

Chapter 2

Retroviral Insertional Mutagenesis in Mouse Models of Leukemia and Lymphoma

David A. Largaespada

Abstract Leukemia and lymphoma are cancers derived from cellular elements of the hematopoietic system. While they make up a minority of human cancer morbidity and mortality, the study of these cancers has illuminated many important aspects of cancer development and biology. In fact, the leukemias and lymphomas are among the best-studied and well understood types of cancer from a genetic perspective. In part, this may derive from the fact that these types of cancer are highly amenable to study using models in which mice are chronically infected with a retrovirus so as to induce or accelerate the disease. In this chapter, I have briefly reviewed the long and rich history of cancer studies using the murine leukemia viruses (MLV). Special attention has been paid to the replication competent MLV that typically cause cancer after a long latency and via insertional mutagenesis. This is followed by a discussion of the limitations of these models and suggestions for future work.

2.1 Introduction to the Murine Leukemia Viruses

The Murine Leukemia Viruses (MLV) are members of a large, naturally occurring group of type C gammaretroviruses that can infect rodents (reviewed in [22, 57]). The MLV were discovered many decades ago by careful observations, indicating a “filterable agent” could induce a malignancy of the lymphatic system in susceptible mice. These experiments were strong early examples that a virus could cause cancer. Like all retroviruses, the MLV carry an RNA genome that is reverse-transcribed into a double stranded DNA copy called a provirus. The integrated provirus is inherited by all daughter cells of the infected cell after cell division. The provirus serves as the template for the production of the messenger RNAs (mRNAs), which encode

D.A. Largaespada (✉)

The Department of Genetics, Cell Biology and Development, The Department of Pediatrics, Masonic Cancer Center, The Center for Genome Engineering, The University of Minnesota, Minneapolis, MN 55455, USA
e-mail: larga002@umn.edu

the proteins required for virion production, and the viral RNA (vRNA) that will be packaged into new viral particles produced by the infected cell. The MLV are typically non-cytopathic viruses that do not result in lysis of virus producing cells. This fact, combined with the effects of integration of the provirus, described in more detail below, create the conditions that can lead to cancer development in infected animals.

The MLV can be divided into subgroups. The reader is referred to other comprehensive reviews on the structure, genetics, and distribution of the MLV, and other retroviruses [22]. Suffice it to say that early research showed MLV could be detected by their ability to induce syncytia formation in a cell line called XC, or focus formation in a cell line called SC [22]. Super-infection interference is a phenomenon that was observed when different isolates of MLV were used to infect the same cells in series. The second infection was blocked because of receptor occupancy by retroviral envelope protein produced by the first infection. This phenomenon, and the cell type tropism of the MLV, was used to show that these viruses could be placed into groups based on common usage of a cell surface receptor. The xenotropic MLV can infect non-rodent, but not rodent cells, despite the fact they exist in the genome of some mice. The ecotropic MLV can infect rodent cells, but not most other cell types including primate cells. The amphotropic MLV infect both rodent and primate cells. In most leukemia or lymphoma induction experiments done in laboratory mice, the ecotropic MLV are used. However, some of these studies have used amphotropic MLV, particularly in studies on acute myeloid leukemia (AML) induction.

All three types of MLV share a similar overall genome organization, but they have different envelope genes (*env*) that encode proteins, which use different cell surface receptors to gain entry into cells. The MLV are enveloped viruses, meaning they are surrounded by a lipid membrane bilayer derived from the infected cell that produced the virion and studded with *env* protein required to infect target cells. The *env* gene is expressed from a spliced mRNA, while the two other MLV genes—*gag* and *pol*—are expressed from the same full-length RNA that also serves as the genomic viral RNA (vRNA) destined to be packaged into new viral particles. Splicing between a splice donor (SD) and splice acceptor (SA) occurs with an appropriate frequency to generate *env* protein for virion production, while removing a special sequence called ψ , required to package the full-length vRNA into new virions. The *gag* gene encodes nucleocapsid proteins, while the *pol* gene encodes the reverse transcriptase and integrase enzymes required to produce the proviral DNA and integrate it into the genome of an infected host cell after viral infection.

The process of reverse transcription and integration of the proviral DNA is a complex process reviewed elsewhere [74]. However, the integrated provirus consists of the *gag*, *pol*, and *env* genes flanked by long terminal repeats (LTRs), which are sub-divided into the unique 3' (U3), repeat (R) and unique 5' (U5) regions. The U3 contains the enhancer and promoter sequences required to initiate transcription of the vRNA and spliced mRNA. The U5 region contains a polyadenylation signal. The U3 and U5 sequences, as well as the SD and SA sequences, are important in the process of insertional mutagenesis because they can alter the expression and processing of endogenous RNAs.

This review focuses on pressing issues related to the identification and study of leukemia and lymphoma genes identified via insertional mutagenesis by MLV. However, it is worth pointing out that replication-defective, partially deleted forms of the MLV, and other retroviruses, have been discovered that carry processed forms of endogenous proto-oncogenes. These viral oncogenes are found in acute transforming retroviruses (reviewed in [40, 47]). The study of these acute transforming retroviruses was pivotal in cancer research because they helped prove that cancer has a genetic origin and helped in the identification of many important targets of mutation in human cancer, such as the *RAS* oncogenes and *MYC*. The acute transforming retroviruses can be distinguished from the replication competent, slow transforming retroviruses, including many MLV, that are the subject of this review.

2.2 Viral and Host Determinants of Disease

Much prior research on the MLV was done to discover and map viral and host determinants of disease induction by the MLV. These studies are less often performed now, but were very critical from a practical standpoint because they taught lessons about the mechanisms by which the MLV cause cancer in the first place. It is worth pointing out that studies have allowed scientists to create new chimeric MLV useful for studying particular types of leukemia (for example [77]).

The murine host strain can influence disease by restricting infection or replication of the MLV. The best studied example is the *Fv1* gene, which exists in laboratory strains of mice in at least two different forms: *Fv1^b*, which restricts infection by b-ecotropic MLV; and *Fv1ⁿ*, which restricts infection by n-tropic MLV [3]. Thus, most laboratory strains of mice carry one or the other allele and can be infected by only one or the other type of ecotropic MLV. Beyond the effects of *Fv1* and other restriction factors, the host strain background interacts with the strain of MLV to dictate the ultimate disease course. Thus, the same MLV can induce different diseases in the different inbred strains of mice. These differences could reflect inherent differences in susceptibility to various forms of leukemia in different inbred strains of mice. For example, a typical b-ecotropic MLV causes acute myeloid leukemia in BXH-2 strain mice [3, 25]. The susceptibility of BXH-2 strain to AML induction in this model, rather than other forms of leukemia, has been tentatively traced to a unique BXH-2-specific germline mutation in the *Icsbp1* gene in this strain [62].

One often ignored host strain determinant of leukemia induction by MLV is the development of other replication-competent viruses in the infected host via recombination. It is common for mice that are chronically infected with an ecotropic MLV to develop recombined MLV called mink cell focus forming (MCF) viruses [57]. The MCF are generated by recombination events between ecotropic MLV and vRNAs produced by endogenous MLV which lead to replacement of env gene and LTR sequences. The MCF utilize a different cell surface receptor for infection than do ecotropic MLV and can escape superinfection interference by ecotropic MLV. The MCF may drive leukemogenesis by allowing more proviral insertion mutations

to accumulate in pre-neoplastic clones. The details of the strain-specific effects of some MLV are in general not known and useful models of MLV-driven leukemia are generally found, or tested empirically, rather than being deliberately created.

Viral determinants of disease specificity have also been discovered. Often this has been done by switching sequences and genes within one MLV, with those of another MLV to determine what sequences determine disease specificity. The determinants of leukemia specificity have often been shown to map to the LTRs [51] and within an LTR, to the enhancer sequences in the U3 region [77]. These studies have suggested that the ability of some MLV to specifically induce one type of leukemia, and not another, may be due to the activity of the LTR. If an LTR promoter were especially active in one cell type versus another, it is possible that viral spread within that cell type would be greater, and that activating mutations would be more likely. The MLV LTR is a major driver of gene activation in leukemia induction studies because it can enhance transcription from endogenous promoters and drive chimeric RNAs composed of viral and endogenous mRNA sequences. These events are common in the activation of many endogenous genes by MLV insertion. It should be noted that some mechanisms of MLV mutagenesis do not seem to require LTR activity. In some cases, MLV insertions seems to alter the splicing and processing of an endogenous mRNA encoded by genes that have suffered proviral insertions. These events can decrease the expression of a tumor suppressor gene [35], or stabilize an oncogene mRNA by replacing 3' UTR sequences that contains destabilizing motifs [67, 69]. The viral determinants of MLV leukemia induction are in general not well enough known to predict *a priori* what sequence alterations to the viral sequence should be made to produce specific types of leukemia, let alone specific types of cancer in mice. Since MLV, and other gamma retroviruses, can only efficiently infect cells undergoing mitosis [43], there must be inherent limitations in the tissue tropism that can be built into the MLV. Nevertheless, it would be highly desirable to generate new recombinant MLV with useful properties, such as the ability to induce each of the forms of leukemia and lymphoma that characterize human disease. For example, among the eight or more French-American-British (FAB) subtypes of AML known to occur in humans, not all can be induced in mice. It would be useful to have MLV-induced models of M6 and M7 AML, but the alterations that might allow this are not easy to predict. Despite these issues, there have been engineered MLV that can induce a spectrum of leukemia not usually seen with the parental viruses and which have other useful properties.

The MOL4070LTR virus is a chimeric MLV containing U3 LTR sequences from the amphotropic MLV called 4070A and other sequences from the Moloney-MLV [77]. The 4070A virus has been shown to induce AML in some strains of mice [76]. However, this virus is restricted in the number of common laboratory strains in which it can be used, as 4070A is an n-tropic MLV. Moloney-MLV is an nb-tropic MLV and so can infect and replicate in mice that carry either the *FvIⁿ* or *FvI^b* alleles common to various laboratory strains [77]. However, Moloney-MLV induces almost exclusively T cell leukemias in mice [20]. The MOL4070LTR virus has the nb-tropism of Moloney MLV, which is determined in the gag gene sequences [54], and the ability to induce AML in some strains of mice, from the 4070A

parent virus, a trait apparently largely influenced by the U3 sequences [77]. This virus has been shown to induce AML and ALL in various laboratory strains of mice, such as FVB/n and 129/SvJ, and in various leukemia-susceptible transgenic mice [6]. The MOL4070LTR virus thus seems to be generally useful for finding cooperating mutations in AML and ALL using strains commonly used to make transgenic mice. For example, this virus was used to find gene mutations that could cooperate with expression of the *NUP98-HOXD13* oncogene [59]. We have used the MOL4070LTR virus to find gene mutations that could cooperate with expression of the *MLL-AF9* oncogene in AML and ALL development (unpublished observations).

Others have replaced the enhancers within an MLV to alter its disease specificity. The PyF101 + Mo M-MuLV consists of a Moloney-MLV in which the viral enhancers in the U3 region are replaced with those from the SV40 early promoter [21]. This virus induces myeloid and T lymphoid leukemia in mice. These kinds of experiments are perhaps hampered by our inability to fully predict what kind of disease will result, and from the fact that the alterations that are made could impair the ability of the virus to replicate. Despite these concerns, the example provided by the PyF101 + Mo M-MuLV suggest that perhaps hematopoietic lineage-specific enhancers could be included in the U3 region of a recombinant MLV to achieve specific leukemia phenotypes in mice. Transposable elements, such as *Sleeping Beauty*, provide an alternative to this approach [60]. These are described in a different chapter of this book.

2.3 Challenges in MLV Mutagenesis Studies

There are many challenges in MLV mutagenesis studies. Among the first to consider, when starting a project, is the number of independent leukemia samples that should be generated for study. As has been observed before, “quantity has a quality all its own.” Unfortunately, experience has shown that more statistically significant CIS and associated genes are discovered as more leukemias are studied, even when one approaches and surpasses 100 individual leukemias. The large number of leukemias required for identification of rare CIS—involved in <10% of cases, can be cost prohibitive to generate. Our usual goal is to obtain 60+ mice of each experimental group expected to get leukemia. This number is based on having the precision to detect any common insertion site present in 10% of the tumors balanced against the cost and expense of generating and aging these cohorts. If we assume that each mouse will develop a malignancy with on average one cloned insertion per tumor the R statistical package (<http://www.r-project.org>) calculates the binomial probability of missing a common insertion site, with a 10% true insertion frequency, as 1.38% at this sample size. This is a conservative estimate, as usually many more than one insertion per leukemia is recovered. As we have seen, not only are many mice needed, but slow transforming retroviruses typically induce leukemia after 5

or more months. It can often cost \$20-30,000 just in mouse animal housing charges for a project like this one—assuming it lasts 1 year.

One other complication of studies of this sort has to do with recovery of proviral insertion sites. Many different approaches to accomplish this have been described in the literature in the past. Methods based on the use of primers specific to the MLV and which use the polymerase chain reaction (PCR) are currently preferred [37]. This issue is complicated by the fact that any MLV induced leukemia is actually a collection of some number of related, but genetically distinct clones. Some insertion mutations are present in only a subset of leukemia cells. The significance of these sub-clonal insertion mutations is not entirely clear. Work by investigators, such as Dr. Philip Tsichlis, clearly shows that new subclones with new growth-promoting proviral insertion mutations can be identified in MLV induced leukemias as they are passaged in vivo or in vitro [55]. Recently, ligation-mediated PCR (LM-PCR) based methods for proviral or transposon insertion site recovery have been linked to “single molecule” or “pyrosequencing” approaches, such as that made possible by the Roche GS-FLX machine [60, 64]. Thus, hundreds of thousands of amplified provirus/cellular genomic DNA junction PCR products can be sequenced, from 100 or more leukemias, en masse, without the need for cloning them beforehand. By using bar-coded PCR primers for the secondary PCR in the LM-PCR reaction, it is possible to determine, after sequence analysis, which insertions were derived from which individual leukemia. A benefit of this approach is that the number of times any one specific insertion sequence is read may be related to the clonal abundance of that insertion mutation in the leukemia. So, two sorts of data are obtained: the identity of all insertions and their relative abundance in the tumor clone. However, for these data to be most meaningful, it is most desirable for the recovery of insertion sites to be saturating. Therefore, we recommend that insertions are amplified and sequenced from both ends of the integrated provirus and that multiple restriction enzymes be used to generate the junction fragments that will be amplified. We refer the reader to several methods articles for details on this kind of approach [34, 64, 80]. A few complicating issues bear mentioning, however. It is important to keep in mind that when amplifying and sequencing proviral insertion sites the fact that they are flanked by LTRs presents a special problem. One must introduce a restriction enzyme digestion step to avoid recovery of internal proviral sequences downstream of the 5' LTR, when amplifying “right-hand” junction fragments downstream of the 3' LTR. Similarly, a restriction enzyme digestion step is introduced to avoid recovery of internal proviral sequences upstream of the 3' LTR, when amplifying “left-hand” junction fragments upstream of the 5' LTR. We have developed methods for proviral insertion site amplification using LM-PCR that rely on the use of blunt-ended restriction enzymes, or sticky-ended restriction enzyme digestion followed by filling in the overhangs with *Taq* polymerase, thus generating products with a single adenosine (A) overhang [80]. In each case, after generation of junction fragments, they are then ligated to linkers and amplified using two rounds of PCR. In nearly all of the techniques used to generate amplified LM-PCR products the linkers are double-stranded and some provision is made to prevent PCR amplification of “linker to linker” fragments and allow “linker to proviral LTR” PCR amplification. This

is accomplished by using as a linker primer, a sequence that has no homologous sequence to anneal to until and unless the LTR-specific primer has annealed to a template junction fragment allowing the PCR polymerase to generate a product that extends into the linker and generate sequence complementary to the linker-specific primer. This can be done using two primers of unequal length to create a partially double-stranded linker, in which the 3' end of the shorter primer is blocked to prevent the polymerase from using it to create a fully double-stranded primer during the PCR reaction, which would cause the generation of many contaminating "linker to linker" PCR products.

It is possible that if one shears leukemia genomic DNA, "blunt ends" all the fragments using a polymerase, and then does LM-PCR using blunt-ended primers, it may be possible to recover more insertion sites than using restriction enzymes to generate fragments. Innovations such as this one will allow greater recovery of all MLV proviral insertions that are present. In the future, it seems possible that MLV mutagenesis studies would benefit from whole genome re-sequencing, once sequencing technology becomes cheap enough to perform on a large number of samples. Such data may reveal cooperating mutations not induced by MLV insertions, but by other mechanisms such as translocations, point mutations or deletions. In any case, very robust bioinformatics analysis of proviral insertion site sequencing results is required before determining what significant, recurrently occurring, insertion site mutations have been discovered. It is important to remember that LM-PCR may generate products from endogenous retroviruses. In some cases, these are well known retroviruses as in the endogenous b-tropic MLV in BXH-2 strain mice [80]. In other cases, they can be recognized because identical insertion sites are amplified from multiple independent mice. In addition, the sequence analysis algorithm must be designed to recognize bona fide MLV insertion sites by recognition of LTR sequences downstream of the second LTR-specific primer, followed by genomic sequence mapping with high confidence to one unique position in the mouse genome. Once a non-redundant set of insertions is defined, the definition of common sites of insertion (CIS) follows.

Several methods for determining which genomic loci have suffered a statistically significant number of independent insertion mutations have been published [13, 78]. However, no clear consensus exists for the correct methods. It was assumed in the past that MLV integration was a random process, not influenced by the genomic architecture. This meant that Monte Carlo simulations could be performed using a number of randomly selected genomic positions and calculating the expected fraction of random or background CIS detected in different window sizes and these were used to select cut-offs for significant CIS observed in actual data [39]. Other approaches calculated CIS based on the assumption of a Poisson distribution of insertion mutations in the genome [42]. However, more recent analyses show that MLV insertion is not random and is heavily influenced by the presence of genes, with insertions near the start site of gene transcription being preferred. Thus, a null set of MLV insertions has been obtained which provides a different estimate of statistically significant CIS [78]. It has been estimated that up to 2/3rds of the CIS defined in the Retroviral Tagged Cancer Genome Database [1] can be explained

by the null hypothesis, that is, the lack of any selective pressure for their mutation by MLV insertion [78]. It was recommended that CIS defined by four or fewer insertion mutations be considered cautiously [78]. The problem of background false positive CIS is compounded by the large data sets obtained using LM-PCR and high throughput sequencing—resulting in definitions of CIS requiring large numbers of insertions within small regions of the genome. Thus, CIS in tumor suppressor genes that are very large and where insertions in various parts of the gene could disrupt them might be missed. Similarly, if MLV can effect the transcription of a gene at a long distance, CIS might be missed in a large dataset of insertions because they are spread out over a large area. To correct for these problems, Dr. Wessels and colleagues in the Netherlands have defined CIS using an entirely different and sophisticated method [13]. In this method, an expected distribution, called a kernel density estimate, of insertion mutation density was defined across the entire genome based on permutations of all the data and which can be modified based on assumptions about any non-biased insertion preferences of MLV, such as the tendency to insertion near transcription start sites of genes. This approach allows one to discover CIS defined by a peak of insertions that exceed an alpha value for significance without suffering from high false positive rates due to large numbers of insertions. Since the method is scalable to different intervals, significant CIS defining small or larger intervals can be defined, which is important since it is possible that biologically significant intervals vary depending on the genomic context of a given region. For example, CIS could be defined in large tumor suppressor genes using this method that might be missed if CIS over very large intervals are not sought.

These same researchers have gone on to apply this method for detection of statistically significant pairs of common co-occurrence of insertions (CCIs), which could define pairs of cooperating insertion mutations [12]. This method uses a two-dimensional version of the kernel convolution density estimate method described above. These authors defined 86 significant CCIs using published data, among which were previously noted collaborating oncogenes discovered using MLV mutagenesis, such as *Hoxa9* with *Meis1* and *Myc* with *Pim1* [12]. This is among the most exciting attributes of MLV mutagenesis studies, that is, the ability to detect networks of cooperating oncogenes. Such cooperating oncogenes could reveal targets for therapy. For example, the cooperative relationship between *Pim1* and *Myc* could suggest that transformation by *Myc* oncogenes requires Pim1 kinase activity, a prediction that was recently demonstrated experimentally [82]. Therefore, Pim1 kinase inhibition may be of therapeutic value in the many forms of human cancer that have aberrant *Myc* activity. Attempts to define CCI are just the beginning of more sophisticated network analyses that should be possible using data from MLV mutagenesis studies.

It seems likely that a network of CCI relationships could be assembled by linking all such associations detected in an MLV mutagenesis screen. Data published by de Ridder et al. already establishes that *Myc* exists at a “node” having cooperative relationships with several other oncogenes [12]. By linking all CCIs into a network one might define complexes or signaling pathways. These authors have also suggested that some genes associated with insertion mutations could be clustered into

functional classes or pathways prior to analysis to determine if a gene interacts with any of a member of a functional class or pathway. They detected a potential interaction between *Sox4* and any of a number of cyclin dependent kinase genes in this way [12]. In the future, it would also be highly desirable to link these genetic data with specific phenotypes such as the subtype of leukemia (e.g. lymphoid versus myeloid leukemia), disease latency, and gene expression profiling.

We have started to develop methods for analysis of CIS-associated genes and leukemia phenotypes. We find that reiterating Fisher's Exact test can be used to associate insertion mutations at specific loci and phenotypes such as surface immunophenotype. Other analytical approaches should be developed to visualize and analyze a systems network for MLV mutagenesis experiments. One complication of this kind of analysis, and CCI analyses, is that we usually assume that any amplified and sequenced insertion mutation is present in all or nearly all cells of the tumor mass. Thus, we assume that any two insertion mutations are present in the same cell and could potentially be cooperating. Similarly, we assume that if a leukemia clone expresses myeloid surface markers, those cells also harbor a particular insertion mutation that was amplified and sequenced from this sample. However, the insertion mutation in question might have been present only in a rare subclone that expresses lymphoid markers and is distantly or unrelated to the majority clone. One method for getting around this problem is to introduce a specific mutation into the germline of mice, as a transgene, and then using MLV mutagenesis to accelerate disease. Then one can potentially identify insertion mutations more prevalent in the transgenic cohort of leukemias compared to non-transgenic control leukemia populations. Many such "sensitized" MLV mutagenesis screens have been performed and are the most common type underway today.

Sensitized MLV mutagenesis screens were pioneered by Dr. Anton Berns at the Netherlands Cancer Institute. He and his collaborators noticed that among T cell leukemias induced by Moloney MLV certain insertion mutations tended to be present within the same clones [4]. Notably, an association was seen between activation of the *Cmyc* gene and the *Pim1* kinase. Indeed, when MLV was used to accelerate T cell leukemia in *Cmyc* overexpressing transgenic mice, the *Pim1* gene was a frequent target of MLV activation [68]. In a brilliant experiment, the MLV infection was used to accelerate disease in *Cmyc* transgenic, *Pim1*^{-/-} mice and the resultant leukemias now had MLV insertion mutations in the related *Pim2* gene [66]. This experiment revealed that it should be possible to use MLV mutagenesis to isolate genes that play a role downstream of another oncogene, or which can act in a parallel pathway, much the way in which *Drosophila* genetics has been used to define multiple members of a pathway important for some developmental process.

More recent sensitized MLV mutagenesis screens have utilized genetic backgrounds of mice carrying alterations similar to known human leukemia-associated mutations, such as tumor suppressor gene loss or expression of specific fusion oncogene products. Several recurrent chromosomal translocations create fusion oncoprotein genes in human myeloid leukemia. Some of these have been engineered into mice using homologous recombination in mouse embryonic stem cells. Perhaps not surprisingly, most of these fusion oncoproteins by themselves cannot produce rapid,

highly penetrant leukemia in mice. Therefore, investigators have used MLV mutagenesis in attempts to define genes and pathways that can cooperate with expression of these fusion genes. We have investigated genetic events that can cooperate with loss of the *Nf1* tumor suppressor gene in AML development using MLV mutagenesis. The *Myb* and *Bcl11a* genes, among others, were found to cooperate with loss of the *Nf1* gene [79]. These experiments were done by backcrossing an *Nf1* loss of function allele to BXH-2 strain mice, which develop spontaneous AML due to chronic infection with a typical MLV that is passed from mother to offspring [33]. This is a laborious process, and so alternative MLV have been sought for use in strains of mice that are typically used to generate transgenic mice such as the FVB/n and 129/sv strains. For example, the amphotropic MLV, 4070A, has been used to accelerate AML in mice expressing a *CBFbeta-MYH11* transgene, which mimics the effect of the *inv(16)* event in human AML [6]. Activation of the *Plg1* and *Plg2* genes were identified as cooperating genetic events, shown to be overexpressed in some human AML, and capable of cooperating with *CBFbeta-MYH11* in a mouse model [32]. The 4070A virus is unfortunately n-tropic and so not useful in the many laboratory strains of mice that carry the *Fv1^b* allele, such as C57BL/6 [76]. In addition, 4070A can infect human cells and so safety issues are raised. Dr. Linda Wolff has created a chimeric MLV retaining beneficial properties of Moloney and 4070A MLV [77]. This virus, called MOL4070LTR, can induce AML, like 4070A, or T cell ALL and is nb-tropic, like Moloney MLV, and can therefore be used in most laboratory strains of mice. We have used MOL4070LTR to investigate genetic events that cooperate with a prototypical fusion oncogene involving the *MLL* gene, called *MLL-AF9*. The *MLL* gene encodes a Trithorax-related chromatin protein that seems to be required for expression of specific target genes such as some of the *HOX* genes [9]. While some of the fusion oncogenes encode chimeric transcription factors whose activity may be hard to target pharmacologically, it is possible that MLV mutagenesis will uncover critical cooperating genetic events that can be targeted pharmacologically. The same could be said for loss of tumor suppressor genes.

Recently, Berns and colleagues used MLV mutagenesis to investigate genetic events that could cooperate with loss of *p53* or *p19^{Arf}* function in T cell leukemia genesis [63]. This work is notable for several reasons. The first being that *p53* and *p19Arf* had often been assumed to suppress tumor formation by acting in the same pathway only, and so one might expect identical cooperating events would be identified. However, different cooperating genetic events were found in the two groups. The work also demonstrated the power of using very large numbers of leukemia specimens, several hundred, and using high throughput amplification and sequencing methods to identify MLV insertion mutations. Certainly, greater numbers of genes can be implicated on cancer development by doing so.

Other desirable sensitized, or context-specific, genetic screens can be envisioned using MLV mutagenesis. Of course, the issue of multi-clonality described above complicates all studies in which the tumor genotype (in this case specific insertion mutations) is to be associated with specific phenotypes (e.g., gene expression patterns, surface marker expression) or behaviors (e.g., response to chemotherapy). However, this issue can be ameliorated in part by the study of very large data sets.

Nevertheless, it would be ideal to unequivocally separate all genetically distinct tumor clones prior to any type of analysis. This issue was reviewed recently by Kool and Berns [30]. Methods to do this could involve creating permanent cell lines from MLV-induced leukemias, passaging limited numbers of leukemia cells in mice, or developing methods for analysis of insertions and gene expression patterns from one or very few cells in a section of an MLV-induced leukemia. None of these is a perfect solution to the problem, but some attempt should be made to more reliably allow one to correlate leukemia phenotype with insertion mutation genotype. MLV mutagenesis studies have been most often done in order to discover new leukemia genes and pathways. In the future, it would be desirable to be able to associate specific insertion mutations with specific clinically relevant traits, such as chemotherapy resistance, disease aggressiveness or ability to synergize with specific germline mutations or exposure histories. Several examples of work such as this are described below.

Using MLV-accelerated, *Nf1*-deficient, primary AML samples, Lauchle et al., were able to select *in vivo* for AML cells resistant to a MEK inhibitor [36]. Interestingly, the authors were able to find MEK inhibitor resistant forms of the primary AMLs that harbored insertions affecting *Rasgrp1*, *Rasgrp4* and *Mapk14*. The authors found that expression of the *Rasgrp* genes at high levels, or downregulation of *Mapk14*, could induce MEK inhibitor resistance. Another group used MLV mutagenesis to select for lymphoma cells with enhanced ability to invade and crawl underneath a monolayer of fibroblasts in culture [23]. This led to the identification of the *Tiam1* gene, encoding a guanine nucleotide exchange factor (GEF) for Rac GTPases, helping to establish a role for Rac proteins in cell migration and tumor dispersal. Many similar studies should be undertaken. In general, models of MLV-induced leukemia and lymphoma have been underutilized for treatment studies using conventional chemotherapy or molecularly targeted therapies. Indeed, mouse modeling of human cancer has been mostly limited to the study of cancer as it might first present, but not as it exists in patients after responding to chemotherapy and/or radiation. It will be very interesting to examine the evolution of a transformed clone in response to selective pressure applied by therapeutic agents. Such studies could allow the identification of resistance mechanisms and/or define the genetic subtypes that will be most sensitive to a given treatment.

2.4 Comparative Oncogenomics

Once the screen for cancer-related genes has been undertaken using MLV mutagenesis and significant CIS-associated genes have been identified, many challenges remain. Among the most challenging is interpreting the results of viral insertional mutagenesis screens to determine what the results have to say about the process of cancer development in humans. Insertional mutagenesis screens have led to the identification of genes that play important biological roles in the normal development of target tissues. However, the prime motivation of these studies is the identification

of novel human cancer genes and pathways. A second motivation, only recently studied on a large scale, is the identification of specific patterns of concurrent gene mutation. It can be challenging to achieve these goals because the corresponding tumor type which should be examined in human patients is not always clear, based only on the MLV-induced leukemias. Also, the MLV leukemias may not recapitulate certain types of mutagenic alterations that are common in human leukemia. For example, the *Ras* genes are not common targets of MLV insertional mutagenesis, but are frequently mutated by activating point mutations in specific codons in human leukemias. It is possible that *Ras* genes are not identified at CIS because *Ras* gene overexpression is not sufficient to induce a strong enough transforming signal to be selected for in hematopoietic cells. However, the guanine nucleotide exchange factors, such as *Rasgrp1*, are frequently targeted at CIS [16, 28, 37]. These GEFs activate Ras proteins downstream and can transform cells simply by overexpression [16, 29, 53, 56]. So, in addition to considering the specific genes associated with CIS, one must consider the pathways revealed by MLV insertional mutagenesis screens. Nevertheless, certain examples in which CIS-associated target genes play a direct role in human leukemia as targets of mutation deserve mention.

Several important transcription factors involved in leukemia development have been identified at common sites of MLV integration. The *Myc* gene itself was identified both as an acute transforming viral oncogene in an acute transforming retrovirus from chicken and murine retroviruses and as a recurrent target of MLV integration [70]. The *Myb* gene is a well known target of insertional mutagenesis in myeloid and lymphoid leukemia [28, 44, 76]. This transcription factor is required for normal hematopoiesis [45], and was long thought to not play a direct role in human leukemia. However, recent studies show that these conclusions were premature as *MYB* is often amplified and overexpressed in human leukemia [8, 31, 46]. These studies show that even well known CIS associated genes should be carefully examined for their potential role in human leukemia, in light of new genome wide approaches for characterizing cancer genomes. A large public catalog of CIS-associated genes identified by MLV mutagenesis studies is maintained by Dr. Keiko Agaki, now at the Ohio State University, which lists over 300 CIS genes [1]. Although recent analyses suggest that many of these may not be statistically significant [13, 78] given the fact that MLV do not actually randomly integrate into the genome, this list is still highly likely to contain many loci directly involved in human leukemia via the acquisition of somatically acquired genetic or epigenetic changes. For example, studies on the *Evi2* CIS helped identify the human gene for Neurofibromatosis Type 1 (NF1), a tumor suppressor gene that causes cancer susceptibility when inherited in a mutant form [52]. The fact that human *HOX* genes can be direct leukemia genes was made possible by studies of BXH-2 strain MLV-induced AMLs [49, 50]. The *HOX* co-factor *Meis1* was discovered as a *Hoxa9* cooperating AML oncogene in studies of the same strain. *Hoxa* and *Meis1* gene activity characterizes the large subset of human leukemias with *MLL* gene rearrangements and the activity of these proteins seems to play a major role in transformation by *MLL* fusion oncoproteins [2]. This finding was an important

outcome of the MLV mutagenesis studies done in BXH-2 strain mice and subsequent validation studies.

Evidence that MLV screens can identify relevant pathways for human leukemia has been obtained by large-scale comparative studies using gene expression profiling of human leukemia genes also. Scientists at Erasmus University in Rotterdam, The Netherlands, have compared CIS-associated genes with gene expression profiling from human AML [18]. Their results show that genes near CIS from the RTCDG database are very significantly enriched for those that are differentially expressed among human AML subtypes, a feature not seen for genes farther away from CIS. Therefore, the expression of these genes may underlie specific programs of leukemic transformation, perhaps downstream of certain well-known fusion oncoproteins. Their level of expression could also be determinants of the varying clinical behavior of different cases of leukemia.

Much attention has recently been paid to the role of microRNA genes in human cancer development. Not surprisingly, MLV mutagenesis has recently been shown to activate miRNA genes in leukemia also [10, 59, 71]. It seems possible that MLV mutagenesis could also pinpoint other types of encoded genomic information that, when altered, could cause cancer. Such encoded information could include other non-protein encoding RNA species or CIS regulatory elements. As more is learned about the architecture of the genome, the results of CIS mutagenesis projects will need to be re-evaluated.

2.5 Validation of Candidates

Perhaps one of the most difficult challenges faced by an investigator once a large number of CIS-associated genes has been identified in an MLV mutagenesis study is to decide which of these genes deserve further study. Two central questions come to mind: Which of these genes can play a role in cellular transformation when altered? And which of the genes, or pathways they act in, are directly involved in human cancer development? The advent of human whole cancer genome re-sequencing, methylation arrays, and extensive mRNA and miRNA expression profiling make it possible to address the second question. However, the answer to the first question remains a challenge for the field of cancer genomics in general.

Despite the seemingly daunting task of studying the function of hundreds of genes in cultured cells, living mice, or other model organisms such as the zebrafish, there are now examples that we can point to as evidence of the practicality of such approaches. One method that is finding widespread utility in the study of cancer genes in maintenance of transformation is based on the use of RNAi [11, 17, 73]. The set of CIS-associated genes identified in a screen can be tested for their role in tumor cell maintenance using cultured leukemia cell lines. However, it is likely that some genes recovered in MLV screens are tumor suppressor genes and would thus not score in such an assay. So for this reason, the validation of some genes may require that they be overexpressed from cDNA expression vectors in cell lines that

are partially, but not fully, transformed. While this is conceptually straightforward, the choice of such a cell line is not a simple matter. However, large scale projects to produce useful cDNA vectors for all genes in the human and mouse genomes are underway [48]. The projects will of course be very valuable assets in the goal of validating cancer genes and placing them in known signaling pathways. In any case, a role in transformation is not the same as a role in leukemia cell maintenance. Moreover, the effects of knocking down expression of any one gene may have a context dependent effect, and so the phenotype under examination may only be revealed in cell lines that provide the appropriate context. Thus it would perhaps be best to use cell lines derived from the MLV model from which the candidate gene was identified. Finally, some oncogenes probably make a contribution to leukemia development that is not easily revealed in a cell culture assay designed to measure a decrease in cell proliferation or viability. Such phenotypes could have to do with self-renewal, differentiation, interaction with non-tumor cells (i.e., other cells of the immune system), or the response to endogenous growth factors which may not be present in the cell culture system. For this reason, the closer the assay is to true leukemia formation *in vivo*, the better.

The well known bone marrow transduction and transplantation (BMTT) assay would seem to be ideally suited to validating candidate oncogenes from MLV mutagenesis screens and indeed the assay has been proposed for this use [14], and used by many for such a purpose [79]. However, the assay is technically demanding and the generation of 100 or more high titer MLV based retroviral pools for a project like this is not a simple matter. Even so, Dr. Scott Lowe's lab has succeeded in using pools of large numbers of retroviral shRNA vectors to screen for genes that, when "knocked down" in p53-deficient hepatoblasts, can induce hepatocellular carcinoma in mice [81]. This and other projects make it clear that retroviral transduction studies are a useful method for validating cancer genes *in vivo* in mice.

Other approaches to validate candidate cancer genes in transgenic mice may be more flexible and powerful. One creative method, developed by Dr. Pentao Liu, involves cloning candidate oncogenes into transposon vectors in which the cDNA can only be expressed after mobilization into an actively transcribed gene [61]. By creating mice carrying many such vectors with a variety of oncogenes and mobilizing them tissue-specifically to induce tumors, one can determine which have been activated by insertion into an active gene. This method has the advantage of allowing sets of candidate oncogenes to be simultaneously tested in a variety of tissues using a Cre/LoxP-regulated transposase transgene and allows for the possibility that only a certain combination of oncogenes might collaborate to induce cancer. While clever, this approach still requires the production of gene-trap cDNA vectors and a transgenic line of mice. It is also still unclear how many different oncogene candidates would be practical to include in a single transgenic mouse strain. In any case, this technology is emblematic of the kind of new thinking that should be undertaken in projects aimed at *in vivo* functional validation of oncogene candidates.

One approach is to use transposon-based methods of gene delivery, which use transfected plasmids, rather than retroviral vectors. This approach has the advantage that one need only produce plasmid expression vectors for each cDNA or shRNA

to be tested. We, and others, have shown that the *Sleeping Beauty* transposon system can be used to deliver of cDNAs and shRNAs to cause tumors in mice when delivered along with a source of the *Sleeping Beauty* transposase [5, 27, 72]. While these were solid tumors, recent data shows that it is possible to use *Sleeping Beauty* to efficiently deliver expressed cDNAs to early hematopoietic precursors via electroporation [41]. This approach has strong potential because cDNA delivery via transposition also allows efficient co-delivery of multiple vectors to the same cell in vivo, as has been demonstrated for glioma induced in mice [75].

2.6 Do Murine Retroviruses Cause Human Cancer?

Recent studies have indicated that a class of xenotropic murine retroviruses (XMRV) are present in some human tissue samples [15, 65]. This virus is closely related to the MLV. XMRV has been linked to both prostate cancer and chronic fatigue development [38]. Interestingly, some reports suggest that XMRV infection is present in prostate tumors, but only common in prostate tumors from patients with a polymorphism in the *Rnase1* gene, which seems to render them susceptible to viral infections [65]. Controversy surrounds the question of how XMRV may contribute to prostate cancer development. Some reports show that XMRV sequences are not common in human prostate cancer samples [24] or lymphocytes from chronic fatigue syndrome patients [19]. Other reports suggest that the epithelial tumor cells themselves are infected with XMRV and that among the insertion sites are genes with roles in growth regulation [58]. It is therefore unclear whether XMRV infection might contribute to prostate cancer development by an insertional mutagenesis mechanism. If XMRV can contribute to human disease, the implications are profound and the lessons learned from the study of MLV-induced leukemia and lymphoma in rodents will be useful in interpreting the situation with XMRV.

2.7 Summary and Future Perspectives

The future of MLV mutagenesis studies promises to be very exciting. It should involve selection for novel and more specific phenotypes. The history of *Drosophila* genetics shows that screens for more specific phenotypes was the key to unraveling specific and critical pathways in developmental processes. Although difficult to design and implement, screens for such specific subtypes of leukemia, or leukemic phenotypes, are the best way to generate the most useful lists of CIS-associated genes. Insertion mutations that allow survival and expansion in a specific setting would seem to offer the best chance to understand how the leukemia phenotype evolves and is maintained. However, several problems must be addressed if MLV mutagenesis is to achieve this goal. This includes the best definition of a significant CIS and co-occurring CIS, methods to ensure saturation recovery of MLV insertion mutations, and identification of the targeted gene(s) at each CIS. Moreover, methods to deal with the clonal heterogeneity of MLV-induced leukemias, which complicates

associations of genotype with phenotype, must be addressed. The best methods for comparisons to human leukemia genome alterations must be sought. Finally, new higher throughput methods for functional validation of CIS-associated genes should be developed so that more candidates can be tested in a reasonable period of time.

References

1. Akagi, K., Suzuki, T., Stephens, R. M., Jenkins, N. A., & Copeland, N. G. (2004). RTCGD: Retroviral tagged cancer gene database. *Nucleic Acids Research*, *32*, D523–D527.
2. Ayton, P. M., & Cleary, M. L. (2003). Transformation of myeloid progenitors by MLL oncoproteins is dependent on Hoxa7 and Hoxa9. *Genes Dev*, *17*, 2298–2307.
3. Bedigian, H. G., Taylor, B. A., & Meier, H. (1981). Expression of murine leukemia viruses in the highly lymphomatous BXH-2 recombinant inbred mouse strain. *Journal of Virology*, *39*, 632–640.
4. Berns, A. (1988). Provirus tagging as an instrument to identify oncogenes and to establish synergism between oncogenes. *Archives of Virology*, *102*, 1–18.
5. Carlson, C. M., Frandsen, J. L., Kirchhof, N., McIvor, R. S., & Largaespada, D. A. (2005). Somatic integration of an oncogene-harboring Sleeping Beauty transposon models liver tumor development in the mouse. *Proceedings of the National Academy of Sciences of the USA*, *102*, 17059–17064.
6. Castilla, L. H., Perrat, P., Martinez, N. J., Landrette, S. F., Keys, R., Oikemus, S., et al. (2004). Identification of genes that synergize with Cbfb-MYH11 in the pathogenesis of acute myeloid leukemia. *Proceedings of the National Academy of Sciences of the USA*, *101*, 4924–4929.
7. Caudell, D., Harper, D. P., Novak, R. L., Pierce, R. M., Slape, C., Wolff, L., et al. Retroviral insertional mutagenesis identifies Zeb2 activation as a novel leukemogenic collaborating event in CALM-AF10 transgenic mice. *Blood*, *115*, 1194–1203.
8. Clappier, E., Cuccini, W., Kalota, A., Crinquette, A., Cayuela, J. M., Dik, W. A., et al. (2007). The C-MYB locus is involved in chromosomal translocation and genomic duplications in human T-cell acute leukemia (T-ALL), the translocation defining a new T-ALL subtype in very young children. *Blood*, *110*, 1251–1261.
9. Collins, E. C., & Rabbitts, T. H. (2002). The promiscuous MLL gene links chromosomal translocations to cellular differentiation and tumour tropism. *Trends in Molecular Medicine*, *8*, 436–442.
10. Cui, J. W., Li, Y. J., Sarkar, A., Brown, J., Tan, Y. H., Premyslova, M., et al. (2007). Retroviral insertional activation of the Fli-3 locus in erythroleukemias encoding a cluster of microRNAs that convert Epo-induced differentiation to proliferation. *Blood*, *110*, 2631–2640.
11. DasGupta, R., Kaykas, A., Moon, R. T., & Perrimon, N. (2005). Functional genomic analysis of the Wnt-wingless signaling pathway, *308*, 826–833 (Epub 2005 Apr 2007).
12. de Ridder, J., Kool, J., Uren, A., Bot, J., Wessels, L., & Reinders, M. (2007). Co-occurrence analysis of insertional mutagenesis data reveals cooperating oncogenes. *Bioinformatics*, *23*, i133–i141.
13. de Ridder, J., Uren, A., Kool, J., Reinders, M., & Wessels, L. (2006). Detecting statistically significant common insertion sites in retroviral insertional mutagenesis screens. *PLoS Computational Biology*, *2*, e166.
14. DePinho, R. A., & Jacks, T. (1999). A bumper crop of cancer genes. *Nature Genetics*, *23*, 253–254.
15. Dong, B., Kim, S., Hong, S., Das Gupta, J., Malathi, K., Klein, E. A., et al. (2007). An infectious retrovirus susceptible to an IFN antiviral pathway from human prostate tumors. *Proceedings of the National Academy of Sciences of the USA*, *104*, 1655–1660.

16. Dupuy, A. J., Morgan, K., von Lintig, F. C., Shen, H., Acar, H., Hasz, D. E., et al. (2001). Activation of the Rap1 guanine nucleotide exchange gene, CalDAG-GEF I, in BXH2 murine myeloid leukemia. *Journal of Biological Chemistry*, 276, 22, 22.
17. Ebert, B. L., Pretz, J., Bosco, J., Chang, C. Y., Tamayo, P., Galili, N., et al. (2008). Identification of RPS14 as a 5q- syndrome gene by RNA interference screen. *Nature*, 451, 335–339.
18. Erkeland, S. J., Verhaak, R. G., Valk, P. J., Delwel, R., Lowenberg, B., & Touw, I. P. (2006). Significance of murine retroviral mutagenesis for identification of disease genes in human acute myeloid leukemia. *Cancer Research*, 66, 622–626.
19. Erlwein, O., Kaye, S., McClure, M. O., Weber, J., Wills, G., Collier, D., et al. Failure to detect the novel retrovirus XMRV in chronic fatigue syndrome. *PLoS One*, 5, e8519.
20. Fan, H. (1997). Leukemogenesis by Moloney murine leukemia virus: A multistep process. *Trends in Microbiology*, 5, 74–82.
21. Fan, H., Mittal, S., Chute, H., Chao, E., & Pattengale, P. K. (1986). Rearrangements and insertions in the Moloney murine leukemia virus long terminal repeat alter biological properties in vivo and in vitro. *Journal of Virology*, 60, 204–214.
22. Gardner, M. B. (1978). Type C viruses of wild mice: Characterization and natural history of amphotropic, ecotropic, and xenotropic MuLV. *Current Topics in Microbiology and Immunology*, 79, 215–259.
23. Habets, G. G., Scholtes, E. H., Zuydgeest, D., van der Kammen, R. A., Stam, J. C., Berns, A., et al. (1994). Identification of an invasion-inducing gene, Tiam-1, that encodes a protein with homology to GDP-GTP exchangers for Rho-like proteins. *Cell*, 77, 537–549.
24. Hohn, O., Krause, H., Barbarotto, P., Niederstadt, L., Beimforde, N., Denner, J., et al. (2009). Lack of evidence for xenotropic murine leukemia virus-related virus(XMRV) in German prostate cancer patients. *Retrovirology*, 6, 92.
25. Jenkins, N. A., Copeland, N. G., Taylor, B. A., Bedigian, H. G., & Lee, B. K. (1982). Ecotropic murine leukemia virus DNA content of normal and lymphomatous tissues of BXH-2 recombinant inbred mice. *Journal of Virology*, 42, 379–388.
26. Jin, S., Zhao, H., Yi, Y., Nakata, Y., Kalota, A., & Gewirtz, A. M. (2010). c-Myb binds MLL through menin in human leukemia cells and is an important driver of MLL-associated leukemogenesis. *Journal of Clinical Investigation*, 120, 593–606.
27. Keng, V. W., Villanueva, A., Chiang, D. Y., Dupuy, A. J., Ryan, B. J., Matise, I., et al. (2009). A conditional transposon-based insertional mutagenesis screen for genes associated with mouse hepatocellular carcinoma. *Nature Biotechnology*, 27, 264–274.
28. Kim, R., Trubetskoy, A., Suzuki, T., Jenkins, N. A., Copeland, N. G., & Lenz, J. (2003). Genome-based identification of cancer genes by proviral tagging in mouse retrovirus-induced T-cell lymphomas. *Journal of Virology*, 77, 2056–2062.
29. Klinger, M. B., Guilbault, B., Goulding, R. E., & Kay, R. J. (2005). Deregulated expression of RasGRP1 initiates thymic lymphomagenesis independently of T-cell receptors. *Oncogene*, 24, 2695–2704.
30. Kool, J., & Berns, A. (2009). High-throughput insertional mutagenesis screens in mice to identify oncogenic networks. *Nature Review Cancer*, 9, 389–399.
31. Lahortiga, I., De Keersmaecker, K., Van Vlierberghe, P., Graux, C., Cauwelier, B., Lambert, F., et al. (2007). Duplication of the MYB oncogene in T cell acute lymphoblastic leukemia. *Nature Genetics*, 39, 593–595.
32. Landrette, S. F., Kuo, Y. H., Hensen, K., S., Perrat, P. N., Van de Ven, W. J., et al. (2005). Plagl1 and Plagl2 are oncogenes that induce acute myeloid leukemia in cooperation with Cbfb-MYH11. *Blood*, 105, 2900–2907.
33. Largaespada, D. A., Brannan, C. I., Jenkins, N. A., & Copeland, N. G. (1996). Nf1 deficiency causes Ras-mediated granulocyte/macrophage colony stimulating factor hypersensitivity and chronic myeloid leukaemia. *Nature Genetics*, 12, 137–143.
34. Largaespada, D. A., & Collier, L. S. (2008). Transposon-mediated mutagenesis in somatic cells: Identification of transposon-genomic DNA junctions. *Methods in Molecular Biology*, 435, 95–108.

35. Largaespada, D. A., Shaughnessy, J. D., Jenkins, N. A., & Copeland, N. G. (1995). Retroviral insertion at the Evi-2 locus in BXH-2 myeloid leukemia cell lines disrupts Nf1 expression without changes in steady state ras-GTP levels. *Journal of Virology*, 69, 5095–5102.
36. Lauchle, J. O., Kim, D., Le, D. T., Akagi, K., Crone, M., Krisman, K., et al. (2009). Response and resistance to MEK inhibition in leukaemias initiated by hyperactive Ras. *Nature*, 461, 411–414.
37. Li, J., Shen, H., Himmel, K. L., Dupuy, A. J., Largaespada, D. A., Nakamura, T., et al. (1999). Leukaemia disease genes: Large-scale cloning and pathway predictions [see comments]. *Nature Genetics*, 23, 348–353.
38. Lombardi, V. C., Ruscetti, F. W., Das Gupta, J., Pfof, M. A., Hagen, K. S., Peterson, D. L., et al. (2009). Detection of an infectious retrovirus, XMRV, in blood cells of patients with chronic fatigue syndrome. *Science*, 326, 585–589.
39. Lund, A. H., Turner, G., Trubetskoy, A., Verhoeven, E., Wientjens, E., Hulsman, D., et al. (2002). Genome-wide retroviral insertional tagging of genes involved in cancer in Cdkn2a-deficient mice. *Nature Genetics*, 32, 160–165.
40. Maeda, N., Fan, H., & Yoshikai, Y. (2008). Oncogenesis by retroviruses: Old and new paradigms. *Reviews in Medical Virology*, 18, 387–405.
41. Mates, L., Chuah, M. K., Belay, E., Jerchow, B., Manoj, N., Acosta-Sanchez, A., et al. (2009). Molecular evolution of a novel hyperactive Sleeping Beauty transposase enables robust stable gene transfer in vertebrates. *Nature Genetics*, 41, 753–761.
42. Mikkers, H., Allen, J., Knipscheer, P., Romeijn, L., Hart, A., Vink, E., et al. (2002). High-throughput retroviral tagging to identify components of specific signaling pathways in cancer. *Nature Genetics*, 32, 153–159.
43. Miller, D. G., Adam, M. A., & Miller, A. D. (1990). Gene transfer by retrovirus vectors occurs only in cells that are actively replicating at the time of infection. *Molecular and Cellular Biology*, 10, 4239–4242.
44. Mukhopadhyaya, R., & Wolff, L. (1992). New sites of proviral integration associated with murine promonocytic leukemias and evidence for alternate modes of c-myc activation [published erratum appears in *J Virol* 1993 May;67(5):2960]. *Journal of Virology*, 66, 6035–6044.
45. Mukoyama, Y., Chiba, N., Mucenski, M. L., Satake, M., Miyajima, A., Hara, T., et al. (1999). Hematopoietic cells in cultures of the murine embryonic aorta-gonad-mesonephros region are induced by c-Myb. *Current Biology*, 9, 833–836.
46. Murati, A., Gervais, C., Carbuccia, N., Finetti, P., Cervera, N., Adelaide, J., et al. (2009). Genome profiling of acute myelomonocytic leukemia: Alteration of the MYB locus in MYST3-linked cases. *Leukemia*, 23, 85–94.
47. Muriaux, D., & Rein, A. (2003). Encapsidation and transduction of cellular genes by retroviruses. *Frontiers in Bioscience*, 8, d135–d142.
48. Nagase, T., Yamakawa, H., Tadokoro, S., Nakajima, D., Inoue, S., Yamaguchi, K., et al. (2008). Exploration of human ORFeome: High-throughput preparation of ORF clones and efficient characterization of their protein products. *DNA Research*, 15, 137–149.
49. Nakamura, T., Largaespada, D. A., Lee, M. P., Johnson, L. A., Ohyashiki, K., Toyama, K., et al. (1996a). Fusion of the nucleoporin gene NUP98 to HOXA9 by the chromosome translocation t(7;11)(p15;p15) in human myeloid leukaemia. *Nature Genetics*, 12, 154–158.
50. Nakamura, T., Largaespada, D. A., Shaughnessy, J. D., Jr., Jenkins, N. A., & Copeland, N. G. (1996b). Cooperative activation of Hoxa and Pbx1-related genes in murine myeloid leukaemias. *Nature Genetics*, 12, 149–153.
51. Nishigaki, K., Hanson, C., Thompson, D., Yugawa, T., Hisasue, M., Tsujimoto, H., et al. (2002). Analysis of the disease potential of a recombinant retrovirus containing Friend murine leukemia virus sequences and a unique long terminal repeat from feline leukemia virus. *Journal of Virology*, 76, 1527–1532.

52. O'Connell, P., Viskochil, D., Buchberg, A. M., Fountain, J., Cawthon, R. M., Culver, M., et al. (1990). The human homolog of murine Evi-2 lies between two von Recklinghausen neurofibromatosis translocations. *Genomics*, 7, 547–554.
53. Oki-Idouchi, C. E., & Lorenzo, P. S. (2007). Transgenic overexpression of RasGRP1 in mouse epidermis results in spontaneous tumors of the skin. *Cancer Research*, 67, 276–280.
54. Ou, C. Y., Boone, L. R., Koh, C. K., Tennant, R. W., & Yang, W. K. (1983). Nucleotide sequences of gag-pol regions that determine the Fv-1 host range property of BALB/c N-tropic and B-tropic murine leukemia viruses. *Journal of Virology*, 48, 779–784.
55. Patriotis, C., Makris, A., Bear, S. E., & Tsichlis, P. N. (1993). Tumor progression locus 2 (Tpl-2) encodes a protein kinase involved in the progression of rodent T-cell lymphomas and in T-cell activation. *Proceedings of the National Academy of Sciences of the USA*, 90, 2251–2255.
56. Reuther, G. W., Lambert, Q. T., Rebhun, J. F., Caligiuri, M. A., Quilliam, L. A., & Der, C. J. (2002). RasGRP4 is a novel Ras activator isolated from acute myeloid leukemia. *Journal of Biological Chemistry*, 277, 30508–30514.
57. Risser, R., Horowitz, J. M., & McCubrey, J. (1983). Endogenous mouse leukemia viruses. *Annual Review of Genetics*, 17, 85–121.
58. Schlager, R., Choe, D. J., Brown, K. R., Thaker, H. M., & Singh, I. R. (2009). XMRV is present in malignant prostatic epithelium and is associated with prostate cancer, especially high-grade tumors. *Proceedings of the National Academy of Sciences of the USA*, 106, 16351–16356.
59. Slape, C., Hartung, H., Lin, Y. W., Bies, J., Wolff, L., & Aplan, P. D. (2007). Retroviral insertional mutagenesis identifies genes that collaborate with NUP98-HOXD13 during leukemic transformation. *Cancer Research*, 67, 5148–5155.
60. Starr, T. K., & Largaespada, D. A. (2005). Cancer gene discovery using the sleeping beauty transposon. *Cell Cycle*, 4, 1744–1748.
61. Su, Q., Prosser, H. M., Campos, L. S., Ortiz, M., Nakamura, T., Warren, M., et al. (2008). A DNA transposon-based approach to validate oncogenic mutations in the mouse. *Proceedings of the National Academy of Sciences of the USA*, 105, 19904–19909.
62. Turcotte, K., Gauthier, S., Tuite, A., Mullick, A., Malo, D., & Gros, P. (2005). A mutation in the Icsbp1 gene causes susceptibility to infection and a chronic myeloid leukemia-like syndrome in BXH-2 mice. *Journal of Experimental Medicine*, 201, 881–890.
63. Uren, A. G., Kool, J., Matentzoglou, K., de Ridder, J., Mattison, J., van Uiter, M., et al. (2008). Large-scale mutagenesis in p19(ARF)- and p53-deficient mice identifies cancer genes and their collaborative networks. *Cell*, 133, 727–741.
64. Uren, A. G., Mikkers, H., Kool, J., van der Weyden, L., Lund, A. H., Wilson, C. H., et al. (2009). A high-throughput splinkerette-PCR method for the isolation and sequencing of retroviral insertion sites. *Nature Protocols*, 4, 789–798.
65. Urisman, A., Molinaro, R. J., Fischer, N., Plummer, S. J., Casey, G., Klein, E. A., et al. (2006). Identification of a novel Gammaretrovirus in prostate tumors of patients homozygous for R462Q RNASEL variant. *PLoS Pathogens*, 2, e25.
66. van der Lugt, N. M., Domen, J., Verhoeven, E., Linders, K., van der Gulden, H., Allen, J., et al. (1995). Proviral tagging in E mu-myc transgenic mice lacking the Pim-1 proto-oncogene leads to compensatory activation of Pim-2. *EMBO Journal*, 14, 2536–2544.
67. van Lohuizen, M., Breuer, M., & Berns, A. (1989). N-myc is frequently activated by proviral insertion in MuLV-induced T cell lymphomas. *EMBO Journal*, 8, 133–136.
68. van Lohuizen, M., Verbeek, S., Scheijen, B., Wientjens, E., van der Gulden, H., & Berns, A. (1991). Identification of cooperating oncogenes in E mu-myc transgenic mice by provirus tagging [see comments]. *Cell*, 65, 737–752.
69. van Ooyen, A., & Nusse, R. (1984). Structure and nucleotide sequence of the putative mammary oncogene int-1; proviral insertions leave the protein-encoding domain intact. *Cell*, 39, 233–240.
70. Varmus, H. E. (1983). Using retroviruses as insertional mutagens to identify cellular oncogenes. *Progress in Clinical and Biological Research*, 119, 23–35.

71. Wang, C. L., Wang, B. B., Bartha, G., Li, L., Channa, N., Klinger, M., et al. (2006). Activation of an oncogenic microRNA cistron by provirus integration. *Proceedings of the National Academy of Sciences of the USA*, 103, 18680–18684.
72. Wangenstein, K. J., Wilber, A., Keng, V. W., He, Z., Matise, I., Wangenstein, L., et al. (2008). A facile method for somatic, lifelong manipulation of multiple genes in the mouse liver. *Hepatology*, 47, 1714–1724.
73. Westbrook, T. F., Martin, E. S., Schlabach, M. R., Leng, Y., Liang, A. C., Feng, B., et al. (2005). A genetic screen for candidate tumor suppressors identifies REST. *Cell*, 121, 837–848.
74. Whitcomb, J. M., & Hughes, S. H. (1992). Retroviral reverse transcription and integration: Progress and problems. *Annual Review of Cell Biology*, 8, 275–306.
75. Wiesner, S. M., Decker, S. A., Larson, J. D., Ericson, K., Forster, C., Gallardo, J. L., et al. (2009). De novo induction of genetically engineered brain tumors in mice using plasmid DNA. *Cancer Research*, 69, 431–439.
76. Wolff, L., Koller, R., & Davidson, W. (1991). Acute myeloid leukemia induction by amphotropic murine retrovirus (4070A): Clonal integrations involve c-myc in some but not all leukemias. *Journal of Virology*, 65, 3607–3616.
77. Wolff, L., Koller, R., Hu, X., & Anver, M. R. (2003). A Moloney murine leukemia virus-based retrovirus with 4070A long terminal repeat sequences induces a high incidence of myeloid as well as lymphoid neoplasms. *Journal of Virology*, 77, 4965–4971.
78. Wu, X., Luke, B. T., & Burgess, S. M. (2006). Redefining the common insertion site. *Virology*, 344, 292–295.
79. Yin, B., Delwel, R., Valk, P. J., Wallace, M. R., Loh, M. L., Shannon, K. M., et al. (2009). A retroviral mutagenesis screen reveals strong cooperation between Bcl11a overexpression and loss of the Nf1 tumor suppressor gene. *Blood*, 113, 1075–1085.
80. Yin, B., & Largaespada, D. A. (2007). PCR-based procedures to isolate insertion sites of DNA elements. *Biotechniques*, 43, 79–84.
81. Zender, L., Xue, W., Zuber, J., Semighini, C. P., Krasnitz, A., Ma, B., et al. (2008). An oncogenomics-based in vivo RNAi screen identifies tumor suppressors in liver cancer. *Cell*, 135, 852–864.
82. Zippo, A., De Robertis, A., Serafini, R., & Oliviero, S. (2007). PIM1-dependent phosphorylation of histone H3 at serine 10 is required for MYC-dependent transcriptional activation and oncogenic transformation. *Nature Cell Biology*, 9, 932–944.

Insertional Mutagenesis Strategies in Cancer Genetics

Dupuy, A.J.; Largaespada, D.A. (Eds.)

2011, VIII, 200 p., Hardcover

ISBN: 978-1-4419-7655-0