Identification and investigation of new viruses in tumor-derived skin lesions

Ph.D. Thesis

Katalin Barbara Horváth-Szőgyi M.D.

Supervisor: Gábor Reuter, M.D., Ph.D., Med. Habil., D.Sc. [1,2]

Head of the Doctoral School: Dóra Reglődi**, M.D., Ph.D., Med. Habil., D.Sc.**
Head of the Program: Monika Kerényi**, M.D., Ph.D.**

[1]Government Office for Baranya County, Regional Laboratory of Virology,
National Reference Laboratory of Gastroenteric Viruses, Pécs,
[2]Department of Medical Microbiology and Immunology, Medical School,
University of Pécs

Pécs, 2022

I. INTRODUCTION

Various endogenous and exogenous factors may be responsible for the development of malignancies. In addition to carcinogenic chemicals and physical radiation, tumorigenic microorganisms, especially oncogenic viruses, play a significant role among exogenous factors. Extensive human studies confirm that the pathogenic role of an infectious agent in the background of every sixth tumor worldwide can be demonstrated. Although this frequency is only an average and probably underestimated, we can see significant geographical differences. Developing countries in particular are more severely affected by infectious cancers, where they may play a role in every second cancer and affect the younger population.

Oncogenic tumor viruses are subclassified as either DNA viruses, which include Epstein–Barr virus (EBV, family *Herpesviridae*), Kaposi's sarcoma–associated herpesvirus (KSHV, family *Herpesviridae*), human papillomavirus (HPV, family *Papillomaviridae*), hepatitis B virus (HBV, family *Hepadnaviridae*), and Merkel cell polyomavirus (MCV, family *Polyomaviridae*), or RNA viruses, such as hepatitis C virus (HCV, family *Flaviviridae*) and human T-cell lymphotropic virus (HTLV-1, family *Retroviridae*).

Merkel cell carcinoma and polyomavirus

Merkel cell carcinoma (MCC) is a rare and aggressive neoplasia that was first described in 1972 by Toker. Its origin is uncertain, but it is thought that this (neuroendocrine?/epithelial?/pre/pro B cell?) carcinoma arises from the stratum basale. MCC affects mostly the elderly and immunosuppressed people. MCC presents as a rapidly growing, solitary skin tumor that is located mostly on sun-exposed areas, especially at the head-neck region. Lesions are asymptomatic, red-to-violet or skin-colored nodules that might be clinically misconstrued as benign, semimalignant or malignant lesions (such as basal cell carcinoma, cutaneous squamous cell carcinoma, pyogenic granuloma, keratoacanthoma, amelanotic melanoma, benign cyst, cutaneous lymphoma and other metastatic neoplasia).

A DNA tumor virus, a Merkel cell polyomavirus (MC), has been identified in MCC in 2008 which open up new perspectives for the cause of the disease. Feng studied ten tumors by digital transcriptome subtraction and detected the virus in eight cases. In six of

the eight positive tumors, the virus was integrated into the tumor genome. The infectious – viral – origin of the disease results in attitude change not only in the aetiology, but also in the fields of pathogenesis, diagnostics, and then in future prevention and therapy. The importance of the prognosis is further emphasized by the fact that the lethality of MCC is more than three times higher (46%) than those of the malignant melanoma (12%) related to all diagnoses. Since its discovery, MC polyomavirus is detected between 71% and 94% of the MCC - despite the differences in the methodologies used - worldwide. The incidence of MCC is increasing due to the advancing age of the population, a higher incidence of damaging sun exposure, an increase in the number of immunocompromised individuals and partly with improved diagnostics. As it can metastasize as a small tumor, it is of particular importance for its rapid detection and isolation from other skin tumors (e.g. pyogenic granuloma, basal cell carcinoma), as well as for professional treatment.

The diagnosis of MCC is very rare based on clinical examination. Usually, the suspicion of MCC is confirmed by microscopic examination and a well-founded diagnosis can be made by immunohistochemistry. Similar to other neuroendocrine tumors, small round to ovoid cells with hyperchromatic nuclei and frequent mitosis or apoptosis are commonly observed. Three main types of MCC have been described - small-cell, trabecular and intermediate - but most cases present with overlapping features. Sometimes, it is difficult to differentiate MCC from other small cell tumors, and in these cases, immunohistochemical staining must be used. MCC stains positive for neuron-specific enolase, synaptophysin, chromogranin A, cytokeratin 20 (in a characteristic dot-like pattern) and is negative for S100, TTF-1 and CK-7.

Although the presence of MC polyomavirus and the integration of its genome into the chromosome of MCC cells appear to confirm its role in tumorigenesis, its exact pathomechanism raises even more questions. It is likely to be a complex, multi-step, host-dependent process, as the virus can be detected in skin areas that appear healthy in patients with MCC or other skin cancers, but also in the oral and respiratory epithelium. Furthermore, MCV, like other polyomaviruses, is found in the healthy population. However, MC polyomavirus is not detectable in one-fifth of MCCs, which may be a distinct clinicopathological entity that is similar in histological picture but separable in its course from the virus-carrying variant. An important initial step in the development of MCC may be the clonal integration of viral DNA, which can occur randomly in the host cell genome.

However, the most important role in tumorigenesis is the extent and mode of expression of the viral LT protein in the tumor. This is because this oncoprotein carries

conserved regions (e.g., DnaJ, pocket protein-binding LXCXE, and pp2A-binding domains) that have been shown to have oncogenic activity in other polyomaviruses (e.g., SV40). UV radiation can promote viral genome integration and mutations that truncate the LT region, which together eliminate viral replication. In tumor cells, the LT protein helicase activity characteristic of viral replication disappears (the “virus” is no longer infectious), while the damaged LT protein directs the processes toward tumor formation. Among other things, the LT protein inhibits the function of the cellular tumor suppressor proteins, in addition to which the cellular survivin (= baculoviral inhibitor of apoptosis repeat-containing 5: BIRC5a) oncoprotein appears and its expression is multiplied in the cells. The appearance of this protein that inhibits apoptosis plays a role in tumor maintenance, and its nuclear localization is associated with poor prognosis.

Because MCC can develop in combination with other malignant and semimalignant skin tumors, such as Bowen's disease, basal cell carcinoma, and most commonly squamous cell carcinoma, their characteristics may also color the histological picture. Squamous cell carcinoma mostly appears in an arrangement covering actinic keratosis or Bowen's disease, distancing itself from the dermal MCC component.

The transition between the two cell types is usually not visible. Presumably, the pathomechanism of this co-malignancy is different from that of other Merkel cell carcinomas.

Due to its rare incidence, there is currently a lack of standard therapeutic procedures ascribed for the treatment of MCC, especially when the disease is locally advanced. However, due to the increasing incidence of this tumor, modernization of treatment is an essential issue. Surgical excision and radiotherapy remain the mainstays of therapy in the initial stages.

MCC should be treated in highly specialized centers by an experienced multidisciplinary team. Surgical treatment is the mainstay of therapy in local and loco-regional primary MCC. For local disease, excision should be done with 1–2 cm margins and down to the fascia or periosteum. Patients with the clinical node-positive disease should undergo complete lymph node dissection. For clinical node-negative cases, sentinel lymph node biopsy is required. In some cases, complete lymph node dissection should be followed by radiotherapy. Depending on the size of the primary MCC, the 5-year survival rate is 66-75% for tumors with up to 2 cm and 50-60% for tumors above 2 cm. For patients with lymph node metastasis the 5-year survival rate is 42-52%, and for patients with metastasis, it is 17-18%. Until 2016, before the introduction of immunotherapy, the most common

treatments for metastatic MCC not amenable to surgery were chemotherapeutic regimens often used for other small-cell carcinomas. In 2017, avelumab (anti–programmed death-ligand 1 (PD-L1)) became the first approved treatment for patients with metastatic MCC, based on the occurrence of durable responses in a subset of patients.

Retroviruses and their role in oncogenesis

The members of the family *Retroviridae* are important pathogens of humans and animals. The ability of some retroviruses to induce tumors has been known since the turn of the 20th century. In 1908, Ellerman and Bang described a chicken erythroleukemia that was caused by a retrovirus followed by isolation of Rous sarcoma virus from a chicken fibrosarcoma by Peyton Rous. These discoveries mark the beginning of experimentation that led to our current understanding of retroviruses as cancer-causing agents. The list of animals affected by oncogenic retroviruses expanded as the 20th century progressed to include cats, cows, rats, mousses, sheep, goats, several primates, etc. and some fish. The discovery of Bittner and Gross revealed that betaretrovirus (mouse mammary tumor virus, MMTV) was associated with mammary epithelial tumors and thymic lymphomas in mice. MMTV, formerly known as Bittner virus, has a causal role and the most important etiology of mouse mammary tumors in feral and experimental mice. MMTV is a prototype species of the genus *Betaretrovirus* in the family *Retroviridae*. Multiple double-stranded MMTV DNA copies are also found in the chromosomal DNA (called integrated or endogenous proviruses) of most commonly used laboratory mice. The identification of a retrovirus that causes breast cancer in mice created great interest in determining whether similar viruses exist in other species, particularly in humans.

A number of studies showed that human breast cancer cells and tissues had proteins that cross-reacted with anti-MMTV antisera. Betaretrovirus was also detected in human milk from patients with an established diagnosis of breast cancer. The human mammary tumor virus (HMTV) 90-98% homologous to MMTV, the causative agent for breast cancer in mice, has been detected by PCR method in breast cancers approximately 40% of American women.

Human endogenous retroviral elements comprise 8% of human genome. Modern day exogenous retroviruses, as well as the infectious predecessors of endogenous retroviruses (ERVs), are demonstrably oncogenic. Furthermore, replication-competent ERVs contribute

to cancer development. Moreover, human cancers are characterized by transcriptional activation of human ERVs. These observations conspire to incriminate human ERVs as potential causative agents of human cancer.

Hamsters are popular pet animals with Syrian hamsters (*Mesocricetus auratus*) and Djungarian hamsters (*Phodopus sungaros*) being among the most frequently kept species. Reports on spontaneous tumors in domestic hamsters are scarce, and most are individual case reports. There are only limited and old (from the 1960-1970) knowledge related to retroviruses in hamster.

AIMS

The aim of our research was to search for viral genomic sequences in different skin and human tumor types using different molecular virological (PCR, gel electrophoresis, sequencing, viral sequence identification using metagenomics, *in situ* hybridization) and pathological (sample fixation, staining, sectioning), immunoassays. Our studies can be divided into two parts: MC polyomavirus from MCC tumors in humans and a new endogenous retrovirus from hamster skin tumor.

I. The study of MC polyomavirus from MCC tumors

- The aim was to search for and molecular biological detection of MC polyomavirus discovered in 2008 from histologically diagnosed samples of MCC archived in South Transdanubia between 2007 and 2012.
- The aim was to explore and analyze the relationships between the presence (or absence) of MC polyomavirus and the clinicopathological characteristics of the cases.
- The aim was to determine and study the detected MC polyomavirus viral genome.
- The aim was to search for additional new polyomaviruses (HPyV6, HPyV7, TSV, HPyV9) in skin tissue samples from the same patients.

II. Detection of a new endogenous retrovirus from a hamster skin tumor

- The aim was to investigate hamster skin tumor of unknown etiology using viral metagenomics.
- The aim was to analyze viral metagenomic results and to analyze potential viral sequences.
- The aim was to determine the entire genome of the retrovirus identified from the sample by molecular biological methods, and to analyze and phylogenetically compare the viral genome with known related retrovirus reference sequences.
- The aim was to investigate the presence of the retroviral genome (DNA and RNA) using *in situ* hybridization method in a tumor tissue sample.
- The aim was to search for a new retroviral genome sequence from different samples of additional hamster individuals using qualitative and quantitative PCR methods.

MATERIALS AND METHODS

I. Merkel cell carcinoma and Merkel cell polyomavirus

Samples selection

Archived skin samples collected retrospectively between 2007 and 2012 in the Department of Pathology, University of Pécs. The histological diagnosis relied on haematoxylin-eosin morphology and imunophenotype, using monoclonal antibodies as CK20, CD56, Synaptophysin, Chromogranin A and TTF1. All samples were reanalyzed histologically. Beside the primary MCC, the available skin and lymph node metastases, and the basal and squamous cell carcinomas adjacent to them were also tested. Samples with inadequate DNA quality were excluded from the study.

Epidemiological and clinical data

The patients' data have been collected from the electronic medical reports (permission number: KK/718-1/2020) with special attention to the case history, co-morbidities, potential immunosuppressive factors, location of MCC and the time spent between the appearance of symptoms and the diagnosis. Clinicopathological characteristics (age, histological picture and lymphovascular invasion, co-morbidities) have been evaluated in MC-positive and MC-negative cases.

Deparaffinization of tissue samples embedded in paraffin

If possible, 5-6 sections of 10-20 μm thick prepared from necrosis-free paraffinic block containing pure tumor tissue, which deparaffined with xylene and ethyl alcohol.

Viral DNA isolation and PCR

Viral DNA was isolated from native (N=6) or in paraffin embedded (N=9) tissue samples after deparaffinization using BKV Real-time PCR Kit (Shanghai Biotech, Shanghai, China). MC polyomavirus was detected by PCR method with a primer pair designed for the T region of MC polyomavirus. From a native sample of a lymph node metastasis, the total genome sequence of the MC polyomavirus was determined by primer walking method. Newly discovered polyomaviruses, polyomavirus 6, 7, 9 and trichodysplasia spinulosa associated polyomavirus were also screened by PCR method in MCCs samples.

Gel electrophoresis, sequencing and phylogenetic analysis

PCR-products were run in 1.5% agarose gel buffered with Tris-Borate-EDTA and stained with ethidium-bromide. PCR-products were sequenced directly using BigDye Terminator Kit and run on automated capillary DNA sequencer according to the manufacturer's instruction. GenBank database, ClustalX, GeneDoc and MEGA softwares were used for phylogenetic analysis.

II. Black Syrian hamsters skin tumor and BSHRV

Samples

Few-day-old male black Syrian hamster (*Mesocricetus auratus*) was purchased in January 2013 and kept as a pet in Pécs. Three months later skin tumor developed on the face and neck area of the hamster. During the following three weeks the neoplasia reached 25 mm in diameter and a skin metastasis on the chest. The hamster survived for 6 months and was euthanized by a veterinary because of unfavorable prognosis. Tumor specimen and tumor-free tissue samples from the tail, liver and lung were collected and used for molecular and histo/immunopathological investigation.

To acquire cellular nucleic acids by a non-invasive way, fresh faecal specimens which are known to contain intestinal epithelial cells were collected from another three, clinically healthy, 12-21-month-old black Syrian hamsters kept as a house pet by three different owners (medical students) in town of Pécs, in May and June 2018 to test for the study virus. One of these animals (a 1.5-year-old female) died – because of the old age - on November 22, 2018 with skin fibroma. Different *post mortem* tissue specimens (skin fibroma, lung and liver) were collected from this animal were also tested by (RT/) nested PCR and real-time PCR for the study virus.

Viral metagenomics

A tumor sample from the face was subjected to viral metagenomic and next generation sequencing analysis using random nucleic acid amplification of enriched tissue associated viral particles.

Complete viral genome determination

The freshly dissected tumor specimen and tumor-free tissue samples from the tail, skin, liver and lung were homogenized. Total RNA and DNA were extracted in separate reactions from freshly homogenized tissue samples and faecal suspensions for RNA and DNA isolations.

The complete genome of the black Syrian hamster retrovirus (BSHRV/2013/HUN, MK304634) was determined by primer-walking method and PCR techniques using DNA

extract from the tumor tissue. Sequence-specific oligonucleotide primers were designed using the retroviral reads and contigs determined by viral metagenomics. Both ends of the genome and the flanking chromosomal regions of the provirus were determined using Thermal Asymmetric Interlaced PCR (TAIL-PCR) method. The sequences determined by TAIL-PCR and viral metagenomics were verified by conventional and Long Range PCR reactions using reagents and methods described previously.

PCR-products were sequenced directly (Sanger method) with the specific primers and run on an automated sequencer.

Sequence- and phylogenetic analysis

ClustalX and GeneDoc software were used to align, assemble and compare the study strain and the prototype betaretroviral sequences available in GenBank. The phylogenetic tree from the aligned amino acid sequences of the *pol* and *env* gene was constructed by the neighbor-joining method using MEGA6. Bootstrap analysis of 1000 replicates was done to measure the significance of branching.

Histopathology

Samples for histopathological examination were fixed in 10% neutral buffered formalin, embedded in paraffin wax, sectioned at 5 µm thick slices and stained with haematoxylin and eosin.

***In situ* hybridization**

Two, 191-nt-long sense, and 193-nt-long antisense Digoxigenin-11-UTP labeled RNA probes were produced using TranscriptAid T7 High Yield Transcription Kit and T7-promoter-tagged silica-column purified PCR-products as templates. The sense and antisense probes were located in the presumed dUTPase (sense) and the retroviral envelope protein-coding (antisense) regions. The produced DIG-labeled RNA probes were purified. For the evaluation of the integrity, length and yield of the transcripts the purified samples were run on a 1% agarose gel with ethidium bromide. The produced DIG-labeled RNA probes were used for the detection of retroviral DNA and/or RNA using *in situ* hybridization on formalin fixed paraffin embedded slides.

Integrated viral genome copy number determination using quantitative PCR (qPCR)

For the absolute quantification of integrated viral copy number of tumor, tail, liver and lung tissue samples, the SYBR Green–based qPCR method was used.

RESULTS

I. Merkel cell carcinoma and Merkel cell polyomavirus

Altogether, 11 patients have been diagnosed histologically with MCC in the Department of Pathology, University of Pécs between 2007 and 2012. Three patients and their samples were excluded from the study because of the low sample quality.

Out of the eight patients, two had three different samples, three had two different samples and three patients had one sample each. MC polyomavirus was successfully detected by PCR in samples from three (37.5%) patients' primary tumors (2 samples), skin (1 sample) and lymph node (2 samples) metastases. In one patient, the virus was present both in the primary skin tumor, and the skin and lymph node metastases. None of the 15 samples were positive for polyomavirus 6, 7, TSV and 9 by PCR.

None of the preliminary clinical diagnoses was MCC before the histological examination. Pyogenic granuloma, keratoacanthoma, basal cell carcinoma and squamous cell carcinoma were established as primary clinical diagnoses. Out of the eight patients, six (75%) were male and two (25%) were female. In the case of the MC polyomavirus PCR-positive patients, 66% percent were female and 33% were male. The year-wise distribution of the histological diagnoses was as follows: one patient in 2007, 2008, and 2012 in each, two patients in 2010, and three patients in 2011. The youngest patient was 55 years old, the oldest one was 86. The patients' average age was 73.8 years (this was 64.3 years in the case of the MC polyomavirus PCR-positive patients). With the exception of the youngest patient (the MC-positive primary tumor appeared on his leg), all tumors were found at the head-neck region (face, forehead, ala of nose and neck). Factors predisposing to MCC such as

immunosuppression (e.g. prolonged oral steroid therapy, chronic lymphoid leukemia, and chronic obstructive pulmonary disease) and/or old age were found in all cases.

Morphological differences could be found in the histological pattern of the MC-positive and MC-negative groups. In the MC polyomavirus-associated cases, the histological samples have much more ordered structure (round or oval cells with vesicular nuclei and homogeneous cytoplasm), while the structure of MC-negative cases was much more disordered (irregular, heterogeneous tumor cells having polygonal nuclei and heterogeneous cytoplasm).

Additionally, there are differences in the lymphovascular invasion of the primary tumor. While the invasion of the vascular and lymphatic tissues could be observed at the edge of the tumor in the MC-negative cases at the appearance of the primary tumor, this could not be observed in the MC-positive cases.

The entire genome of MC polyomavirus, 5392 nucleotides in length, was determined from a native tissue sample from the lymph node metastasis of a 70-year-old female patient (GenBank: KC202810). The polyomavirus differs by 10 nucleotides (0.19%) from the prototype MC polyomavirus (EU375803). We found a truncated point mutation (from cytosine → thymine) in the coding portion of the large T (LT) protein of the viral genome at nucleotide position 1461, which is prominent in viral oncogenesis, resulting in abnormal cell proliferation and stopping virus replication.

MC polyomavirus is phylogenetically distinct from previously known human polyomaviruses and is most closely related to the first mouse polyomavirus (MPyV) capable of inducing mouse parotid tumor.

II. Black Syrian hamsters skin tumor and BSHRV

Case study and clinical pathology

At *post mortem* pathology the black Syrian hamster body weight was 40 g. The tumor was circumscribed and white-brown in color. On the cut surface a large central necrosis circumscribed by homogeneous grayish tissue was apparent. The metastasis on the chest was solid, well circumscribed, measured 0.7 cm in diameter and of white color on the cut surface. No other metastasis was observed in other parts of the body.

Detection of viral sequences by viral metagenomics

Primer tumor tissue from the face was subjected to viral metagenomics analysis. After *in silico* analysis and *de novo* assembly, 5593 sequence reads were obtained showing similarity to viruses (BLASTx cut-off E score $\leq 10^{-10}$). The detected sequences containing more than 100 reads were from virus family *Retroviridae* (N=2284), human endogenous retrovirus elements (N=613), other (N=1622) and unclassified (N=1074) viruses. The reads corresponding to virus family *Retroviridae* were further analyzed.

Determination of the retrovirus sequence

Using multiple sequence-specific primer-pairs designed for the metagenomic retroviral reads the viral DNA sequence was detectable from the DNA extracts of the tumor tissue as well as the additional analyzed tissue samples of liver, lung, tail by PCR. These results suggested that retroviral genome(s) are present in the DNA form likely integrated as a provirus in the hamster genome. Viral RNA was only detectable in the 2nd PCR round of RT nested-PCR in the DNase-treated RNA extract of the tumor but not in the other RNA samples of tumor-free organs.

Using a primer walking method, TAIL-PCR and multiple control PCR-reactions the continuous, complete genome of provirus, tentatively named black Syrian hamster retrovirus (BSHRV/2013/HUN, GenBank: MK304634), was determined from the DNA sample of the tumor tissue. The BSHRV genome is 8784 nucleotide (nt) in length and the genome organization follows the typical retroviral genome structure, 5'LTR-*gag-pro-pol-env*-3'LTR. BSHRV encodes the capsid/nucleocapsid (*gag*) proteins; the enzymes (reverse transcriptase and integrase) needed for genome replication (*pro/pol*) and the envelope proteins (*env*) that bind the cell surface molecules used for virus entry. Because BSHRV genome encodes at least all known fundamental retrovirus genome elements and proteins it is considered a complex betaretrovirus.

Phylogenetic analysis

Phylogenetic analysis of the translated *pol* and *env* genes of BSHRV were performed using the neighbor-joining option in MEGA6 and reference sequences from betaretroviruses. BSHRV phylogenetically situated in a similar cluster than MMTV confirming the genetic

relationship, however, it could represent a potential, taxonomically novel betaretrovirus species.

Histological findings

The haematoxylin and eosin stained histological sections of the tumor demonstrate a well-differentiated squamous cell carcinoma. The fibroblastic neostroma is infiltrated with typical flame-like tongues of centrally keratinized tumor cell nests.

***In situ* hybridization**

BSHRV-specific hybridization was predominantly restricted to the proliferating tumor cells of the squamous epithelium which formed numerous concentric keratin nodules. The hybridizations were observable mainly in the nuclei of the tumor cells in sections hybridized with either sense or antisense probes. Furthermore the localization of the hybridization signals were essentially the same in the sections labeled with the sense or the antisense probes. These results suggest that the probes were detecting the integrated proviral DNA of BSHRV not the viral RNA.

Detection of BSHRV-like betaretrovirus sequence from further black Syrian hamsters

One of the three DNA extracts from faecal specimens collected from further three animals were strong positive for BSHRV-like betaretrovirus by nested-PCR and sequencing. The amplified 538nt long *gag* gene sequence had 99.4% nt identity to the corresponding regions of BSHRV/2013/HUN. This animal (a 1.5-year-old female, animal 2) died – because of the old age - on November 22, 2018. Different *post mortem* tissue specimens (skin, lung and liver) were collected and all specimens tested also positive by nested PCR for the novel betaretrovirus. Based on the results of the qPCR analysis, tumor, liver, and lung samples from black Syrian hamster (animal 1) contained similar amounts of viral copy number (~ $1.18x10^4$/ng DNA), while the viral copy number ($2.03x10^4$/ng DNA) was almost twice in tail sample. Interestingly, the viral copy number in the tissue samples of female animal 2 was one order of magnitude higher (between $1.11x10^5$/ng and $1.76x10^5$/ng DNA) than in animal 1 with squamous cell skin tumor.

DISCUSSION

Oncogenic viruses may be one of the pathogens of cancer. It is estimated that 15-20% of human cancers are caused by microorganisms, including viruses. The currently known oncogenic viruses include viruses with DNA (EBV, HPV, HBV, HHV-8, and MCV) and RNA (HCV and HTLV-1) genomes. Presumably, the range of known human oncogenic tumor viruses is not yet complete. Numerous studies are currently underway to investigate and describe known, potentially carcinogenic viruses, as well as previously unknown carcinogens, in tumor diseases of human and animal origin. It is significantly more difficult to prove a causal relationship between a viral infection and a given tumor, as in many cases a longer time may elapse between events, the tumorigenic effect may become multifactorial as events progress. It adds color and excitement, and at the same time makes it difficult to see clearly the fact that in many cases the given viral genome can be present and function as an exogenous virus and as an endogenous genome sequence integrated into the genetic material of the host cell.

In our work, we searched for the viral etiology of a human and an animal skin tumor using a variety of advanced sequence-independent and sequence-dependent molecular biological methods. In the case of human Merkel cell carcinoma, we investigated the clinicopathological role of a newly described tumor-causing virus, MC polyomavirus, in human skin tumors; and in the case of skin tumor of unknown origin in the Syrian hamster, we tried to reveal the possible viral etiology by sequence-independent viral metagenomics.

I. Merkel cell carcinoma and Merkel cell polyomavirus

Merkel cell carcinoma is a novel clinical entity which is probably caused by a cancer-causing microbe. Although this previously rare disease is known for fifty years, the increasing number of the acquired immunosuppressive conditions has led to the increasing occurrence of the disease, and later to the confirmation of the infectious origin. In this study, MC polyomavirus was detected partly in histological samples characteristic for MCC by molecular techniques, and in one case the location of the truncating point mutation having crucial role in tumor formation was determined.

None of the preliminary clinical diagnoses raised the suspicion of MCC before the histological examination in cases above, i.e. it cause differential diagnostic problem causing

incorrect or incomplete diagnoses in clinical practice. Some of the conditions do not require urgent, wide surgical excision as opposed to the rapidly progrediating MCC. The early recognition of MCC is especially important, since the expected 5-year survival is 60-79% in the early stage (I/A), but it is only 18% by the time of the appearance of the first distant metastasis (stage IV).

The improvement of the knowledge, minimization of the risk factors, reduction of the uncertainty of early recognition and effective treatment as far as possible highly contribute to the increase in the patients' survival in MCC. We consider the case of the young 55-year-old patient especially instructive, since the development of the MCC and the patient's (early) death was probably activated by a medical treatment (PUVA) of a basically non-lethal disease (psoriasis), which therapy increases the risk of MCC nearly 100-fold. This raises the attention of the fact that this strong risk factor should be known and avoided as possible, the patients treated by this method should be checked up regularly, and in the case of suspicious skin changes, an early intervention is necessary. This also means that similarly to other tumor diseases, MCC also requires cooperation of different specialists. Team-working between the observing physician and the dermatologist, pathologist, oncologist, radiologist, surgeon, and in some cases the microbiologist may also be needed.

Although we worked with small sample size, and our results are based on retrospective data and archive histological samples, important conclusions can be seen clearly. Considering the epidemiological characteristics of our patients, we should highlight that the average age (73.8 years) was in accordance with the international data (70 years), and the male population (66%) predominates over the female population (similarly to the international tendency, 59%). Considering the localization of the tumor, there was no difference between the virus-associated and virus non-associated cases. Similarly to the literature data, the tumor appeared at the head-neck region in all cases, except for one; the youngest patient where the primary tumor presented on a naturally sun-protected area of the leg. All anamnesis revealed some factors predisposing to MCC like immunosuppression and/or old age. In our patients, the most frequent locations of the metastases were the skin and the lymph nodes; nevertheless the lymph node metastasis has changing frequency in the literature. This could be observed in six (75%) out of the eight cases, which can be partially explained by the fact that the tumors were diagnosed in late stages, and the treatments were also not adequate.

Results can be useful not only for clinical practice but also for laboratory diagnostics. The well-defined selection of patients/disease groups and better characterization of

differences between MC-polyomavirus positive and negative cases is an important step towards in the recognition of the etiology and pathogenesis of all MCCs. The diagnosis of MCC is based on histological examination. According to our observations, the presence of MC polyomavirus can already be suggested based on the histological examination. The histological pattern of the MC-negative tumor tissues shows much more irregular, heterogeneous tumor cells with polygonal nuclei and sometimes light-colored, less uniform cytoplasm. The MC-positive tumor cells are round or oval; they have vesicular nuclei and homogeneous cytoplasm. This study supports the theory that among the histologically proven MCC (more homogeneous histological pattern, lack of lymphovascular invasion) the presence of MC polyomavirus is associated with better prognosis. Mortality was 33% in our MCV-positive and 80% in MC-negative groups. Based on the international and national experiences, it is even certain that the virus non-associated cases might be an independent clinicopathological entities, which are similar to virus-associated types in clinical manifestation, but differ from them in histological picture and development. Other types of skin cancer are also commonly found in MC-negative cases, of which pathomechanism may be better understood with the help of a new research finding that Merkel cells have epidermal (and not neural crest) origin, i.e. the epidermal premalignant progenitor cells may develop into epithelial or neuroendocrine structures in principle. In MC polyomavirus-negative MCC cases, other factors, such as ultraviolet irradiation alone, immunosenescence and use of immunosuppressants may be involved in the carcinogenesis. Ionizing radiation may also be related to the occurrence of MCC. Consequently, the tumor presumably originates from unripe cells in these cases, which may explain the unfavorable progression of the MC-negative cases by itself. Although the presence of the MC polyomavirus and the integration of its genome into the chromosome of the MC cells seem to prove its role in the development of the tumor, the exact pathomechanism still raises many questions and requires further investigations.

It also deserves attention that the detection of the nucleic acid of MC polyomavirus is not equal to the diagnosis of the carcinoma. According to the present knowledge, only that viral nucleic acid can be responsible for the tumor, which integrated into the genome of the host and mutate oncogenic viruses may be one of the pathogens of cancer. This phenomenon is similar to those of papillomavirus, the closest relative of MC polyomavirus. The virus may be present “innocently” (i.e. it can be detected) e.g. in the cervix, but only that papillomavirus can be hold responsible for cervix carcinoma, which integrated into the genome of the infected host cell.

There is only small nucleotide differences in the genome of the MC polyomaviruses investigated in different geographical areas. Only every 540th nucleotide of MC polyomavirus found in our patient differs from those of the prototype virus (including the truncating point mutation results in a stop codon which has a prominent role in the viral oncogenesis, stops the viral replication and induces the pathological cell proliferation); that is the evolutionary mutation rate of these viruses having double-stranded DNA genome – related to the whole genome – may be low. This may also mean that the difficulties in the molecular diagnostics originating from the genetic variability should be less considered than that of the RNA viruses characterized by high mutation rate. Well-developed, sensitive, differentiating molecular biology techniques (based on either quantitative or mutational measurements) will be of significant diagnostic and prognostic importance in the detection of both the virus and the antibodies against the large T antigen (used as a biomarker) in MCC diagnostics in the future.

MC polyomavirus is the first known human cancer-causing polyomavirus, which associated with MCC. The potential infectious origin of this aggressive carcinoma leads to a change in attitude in the field of the pathogenesis and diagnostics. However, only detailed disclosure of the pathomechanism of the MCC may lead to effective prevention and specific treatment methods. MCC with viral origin – similarly to tumors caused by hepatitis B and papillomaviruses – may be preventable by vaccines in the future.

II. Black Syrian hamsters skin tumor and BSHRV

In this study, using viral metagenomics, histological analysis, *in situ* hybridization, PCR and RT-PCR approaches, we report the serendipitous identification, complete genome characterization and analysis of a novel endogenous retrovirus sequence, that was named black Syrian hamster retrovirus (BSHRV, MK304634), in black Syrian hamster (*Mesocricetus auratus*) initially with differentiated squamous cell skin cancer.

The BSHRV retroviral DNA sequence was detectable from the DNA extracts of the tumor tissue as well as the additional analyzed tissue samples of liver, lung and tail collected from the same animal by conventional PCR. These results suggested that the retroviral sequence(s) is possible present in integrated form in this host germ-line, as a provirus. Interestingly, the BSHRV retroviral RNA was detectable only in the 2nd PCR round of RT nested-PCR in the DNase-treated RNA sample of the tumor but not in the other RNA samples of tumor-free organs which suggest a low level viral RNA expression in the tumor

but not in the other analyzed tissues. The BSHRV retroviral DNA sequence was also identified in a faecal (most likely from host cells present in the faeces) and other tissue samples collected from an additional black Syrian hamster that was clinically healthy at the time of the specimen's collection. The association between the BSHRV and the skin tumor in hamster remains unknown. To prove that BSHRV, as do oncogenic retroviruses, influences or contributes to tumor genesis and therefore capable of transforming normal cells into malignant cells will require further studies.

The BSHRV retrovirus genome had full length and typical betaretroviral genome features and organization of 5'LTR-*gag-pro-pol-env*-3'LTR with dUTPase domain but without superantigen (*sag*) gene in 3'LTR. In the translated region, the *gag* gene encodes the characteristic betaretrovirus elements p10, p24 (capsid), zinc knuckle and the zinc finger. The *pro* gene encodes the dUTPase. The pro/*pol* gene includes the retropepsin, the reverse transcriptase (RT), the RNaseH and the integrase. The *env* gene encodes the envelope proteins. The complete and intact LTR/coding proviral genome regions of BSHRV are remarkable and one of the prerequisite for a viral replication and production of potential viable retrovirus particles.

By sequence- and phylogenetic analysis the proteins of the BSHRV had low (only 50-63%) amino acid sequence identity to the known betaretroviruses. BSHRV represents a potential novel betaretrovirus species but its proteins had the closest relationship to the corresponding proteins of mouse mammary tumor virus. MMTV is an important prototype member of the betaretroviruses associated with tumor in mice and conflicting results in breast cancer in human. In this context the determination of a novel sequence relative of MMTV-like betaretrovirus in a novel mammal host species increase the knowledge of the sequence diversity of these retroviruses. It would be worth considering this betaretrovirus sequence variant in screening primer design for PCR in human and animal "MMTV" prevalence studies in the future.

There are several important open questions related to BSHRV. Based on the methods used in this study the DNA copy number(s) in the host genome and the integration site(s) are remains unknown in either the tumor or the germ-line. The viral particle nature and the transmission mode(s) of the potential exogenous form of BSHRV virions are also undetermined. BSHRV or BSHTV-like betaretrovirus sequences have not been reported from hamsters. Further study are required to determine the prevalence and biological significance – including in tumorigenesis - of BSHRV and BSHRV-like betaretroviruses in hamsters and other mammals.

The role of microorganisms and the importance of infectious diseases are by no means a closed issue. We are confronted almost daily with newer and newer, previously unknown aspects of the microbial world. The relationship between microorganisms and cancers in this regard has always abounded in contradictions, slow acceptance or forgetting of results, and dogmas and dead ends. Due to multifactorial, time-consuming and complex disease processes, it is difficult to prove the cause. In our studies, we tried to expand our knowledge about a newly described human tumor virus without preconceptions, and we looked for a potentially viral background from a tumor lesion in a hamster kept as a pet. A science-based mapping of the world around us can lead us to describe and understand the living world around us.

ACKNOWLEDGEMENTS

I would like to thank all those who helped and contributed to my work to make this dissertation possible.

First I would like to express my gratitude to my mentor, Prof. Dr. Gábor Reuter, who, in addition to his professional advice and ideas, also created the financial conditions necessary for my work. I am grateful for your selfless help, for your unbroken support, for the transfer of your professional theoretical and practical knowledge, and for the trust you have placed in me in advance.

I am especially grateful for Dr. Ákos Boros and Dr. Péter Pankovics for the support in the practical implementation of my work. I am also grateful to all the staff of the Government Office for Baranya County, Regional Laboratory of Virology and the Department of Medical Microbiology and Immunology, Faculty of Medicine University of Pécs.

I am grateful to Dr. Endre Kálmán, senior clinical doctor (Institute of Pathology, University of Pécs), who helped both of my topics with his professional advice, remarks and ideas and provided professional assistance in the preparation of the autopsy and histology for my help.

I am grateful to Dr. Zsolt Kádár, senior clinical doctor and Prof. Dr. Zita Battyányi, dermatologist, for sharing patient samples, photographs, and experiences of their clinical work.

Thanks are due to Dr. Eric Delwart, Head of Laboratory (Vitalant Research Institute, San Francisco, CA, USA), who provided the opportunity to perform viral metagenomic and next-generation sequencing studies in his laboratory.

I would like to express my special thanks to my parents for supporting and encouraging me during my Ph.D. studies.

At last but not least, I wish to thank the rest of my family as well and my husband, Dr. Donát Szőgyi for encouraging my work by providing a solicitude background and our child, Viktória Szőgyi for their understanding and patience.

LIST OF PUBLICATIONS

PUBLICATIONS RELATED TO THE THESIS

1. **Horváth KB**, Pankovics P, Battyáni Zita, Kálmán E, Reuter G. Merkel-sejtes polyomavírus, mint a Merkel-sejtes carcinoma valószínű kóroka.
   Orvosi Hetilap 2013;154(3):102-112. doi: 10.1556/OH.2013.29525
   (IF: **0**)

2. **Horváth KB**, Pankovics P, Kálmán E, Kádár Z, Battyáni Z, Lengyel Z, Reuter G. Epidemiological, clinicopathological and virological features of Merkel cell carcinomas in Medical Center of University of Pecs, Hungary (2007-2012).
   Pathology and Oncology Research, 2016;22(1):71-77. doi: 10.1007/s12253-015-9974-z.
   (IF: **1,736**)

3. **Horváth KB**, Boros Á., Kálmán E, Pankovics P, Delwar E, Reuter G. Characterization of an integrated, endogenous mouse mammary tumor virus-like (MMTV) betaretrovirus genome in a black Syrian hamster (*Mesocricetus auratus*).
   Infection, Genetics and Evolution, 2019;75:103995. doi: 10.1016/j.meegid.2019.103995.
   (IF: **2,773**)

ABSTRACTS AND PRESENTATIONS THAT CAN BE CITED RELATED TO THE THESIS

1. **Horváth KB**. A Merkel-sejtes polyomavírus, mint a Merkel-sejtes karcinóma valószínű kóroka – kliniko-patológiai és virológiai vizsgálataink tapasztalatai. Tudományos Diákköri Konferencia (TDK), Pécsi Tudományegyetem, Általános Orvostudományi Kar, 2013. február 7-8, Pécs (bőrgyógyászat szekció: I. helyezés)
2. **Horváth KB**. A Merkel-sejtes polyomavírus, mint a Merkel-sejtes karcinóma valószínű kóroka – kliniko-patológiai és virológiai vizsgálataink tapasztalatai.

XXXI. Országos Tudományos Diákköri Konferencia (OTDK), Szegedi Tudományegyetem, 2013. április 2-5, Szeged (mikrobiológia szekció: I. helyezés)

3. **Horváth KB**. A Merkel-sejtes polyomavírus, mint a Merkel-sejtes karcinóma valószínű kóroka – kliniko-patológiai és virológiai vizsgálataink tapasztalatai. V. Nemzetközi és XI. Országos Interdiszciplináris Grastyán Konferencia, 2013. április 17-19, Pécs (a legközérthetőbb előadás díja).
4. **Horváth KB**, Pankovics P, Kálmán E, Reuter G. A Merkel-sejtes polyomavírus, mint a Merkel-sejtes karcinóma valószínű kóroka – kliniko-patológiai és virológiai vizsgálataink tapasztalatai. Magyar Infektológiai és Klinikai Mikrobiológiai Társaság, Fiatal Infektológusok Fóruma, 2013. június 6, Budapest (I. helyezés).
5. **Horváth KB**, Pankovics P, Kálmán E, Reuter G. A Merkel-sejtes polyomavírus, mint a Merkel-sejtes karcinóma valószínű kóroka – kliniko-patológiai és virológiai vizsgálataink tapasztalatai, Amerikai Magyar Orvosszövetség (HMAA) Konferenciája, 2013. augusztus 17, Balatonfüred. (poszter)
6. **Horváth KB**, Pankovics P, Kálmán E, Reuter G. A Merkel-sejtes polyomavírus, mint a Merkel-sejtes karcinóma valószínű kóroka – kliniko-patológiai és virológiai vizsgálataink tapasztalatai. Magyar Onkológusok Társasága XXX. Kongresszus, 2013. november 14-16, Pécs.
7. **Horváth KB** A Merkel-sejtes karcinóma, kliniko-patológiai és virológiai vizsgálatainak tapasztalatai. Doktoranduszok Országos Szövetsége, Tavaszi Szél Konferencia, 2014. március 22.

OTHER PUBLICATIONS

1. Pankovics P, Boros Á, Bíró H, **Horváth KB**, Phan TG, Delwart E, Reuter G. Novel picornavirus in domestic rabbits (*Oryctolagus cuniculus var. domestica*). **Infection, Genetics and Evolution**; 2016;37:117-22. 10.1016/j.meegid.2015.11.012. (IF: **2,885**)