Chapter

Basic Research and Clinical Examination of Tumor Virus

Tatsuaki Tsuruyama, Shuichiro Ogawa and Norishige Yamada

Abstract

Tumor viruses cause cancer; thus, extensive studies are being conducted on them. In this article, we will review the basic medical research on the current clinical genetic tests for tumor viruses such as human papilloma virus, hepatitis B virus, and T-cell leukemia virus. Recently, clinical genetic tests have been developed for quick diagnosis of the tumor virus infection. Additionally, we will review the mutagenesis of murine leukemia retrovirus. In particular, we will focus on the insertional mutagenesis. This will help in deciding the direction of future virus research by combining clinical and basic research.

Keywords: tumor virus, gene analysis, clinical examination, cancer diseases

1. Introduction

1.1 History of tumor virus research

Tumor virus research is an important subject to understand the molecular biology of tumor viruses and carcinogenesis. Dr. Katsusaburo Yamagiwa, a Japanese pathologist, is famous for the experiment in which coal tar was applied to rabbit ears to induce carcinogenesis in 1915. He was also a candidate for the Nobel Prize in Medicine. The following year, his student Dr. Akira Fujinami was appointed to study carcinogenesis in Germany for 4 years. After returning to Japan, in 1900, he became the first professor at the Department of Pathology, Kyoto Imperial University Medical College. In 1905, he started studying chicken tumors, and as early as 1910, he established a chicken breast tumor transplant system. He had suggested the possibility of carcinogenesis by filterable pathogens, which were later called Fujinami sarcoma virus (FSV). He is also well known for the discovery of Japan schistosomiasis.

Interestingly, in the same year, Peyton Rous, who was studying at the Rockefeller laboratory in the United States, also established the avian sarcoma transplant system. In 1966, Rous received the Nobel Prize in Physiology and Medicine for the discovery of carcinogenic viruses. Hanafusa et al. elucidated the FSV gene structure and identified the similarity with the Rous sarcoma virus gene structure [1]. Both FSV and Rous sarcoma virus belong to the same alpha-retrovirus. FSV has the oncogene v-fps in its genome [2]. It also has long terminal repeats (LTRs) at the ends of the genome that stabilizes as parts of the host cell genome when inserted into host mesenchymal cells after reverse transcription. In addition to the host cell division, the viral genome also replicates during the DNA replication cycle. Tyrosine-protein

kinase transforming protein Fps encoded by v-fps has the ability to induce transformation by promoting tyrosine phosphorylation of the PDGFbeta receptor [3, 4]. Thus, the Japanese scientists have greatly contributed to the field during the early days of oncology research.

1.2 Viral genome integration

Retroviral integration is a type of insertional mutagenesis. Integration processes are widely studied from biochemistry, and their molecular mechanisms are elucidated. Retrovirus-induced tumorigenesis is divided into two types. First, the retroviral genome carries the oncogene such as *v-onc*, and the inserted viral element may induce malignant tumors. Second, the inserted retroviral element has the promoter or enhancer activity of the host proto-oncogene. The genome of human papilloma virus (HPV), hepatitis B virus (HBV), human T-cell adult leukemia virus (HTLV) [5], avian sarcoma virus (ASV) [6], feline leukemia virus (FeLV) [7], murine leukemia virus (MLV) [8–13], and mouse mammary tumor virus (MMTV) can be integrated in the host genome. As a survival strategy, the inserted retroviral genome (provirus) is thereafter replicated during host cell division.

The retroviral promoter and enhancer elements in LTRs originally promote the expression of retroviral genes such as *gag*, *env*, and *pol* [14] (**Figure 1**). However, once integrated in the host genome, LTR elements can promote host gene expression in place of the host genome promoters and enhancers. After infection, the retroviral RNA genome is reverse-transcribed into DNA, which forms a pre-insertion complex (PIC) with the oligomeric integrase enzyme [15], which catalyzes the insertion reaction. The PIC crosses the nuclear envelope through the nuclear pore complex to access host cell DNA. The retroviral DNA ends then attack and nick the host DNA with target nucleotides. These ends are ligated to the truncated host DNA to complete the insertion process involving replication of the target nucleotides.

To date, extensive research has been conducted on target motifs for retroviral insertion to better understand tumorigenesis, neuronal damage, aging, and immunodeficiency including acquired immunodeficiency syndrome (AIDS). Retroviral-host DNA interactions support specific host DNA structures [16–21] but are less likely to be inserted into particular base sequences. Previous studies on human immunodeficiency virus (HIV)-1 insertion have shown that the insertion sites are distributed in 10-bp cycles on the nucleosome surface of the outward DNA major groove of chromatin [16, 22].

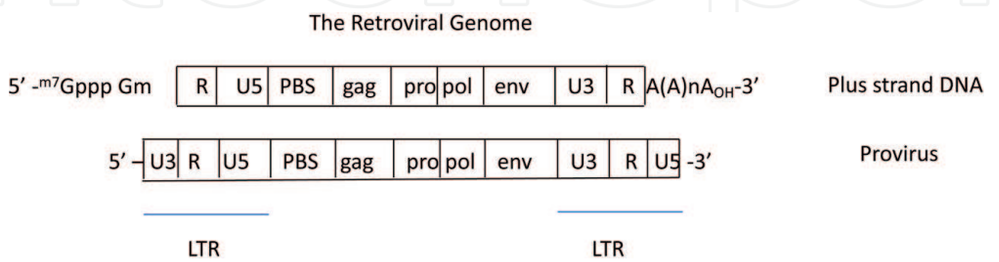

Figure 1.
A common structure of the retroviral genome. Plus strand RNA is shown. R, terminal redundant sequence; U5, unique 5′ sequence; PBS, tRNA primer-binding site; gag, group-specific antigen encoding capsid protein MA (matrix), CA, capsid protein; NC, nucleocapsid; pro, protease; pol, polymerase including reverse transcriptase (RT) and integrase (IN) catalyzing integration; env, envelope glycoproteins; U3, unique 3′ sequence. U3, R, and U5 are duplicated at both ends of the DNA provirus genome in the host genome sequence to form the long terminal repeat (LTR). LTR is essential for the initiation of viral DNA syntheses by reverse transcription as well as the integration of proviral DNA and the regulation of viral gene expression.

2. HPV

HPV is one of the most common viruses that infect more than half women once in life through sexual intercourse. It is a small DNA virus with approximately 8000 bp circular double-stranded DNA in its genome. The viral particle has icosahedral capsid structure as its outer shell. The virus has been known to be involved in the development of cervical, anal, and vaginal cancer and condyloma acuminatum. Many types of HPV are known, and high-risk HPV based on this cancer risk is concentrated in specific subgroups epidemiologically. In particular, types 16 and 18 have high carcinogenic risks of cervical intraepithelial neoplasia (CIN) and cancer. Type 16 has been detected in the Japanese 5-year-old girls [23]. High-risk group (types 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68) causes bladder cancer and pharyngeal cancer in addition to vulvar and cervical cancer. Low-risk group (HPV2, 27, and 57) causes verruca vulgaris that develops in the hands and feet [24].

HPV genome has early genes (E1, E2, E4, E5, E6, and E7) and late genes encoding viral capsid proteins (L1 and L2). E6 and E7 are considered to be involved in carcinogenesis. E6 binds to a tumor-suppressor gene product p53 and degrades p53 through promoting the ubiquitination. It also contributes to carcinogenesis by reactivating telomerase reverse transcriptase (hTERT) and degrading proteins with PDZ domain [shared by—post-synaptic density protein (PSD95), Drosophila disc large tumor suppressor (Dlg1), and Zonula occludens-1 protein (zo-1)]. E7 binds to the tumor suppressor gene pRb and contributes to carcinogenesis by releasing and activating the transcription factor E2F bound to pRb. It is also involved in carcinogenesis by interacting with cyclin-dependent kinase (Cdk) inhibitors p21 and p27 to evade cell cycle checkpoint mechanisms and apoptosis. However, in many cases, immortalization does not occur only by the actions of human cervix mucosal cells E6 and E7. In the well-known HeLa cell line, telomerase is activated, preventing the gradual shortening of telomeres in cell death.

High-risk HPV can be detected in more than 90% patients suffering from cervical cancer and dysplasia, which is a precancerous lesion, and in the vulva of 5% normal women. Although HPV is also integrated into the genome of host cells, its integration pattern has unique features. The E6 and E7 genes are retained in the provirus with frequent deletion mutations in E2 gene. Thus, the viral genome is also found to be truncated in the complex reaction of integration, but its carcinogenicity is rather enhanced. Low-grade dysplasia of the cervix and cervical intraepithelial neoplasia (CIN) is well known. These pathological changes are reversible and are known to spontaneously disappear within about 3 years. However, about 10% infected patients sustain the infection for more than 3 years, and cervical cancer occurs after several months to several decades in some of them. HPV genome is integrated into the host cell genome during the progression of CIN, and the progression becomes irreversible.

2.1 Clinical test for HPV detection

HPV-DNA typing tests have been conducted using real-time PCR, loop-mediated isothermal amplification (LAMP) [25, 26], invader, and hybridization methods. In recent years, the Clinichip method [27] combining LAMP is broadly available in advanced countries including Japan due to the ability to amplify genes at a constant temperature. The LAMP method is very specific because it uses four primers to recognize the six regions of the HPV genome. Nonspecific amplification does not occur because the amplification reaction depends on the order of the six regions of the HPV gene. In addition, since the amplification product constitutes

the target gene sequence, the target gene can be detected immediately by the presence or absence of amplification.

In an environment where electricity availability is limited, easy-to-use and durable reagents and equipment are desirable. In recent years, attempts have been made to spread virus testing as point-of-care tests (POCT) [28]. For some pathogens such as malarial parasite (*Plasmodium*), HIV, Ebola virus, and Zika virus, development of the testing system by fluid device is in progress; however, HPV detection methods have not been put to practical use as POCT yet.

3. HBV

HBV has been identified as the pathogen that causes acute and chronic hepatitis B [29, 30]. HBV is a DNA hepatitis virus and has a double structure of an envelope and a core. In persistent infections, the HBV genome is often inserted into the hepatocyte genome and is not eliminated by the host's immune system. Several studies on many common insertion sites have identified common integration sites in the HBV genome by full genome sequence of hepatocellular carcinoma [31]. The envelope protein of HBV covers the HBs antigen, and the core protein is called the HBc antigen. HBV DNA encodes HBs antigen, HBc antigen, X protein, and DNA polymerase. The HBe antigen is translated from the pre-C gene and the downstream HBc antigen. It is released into blood as a soluble protein during HBV virus growth and therefore considered a marker for the virus growth. In asymptomatic carriers positive for HBe antigen, the host immune system continuously reacts with HBV-infected hepatocytes, and biopsy reports show persistent lymphocyte infiltration in the portal vascular area of liver tissues. If the damage persists for a long time, remodeling of the hepatocytes occurs and progresses to cirrhosis with obvious fibrotic crosslinks. After all, when seroconversion occurs in HBV carriers negative for HBe antigen and positive for HBe antibody, the patient becomes an inactive carrier. HBV particles remain alive during periods of inactivity. Host immune system can attack infected hepatocytes, and hepatocyte regeneration may occur with genomic mutations, and the effects of the inserted HBV genome persist. Hepatocellular carcinoma can develop with persistent liver dysfunction associated with liver cirrhosis.

3.1 Clinical test for HBV detection

For the diagnosis of hepatitis B, HBs antigen/antibody, HBc antibody, HBe antigen/antibody, and HBV DNA are detected, and HBV DNA polymerase activity is measured. HBs antigen is produced continuously in patients with HBV infection; thus, positive results for HBe antigen can diagnose current infection with HBV. Recently, the HBs antigen measurement system has become very sensitive, and infected people have been found to be positive for HBs. In addition to antigen or antibody detection method, viral DNA detection system (quantitative PCR, LAMP method) has been used for clinical tests. Liver fibrosis is seen as a precancerous stage of HBV and hepatitis C virus (HCV) infection. Liver biopsy and qualitative analysis of hyaluronic acid and type IV collagen in blood are well-known methods to evaluate liver fibrosis progress. Monitoring of glycolysis of serum M2BP (Macingo binding biomarker) is used to evaluate liver fibrotic stage [32]. Recent studies have suggested that the most common cause of nonexclusion of HBV is covalently closed circular DNA (cccDNA). When HBV infects hepatocytes, the incomplete duplex circular DNA genome becomes a complete duplex and is retained in the nucleus in the form of cccDNA. This HBV gene is the source of HBV

replication. Since cccDNA is an extremely stable biochemical structure, it can be used as a monitoring marker after the infection. As the nucleic acid analog drug has little effect on reducing HBV cccDNA, a high risk of reactivation remains. If host immunity is suppressed, the reactivation occurs from this cccDNA as a starting transcriptional point. Therefore, elimination or suppression of reactivation will be central to future therapeutic strategies [33].

In the clinical tests for HBV detection, the detection sensitivity of the enzyme-linked immunosorbent assay (ELISA) method for HBs antigen is higher than those of other tests. Each available method depends on the titer of the used antibody, and the sensitivity varies. In particular, the mutation in the gene encoding the envelope protein HBs antigen may cause false negative results. Many test kits increase sensitivity by using multiple monoclonal antibodies or a polyclonal antibody that has an affinity for multiple epitopes. In some cases, antibodies that recognize epitopes within the viral particle are used instead of surface proteins with many mutations. However, the high sensitivity of many kits may instead cause false positive results. Accordingly, actual early diagnosis requires careful evaluation of test results in combination with other clinical tests.

4. HTLV type 1 (HTLV-1)

Another important human carcinogenic virus is human T-cell leukemia virus type 1 (HTLV-1). This virus has the potential to cause adult T-cell leukemia/lymphoma (ATL), HTLV-1-associated myelopathy (HAM), and HTLV-1 uveitis (HU). These diseases develop in HTLV-1 carriers after prolonged latent infection. HTLV-1 carriers are particularly prevalent in southwest Japan, including Kyushu, Shikoku, and Okinawa. HTLV-1 was isolated and reported by Gallo R as the first human pathogenic retrovirus in 1980. Yorio Hinuma of Kyoto University in Japan precisely analyzed the genomic structure and identified the HTLV-1 [34]. HTLV-1 is a weakly infective virus particle, and cell-to-cell transmission is the main route of infection.

4.1 Molecular biology of HTLV-1

ATL is a hematologic malignancy in which abnormal lymphocytes appear in peripheral blood and infiltrate various organs throughout the body. Kiyoshi Takatsuki of Kumamoto University in Japan and Taku Uchiyama of Kyoto University collected samples. In 1981, Hinuma identified HTLV-1 to be the causative virus of ATL; this was the first report on tumor-induced retrovirus in humans. Most infected cells are CD4 + CD25 + T cells. Viral receptors are thought to involve other proteins and heparan sulfate. HTLV-1 is an approximately spherical virus particle about 100 nm in diameter and has an envelope. It has an RNA genome composed of the genes common to retroviruses such as *gag*, *pro*, *pol*, and *env*, as well as the gene encoding pX. It selectively uses the splicing and translation initiation sites in a region less than 2 kb called pX located between the env gene and the 3′ LTR [35] and uses the plus and minus strands relative to the transcription direction and efficiently. It often encodes several regulatory and accessory genes. As described later, this pX is involved in the action of HTLV-1 encoding HTLV-1 bZIP factor (HBZ).

Tax and HBZ may be the carcinogenic potentials of HTLV-1. Tax is a potent transactivator essential for the activation of the viral promoter 5′ LTR. After entering the cell by infection, the virus particle capsid is lost, and the HTLV-1 genomic RNA is reverse transcribed to form genomic duplex DNA. Thereafter, the integrase encoded by the viral genome and the DNA end forms a complex, which contacts the genomic DNA on the cell side, and inserts by catalytic reaction. The 5-base-long

sequence of the targeted host cell genome is duplicated at both ends of the inserted viral genome in the insertion process. Tax, the transcript of TLV-1 regulatory gene tax, not only binds to the HTLV-1 promoter to promote viral transcription and expression but also interacts with cell-side transcription factors to activate host gene transcription activity. Therefore, *tax* is implicated in the early stages of tumorigenesis of infected cells. However, in peripheral blood of most ATL patients, Tax is suppressed by its promoter 5′ LTR deletion and epigenetic modification. Thus, Tax is not essential, at least at the final stage of carcinogenesis. Recently, Matsuoka has suggested that Tax is constitutively expressed in a small fraction of leukemia cells, and this fraction triggers an anti-apoptotic mechanism to establish sustained infection and induces oncogenesis [36]. Tax is induced by a variety of cytotoxic stresses and also promotes HTLV-1 replication. Thus, it protects infected cells from apoptosis and increases the chance of viral infection at the critical phase of disease in the carrier. Another HBZ gene is expressed in all ATL patients causing and promoting the growth of ATL cells; therefore, it may be the gene responsible for ATL.

4.2 Insertion site of HTLV-1

At present, in addition to the carcinogenesis by the proteins of these viruses themselves, attention has been focused on the activation of host genes at the insertion site. The integrated HTLV-1 provirus has the LTR [35] in both directions for the transcription. This sequence has many motifs that bind to transcription factors on the host and have strong promoter/enhancer activities. Therefore, there may be an increase in constitutive expression of the gene at the insertion site. Ogawa performed systematic genome analysis of ATL cells in 426 ATL patients and showed that genes interacting with Tax, T-cell receptor-NF-κB signal transduction, T-cell transport, and other T-cell-related pathways as well as genes related to immune surveillance are injured by the integration. Also, the expressions of VAV1, IRF4, and FYN related to lymphocyte maturation and signal transduction, chemokine receptors CCR4 and CCR7, and gene fusions (CTLA4-CD28 and ICOS-CD28) were enhanced in the infected cells [37]. Viral transcripts were mainly derived from the antisense strand. Also, the suppression of *Tax* expression and the constitutive expression of *HBZ* were observed in almost all patients. These insertion sites have not been known to be likely to occur in any particular sequence motif.

4.3 Clinical test for HTLV-1 detection

To date, no chromosomal abnormality specific to HTLV-1 has been observed, and Southern blot or inverse PCR methods are used for diagnosis; however, in recent years, multiplex LAMP (RT-LAMP) using universal probe has been developed. The method detects both HIV and HTLV-1 RNA from the same sample [39]. HAS HTLV-1 analyzing system (HAS)-Flow method has also been proposed to evaluate the progression of ATL stage by applying flow cytometric analysis focusing on surface markers of infected immune cells [38].

5. Integration model in mice lymphomas

5.1 The specificity of insertion sequence of the retrovirus genome

As experimental models of MLV integration, BXH2 [40], AKXD [41–44], and SL/Kh [20, 45–52] are well-known mouse strains into which endogenous provirus

genomes have been inserted. The genome of an endogenous MLV, AKV [14], was integrated into *gltscr 1* [20], which has been characterized as a tumor suppressor gene. This locus was identified as *emv11* on chr 7. In our previous work, the inserted AKV genome was shown to contain the full-length MLV genome and expressed to form a retroviral particle, which had the ability to infect lymphocytes according to its retroviral tropism. In the consequent tumorigenesis, the phenotype of the murine tumor is generally variable in inbred strains. For example, AKR/J mice are susceptible to T-cell lymphoma/leukemia [53–55]; however, in BXH2 mice [39, 40], tumors resemble myeloid and lymphocytic tumors. Even if the MLV is integrated into identical genes, the tumor phenotypes are variable [56]. It is likely that the individual genetic backgrounds of the inbred mouse strains determine the phenotype in combination with the integration-induced expression of oncogenes. Thus, oncogenes are responsible for the tumorigenesis; however, the phenotypes of the tumors are determined by other elements, including the genetic background. The insertion site of MLV has been studied using model mouse strains. SL/Kh strain is known to be a model of acute lymphocytic leukemia/lymphoma by MLV infection, and the target sequence of integration has been identified. This phenomenon may not always be generalized in human retroviruses as MLV LTR is very long and has palindromic motifs unlike HTLV-1 LTR. This sequence motif may contribute to the stabilization of the structure as the tetrameric integrase binds to the LTR. Such a stabilizing structure is not found in HTLV-1, and the retroviral genome may differ in HTLV-1 and MLV in this regard. The existence of a specific sequence motif in the genome insertion of the virus has been the subject of much debate and was initially resolved negatively in early days. The conclusion was that a virus integration is a random event [57] (by Howard Temin, who discovered reverse transcriptase and won the Nobel Prize in 1975). The randomness was proved through the genome analysis of infecting fibroblasts with integrated MLV in the culture state artificially. No common sequence motif for the insertion site was found in this way. However, there has been a report that active transcriptional genes can be the target of integration [21]. Besides, there are reports stating that palindrome sequence is favored by integration [16, 58]; however, the target sequence itself is yet to be identified. The idea is that viral insertion affects the transcriptional activity of the host cell genome by the three-dimensional structure of chromatin. A model may be considered in which the integrase complex binds to the site where DNA is locally released during transcription.

To study MLV genome insertion *in vivo,* inbred strains of mice harboring many copies of endogenous MLV genomes were developed. SL/Kh is one of the inbred strains (Tsukuba, Japan; http://www.riken.jp/engn/index.html) [14, 45]. In this strain, the harbored MLV genome is spontaneously and selectively integrated into the genome of the B-cell genome DNA. In reference to the data about MLV insertion sites in the genome of SL/Kh mice, whether MLV insertion occurred preferentially into the specific sequence motifs in the host cell genome was investigated. *Stat5a* [20, 51, 59, 60] and *c-myc* [47] were the target genes for MLV insertion. More than 95% studied SL/Kh mice had such the MLV genome (AKV) insertion in zinc finger protein 521 gene (*Zfp521*) in the genomic DNA of B-cell lineage at 7 weeks of age [61] (**Figure 2**).

5.2 Integration target sequence

To date, such significantly selective insertion into the specific gene has not been known. When insertion further occurs in an oncogene (e.g., *Stat5a* [51] or *c-myc* [47]) in addition to *Zfp521* in the genome of SL/Kh mice, B lymphocytic leukemia develops.

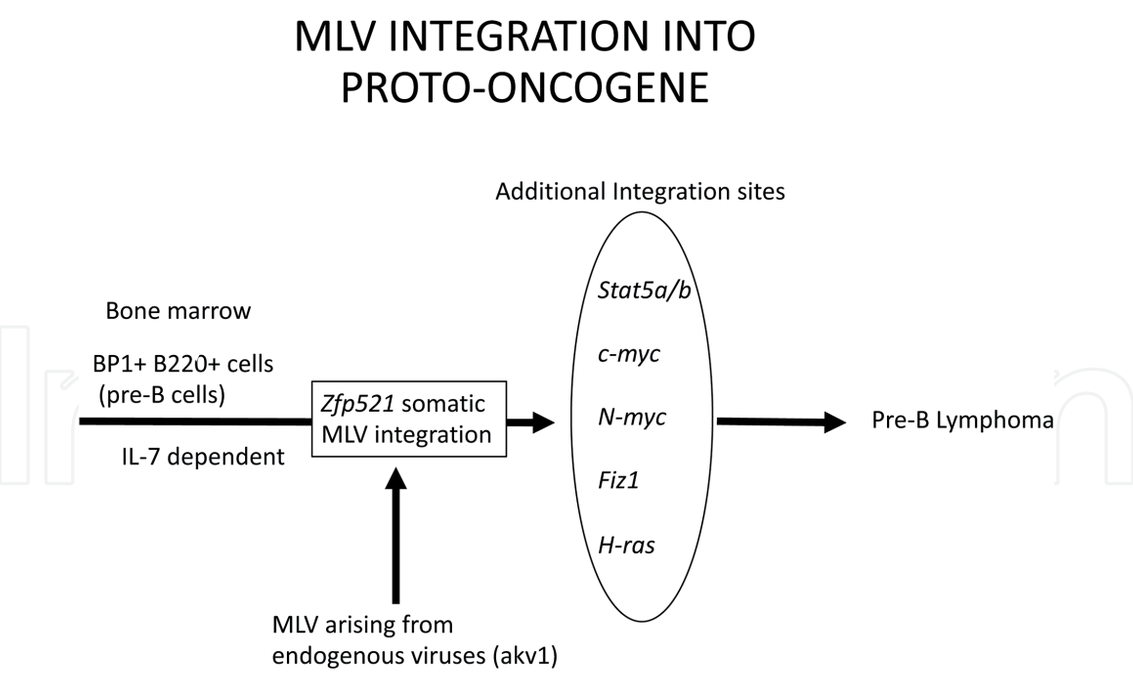

Figure 2.
Murine leukemia retrovirus (MLV) integration into oncogenes in SL/Kh lymphoma. In most cases of lymphoma, MLV integrations into more than one oncogenes, including Stat5a, c-myc, Fiz1, Hipk2, and others, are observed. It is not clear why the tumor cells are restricted to pre-B lymphoma phenotype. To date, genetic background dependent on Bomb1 is critical for conferring the phenotype. In addition, Stat5a and Zfp521 genes cooperate with IL-7 and pre-B cell receptor signal pathways, respectively. Host cell phenotype at the integration may contribute to the phenotypes. Differentiation stage of the host cell at the mutagenesis may be one of the critical factors of the phenotype.

In AKXD and SL/Kh lymphoma genome, *Zfp521* is one of the most frequent targets of MLV integration [62]. This gene locus was first reported as a common integration site, ecotropic retrovirus integration site 3 (*Evi3*) [61–64]. In more than 400 male SL/Kh mice, >90% pre-B-lymphoma genomes acquired the integrated proviral genome within a 50-bp segment located in the region of the second to third exon of *Zfp521*. This target gene carries alternating palindrome sequences [16], which was frequently targeted by MLV insertion. *Zfp521* is one the novel identified target genes. It probably contributes to chondrocyte development [65], in addition to B-cell development that requires an abnormal chimeric gene [64]. Each insertion occurs once or twice during the development of lymphoma within a 50-bp segment located in the second intron to the third exon of *Zfp521*. Most commonly, the insertion occurs at the ":" site in the sequence "CTGAATTG*AAAC:* AACTTCAGC*TGTTT*," where the pair of underlined sequences and the pair of sequences in italics are palindromic sequences [61–63, 66]. The insertion does not occur randomly; they frequently occur at approximately 10-bp intervals and are symmetrically distributed within the shown 50-bp segment (**Figure 3**).

SL/Kh mouse has multiple copies of the endogenous virus, which produces viral particles and infects immature B-cell lineages suitable for tropism. Several common insertion sites such as Stat5a and Stat5b and well-known genes such as c-myc have been identified till date. Interestingly, many of these insertions occur upstream of the translational initiation site of the protein, so that many intact targeted proteins are produced. The fact that many insertions occur in such specific genes suggests that DNA structures such as palindromic sequences are targeted. However, insertions in a short 50-bp segment cannot have such secondary structure. MLV insertion was observed in almost the entire genome of this SL/Kh acute lymphocytic leukemia/lymphoma. The integration target segment is located just before the translation initiation

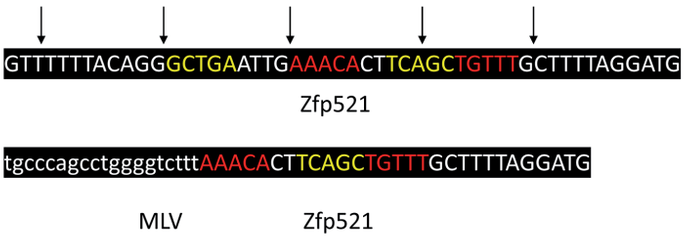

Figure 3.
Alternate palindrome sequence in Zfp521. The upper and lower sequences represent the native sequence of host murine Zfp521 gene and the junction sequence of host murine Zfp521 gene with the integrated MLV genome, respectively. Green and red letters represent pairs of alternating palindromes. The sequences in the same colors are palindromes. The downward arrows represent the hot spot of the MLV integration.

origin of the Zfp521 gene, which was first reported as a retrovirus insertion sequence. As a result of the insertion, downstream Zfp521 gene is constitutively expressed, and the upregulation may be responsible for acute lymphocytic leukemia/lymphoma.

5.3 Interaction partners for MLV integrase and recent developments in the biosafety of gene therapy

Recently, several cellular proteins including transcription regulatory factors and chromatin and RNA-binding proteins have been discovered as potential interaction partners for MLV integrase [67]. The bromodomain and extra-terminal domain (BET)-containing family of proteins (BRD2, BRD3, BRD4, and BRDT) act as cellular-binding partners of MLV integrase and preferentially bind to open chromatin regions that are enriched for transcription start sites, CpG islands, DNase I-hypersensitive sites, and proto-oncogenes [68–73]. The BET family of proteins plays a critical role in regulating gene transcription through epigenetic interactions between dual bromodomains and acetylated lysines on histones H3 and H4 during cellular proliferation, cell cycle progression, and differentiation processes [8, 9]. *In vitro* interaction analysis in human cells revealed that the ET domain of BRDs, BRD4 in particular, interacts with a motif in the unstructured C-terminal tail of MLV integrase, and these interactions are crucial for the recruitment of the MLV PIC to host DNA during integration [10, 11]. Through the recognition of open chromatin structure, BET proteins have been suggested to contribute to the tethering of the MLV PIC to the host chromatin [10, 12]. Furthermore, it has been shown that the cell-permeable small molecule JQ-1, a pan-BET bromodomain inhibitor, prevents the BRD4-acetylated lysine interaction by competitively binding to BRD4 and reduces MLV integration frequencies at transcription start sites [4, 13, 14]. These findings suggest that BET proteins navigate the MLV genome and promote efficient MLV integration around transcription start sites associated with a chromatin structure characteristic of an open, transcriptionally active domain [3].

The dual bromodomain protein of BRD2 promotes B-cell expansion and plays a critical role in the regulation of mitogenic response genes, particularly cyclin A, in B cells [15]. In addition, it has been shown that the inhibition of BRD4 leads to the suppression of germinal center B-cell differentiation and antibody responses through the down-regulation of nuclear factor-κBp65 activation and B-cell lymphoma 6 expression [16]. Hence, these BET proteins might broadly contribute to tumorigenesis in cooperation with MLV integration in spontaneous B-cell lymphoma development, and these findings likely provide novel standpoints of proto-oncogene function as well as a better understanding of MLV insertional mutagenesis [15, 17].

The stable integration of retroviral vectors is suitable to deliver therapeutic genes into cells to correct genetic diseases. MLV-based retroviral vectors are one of the most frequently used gene delivery vehicles in gene therapy studies for primary immunodeficiencies such as adenosine deaminase deficiency-severe combined immunodeficiency [10, 18, 19]. However, a subset of patients developed serious adverse events including leukemia that could be attributed to the integration pattern and vector design [10, 18]. MLV-based retroviral vectors integrate in the vicinity of transcription start site through the interaction between the C-terminal tail of MLV integrase (amino acids 381–408) and the ET domain of BET family of proteins [18, 20]. To resolve side effects such as leukemia, BET-independent MLV vectors have been developed by truncation of the C-terminal tail of MLV integrase or the single W390A mutation [18, 20]. Most recently, next-generation BET-independent MLV vectors have been engineered with the addition of the chromodomain of Chromobox protein homolog 1 (CBX1) to MLV integrase$_{_W390A}$. The MLV integrase$_{_W390A\text{-}CBX1}$ efficiently targets integration away from traditional markers of MLV integration including gene regulatory elements [10]. These findings open new avenues to improve the biosafety of gammaretroviral vectors for gene therapy.

6. Summary

We followed the history of viral carcinogenic research in this review. The topic of viral carcinogenesis mentioned here is a case where the viral genome is inserted into the host genome. Because the molecular mechanism of viral carcinogenesis has been elucidated, significant novel technical methods will be developed in the future.

Clinical sequence analyzes gene changes occurring at diseased sites and provides information useful for diagnosis and treatment of diseases based on the results of gene mutations.

In addition to the various nucleic acid analysis methods such as conventional PCR and LAMP, next-generation sequencing and other rapid technological innovation will aid in the measurement and analysis of numerous gene sequences in a short period. Furthermore, these technologies are expected to be used in the future to analyze infectious diseases such as cancer virus types, HPV high-risk groups, and virus insertion sites. Information on cancer virus infection will also be valuable for designing personalized drugs for cancer patients in the future.

Acknowledgements

This work was supported by a Grant-in-Aid from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (no. 15H05721) and a Grant-in-Aid from the Scientific Research on Innovative Areas, MEXT, Japan (no. P2013-201).

Conflicts of interest

The authors declare no conflicts of interest.

Abbreviations

HPV	human papilloma virus
HBV	hepatitis B virus
HTLV-1	human T-cell leukemia virus type 1
MLV	murine leukemia virus

Author details

Tatsuaki Tsuruyama*, Shuichiro Ogawa and Norishige Yamada
Department of Drug and Discovery Medicine, Clinical Bioresource Center of Kyoto University Hospital, Sakyo-ku, Kyoto, Japan

*Address all correspondence to: tsuruyam@kuhp.kyoto-u.ac.jp

References

[1] Hanafusa T, Wang LH, Anderson SM, Karess RE, Hayward WS, Hanafusa H. Characterization of the transforming gene of Fujinami sarcoma virus. Proceedings of the National Academy of Sciences of the United States of America. 1980;**77**(5):3009-3013. DOI: 10.1073/pnas.77.5.3009

[2] Mathey-Prevot B, Shibuya M, Samarut J, Hanafusa H. Revertants and partial transformants of rat fibroblasts infected with Fujinami sarcoma virus. Journal of Virology. 1984;**50**(2):325-334

[3] Anderson DH, Ismail PM. v-fps causes transformation by inducing tyrosine phosphorylation and activation of the PDGFbeta receptor. Oncogene. 1998;**16**(18):2321-2331. DOI: 10.1038/sj.onc.1201780

[4] Wang Y, Li J, Li Y, Fang L, Sun X, Chang S, et al. Identification of ALV-J associated acutely transforming virus Fu-J carrying complete v-fps oncogene. Virus Genes. 2016;**52**(3):365-371. DOI: 10.1007/s11262-016-1301-6

[5] Derse D, Crise B, Li Y, Princler G, Lum N, Stewart C, et al. Human T-cell leukemia virus type 1 integration target sites in the human genome: Comparison with those of other retroviruses. Journal of Virology. 2007;**81**(12):6731-6741. DOI: 10.1128/Jvi.02752-06

[6] Aiyar A, Hindmarsh P, Skalka AM, Leis J. Concerted integration of linear retroviral DNA by the avian sarcoma virus integrase in vitro: Dependence on both long terminal repeat termini. Journal of Virology. 1996;**70**(6):3571-3580

[7] Levy LS, Lobelle-Rich PA, Overbaugh J. flvi-2, a target of retroviral insertional mutagenesis in feline thymic lymphosarcomas, encodes bmi-1. Oncogene. 1993;**8**(7):1833-1838

[8] Beverly LJ, Capobianco AJ. Perturbation of Ikaros isoform selection by MLV integration is a cooperative event in Notch(IC)-induced T cell leukemogenesis. Cancer Cell. 2003;**3**(6):551-564

[9] Hematti P, Hong BK, Ferguson C, Adler R, Hanawa H, Sellers S, et al. Distinct genomic integration of MLV and SIV vectors in primate hematopoietic stem and progenitor cells. PLoS Biology. 2004;**2**(12):e423. DOI: 10.1371/journal.pbio.0020423

[10] Jeanteur P. Differences in the MLV and HIV-1 integration sites could be important for gene therapy. Bulletin du Cancer. 2003;**90**(8-9):671

[11] LaFave MC, Varshney GK, Gildea DE, Wolfsberg TG, Baxevanis AD, Burgess SM. MLV integration site selection is driven by strong enhancers and active promoters. Nucleic Acids Research. 2014;**42**(7):4257-4269. DOI: 10.1093/nar/gkt1399

[12] Schambach A. Toward a safer integration profile of MLV-based retroviral vectors. Molecular Therapy. 2014;**22**(8):1405-1406. DOI: 10.1038/mt.2014.124

[13] Tsuruyama T, Hiratsuka T, Yamada N. Hotspots of MLV integration in the hematopoietic tumor genome. Oncogene. 2017;**36**(9):1169-1175. DOI: 10.1038/onc.2016.285

[14] Lenz J, Crowther R, Straceski A, Haseltine W. Nucleotide sequence of the Akv env gene. Journal of Virology. 1982;**42**(2):519-529

[15] Elis E, Ehrlich M, Prizan-Ravid A, Laham-Karam N, Bacharach E. p12 tethers the murine leukemia virus pre-integration complex to mitotic

chromosomes. PLoS Pathogens. 2012;**8**(12):e1003103. DOI: 10.1371/journal.ppat.1003103

[16] Holman AG, Coffin JM. Symmetrical base preferences surrounding HIV-1, avian sarcoma/leukosis virus, and murine leukemia virus integration sites. Proceedings of the National Academy of Sciences of the United States of America. 2005;**102**(17):6103-6107. DOI: 10.1073/pnas.0501646102

[17] Lewinski MK, Yamashita M, Emerman M, Ciuffi A, Marshall H, Crawford G, et al. Retroviral DNA integration: Viral and cellular determinants of target-site selection. PLOS Pathogens. 2006;**2**(6):e60. DOI: 10.1371/journal.ppat.0020060

[18] Pruss D, Reeves R, Bushman FD, Wolffe AP. The influence of DNA and nucleosome structure on integration events directed by HIV integrase. The Journal of Biological Chemistry. 1994;**269**(40):25031-25041

[19] Schröder AR, Shinn P, Chen H, Berry C, Ecker JR, Bushman F. HIV-1 integration in the human genome favors active genes and local hotspots. Cell. 2002;**110**(4):521-529. DOI: 10.1016/s0092-8674(02)00864-4

[20] Tsuruyama T, Hiratsuka T, Jin G, Imai Y, Takeuchi H, Maruyama Y, et al. Murine leukemia retrovirus integration induces the formation of transcription factor complexes on palindromic sequences in the signal transducer and activator of transcription factor 5a gene during the development of pre-B lymphomagenesis. American Journal of Pathology. 2011;**178**(3):1374-1386. DOI: 10.1016/j.ajpath.2010.12.012

[21] Wu XL, Li Y, Crise B, Burgess SM. Transcription start regions in the human genome are favored targets for MLV integration. Science. 2003;**300**(5626):1749-1751

[22] Wang GP, Ciuffi A, Leipzig J, Berry CC, Bushman FD. HIV integration site selection: Analysis by massively parallel pyrosequencing reveals association with epigenetic modifications. Genome Research. 2007;**17**(8):1186-1194. DOI: 10.1101/gr.6286907

[23] Kojima A, Maeda H, Kurahashi N, Sakagami G, Kubo K, Yoshimoto H, et al. Human papillomaviruses in the normal oral cavity of children in Japan. Oral Oncology. 2003;**39**(8):821-828

[24] de Martel C, Ferlay J, Franceschi S, Vignat J, Bray F, Forman D, et al. Global burden of cancers attributable to infections in 2008: A review and synthetic analysis. The Lancet Oncology. 2012;**13**(6):607-615. DOI: 10.1016/S1470-2045(12)70137-7

[25] Nagamine K, Kuzuhara Y, Notomi T. Isolation of single-stranded DNA from loop-mediated isothermal amplification products. Biochemical and Biophysical Research Communications. 2002;**290**(4):1195-1198. DOI: 10.1006/bbrc.2001.6334

[26] Notomi T, Okayama H, Masubuchi H, Yonekawa T, Watanabe K, Amino N, et al. Loop-mediated isothermal amplification of DNA. Nucleic Acids Research. 2000; **28**(12):E63. DOI: 10.1093/nar/28.12.e63

[27] Satoh T, Matsumoto K, Fujii T, Sato O, Gemma N, Onuki M, et al. Rapid genotyping of carcinogenic human papillomavirus by loop-mediated isothermal amplification using a new automated DNA test (Clinichip HPV™). Journal of Virological Methods. 2013;**188**(1-2):83-93. DOI: 10.1016/j.jviromet.2012.10.014

[28] Chen H, Liu K, Li Z, Wang P. Point of care testing for infectious diseases. Clinica Chimica Acta. 2019;**493**:138-147. DOI: 10.1016/j.cca.2019.03.008

[29] Purcell RH, Gerin JL, Almeida JB, Holland PV. Radioimmunoassay for the detection of the core of the Dane particle and antibody to it. Intervirology. 1974;**2**(4):231-243. DOI: 10.1159/000149428

[30] Zavalin A, Todd EM, Rawhouser PD, Yang J, Norris JL, Caprioli RM. Direct imaging of single cells and tissue at sub-cellular spatial resolution using transmission geometry MALDI MS. Journal of Mass Spectrometry. 2012;**47**(11):i. DOI: 10.1002/jms.3132

[31] Furuta M, Tanaka H, Shiraishi Y, Unida T, Imamura M, Fujimoto A, et al. Characterization of HBV integration patterns and timing in liver cancer and HBV-infected livers. Oncotarget. 2018;**9**(38):25075-25088. DOI: 10.18632/oncotarget.25308

[32] Kuno A, Ikehara Y, Tanaka Y, Ito K, Matsuda A, Sekiya S, et al. A serum "sweet-doughnut" protein facilitates fibrosis evaluation and therapy assessment in patients with viral hepatitis. Scientific Reports. 2013;**3**:1065. DOI: 10.1038/srep01065

[33] Gane EJ. Future anti-HBV strategies. Liver International. 2017;**37**(Suppl 1):40-44. DOI: 10.1111/liv.13304

[34] Ishida T, Hinuma Y. The origin of Japanese HTLV-I. Nature. 1986;**322**(6079):504. DOI: 10.1038/322504a0

[35] Ruiz-Herguido C, Guiu J, D'Altri T, Inglés-Esteve J, Dzierzak E, Espinosa L, et al. Hematopoietic stem cell development requires transient Wnt/beta-catenin activity. The Journal of Experimental Medicine. 2012;**209**(8):1457-1468. DOI: 10.1084/jem.20120225

[36] Caginalp C, Caginalp G. Opinion: Valuation, liquidity price, and stability of cryptocurrencies. Proceedings of the National Academy of Sciences of the United States of America. 2018;**115**(6):1131-1134. DOI: 10.1073/pnas.1722031115

[37] Kataoka K, Nagata Y, Kitanaka A, Shiraishi Y, Shimamura T, Yasunaga J, et al. Integrated molecular analysis of adult T cell leukemia/lymphoma. Nature Genetics. 2015;**47**(11):1304-1315. DOI: 10.1038/ng.3415

[38] Becherer L, Bakheit M, Frischmann S, Stinco S, Borst N, Zengerle R, et al. Simplified real-time multiplex detection of loop-mediated isothermal amplification using novel mediator displacement probes with universal reporters. Analytical Chemistry. 2018;**90**(7):4741-4748. DOI: 10.1021/acs.analchem.7b05371

[39] Kobayashi S, Tian Y, Ohno N, Yuji K, Ishigaki T, Isobe M, et al. The CD3 versus CD7 plot in multicolor flow cytometry reflects progression of disease stage in patients infected with HTLV-I. PLoS One. 2013;**8**(1):e53728

[40] Shaughnessy JD Jr, Largaespada DA, Tian E, Fletcher CF, Cho BC, Vyas P, et al. Mrvi1, a common MRV integration site in BXH2 myeloid leukemias, encodes a protein with homology to a lymphoid-restricted membrane protein Jaw1. Oncogene. 1999;**18**(12):2069-2084. DOI: 10.1038/sj.onc.1202419

[41] Justice MJ, Morse HC 3rd, Jenkins NA, Copeland NG. Identification of Evi-3, a novel common site of retroviral integration in mouse AKXD B-cell lymphomas. Journal of Virology. 1994;**68**(3):1293-1300

[42] Liao X, Buchberg AM, Jenkins NA, Copeland NG. Evi-5, a common site of retroviral integration in AKXD T-cell lymphomas, maps near Gfi-1 on mouse chromosome 5. Journal of Virology. 1995;**69**(11):7132-7137

[43] Mucenski ML, Taylor BA, Ihle JN, Hartley JW, Morse HC 3rd, Jenkins NA, et al. Identification of a common ecotropic viral integration site, Evi-1, in the DNA of AKXD murine myeloid tumors. Molecular and Cellular Biology. 1988;**8**(1):301-308

[44] Mucenski ML, Taylor BA, Jenkins NA, Copeland NG. AKXD recombinant inbred strains: Models for studying the molecular genetic basis of murine lymphomas. Molecular and Cellular Biology. 1986;**6**(12):4236-4243

[45] Hiai H, Tsuruyama T, Yamada Y. Pre-B lymphomas in SL/Kh mice: A multifactorial disease model. Cancer Science. 2003;**94**(10):847-850

[46] Hiratsuka T, Tsuruyama T, Kaszynski R, Kometani K, Minato N, Nakamura T, et al. Bone marrow pre-B expansion by SL/Kh-Bomb1 locus: Not sufficient for lymphomagenesis. Leukemia Research. 2008;**32**(2):309-314. DOI: 10.1016/j.leukres.2007.05.013

[47] Jin G, Tsuruyama T, Yamada Y, Hiai H. Svi3: A provirus common integration site in c-myc in SL/Kh pre-B lymphomas. Cancer Science. 2003;**94**(9):791-795

[48] Lu LM, Shimada R, Higashi S, Zeng Z, Hiai H. Bone marrow pre-B-1 (Bomb1): A quantitative trait locus inducing bone marrow pre-B-cell expansion in lymphoma-prone SL/Kh mice. Cancer Research. 1999;**59**(11):2593-2595

[49] Okamoto K, Yamada Y, Ogawa MS, Toyokuni S, Nakakuki Y, Ikeda H, et al. Abnormal bone marrow B-cell differentiation in pre-B lymphoma-prone SL/Kh mice. Cancer Research. 1994;**54**(2):399-402

[50] Shimada MO, Yamada Y, Nakakuki Y, Okamoto K, Fukumoto M, Honjo T, et al. SL/KH strain of mice: A model of spontaneous pre-B-lymphomas. Leukemia Research. 1993;**17**(7):573-578

[51] Tsuruyama T, Nakamura T, Jin G, Ozeki M, Yamada Y, Hiai H. Constitutive activation of Stat5a by retrovirus integration in early pre-B lymphomas of SL/Kh strain mice. Proceedings of the National Academy of Sciences of the United States of America. 2002;**99**(12):8253-8258. DOI: 10.1073/pnas.112202899

[52] Yamada Y, Matsushiro H, Ogawa MS, Okamoto K, Nakakuki Y, Toyokuni S, et al. Genetic predisposition to pre-B lymphomas in SL/Kh strain mice. Cancer Research. 1994;**54**(2):403-407

[53] Haran-Ghera N, Peled A, Wu L, Shortman K, Brightman BK, Fan H. The effects of passive antiviral immunotherapy in AKR mice: I. The susceptibility of AKR mice to spontaneous and induced T cell lymphomagenesis. Leukemia. 1995;**9**(7):1199-1206

[54] Ikeda H, Rowe WP, Boyse EA, Stockert E, Sato H, Jacobs S. Relationship of infectious murine leukemia virus and virus-related antigens in genetic crosses between AKR and the Fv-1 compatible strain C57L. The Journal of Experimental Medicine. 1976;**143**(1):32-46

[55] Lenz J, Crowther R, Klimenko S, Haseltine W. Molecular cloning of a highly leukemogenic, ecotropic retrovirus from an AKR mouse. Journal of Virology. 1982;**43**(3):943-951

[56] Akagi K, Suzuki T, Stephens RM, Jenkins NA, Copeland NG. RTCGD: Retroviral tagged cancer gene database. Nucleic Acids Research. 2004;**32**(Database issue):D523-D527. DOI: 10.1093/nar/gkh013

[57] Fitts R, Temin HM. Cellular DNA surrounding integration sites of an avian retrovirus. The Journal of General Virology. 1983;**64**(2):267-274. DOI: 10.1099/0022-1317-64-2-267

[58] Coffin JM. Retroviral DNA integration. Developments in Biological Standardization. 1992;**76**:141-151

[59] Tsuruyama T, Hiratsuka T, Aini W, Nakamura T. STAT5A modulates chemokine receptor CCR6 expression and enhances pre-B cell growth in a CCL20-dependent manner. Journal of Cellular Biochemistry. 2016;**117**(11):2630-2642. DOI: 10.1002/jcb.25558

[60] Tsuruyama T, Imai Y, Takeuchi H, Hiratsuka T, Maruyama Y, Kanaya K, et al. Dual retrovirus integration tagging: Identification of new signaling molecules Fiz1 and Hipk2 that are involved in the IL-7 signaling pathway in B lymphoblastic lymphomas. Journal of Leukocyte Biology. 2010;**88**(1):107-116. DOI: 10.1189/jlb.1109748

[61] Hiratsuka T, Takei Y, Ohmori R, Imai Y, Ozeki M, Tamaki K, et al. ZFP521 contributes to pre-B-cell lymphomagenesis through modulation of the pre-B-cell receptor signaling pathway. Oncogene. 2016;**35**(25):3227-3238. DOI: 10.1038/onc.2015.385

[62] Warming S, Liu P, Suzuki T, Akagi K, Lindtner S, Pavlakis GN, et al. Evi3, a common retroviral integration site in murine B-cell lymphoma, encodes an EBFAZ-related Kruppel-like zinc finger protein. Blood. 2003;**101**(5):1934-1940. DOI: 10.1182/blood-2002-08-2652

[63] Hentges KE, Weiser KC, Schountz T, Woodward LS, Morse HC, Justice MJ. Evi3, a zinc-finger protein related to EBFAZ, regulates EBF activity in B-cell leukemia. Oncogene. 2005;**24**(7):1220-1230. DOI: 10.1038/sj.onc.1208243

[64] Yamasaki N, Miyazaki K, Nagamachi A, Koller R, Oda H, Miyazaki M, et al. Identification of Zfp521/ZNF521 as a cooperative gene for E2A-HLF to develop acute B-lineage leukemia. Oncogene. 2010;**29**(13):1963-1975. DOI: 10.1038/onc.2009.475

[65] Correa D, Hesse E, Seriwatanachai D, Kiviranta R, Saito H, Yamana K, et al. Zfp521 is a target gene and key effector of parathyroid hormone-related peptide signaling in growth plate chondrocytes. Developmental Cell. 2010;**19**(4):533-546. DOI: 10.1016/j.devcel.2010.09.008

[66] Bond HM, Mesuraca M, Carbone E, Bonelli P, Agosti V, Amodio N, et al. Early hematopoietic zinc finger protein (EHZF), the human homolog to mouse Evi3, is highly expressed in primitive human hematopoietic cells. Blood. 2004;**103**(6):2062-2070. DOI: 10.1182/blood-2003-07-2388

[67] Studamire B, Goff SP. Host proteins interacting with the Moloney murine leukemia virus integrase: Multiple transcriptional regulators and chromatin binding factors. Retrovirology. 2008;**5**:48. DOI: 10.1186/1742-4690-5-48

[68] Kvaratskhelia M, Sharma A, Larue RC, Serrao E, Engelman A. Molecular mechanisms of retroviral integration site selection. Nucleic Acids Research. 2014;**42**(16):10209-10225. DOI: 10.1093/nar/gku769

[69] Sharma A, Larue RC, Plumb MR, Malani N, Male F, Slaughter A, et al. BET proteins promote efficient murine leukemia virus integration at transcription start sites. Proceedings of the National Academy of Sciences of the United States of America. 2013;**110**(29):12036-12041. DOI: 10.1073/pnas.1307157110

[70] Dey A, Chitsaz F, Abbasi A, Misteli T, Ozato K. The double

bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis. Proceedings of the National Academy of Sciences of the United States of America. 2003;**100**(15):8758-8763. DOI: 10.1073/pnas.1433065100

[71] Vollmuth F, Blankenfeldt W, Geyer M. Structures of the dual bromodomains of the P-TEFb-activating protein Brd4 at atomic resolution. The Journal of Biological Chemistry. 2009;**284**(52):36547-36556. DOI: 10.1074/jbc.M109.033712

[72] Umehara T, Nakamura Y, Jang MK, Nakano K, Tanaka A, Ozato K, et al. Structural basis for acetylated histone H4 recognition by the human BRD2 bromodomain. The Journal of Biological Chemistry. 2010;**285**(10):7610-7618. DOI: 10.1074/jbc.M109.062422

[73] Zhang W, Prakash C, Sum C, Gong Y, Li Y, Kwok JJ, et al. Bromodomain-containing protein 4 (BRD4) regulates RNA polymerase II serine 2 phosphorylation in human CD4+ T cells. The Journal of Biological Chemistry. 2012;**287**(51):43137-43155. DOI: 10.1074/jbc.M112.413047