Vector integration
. . . .

Location, location, location

HP Mok and AML Lever

A new analysis of vector integration has shown how integration site varies among three different groups of retroviruses[1]

Integration sites of retroviruses

Even before we knew what retroviruses were we knew they could induce malignancy.[2] We now know that a number of them are directly associated with cell transformation through mechanisms such as insertional mutagenesis.[3,4] Nonetheless, the utility of retroviruses as gene vectors has led to their widespread use *in vivo* and *in vitro*. So the relatively recent appearance of leukaemia in two child recipients of retroviral vectors was perhaps less of a surprise to retrovirologists than it was to gene therapists. However, the effect of these sad occurrences on the field has been profound: in particular, it has installed the goal of the ability to control the sites of vector integration as a holy grail for gene therapy researchers.

Early studies that attempted to identify integration sites used laborious techniques such as positional cloning,[5] isochore analysis[6] and FISH.[7] However, powerful PCR methodologies and the completion of the human genome sequence have revolutionized the identification of vector integration sites. Several large studies have shown that retroviral integration (and hence retroviral vector delivery) is far from random: factors such as base composition[6,8] and the presence of alphoid[9] or Alu repeats[10] as well as transcriptional activity all influence the integration site.

Now Rick Mitchell and co-workers have extended these previous findings in a large and well-controlled analysis of new and previously identified integration sites, which compares site preference of murine (MLV), avian (ASLV) and human (HIV) retroviruses. Chromosomal distribution, GC content, transcriptional activity and distance from cellular gene transcription start sites were compared using appropriate controls.

The authors confirmed that HIV has a tendency to integrate in regions rich in expressed genes and they also showed that it does not favour integration in the intervening CpG islands. These results were cell type independent and it is likely that other viral vectors will demonstrate similar patterns. They also confirmed previous findings that MLV favours transcriptional start sites. By contrast, their new data showed that ASLV has only a slight predilection for transcriptionally active regions and no start site preference. Moreover, many ASLV integrations occur in intergenic regions.

It would be interesting to extend these studies to cells with different cycling properties than those that these authors studied. Vector integration is complete within 24 h of infection and often considerably sooner in rapidly cycling cells,[11] so the authors' analysis at 48 h might have marginally over-represented integrations in such cells. It would also be interesting to compare integration sites in growth-arrested cells to those in otherwise identical cycling cells. This would be of particular relevance to those that use lentiviruses (including HIV as studied here) to deliver genes to non-dividing cells.

What guides the integration event?

The clear inference from these new data is that different retroviral pre-integration complexes (PICs) recognize different features of genomic nucleoproteins, but what is the crucial link in this process? Integrase itself has only a weak nucleotide sequence preference[12] so it is much more likely that it is the DNA associated proteins.

Nucleosome structure is important for both ASLV integration, where DNA ligation occurs more efficiently in compact chromatin, and for HIV, where the reverse is true.[13] Addition of the transcription factor HNF3 inhibited ASLV end joining at the HNF3 sites.[13] This together with previous data that showed that ASLV integration does not favour transcriptionally active sites[14] would have predicted the Mitchell findings and it is nice to see this so comprehensively confirmed.

Could covalent modifications of Histone proteins be important for target site selection? Lysine 4 and 9 of H3 are methylated at transcriptional start sites[15] and both H3 and H4 have varying acetylation patterns depending on whether they are in open reading frames or intergenic regions[16] Other chromosomal proteins also redistribute with transcription and at least one, the INI/SNF5 complex, has been shown to interact with HIV-1 integrase.[17] Any or all of these factors (and others) might affect integration site choice.

What are the implications for gene therapy?

Mitchell and co-workers suggest that the large number of intergenic integrations seen with ASLV might make it a good gene vector candidate since it could avoid the risk of insertional mutagenesis. The assumption that noncoding regions are 'safe' and do not play a significant role in cellular regulation is however not completely secure and there is good published data on their importance in influencing heterochromatin formation (for reviews see Henikoff,[18] Fischle *et al* and [19] Maison and Almouzni[20]). Retrotransposons, endogenous relatives of retroviruses, certainly do affect heterochromatin formation.[21,22] Thus noncoding regions might play important regulatory roles in the genome.

There are also doubts about the subsequent transcriptional activity of an ASLV inserted in an intergenic region. Spreading chromatinization might silence the provirus. So multiple integrations might be needed to achieve at least one that is regularly transcribed, which in turn make it

more likely that one of these would affect a critical site.

Overall, Mitchell and co-workers have completed a well-executed and comprehensive study that adds to our knowledge of retroviral integration and the factors that affect it. Bushman's group is one of the major players in this field and the work takes us one step further towards understanding and ultimately controlling retroviral integration. However, there is more to be discovered before we can make definitive decisions about our choice of retroviral gene vectors. It is perhaps significant that despite the apparent predilection of HIV for transcriptionally active sites, and its abundance, only one case of probable insertional mutagenesis has been documented reasonably securely[23] and a second with another lentivirus, Feline immunodeficiency virus.[24] Many of the target cells for HIV are lymphocytes with a short lifespan, but there are also many infected longer lived cells of the monocyte macrophage series in every carrier. For this reason, if the virus had some potential for oncogenicity one might have expected this to appear in at least some of the 40 million infected individuals who are each producing millions of integrating viruses each day. Arguably the lytic nature of HIV has concealed such a tendency that might be revealed as they are used as vectors. However, there is no room for complacency. If evidence of insertional mutagenesis from the use of any specific retro or lentiviral vector comes to light subsequently, we will have to think seriously about our choices and indeed whether with our current state of knowledge we can risk any integrating vectors in human studies.

Acknowledgements

HP Mok and AML Lever are at the University of Cambridge Department of Medicine, Addenbrooke's Hopital, Hills Road, Cambridge, Cambs BC2 2QQ, UK.
E-mail: amll1@mole.bio.cam.ac.uk
Published online 28 October 2004

1 Mitchell RS *et al*. Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol* 2004; **2**: 1–11.

2 Rous P. A transmissible avian neoplasm: Sarcoma of the common fowl. *J Exp Med* 1910; **12**: 696–705.

3 Coffin JM, Hughes SH, Varmus HE. *Retroviruses*. Cold Spring Harbor Laboratory: New York, 1997: 843.

4 Weiss R, Teich N, Varmus H, Coffin J. *RNA Tumour Viruses*. Cold Spring Harbor Laboratory: New York, 1984.

5 Seiki M, Eddy R, Shows TB, Yoshida M. Nonspecific integration of the HTLV provirus into adult T-cell leukaemia cells. *Nature* 1984; **309**: 640–642.

6 Zoubak S *et al*. Regional specificity of HTLV-1 proviral integration in the human genome. *Gene* 1994; **143**: 155–163.

7 Richardson JH *et al*. Chromosomal positioning of human T-lymphotropic type 1 proviruses by fluorescence *in situ* hybridisation. *J Virol Methods* 2001; **93**: 65–74.

8 Elleder D, Pavlicek A, Paces J, Hejnar J. Preferential integration of human immunodeficiency virus type 1 into genes, cytogenetic R bands and GC-rich DNA regions: insight from the human genome sequence. *FEBS Lett* 2002; **51**: 285–286.

9 Carteau S, Hoffmann C, Bushman F. Chromosome structure and human immunodeficiency virus type 1 cDNA integration: centromeric alphoid repeats are a disfavoured target. *J Virol* 1998; **72**: 4005–4014.

10 Stevens SW, Griffith JD. Sequence analysis of the human DNA flanking sites of human immunodeficiency virus type 1 integration. *J Virol* 1996; **70**: 6459–6462.

11 Vandegraaff NK, Burrell R, C J, Li P. Kinetics of human immunodeficiency virus type 1 (HIV) DNA integration in acutely infected cells as determined using a novel assay for detection of integrated HIV DNA. *J Virol* 2001; **75**: 11253–11260.

12 Harper AL, Sudol M, Katzman M. An amino acid in the central catalytic domain of three retroviral integrases that affects target site selection in nonviral DNA. *J Virol* 2003; **77**: 3838–3845.

13 Taganov KD *et al*. Integrase-specific enhancement and suppression of retroviral DNA integration by compacted chromatin structure *in vitro*. *J Virol* 2004; **78**: 5848–5855.

14 Weidhaas JB, Angelichio EL, Fenner S, Coffin JM. Relationship between retroviral DNA integration and gene expression. *J Virol* 2000; **74**: 8382–8389.

15 Liang G *et al*. Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome. *Proc Natl Acad Sci USA* 2004; **101**: 7357–7362.

16 Kurdistani SK, Tavazoie S, Grunstein M. Mapping global histone acetylation patterns to gene expression. *Cell* 2004; **117**: 721–733.

17 Yung E *et al*. Specificity of interaction of INI1/hSNF5 with retroviral integrases and its functional significance. *J Virol* 2004; **78**: 2222–2231.

18 Henikoff S. Heterochromatin function in complex genomes. *Biochim Biophys Acta* 2000; **1470**: 1–8.

19 Fischle W, Wang Y, Allis CD. Histone and chromatin cross-talk. *Curr Opin Cell Biol* 2003; **15**: 172–183.

20 Maison C, Almouzni G. HP1 and the dynamics of heterochromatin maintenance. *Nat Rev Mol Cell Biol* 2004; **5**: 296–304.

21 Lippman Z *et al*. Role of transposable elements in heterochromatin and epigenetic control. *Nature* 2004; **430**: 471–476.

22 Schramke V, Allshire R. Hairpin RNAs and retrotransposon LTRs effect RNAi and chromatin-based gene silencing. *Science* 2003; **301**: 1069–1074.

23 Herndier BG *et al*. immunodeficiency syndrome-associated T-cell lymphoma: evidence for human immunodeficiency virus type 1-associated T-cell transformation. *Blood* 1992; **79**: 1768–1774.

24 Beatty J *et al*. Feline immunodeficiency virus integration in B-cell lymphoma identifies a candidate tumor suppressor gene on human chromosome 15q15. *Cancer Res* 2002; **62**: 7175–7180.