from a viral enhancer-activated gene initiate at the normal transcriptional start sites, suggesting that viral enhancer elements have increased the efficiency of the normal promoter. Enhancer insertion involves integration of a provirus at the 5' end of a gene in the reverse transcriptional orientation, or at the 3' end in the same transcriptional orientation. In this way, the positioning of the viral promoter between the viral enhancer and the host gene promoter is avoided, suggesting a model in which the bi-directional viral enhancer acts primarily on the closest promoters (Dickson et al., 1984; Nusse et al., 1984). However, exceptions have been described as both *Wnt-1/int-1* and *Fgf-1/int-2* can be activated by downstream insertions of mouse mammary tumour virus (MMTV) in the opposite transcriptional orientation (Clausse et al., 1993). Since in both cases the integrated proviruses lack the 5' LTR, an alternative enhancer insertion model is suggested in which the viral enhancer can only function if it is not transcribed.

When the proviral insertion site is located within the 3' untranslated region of the proto-oncogene and in the same transcriptional orientation, the transcript is cleaved at the polyadenylation site of the 5' LTR. This premature polyadenylation may result in production of transcripts with increased stability due to removal of destabilizing sequences, such as AUUUA motifs, which are commonly located in the 3' untranslated region (UTR) of short-lived messages (Shaw and Kamen, 1986). Representative examples are *Gfi-2/IL-9R*, *Pim-1*, *Pim-2* and *N-myc*.

The integration of a provirus into the transcription unit can affect the resulting protein. If a viral insertion disrupts coding domains, the protein sequence can be completely inactivated, or mutated in such a way that an aberrant gene product with abnormal biologic activity is produced. This may occur in concert with enhancer or promoter activation, thereby generating high levels of a mutant protein. Examples of this case include c-*myb*, *Tpl-2*, *Tiam-1*, *int-3* and *int-6*.

Although *cis*-activation of the host genes appear to be the prominent mechanism by which slow-transforming retroviruses exert their transforming potential, certain retroviruses of this class, namely bovine leucosis virus (BLV) and human T-cell lymphotropic virus (HTLV), are capable of *trans*-activating viral genes as well as distinct host genes via the virally encoded transactivator protein Tax (Sodroski et al., 1985). Interestingly, it has been reported that specific sequences within the U3 region of the Moloney murine leukemia virus (MuLV) LTR also encode a transcriptional *trans*-activator, despite the fact that this region lacks an extensive open reading frame (Choi and Faller, 1994). Transfection experiments indicated that this region in the U3 LTR, through activation of the transcription factor AP-1, is capable of regulating the expression of cellular genes that contain tetraphorbolester (TPA)-responsive elements (TRE) in their promoter region (Weng et al., 1995).

Gene inactivation

Proviral integration into, or close to a gene, is more likely to impair its function, but since this occurs on only one of any chromosome pair it will generally be compensated for by the unaffected allele on the other chromosome. Only a few examples have been reported where a gene has been inactivated by insertional mutagenesis. The *p53* gene was inactivated by proviral insertion of the Friend murine leukaemia virus (FMuLV) in an erythroleukaemia cell line. In this case it was reported that viruses were inserted at the coding sequence of both alleles of the *p53* gene (Hicks and Mowat, 1988).

Another example of gene inactivation by proviral insertion was that of the dilute coat colour mutation of mice. Insertion of an endogenous provirus, Emv-3, occurred in the chromosome (Jenkins et al., 1981). Other examples include the hairless mutation of mice (Stoye et al., 1988) and the slow-feathering mutation of chickens (Bacon et al., 1990).

Retroviruses as a genetic tool for oncogene identification

Viruses are aetiologically linked to approximately 20% of all malignancies worldwide (Blattner, 1999). Retroviruses account for approximately 8–10% of their total. Human retroviruses cause malignancy either via direct effects or through interactions with other oncogenic viruses. Retroviruses can cause cancer, even though they lack a dominantly acting oncogene. Due to variations in the two principal viral genetic determinants of leukaemogenesis, namely the *env* region and the viral LTR, a number of leukaemogenic retroviruses have been isolated which induce malignancy in a variety of haematopoietic cell types. Other strains of slow-transforming retroviruses cause carcinomas of the mammary gland, instead of haematopoietic tumours.

Many genes have been shown to be activated or otherwise mutated by insertional mutagenesis, and proviral tagging has led to identification of several novel genes which are implicated in tumorigenesis. The major contribution to the identification of cellular oncogenes has come from transduction of cellular oncogenes by retroviruses. The first example of interaction between retroviruses and oncogenes was the insertion of avian leucosis virus (ALV) at the c-*myc* locus (Hayward et al., 1981). The *myc* oncogene was also found to be captured by ALV and to induce transformation. These observations first uncovered oncogenes. The induction and isolation of mutants of avian retroviruses have contributed considerably to our understanding of the viral gene functions involved in replication and cell transformation. In avian retroviruses it has been shown that some rare isolates have acquired two cellular sequences that are separate in the normal cell. Examples are the avian erythroblastosis virus that contains *erb*-A and *erb*-B (Vennstrom and Bishop, 1982), the MH2 avian carcinoma virus that

contains *myc* and *mil* (Coll et al., 1983) and the E6 avian myeloblastosis virus that contains *myb* and *ets* (Leprince et al, 1983). Therefore, oncogene transduction by retroviruses has given insight into oncogenic mechanisms.

Retroviruses are also useful as molecular tags for identifying novel potential oncogenic sequences. During the retroviral life cycle the viral reverse-transcribed DNA is being inserted into the cellular genome, disrupting the normal function of the gene located in this region. Molecular cloning of these regions has helped us to identify numerous oncogenes such as *int-1, int-2, Pim-1, bmi-1, Tpl-1, Tpl-2*. Analysis of the pattern of proviral insertions in a particular tumour can indicate oncogene co-operation. In the case of feline leukaemia virus (FeLV), a significant coincidence between activation of the *myc* oncogene and the *fit-1* or *bmi-1* loci was found, indicating gene complementation for tumour formation (Tsatsanis et al., 1994).

Leukaemias associated with retroviral infection

Viruses have long been recognized as primary causes of leukaemia in many species of animals, including man. Therefore, a practical reason for the study of leukaemogenic viruses is definition of such targets in order to interfere with the cycle of viral infection and consequently prevent development of leukaemia by targeting specific prophylactic or therapeutic measures against the disease. It has been considered that multiple steps may be involved in the disease development. In at least some instances, the activation of cellular proto-oncogenes due to proviral insertion may be a relatively late event in leukaemogenesis (O'Donnell et al., 1985). In several virus-induced leukaemias, the early pre-leukaemic state is characterized by enhanced proliferation of specific cell types (Davis et al., 1987; D'Andrea et al., 1992).

Retroviruses and herpes viruses cause naturally occurring leukaemia in animals and man (Blattner, 1999). Such examples are HTLV (Watanabe, 1997; Ferreira et al., 1997), ALV (Beug et al., 1996) and FeLV (Rohn et al., 1996) among the retroviruses, and the Epstein-Barr virus (EBV) (Ohtsubo et al., 1999) and Marek's disease virus (MDV) among the herpes viruses (Hihara et al., 1998). Many common features of these viral infections are readily discerned. For example, each may establish a life-long infection in its host. It is in those persistently infected individuals that the pathogenic effects of the virus are apparent, often after a long incubation period during which the host is clinically healthy. However, it is in this period that critical events may occur which contribute eventually to the development of leukaemia.

Characteristically, the viruses are transmitted in contact with infected body fluids or congenitally from these healthy carriers to susceptible individuals. Several members of the retroviral family are involved in haematopoietic tumours. There are three separate subfamilies of retroviruses: the oncoviruses, lentiviruses and spumaviruses. Among these, the oncoviruses are most frequently associated with the development of tumours, while the lentiviruses are usually implicated in degenerative diseases such as HIV that causes the acquired immunodeficiency syndrome (AIDS). The oncovirus subfamily can be divided into two subgroups. The first includes FeLV, ALV and MuLVs; they are characterized by a simple genomic structure and often replicate *in vivo* to produce cell-free infectious viruses. The second consists of HTLV and BLV, which have a more complex genomic structure and are usually maintained *in vivo* in a latent non-productive form.

Diseases that are associated with HTLV infection are adult T-cell leukaemia (ATL) and tropical spastic paraparesis (TSP), which is an HTLV-associated myelopathy in Japan (Manns et al., 1999). ATL is an aggressive lymphoproliferative condition of mature T cells, which usually have the phenotype $CD4^+$, $CD8^-$, $CD25^+$. Middle-aged or elderly patients are frequently presented with an advanced stage of the disease with a leukaemia characterized by cells with pleomorphic lobulated nuclei, lymphadenopathy and hepatosplenomegaly. A striking feature is the high incidence of hypercalcaemia, sometimes accompanied by lytic bone lesions.

MuLVs induce disease with a relatively long latency. MuLV is implicated in both preleukaemic events as well as in later stages in MuLV-induced leukaemia. A variety of studies have provided compelling evidence that the later stages of this disease are predominated by the occurrence of somatic mutations due to proviral insertion and subsequent clonal expansion, triggered by mutations in specific genes. A vast number of common insertional sites have been identified and among them are many loci that contain genes which become deregulated due to proviral integration. The hitherto characterized gene products are invariably regulatory proteins, including transcription factors such as Bmi-1, c-fos, c-myb, Evi-1, N-myc, p53 and Tpl-2 (Mucenski et al., 1988; van Lohuizen et al., 1989; Haupt et al., 1991; Johnson et al., 1992; Mukhopadhyaya et al., 1992), growth factors and cytokines such as CSF-1, GM-CSF, IL-2, IL-3, IL-5 (Pierce et al., 1985; Gisselbrecht et al., 1987; Beaty et al., 1999), growth factor receptors such as EGF-R/c-erbB, EPO-R, IL-2Rβ, IL-6R (Callaghan et al., 1993; Chretien et al., 1994; Sugiyama et al., 1996), cytoplasmic protein kinases such as c-mos, Pim-1, Pim-2, Tpl-2 (Cuypers et al., 1984; Patriotis et al., 1993), and GTPases such as c-H-ras, c-K-ras (Gougopoulou et al., 1996).

MMTV-induced mammary carcinomas

MMTV is a type B retrovirus that usually causes adenocarcinomas of mammary epithelial tissue after a long latency period. The targets of primary infection are most often the B cells present either within the intestinal environment or in the associated Peyer's patches. An