Supplementary Material & Methods

LV Plasmids

SIN.LV.SF.PRE transfer vector originated from the SIN.LV.SF.GFP.PRE plasmid cut with BamHI/SalI to excise the GFP coding sequence. SIN.LV.SF vector was produced from SIN.LV.SF.GFP.PRE that was AgeI/EcoRI cut to excise the GFP.PRE sequence.

The INS.SIN.LV.SF:GFP.PRE construct was cloned starting from the pCCL.SIN.SF.GFP.PRE plasmid cut with EcoRI/SfiI, to excise the 3'region of the LV harboring delta-nef, the LV SIN LTR and the SV40 polyA site. Next, from the D.Caro4.23 plasmid digested with HincII/SfiI enzymes we isolate the band containing the 4xCTF insulator elements into the LV SIN LTR and the SV40 polyA. The insert and the plasmid backbone obtained from pCCL.SIN.SF.GFP.PRE digestion were blunt-filled, dephosphorylated and ligated to generate the INS.SIN.SF.GFP.PRE plasmid that was used for vector production.

Mouse treatment, sample collection and histopathology

All animal procedures were performed according to protocols approved by the Animal Care and Use Committee of the San Raffaele Institute (IACUC 320 and 451) and communicated to the Ministry of Health and local authorities according to Italian law. All mice were bred and kept in a dedicated pathogen-free animal facility, and were euthanized when they showed signs of severe sickness.

By grossly autoptical analysis, we could identify several tumor-infiltrating organ, in particular liver and spleen that were collected for immunofluorescence, histological and molecular analyses (DNA and RNA extractions). Histological analysis was also performed on lung, heart, thymus, sternum, kidney and gut. Hematoxylin-eosin staining was performed on 4-µm-thick sections of formalin-fixed, paraffin-embedded tissues. Specimens were evaluated in blinded fashion, independently by two investigators (F.S. and M.P.); in discordant cases, a consensus was reached after discussion at multi-head microscope. To better define the tumor origin, immune-histochemistry analyses was performed on selected samples. Antigen immune-localizations were performed using specific antibodies on formalin-fixed, paraffin-embedded 4 µm-thick sections after antigen retrieval with microwave using Tris-EDTA pH 9. Rat anti-human CD3 (AbD Serotec), rat anti-mouse B220 (AbD Serotec), rat anti-mouse F4/80 (AbD Serotec) and rabbit anti-human myeloperoxidase (DAKO), were used. The immunoreaction was revealed by byotinilated-conjugated anti-rat antibody (Vector), horseradish peroxidase (HRP)-conjugated streptavidin, and using 3,3 diaminobenzidine (DAB) as chromogen (Biogenex, SanRamon). Photomicrographs were taken using the AxioCam HRc (Zeiss) with the AxioVision System 6.4 (Zeiss).

Immunofluorescence Analysis

To check the tumor origin and presence of vector in tumor cells, we performed immune-fluorescence analyses on the spleen and the liver of selected samples. Samples were fixed in 4% paraformaldehyde, equilibrated in

sucrose gradients (10-20-30%) at 4°C, embedded in optimal-cutting-temperature compound for quick freezing and stained with the antibody. See Supplemental Methods for details on Antibody staining.

Samples were embedded in optimal-cutting-temperature compound for quick freezing. 16µm slice sections were blocked with PBS containing 1% bovine serum albumin, 5% fetal bovine serum and incubated with coniugated (Rat Anti-CD45 PE, 0.2 mg/ml (1:50) ; Rat Anti-CD11b PE 0.2 mg/ml (1:50); Rat Anti-F4/80 PE 0.25 mg/ml (1:50)) or unconiugated (Rabbit Anti-GFP, 2mg/ml (1:500)) primary antibodies, washed and when required stained with secondary conjugated antibody (Goat anti-Rabbit Alexa Fluor 488 (1:200)) (Molecular Probes). Cell nuclei were stained with TOPRO-3 (1:2000) after 1h of RNAse treatment (Molecular Probes). Stained sections were analyzed by three-laser confocal microscope (Radiance 2100; BioRad). Fluorescent signals from single optical sections were sequentially acquired and analyzed by Corel (Adobe).

Vector copy number analysis

VCN was determined as the ratio between the relative amounts of LV versus total DNA (number of diploid genome) evaluated by β-actin. A standard curve was made using dilutions from murine DNA with a known LV VCN determined by Southern blot Reactions were carried out according to manufacturer's instructions in Optical 96-well Fast Thermal Cycling Plates on ABI PRISM 7900 HT or on Viia7 Sequence Detector System and analyzed using the ABI Prism 7900 HT Sequence Detection System (Applera – Life Technologies). Sequences of primers and probes are available upon request. VCN was determined as the ratio between the relative amounts of LV versus total DNA (number of diploid genome) evaluated by β-actin. Since hematopoietic tumoral cells infiltrate the mouse liver parenchima that contain hepatocyte and stromal cells that could not be LV-transduced, the VCN that we measured may be an inaccurate estimation of the VCN of tumoral cells.

LAM-PCR and genomic integration site analysis

We used different amounts of DNA as template for LAM-PCR, according to the VCN that was detected in the sample by Q-PCR: 100ng if VCN<1; 50ng if VCN between 1 and 3; 10ng if VCN>3. LAM-PCR was initiated with a 25-cycle linear PCR and restriction digest using Tsp509I, or HpyCHIV4. LAM-PCR primers for LV were previously described. LAM-PCR amplicons were separated on spreadex gels (Elchrom Scientific) to evaluate PCR efficiency and the bands pattern for each sample. Products of the second exponential amplification were tagged with six nucleotide-long tags and then pooled and subjected to pyro-sequencing with the 454 GS Flx platform (Roche) by GATC Biotech.

Sequences were aligned to the mouse genome (assembly July 2007, mm9) using the NCBI BLAST genome browser (http://blast.ncbi.nlm.nih.gov/Blast.cgi) coupled to bioinformatic analyses. Identification of the nearest gene was performed by bioinformatic analyses.

RNA isolation and gene expression analysis

The relative expression level of each gene was calculated by the ΔΔCt method (Pfaffl, 2001), normalized to *Hprt* and *Pgk* expression (housekeeping gene controls), and represented as fold change relative to mock-transduced samples (calibrator). The qBase software program was used to measure the relative expression level for each gene.

Supplemental Statistical methods

Estimation of the Sample size

To detect reliable differences between treatment groups we estimated the sample size of each group based on α level of 0.05 and a power (β) of 80%.

The formula used to calculate the sample size was the following:

$$N = f(\alpha, \beta)\frac{\sigma^2}{(\mu a - \mu b)^2}$$

Where: N= sample size; f(α,β)=15.8 (for α=0.05 and β=80%); μa= median survival of control group; μb= median survival of treatment group.

Following these power calculations, to detect the genotoxicity of a vector treatment that will result on a 15-20% reduction of the median survival time of with respect to control mice, we will need to test between 14-25 mice for each treatment group.

Using these numbers we effectively detected significant differences in survival between different treatment groups.

In the table below, we estimated the number of mice that we should be tested hypothesizing a 5, 10, 15 and 20% reduction in median survival with respect to untreated control mice.

From these estimations, to observe significant differences between treatments that will result in a reduction in median survival of 5% with respect untreated controls will require a group of at least 225 mice. If the differences between treatments are of 10% then the number of mice per group will decrease to 56.

Genotoxicity	Median survival	n
5%	221.4	225
10%	209.7	56
15%	198.1	25
20%	186.4	14
30%	163.1	6

CIS identification and analyses

To investigate the presence of CIS in our study we used a region-based approach based on sliding windows (*Suzuki T, et al. 2002, Nat. gen. 32(1):166-174; Abel U, et al. 2011, PLoS One 6(10)*) and a new method for CIS identification based on a new genome-wide Grubbs test for outliers' analysis (*Biffi A, et al. 2011, Blood 117(20):5332-5339*).

The Grubbs test for outliers is a gene-centered approach and corrects the integration frequency by the size of the targeted gene rather than by user-defined genomic intervals. The rationale of the Grubbs test for outliers for CIS validation relies on the postulate that a significant CIS gene (identified by any statistical method) will be targeted at a significantly higher frequency than the average. Because the genes differ in size resulting in a different probability to be hit by vector integrations, the number of integrations targeting each gene was divided by the gene size (gene integration frequency). The basic calculations for genome-wide Grubbs test are described in Biffi et al., 2011. Briefly, the gene integration frequency is calculated by dividing the number of integrations targeting the same gene on genomic intervals defined by the boundaries of the targeted genes. Given that genes with no integrations can be non-targeted or non-sampled due to possible sub-sampling and saturation issues, LV targeted genes are a conditioning variable and we considered only the genes targeted by at least one integration. The gene integration frequency values are transformed by the minus logarithm base 2 to obtain a statistically evaluable normal distribution of the data. Indeed, the gene integration frequency values do not follow a normal distribution as they may vary from 0 to 1. The –log2 transformed gene integration frequency shows that the data follow a normal distribution by the D'Agostino & Pearson omnibus normality test. To reduce the background of high integration frequency values originating by small genes, an additional 100 Kb distance was added to each gene interval. Thus, the Z-score, t-studentization and a raw p-value are calculated considering a minimal gene size of at least 100 Kb. Bona fide CIS genes are those that are identified by the Abel's method (Supplemental Table 4) and whose integration frequency provided a raw p-value <0.05 (Table 1).

Role of the CISs in human hematopoietic tumors

To assess if the newly discovered CIS genes (targeted by >3 integrations N=34) are deregulated in human cancer, we analyzed independent gene expression studies of hematopoietic tumors with the Oncomine-datamining-tool (www.oncomine.org; Supplemental Methods). We interrogated datasets of human hematopoietic data, specifically "Leukemia", "Lymphoma" and "Myeloma" datasets, for a total of 139 different datasets and 15859 analyzed samples. We used the unpaired t-test statistics provided by Oncomine, applying the following tresholds: fold change >2; p-value <0.0001; gene rank: all genes. In order to score for gene deregulation, we considered both upregulation and downregulation. Among the 34 CIS genes, 32 could be successfully interrogated by Oncomine. Significant upregulation and/or downregulation (p<0.0001 by unpaired-

T-test) of 25/32 CIS gene in tumors vs non-tumoral cells was found ($p<0.0001$, Supplemental Table 5 and Supplemental Fig 4).

We then investigated the expression levels of the CIS-genes identified by our insertional mutagenesis screening in two representative cohorts of 224 (Cohort1) and 165 (Cohort2) Acute Myeloid Leukemia (AML) patients further sub-divided in different molecular subclasses (*Valk PJ, et al The New England journal of medicine. 2004; 350(16):1617-1628; Verhaak RG, et al Haematologica. 2009;94(1):131-134*) (Supplemental Table7). For microarray gene expression analyses of AML samples, nonparametric Kruskal–Wallis Analysis Of Variance (ANOVA) was performed to identify overall differences in gene expression between AML groups. Whenever significant, Wilcoxon rank sum test for pairwise comparisons and adjusted with Bonferroni correction was applied. A p-value less than 0.05 was considered significant.

We found that the expression levels of 6 of the newly identified CIS genes (19%) were differentially expressed among different AML subgroups ($p<0.05$ non parametric Anova, Supplemental Table 6 and Supplemental Fig 5), suggesting that they may have a distinctive role in these different subtypes of AML.