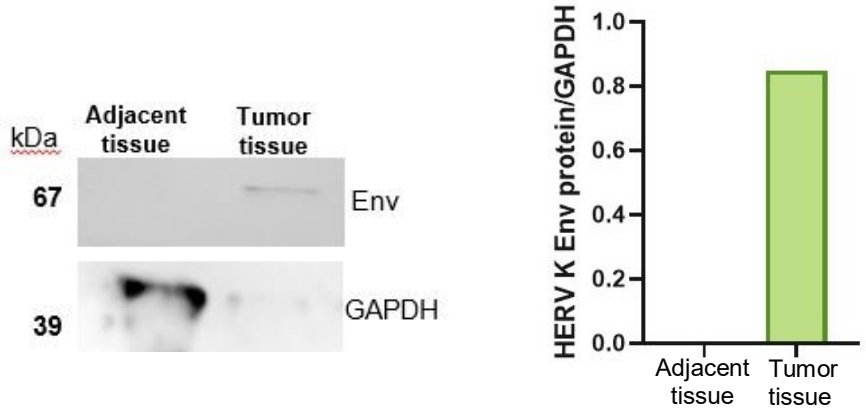


Figure 32: *HERV-K Env and GAPDH expression in the adjacent normal and in the tumor tissue of patient #7 (left). HERV-K Env expression normalised on the GAPDH expression (right).*

Taken together our results, in the patients the HERV-K Env protein was only expressed in the tumor tissue while no expression was observed in the adjacent normal tissue, but the difference was no significant (p-value>0.05); the HERV-K Env expression level in the tissues was reported in the Figure 33.

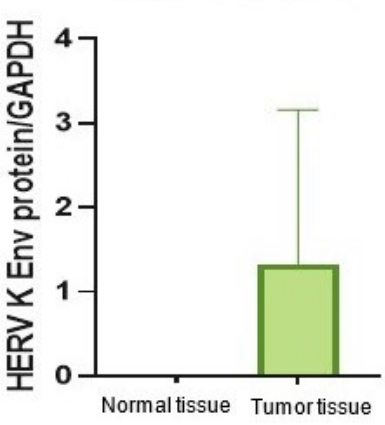


Figure 33: *Mean expression values of HERV-K Env protein in the adjacent normal and in the tumor tissue of the enrolled patients.*

5. DISCUSSION

HERV are remnants of ancient exogenous retroviral infection [6], fixed in the human genome, representing about the 8% of our genome [10]. They can be divided in at least 40 families [13] [14], named after the amino acid carried by the tRNA complementary to the primer binding site of the HERV genome. The HERV genome is composed of four coding genes, *gag*, *pro*, *pol*, and *env* gene, flanked at 3' and at 5' by two regulatory regions, named LTRs [16]. HERV are not able to produce complete virion to infect the cells due to the accumulation of mutation in the ORFs; instead, the LTRs maintain their functions. The HERV transcription is regulated by different mechanisms, such as the methylation of the regulatory regions. HERV play an important role in the development of the placenta but they are also associated to several diseases, such as neurodegenerative pathologies or autoimmune diseases. Several studies reported also the association between HERV and cancer, such as melanoma, breast cancer, and leukemia, but there are few studies regarding the association between HERV and colon cancer, the fourth malignancy in incidence and mortality in the developed countries [113]. HERV can be involved in oncogenesis with many mechanisms. HERV can inducing gene silencing with the production of lnc RNA; the LTRs cause genomic instability and can regulate the transcription of the adjacent genes, such as the oncogenes or tumor-suppressor genes; HERV mediated insertional mutagenesis due to the non-allelic-recombination and they encode for two oncogenic proteins, Rec and Np9; moreover HERV can induce the cell fusion, promoting the cell growth and their sequences can be also the cargo of the tumor-derived retroviral-like microvesicles, involved in the tumour progression [16].

It was reported that some HERV are re-activated in the tumor, resulting in the increase of the HERV expression in the tumor tissue [169]. Regarding the colon cancer, in contrast with the studies that assess the increase of

HERV expression [170], but according with Kim and colleagues [150] [171] [146], in our study the mean expression levels of HERV *env* and *pol* genes were similar in the blood, in the normal adjacent and in the tumor tissue in the Italian population. Interestingly, the results were confirmed by the analysis of the HERV expression in the adjacent normal and in the tumor tissues of other colon cancer patient population, the Tunisian population. But the studies related to the association of HERV and colon cancer were few and they were also in contrast: these data will need further investigation because probably the role of HERV in the tumorigenesis is not only related to the alteration of their expression.

Our population was composed of elderly patients and the majority of these patients were affected with advanced colorectal cancer, stage III and IV. HERV-H expression has been associated with the adenoma, a non-cancerous polyps that can become a tumor: we supposed that HERV expression may increase in the early stage of the transformation; thus, HERV may not play a role in following phases of tumor development [103]. A definite age-related pattern of HERV expression in the population was observed, concluding that HERV are less expressed in the late stage of the life [172] [9]. The stage of CRC and the age of the enrolled patients may be biases of this study.

It is interesting that, in the Italian patients, the difference in the expression of the four HERV families, HERV-H, -K, -P, and -R, with significant values for HERV-P and -R, was observed in the tumor specimens collecting from right-sided (RS) CRC patients compared to those collected from left-sided (LS) CRC patients. It is well known that RS colon tumors present different molecular characteristics from colon tumors located on the left side [173]. In particular, RS colon tumors are mainly characterize by the MSI, and are considered hypermutated tumors, due to the defects in the DNA mismatch repair system, that increases the single nucleotide mutations and alters the

length of repetitive microsatellite sequences [174]. The main mutated genes in the RS colon tumors are for example BRAF, and RAS; these genes are involved in the DNA mismatch repair and in the cell survival and proliferation pathways [175]. On the contrary, LS colon tumors show high frequency of DNA CIN, and non-hypermutated tumors [176]. Interestingly, RS colon tumors show often a CIMP-high, that induce the silencing of the gene function with aberrant hypermethylation and the following involvement in the carcinogenetic progression [177] [178] [179]. This may not be the case of the HERV families in our Italian population, because they were more expressed, despite a higher level of methylation in the right side compared to the tumor tissue located at the left side. Unfortunately, the clinical analyses conducted by the two Italian clinics and by the Tunisian clinics, where the patients underwent surgery, did not include the molecular characterization of the tumor biopsies.

To the best of our knowledge, this is the first time that a difference in expression of HERV is reported in RS compared to LS tumor tissues. The meaning of this differential expression, as well as its consequences in terms of susceptibility to immune therapies, needs to be further analyzed.

The HERV expression, especially HERV-K and -H, were detected in the EVs isolated from patients' plasma, at higher levels compared to the other biological samples (blood, adjacent normal tissue, and tumor tissue) collected from the same patients. In a study conducted by Balaj and colleagues it has been reported that tumor-derived EVs are characterized by the presence of the retrotransposon elements, such as LINE-1 and Alu, and of HERV-K [97]. These results were derived from *in vitro* observations [97], but, to the best of our knowledge, there are no reports regarding the presence of HERV sequences in the EVs derived from clinical samples. It is interesting that the HERV sequences were able to be packaged in the plasmatic EVs, which may potentially be transferred from one cell to another.

This fact confirmed that retrotransposable elements can create a dynamic environment, inducing new insertions and, potentially, mutations, deletions, rearrangements, and changes in gene expression and in cell mobility [180].
In the tumor tissue, the hypomethylation of the LTRs sequences, especially for the HERV-H and HERV-K, did not result in an overexpression of the HERV *env* and *pol* gene: in fact, HERV *env* and *pol* gene were expressed at low level in the tumor tissue.
As expected, and as confirmed by numerous previous reports, the methylation level of the repeat elements Alu and LINE-1 was lower in the tumor tissues compared to the adjacent normal tissue, showing the well-known epigenetic alteration linked to the colon cancer [134] [181]. The LTRs methylation showed a very specific trend, especially for HERV-H and HERV-K, where the LTRs in the tumor tissues were demethylated compared to the LTRs in the normal adjacent tissues. The HERV hypomethylation could be a consequence of global hypomethylation in the cancer cells. These results were not confirmed by the analysis of the methylation status of the HERV-W LTRs, an HERV that seems to be not involved in the cancer pathogenesis, and of a control gene, *psen*. The methylation levels of HERV-W LTRs and *psen* were similar and at high levels in the blood, in the adjacent tissue and in the tumor tissue, strongly supporting the hypothesis that alterations in the methylation status are sequence-specific.
Most published studies described a correlation between LTR methylation and HERV expression: for example, Liang et al. observed as the HERV-H expression of transcripts were different in the colon cancer cell lines could be affected by the changes in the DNA methylation levels [149]. In particular, Liang and colleagues modified the methylation level of HERV using specific inhibitors: this inhibition induced the change in the expression patter of HERV-H in two different cell lines, the SW480 cells, isolated from colorectal adenocarcinoma, and RKO cells, isolated from colon carcinoma [182]. These

results suggested that the epigenetic alteration affected the expression of HERV elements. In contrast, there are studies regarding the changes in the DNA methylation, especially the hypomethylation, that do not affect the expression of the retro-elements and they do not lead to HERV over-expression. In particular, the studies reported the hypomethylation of HERV retroelements in bladder cancer cell lines and tissues and in neuroblastoma cells without any activation of the retroviruses [183] [184]. In addition, it was demonstrated that the chemical treatment of neuroblastoma cell lines with 5-azacytidine, a demethylating agent, induced the expression of multiple HERV-W loci [185]: this result assessed that the HERV expression the cancer cell line could be suppressed by CpG island methylation.

Furthermore, LTRs can serve as promoters and they can alter the expression of host cellular genes. We cannot exclude that the hypomethylation of the HERV LTRs may affect adjacent oncogenic sequences that are differently subjected to silencing [87]. In fact, Romanish and colleagues demonstrated that the LTRs affect the expression of the neuronal apoptosis inhibitory protein, a member of the inhibitor of apoptosis family, acting as a promoter of this gene [87].

Regarding the HERV protein expression, in our study we analyzed the expression of Pol and Env HERV-K proteins. The Pol protein in our samples were processed by the proteolytic cleavage, mediated by the Pro protein, in three different complexes: the Pol intermediates, the RT-RH and RT. The difference of HERV-K Pol was significant among the biological samples analyzed. In particular, the Polymerase protein of HERV-K was more expressed in a significant way in the adjacent normal tissue compared to the tumor tissue, even if the HERV-K *pol* gene expression was similar in the tissues. Probably there were alteration in the post-trascriptional mechanisms and/ or in the post-tradutional changes that interested the protein half-life [186].

In literature there are few reports regarding the HERV-K Pol protein expression, localization and the composition of the isoform, in cancer patients or in healthy subjects [168] and its role in cancer pathogenesis is still unclear. The higher expression in the adjacent normal tissue may be involved in the reverse-transcription of HERV RNA into DNA and in the integration into the host genome, contributing to the cellular transformation.
Instead, the Envelope protein of HERV-K was only expressed in the tumor tissue of the analyzed patients, while in the adjacent normal tissue there was no expression. Also, in this case, the *env* gene expression was similar among the adjacent normal and the tumor tissue. The Env protein of HERV-K is composed of two subunits: the surface subunit and the transmembrane subunit; the first domain is involved in the cell to cell fusion and in the syncytiotrophoblast formation, while the second domain presents immunosuppressive activity [187]. The fusion of the tumor cells with normal cells induces the cancer growth and the formation of metastasis. The transmembrane domain instead facilitates the immune-evasion of the tumor cell and preventing the apoptosis of the tumor cells. Our results were confirmed by several reports in literature. In particular, it was reported the over-expression of the Env protein in the surface and in the cytoplasm of the ovarian cancer cells compared to the normal ovarian cells [76]. The HERV-K Env protein expression was also reported in the 88% of tissues collected from the breast cancer patients, while no expression was observed in the adjacent normal tissue [106]. The expression of the Env protein only in the tumor cells, especially if protein is expressed on the surface of the cells, could be used as a target for the immunotherapy, even if the HERV protein does not act directly in the tumor pathogenesis [187].

6. CONCLUSION

The role of HERV in colon cancer pathogenesis is still unclear and further studies are needed in order to clarify if the presence of HERV elements has a role in cancer initiation and/or in the progression of the colon cancer. Our study was focused on unraveling the role of HERV in colon cancer. In particular we aimed to discover whether HERV are differentially regulated and expressed in colon cancer patients and whether HERV could “go around” through extracellular vesicles, contributing to the malignant transformation. Based on our results, we drew a hypothetical model for HERV involvement in colorectal cancer pathogenesis, as represented in the Figure 34; first, we supposed that HERV expression may increase in the early stage of colon tumorigenesis, thus HERV may not play a role in subsequent phases of tumor progression. To this regard, in a study conducted by Peròt and colleagues, it was assessed that the overexpression of HERV-H was associated with the adenoma, a non-cancerous polyps that can become a tumor: we supposed that HERV expression may increase in the early stage of colon tumorigenesis, while lower HERV-H expression levels were detected in the CRC specimens and in the metastases [103].

In the advanced phases probably HERV overexpression may not be needed anymore, because the complete transformation of the cells has already occurred. Studies are needed to clarify the changes in the HERV expression during the single tumor stages: it will be also important to enroll patients affected by early stages of the tumor and patients affected by polyps to investigate the role of HERV in the cancer pathogenesis.

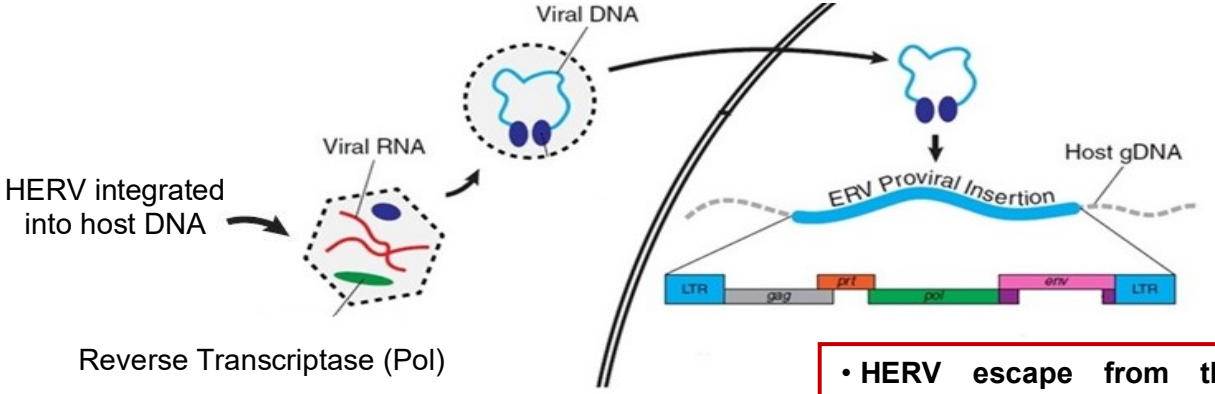
Moreover, in our model we also hypothesized that in the early stages of the pathogenesis, and in the normal cells, HERV may be reverse-transcribed by the Pol protein, that seems to be overexpressed. Then the HERV DNA sequences may be integrated into the host genome, triggering the transformation of the cells, due to the induction of genomic instability. For

example, HERV may be integrated in oncogene o onco-suppressor gene sequences, affecting the molecular pattern into the cells. Additionally, few copies of HERV, as RNA, might be packaged into EVs and transferred to other cells, where they can contribute to cellular genomic instability with the integration and the recombination. In the tumor cells, where the malignant transformation has already occurred, HERV may escape from the epigenetic mechanism; thus the hypomethylation of the LTRs does not act as a promoter of the HERV expression. In addition, in the tumor cells, the expression of the HERV Env protein may contribute to the promotion and to the growth of the tumor: in literature, it is known that the Env protein is associated to the cell to cell fusion [42]. This protein may play the same role also during the tumor progression. Increasing the number of the tissue in which analyzed the protein expression and extending the analysis to other HERV proteins will be the important steps for the future analysis to confirm the results. The different HERV proteins may have strong potentialities as biomarkers of diseases, and/or as tumor-associated antigens, and for immunotherapeutic approach, including vaccines. To evaluate the potentiality of HERV as target of the immunotherapy, the primary cell lines, isolated from the tissue of the patients will be co-cultured with dendritic cells-loaded with HERV Env and Pol proteins, and CD8 positive T cells obtained from colon cancer patients. Based on the results, HERV proteins would be considered as possible tumor-associated antigens. Other important step for the future analysis will be to verify whether HERV mRNA found in colon cancer plasmatic EVs may act as biomarker for diagnosis and/or for prognosis.

Then, the next perspectives of this project will be focused on define the basis for future diagnostic and therapeutic studies, considering HERV as a potential novel biomarker for colon cancer diagnosis and/or prognosis, using readily available body fluids and a minimally invasive procedure.

Normal cell

Tumor cell



- **HERV escape from the epigenetic mechanisms**
- **Production of Env protein, promoting cell to cell fusion**

Modified from T.J.Meyer et al., 2017

Figure 34: *Proposed model of HERV involvement in colorectal cancer pathogenesis. In the normal cells, HERV may be retro-transcripted by the Pol protein into DNA and integrated into the host genome, inducing the cells transformation. In the transformed cells, LTR hypomethylation seems to be not result in promoter activation of the HERV expression. The expression of Env protein in the tumor cells may promotes the tumorigenesis through the cell to cell fusion.*

7. REFERENCES

[1] "The European Code of Conduct for Research Integrity_ALTEA."

[2] S. Delbue, L. Signorini, S. Villlani, M. Bregni, and P. Ferrante, "Do the Human Endogenous Retroviruses Play a Role in Colon Cancer?," *Adv. Tumor Virol.*, no. September, p. 11, 2016.

[3] C. R. Beck, J. L. Garcia-Perez, R. M. Badge, and J. V. Moran, "LINE-1 Elements in Structural Variation and Disease," *Annu. Rev. Genomics Hum. Genet.*, vol. 12, no. 1, pp. 187–215, 2011.

[4] G. R. Young, J. P. Stoye, and G. Kassiotis, "Are human endogenous retroviruses pathogenic? An approach to testing the hypothesis," *BioEssays*, vol. 35, no. 9, pp. 794–803, 2013.

[5] F. Zhu, "The role of human endogenous retroviral long terminal repeat sequences in human cancer (Review)," *Int. J. Mol. Med.*, pp. 755–762, 2013.

[6] G. Kassiotis, J. P. Stoye, G. Kassiotis, and J. P. Stoye, "Making a virtue of necessity : the pleiotropic role of human endogenous retroviruses in cancer," vol. 3, 2017.

[7] G.-C. María, I. Paola, K. Niki, S. Mariacarmela, B. Julià, and R. Rafael, "Human endogenous retroviruses and cancer," *Cancer Biol. Med.*, vol. 13, no. 4, p. 483, 2016.

[8] N. Grandi and E. Tramontano, "Type W human endogenous retrovirus (HERV-W) integrations and their mobilization by L1 machinery: Contribution to the human transcriptome and impact on the host physiopathology," *Viruses*, vol. 9, no. 7, pp. 1–37, 2017.

[9] E. Balestrieri *et al.*, "Transcriptional Activity of Human Endogenous Retroviruses in Human Peripheral Blood Mononuclear Cells," vol. 2015, 2015.

[10] Q. Liang, J. Ding, R. Xu, Z. Xu, and S. Zheng, "Identification of a novel human endogenous retrovirus and promoter activity of its 5′ U3," *Biochem. Biophys. Res. Commun.*, vol. 382, no. 2, pp. 468–472, 2009.

[11] G. R. Young, J. P. Stoye, and G. Kassiotis, "Prospects & Overviews Are

human endogenous retroviruses pathogenic? An approach to testing the hypothesis," *BioEssays*, vol. 35, no. 9, pp. 794–803, 2013.

[12] X. Wang, J. Huang, and F. Zhu, "Human endogenous retroviral envelope protein Syncytin-1 and inflammatory abnormalities in neuropsychological diseases," *Front. Psychiatry*, vol. 9, no. SEP, pp. 1–9, 2018.

[13] R. Microarray *et al.*, "Comprehensive Analysis of Human Endogenous Retrovirus Transcriptional Activity in Human Tissues with a," *Society*, vol. 79, no. 1, pp. 341–352, 2005.

[14] N. Bannert and R. Kurth, "The Evolutionary Dynamics of Human Endogenous Retroviral Families," *Annu. Rev. Genomics Hum. Genet.*, vol. 7, no. 1, pp. 149–173, 2006.

[15] J. Chen, M. Foroozesh, and Z. Qin, "Transactivation of human endogenous retroviruses by tumor viruses and their functions in virus-associated malignancies," *Oncogenesis*, vol. 8, no. 1, 2019.

[16] M. Zhang, J. Q. Liang, and S. Zheng, "Expressional activation and functional roles of human endogenous retroviruses in cancers," *Rev. Med. Virol.*, no. November 2018, pp. 1–11, 2019.

[17] A. A. Buzdin, V. Prassolov, and A. V. Garazha, "Friends-Enemies: Endogenous Retroviruses Are Major Transcriptional Regulators of Human DNA," *Front. Chem.*, vol. 5, no. June, pp. 1–8, 2017.

[18] M. Garcia-Montojo, T. Doucet-O'Hare, L. Henderson, and A. Nath, "Human endogenous retrovirus-K (HML-2): a comprehensive review," *Crit. Rev. Microbiol.*, vol. 44, no. 6, pp. 715–738, 2018.

[19] R. F. Downey, F. J. Sullivan, F. Wang-Johanning, S. Ambs, F. J. Giles, and S. A. Glynn, "Human endogenous retrovirus K and cancer: Innocent bystander or tumorigenic accomplice?," *Int. J. Cancer*, vol. 137, no. 6, pp. 1249–1257, 2015.

[20] K. Schmitt, K. Heyne, K. Roemer, E. Meese, and J. Mayer, "HERV-K(HML-2) Rec and Np9 transcripts not restricted to disease but present in many normal human tissues," *Mob. DNA*, vol. 6, no. 1, pp. 1–13, 2015.

[21] A. Buzdin, E. Kovalskaya-Alexandrova, E. Gogvadze, and E. Sverdlov, "At Least 50% of Human-Specific HERV-K (HML-2) Long Terminal Repeats

Serve In Vivo as Active Promoters for Host Nonrepetitive DNA Transcription," *J. Virol.*, vol. 80, no. 21, pp. 10752–10762, 2006.

[22] A. Manuscript, "Elements in the Placenta," vol. 45, no. 3, pp. 325–329, 2013.

[23] M. Suntsova, A. Garazha, A. Ivanova, D. Kaminsky, A. Zhavoronkov, and A. Buzdin, "Molecular functions of human endogenous retroviruses in health and disease," *Cell. Mol. Life Sci.*, vol. 72, no. 19, pp. 3653–3675, 2015.

[24] J. M. Young *et al.*, "DUX4 Binding to Retroelements Creates Promoters That Are Active in FSHD Muscle and Testis," *PLoS Genet.*, vol. 9, no. 11, 2013.

[25] P. Dessen and T. Heidmann, "Identification , Phylogeny , and Evolution of Retroviral Elements Based on Their Envelope Genes," vol. 75, no. 23, pp. 11709–11719, 2001.

[26] I. Katoh and S. Kurata, "Association of Endogenous Retroviruses and Long Terminal Repeats with Human Disorders," *Front. Oncol.*, vol. 3, no. September, pp. 1–8, 2013.

[27] T. P. Hurst and G. Magiorkinis, "Epigenetic control of human endogenous retrovirus expression: Focus on regulation of long-terminal repeats (LTRs)," *Viruses*, vol. 9, no. 6, pp. 1–13, 2017.

[28] J. B. Johnston, C. Silva, J. Holden, K. G. Warren, A. W. Clark, and C. Power, "Monocyte activation and differentiation augment human endogenous retrovirus expression: Implications for inflammatory brain diseases," *Ann. Neurol.*, vol. 50, no. 4, pp. 434–442, 2001.

[29] J.-R. Lee, K. Ahn, Y.-J. Kim, Y.-D. Jung, and H.-S. Kim, " Radiation-Induced Human Endogenous Retrovirus (HERV)-R env Gene Expression by Epigenetic Control ," *Radiat. Res.*, vol. 178, no. 5, pp. 379–384, 2012.

[30] R. Contreras-Galindo, P. López, R. Vélez, and Y. Yamamura, " HIV-1 Infection Increases the Expression of Human Endogenous Retroviruses Type K (HERV-K) in Vitro ," *AIDS Res. Hum. Retroviruses*, vol. 23, no. 1, pp. 116–122, 2007.

[31] H. J. Kwun, H. J. Han, W. J. Lee, H. S. Kim, and K. L. Jang, "Transactivation of the human endogenous retrovirus K long terminal repeat by herpes simplex virus type 1 immediate early protein 0.," *Virus Res.*, vol. 86, no. 1–2, pp. 93–100, 2002.

[32] M. Manghera and R. N. Douville, “Endogenous retrovirus-K promoter: A landing strip for inflammatory transcription factors?,” *Retrovirology*, vol. 10, no. 1, pp. 1–11, 2013.

[33] S. Szpakowski *et al.*, “NIH Public Access,” vol. 448, no. 2, pp. 151–167, 2010.

[34] T. Hurst *et al.*, “Human endogenous retrovirus (HERV) expression is not induced by treatment with the histone deacetylase (HDAC) inhibitors in cellular models of HIV-1 latency,” *Retrovirology*, vol. 13, no. 1, pp. 1–11, 2016.

[35] C. H. White *et al.*, “Transcriptional modulation of human endogenous retroviruses in primary CD4+ T cells following vorinostat treatment,” *Front. Immunol.*, vol. 9, no. APR, pp. 1–10, 2018.

[36] M. Daskalakis *et al.*, “Reactivation of endogenous retroviral elements via treatment with DNMT- and HDAC-inhibitors,” *Cell Cycle*, vol. 4101, pp. 1–12, 2018.

[37] R. Human *et al.*, “Cytokine Regulation of env Gene Expression of Human Endogenous,” vol. 93, no. 1, pp. 75–80, 1999.

[38] T. Nevalainen, A. Autio, B. H. Mishra, S. Marttila, M. Jylhä, and M. Hurme, “Aging-associated patterns in the expression of human endogenous retroviruses,” *PLoS One*, vol. 13, no. 12, pp. 1–11, 2018.

[39] B. Cakmak Guner, E. Karlik, S. Marakli, and N. Gozukirmizi, “Detection of HERV-K6 and HERV-K11 transpositions in the human genome,” *Biomed. Reports*, pp. 53–59, 2018.

[40] M. Mangeney *et al.*, “Placental syncytins: Genetic disjunction between the fusogenic and immunosuppressive activity of retroviral envelope proteins,” *Proc. Natl. Acad. Sci.*, vol. 104, no. 51, pp. 20534–20539, 2007.

[41] H. Finnerty *et al.*, “Syncytin is a captive retroviral envelope protein involved in human placental morphogenesis,” *Nature*, vol. 403, no. 6771, pp. 785–789, 2002.

[42] “https://www.uniprot.org/uniprot/Q9UQF0.”

[43] J.-L. Frendo *et al.*, “Direct Involvement of HERV-W Env Glycoprotein in Human Trophoblast Cell Fusion and Differentiation,” *Mol. Cell. Biol.*, vol. 23, no. 10, pp. 3566–3574, 2003.

[44] A. Malassiné *et al.*, “Expression of the Fusogenic HERV-FRD Env Glycoprotein (Syncytin 2) in Human Placenta is Restricted to Villous Cytotrophoblastic Cells,” *Placenta*, vol. 28, no. 2–3, pp. 185–191, 2007.

[45] N. Grandi and E. Tramontano, “HERV Envelope Proteins : Physiological Role and Pathogenic Potential in Cancer and Autoimmunity,” *Front. Microbiol.*, vol. 9, no. March, pp. 1–26, 2018.

[46] M. Suntsova *et al.*, “Human-specific endogenous retroviral insert serves as an enhancer for the schizophrenia-linked gene PRODH,” *Proc. Natl. Acad. Sci.*, vol. 110, no. 48, pp. 19472–19477, 2013.

[47] M. Ohnuki *et al.*, “Dynamic regulation of human endogenous retroviruses mediates factor-induced reprogramming and differentiation potential,” *Proc. Natl. Acad. Sci.*, vol. 111, no. 34, pp. 12426–12431, 2014.

[48] T. Brudek, T. Christensen, L. Aagaard, T. Petersen, H. J. Hansen, and A. Møller-Larsen, “B cells and monocytes from patients with active multiple sclerosis exhibit increased surface expression of both HERV-H Env and HERV-W Env, accompanied by increased seroreactivity,” *Retrovirology*, vol. 6, pp. 1–13, 2009.

[49] T. Christensen, “Human endogenous retroviruses in neurologic disease,” *APMIS*. 2016.

[50] S. Sotgiu *et al.*, “Multiple sclerosis-associated retrovirus and MS prognosis: An observational study,” *Neurology*, vol. 59, no. 7, pp. 1071–1073, 2002.

[51] B. A. Nexø *et al.*, “The etiology of multiple sclerosis: Genetic evidence for the involvement of the human endogenous retrovirus HERV-Fc1,” *PLoS One*, vol. 6, no. 2, pp. 1–5, 2011.

[52] J. M. Antony *et al.*, “Human endogenous retrovirus glycoprotein-mediated induction of redox reactants causes oligodendrocyte death and demyelination,” *Nat. Neurosci.*, vol. 7, no. 10, pp. 1088–1095, 2004.

[53] D. Kremer *et al.*, “Human endogenous retrovirus type W envelope protein inhibits oligodendroglial precursor cell differentiation,” *Ann. Neurol.*, vol. 74, no. 5, pp. 721–732, 2013.

[54] D.-W. Rhyu *et al.*, “Expression of Human Endogenous Retrovirus env Genes in the Blood of Breast Cancer Patients,” *Int. J. Mol. Sci.*, vol. 15, no. 6, pp.

9173–9183, 2014.

[55] T. Muster *et al.*, “An Endogenous Retrovirus Derived from Human Melanoma Cells An Endogenous Retrovirus Derived from Human Melanoma Cells,” *Cancer Res.*, vol. 63, pp. 8735–8741, 2003.

[56] T. Ishida *et al.*, “Identification of the HERV-K gag antigen in prostate cancer by SEREX using autologous patient serum and its immunogenicity,” *Cancer Immun.*, vol. 8, no. November, pp. 1–10, 2008.

[57] K. Büscher, U. Trefzer, M. Hofmann, W. Sterry, R. Kurth, and J. Denner, “Expression of human endogenous retrovirus K in melanomas and melanoma cell lines,” *Cancer Res.*, vol. 65, no. 10, pp. 4172–4180, 2005.

[58] K. Buscher *et al.*, “Expression of the human endogenous retrovirus-K transmembrane envelope, Rec and Np9 proteins in melanomas and melanoma cell lines.,” *Melanoma Res.*, vol. 16, no. 3, pp. 223–234, Jun. 2006.

[59] F. Schiavetti, J. Thonnard, D. Colau, T. Boon, and P. G. Coulie, “A Human Endogenous Retroviral Sequence Encoding an Antigen Recognized on Melanoma by Cytolytic T Lymphocytes A Human Endogenous Retroviral Sequence Encoding an Antigen Recognized on Melanoma by Cytolytic T Lymphocytes 1,” *Cancer Res.*, pp. 5510–5516, 2002.

[60] M. Sauter *et al.*, “Human endogenous retrovirus K10: expression of Gag protein and detection of antibodies in patients with seminomas.,” *J. Virol.*, vol. 69, no. 1, pp. 414–21, 1995.

[61] O. Frank *et al.*, “Variable Transcriptional Activity of Endogenous Retroviruses in Human Breast Cancer,” *J. Virol.*, vol. 82, no. 4, pp. 1808–1818, 2008.

[62] R. Contreras-Galindo *et al.*, “Human Endogenous Retrovirus K (HML-2) Elements in the Plasma of People with Lymphoma and Breast Cancer,” *J. Virol.*, vol. 82, no. 19, pp. 9329–9336, 2008.

[63] F. Wang-Johanning *et al.*, “Detecting the expression of human endogenous retrovirus E envelope transcripts in human prostate adenocarcinoma,” *Cancer*, vol. 98, no. 1, pp. 187–197, 2003.

[64] M. Golan *et al.*, “Human Endogenous Retrovirus (HERV-K) Reverse Transcriptase as a Breast Cancer Prognostic Marker,” *Neoplasia*, vol. 10, no.

6, pp. 521-IN2, 2008.

[65] H. Davari Ejthadi *et al.*, “A novel multiplex RT-PCR system detects human endogenous retrovirus-K in breast cancer,” *Arch. Virol.*, vol. 150, no. 1, pp. 177–184, 2005.

[66] H. Iwabuchi *et al.*, “A gene homologous to human endogenous retrovirus overexpressed in childhood acute lymphoblastic leukemia.,” *Leuk. Lymphoma*, vol. 45, no. 11, pp. 2303–2306, Nov. 2004.

[67] M. Simon, M. Haltmeier, G. Papakonstantinou, T. Werner, R. Hehlmann, and C. Leib-Mosch, “Transcription of HERV-K-related LTRs in human placenta and leukemic cells.,” *Leukemia*, vol. 8 Suppl 1, pp. S12-7, Apr. 1994.

[68] B. K. Prusty, H. zur Hausen, R. Schmidt, R. Kimmel, and E.-M. de Villiers, “Transcription of HERV-E and HERV-E-related sequences in malignant and non-malignant human haematopoietic cells.,” *Virology*, vol. 382, no. 1, pp. 37–45, Dec. 2008.

[69] S. Patzke, M. Lindeskog, E. Munthe, and H. C. Aasheim, “Characterization of a novel human endogenous retrovirus, HERV-H/F, expressed in human leukemia cell lines.,” *Virology*, vol. 303, no. 1, pp. 164–173, Nov. 2002.

[70] G. Mameli *et al.*, “Brains and peripheral blood mononuclear cells of multiple sclerosis (MS) patients hyperexpress MS-associated retrovirus/HERV-W endogenous retrovirus, but not human herpesvirus 6,” *J. Gen. Virol.*, vol. 88, no. 1, pp. 264–274, 2007.

[71] N. Tomita *et al.*, “Transcription of human endogenous retroviral long terminal repeat (LTR) sequence in a lung cancer cell line.,” *Biochem. Biophys. Res. Commun.*, vol. 166, no. 1, pp. 1–10, Jan. 1990.

[72] A. C. Andersson, A. C. Svensson, C. Rolny, G. Andersson, and E. Larsson, “Expression of human endogenous retrovirus ERV3 (HERV-R) mRNA in normal and neoplastic tissues.,” *Int. J. Oncol.*, vol. 12, no. 2, pp. 309–313, Feb. 1998.

[73] F. H. Schmitz-Winnenthal *et al.*, “Potential target antigens for immunotherapy in human pancreatic cancer,” *Cancer Lett.*, vol. 252, no. 2, pp. 290–298, 2007.

[74] N. Wentzensen *et al.*, “Expression of an endogenous retroviral sequence

from the HERV-H group in gastrointestinal cancers," *Int. J. Cancer*, vol. 121, no. 7, pp. 1417–1423, 2007.

[75] A. Willer *et al.*, "Two groups of endogenous MMTV related retroviral env transcripts expressed in human tissues.," *Virus Genes*, vol. 15, no. 2, pp. 123–133, 1997.

[76] F. Wang-Johanning *et al.*, "Expression of multiple human endogenous retrovirus surface envelope proteins in ovarian cancer," *Int. J. Cancer*, vol. 120, no. 1, pp. 81–90, 2007.

[77] L. Hu, D. Hornung, R. Kurek, H. Östman, J. Blomberg, and A. Bergqvist, "Expression of Human Endogenous Gammaretroviral Sequences in Endometriosis and Ovarian Cancer," *AIDS Res. Hum. Retroviruses*, vol. 22, no. 6, pp. 551–557, 2006.

[78] N. Götzinger, M. Sauter, K. Roemer, and N. Mueller-Lantzsch, "Regulation of human endogenous retrovirus-K Gag expression in teratocarcinoma cell lines and human tumours," *J. Gen. Virol.*, vol. 77, no. 12, pp. 2983–2990, 1996.

[79] R. Lower, J. Lower, C. Tondera-Koch, and R. Kurth, "A general method for the identification of transcribed retrovirus sequences (R-U5 PCR) reveals the expression of the human endogenous retrovirus loci HERV-H and HERV-K in teratocarcinoma cells.," *Virology*, vol. 192, no. 2, pp. 501–511, Feb. 1993.

[80] R. Strick *et al.*, "Proliferation and cell-cell fusion of endometrial carcinoma are induced by the human endogenous retroviral Syncytin-1 and regulated by TGF-beta.," *J. Mol. Med. (Berl).*, vol. 85, no. 1, pp. 23–38, Jan. 2007.

[81] S. Rakoff-Nahoum *et al.*, "Detection of T Lymphocytes Specific for Human Endogenous Retrovirus K (HERV-K) in Patients with Seminoma," *AIDS Res. Hum. Retroviruses*, vol. 22, no. 1, pp. 52–56, 2006.

[82] H. Herbst *et al.*, "Human endogenous retrovirus (HERV)-K transcripts in gonadoblastomas and gonadoblastoma-derived germ cell tumours.," *Virchows Arch.*, vol. 434, no. 1, pp. 11–15, Jan. 1999.

[83] T. Vinogradova, L. Leppik, E. Kalinina, P. Zhulidov, K. H. Grzeschik, and E. Sverdlov, "Selective Differential Display of RNAs containing Interspersed Repeats: Analysis of changes in the transcription of HERV-K LTRs in germ cell tumors," *Mol. Genet. Genomics*, vol. 266, no. 5, pp. 796–805, 2001.

[84] L. Li, T. Feng, Y. Lian, G. Zhang, A. Garen, and X. Song, “Role of human noncoding RNAs in the control of tumorigenesis,” *Proc. Natl. Acad. Sci.*, vol. 106, no. 31, pp. 12956–12961, 2009.

[85] J. Lv and Z. Zhao, “Binding of LINE-1 RNA to PSF transcriptionally promotes GAGE6 and regulates cell proliferation andtumor formation in vitro,” *Exp. Ther. Med.*, vol. 14, no. 2, pp. 1685–1691, 2017.

[86] E. Oricchio, I. Sciamanna, R. Beraldi, G. V. Tolstonog, G. G. Schumann, and C. Spadafora, “Distinct roles for LINE-1 and HERV-K retroelements in cell proliferation, differentiation and tumor progression,” *Oncogene*, vol. 26, no. 29, pp. 4226–4233, 2007.

[87] M. T. Romanish, W. M. Lock, L. N. Van De Lagemaat, C. A. Dunn, and D. L. Mager, “Repeated recruitment of LTR retrotransposons as promoters by the anti-apoptotic locus NAIP during mammalian evolution,” *PLoS Genet.*, vol. 3, no. 1, pp. 0051–0062, 2007.

[88] A. Babaian and D. L. Mager, “Endogenous retroviral promoter exaptation in human cancer,” *Mob. DNA*, vol. 7, no. 1, pp. 1–21, 2016.

[89] T. Wang *et al.*, “Species-specific endogenous retroviruses shape the transcriptional network of the human tumor suppressor protein p53,” *Proc. Natl. Acad. Sci.*, vol. 104, no. 47, pp. 18613–18618, 2007.

[90] N. Bannert, H. Hofmann, A. Block, and O. Hohn, “HERVs new role in cancer: From accused perpetrators to cheerful protectors,” *Front. Microbiol.*, vol. 9, no. FEB, pp. 1–8, 2018.

[91] S. A. Tomlins *et al.*, “Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer,” *Nature*, vol. 448, 2007.

[92] J. Fahnenstich, A. Nandy, K. Milde-Langosch, T. Schneider-Merck, N. Walther, and B. Gellersen, “Promyelocytic leukaemia zinc finger protein (PLZF) is a glucocorticoid- and progesterone-induced transcription factor in human endometrial stromal cells and myometrial smooth muscle cells,” *Mol. Hum. Reprod.*, vol. 9, no. 10, pp. 611–623, 2003.

[93] M. Denne, M. Sauter, V. Armbruester, J. D. Licht, K. Roemer, and N. Mueller-Lantzsch, “Physical and Functional Interactions of Human Endogenous Retrovirus Proteins Np9 and Rec with the Promyelocytic Leukemia Zinc

Finger Protein," *J. Virol.*, vol. 81, no. 11, pp. 5607–5616, 2007.

[94] H. Gross *et al.*, "The NP9 protein encoded by the human endogenous retrovirus HERV-K(HML-2) negatively regulates gene activation of the Epstein-Barr virus nuclear antigen 2 (EBNA2)," *Int. J. Cancer*, vol. 129, no. 5, pp. 1105–1115, 2011.

[95] D. J. Huang G1, Li Z, Wan X, Wang Y, "Human endogenous retroviral K element encodes fusogenic activity in melanoma cells," *J. Carcinog.*, vol. 12, no. 5, 2013.

[96] J. Rak and A. Guha, "spread cancer genes," pp. 489–497, 2012.

[97] L. Balaj *et al.*, "Tumour microvesicles contain retrotransposon elements and amplified oncogene sequences," 2011.

[98] C. Kudo-Saito, M. Yura, R. Yamamoto, and Y. Kawakami, "Induction of immunoregulatory CD271+ cells by metastatic tumor cells that express human endogenous retrovirus H," *Cancer Res.*, 2014.

[99] B. Heemskerk, P. Kvistborg, and T. N. M. Schumacher, "The cancer antigenome," *EMBO J.*, vol. 32, no. 2, pp. 194–203, 2013.

[100] G. Kassiotis and J. P. Stoye, "Immune responses to endogenous retroelements: Taking the bad with the good," *Nat. Rev. Immunol.*, vol. 16, no. 4, pp. 207–219, 2016.

[101] F. Schiavetti, J. Thonnard, D. Colau, T. Boon, and P. G. Coulie, "A human endogenous retroviral sequence encoding an antigen recognized on melanoma by cytolytic T lymphocytes.," *Cancer Res.*, vol. 62, no. 19, pp. 5510–6, 2002.

[102] N. Geller *et al.*, "Regression of human kidney cancer following allogeneic stem cell transplantation is associated with recognition of an HERV-E antigen by T cells," *J. Clin. Invest.*, vol. 118, no. 3, 2008.

[103] P. Pérot *et al.*, "Expression of young HERV-H loci in the course of colorectal carcinoma and correlation with molecular subtypes.," *Oncotarget*, vol. 6, no. 37, pp. 40095–111, 2015.

[104] G. Kassiotis and J. P. Stoye, "Making a virtue of necessity: The pleiotropic role of human endogenous retroviruses in cancer," *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 372, no. 1732, 2017.

[105] N. Harashima, H. Yokoyama, and R. Srinivasan, “HHS Public Access,” vol. 76, no. 8, pp. 2177–2185, 2017.

[106] F. Wang-Johanning *et al.*, “Human endogenous retrovirus K triggers an antigen-specific immune response in breast cancer patients,” *Cancer Res.*, vol. 68, no. 14, pp. 5869–5877, 2008.

[107] F. Wang-Johanning *et al.*, “Immunotherapeutic potential of anti-human endogenous retrovirus-k envelope protein antibodies in targeting breast tumors,” *J. Natl. Cancer Inst.*, vol. 104, no. 3, pp. 189–210, 2012.

[108] D. F. Grabski, Y. Hu, M. Sharma, and S. K. Rasmussen, “Close to the Bedside: A Systematic Review of Endogenous Retroviruses and Their Impact in Oncology,” *J. Surg. Res.*, vol. 240, pp. 145–155, 2019.

[109] A. Kleiman *et al.*, “HERV-K(HML-2) GAG/ENV antibodies as indicator for therapy effect in patients with germ cell tumors,” *Int. J. Cancer*, vol. 110, no. 3, pp. 459–461, 2004.

[110] J. Zhao *et al.*, “Expression of Human Endogenous Retrovirus Type K Envelope Protein is a Novel Candidate Prognostic Marker for Human Breast Cancer,” *Genes and Cancer*, vol. 2, no. 9, pp. 914–922, 2011.

[111] S. Hahn *et al.*, “Serological Response to Human Endogenous Retrovirus K in Melanoma Patients Correlates with Survival Probability,” *AIDS Res Hum Retroviruses.*, vol. VOL. 24, no. NO. 5, 2008.

[112] S. Singh *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.*, vol. 26, no. 3, pp. 426–428, 2013.

[113] “www.wcrf.org.”

[114] “www.cancer.gov.”

[115] M. Arnold, M. S. Sierra, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, “Global patterns and trends in colorectal cancer incidence and mortality,” *Gut*, vol. 66, no. 4, pp. 683–691, 2017.

[116] “World Health Organization,” *www.who.int*.

[117] H. Khiari, A. H. W. Ben, K. H. Ben, and M. Hsairi, “Colorectal Cancer Incidence Trend and Projections in Tunisia (1994 - 2024),” *Asian Pac J Cancer Prev*, vol. 18, no. 10, pp. 2733–2739, 2017.

[118] L. R, “Variations du risque de cancer colorectal selon les ressources économiques.,” *Cancéro Dig.*, vol. 4, pp. 140–4, 2012.

[119] H. Belfki *et al.*, “Prevalence and determinants of the metabolic syndrome among Tunisian adults: Results of the Transition and Health Impact in North Africa (TAHINA) project,” *Public Health Nutr.*, vol. 16, no. 4, pp. 582–590, 2013.

[120] J. El Ati *et al.*, “Gender Obesity Inequities Are Huge but Differ Greatly According to Environment and Socio-Economics in a North African Setting: A National Cross-Sectional Study in Tunisia,” *PLoS One*, vol. 7, no. 10, 2012.

[121] H. Vogelstein, B., Fearon, E.R., A. C. S. R., Kern, S.E., Preisinger, and E. Leppert, M., “Genetic alterations during colorectal-tumor development,” *N. Engl.J.Med*, no. 319, pp. 525–532, 1988.

[122] A. S. Sameer, “Colorectal Cancer: Molecular Mutations and Polymorphisms,” *Front. Oncol.*, vol. 3, no. May, pp. 1–8, 2013.

[123] M. A. Kerachian and M. Kerachian, “Long interspersed nucleotide element-1 (LINE-1) methylation in colorectal cancer,” *Clin. Chim. Acta*, vol. 488, no. September 2018, pp. 209–214, 2019.

[124] C. Ji-Jun, W. Ai-Qin, C. Qing-Qi, C. Ji-Jun, W. Ai-Qin, and C. Qing-Qi, “DNA methylation assay for colorectal carcinoma,” *Cancer Biol. Med.*, vol. 14, no. 1, pp. 42–49, 2017.

[125] S. N. Hong, “Genetic and epigenetic alterations of colorectal cancer,” *Intest. Res.*, vol. 16, no. 3, p. 327, 2018.

[126] E. Sancho *et al.*, “The b-Catenin/TCF-4 Complex Imposes a Crypt Progenitor Phenotype on Colorectal Cancer Cells,” *Cell*, vol. 111, pp. 241–250, 2002.

[127] E. Sakai, A. Nakajima, and A. Kaneda, “Accumulation of aberrant DNA methylation during colorectal cancer development,” *World J. Gastroenterol.*, vol. 20, no. 4, pp. 978–987, 2014.

[128] M. Hajdúch, S. Jančík, J. Drábek, and D. Radzioch, “Clinical relevance of KRAS in human cancers,” *J. Biomed. Biotechnol.*, vol. 2010, 2010.

[129] “www.nih.gov.”

[130] F. D. R. Hobbs, “The role of primary care early diagnosis and referral guidelines population screening in primary care,” *October*, vol. 321, no. 7268,

pp. 1068–1070, 2000.

[131] P. M. Calvert and H. Frucht, “The Genetics of Colorectal Cancer,” *Ann. Intern. Med.*, vol. 137, no. 7, pp. 603–12, 2002.

[132] T. Armaghany, J. D. Wilson, Q. Chu, and G. Mills, “Genetic alterations in colorectal cancer,” *Gastrointest. Cancer Res.*, vol. 5, no. 1, pp. 19–27, 2012.

[133] S. B. Jones, Peter A. Baylin, “The Epigenomics of Cancer,” vol. 128, no. 4, pp. 683–692, 2007.

[134] H. Ashktorab *et al.*, “Clinicopathological features and microsatellite instability (MSI) in colorectal cancers from African Americans,” vol. 116, no. 6, pp. 914–919, 2014.

[135] M. Barchitta, A. Quattrocchi, A. Maugeri, M. Vinciguerra, and A. Agodi, “LINE-1 hypomethylation in blood and tissue samples as an epigenetic marker for cancer risk: A systematic review and meta-analysis,” *PLoS One*, vol. 9, no. 10, 2014.

[136] S. Ogino *et al.*, “A cohort study of tumoral LINE-1 hypomethylation and prognosis in colon cancer,” *J. Natl. Cancer Inst.*, vol. 100, no. 23, pp. 1734–1738, 2008.

[137] E. Sunami, M. de Maat, A. Vu, R. R. Turner, and D. S. B. Hoon, “LINE-1 hypomethylation during primary colon cancer progression,” *PLoS One*, vol. 6, no. 4, pp. 2–8, 2011.

[138] M. Świderska *et al.*, “The diagnostics of colorectal cancer,” *Wspolczesna Onkol.*, vol. 18, no. 1, pp. 1–6, 2014.

[139] S. P. Arora and D. Mahalingam, “Immunotherapy in colorectal cancer: For the select few or all?,” *J. Gastrointest. Oncol.*, vol. 9, no. 1, pp. 170–179, 2018.

[140] K. Schmitt, J. Reichrath, A. Roesch, E. Meese, and J. Mayer, “Transcriptional profiling of human endogenous retrovirus group HERV-K(HML-2) loci in melanoma,” *Genome Biol. Evol.*, vol. 5, no. 2, pp. 307–328, 2013.

[141] J. P. Pichon, B. Bonnaud, P. Cleuziat, and F. Mallet, “Multiplex degenerate PCR coupled with an oligo sorbent array for human endogenous retrovirus expression profiling,” *Nucleic Acids Res.*, vol. 34, no. 6, 2006.

[142] Y. Stauffer, G. Theiler, P. Sperisen, Y. Lebedev, and C. V. Jongeneel, “Digital

expression profiles of human endogenous retroviral families in normal and cancerous tissues.," *Cancer Immun.*, vol. 4, p. 2, Feb. 2004.

[143] J. A. M. G. D. L. R. C. C.Huang, "mRNA from human colon tumor and mucosa related to the pol gene of an endogenous A-type retrovirus," *Biochem. Biophys. Res. Commun.*, vol. 139, no. 3, pp. 1071–1077, 1986.

[144] Q.-Y. Liang, Z.-F. Xu, R.-Z. Xu, S. Zheng, and J.-Y. Ding, "[Deletion of the env region in HERV-H-X gene and its expression in colon cancer].," *Ai Zheng*, vol. 26, no. 9, pp. 952–956, Sep. 2007.

[145] P. Jern, G. O. Sperber, G. Ahlsen, and J. Blomberg, "Sequence Variability, Gene Structure, and Expression of Full-Length Human Endogenous Retrovirus H," *J. Virol.*, vol. 79, no. 10, pp. 6325–6337, 2005.

[146] P. M. S. Alves *et al.*, "Identification of tumor-associated antigens by large-scale analysis of genes expressed in human colorectal cancer," *Cancer Immun.*, vol. 8, no. May, pp. 1–11, 2008.

[147] M. Mangeney, N. De Parseval, G. Thomas, and T. Heidmann, "The full-length envelope of an HERV-H human endogenous retrovirus has immunosuppressive properties," *J. Gen. Virol.*, no. 2001, pp. 2515–2518, 2017.

[148] Q. Liang, J. Ding, and S. Zheng, "Identification and detection of a novel human endogenous retrovirus-related gene , and structural characterization of its related elements," *Genet. Mol. Biol.*, vol. 708, pp. 704–708, 2009.

[149] Q. Liang, Z. Xu, R. Xu, L. Wu, and S. Zheng, "Expression patterns of non-coding spliced transcripts from human endogenous retrovirus HERV-H elements in colon cancer," *PLoS One*, vol. 7, no. 1, 2012.

[150] S.-H. Lee *et al.*, "Elevation of human ERV3-1 env protein expression in colorectal cancer.," *Journal of clinical pathology*, vol. 67, no. 9. England, pp. 840–844, Sep-2014.

[151] H. S. Kim, K. Ahn, and D. S. Kim, "Quantitative expression of the HERV-W env gene in human tissues," *Arch. Virol.*, vol. 153, no. 8, pp. 1587–1591, 2008.

[152] K. Ahn and H.-S. Kim, "Structural and quantitative expression analyses of HERV gene family in human tissues," *Mol. Cells*, vol. 28, no. 2, pp. 99–103,

2009.

[153] C. Théry, M. Ostrowski, and E. Segura, “Membrane vesicles as conveyors of immune responses.,” *Nat. Rev. Immunol.*, vol. 9, no. 8, pp. 581–93, 2009.

[154] B. Février and G. Raposo, “Exosomes: Endosomal-derived vesicles shipping extracellular messages,” *Curr. Opin. Cell Biol.*, vol. 16, no. 4, pp. 415–421, 2004.

[155] H. Kalra *et al.*, “Vesiclepedia: A Compendium for Extracellular Vesicles with Continuous Community Annotation,” *PLoS Biol.*, vol. 10, no. 12, pp. 8–12, 2012.

[156] C. Théry, A. I. Curie, and U. Inserm, “Exosomes : secreted vesicles and intercellular communications,” vol. 8, no. July, pp. 1–8, 2011.

[157] G. Camussi, M. C. Deregibus, S. Bruno, V. Cantaluppi, and L. Biancone, “Exosomes/microvesicles as a mechanism of cell-to-cell communication,” *Kidney International*. 2010.

[158] M. P. Bebelman, M. J. Smit, D. M. Pegtel, and S. R. Baglio, “Biogenesis and function of extracellular vesicles in cancer,” *Pharmacol. Ther.*, vol. 188, pp. 1–11, 2018.

[159] K. Al-Nedawi *et al.*, “Intercellular transfer of the oncogenic receptor EGFRvIII by microvesicles derived from tumour cells,” *Nat. Cell Biol.*, vol. 10, no. 5, pp. 619–624, 2008.

[160] H. W. King, M. Z. Michael, and J. M. Gleadle, “Hypoxic enhancement of exosome release by breast cancer cells,” *BMC Cancer*, vol. 12, no. 1, p. 1, 2012.

[161] R. Sullivan *et al.*, “The emerging roles of extracellular vesicles as communication vehicles within the tumor microenvironment and beyond,” *Front. Endocrinol. (Lausanne).*, vol. 8, no. AUG, pp. 1–11, 2017.

[162] P. D. Robbins and A. E. Morelli, “Regulation of Immune Responses by Extracellular Vesicles,” *Nat Rev Immunol.*, vol. 14, no. 3, pp. 195–2018, 2014.

[163] E. U. Wieckowski, C. Visus, M. Szajnik, M. J. Szczepanski, W. J. Storkus, and T. L. Whiteside, “Tumor-Derived Microvesicles Promote Regulatory T Cell Expansion and Induce Apoptosis in Tumor-Reactive Activated CD8+ T Lymphocytes,” *J. Immunol.*, vol. 183, no. 6, pp. 3720–3730, 2009.

[164] T. Lener *et al.*, “Applying extracellular vesicles based therapeutics in clinical trials - an ISEV position paper.,” *J. Extracell. vesicles*, vol. 4, p. 30087, 2015.

[165] R. O. Wright, J. Schwartz, R. J. Wright, V. Bollati, L. Tarantini, and S. K. Park, “Biomarkers of Lead Exposure and DNA Methylation within Retrotransposons,” no. 6, pp. 790–795, 2010.

[166] L. Pergoli *et al.*, “Blood DNA methylation , nevi number , and the risk of melanoma,” *Melanoma Res.*, vol. 24, no. 5, pp. 480–487, 2014.

[167] M. Manghera, J. Ferguson, and R. Douville, “ERVK polyprotein processing and reverse transcriptase expression in human cell line models of neurological disease,” *Viruses*, vol. 7, no. 1, pp. 320–332, 2015.

[168] M. Manghera, J. Ferguson, and R. Douville, “ERVK polyprotein processing and reverse transcriptase expression in human cell line models of neurological disease,” *Viruses*, 2015.

[169] M. S. Rooney, S. A. Shukla, C. J. Wu, G. Getz, and N. Hacohen, “Molecular and genetic properties of tumors associated with local immune cytolytic activity,” *Cell*, vol. 160, no. 1–2, pp. 48–61, 2015.

[170] Y. Stauffer, G. Theiler, P. Sperisen, Y. Lebedev, and C. V. Jongeneel, “Digital expression profiles of human endogenous retroviral families in normal and cancerous tissues,” *Cancer Immun.*, vol. 4, no. 2, 2004.

[171] K. Ahn and H.-S. Kim, “Structural and quantitative expression analyses of HERV gene family in human tissues,” *Mol. Cells*, 2009.

[172] A. Serafino *et al.*, “The activation of human endogenous retrovirus K (HERV-K) is implicated in melanoma cell malignant transformation,” *Exp. Cell Res.*, vol. 315, no. 5, pp. 849–862, 2009.

[173] C. Gallois, S. Pernot, A. Zaanan, and J. Taieb, “Colorectal Cancer: Why Does Side Matter?,” *Drugs*, vol. 78, no. 8, pp. 789–798, 2018.

[174] B. Baran, N. Mert Ozupek, N. Yerli Tetik, E. Acar, O. Bekcioglu, and Y. Baskin, “Difference Between Left-Sided and Right-Sided Colorectal Cancer: A Focused Review of Literature,” *Gastroenterol. Res.*, vol. 11, no. 4, pp. 264–273, 2018.

[175] Woalder, “乳鼠心肌提取 HHS Public Access,” *Physiol. Behav.*, vol. 176, no. 1, pp. 139–148, 2017.

[176] E. Fearon and B. Vogelstein, “A genetic model for colorectal tumorigenesis,” *Cell*, vol. 61, pp. 759–67, 1990.

[177] D. L. Worthley and B. A. Leggett, “Colorectal cancer: molecular features and clinical opportunities.,” *Clin. Biochem. Rev.*, vol. 31, no. 2, pp. 31–8, 2010.

[178] J. H. Yi, J. Liu, and K. H. Wang, “CpG island methylator phenotype in colorectal cancer,” *World Chinese J. Dig.*, vol. 24, no. 4, pp. 558–565, 2016.

[179] D. J. Weisenberger *et al.*, “CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer,” *Nat. Genet.*, vol. 38, no. 7, pp. 787–793, 2006.

[180] J. Pothlichet, M. Mangeney, and T. Heidmann, “Mobility and integration sites of a murine C57BL/6 melanoma endogenous retrovirus involved in tumor progression in vivo,” *Int. J. Cancer*, vol. 119, no. 8, pp. 1869–1877, 2006.

[181] R. Nayani, H. Ashktorab, H. Brim, and A. O. Laiyemo, “Genetic Basis for Colorectal Cancer Disparities,” *Physiol. Behav.*, vol. 176, no. 3, pp. 139–148, 2017.

[182] Q. Liang, Z. Xu, R. Xu, L. Wu, and S. Zheng, “Expression patterns of non-coding spliced transcripts from human endogenous retrovirus HERV-H elements in colon cancer,” *PLoS One*, vol. 7, no. 1, 2012.

[183] A. R. Florl, R. Löwer, B. J. Schmitz-Dräger, and W. A. Schulz, “DNA methylation and expression of LINE-1 and HERV-K provirus sequences in urothelial and renal cell carcinomas,” *Br. J. Cancer*, vol. 80, no. 9, pp. 1312–1321, 1999.

[184] L. Hu, D. Uzhameckis, F. Hedborg, and J. Blomberg, “Dynamic and selective HERV RNA expression in neuroblastoma cells subjected to variation in oxygen tension and demethylation,” *Apmis*, vol. 124, no. 1–2, pp. 140–149, 2016.

[185] L. Hu, D. Uzhameckis, F. Hedborg, and J. Blomberg, “Dynamic and selective HERV RNA expression in neuroblastoma cells subjected to variation in oxygen tension and demethylation,” *APMIS*, 2016.

[186] D. Greenbaum, C. Colangelo, K. Williams, and M. Gerstein, “Comparing protein abundance and mRNA expression levels on a genomic scale,” *Genome Biol.*, vol. 4, no. 9, 2003.

[187] N. Grandi and E. Tramontano, “HERV envelope proteins: Physiological role and pathogenic potential in cancer and autoimmunity,” *Frontiers in Microbiology*. 2018.

[188] R. O. Wright *et al.*, “Biomarkers of lead exposure and DNA methylation within retrotransposons,” *Environ. Health Perspect.*, 2010.

[189] L. Tarantini *et al.*, “Effects of particulate matter on genomic DNA methylation content and iNOS promoter methylation,” *Environ. Health Perspect.*, vol. 117, no. 2, pp. 217–222, 2009.

[190] A. Manuscript, W. Blood, and C. Count, “NIH Public Access,” vol. 49, no. 18, pp. 1841–1850, 2009.

APPENDIX

APPENDIX A- FICOLL

The isolation of the plasma and of PBMCs was done using Ficoll-Paque™ Plus method (GE Healthcare Bio-Sciences AB, Sweden). The blood was diluted 1:1 with PBS (EuroClone, Milano, Italy): 5ml of the solution was added to 3ml of Ficoll and a centrifugation at 1300rpm for 30 minutes were performed. Differential migration during centrifugation results in the separation of cell types into different layers (Figure 10). The upper layer was composed of plasma and platelets, then the PBMCs, the Ficoll and in the bottom layer the erythrocytes.

1ml of plasma and the PBMCs were removed: the PBMCs were subjected to a wash, adding PBS to a final volume of 13 ml. After a centrifugation of 10 minutes at 1200 rpm, the obtained pellet was incubated for 10 minutes in ice with 10 ml of lysis buffer, to eliminate the erythrocytes. The solution was subjected to a centrifugation of 10 minutes at 1200 rpm at 4°C. The isolated cells were stained with 0.4% Trypan blue (EuroClone, Milano, Italy) and counted in a Burker chamber. The aliquots of PBMCs, plasma and blood were stored at -80°C for the future analysis.

APPENDIX B- Methylation analysis

B1- DNA isolation from tissue and blood

DNA was isolated from 20mg of tumor and the adjacent normal tissues and from 200µl of blood, using DNA Mini and DNA Blood Mini kit (Qiagen, Hilden, Germany), respectively. For DNA isolation from tissue, 20mg of tissue were crumbled with the scalpel and 180µl of ATL buffer and 20µl of proteinase K (Qiagen, Hilden, Germany) were added to the sample; the solution was

incubated overnight at 56°C for the lysis of the tissue. Then, 200µl of lysis buffer (AL buffer) were added and the sample was mixed.
Instead for the blood, 200µl of AL buffer and 20µl of proteinase K (Qiagen, Hilden, Germany) were added to 200µl of blood to the lysis of the cells. The sample was incubated for 10 minutes at 56°C in the thermo block.
For both the tissue and the blood, 200µl of 100% EtOH was added and the solution was added to a spin column with a membrane with positive charge. The solution was centrifuged for 1 minute at 8000rpm. Two different washes, with 500µl of AW1 and AW2 buffers, were done and finally the DNA was eluted in 50µl of RNase free water, after an incubation of 5 minutes. The DNA concentration was evaluated using Nanophotometer NP80 (Implen, Munich, Germany) and the DNA was stored at -20°C for the next analysis.

B2- Bisulfite treatment

500 ng of DNA, isolated from the tissue and from the blood of each patients, were transferred to a PCR tubes and treated with bisulfite. The treatment was performed using the EZ DNA Methylation-Gold Kit (Zymo Research, California, USA). 130 µl of CT conversion reagent, containing water, M-Diluition Buffer and M-Dissolving Buffer, was added to 20 µl of DNA (500 ng) and then incubated in the thermal cycler, following these steps: 10 minutes at 98°C, 2.5 hours at 64°C and up to 20 hours at 4°C.
600 µl of M-Binding Buffer to a Zymo-Spin™ IC Column was added and placed the column into a collection tube, and the DNA treated with CT conversion buffer was loaded into Zymo-Spin™ IC Column containing the M-Binding Buffer and mixed: the buffer allows the binding of the DNA to a membrane. The plate was centrifuged at full speed for 30 seconds and then the first wash was done with 100µl of M-Wash Buffer, and the plate was subjected to a centrifuge at the full speed for 30 seconds. 200µl of M-Desulphonation Buffer was added to each tube and incubated at room

temperature for about 20 minutes and then centrifuged for 30 seconds: Desulphonation Buffer, which with an alkaline treatment removes the sulfonate group to allow the subsequent addition of uracil. Finally, other wash was performed with 200µl of M-Wash Buffer, and the DNA was eluted using 200µl of M-Elution Buffer. Bisulfite-treated DNA was stored at -80°C.

B3- Bisulfite PCR

A bisulfite-PCR was conducted using specific primers for the HERV-H, -K, -P, -W LTRs, Alu and LINE-1 and *psen*: a 50µl PCR was carried out in 25µl GoTaq Green Master Mix (Promega, Wisconsis, USA), (containing Hotstar Taq DNA Polymerase, PCR Buffer 10X, Q-Solution 5X, and $MgCl_2$ 25 mM), 1 pmol forward primers and 1 pmol reverse primers, 50ng bisulfite-treated genomic DNA and water. One of the two primers was biotinylated to allow the following separation of the DNA into a single strand. The sequences of the primers for HERV-H, -K, and-P LTRs were draw using specific software (Table 6) and they were synthetized by Eurofins Genomics (Milano, Italy). The sequences of the primers are reported in the Table 6. The HERV-W, Alu e LINE-1 primers sequences were previous published [188] [189]. The composition of the reaction mix was the following (Table B1):

Table B1: *Composition of bisulfite PCR mix.*

	Initial concentration	**Final concentration**	**Volume**
Go Taq Green Master Mix (2x)	2x	1x	25µl
Forward Primer	10 pmol/µl	0.2 pmol/µl	1µl
Reverse Primer	10 pmol/µl	0.2 pmol/µl	1µl
Bisulfite trated DNA			up to 10µl (50ng)
H_2O			up to 13µl
Final volume			50µl

In each well, 40µl of mix were added with 10µl of bisulfite-treated DNA. The thermal profile of the bisulfite PCR reactions is shown in Table B2 and Table B3. Also, in this case, the thermal profile of the reactions for HERV-W LTRs, and Alu and LINE-1 have been previously published [190].

Table B2: *Thermal profile of HERV-H, -K. -P, and -W LTRs PCR.*

HERV H	HERV K	HERV P	HERV W	
95°C 5’				
95°C 30” 50°C 30” 72°C 30”	95°C 30” 52°C 30” 72°C 30”	95°C 30” 52°C 30” 72°C 30”	95°C 60” 55°C 60” 72°C 60”	45 cycles
72°C 5’				
4°C				

Table B3: *Thermal profile of Alu and LINE-1 PCR.*

Alu	LINE-1	
95°C 5’		
96°C 90” 43°C 60” 72°C 120”	95°C 30” 50°C 30” 72°C 30”	45 cycles
72°C 5’		
4°C		

B4- Isolation of single strands of DNA using streptavidin

In the amplified sample, one of the two strands were biotinylated at 5' and the binding with streptavidin allowed to isolate the single strand of interest. The binding mix was composed of 1µl streptavidin, that mediates the binding with the biotin on the single strand, 40µl of Binding Buffer and 24µl of H_2O, for a final volume of 65µl. Then 15µl of the PCR product was added. The PCR plate was shaken at 1140 rpm for at least 10 minutes, to allow the bound

of the beads to the biotin (BenchMark ORB-Shaker JRTM) (Sayreville, New Jersey, United States).

Then, to isolate the biotinylated DNA single strand, the Pyromark Q96 Workstation (Qiagen, Hilden, Germany) was set up, as shown in the Figure B1.

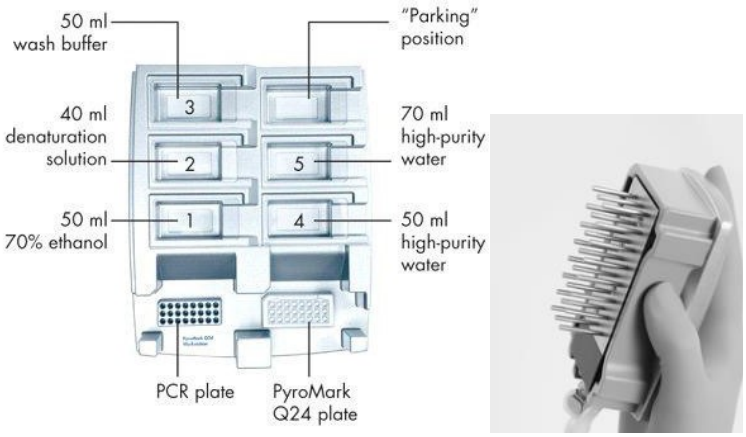


Figure B1: *The Pyromark Q96 Workstation.*

The Workstation was composed of:

a. The vacuum prep tool (VPT), (Biotage, Uppsala Sweden) connected to a vacuum pump. VPT isolating the DNA strand by filters that bind the beads bound to the strand with biotin;
b. The denaturing solution, consisting of 4 g soda diluted in 500ml of MilliQ water, to denature the two DNA strands;
c. 70% ethanol (150 ml of milliQ water and 350 ml 100% EtOH);
d. Water;
e. The washing solution (900 ml of water milliQ and washing buffer 10X (Qiagen, Hilden, Germany);
f. Pyromark Q96 HS Plate (PSQ plate): 14.2µl of annealing Buffer is added to each well with 0.8µl of the sequencing primer.

The sequences of the pyrosequencing primers are reported in the Table 8.

The vacuum pump is turned on and the VPT is placed first in the PCR plate to isolate and to bind the biotinylated strand; the VPT was moved to the support containing 70% EtOH, then to the denaturing solution, to the washing solution and finally to the water. When the vacuum pump was removed, the VPT was placed in the PSQ plate to allow the release of the isolated biotin-streptavidin strand. The plate containing the annealing mix and the isolated strand was moved into the Dry-Bath-Heater (Biotage, Uppsala, Sweden), and incubated at 85°C for 2 minutes.

APPENDIX C- RNA isolation and Real Time PCR

C1- RNA isolation from blood

The RNA was isolated from the biological sample using QIAmp RNA Blood Mini Kit (Qiagen, Hilden, Germany). The RNA isolation is allowed by the using of column with a silica membrane, with a positive charge and specific affinity of RNA that bind the RNA, with a negative charge. Blood cells were lysed in two separate procedures, erythrocyte lysis and peripheral blood mononuclear cells (PBMCs) lysis. First the erythrocytes were lysis: one volume of blood was mixed with five volume of buffer EL and incubated for 15 minutes in ice. After a centrifugation, the PBMCs were pelleted on the bottom of the tube. The lysis of the cellular membrane of PBMCs was performed using 350 µl of lysis buffer (RLT) and then add to QIAshredder spin column. After centrifugation a volume of ethanol at 70% was added to the homogenized lysate to induce the precipitation of RNA. The solution was added into the QIAamp spin column and subjected to a centrifugation. 10µl of DNase I stock solution (Qiagen, Hilden, Germany) was mix with 70µl of RDD buffer (Qiagen, Hilden, Germany) and DNase digestion was performed to remove the DNA bind to the column: the column was stand for 15 minutes at room temperature. Then the membrane was subjected to three washes, one with buffers RW1 and two with RPE buffer, to eliminate the

contaminants, such as protein. RNA was finally eluted in RNase free water and it was quantified using Nanophotometer NP80 (Implen, Munich, Germany). The RNA was stored at -80°C for the next analysis.

C2- RNA isolation from EVs

The RNA was also isolated from the EVs using QIAzol lysis reagent (Qiagen, Hilden, Germany) to the lysis of the cells and RNeasy MinElute Cleanup kit to concentrate the RNA (Qiagen, Hilden, Germany). The pellet of EVs, obtained from the ultracentrifugation, was resuspended with 700µl of QIAzol lysis reagent and then 140µl of clorophormio was added. The mixed solution was centrifugated for 15 minutes at 12,000 g and the upper phase, the aqueous phase, that contained the RNA, was transferred to a new tube. After the addition of a volume of 70% EtOH, the solution was added to RNeasy MinElute spin column (Qiagen, Hilden, Germany), stored at -4°C. The sample was subjected to a centrifugation for 15 seconds at 11000rpm was performed (for two times). 500µl of RPE buffer and 500µl of EtOH were added to wash the membrane, and finally a centrifuge at full speed for 5 minutes was done. The RNA was eluted in 22µl of RNeasy free water, after an incubation at room temperature for 5 minutes. The RNA isolated from the EVs was quantified using Nanophotometer NP80 (Implen, Germany). The RNA was stored at -80°C for the next analysis.

C3- Real Time PCR

To evaluate the HERV expression in the biological samples, a quantitative Sybr Green Real Time PCR was performed, especially were evaluate the expression of the HERV-H, -K, -R and -P *env*, HERV-K *pol* gene, and Alu and LINE-1 using the 7500 Real Time PCR system (Applied Biosystem, California, USA).

The Sybr Green is a fluorescent dye that binds the double stranded DNA, intercalating between the bases. The quantification is relative: all the *env/pol*

genes and Alu and LINE-1 expressions were related to the expression of a housekeeping gene, Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) to normalize the different expressions. The reaction was conducted in a final volume of 25 µl, containing 0.3 µM forward primer, 0.3 µM reverse primer, 12.5µl of QuantiTec Sybr Green Master Mix 2x (composed of reaction buffer, Taq Polymerase, dNTPs, Magnesium, Rox and Sybr Green dyes) and 2 µl of cDNA. The composition of the mix is reported in the Table 10 and the sequences of the primers were previous published [54] (Table 11).

The reaction steps were: 2 minutes at 50°C, the holding stage for 15 minutes at 95°C and 45 cycle, each of 95°C for 15 seconds (denaturation of the double stranded DNA), 54°C for 30 seconds for HERV and GAPDH while 60°C for 30 seconds for Alu and LINE-1 (annealing of the primers), and the final step at 72°C for 30 seconds (elongation of the new filaments and detection of the fluorescence).

When the Real Time PCR reaction was performed, the amplification plot was analysed: threshold was set at 0.02 to determine the Ct of HERV and GAPDH of each patient and in each sample.

The quantification of the HERV-H, -K, -P and -R *env and* HERV-K *pol* gene, and Alu and LINE-1 expression was performed using the comparative Ct method, and the differences between the levels of *env/pol* gene and Alu and LINE-1 expression in the biological samples were calculated by relative quantification (RQ).

$$\Delta\text{Ct HERV} = Ct\ HERV\ (or\ Alu\ and\ LINE-1) - Ct\ GAPDH$$

$$\Delta\text{Ct mean HERV} = \frac{\sum Ct\ HERV\ (or\ Alu\ and\ LINE-1)}{number\ of\ samples}$$

$$\text{RQ expression} = 2^{-(\Delta\text{Ct HERV(or Alu and LINE}-1) - \Delta\text{Ct mean HERV (or Alu and LINE}-1))}$$

PAPERS

L. Ferrari, M. Cafora, F. Rota, M. Hoxha, S. Iodice, L. Tarantini, **M. Dolci**, S. Delbue, A. Pistocchi, V. Bollati. *Extracellular vesicles released by colorectal cancer cell lines modulate innate immune response in zebrafish: the possible role of human endogenous retroviruses.* International Journal of Molecular Sciences 2019, 20 (15), 3669.

M. Dolci, C. Favero, V. Bollati, L. Campo, A. Cattaneo, M. Bonzini, S. Villani, R. Ticozzi, P. Ferrante, S. Delbue. *Particulate matter exposure increases JC polyomavirus replication in the human host.* Environmental Pollution 241 (2018) 234e239.

L. Signorini, S. Villani, R.M. Ticozzi, F. Ambrogi, **M. Dolci**, R. Boldorini, M. Ciotti, P. Ferrante, S. Delbue *Merkel cell polyomavirus DNA in the blood of patients with neurological diseases and healthy controls* Future Virology. 12:12(2017 Dec), pp. 713-720.

Villani S, Zanotta N, Ambrogi F, Comar M, Franciotta D, **Dolci M**, Cason C, Ticozzi R, Ferrante P, Delbue S. *Multiplex array analysis of circulating cytokines and chemokines in natalizumab-treated patients with multiple sclerosis*. J Neuroimmunol. 2017 Sep 15; 310:91-96.

R. Bella, **M. Dolci**, M. Ferraresso, R. Ticozzi, L. Ghio, J. Rizzo, L.Signorini, S. Villani, F. Elia, P. Ferrante, S. Delbue *Human herpesvirus-6 and polyomavirus DNAemia in children and young adult patients after kidney transplantation*, Future Virology 2015*, 10(12),* 1275-1284.

SCIENTIFIC PRODUCTS OF THE PHD PERIOD

M. Dolci, C. Favero, L. Tarantini, S. Villani, M. Bregni, L. Signorini, A. Della Valle, F. Crivelli, T.Casati, P. Ferrante, V. Bollati, S. Delbue *Human endogenous retroviruses (HERV) methylation and expression in colorectal cancer patients. 47° congresso nazionale Società Italiana Microbiologia (SIM), 18-21 September 2019, Roma, Italy.* Poster Presentation.

M. Dolci, C. Favero, L. Tarantini, S. Villani, M. Bregni, L. Signorini, A. Della Valle, F. Crivelli, P. Ferrante, V. Bollati, S. Delbue *Human endogenous retroviruses (HERV) methylation and expression in colorectal cancer patients.* 22nd European Society for Clinical Virology (ESCV), 11-14 September 2019, Copenhagen, Denmark. Oral presentation and travel grant awarded.

M. Dolci, C. Favero, L. Tarantini, S. Villani, M. Bregni, L. Signorini, A. Della Valle, F. Crivelli, P. Ferrante, V. Bollati, S. Delbue *Human endogenous retroviruses Long Terminal Repeat methylation and env gene expression in colorectal cancer patients.* 29th European Congress of Clinical Microbiology & Infectious Diseases (ECCMID), 13-16 April 2019, Amsterdam, The Netherlands. Poster Presentation.

M. Dolci, C. Favero, L. Tarantini, S. Villani, M. Bregni, L. Signorini, A. Della Valle, F. Crivelli, P. Ferrante, V. Bollati, S. Delbue *Human Endogenous Retroviruses are overexpressed and demethylated in right colon cancer tissue, and packaged in plasma extracellular vesicles of colon cancer patients.* 21° European Society for Clinical Virology (ESCV), 22-26 September 2018, Athens, Greece. Oral presentation and travel grant awarded.

M. Dolci, C. Favero, V. Bollati, L. Campo, A. Cattaneo, M. Bonzini, S. Villani, R. Ticozzi, P. Ferrante, S. Delbue. *Particulate Matter 10 (PM10) exposure induces JC polyomavirus replication.* 20th Annual Meeting of European Society of Clinical Virology (ESCV), 13-16 September 2017: Stresa (Italia). Poster presentation and travel grant awarded.

ACKNOWLEDGMENT

The study was supported by the Italian Ministry of University – PRIN 2015 to Pasquale Ferrante; contract grant number 2015w729wh and by the University of Milan – Piano Sostegno alla Ricerca (PSR) 2017 and 2018, to Serena Delbue.

I would like to thank Professor Pasquale Ferrante, Professor Serena Delbue, and my colleagues for their scientific support. I would like to thank the Epidemiology, Epigenetics and Toxicology lab (EPIGET) of the University of Milano, directed by Professor Valentina Bollati, for the methylation and statistical analysis. I would like to thank the hospitals with the clinicians that enrolled the patients (Istituto Clinico Città Studi, Milano, Italy. Ospedale di Circolo di Bursto Arstizio, Varese, Italy. Habib Yhameur Hospital, Tunis, Tunisia).

This thesis was evaluated by one independent reviewer: Dr Marco Bregni and by one anonymous reviewer.