Research

A murine leukemia virus with Cre-LoxP excisible coding sequences allowing superinfection, transgene delivery, and generation of host genomic deletions

Clifford L Wang*[1], J Graeme Hodgson[2], Tiffany Malek[4], Finn Skou Pedersen[3] and Matthias Wabl[1]

Address: [1]Department of Microbiology and Immunology, University of California, San Francisco, CA, USA, 94143-0414, [2]Department of Laboratory Medicine, University of California, San Francisco, CA, USA, 94143-0808, [3]Department of Molecular Biology and Department of Medical Microbiology and Immunology, University of Aarhus, Denmark, DK-8000 and [4]Fred Hutchinson Cancer Research Centre, Department of Human Biology, Seattle, WA 98109, USA

Email: Clifford L Wang* - cliffw@itsa.ucsf.edu; J Graeme Hodgson - ghodgson@cc.ucsf.edu; Tiffany Malek - tmalek@fhcrc.org; Finn Skou Pedersen - fsp@mb.au.dk; Matthias Wabl - mutator@itsa.ucsf.edu

* Corresponding author

Abstract

Background: To generate a replication-competent retrovirus that could be conditionally inactivated, we flanked the viral genes of the Akv murine leukemia virus with LoxP sites. This provirus can delete its envelope gene by LoxP/Cre mediated recombination and thereby allow superinfection of Cre recombinase expressing cells.

Results: In our studies, the virus repeatedly infected the cell and delivered multiple copies of the viral genome to the host genome; the superinfected cells expressed a viral transgene on average twenty times more than non-superinfected cells. The insertion of multiple LoxP sites into the cellular genome also led to genomic deletions, as demonstrated by comparative genome hybridization.

Conclusion: We envision that this technology may be particularly valuable for delivering transgenes and/or causing deletions.

Background

Resistance to superinfection by receptor interference has been established for several groups of retroviruses when the infected cell produces the viral envelope protein [1-7]. In addition to being a component of burgeoning viruses, the envelope is also believed to bind to the cellular receptor of the virus, either intracellularly or at the cell membrane. This interaction prevents adequate surface display of the receptor, without which other retroviruses dependent on the same receptor cannot enter and infect the cell [8-10].

Resistance to superinfection may present an obstacle to applications of retroviral gene technology where multiple hits are required. Such applications include delivery of transgenes in multiple copies and introduction of multiple target sites for DNA recombinases to create deletions or rearrangements of host cell DNA. While gammaretroviral packaging cell lines can be manipulated to generate a high multiplicity of infection in cell culture [11,12], it may be useful, especially for studies in animal models, to have a retrovirus that circumvents superinfection barriers without such lines.

Akv is a well-characterized ecotropic murine leukemia virus (MLV) [13-18]; the cell surface receptor for Akv and other ecotropic MLVs is a cationic amino acid transporter, CAT-1 [19]. Among various designs of replication-competent vectors based upon gammaretroviruses, the most stable transgene maintenance has been achieved by inserting the gene with an internal ribosomal entry site (IRES) into the U3 region or 3'UTR [20,21]. Here we utilize an Akv virus that delivers a transgene located in the U3 region. Because we flanked the structural genes of the virus with LoxP sites, the virus can delete its envelope gene and thus ought to be able to permit superinfection of cells that express Cre recombinase. Furthermore, because the virus can randomly distribute LoxP sites over the host genome, we investigated the potential of the virus for deletional mutagenesis.

Results and Discussion
Design of a replication competent superinfecting retroviral vector
The objective of this study was to engineer an Akv-based retroviral vector capable of superinfection, and to assess its potential for both transgene overexpression and deletional mutagenesis. Because we wanted the vector to undergo multiple rounds of replication, in particular to allow efficient delivery in animal model settings, we based our design upon a virus with the full set of genes needed for replication. To easily monitor infection and transgene expression by flow cytometry, we inserted the Enhanced Yellow Fluorescent Protein (EYFP) downstream of an IRES into the U3 region of the LTR (Fig. 1). To allow superinfection, we flanked the envelope gene with two LoxP sites so that it could be excised by Cre recombinase. The exact positioning of the LoxP sites (Fig. 1) was based upon the following considerations. If the LoxP site were to be placed in the LTR [22], Cre mediated deletion of the proviral LoxP fragment would result in a solo LTR, from which transgene expression would be difficult to achieve. We therefore inserted the downstream LoxP site at a position 9 bp after the TAA stop codon of *env*, just before the LTR. Because of the overlapping regions of *gag-pol* and *env*, it would be complicated to engineer the upstream LoxP site into the viral genome immediately in front of the coding sequence of *env* without compromising proper expression of viral proteins. We therefore placed the second LoxP site upstream of *gag*, but downstream of the major *cis*-elements for RNA packaging [23]. For convenience of construction, we chose position 413 of the viral RNA, 232 nt upstream of the ATG start codon of *gag*.

Normal NIH/3T3 cells infected with Akv-Y (virus containing EYFP, but no LoxP sites) and Akv-2XY (virus containing EYFP and two LoxP sites) are expected to be resistant to superinfection, since infected cells constantly produce envelope protein (Fig. 2). Consistent with this notion, the

intensity of EYFP fluorescence of the infected cells was unchanged over a period of more than two months (data not shown).

The viral titer from PhoenixEco cells, a retroviral packaging line, used to initially infect the NIH/3T3 cells corresponded to a multiplicity of infection of <0.05. Because Akv-Y and Akv-2XY were fully competent viruses, the NIH/3T3 population infected by PhoenixEco-generated viral supernatant was able to propagate the infection to the rest of the culture. Because of the low multiplicity of infection and the stable transgene expression level over time, the superinfection resistant population can be assumed to contain close to just one active copy of the EYFP gene per cell. Therefore, the EYFP fluorescence from NIH/3T3 cells infected by Akv-Y or Akv-2XY, respectively, served as a reference for quantitative comparison in this study.

Cre-mediated inactivation of env production
Clone GC4 is an NIH/3T3-derived cell clone transduced with a replication-defective retroviral vector that encodes a fusion protein consisting of EGFP and Cre recombinase (see Materials and Methods). GC4 cells were infected by adding supernatant containing the complete retroviruses, Akv-Y or Akv-2XY. After 10 days in culture, the infection by Akv-Y had spread to nearly all GC4 cells. This can be deduced from the fact that almost all cells expressed viral envelope protein on the surface and EYFP (Fig. 3a). Since the EYFP intensity of the Akv-Y infected GC4 population did not increase appreciably over time, this indicated that the GC4 cells resisted superinfection.

In contrast, after 10 days only 6.8 % of the GC4 culture infected with Akv-2XY expressed EYFP (Fig. 3b; Table 1). Assuming a Poisson distribution of infection, the low multiplicity of infection allowed us to calculate that the infected population contained one active copy of the EYFP gene per cell (Table 1). The lack of surface envelope protein by this same population (Fig. 3b) suggested that GFP-Cre had excised the envelope gene, and thus had inactivated the virus' ability to replicate. Because almost no cells on days 5 and 10 expressed surface envelope, it is likely that the Cre-mediated deletion occurred quickly after reverse transcription and rendered the virus incapable of replication. However, after 16 days some envelope expressing cells were detected, suggesting that a mutant, either viral or cellular, had emerged (Fig. 3b).

Infection by mixed culture
Because our LoxP containing virus was so quickly made replication-incompetent in Cre recombinase active cells, we used infected NIH/3T3 cells ("feeder cells") as a steady generator of viruses and co-cultured them with initially uninfected GC4 cells ("host cells"). Because they fluoresce
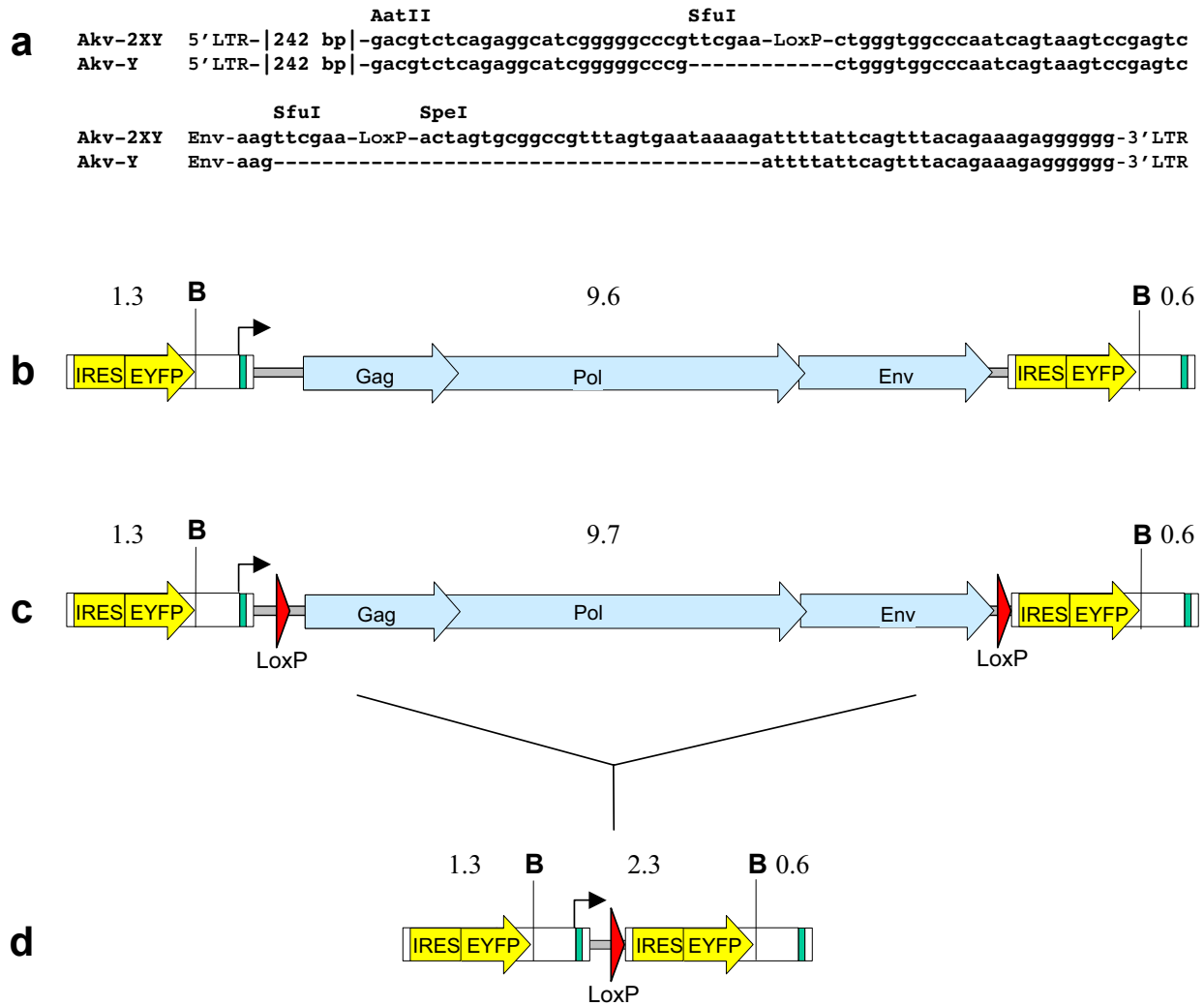
a

```
                              AatII                        SfuI
Akv-2XY   5'LTR-|242 bp|-gacgtctcagaggcatcgggggcccgttcgaa-LoxP-ctgggtggcccaatcagtaagtccgagtc
Akv-Y     5'LTR-|242 bp|-gacgtctcagaggcatcgggggcccg------------ctgggtggcccaatcagtaagtccgagtc

          SfuI        SpeI
Akv-2XY   Env-aagttcgaa-LoxP-actagtgcggccgtttagtgaataaaagatttattcagtttacagaaagagggggg-3'LTR
Akv-Y     Env-aag-------------------------------------attttattcagtttacagaaagagggggg-3'LTR
```


Figure 1

Akv constructs. (a) LoxP sites in Akv-2XY were inserted between the 5'LTR and *gag* and between the *env* and the 3'LTR. (b) The Akv-Y provirus contained IRES-EYFP in the U3 region of the LTR. (c) The Akv-2XY provirus contained IRES-EYFP and two LoxP sites flanking the Akv structural genes. (d) Cre-mediated recombination of the LoxP sites in Akv-2XY results in deletion of the viral genes. Each of the proviruses contained two BsrG1 sites, B. The sizes (kb) of DNA segments generated by BsrG1 are above each construct.

green, the host cells can be distinguished from the feeder cells by flow cytometry.

The cultures were seeded with 10% NIH/3T3 cells stably infected with Akv-Y or Akv-2XY and 90% GC4 cells. After 10 days, the GC4 cells were almost completely infected by Akv-Y, as shown by envelope and EYFP expression (Fig. 4a). In this co-culture, the stable EYFP fluorescence over

time, and the near equal EYFP fluorescence between the host and feeder populations demonstrated superinfection resistance. In contrast, the EYFP expression of GC4 cells infected with Akv-2XY increased over a period of 16 days, after which the average EYFP intensity was forty-three times greater than in the feeder cells (Fig. 4b; Table 1). This demonstrated that Akv-2XY was able to superinfect cells with Cre recombinase activity. We estimated that
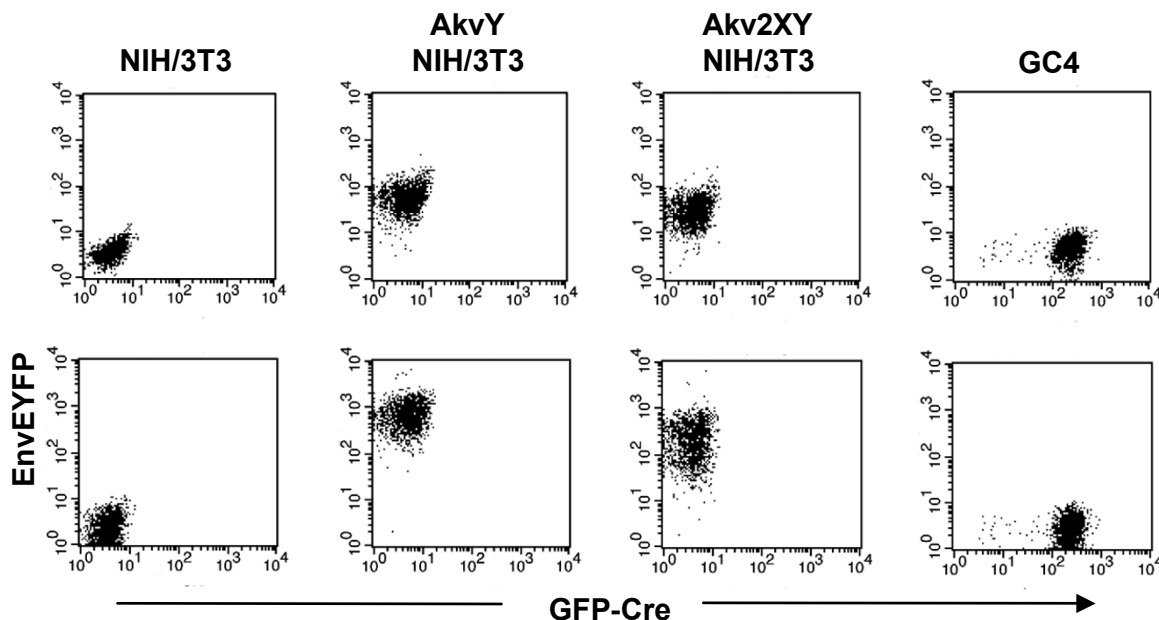
Figure 2
Flow cytometry analysis of control cells, demonstrating GFP-Cre, EYFP, and surface envelope (Env) expression. These control cells consist of NIH/3T3 cells uninfected or infected with AkvY or Akv2XY and GC4, a GFP-Cre expressing NIH/3T3 line expanded from a single clone.

after 16 days there were, on average, 20 or more actively expressed provirus copies per cell (Table 1). We also noted that, unlike in the supernatant infection experiments, there was some envelope protein present on the surface of the GC4 cells. However, the Env expression was considerably lower than that of the feeder cells. If indeed the envelope gene was deleted from provirus and not produced from the host cell, the envelope protein might arise from "shedded" Env-protein or particles bound or fused to the cell membrane [24]. This would likely require the high concentrations of virus maintained by the mixed culture.

One advantage of using GFP-Cre recombinase fusion protein was that we could monitor the population dynamics of the feeder and host cell populations using flow cytometry. While the Akv-Y mixed culture essentially maintained the initial 10% NIH/3T3 and 90% GC4 cell distribution, in the Akv-2XY mixed culture, the GC4 fraction steadily decreased and after 16 days consisted of only 2.8% of the culture. The multiple integrations of the Akv-2XY provirus apparently bestowed a considerable growth disadvantage. This disadvantage could be due to the bur-

den of the additional proviruses. In addition, when two LoxP containing proviruses integrate on the same chromosome, a deletion of genomic DNA from the host cell can occur (see below). These deletions could kill or slow the growth of a cell. Translocations, though reportedly infrequent in Cre-expressing cells with multiple LoxP sites [25], and inversions of genomic DNA may also adversely affect the growth of the host GC4 cells.

Analysis of subclones

After either 10 or 16 days in the mixed culture, infected cells were subcloned by single cell fluorescence activated cell sorting (FACS). The subclones were expanded and then analyzed by flow cytometry (Fig. 5). The fluorescence intensity was correlated to that of the unsubcloned cells that were presumed to have on average one expressed EYFP gene per cell (Table 2).

After 39 days in culture (10 or 16 days in the mixed culture followed by 29 or 23 days as a subcloned culture) three of the Akv-Y subclones, AYGC.2A2 (Fig. 5; Table 2), AYGC.2A3 and AYGC.2C3 (Table 2), had EYFP intensities
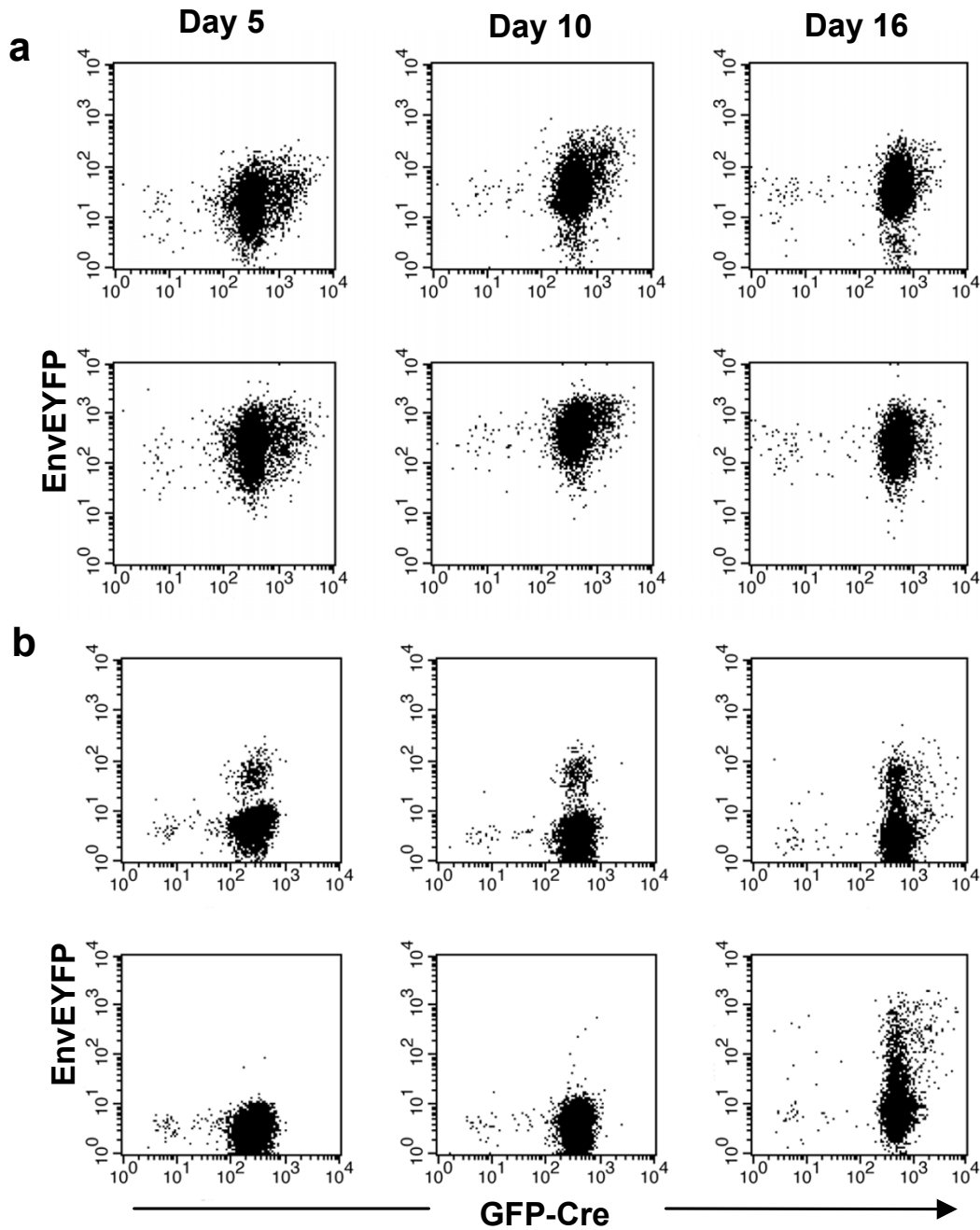
Figure 3
Flow cytometry analysis of GC4 cells infected by virus containing supernatant to measure GFP-Cre, EYFP, and surface envelope (Env) expression. Infection by (a) Akv-Y and (b) Akv-2XY; shown are days 5, 10, and 16. Y-axis, fluorescence intensity on a logarithmic scale of EYFP and Env, respectively. X-axis, fluorescence intensity of GFP-Cre.

Table 1: Proviral expression and estimated copy numbers in infected GC4 cell populations.

Virus	Infection	time (days)	GFP-Cre+ (%)	GFP-Cre+ EYFP+ (%)	GFP-Cre+ EYFP- (%)	Ratio of intensities[a]	Est. # of expressed copies
						(GC4 EYFP)/(NIH/ 3T3 EYFP)	(GC4 EYFP)/(one expressed EYFP)
Akv-Y	supernatant	5	99.2	78.7	20.5	0.5	1[b]
Akv-Y	supernatant	10	99.6	95.7	3.9	0.9	1[b]
Akv-Y	supernatant	16	99.1	95.7	3.5	1.2	1[b]
Akv-2XY	supernatant	5	99.1	5.4	93.7	1.8	
Akv-2XY	supernatant	10	99.6	6.8	92.8	1.7	1[c]
Akv-2XY	supernatant	16	99.4	12.7	86.8	1.6	
Akv-Y	mixed culture	5	88.5	82.5	6.0	0.9	1[b]
Akv-Y	mixed culture	10	93.2	91.8	1.4	0.9	1[b]
Akv-Y	mixed culture	16	92.6	91.4	1.2	1.1	1[b]
Akv-2XY	mixed culture	5	69.3	46.7	22.7	3.9	2[d]
Akv-2XY	mixed culture	10	46.4	45.5	0.9	12.8	8[d]
Akv-2XY	mixed culture	16	2.8	2.8	0.0	43.1	20[d]

[a] For a given virus, the ratio was calculated with fluorescence intensities of that virus to account for differences not related to copy number (see Materials and Methods). [b] One expressed EYFP gene based on unsubcloned NIH/3T3 EYFP fluorescence intensity. [c] One expressed EYFP gene based on the Poisson distribution calculation, using the EYFP- value of non-infected cells. [d] One expressed EYFP gene based on single expressed EYFP fluorescence intensity from supernatant infected culture.

corresponding to the reference population of one EYFP provirus expressed per cell. One Akv-Y subclone, AYGC.3B2, from a cell population with high EYFP fluorescence (Region C of Fig. 5a), expressed EYFP with an intensity two to three times higher (Table 2). Each of the Akv-Y subclones expressed surface envelope protein at a high level, comparable to the unsubcloned, infected NIH/3T3 cells. This indicated that it was the host cell that expressed the envelope gene, directed by the integrated provirus.

In contrast, none of the Akv-2XY subclones had envelope on the cell surface, indicating that the envelope gene indeed had been excised by Cre/LoxP recombination (for example, clones 2B1 and 3B8 in Fig. 5). This also supports the assertion that the low amount of envelope observed on GC4 cells in the mixed culture was not due to envelope gene transcription by the provirus [24]. After 39 days in culture, the average fluorescence intensity of EYFP from the Akv-2XY subclones was 30 to 50 times greater than the unsubcloned, infected NIH/3T3 population. We estimated a range of 18–29 expressed copies of the transgene in the subclones (Table 2).

Analysis of integration sites
Because each cell of a subclone culture should have the same integration site, the sites of various clones could be analyzed by Southern blotting (Fig. 6a). Genomic DNA digested with BsrG1 was examined using EYFP as a probe.

The 9.6 kb bands from the Akv-Y subclones represent the BsrG1 fragment of the original provirus containing the 3' copy of the EYFP gene (Fig. 1b). After Cre-recombinase-mediated deletion in the Akv-2XY subclones, the 9.6 kb fragment is converted into the 2.3 kb BsrG1 fragment of the recombined provirus, which still contains the 3' copy of the EYFP gene (Fig. 1d). Because equal amounts of DNA were loaded into each lane, the greater intensity of the 2.3 kb bands shows that the Akv-2XY subclones contained a greater number of proviruses, providing further evidence that the virus is capable of superinfection. Since bands smaller and larger than 2.3 kb had lower intensities than the 2.3 kb bands, the intensity difference is not likely due to incomplete transfer to the membrane. It is possible that the 2.3 and 9.6 kb bands also represent unintegrated provirus, as unintegrated DNA may be found in superinfected cells [26,27]. However, we believe this to be less likely, since such complexes would have to survive for many cell generations (cells were grown for over 2 months before the Southern blots were performed) after we subcloned the cells (after subcloning, cells were no longer being infected with new virus).

The other bands on the gel represent the fragment containing part of the 5' LTR and the adjacent host genomic DNA, with each band marking a unique integration site. Due to their abundance (>17) the unique bands from the
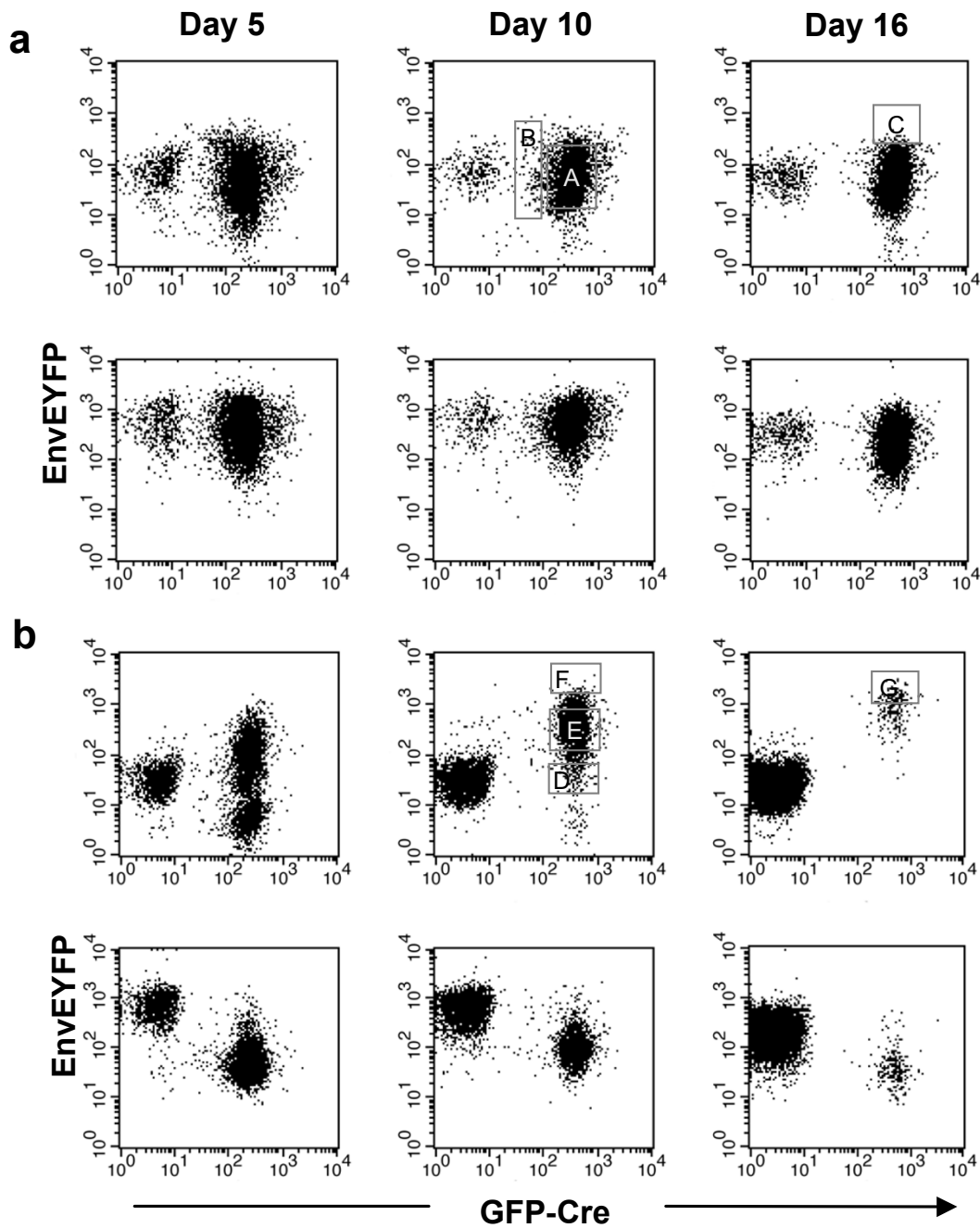
Figure 4
Flow cytometry analysis of GC4 cells infected by virus-producing NIH/3T3 cells in a mixed culture. GFP-Cre, EYFP, and surface envelope (Env) expression were measured. The GC4 populations were GFP-Cre+ and virus producing NIH/3T3 were GFP-Cre-. Infection by (a) Akv-Y and (b) Akv-2XY was monitored over 16 days. Subclones were isolated from boxed regions A-G by single-cell sorting.
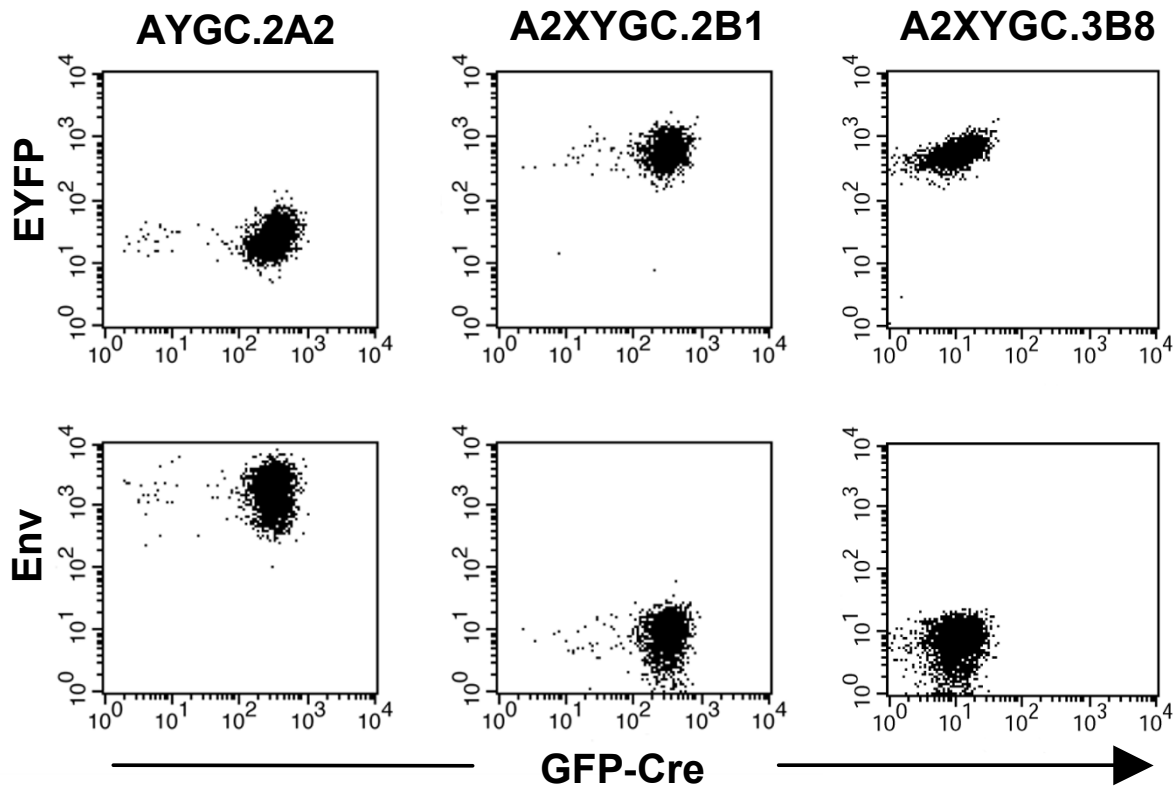
Figure 5
Flow cytometry analysis of clones of GC4 infected by Akv-Y and Akv-2XY. The profile of AYGC.2A2 was representative of all Akv-Y infected GC4 clones, expressing Env and EYFP. The profile of A2XYGC.2B1 was representative of all Akv-2XY infected GC4 clones, except A2XYGC.3B8, in that they expressed high EYFP levels and GFP-Cre but no Env. Of all the clones isolated from the Akv-2XY infected GC4 cells, A2XYGC.3B8 did not demonstrate GFP-Cre fluorescence comparable to the GC4 mother cell line or the other clones.

Table 2: Proviral expression and estimated copy numbers in GC4 subclones

Subclone	Virus	region sorted[b]	Ratio of intensities	Est. # of expressed copies[a]
			(GC4 EYFP)/(NIH/3T3 EYFP)	(GC4 EYFP)/(One expressed EYFP)
AYGC.2A2	Akv-Y	A	0.6	1
AYGC.2A3	Akv-Y	A	1.0	1
AYGC.2C3	Akv-Y	B	0.6	1
AYGC.3B2	Akv-Y	C	2.5	2–3
A2XYGC.2A2	Akv-2XY	D	49.8	29
A2XYGC.2B1	Akv-2XY	E	45.7	27
A2XYGC.2C1	Akv-2XY	F	34.9	21
A2XYGC.3B1	Akv-2XY	G	43.5	26
A2XYGC.3B3	Akv-2XY	G	30.1	18
A2XYGC.3B8	Akv-2XY	G	42.8	25

[a] For the Akv-2XY virus, this value accounts for the difference in EYFP intensity between the unrecombined and recombined provirus. [b] Sorted regions boxed on Fig. 4.
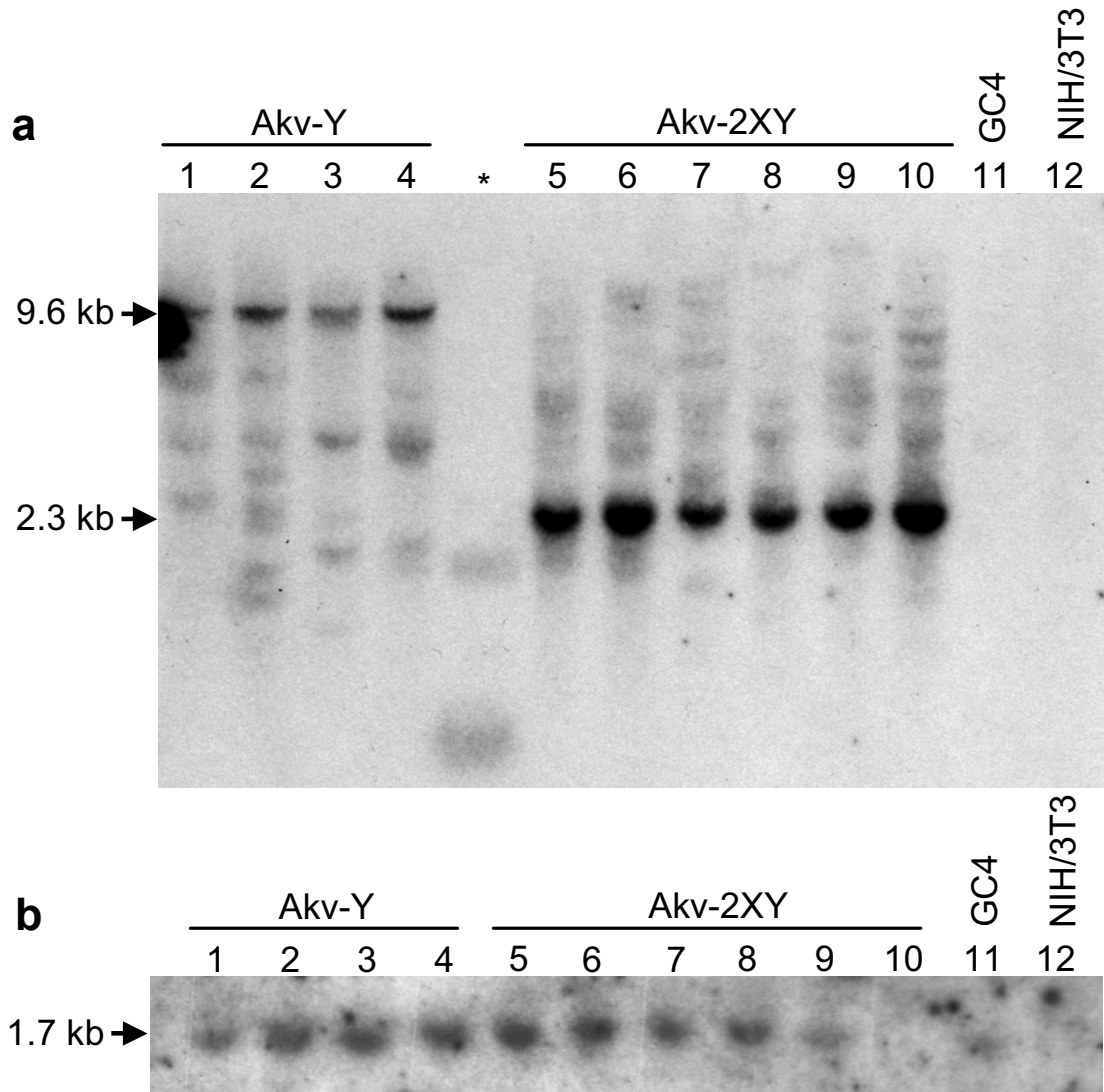
Figure 6
Southern analysis of subclones. Cellular DNA digested with BsrG1 and visualized using an (a) EYFP and (b) Cre probe. Lanes contained DNA (8 μg/lane) from clones (1) AYGC.2A2, (2) AYGC.2A3, (3) AYGC.2C3, (4) AYGC.3B2, (5) A2XYGC.2A2, (6) A2XYGC.2B1, (7) A2XYGC.2C1, (8) A2XYGC.3B1, (9) A2XYGC.3B3, and (10) A2XYGC.3B8, and cell lines (11) GC4 and (12) NIH/3T3. (*) DNA size marker, 1-kb ladder. When more DNA (15 μg/lane) was applied, more EYFP-probed bands can be discerned (not shown). The Cre probe hybridizes to the vector-genomic DNA junction.

Akv-2XY subclones were not easily counted. However, there are more bands present in the clones infected with the LoxP containing virus than with the virus without LoxP.

Consistent with its high EYFP intensity (Table 2), subclone AYGC.3B2 showed two, or perhaps three, unique bands on the Southern blot, possibly representing the number of expressed EYFP genes. In Akv-Y subclones AYGC.2A2, AYGC.2A3, and AYGC.2C3, the Southern blot

showed that there were more than one integration site, but their EYFP fluorescence intensities suggested that these subclones contained only one active EYFP gene per cell. This discrepancy might be due to a position effect on expression [28], thus resulting in a different gene copy prediction, or be due to silencing of retroviral genes [29,30], including the envelope gene.

Evidence for deletions generated in the cellular DNA
By delivering LoxP sites to a genome, the LoxP virus may create large gene deletions, translocations, and inversions. Because the superinfecting retrovirus is able to deliver numerous copies of LoxP, these mutations can be created proficiently. The high multiplicity of infection will favor the integration of two LoxP sites (of the same directional orientation) on the same chromosome and this would result in gene deletion. Because large deletions of DNA (i.e. several megabases) can occur without lethal effects [25,31,32] the superinfecting retrovirus can be used to tag genes where a phenotype results from gene inactivation, e.g. a tumor suppressor gene. Indeed, we found by flow cytometry that one Akv-2XY infected clone, A2XYGC.3B8, lacked GFP-Cre, yet EYFP expression indicated multiple expressed proviral copies (Fig. 5). This may imply that the deletion of the DNA segment containing Cre recombinase must have occurred *after* deletion of the proviral segments containing the viral structural genes. Consistent with the notion that there was deletion leading to the absence of expression, rather than silencing of the gene, on a Southern blot hybridized with a Cre recombinase probe, there was no band detected in the subclone (Fig. 6b, lane 10), while a 1.7 kb band was visible in other cells of GC4 cell lineage (Fig. 6b). Because superinfection did lead to a growth disadvantage before subcloning, it is possible that cells inactivating the GFP-Cre gene, and thus stopping the superinfection process, might be selected for and enriched in the culture. It is also possible that whole chromosomal deletion led to loss of the GFP-Cre; this may or may not have been mediated by LoxP recombination.

We further searched for deletions in these cell lines using array-based comparative genome hybridization (ACGH) [33,34]. ACGH can scan an entire genome for copy number changes including single copy losses and gains. The resolution of ACGH is limited only by the genomic spacing of the clones that are present on the arrays. For the experiments reported here we used an average resolution of 1 clone per 4–5 Mb across all chromosomes, with some regions of the genome covered at much higher resolution. Of the six Akv-2XY clones analyzed, four contained intrachromosomal deletions large enough to be detected with our arrays (Fig. 7), while none of the four Akv-Y clones contained evidence for such deletions (not shown here; array CGH data from all cell lines is available at http://cc.ucsf.edu/gray/public). In addition to the normal, con-

ventional genome comparison (Fig. 7, left side), deletions were confirmed by performing hybridizations in a reverse orientation (Fig. 7, right side) such that deletions appeared as gains. The sizes of the deletions varied. Since one of the deletions occurred in a region of the genome for which we had a high density of arrayed clones, we clearly defined it as an 18 Mb deletion (Fig. 7a). Other deletions were less well defined but ranged in sizes from 46–100 Mb (Fig. 7b), 6–30 Mb (Fig. 7c), and 0.2–21 Mb (Fig. 7d) – the limit of resolution for the arrays used here. In addition to the intrachromosomal deletions, whole chromosome losses were detected relative to the mother cell line GC4. These whole chromosome losses likely occurred during propagation of the polyploid GC4 cell line *in vitro* and are not likely due to a Cre mediated event.

Conclusion
Gene delivery by LoxP viruses
In our experiments the Cre recombinase activity was so high that the envelope gene was excised soon after infection. As a result, the integrated proviruses not only allowed the host cell to be reinfected, but also abolished viral replication. In cases where extinguishing the virus is desired, this is clearly advantageous. Alternatively, it is possible that there exists a level of Cre activity such that some virus is still generated before recombination while superinfectibility is maintained. Under such circumstances the virus would be both replication competent and capable of superinfection. Furthermore, in a given transgenic mouse strain, usually not all cells express Cre and such cells can serve as a reservoir for virus production. In another modification of the virus, it may be possible to engineer a superinfecting virus to also deliver Cre. In this case, recombinase producing cells or organisms would not be required for superinfection.

The LoxP virus could have applications when a high number of integrations of a transgene, provirus, or LoxP site are desired. As done in this study, cell lines expressing multiple copies of a transgene can be created by infection with the engineered virus. A superinfecting virus might also be useful in delivering a transgene to animals – for example, to chickens where exogenous protein can be produced in eggs [35,36] and to mammals where exogenous protein can be produced in milk [37,38]. The virus may be useful in delivering genes to germline or somatic cells. Yet a LoxP based virus would probably be undesirable for human gene therapy; even if Cre could be delivered with virus, the delivery of multiple LoxP might create harmful genomic deletions, inversions, and translocations. Furthermore, because this application utilizes a complete retrovirus, genome packaging constraints may limit the maximum size of the transgene.
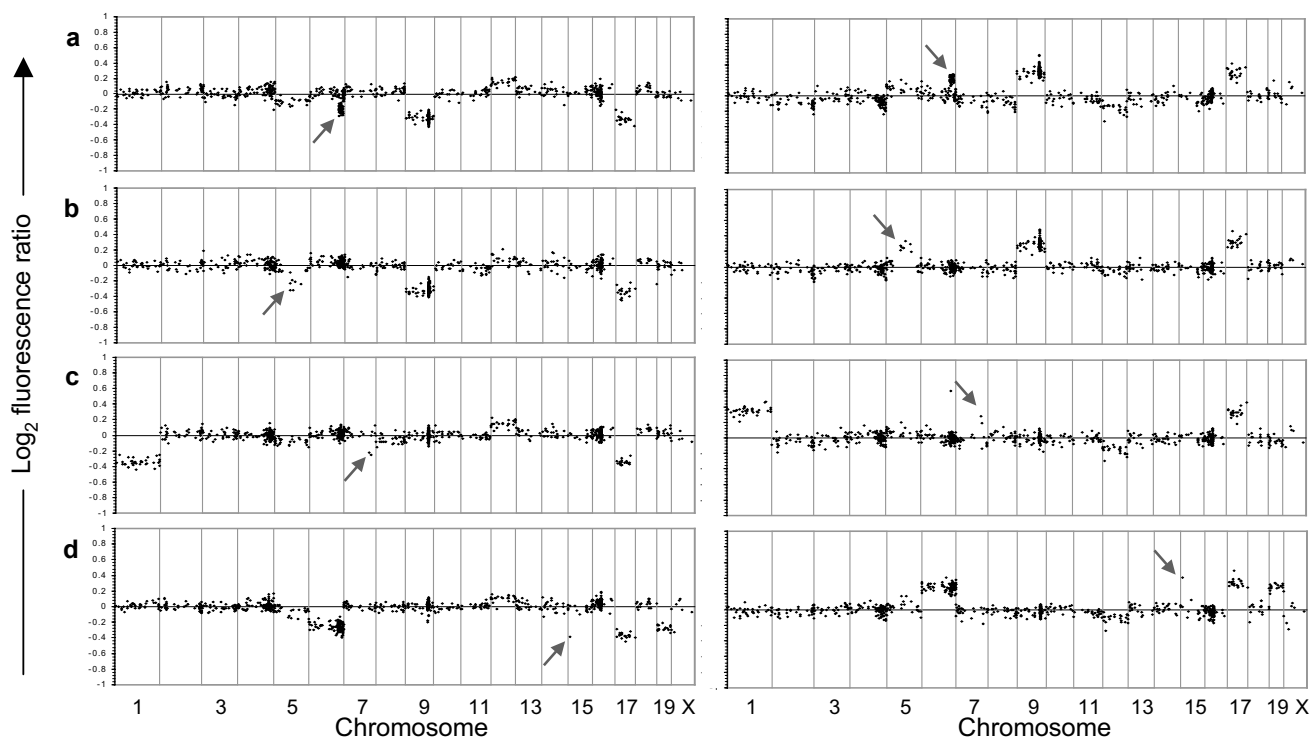
Figure 7
Array comparative genome hybridization showing copy number changes between the mother cell line (GC4), containing Cre recombinase, and four daughter clones infected with the LoxP virus Akv-2XY; (a) A2XYGC.2B1; (b) A2XYGC.2A2; (c) A2XYGC.3B1; (d) A2XYGC.2C1. Each cell line was analyzed twice, as conventional CGH hybridizations (left) and reverse array hybridizations (right) where deletions now appear as gains. Arrows indicate regions of intrachromosomal deletion. The cell lines may or may not have a genealogical relationship in regard to LoxP virus infection. Each black dot represents the average fluorescence ratio of 4 arrayed replicates plotted as function of their position in the genome based on Celera's assembly of the mouse genomic sequence; markers contained within each BAC were blasted against the genome assembly to establish the genomic coordinate. All data have been normalized such that the median fluorescence ratio is 0. Grey vertical bars represent chromosome boundaries. A two-fold reduction in fluorescence ratio (i.e. 1 versus 2 genomic copies) is represented by a value of -1 on this scale. Since these cell lines are aneuploid, ratios are suppressed such that single copy deletions are seen as 2 versus 3 or 3 versus 4 genomic copies.

LoxP virus as a mutator
In addition to gene delivery, the superinfecting LoxP virus might also be useful as a mutagenesis and gene discovery tool. Retroviruses are currently used to tag and identify new genes of interest. The gene discovery technique relies on the fact that the provirus is integrated into random locations of the host genome. When by chance the provirus integrates near a gene, the viral promoter and/or enhancer may lead to over-expression of that gene. When the provirus integrates within a gene, gene expression can be disrupted. By screening infected cells or organisms for the desired phenotype, the gene of interest can be identified by virtue of the proviral tag. Retroviral tagging has been valuable in identifying developmental genes [39,40]

and oncogenes [41-44]. The superinfecting virus would be especially useful when investigating a phenotype that requires multiple genes. For example, tumor formation typically requires that multiple oncogenes be activated. Discovery of oncogenes by retroviral tagging depends on multiple provirus integrations, which require superinfection. In principle, the LoxP virus would be able to activate multiple oncogenes by superinfection, leading to faster tumor formation and to a more comprehensive oncogene discovery screen.

By delivering LoxP sites to a genome, the LoxP virus might also serve to create large gene deletions, translocations, and inversions. Because the superinfecting retrovirus is

able to deliver numerous copies of LoxP, these mutations can be created more proficiently. When used as a tagging tool, the superinfecting LoxP virus would serve as a novel mutator to identify new genes. For example, a high multiplicity of infection will favor the integration of two LoxP sites (of the same directional orientation) on the same chromosome and this would result in gene deletion. Because large deletions of DNA (i.e. several megabases) can occur without lethal effects [25,31,32], the superinfecting retrovirus could be used to tag genes where a phenotype results from gene inactivation. The ability to generate large deletions would allow both alleles of a gene to be inactivated, thus favoring the discovery of a desired gene (e.g. a tumor suppressor gene).

Methods
Vector construction
pAkv-Y
The IRES-EYFP (Internal Ribosomal Binding Site – Enhanced Yellow Fluorescent Protein) fragment was PCR amplified from pIRES-EYFP (Clontech) and cloned into the CelII site in the U3 region of the 3' LTR of pAkv [20]. Plasmid pAkv contains a partially truncated Akv provirus and is able to produce the entire Akv retrovirus when transfected into cells.

pAkv-2XY
By using a primer containing an overhanging LoxP sequence, a LoxP site (ATAACTTCGTATAGCATACAT-TATACGAAGTTAT) was added 5' to the 3' Akv LTR by PCR amplification from pAkv. The PCR fragment was cloned into the SfuI and SdaI sites of pAkvEn-EGFP [45], to create pAkv-EndEnvLoxP2. This plasmid contained the Akv provirus with a LoxP site between the envelope gene and the 3' LTR. Because the LoxP site of pAkv-EndEnvLoxP2 was flanked by repetitive sequences, and because these repetitive sequences might later lead to homologous recombination and subsequent deletion of the LoxP site, the repetition in sequence was shortened by excising the PacI-PmlI fragment of pAkv-EndEnvLoxP2, creating pAkv-EndEnvLoxP3. By using a primer containing an overhanging LoxP sequence, a LoxP site was added between the 5'LTR and gag gene by PCR amplification of the AatII-BsiWI fragment of pAkvPro. The PCR fragment was cloned into the AatII and BsiWI sites of pAkvEndEnvLoxP to create pAkv-2LoxPs3. The CelII fragment containing IRES-EYFP from pAkv-U3EYFP was cloned into the CelII site of pAkv-2LoxPs3, creating pAkv-2LoxPs3-U3EYFP, or, for short, pAkv-2XY (Fig. 1). (Interestingly, when a LoxP site was added directly after the stop codon of the gene (Env-TAA-LoxP), the engineered construct was unable to produce a functional virus. This suggests that the nucleotide regions flanking the 3'LoxP site in Akv-2XY are important for infection.).

Viral infection
PhoenixEco packaged viruses
Plasmids pNIT-GFP-Cre (GFP-Cre in retroviral vector NIT [46]), pAkv-Y, and pAkv-2XY were introduced into PhoenixEco cells by using Fugene transfectant (Roche Biosciences). After 3 days, PhoenixEco culture supernatant containing packaged retroviral genomes were passed through a 0.45 μm filter, supplemented with 4 μg/ml polybrene and added to NIH/3T3 cells. The culture medium was replenished with fresh DMEM after 8–12 h. The NIH/3T3 cells infected with Akv-Y and Akv-2XY were kept in culture until all cells were infected. These cells served to produce the viruses for the supernatant and mixed culture infection experiments. The NIT-GFP-Cre infected NIH/3T3 cells were isolated by fluorescence activated cell sorting (FACS) and subcloned by single-cell sorting. The clone GC4 produced a fusion protein consisting of Enhanced Green Fluorescence Protein (EGFP) and Cre recombinase [47], and it exhibited both EGFP fluorescence and recombinase activity. When grown without other cells or viruses in the same culture, GC4 demonstrated no silencing or inactivation of fluorescence after more than 6 months.

Infection with supernatant
Akv-Y and Akv-2XY infected NIH/3T3 were inoculated at 20–40% confluency in a 175 cm^2 tissue culture flask containing 40 ml of media. After 2 days, 10 ml of virus containing media were passed through a 0.45 μm filter and combined with 10 ml fresh media and 10^5 trypsinized GC4 cells in an 83 cm^2 flask.

Mixed culture infection
9×10^4 trypsinized GC4 cells were combined with 10^4 trypsinized Akv-Y or Akv-2XY infected NIH/3T3 cells. These cells were co-cultured in an 83 cm^2 flask.

Flow cytometry
The virus infection was monitored by flow cytometry. Cells were stained with monoclonal rat antibody to viral envelope (MAb 83A25; of IgG2A subclass) [48], and then with allophycocyanin (APC) conjugated polyclonal goat anti-IgG (PharMingen) as a secondary antibody. Thus, the infected cell population was quantified by measuring APC fluorescence of the stained cells and EYFP. GFP-Cre fluorescence was also measured. GFP-Cre and EYFP could be distinguished on a on a Becton Dickinson FACS Vantage SE equipped with a 200 mW 488 nm laser and two bandpass filters (510/10 nm, GFP-Cre; 550/30 nm, EYFP). Infected cells were subcloned by single cell FACS sorting.

Provirus copy number was estimated from fluorescence intensities by dividing the intensity of a given cell population by a reference value of the intensity of a population known or presumed to contain one copy of active

provirus per cell. The intensity of EYFP fluorescence in NIH/3T3 infected with Akv-Y was consistently double that of cells infected with Akv-2XY (data not shown). Apparently, the EYFP expression of Akv-2XY is slightly hindered by the two added LoxP sites. The recombined Akv-2XY provirus showed a consistently higher (approximately 1.7×) fluorescence than the unrecombined provirus (Table 1), suggesting that excision of the retroviral genes augmented expression of the EYFP transgene. These differences in expression were taken into account when estimating the copy number of expressed genes. Because EYFP fluorescence intensity might not correlate linearly with gene expression, copy number estimations based on fluorescence likely provide lower limit values. Because the expression of an EYFP gene from a clonal population will vary to some degree depending on its integration site, and because different clones will have different integration sites, the relative fluorescence intensity may be affected by clonal variation.

Comparative Genomic Hybridization
Preparation of the BAC arrays, hybridization, and data processing was done as described previously [33]. BAC DNA was PCR amplified using a 5'-amine linked degenerate primer, printed on to amine reactive slides, and slides were processed following manufacturer's recommendations (Motorola). Genomic DNA extracted from Akv-Y or Akv-2XY cell lines (test sample) was labeled with Cy3 by random primed labeling (Bioprime, Invitrogen) and genomic DNA from GC4 cell line (reference sample) was labeled with Cy5. Cy dyes were reversed for reverse hybridizations. Labeled test and reference genomic DNA samples were co-hybridized to the arrays in the presence of excess repeat-blocking Cot-1 DNA (Invitrogen) for two nights, washed, and imaged using a custom CCD imaging system [49]. Image analysis was performed using custom software [33,50]. Detailed genomic DNA extraction and hybridization protocols are available at http://cc.ucsf.edu/gray/public.

Competing interests
None declared.

References
1. Chesebro B, Wehrly K: **Different murine cell lines manifest unique patterns of interference to superinfection by murine leukemia viruses.** *Virology* 1985, **141:**119-129.
2. Mitchell T, Risser R: **Interference established in mice by infection with Friend murine leukemia virus.** *J Virol* 1992, **66:**5696-5702.
3. Rubin H: **A Virus in Chick Embryos Which Induces Resistance in Vitro to Infection with Rous Sarcoma Virus.** *Proc Natl Acad Sci U S A* 1960, **46:**1105-1119.
4. Kai K, Sato H, Odaka T: **Relationship between the cellular resistance to Friend murine leukemia virus infection and the expression of murine leukemia virus-gp70- related glycoprotein on cell surface of BALB/c-Fv-4wr mice.** *Virology* 1986, **150:**509-512.
5. Federspiel MJ, Crittenden LB, Hughes SH: **Expression of avian reticuloendotheliosis virus envelope confers host resistance.** *Virology* 1989, **173:**167-177.
6. Bosselman RA, Hsu RY, Bruszewski J, Hu S, Martin F, Nicolson M: **Replication-defective chimeric helper proviruses and factors affecting generation of competent virus: expression of Moloney murine leukemia virus structural genes via the metallothionein promoter.** *Mol Cell Biol* 1987, **7:**1797-1806.
7. Delwart EL, Panganiban AT: **Role of reticuloendotheliosis virus envelope glycoprotein in superinfection interference.** *J Virol* 1989, **63:**273-280.
8. Stevenson M, Meier C, Mann AM, Chapman N, Wasiak A: **Envelope glycoprotein of HIV induces interference and cytolysis resistance in CD4+ cells: mechanism for persistence in AIDS.** *Cell* 1988, **53:**483-496.
9. Maddon PJ, Dalgleish AG, McDougal JS, Clapham PR, Weiss RA, Axel R: **The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain.** *Cell* 1986, **47:**333-348.
10. Geleziunas R, Bour S, Wainberg MA: **Human immunodeficiency virus type 1-associated CD4 downmodulation.** *Adv Virus Res* 1994, **44:**203-266.
11. Bestwick RK, Kozak SL, Kabat D: **Overcoming interference to retroviral superinfection results in amplified expression and transmission of cloned genes.** *Proc Natl Acad Sci U S A* 1988, **85:**5404-5408.
12. Kozak SL, Kabat D: **Ping-pong amplification of a retroviral vector achieves high-level gene expression: human growth hormone production.** *J Virol* 1990, **64:**3500-3508.
13. Nielsen AL, Pallisgaard N, Pedersen FS, Jorgensen P: **Murine helix-loop-helix transcriptional activator proteins binding to the E-box motif of the Akv murine leukemia virus enhancer identified by cDNA cloning.** *Mol Cell Biol* 1992, **12:**3449-3459.
14. Lovmand S, Kjeldgaard NO, Jorgensen P, Pedersen FS: **Enhancer functions in U3 of Akv virus: a role for cooperativity of a tandem repeat unit and its flanking DNA sequences.** *J Virol* 1990, **64:**3185-3191.
15. Morrison HL, Dai HY, Pedersen FS, Lenz J: **Analysis of the significance of two single-base-pair differences in the SL3-3 and Akv virus long terminal repeats.** *J Virol* 1991, **65:**1019-1022.
16. Etzerodt M, Mikkelsen T, Pedersen FS, Kjeldgaard NO, Jorgensen P: **The nucleotide sequence of the Akv murine leukemia virus genome.** *Virology* 1984, **134:**196-207.
17. Jensen NA, Jorgensen P, Kjeldgaard NO, Pedersen FS: **Mammalian expression-and-transmission vector derived from Akv murine leukemia virus.** *Gene* 1986, **41:**59-65.
18. Buchhagen DL, Pedersen FS, Crowther RL, Haseltine WA: **Most sequence differences between the genomes of the Akv virus and a leukemogenic Gross A virus passaged in vitro are located near the 3' terminus.** *Proc Natl Acad Sci U S A* 1980, **77:**4359-4363.
19. Wang H, Kavanaugh MP, North RA, Kabat D: **Cell-surface receptor for ecotropic murine retroviruses is a basic amino-acid transporter.** *Nature* 1991, **352:**729-731.
20. Jespersen T, Duch M, Carrasco ML, Warming S, Pedersen FS: **Expression of heterologous genes from an IRES translational cassette in replication competent murine leukemia virus vectors.** *Gene* 1999, **239:**227-235.
21. Logg CR, Logg A, Tai CK, Cannon PM, Kasahara N: **Genomic stability of murine leukemia viruses containing insertions at the Env-3' untranslated region boundary.** *J Virol* 2001, **75:**6989-6998.
22. Jin YF, Ishibashi T, Nomoto A, Masuda M: **Isolation and analysis of retroviral integration targets by solo long terminal repeat inverse PCR.** *J Virol* 2002, **76:**5540-5547.

23. Mikkelsen JG, Lund AH, Duch M, Pedersen FS: **Recombination in the 5' leader of murine leukemia virus is accurate and influenced by sequence identity with a strong bias toward the kissing-loop dimerization region.** *J Virol* 1998, **72:**6967-6978.

24. Bahrami S, Jespersen T, Pedersen FS, Duch M: **Mutational library analysis of selected amino acids in the receptor binding domain of envelope of Akv murine leukemia virus by conditionally replication competent bicistronic vectors.** *Gene* 2003, **315:**51-61.

25. Ramirez-Solis R, Liu P, Bradley A: **Chromosome engineering in mice.** *Nature* 1995, **378:**720-724.

26. Marquina S, Libonatti O, Ceballos A, Gomez Carrillo M, Martinez Peralta L, Rabinovich RD: **Different integrated and unintegrated HIV-1 DNA after superinfection and cell to cell transmission.** *Acta Physiol Pharmacol Ther Latinoam* 1997, **47:**245-250.

27. Yoshimura FK, Wang T, Nanua S: **Mink cell focus-forming murine leukemia virus killing of mink cells involves apoptosis and superinfection.** *J Virol* 2001, **75:**6007-6015.

28. Paludan K, Dai HY, Duch M, Jorgensen P, Kjeldgaard NO, Pedersen FS: **Different relative expression from two murine leukemia virus long terminal repeats in unintegrated transfected DNA and in integrated retroviral vector proviruses.** *J Virol* 1989, **63:**5201-5207.

29. Lund AH, Duch M, Pedersen FS: **Transcriptional Silencing of Retroviral Vectors.** *J Biomed Sci* 1996, **3:**365-378.

30. Palmer TD, Rosman GJ, Osborne WR, Miller AD: **Genetically modified skin fibroblasts persist long after transplantation but gradually inactivate introduced genes.** *Proc Natl Acad Sci U S A* 1991, **88:**1330-1334.

31. Zheng B, Sage M, Sheppeard EA, Jurecic V, Bradley A: **Engineering mouse chromosomes with Cre-loxP: range, efficiency, and somatic applications.** *Mol Cell Biol* 2000, **20:**648-655.

32. Su H, Wang X, Bradley A: **Nested chromosomal deletions induced with retroviral vectors in mice.** *Nat Genet* 2000, **24:**92-95.

33. Hodgson G, Hager JH, Volik S, Hariono S, Wernick M, Moore D, Nowak N, Albertson DG, Pinkel D, Collins C, Hanahan D, Gray JW: **Genome scanning with array CGH delineates regional alterations in mouse islet carcinomas.** *Nat Genet* 2001, **29:**459-464.

34. Snijders Antoine M., Nowak Norma, Segraves Richard, Blackwood Stephanie, Brown Nils, Conroy Jeffrey, Hamilton Greg, Hindle Anna Katherine, Huey Bing, Kimura Karen, Law Sindy, Myambo Ken, Palmer Joel, Ylstra Bauke, Yue1 Jingzhu Pearl, Gray Joe W., Jain Ajay N., Pinkel Daniel, Albertson Donna G.: **Assembly of microarrays for genome-wide measurement of DNA copy number.** *Nat Genet* 2001, **29:**263-264.

35. Harvey AJ, Speksnijder G, Baugh LR, Morris JA, Ivarie R: **Consistent production of transgenic chickens using replication- deficient retroviral vectors and high-throughput screening procedures.** *Poult Sci* 2002, **81:**202-212.

36. Harvey AJ, Speksnijder G, Baugh LR, Morris JA, Ivarie R: **Expression of exogenous protein in the egg white of transgenic chickens.** *Nat Biotechnol* 2002, **20:**396-399.

37. McKee C, Gibson A, Dalrymple M, Emslie L, Garner I, Cottingham I: **Production of biologically active salmon calcitonin in the milk of transgenic rabbits.** *Nat Biotechnol* 1998, **16:**647-651.

38. Prunkard D, Cottingham I, Garner I, Bruce S, Dalrymple M, Lasser G, Bishop P, Foster D: **High-level expression of recombinant human fibrinogen in the milk of transgenic mice.** *Nat Biotechnol* 1996, **14:**867-871.

39. Chen W, Burgess S, Golling G, Amsterdam A, Hopkins N: **High-throughput selection of retrovirus producer cell lines leads to markedly improved efficiency of germ line-transmissible insertions in zebra fish.** *J Virol* 2002, **76:**2192-2198.

40. Gaiano N, Amsterdam A, Kawakami K, Allende M, Becker T, Hopkins N: **Insertional mutagenesis and rapid cloning of essential genes in zebrafish.** *Nature* 1996, **383:**829-832.

41. Li J, Shen H, Himmel KL, Dupuy AJ, Largaespada DA, Nakamura T, Shaughnessy J. D., Jr., Jenkins NA, Copeland NG: **Leukaemia disease genes: large-scale cloning and pathway predictions.** *Nat Genet* 1999, **23:**348-353.

42. van Lohuizen M, Verbeek S, Scheijen B, Wientjens E, van der Gulden H, Berns A: **Identification of cooperating oncogenes in E mu-myc transgenic mice by provirus tagging.** *Cell* 1991, **65:**737-752.

43. Berns A: **Provirus tagging as an instrument to identify oncogenes and to establish synergism between oncogenes [published erratum appears in Arch Virol 1989;107(1-2):170].** *Arch Virol* 1988, **102:**1-18.

44. Korczak B, Robson IB, Lamarche C, Bernstein A, Kerbel RS: **Genetic tagging of tumor cells with retrovirus vectors: clonal analysis of tumor growth and metastasis in vivo.** *Mol Cell Biol* 1988, **8:**3143-3149.

45. Duch Mogens, Carrasco Maria L., Jespersen Thomas, Hansen Bettina D., Pedersen Finn Skou: **Transgene stability for three replication-competent murine leukemia virus vectors.** *Gene* 2004, **In Press:**.

46. Suhr ST, Senut MC, Whitelegge JP, Faull KF, Cuizon DB, Gage FH: **Identities of sequestered proteins in aggregates from cells with induced polyglutamine expression.** *J Cell Biol* 2001, **153:**283-294.

47. Gagneten S, Le Y, Miller J, Sauer B: **Brief expression of a GFP cre fusion gene in embryonic stem cells allows rapid retrieval of site-specific genomic deletions.** *Nucleic Acids Res* 1997, **25:**3326-3331.

48. Evans LH, Morrison RP, Malik FG, Portis J, Britt WJ: **A neutralizable epitope common to the envelope glycoproteins of ecotropic, polytropic, xenotropic, and amphotropic murine leukemia viruses.** *J Virol* 1990, **64:**6176-6183.

49. Pinkel Daniel, Segraves Richard, Sudar Damir, Clark Steven, Poole Ian, Kowbel David, Collins Colin, Kuo Wen-Lin, Chen Chira, Zhai Ye, Dairkee Shanaz H., Ljung Britt-marie, Gray Joe W., Albertson & Donna G.: **High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays.** *Nat Genet* 1998, **20:**209-211.

50. Jain Ajay N., Tokuyasu Taku A., Snijders Antoine M., Segraves Richard, Albertson Donna G., Pinkel Daniel: **Fully Automatic Quantification of Microarray Image Data.** *Genome Res* 2002:325-332.