



- (51) International Patent Classification:
C12N 15/86 (2006.01) C07K 14/005 (2006.01)
- (21) International Application Number:
PCT/EP2017/076615
- (22) International Filing Date:
18 October 2017 (18.10.2017)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
1617599.4 18 October 2016 (18.10.2016) GB
- (71) Applicant: KATHOLIEKE UNIVERSITEIT LEUVEN
[BE/BE]; KU Leuven R&D, Waaistraat 6 - box 5105, 3000
Leuven (BE).
- (72) Inventors: GIJSBERS, Rik; Kloosterweg 18, 3052 Oud-
Heverlee (BE). DE RIJCK, Jan; Jan Davidtsstraat 103,
3010 Kessel-Lo (BE).

- (84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, GH,
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ,
UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ,
TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,
EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,
MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,
TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW,
KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- of inventorship (Rule 4.17(iv))

Published:

- with international search report (Art. 21(3))
— before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments (Rule 48.2(h))
— with sequence listing part of description (Rule 5.2(a))

- (81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ,
CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO,
DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN,
HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP,
KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME,
MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ,
OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA,
SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

- (54) Title: MLV-BASED GENE THERAPY WITH MUTATED INTEGRASE COUPLED TO PEPTIDES

A

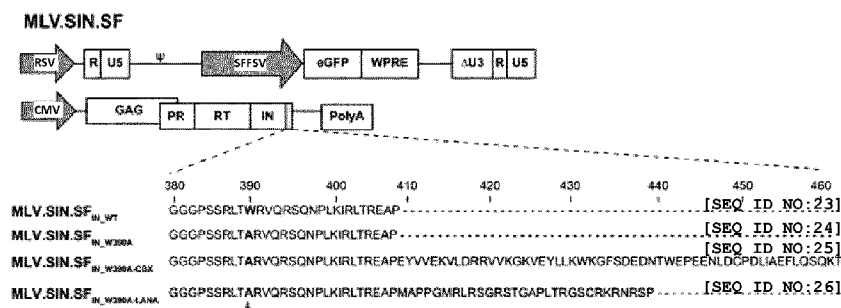


Figure 6

- (57) Abstract: The invention relates to gamma retroviral plasmids comprising a nucleotide sequence which codes for a gamma-retroviral-Integrase (IN), which is independent from bromodomain and extra-terminal (BET) family of proteins, wherein said nucleotide is operably linked to a nucleotide sequence which codes for a chromatin binding domain (CBD) or a chromatin binding fragment thereof.

MLV-BASED GENE THERAPY WITH MUTATED INTEGRASE COUPLED TO PEPTIDES

Field of the invention

5 The application provides methods to improve the safety of retroviral vector-mediated gene therapy, and tools for use in such methods. In the present invention more particularly next-generation BinMLV vectors are engineered with a reduced risk of insertional mutagenesis without compromising transduction efficiency. As such, next-generation BinMLV vectors can improve the safety of gammaretroviral vector gene
10 therapy.

Background

Stable integration of retrovirus-based viral vectors encompassing therapeutic transgene cassettes in patient-derived hematopoietic progenitor cells enables gene
15 therapeutic correction for severe blood and immune disorders. Over the past 25 year, Murine Leukaemia Virus (MLV)-based vectors have shown therapeutic benefits in gene replacement therapy using bone marrow transplantation for primary immunodeficiencies, such as X-linked severe combined immunodeficiency (SCID-X1), adenosine deaminase deficiency-severe combined immunodeficiency (ADA-SCID), X-
20 linked chronic granulomatous disease (X-CGD) and Wiskott-Aldrich syndrome (WAS). Next to inherited disorders, chimeric antigen receptors (CAR) T-cell therapy has been successfully applied for patients with B-cell malignancies.

Compared to haploidentical hematopoietic (HSC) therapy, autologous HSC gene therapy for SCID-X1 showed better clinical outcomes and faster immune
25 reconstitution in infants lacking a matched sibling donor treated with gene therapy. These results show that gene therapy for SCID-X1 is at least an equally good, if not superior, alternative to haploidentical HSCT.

However, adverse events due to insertional mutagenesis were observed in the first clinical trials. Even though the majority of the X-SCID patients treated with gene
30 therapy gained significant clinical benefits, widespread application has been hampered by the development of leukaemia in a subset of SCID-X1 patients (5/19) between two and five years after treatment. Importantly, four of these patients were successfully treated and recovered immunity without the need for further intervention. Clonal expansion was caused by vector-mediated insertional activation
35 of proto-oncogenes, such as integration near the LMO2 promoter (4 out of 5) (Cavazzana-Calvo, M. et al., (2012) *Curr Opinion Immunology*, **24(5)**, 580–584), or

near CCND2. Gene therapy trials for WAS and X-CGD similarly reported insertion-mediated activation of proto-oncogenes, like LMO2, MDS1-EVI1 and PRDM16, leading to myelodysplasia and leukaemia.

The insertional activation of proto-oncogenes occurred as a consequence of the
5 outspoken integration preferences and the presence of strong enhancer sequences in the first generation long terminal repeat (LTR)-driven retroviral vectors. To prevent insertional mutagenesis, modifications in the design of the viral vector genome, such as the enhancer-depleted self-inactivating (SIN) configuration with introduction of physiological promoters and the introduction of insulator improved vector safety
10 without affecting efficacy. A second approach to improve gene therapy safety is to alter the integration pattern, directing proviral integration away from potentially harmful genomic loci, such as proto-oncogene promoters. This can be achieved either by exploiting alternative retroviral vector platforms with an intrinsic more random integration pattern, such as alpha-retroviral vectors (Suerth et al. 2009), or by
15 interference with the integration site selection process through tinkering with the chromatin-tethering of the retroviral pre-integration complex (PIC).

Gammaretroviral integration is not random, but associates with strong enhancers and promoter regions. Integration preference is dictated by cellular cofactors that are co-opted by the viral integrase protein. For gammaretroviruses and their derived
20 viral vectors, the bromodomain and extraterminal domain containing (BET) family of proteins (BRD2, BRD3 and BRD4) serve as anchors on the host chromatin. BET proteins are bipartite. They consist of two N-terminal bromodomains that bind epigenetic chromatin marks and a C-terminal extraterminal (ET) protein-protein interaction domain that binds host-cell proteins, but also the MLV integrase (MLV IN).
25 Together these domains allow BET-proteins to tether the retroviral pre-integration complex to specific chromatin environments and thereby define the retroviral integration profile. A motif in the unstructured C-terminal tail of MLV IN assures interaction with the BRD ET-domain. Deletion of this domain ($\Delta 23$ a.a., IN1-380) or a single substitution (INW390A) uncouples the BET-interaction (El Ashkar, S. et al.,
30 (2014). *Molecular Therapy Nucleic Acids*, **3**:e179), resulting in BET-independent MLV (BinMLV) vectors that transduce target cells as efficiently as wild type (WT) MLV vectors but no longer exhibit the typical MLV integration profile (Ashkar et al. cited above). A disadvantage of this method however, is that, as they distribute more randomly, they can still cause insertional mutagenesis.

35 Retroviral integration site selection is dictated by the interaction between the viral IN as part of the PIC and cellular cofactors. Previously, the HIV-cellular tethering

cofactor LEDGF/p75 was also re-engineered, and efficient redistribution of retroviral integration without compromising transgene expression was demonstrated. These retargeting strategies are however based on (transient) expression of alternative tethers by linking the C-terminal domain of LEDGF/p75 to heterologous chromatin binding domains such as CBX (Vets S, et al. (2013) Mol Ther Nucleic Acids. 5;2:e77.). Although these methods allow retargeting, they are not readily applicable as it requires the introduction (at least transient) of artificial anchors in target cells prior to application of the therapeutic vectors, which is not desirable in a clinical gene therapy setting. Hence, there remains a need for safer methods of gene therapy.

10

Summary of the invention

The present invention overcomes the disadvantages by the design of BET independent gamma retroviral integrase which is fused to a chromatin binding domain which targets retroviral integration to safe regions of the genome, further reducing the risk of insertional mutagenesis.

15

The present invention has the advantage over Al-Ashar et al (2014) in that the random insertion is avoided by re-targeting gamma retroviral integration away from the random integration sites as obtained with the BET deficient constructs, as well as away from the detrimental integration sites as obtained with the wild-type gamma retroviral integrase.

20

The present invention has the advantage over the conceptual study of Vets et al. (Mol Ther Nucleic Acids. (2013) 5, e77) wherein it is required to interfere significantly with the cellular metabolism. Since the authors work with a modified cellular protein LEDGF/p75 of the host cell, it is required to downregulate endogenous wild type LEDGF/p75 to ensure that the lentiviral vector interacts with the modified LEDGF/p75 protein. The authors recognise this problem and suggest the design of a modified integrase that would only bind to the modified LEDGF/p75 hybrid.

25

The present invention overcomes the serious burden of developing modified integrase as well as modified LEDGF proteins, whereby integrase activity should be preserved and the binding of integrase to LEDGF/p75 has to shift from the wild type protein to the, yet unknown, modified LEDGF/p75 protein.

30

The present invention describes the development of a next generation of BinMLV vectors in an effort to alleviate the risk of insertional mutagenesis by interfering with the chromatin-tethering process. The present invention provides for retroviral vectors with an integration site pattern that is directed away from potentially harmful

35

chromosomal sites, such as promoters of proto-oncogenes, by linking peptide sequences that recognize chromatin features that are widely distributed over the genome to the C-terminal end of BinMLV IN (Table 2). The inventors demonstrate that fusion of these peptides to BinMLV IN generates vectors that efficiently transduce
5 laboratory cell lines and primary cells. Moreover, integration site preferences were overall more random relative to epigenetic markers and genomic features, resulting in a commensurate safer integration profile and a reduced risk of hematopoietic cell transformation.

Accordingly, the present invention provides for a retroviral plasmid comprising a
10 nucleotide sequence which codes for a Bromodomain and Extra-Terminal (BET) family of proteins-independent, gamma-retroviral-Integrase (IN), operably linked to a nucleotide sequence which codes for a chromatin binding domain (CBD) or a chromatin binding fragment thereof. In particular embodiments, the gamma-retroviral-IN is a Murine Leukaemia Virus (MLV)-IN. In other embodiments, said
15 gamma-retroviral-IN codes for a C-terminal modified retroviral-IN, wherein said C-terminal amino acid (AA) sequence comprises the BET interaction domain and wherein at least one C-terminal AA is deleted or mutated. In a particular embodiment, said C-terminal AA sequence contains about the last C-terminal 28 AA of MLV IN, or the corresponding amino acids of any gamma-retroviral IN homologue thereof. In a
20 particular embodiment said mutated integrase codes for a single AA-mutation, whereby a conserved W, corresponding to W390 of MLV-IN, is mutated or deleted, or more particularly said conserved W is mutated to A, corresponding to W390A of MLV-IN. In yet another embodiment said mutated retrovirus-IN codes for a C-terminal modified retrovirus-IN, wherein about the last 28 C-terminal (corresponding
25 to MLV-integrase AA382-AA408) AAs are deleted. In further embodiments, the CBD is the latency associated nuclear antigen (LANA) peptide (SEQ ID NO: 1), the heterochromatin-binding protein 1 β (CBX1) (SEQ ID NO: 2), the chromodomain Y-like protein (CDYL) (SEQ ID NO: 3), the human papilloma virus (HPV8) E2 protein (SEQ ID NO: 4), or a variant thereof having at least 80%, preferably at least 85%,
30 more preferably at least 90%, most preferably 95% AA sequence identity to SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3 or SEQ ID NO: 4, and maintaining chromatin binding. The invention further provides for a host cell transfected with said retroviral plasmid. In particular embodiments said retroviral plasmid nucleotide sequence is stably integrated in the cellular genome.

35 The invention further provides an infectious viral vector particle as well as the process for obtaining the infectious viral vector particle, wherein a helper cell is transfected

with a retroviral plasmid according to any one of the embodiments above, a transfer plasmid comprising self-inactivating (SIN) LTR sequences, 5' LTR and 3'LTR sequences and a heterologous nucleotide sequence, and an envelope plasmid comprising an ENV nucleotide sequence and said helper cell is cultured in a suitable medium under conditions for the release of infectious viral vector particles. In particular embodiments the infectious viral vector particle further comprises a promoter sequence, suitable for expression in a mammalian cell and wherein said heterologous nucleotide sequence is a transgene that can be expressed via said promoter. In other embodiments the invention provides for a host cell that has been transduced with said infectious viral vector particle. In yet another embodiment, the invention provides for the in vitro use of said infectious retroviral vector particle for the delivery of a heterologous sequence or transgene to a host cell, preferably in gene therapy, wherein more preferably said gene therapy comprises selecting the host cells in which said heterologous sequence or transgene has been integrated in a "safe" genomic location. In a more particular embodiment the host cell is a haematopoietic cell or a stem cell or more particularly a haematopoietic stem cell or a progenitor cell.

The invention further provides for the use of said infectious retroviral vector particle for preparing a pharmaceutical composition for gene therapy, wherein said infectious retroviral vector particle is present in a form suitable for being introduced into the target cells. In another embodiment, the invention provides for a pharmaceutical composition that comprises said infectious retroviral vector particle together with a pharmaceutically acceptable excipient and/or carrier. In yet another embodiment, the invention further provides for a method of gene therapy, which method comprises administering said infectious retroviral vector particle to said subject.

Other embodiments, objects, features and advantages will be set forth in the detailed description of the embodiments that follows. The summary above is to be considered as a brief and general overview of some of the embodiments disclosed herein, is provided solely for the benefit and convenience of the reader, and is not intended to limit in any manner the scope encompassed by the appended claims.

Numbered statements of the invention are:

1. A retroviral plasmid comprising a nucleotide sequence which codes for a bromodomain and extra-terminal (BET) family of proteins-independent, gamma-retroviral-Integrase (IN), operably linked to a nucleotide sequence which codes for a chromatin binding domain (CBD).

2. The plasmid according to statement 1, wherein said the gamma-retroviral-IN is a Murine Leukemia Virus (MLV)-IN.
3. The plasmid according to statement 1 or 2, wherein said gamma-retroviral-IN codes for a C-terminal modified retroviral-IN, wherein at least one C-terminal amino acid (AA) is deleted or mutated and said C-terminal AA sequence comprises the BET interaction domain.
4. The plasmid according to statement 3, wherein said C-terminal AA sequence contains about the last C-terminal 28 AA of MLV IN, or the corresponding amino acids of any gamma-retroviral IN homologue thereof.
5. The plasmid according to statement 3 or 4, wherein said mutated integrase codes for a single AA-mutation, whereby a conserved W, corresponding to W390 of MLV-IN, is mutated or deleted.
6. The plasmid according to statement 5, wherein said conserved W is mutated to A, corresponding to W390A of MLV-IN.
7. The plasmid according to statement 4, wherein said mutated retrovirus-IN codes for a C-terminal modified retrovirus-IN, wherein about the last 28 C-terminal (corresponding to MLV-integrase AA382-AA408) AAs are deleted.
8. The plasmid according to any one of statements 1 to 7, wherein the CBD is the latency associated nuclear antigen (LANA) peptide (SEQ ID NO: 1), the heterochromatin-binding protein 1 β (CBX1) (SEQ ID NO: 2), or a variant thereof having at least 80% AA sequence identity to SEQ ID NO: 1 or SEQ ID NO: 2.
9. Process for obtaining an infectious viral vector particle, wherein a helper cell is transfected with a retroviral plasmid according to any one of statements 1 to 8, a transfer plasmid comprising self-inactivating (SIN) LTR sequences, 5' LTR and 3'LTR sequences and a heterologous nucleotide sequence, and an envelope plasmid comprising an ENV nucleotide sequence and said helper cell is cultured in a suitable medium under conditions for the release of infectious viral vector particles.
10. An infectious viral vector particle comprising a nucleotide sequence which codes for a BET family of proteins-independent, gamma-retroviral-Integrase, operably linked to a nucleotide sequence which codes for a CBD obtainable by the process according to statement 9.
11. The infectious viral vector particle of statement 10, which further comprises a promoter sequence, suitable for expression in a mammalian cell and wherein said heterologous nucleotide sequence is a transgene that can be expressed via said promoter.

12. A host cell transfected with a retroviral plasmid according to any one of statements 1 to 8.
13. The host cell of statement 12, wherein said retroviral plasmid nucleotide sequence is stably integrated in the cellular genome.
- 5 14. A host cell that has been transduced with an infectious viral vector particle according to statement 10 or 11.
15. In vitro use of the infectious retroviral vector particle according to statement 10 or 11 for the delivery of a heterologous sequence or transgene to a host cell.
16. The use according to statement 15 in gene therapy.
- 10 17. The use according to statement 16, wherein said gene therapy comprises selecting the host cells in which said heterologous sequence or transgene has been integrated in a "safe" genomic location.
18. The use according to any one of statements 15 to 17, wherein the host cell is a haematopoietic cell or a stem cell.
- 15 19. The use according to any one of statements 15 to 18, wherein said host cell is haematopoietic stem cell or a progenitor cell.
20. Use of the infectious retroviral vector particle according to statement 10 or 11 for preparing a pharmaceutical composition for gene therapy, wherein said infectious retroviral vector particle is present in a form suitable for being introduced into the
- 20 target cells.
21. Pharmaceutical composition that comprises the infectious retroviral vector particle according to statement 10 or 11 together with a pharmaceutically acceptable excipient and/or carrier.
22. A method of gene therapy, which method comprises administering the
- 25 infectious retroviral vector particle according to statement 10 or 11 to said subject.
23. The method according to statement 22, wherein a heterologous sequence or transgene is introduced in a subject in need thereof and wherein said subject is a mammal or human.
24. A retroviral plasmid comprising a nucleotide sequence which codes for a
- 30 gamma-retroviral-Integrase (IN), which is independent from bromodomain and extra-terminal (BET) family of proteins, wherein said nucleotide sequence is operably linked to a nucleotide sequence which codes for a chromatin binding domain (CBD) or a chromatin binding fragment thereof.
25. The plasmid according to statement 24, wherein said gamma-retroviral-IN is
- 35 a Murine Leukemia Virus (MLV)-IN.

26. The plasmid according to statement 24 or 25, wherein said gamma-retroviral-IN codes for a C-terminal modified retroviral-IN, wherein at least one amino acid in the BET interaction domain is deleted or mutated, wherein the BET interaction domain corresponds to amino acids 389-405 of MLV-IN.
- 5 27. The plasmid according to statement 26, whereby the conserved W of said BET interaction domain, corresponding to W390 of MLV-IN, is mutated or deleted.
28. The plasmid according to statement 26 or 27, comprising a W390A mutation.
29. The plasmid according to statement 28, wherein said integrase is a C-terminal modified integrase, wherein the region corresponding to amino acids 382 to 408 of
10 MLV-integrase is deleted.
30. The plasmid according to any one of statements 24 to 29, wherein the chromatin binding domain or chromatin binding fragment thereof comprises a chromodomain.
31. The plasmid according to any one of statements 24 to 30, wherein the
15 chromatin binding fragment is the peptide from latency associated nuclear antigen (LANA) peptide with SEQ ID NO: 1, the peptide of heterochromatin-binding protein 1 β (CBX1) with SEQ ID NO: 2, or a chromatin binding variant thereof having at least 80% sequence identity to SEQ ID NO: 1 or SEQ ID NO: 2.
32. A host cell transfected with a retroviral plasmid according to any one of
20 statements 24 to 31.
33. The host cell of statement 32, wherein said retroviral plasmid is stably integrated in the cellular genome.
34. A method for obtaining an infectious viral vector particle, comprising the steps of :
- 25 a) transfecting a helper cell with :
-a retroviral plasmid according to any one of statements 24 to 31,
-a transfer plasmid comprising self-inactivating (SIN) LTR sequences, 5' LTR and 3'LTR sequences and a heterologous nucleotide sequence, and
- an envelope plasmid comprising an ENV nucleotide sequence and
- 30 b) culturing said transfected in a suitable medium under conditions for the release of infectious viral vector particles.
35. An infectious viral vector particle comprising a nucleotide sequence which codes for a gamma-retroviral-Integrase, which is independent from bromodomain and extra-terminal (BET) family of proteins, operably linked to a nucleotide sequence
35 which codes for a chromatin binding domain, or chromatin binding fragment thereof, and comprising a heterologous nucleotide sequence.

36. The infectious viral vector particle according to statement 35, wherein said heterologous nucleotide sequence is a transgene under the control of a promoter sequence, suitable for expression in a mammalian cell.

37. A host cell that has been transduced with an infectious viral vector particle according to statement 35 or 36.

38. In vitro use of the infectious retroviral vector particle according to statement 35 or 36 for the delivery of a heterologous sequence or transgene to a host cell.

39. The in vitro use according to statement 38 in gene therapy.

40. The use according to statement 38 or 39 wherein host cells are selected in which said heterologous sequence or transgene has been integrated in a "safe" genomic location.

41. The use according to any one of statements 38 to 40, wherein the host cell is a hematopoietic cell or a stem cell.

42. The use according to any one of statements 38 to 41, wherein said host cell is hematopoietic stem cell or a progenitor cell.

43. The infectious retroviral vector particle according to statement 35 or 36, wherein said infectious retroviral vector particle is present in a form suitable for being introduced into a target cell, **for use as a medicament** in gene therapy.

44. A method of gene therapy, which method comprises administering the infectious retroviral vector particle according to statement 35 or 36 to said subject.

45. The method according to statement 44, wherein a heterologous sequence or transgene is introduced in a subject in need thereof and wherein said subject is a mammal or human.

25 Figures

Figure 1 : C terminal sequences of integrase of a representative set of gamma retroviruses. The BET interacting part of the sequence underlined. Trp corresponding to W390 of MLV IN is indicated in bold. Conserved amino acids are indicated with *:

30 AKR-MLV : AKR murine leukemia virus (P03356); BaEV: Baboon endogenous virus (P10272) : Cas-BR-E: Cas-Br-E murine leukemia virus (P08361); en-FeLV: endogenous feline leukemia virus (P10273); GaLV: Gibbon ape leukemia virus (P21414); KoRV: Koala retrovirus (Q9TTC1); MLV: murine leukemia virus (P03355); PERV: Porcine endogenous retrovirus (Q8UM96).

35 **Figure 2: Transduction efficiencies of next generation BinMLV vectors.**

- (A) Schematic representation of MLV-based vector production. Structure of the IN_{WT}, IN_{W390A} and peptides fused to IN_{W390A} are highlighted. The N-terminal HHCC zinc binding domain, the catalytic core domain (CCD) and the C-terminal domain (CTD) are indicated. Black arrowhead indicates the position of W390A point mutation. Size of the fused peptides is proportionally represented. LTR; long terminal repeat; ψ ; packaging signal, eGFP; enhanced green fluorescent protein, CMV; cytomegalovirus promoter, VSV-G; vesicular stomatitis virus glycoprotein G, PolyA; polyadenylation signal, GAG; group-specific antigen, Pol; polymerase, PR; protease, RT, reverse transcriptase, IN; integrase.
- (B) FACS analysis of SupT1 cells transduced with equal reverse transcriptase (RT)-units of the indicated vectors. Three days (B) post transduction, the percentage of eGFP-positive cells was determined. Average values and standard deviations of triplicate measurements are shown. Data represent measurements from a representative experiment.
- (C) Mean fluorescence intensity of SupT1 cells transduced with the indicated next generation BinMLV vectors at day 3 post transduction.

Figure 3: Integration site distribution of Next generation BinMLV vectors.

- (A) Murine leukaemia virus (MLV)-based vector integration sites obtained from SupT1 cells and their genomic distribution. Integration percentages in 2 kb windows around TSS, CpG island midpoints and DHS are listed. For comparison, integration of computer generated match random controls (MRC) for MLV_{IN_W390A} are generated. P values(*) show significant departures (**p<0.001, pairwise Fishers test) from MLV_{IN_W390A} and MRC, respectively, separated by 1. TSS; transcription start sites, CpG; CpG-rich islands, DHS; DNase I-hypersensitive sites.
- (B-C) Heat maps summarizing the relation between vector integration site frequency and different genomic (B) or epigenetic (C) features in SupT1 cells. Evaluated vectors are indicated above the columns. Features analysed are shown to the left of the corresponding row of the heat map. Tile colours indicate whether a particular feature is favoured or disfavoured for integration of the respective data sets relative to their MRCs, as detailed in the coloured ROC area scale at the bottom of the panel. P-values (*) show significance of departures from MLV_{IN_W390A} integration sites in SupT1 cells (**p<0.001, Wald statistics referred to χ^2 distribution). In panel B there is a tendency for enrichment, with the exception for items in "gene boundaries" where there is depletion. In panel C there is a tendency for enrichment, with the exception for items in "transcriptionally silent" where there is depletion.

Figure 4: Addition of the fusion peptides in MLV backbone.

(A) Schematic representation of IN_{WT}, IN₁₋₃₈₀ and peptides fused to IN₁₋₃₈₀. The N-terminal HHCC zinc binding domain, the catalytic core domain (CCD) and the C-terminal domain (CTD) are indicated. Size of the fused peptides is proportionally represented.

5 (B) Transduction efficiency (% GFP positive cells) in SupT1 cells three days post transduction. Average values and standard deviations of triplicate measurements are shown. Data represent measurements from a representative experiment.

(C) Mean fluorescence intensity of SupT1 cells transduced with the MLV_{IN_W390A} fusion vectors at day three post-transduction.

10 **Figure 5: Therapeutic potential next generation BinMLV vectors.**

(A) Analysis of next generation BinMLV-based vector integration site with respect to the five safe harbour criteria. P values (*) show significance (***) p<0.001, Pearson's chi-square) from MLV_{IN_WT} and MLV_{IN_W390A} respectively separated by I. Computationally generated matched random controls are shown. Allonco gene list is
15 available at the Bushman Lab cancer gene list website (<http://www.bushmanlab.org/links/genelists>). TSS; transcription start sites, Allonco; oncogenes, TU; transcription units, UCR; ultraconserved region, MRC; matched random controls.

(B) Transduction efficiency of the indicated vectors 3 days and 10 days post
20 transduction in primary CD4⁺ T-cells. Percentage of GFP positive cells and mean fluorescence intensities (MFI) are indicated. Exemplary flow-cytometry plots of untransduced cells are shown for comparison. Representative experiment is shown.

(C-D) Transduction efficiency at day 4 (C) and mean vector copy number at day 5
25 post transduction (D) of next generation BinMLV vectors in mouse hematopoietic lineage depleted bone marrow cells at increasing multiplicity of infections. Lin⁻; lineage depleted cells, VCN; vector copy number, MOI; multiplicity of infection.

Figure 6: Therapeutic potential of next generation BinMLV vectors.

(A) Schematic representation of MLV.SIN.SF used in the IVIM assay. SRS11 plasmid of MLV SIN was used. Indicated are the RSV (Rous sarcoma virus) promoter, the ψ
30 packaging signal, the long terminal repeats (Δ U3, R and U5), spleen focus-forming virus (SFFV) promoter, enhanced green fluorescent protein (eGFP), Woodchuck hepatitis virus posttranscriptional regulatory element (WPRE). Amino acid sequences of the fused peptides and the position of W390A mutation are indicated.

(B) IVIM assay for SIN next generation BinMLV vectors.

35 (C-D) Number of clones (C) and Replating frequencies corrected for the mean vector copy number (D) as measured in DNA of mass cultures transduced with SIN next

generation BinMLV vectors. Horizontal bars indicate the mean of all assays for a given vector.

Figure 7: Transduction efficiencies of next generation BinMLV vector in HeLa cells.

5 (A-B) Transduction efficiency (A) and mean fluorescence intensity (B) of SupT1 cells ten days post transduction with the indicated MLV_{IN_W390A} -fusion vectors. Average values and standard deviations of triplicate measurements are shown. Data represent measurements from a representative experiment.

(C-D) Transduction efficiency (C) and mean fluorescence intensity(D) of the indicated
10 MLV_{IN_1-380} fusion vectors in SupT1 cells 10 days post transduction. Average values and standard deviations of triplicate measurements are shown. Data represent measurements from a representative experiment.

Figure 8: Additional detargeting effect of next generation BinMLV vectors.

(A) MLVIN_WT and MLVIN_1-380 fusion vectors integration sites from SupT1 cells
15 and their genomic distribution in a 2 kb windows around TSS, CpG island midpoints and DHS. TSS; transcription start sites, CpG; CpG-rich islands, DHS; DNase I-hypersensitive sites. All data sets reached significance $p < 0.001$ (pairwise Fishers test), compared to MLVIN_1-380 and matched random controls (MRC for MLVIN_1-380), respectively, separated by |.

20 (B-C) Heat maps summarizing the relation between vector integration site frequency and different genomic (B) and epigenetic (C) features in SupT1 cells. Evaluated vectors are indicated above the columns. Features analysed are shown to the left of the corresponding row of the heat map. Tile colours indicate whether a particular feature is favoured or disfavoured for integration of the respective data sets relative
25 to their MRCs, as detailed in the coloured ROC area scale at the bottom of each panel. P -values (*) show significance of departures from MLVIN_1-380 integration sites in SupT1 cells. (***) $p < 0.001$, Wald statistics referred to χ^2 distribution).

In panel B there is a tendency for enrichment, with the exception for items in "gene boundaries" where there is depletion. In panel C there is a tendency for enrichment,
30 with the exception for items in "transcriptionally silent" where there is depletion.

Figure 9: CBX and LANA fusions do not affect local integration site.

Sequence logos representing nucleotide frequencies at the indicated next generation BinMLV vector integration sites compared to MLV_{IN_WT}. The height of a letter is proportional to the observed frequency of the corresponding nucleotide and the
35 height of each letter stack corresponds to the level of conservation, represented as information content (y-axis, bits).

Detailed description

The present invention will be described with respect to particular embodiments but the invention is not limited thereto. Any reference signs in the claims shall not be construed as limiting the scope thereof.

As used herein, the singular forms "a", "an", and "the" include both singular and plural referents unless the context clearly dictates otherwise.

The terms "comprising", "comprises" and "comprised of" as used herein are synonymous with "including", "includes" or "containing", "contains", and are inclusive or open-ended and do not exclude additional, non-recited members, elements or method steps. The terms "comprising", "comprises" and "comprised of" when referring to recited components, elements or method steps also include embodiments which "consist of" said recited components, elements or method steps.

Furthermore, the terms first, second, third and the like in the description and in the claims, are used for distinguishing between similar elements and not necessarily for describing a sequential or chronological order, unless specified. It is to be understood that the terms so used are interchangeable under appropriate circumstances and that the embodiments of the invention described herein are capable of operation in other sequences than described or illustrated herein.

The term "**about**" as used herein when referring to a measurable value such as a parameter, an amount, a temporal duration, and the like, is meant to encompass variations of +/-10% or less, preferably +/-5% or less, more preferably +/-1% or less, and still more preferably +/-0.1% or less of and from the specified value, insofar such variations are appropriate to perform in the disclosed invention. For example, with regard to Amino acids, it is meant to encompass variations of +/-5 or less, preferably +/-3 or less, and more preferably +/-1 or less of and from the specified value. It is to be understood that the value to which the modifier "about" refers is itself also specifically, and preferably, disclosed.

The recitation of numerical ranges by endpoints includes all numbers and fractions subsumed within the respective ranges, as well as the recited endpoints.

The term "**retrovirus**" shall be used to describe a virus from the family of the Retroviridae and its infections, which term shall be used to embrace human and animal retroviruses. Various genera within the Retroviridae family include Alpharetrovirus, Betaretrovirus, Gammaretrovirus, Deltaretrovirus, Epsilon-retrovirus, Lentivirus and Spumavirus. Members of the Lentivirus genus include

human immunodeficiency virus 1 (HIV-1) and human immunodeficiency virus 2 (HIV-2).

The term "**vector**" generally refers to nucleic acid molecules, usually double-stranded DNA, which may have inserted into it another nucleic acid molecule or transgene (the insert nucleic acid molecule) such as, but not limited to, a cDNA molecule. The vector is used to transport the insert nucleic acid molecule into a suitable host cell. A vector may contain the necessary elements that permit transcribing the insert nucleic acid molecule, and, optionally, translating the transcript into a polypeptide. The insert nucleic acid molecule may be derived from the host cell, or may be derived from a different cell or organism. Once in the host cell, the vector can replicate independently of, or coincidental with, the host chromosomal DNA, and several copies of the vector and its inserted nucleic acid molecule may be generated. The term "vector" may thus also be defined as a gene delivery vehicle that facilitates gene transfer into a target cell.

The term "**viral vector**" as used herein denotes vectors that are derived from viruses including but not limited to: retrovirus, including lentivirus, adeno-associated virus, adenovirus, herpesvirus, hepatitis virus or the like. Typically, but not necessarily, viral vectors are replication-deficient as they have lost the ability to propagate in a given cell since viral genes essential for replication have been eliminated from the viral vector. However, some viral vectors can also be adapted to replicate specifically in a given cell, such as e.g. a cancer cell, and are typically used to trigger the (cancer) cell-specific (onco)lysis. In the context of the present invention a "retroviral vector" is a vector that is retrieved from a retrovirus, such as a gamma retrovirus.

The term "transgene" as used herein refers to particular nucleic acid sequences encoding a polypeptide or a portion of a polypeptide to be expressed in a cell into which the nucleic acid sequence is inserted. However, it is also possible that transgenes are expressed as RNA, typically to lower the amount of a particular polypeptide in a cell into which the nucleic acid sequence is inserted. These RNA molecules include but are not limited to molecules that exert their function through RNA interference (shRNA, RNAi), micro-RNA regulation (miR), catalytic RNA, antisense RNA, RNA aptamers, etc. Of note, expression of the transgene may be restricted to a subset of the cells into which the nucleic acid sequence is inserted. The term transgene is meant to include (1) a nucleic acid sequence that is not naturally found in the cell (i.e., a heterologous nucleic acid sequence); (2) a nucleic acid sequence that is a mutant form of a nucleic acid sequence naturally found in the cell into which it has been introduced; (3) a nucleic acid sequence that serves to add

additional copies of the same (i.e., homologous) or a similar nucleic acid sequence naturally occurring in the cell into which it has been introduced; or (4) a silent naturally occurring or homologous nucleic acid sequence whose expression is induced in the cell into which it has been introduced. By mutant form is meant a nucleic acid sequence that contains one or more nucleotides that are different from the wild-type or naturally occurring sequence, i.e., the mutant nucleic acid sequence contains one or more nucleotide substitutions, deletions, and/or insertions. In some cases, the transgene may also include a sequence encoding a leader peptide or signal sequence such that the transgene product will be secreted from the cell.

5 The term "**promoter**" refers to nucleic acid sequences that regulate, either directly or indirectly, the transcription of corresponding nucleic acid coding sequences (e.g. a transgene) to which they are operably linked. A promoter may function alone to regulate transcription or may act in concert with one or more other regulatory sequences (e.g. enhancers or silencers). In the context of the present application, a promoter is typically operably linked to a transgene to regulate transcription of the transgene.

The term "**operably linked**" as used herein refers to the arrangement of various nucleic acid molecule elements relative to each such that the elements are functionally connected and are able to interact with each other. Such elements may include, without limitation, a promoter, an enhancer, a polyadenylation sequence, one or more introns, and a coding sequence of a gene of interest to be expressed (i.e., the transgene). The nucleic acid sequence elements, when properly oriented or operably linked, act together to modulate the activity of one another, and ultimately may affect the level of expression of the transgene. By modulate is meant increasing, decreasing, or maintaining the level of activity of a particular element. The position of each element relative to other elements may be expressed in terms of the 5' terminus and the 3' terminus of each element, and the distance between any particular elements may be referenced by the number of intervening nucleotides, or base pairs, between the elements. As understood by the skilled person, operably linked implies functional activity, and is not necessarily related to a natural positional link. Indeed, when used in a vector, the regulatory elements will typically be located immediately upstream of the promoter (although this is generally the case, it should definitely not be interpreted as a limitation or exclusion of positions within the vector), but this needs not be the case in vivo. The term "operably linked" can also refer to the arrangement of the retroviral-Integrase and CBD via a linker peptide. As used herein the term "linker peptide" refers to amino acid sequences that connect or

link two polypeptide sequences, e.g., that link two polypeptide domains. In certain embodiments said linker peptide is about 10 amino acids or less than 15 amino acids, such as 12, 10, 8, 6, 4 or 2 amino acids. In a specific embodiment said linker peptide is a traditional Gly/Ser (GS) linker peptide as known in the art.

- 5 As used herein, the terms "**linked**," "**fused**", or "**fusion**", are used interchangeably. These terms refer to the joining together of two more elements or components, by whatever means including chemical conjugation or recombinant means. Methods of chemical conjugation (e.g., using heterobifunctional crosslinking agents) are known in the art. As used herein, the term "genetic fusion" refers to the co-linear, covalent
- 10 linkage or attachment of two or more proteins, polypeptides, or fragments thereof via their individual peptide backbones, through genetic expression of a single polynucleotide molecule encoding those proteins, polypeptides, or fragments. Such genetic fusion results in the expression of a single contiguous genetic sequence. Preferred genetic fusions are in frame, i.e., two or more open reading frames (ORFs)
- 15 are fused to form a continuous longer ORF, in a manner that maintains the correct reading frame of the original ORFs. Thus, the resulting recombinant fusion protein is a single polypeptide containing two or more protein segments that correspond to polypeptides encoded by the original ORFs (which segments are not normally so joined in nature).
- 20 The term "**gene therapy**" refers to a treatment encompassing the administration of a transgene to a patient. The transgene product may be a (e.g. therapeutic or immunogenic) protein, or an RNA molecule to block the expression of a specific gene using RNA interference technology, or the transgene may replace a defective gene in the treatment of a genetic disorder or disease.
- 25 The term "**patient**" or "**subject**" is used herein to describe an animal, especially including a domesticated mammal and preferably a human, to whom a treatment or procedure is performed. For treatment of those infections, conditions or disease states which are specific for a specific animal such as a human patient, the term patient refers to that specific animal. In certain specific embodiments, the patient or
- 30 subject of the present invention is a domesticated/agricultural animal or human patient of either gender.

- As used herein, the terms "**treat**" or "**treatment**" refer to both therapeutic treatment and prophylactic or preventative measures. Beneficial or desired clinical results include, but are not limited to, prevention of an undesired clinical state or disorder,
- 35 reducing the incidence of a disorder, alleviation of symptoms associated with a disorder, diminishment of extent of a disorder, stabilized (i.e., not worsening) state

of a disorder, delay or slowing of progression of a disorder, amelioration or palliation of the state of a disorder, remission (whether partial or total), whether detectable or undetectable, or combinations thereof. "Treatment" can also mean prolonging survival as compared to expected survival if not receiving treatment.

5 As used herein, the terms **"therapeutic treatment"** or "therapy" and the like, refer to treatments wherein the object is to bring a subjects body or an element thereof from an undesired physiological change or disorder to a desired state, such as a less severe or unpleasant state (e.g., amelioration or palliation), or back to its normal, healthy state (e.g., restoring the health, the physical integrity and the physical well-
10 being of a subject), to keep it at said undesired physiological change or disorder (e.g., stabilization, or not worsening), or to prevent or slow down progression to a more severe or worse state compared to said undesired physiological change or disorder.

As used herein the terms **"prevention"**, **"preventive treatment"** or **"prophylactic treatment"** and the like encompass preventing the onset of a disease or disorder,
15 including reducing the severity of a disease or disorder or symptoms associated therewith prior to affliction with said disease or disorder. Such prevention or reduction prior to affliction refers to administration of the gene therapy via a transfected or transduced host cell, or via an infectious viral vector particles
20 envisaged herein to a patient that is not at the time of administration afflicted with clear symptoms of the disease or disorder. "Preventing" also encompasses preventing the recurrence or relapse-prevention of a disease or disorder for instance after a period of improvement.

A **'therapeutic amount'** or **'therapeutically effective amount'** as used herein
25 refers to the amount of an active compound or pharmaceutical agent (e.g., a transgene product) effective to treat a disease or disorder in a subject, i.e., to obtain a desired local or systemic effect. The term thus refers to the quantity of the compound or the agent that elicits the biological or medicinal response in a tissue, system, animal, or human that is being sought by a researcher, veterinarian, medical
30 doctor or other clinician. Such amount will typically depend on the compound or the agent and the severity of the disease, but can be decided by the skilled person, possibly through routine experimentation. The term "prophylactically effective amount" refers to an amount of an active compound or pharmaceutical agent (e.g., a transgene product) that inhibits or delays in a subject the onset of a disorder as
35 being sought by a researcher, veterinarian, medical doctor or other clinician. An **"immunologically effective amount"** as used herein refers to the amount of

transgene product effective to enhance the immune response of a subject against a subsequent exposure to the immunogen encoded by the transgene. Levels of induced immunity can be determined, e.g. by measuring amounts of neutralizing secretory and/or serum antibodies, e.g., by plaque neutralization, complement fixation, enzyme-linked immunosorbent, or microneutralization assay.

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art in the field of the invention. Any methods and materials similar or equivalent to those described herein can also be used in the practice or the present invention, but the preferred methods and products are described herein.

The present application provides safer methods for gene therapy, in particular retroviral gene therapy. Retroviral gene therapy uses retrovirus-derived vectors for modifying the genome of a host cell, such as, but not limited to by delivering a transgene into a host cell. Retroviruses are RNA viruses that have the ability to insert their genes into host cell chromosomes after infection. Various genera within the Retroviridae family include Alpharetrovirus, Betaretrovirus, Gammaretrovirus, Deltaretrovirus, Epsilonretrovirus, Lentivirus and Spumavirus. Members of the Lentivirus genus include human immunodeficiency virus 1 (HIV-1) and human immunodeficiency virus 2 (HIV-2).

Retroviral vectors have been developed that lack the genes encoding viral proteins, but retain the ability to infect cells and insert their genes into the chromosomes of the target cell.

The retroviral vectors envisaged herein typically comprise a nucleic acid sequence or transgene operably linked to a promoter. For example, the nucleic acid sequence may encode a secretable protein or a protein which is defective in the host cell, or ensure disruption of a defective gene in the host cell. In particular embodiments, said retroviral vectors are gamma-retroviral vectors.

Other sequences may be incorporated in the retroviral vectors envisaged herein as well, typically to further increase or stabilize the expression of the transgene product (e.g. introns and/or polyadenylation sequences). The promoter does not need to be the promoter of the transgene in the viral vector, although it is possible that the transgene is transcribed from its own promoter. The promoter may be homologous (i.e. from the same species as the subject to which the infectious viral vector particle is administered) or heterologous (i.e. from a source other than the species of the subject to which the infectious viral vector particle is administered). The promoter may be an inducible or constitutive promoter.

In particular embodiments, the retroviral vectors envisaged herein are configured to enhance expression levels of the transgene comprised in the vector. This can be achieved by appropriate vector design, including the use of cis-acting elements such as promoters, introns, post-transcription regulatory elements such as Woodchuck hepatitis virus posttranscriptional regulatory element (WPRE), polyadenylation signals, and the CMV enhancer as known in the art (Powell et al. (2015) *Discov Med.* **102**:49-57).

The possibility of oncogenic transformation of host cells as a result of integration of retroviral DNA into a "wrong spot" of the host chromosome has always been a concern. The present invention provides for a retroviral plasmid or retroviral vector comprising a nucleotide sequence which codes for a bromodomain and extra-terminal (BET) family of proteins-independent, gamma-retroviral-Integrase (IN), operably linked to a nucleotide sequence which codes for a chromatin binding domain (CBD), that retargets integration of retroviruses towards safe sites. In certain embodiments, the linker peptide linking the IN to the CBD via about 2 amino acids. In other embodiments said linker peptide is a traditional Gly/Ser (GS) linker peptide as known in the art.

Accordingly, said integrase operably linked to a chromatin binding domain as envisaged herein ensures retargeting of retroviral vector integration towards a "safe site" in the host genome, thereby effectively providing safer retroviral gene therapy. With a "safe site" is meant herein a site in the host cell genome where the integrated material is adequately expressed without perturbing endogenous gene structure or function. Within the context of the present invention, integration of the infectious viral vector particle into the host cell genome is considered "safe" if the following five criteria are met: (i) distance of at least 50 kb from the 5' end of any gene, (ii) distance of at least 300 kb from any cancer-related gene, (iii) distance of at least 300 kb from any microRNA (miRNA), (iv) location outside a transcription unit and (v) location outside ultraconserved regions (UCRs) of the human genome.

Additionally or alternatively, a site can be identified as "safe" if it is recognized to be a site which does not lead to clonal proliferation or malignancy.

The invention relates to retroviral plasmids with integrase proteins with deficient binding to BET family proteins (BRD2, BRD3 and BRD4), in other words they are independent from the BET family of proteins. This is achieved by modification of the BET interaction motif in the C terminal region of the integrase protein. Different approaches are possible to render integrase BET independent. This is exemplified by reference to the integrase MLV virus and its amino acid numbering as shown in Figure

1. Similar considerations are possible for other retroviral integrases whereby a sequence alignment with MLV allows to determine the BET interaction motif. As can be seen in this figure the C terminal region of the integrase proteins differ from each other. It is nevertheless clear to identify in each sequence the BET interaction motif
5 and the C terminal parts of the sequence that are dispensable without losing integrase activity.

A BET independent integrase can be obtained by deletion of a part of the integrase c-terminal tail whereby at least also one amino acid of the BET interaction motif is removed. Typically the entire BET interacting motif is deleted. In addition one or
10 more of the amino acids between the C-terminal domain and the BET interacting motif are deleted and/ or one or more of the amino acids C-terminal of the BET interacting motif are deleted.

A BET independent integrase can also be obtained by mutating one or more amino acid in the BET interacting motif. In this approach the overall structure of the C
15 terminal part of integrase remains untouched. Most likely candidates for obtaining a BET independent integrase are mutations in the highly conserved residues indicated with an asterisk in figure 1. Alanine substitution, or a dramatic change in size, charge or hydrophobicity are typical approaches.

The effect of a particular mutation or deleted on BET interaction can be determined
20 as shown in Al Ashkar (2014), cited above.

In particular embodiments, the gamma-retroviral-IN is a Murine Leukaemia Virus (MLV)-IN. In other embodiments, said gamma-retroviral-IN codes for a C-terminal modified retroviral-IN, wherein at least one C-terminal amino acid (AA) is deleted or
25 mutated and said C-terminal AA sequence comprises the BET interaction domain. In a particular embodiment, said C-terminal AA sequence contains about the last C-terminal 28 AA of MLV IN, or the corresponding amino acids of any gamma-retroviral IN homologue thereof. In a particular embodiment said mutated integrase codes for a single AA-mutation, whereby a conserved W, corresponding to W390 of MLV-IN, is
30 mutated or deleted, or more particularly said conserved W is mutated to A, corresponding to W390A of MLV-IN. In yet another embodiment said mutated retrovirus-IN codes for a C-terminal modified retrovirus-IN, wherein about the last 28 C-terminal (corresponding to MLV-integrase AA382-AA408) AAs are deleted or wherein about the last 18 C-terminal (corresponding to MLV-integrase AA390-AA408)
35 AAs are deleted. In another embodiment the BET interaction motif (AA390-AA305) is deleted.

Chromatin binding domains (CDB) in the context of the present of the present invention relates to proteins or fragments thereof which bind to chromatin, more specifically to chromatin in silent regions of the genome. The sequence identity
 5 between different chromatin binding proteins and the chromatin binding portion therein may differ significantly. The suitability of a given chromatin binding domain for the purpose of the present invention can be assessed by performing the examples of the present application

Typically, chromatin binding domains contain a chromo (CHRromatin Organization
 10 MOdifier) structural portion, which consists of an SH3-like beta-barrel capped by a C-terminal helix. These domain is described in the InterPro EMBL-EBI database. Chromodomains are conserved modules that are implicated in the recognition of lysine-methylated histone tails and nucleic acids. Chromodomains were originally identified in *Drosophila* modifiers of variegation, proteins that alter the structure of
 15 chromatin to the condensed morphology of heterochromatin.

Chromodomains can be found in various nuclear proteins, including heterochromatin protein 1 (HP1) (N-terminal chromo domain and C-terminal chromo shadow domain), where the chromodomain recognises histone tails with specifically methylated lysines polycomb protein Pc, which is essential for maintaining the silencing state of homeotic
 20 genes during development (chromodomain important for chromatin targeting); histone methyltransferase clr4, which regulates silencing and switching at the mating-type loci and to affect chromatin structure at centromeres and the ATP-dependent helicase CHD1, which regulates ATP-dependent nucleosome assembly and mobilisation through conserved double chromo domains and a SWI2/SNF2
 25 helicase/ATPase domain

Based upon a limited set of 5 proteins a consensus sequence for chromodomains is defined in Sing (1991) *Nucleic Acids Research* **19**, 789 namely

yv vEKvldrRVv kGkVEYIIKW KGfsdedNTW KPKeN [SEQ ID NO: 13], wherein amino acids in capital letters are absolutely conserved (note amino acids 7, 19 and 20 are
 30 lower case Isoleucine). The invention envisages constructs comprising chromodomains with more than 80, 85, 90 or 95 sequence identity with SEQ ID NO: 13, whereby 1, 2, 3 or 4 of the absolute conserved of SEQ ID NO: 13 amino acids have been modified and wherein the chromodomain retains its chromatin binding properties.

The CBD which are used for the purpose of the present invention can be derived from proteins of the organism which will be targeted with the gene transfer constructs of the invention. Thus for human gene transfers the modified integrase of the present invention will be a fusion between a retroviral protein and a human protein. If the sequences of the CBD are strongly conserved between species, the link between the origin of the CBD and the target cell to be treated may be less strict.

5 CBD are in the context of the present invention typically peptides that bind markers or proteins that are widespread across the chromatin. This can be cellular proteins that recognize specific epigenetic histone modifications, such as the chromodomain of heterochromatin-binding protein 1 β and chromodomain Y-like protein
10 Alternatively, virus-derived peptides can be used, such as the tethering domain of the human papilloma virus (HPV8) E2 and the N-terminal end of Kaposi sarcoma's latency associated nuclear antigen that bind to core histone 2A.

15 In typical embodiments, the CBD is the latency associated nuclear antigen (LANA) peptide (SEQ ID NO: 1), the heterochromatin-binding protein 1 β (CBX1) (SEQ ID NO: 2), the chromodomain Y-like protein (CDYL) (SEQ ID NO: 3), the human papilloma virus (HPV8) E2 protein (SEQ ID NO: 4) or a variant thereof having at least 80% AA, preferably at least 85%, more preferably at least 90%, most preferably
20 95% sequence identity to SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3 or SEQ ID NO: 4. In a further specific embodiment, said CBD is the latency associated nuclear antigen (LANA) peptide (SEQ ID NO: 1) or the heterochromatin-binding protein 1 β (CBX1) (SEQ ID NO: 2) or a variant thereof having at least 80% AA, preferably at least 85%, more preferably at least 90%, most preferably 95% sequence identity to
25 SEQ ID NO: 1 or SEQ ID NO: 2.

Variants of CBD such as recited above and/or truncated versions thereof are envisaged to the extent that they still have chromatin binding properties.

The invention further provides for a host cell transfected with said retroviral plasmid. In particular embodiments said retroviral plasmid nucleotide sequence is stably
30 integrated in the cellular genome.

The invention further provides an infectious viral vector particle as well as the process for obtaining the infectious viral vector particle, wherein a helper cell is transfected with a retroviral plasmid according to any one of the embodiments above, a transfer plasmid comprising self-inactivating (SIN) LTR sequences, 5' LTR and 3'LTR
35 sequences and a transgene or a heterologous nucleotide sequence, and an envelope

plasmid comprising an ENV nucleotide sequence and said helper cell is cultured in a suitable medium under conditions for the release of infectious viral vector particles. In particular embodiments the infectious viral vector particle further comprises a promoter sequence, suitable for expression in a mammalian cell and wherein said
5 heterologous nucleotide sequence is a transgene that can be expressed via said promoter.

The transgene that may be contained in the infectious viral vector particle envisaged herein typically encodes a gene product such as RNA or a polypeptide (protein). Those skilled in the art will appreciate a variety of transgenes that are suitable for
10 use with the invention. In certain embodiments, the transgene encodes a therapeutic protein. Non-limiting examples of therapeutic proteins include clotting factors, such as factor VIII or factor IX, insulin, erythropoietin, lipoprotein lipase, antibodies or nanobodies, growth factors, cytokines, chemokines, plasma factors, etc. In certain other embodiments, the transgene encodes an immunogenic protein. Non-limiting
15 examples of immunogenic proteins include epitopes and antigens derived from a pathogen.

It is envisaged that the retroviral gene therapy methods envisaged herein can be applied e.g. to express a therapeutic amount of a transgene product (such as a polypeptide, in particular a therapeutic protein, or RNA) for therapeutic purposes, to
20 express a prophylactically effective amount of a transgene product to prevent the onset of a disease or disorder, or to express an immunological amount of a transgene product (such as a polypeptide, in particular an immunogenic protein, or RNA) for vaccination purposes. Exemplary diseases and disorders that may benefit from the retroviral gene therapy described herein include genetic disorders (such as
25 haemophilia, including haemophilia A and B, β -thalassemia, muscular dystrophy (e.g. Duchenne muscular dystrophy (DMD))), diabetes, cancer, infectious diseases and the like.

Alternatively, the retroviral gene therapy methods envisaged herein can be applied to disrupt a defective gene. In particular embodiments the gene therapy methods
30 envisaged herein introduce elements which ensure DNA editing, using techniques such as zinc finger nucleases and CRISPR. The vector incorporates genes encoding the elements required for editing into chromosomes and these elements then edit the relevant chromosome in a specific way.

The retroviral gene therapy methods envisaged herein includes the delivery of a
35 transgene into a host cell and may be performed ex vivo or in vivo. The ex vivo approach requires harvesting of the host cells from a subject, in vitro transduction of

the host cells with an infectious viral vector particle comprising the transgene, and re-introduction of the transduced host cells into the subject. The in vivo approach requires the administration of an infectious viral vector particle comprising the transgene directly to a subject. In particular embodiments of the methods envisaged
5 herein, the retroviral gene therapy is in vivo gene therapy.

Accordingly, in particular embodiments, the methods provided herein involve contacting a host cell with an infectious viral vector particle comprising a transgene. Where it is envisaged that the host cell is contacted with the infectious viral vector particle ex vivo, the host cell may be part of an isolated host cell population (e.g.
10 blood or a fraction thereof) or an isolated tissue or organ. Thus in particular embodiments the methods may comprise the step of providing a host cell. While this step is not critical to the present invention, in particular embodiments the method may involve isolating the host cell or host cell population or tissue/organ from the host.

15 In other embodiments the invention provides for a host cell that has been transduced with said infectious viral vector particle. In yet another embodiment, the invention provides for the in vitro use of said infectious retroviral vector particle for the delivery of a heterologous sequence or transgene to a host cell, preferably in gene therapy, wherein more preferably said gene therapy comprises selecting the host cells in which
20 said heterologous sequence or transgene has been integrated in a "safe" genomic location. In a more particular embodiment the host cell is a haematopoietic cell or a stem cell or more particularly a haematopoietic stem cell or a progenitor cell.

The invention further provides for the use of said infectious retroviral vector particle for preparing a pharmaceutical composition for gene therapy, wherein said infectious
25 retroviral vector particle is present in a form suitable for being introduced into the target cells. In another embodiment, the invention provides for a pharmaceutical composition that comprises said infectious retroviral vector particle together with a pharmaceutically acceptable excipient and/or carrier. In yet another embodiment, the invention further provides for a method of gene therapy, which method comprises
30 administering said infectious retroviral vector particle to a patient. Where it is envisaged that the target cell is contacted with an infectious retroviral vector particle in vivo, this methods encompass administering the infectious retroviral vector particle to the patient comprising said target cell. The infectious retroviral vector particle may be delivered in vivo to the subject in a formulation or a pharmaceutical composition
35 with a pharmaceutically acceptable carrier, i.e., one or more pharmaceutically acceptable excipients and/or additives, e.g., buffers, carriers, excipients, stabilizers,

etc. The term "pharmaceutically acceptable" as used herein is consistent with the art and means compatible with the other ingredients of a pharmaceutical composition and not deleterious to the recipient thereof. In vivo gene therapy protocols have been extensively described in the art. These include, but are not limited to, intramuscular
5 injection, microinjection, hydrodynamic gene delivery in various tissues, interstitial injection, instillation in airways, application to endothelium, intra-hepatic parenchyme, and intravenous or intra-arterial administration. Various devices have been developed for enhancing the availability of DNA to the target cell. A simple approach is to contact the target cell physically with catheters or implantable
10 materials containing DNA. Another approach is to utilize needle-free, jet injection devices which project a column of liquid directly into the target tissue under high pressure.

Methods of determining the most effective means and dosage of retroviral vector administration are well known to those of skill in the art and will vary with the vector
15 used for therapy, the purpose of the therapy, the host cell being treated, and the subject being treated.

The methods are envisaged herein for genetic modification of a host cell of an animal, more particularly a mammal, such as but not limited to a human.

The invention will be illustrated by the following non-limiting examples.

20

Examples

Materials and methods

Plasmids

pcDNA3_MLV_gagpol packaging plasmids was a kind gift from Prof. Axel Schambach,
25 Hannover, Germany. BinMLV integrase are cloned as previously described (Ashkar et al. cited above). Chromodomains of CBX and CDYL fusions were cloned with gblocks (IDT, Haasrode, Belgium) in PacI and NotI digested pC1_608 packaging plasmid. HPV8 E2 and LANA peptide fusions were introduced into the indicated vectors by oligonucleotide annealing strategy with the same restriction sites. Oligonucleotide
30 sequences are listed in table 1. All enzymes were purchased from Thermo Fischer Scientific, St. Leon-Rot, Germany). The integrity of all plasmids was verified by DNA sequencing.

Table 1: Oligonucleotide sequences

Name	Sequence (5'-3')
CD CBX	<i>cataacgatcgagagtacgtggtggagaagggtgctggaccgccgctggtgaagg gcaagggtggagtacctgctgaagtggaagggcttcagcgacgaggacaacacctgggag cccgaggagaacctggactgccccgacctgatcgccgagttcctgcagagccagaagacct agcggccgccataa [SEQ ID NO:14]</i>
CD CDYL	<i>cataacgatcgaatgacctccaggccagccaccgcagcgcctggggcaagagcc gcaagaagaactggcagtagagggccccaccagaagctgttctgaagcgcaacaac gtgagcgccccgacggccccagcgacccagcatcagcgtgagcagcgaGcagagcgg cgcccagcagccccccgcctagcggccgccataa [SEQ ID NO:15]</i>
E2_Fw	<i>cagaccgagaccaagggccgcccgtacggccgcccagcagccgcta [SEQ ID NO:16]</i>
E2_Rv	<i>ctagcggctgctggggcgggcgccgtagcggcgcccttggtctcggtctgat [SEQ ID NO:17]</i>
LANA_Fw	<i>atggcccctcctggcatgcgctgcgcagtggccgcagcaccggcgcgcccttgacccgcg gcagctgccgcaagcgcaaccgcagccccgagta [SEQ ID NO:18]</i>
LANA_Rv	<i>ggcctactcggggctgcggttgctgcttcggcagctgccgcgggtcagaggcgcgccggt gctgcggccactgcgcaggcgcatgccaggagggccatatg [SEQ ID NO:19]</i>
eGFP_FW	<i>ggagcgcaccatcttctca [SEQ ID NO:20]</i>
eGFP_Rv	<i>aggggtgtcgccctcgaa [SEQ ID NO:21]</i>
eGFP_probe	<i>gctacaagacccgcgcccagg [SEQ ID NO:22]</i>

Cell culture

- 5 SupT1 cells were cultured in Roswell Park Memorial Institutes medium (RPMI-1640, Gibco BRL, Life Technologies) supplemented with 10% heat-inactivated foetal bovine serum (Gibco BRL) and gentamicin (50µg/ml, Gibco-BRL). HeLa cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Gibco-BRL) supplemented with 8% heat inactivated foetal bovine serum and gentamicin. All cells are grown in a
- 10 humidified atmosphere with 5% CO₂ at 37°C.

T-cell purification

Peripheral blood mononuclear cells were purified from a buffy coat using density-gradient centrifugation (Lymphoprep; Axis-Shield PoC AS, Oslo, Norway). Primary CD4⁺ T-cells were isolated using negative selection (MACS; Miltenyi Biotec, Leiden, the Netherlands) and stimulated with CD2, CD3 and CD28 beads (MACS).

Retroviral vector production and transduction

Viral vectors were produced as previously described (Ibrahimi, A. et al., (2009) *Human Gene Therapy*, **20(8)**, 845–860). Briefly, MLV-based vectors were produced by a triple PEI based transfection of 293T cells with pVSV-G envelope, pcDNA3_MLV_gagpol packaging plasmids or its derived fusions (see above) and p450-GFP transfer plasmid (kindly provided by F.D. Bushman, Philadelphia, NJ, USA). Vector titres, represented as reverse transcriptase units (RTU), were determined by the SYBRGreen-I product-enhanced reverse transcriptase assay (SG-PERT) (Pizzato, M. et al., (200) *J. Virol. Methods* **156(1-2)**, 1–7). For transduction, SupT1 cells (12*10⁴/well), HeLa cells (2*10⁴/well) and CD4⁺ T-cells (20*10⁴/well) were seeded in 96-well plates and subsequently transduced with a dilution series of the respective vectors. Seventy-two hours post transduction, 50% of the cells were harvested for fluorescence-activated cell sorting (FACS) analysis, while the remaining 50% were cultured for 10 days for FACS analysis, to determine integrated copies and to perform integration site analysis.

gDNA isolation and quantitative PCR

gDNA isolation and qPCR are performed as previously described (Ashkar et al. cited above). Briefly, two million cells were pelleted and genomic DNA was extracted using a mammalian genomic DNA miniprep kit (Sigma-Aldrich, Bornem, Belgium). Samples corresponding to 700ng genomic DNA were used for analysis. Each reaction contained 12.5µl iQ Supermix (Biorad, Nazareth, Belgium), 40nM forward and reverse eGFP primer and 40nM of eGFP probe in a final volume of 25µl. RNaseP or β actin were quantified as endogenous controls (TaqMan RNaseP control reagent, Applied Biosystems, The Netherlands). Samples were run in triplicate for 3 minutes at 95°C followed by 50 cycles of 10 seconds at 95°C and 30 seconds at 55°C in a LightCycler® 480 (Roche-applied-science, Vilvoorde, Belgium). Analysis was performed using the LightCycler® 480 software supplied by the manufacturer.

Recovery of integration sites and analysis of integration site distributions

Recovery of integration sites was performed as previously described (De Rijck, J. et al. (2013) *Cell reports*, **5(4)**, 886–894). Briefly, linkers were ligated to restriction

enzyme-digested (MseI) genomic DNA isolated from transduced cells and virus-host DNA junctions were amplified by nested PCR. Samples were individually barcoded with the second pair of PCR primers to generate 454 libraries. PCR products were purified and sequenced using 454/Roche pyrosequencing (titanium technology).

5 Reads were quality-filtered by requiring perfect matches to the long terminal repeat (LTR) linker, barcode, and flanking LTR and subsequently mapped to the human/mouse genome. All sites were required to align to the reference genome within 3bp of the LTR edge. In order to control for possible biases in the datasets due to the choice of the MseI restriction endonuclease in cloning integration sites,

10 matched random control (MRC) sites were generated in silico. To do so, each experimental integration site was paired with three sites in the genome, locating at the same distance from a randomly selected MseI site in the genome.

Analyses were carried out as described (Marshall, H.M. et al., (2007) *PLoS ONE*, **2(12)**, e1340). A detailed account of the statistical methods used and the methods

15 for forming and analysing heat maps using ROC curves can be found in (Brady, T. et al., (2009) *Genes & Development*, **23(5)**, 633–642). Consensus sequence analysis at the point of integration was performed using WebLogo (<http://weblogo.Berkeley.edu/logo.cgi>; Crooks GE, et al. (2004) *Genome Research*, 14; 1188-1190). For association with specific genomic features, the distance of each

20 integration site (in kb) to the respective genomic feature was calculated (midpoint of CpG island or DHS, and X5-end of genes as a measure for TSS). Integration sites left of the genomic feature were given negative kb values, while integration sites towards the right were calculated as positive. For heat maps, comparisons were carried out over three different interval sizes surrounding each integration site (5kb, 10kb, and

25 50kb), since previous studies have shown that the interval sizes chosen for comparison can influence the conclusions. Results of statistical tests comparing the distributions of integration sites to the reference dataset are summarized as asterisks on each tile of the heat map.

30 **In vitro immortalization (IVIM) assay**

The IVIM assay was performed as described earlier (Modlich, U. et al., (2006) *Blood*, **108(8)**, 2545–2553). Briefly, bone marrow cells were isolated from the tibias and femurs of C57Bl/6 mice and enriched for stem and progenitor cells (lineage depletion kit, Miltenyi Biotec) and frozen in aliquots. After thawing and 48 hours of

35 prestimulation, 1x10⁵ cells were transduced on two consecutive days on Retronectin coated wells with MOI of 5, 10 or 20 each day. Transduction efficiency was analysed

by flow cytometry four days thereafter. Cells were expanded for two weeks in IMDM, 10% FCS, 1mM glutamine, 1% (v/v) Penicillin/Streptomycin, mSCF (50 ng/ml), hFlt3L (100 ng/ml), mIL-3 (20 ng/ml) and hIL-11 (100 ng/ml, all cytokines purchased from Peprotech, Hamburg, Germany) and diluted to a cell density of 500 000 cells/ml approximately twice a week. Cells were the seeded on 96-well suspension plates at a density of 100 cells per well (48 wells seeded from each culture). Replating clones were detected by microscopic scoring. The replating frequency (according to Poisson distribution) was calculated with L-calc (StemCell Technologies, Vancouver, Canada) and normalized by vector copy number as determined five days post transduction.

10

Efficient transduction and integration of next generation BinMLV vectors

To direct BinMLV integration away from potentially harmful chromosomal sites, we interfered with the chromatin-tethering process by fusing tethering peptides (between 16-61 amino acids (a.a.) long) to the C-terminal end of INW390A in the MLV packaging plasmid. We opted for peptides that bind markers or proteins that are widespread across the chromatin (Table 2 and Figure 2A). On the one hand, we used peptides derived from cellular proteins that recognize specific epigenetic histone modifications, such as the chromodomain of heterochromatin-binding protein 1 β (CBX1, a.a. 20-73) and chromodomain Y-like protein (CDYL, a.a. 1-60) (Kaustov, L. et al., (2011) *J. Biol. Chemistry*, **286(1)**, 521–529), giving rise to INW390A-CBX and INW390A-CDYL, respectively. Alternatively, we made use of virus-derived peptides, such as the tethering domain of the human papilloma virus (HPV8) E2 protein (a.a. 240-255) (Sekhar, V. et al. (2010) *J. Virology* **84(1)**, 543–557) and the N-terminal end of Kaposi sarcoma's latency associated nuclear antigen (LANA, a.a. 1-31) that bind to core histone 2A and 2B (Barbera, A.J. et al., (2006). *Science*, **311(5762)**, 856–861), resulting in INW390A-E2 and INW390A-LANA, respectively.

25

Table 2: Overview of the tethering domains fused to BinMLV vectors.

Protein [Accession Number]	Selected Peptide	Binding Site	Sequence
CBX1 [P83916]	CD CBX	H3K9me2 H3K9me3	EYVVEKVLDRRVVKGKVEYLLKWKGFSDEDNT WEPEENLDCPDLIAEFLQSQKT [SEQ ID NO: 2]
CDYL [Q9Y232]	CD CDYL	H3K9me2 H3K27me2 H3K27me3	MTFQASHRSAWGKSRKKNWQYEGPTQKLFLK RNNVSAPDGPSPSISVSSEQSGAQPPA [SEQ ID NO: 3]
HPV8 E2 [P06422]	E2	H2A/H2B	QTETKGRRYGRRPSSR [SEQ ID NO: 4]
KSHV [E5LC01]	LANA	H2A/H2B	MAPPGMRLRSGRSTGAPLTRGSCRKRNRSP [SEQ ID 1]

Summary of the amino acid sequences and binding sites of CBX1 and CDYL chromodomains, HPV8 E2 and KSHV LANA used in this study. Amino acid position of each peptide in the respective protein is indicated in superscript. CD; chromodomain, CBX1; chromobox homolog 1, CDYL; chromodomain protein Y-Like, HPV; human papilloma virus, KSHV; Kaposi sarcoma human virus, LANA; latency associated nuclear antigen. Protein sequences were downloaded from the UniProt database.

The respective packaging plasmids were subsequently used to produce VSV-G pseudotyped MLV-based vectors encoding eGFP (MLVIN_W390A-CBX, MLVIN_W390A-CDYL, MLVIN_W390A-E2 and MLVIN_W390A-LANA; Figure 2A). In line with previous results (Ashkar et al. cited above), transduction efficiency of MLVIN_W390A in SupT1 cells was comparable for MLVIN_W390A and MLVIN_WT (Figure 2B). Addition of peptide sequences to the C-terminal end of MLV INW390A resulted in BinMLV-vectors that transduced as efficiently as WT vector particles at different vector dilutions (Figure 2B), and resulted in comparable expression levels (measured as mean fluorescent intensities at day 3; Figure 2C). Transduction data and MFI were corroborated at 10 days post transduction, underscoring stable expression for the respective integrated vector constructs (Figure 7A-B). Similar data were obtained following transduction of HeLa cells (data not shown). Collectively, these results indicate that addition of peptide-sequences at the C-terminal end of MLVIN_W390A does not impair vector integrity nor transduction efficiency compared to MLVIN_WT or MLVIN_W390A.

MLV IN_W390A-peptide fusions redistribute BET-independent (Bin) vector integration

We next set out to investigate whether the respective IN-chimeras redistribute MLV vector integration. Integration sites were amplified and sequenced, yielding a total of 43676 unique sites and their computationally generated matched random control (MRC) sites. Initially, we evaluated integration frequencies relative to transcription start sites (TSS), CpG islands (typically enriched in/near housekeeping gene promoters) and DNase hypersensitive sites (DHS), both surrogate markers for open areas of active chromatin. As expected, MLVIN_WT integration sites were enriched in all three features (Figure 3A and Table 3A). In accordance with previous work (Ashkar et al. cited above), MLVIN_W390A integration near TSSs and CpG islands decreased ~2-fold (Figure 3A and Table 3A). Whereas integration frequencies for MLVIN_W390A-CDYL and MLVIN_W390A-E2 did not differ from MLVIN_W390A near TSSs or CpG islands, the integration frequencies for both MLVIN_W390A-CBX and MLVIN_W390A-LANA were ~4-fold and ~2-fold lower when compared to MLVIN_WT and MLVIN_W390A, respectively (Figure 3A and 7A, $p < 0.001$, compared to MLVIN_W390A). Nonetheless, integration was still enriched near these features when compared to MRC ($p < 0.001$). The detargeting effect of MLVIN_W390A-CBX and MLVIN_W390A-LANA was also evident near DHS; whereas 33% of MLVIN_W390A integrations occurred in a 2kb window around DHS (~13% less than MLVIN_WT), only 25.66% and 27.84% of MLVIN_W390A-CBX and MLVIN_W390A-LANA integrations occurred in this window (Figure 3A and 7A). Comparable data were obtained for larger window sizes (4 kb window, data not shown). Together, the data confirm that these peptide additions to IN generally shift vector integration away from the traditional markers associated with MLV integration.

Table 3: Integration site distribution of next generation BinMLV vectors.
Integration sites of BinMLV vectors engineered by peptide fusion to MLVIN_W390A (A) or MLVIN_1-380 (B). Genomic distribution in 2 kb windows around TSS, CpG island midpoints and DHS are listed. For comparison, integration of computationally generated match random controls (MRC) for MLVIN_W390A and MLVIN_1-380 are shown. P values () show significant departures (***) $p < 0.001$, pairwise Fishers test) from MLVIN_W390A or MLVIN_1-380 and MRC, respectively, separated by |. TSS; transcription start sites, CpG; CpG-rich islands, DHS; DNase I-hypersensitive sites.*

A

Vector	Total Sites	%TSS within 2kb	%CpG within 2kb	%DHS within 2kb
MLV _{IN_WT}	9963	20.65 *** ***	20.87 *** ***	45.96 *** ***
MLV _{IN_W380A}	10088	10.17 ***	10.16 ***	33.05 ***
MLV _{IN_W380A-CBX}	5258	6.12 *** ***	6.41 *** ***	25.66 *** ***
MLV _{IN_W380A-CDYL}	2598	10.93 ***	10.35 ***	33.83 ***
MLV _{IN_W380A-E2}	6595	10.46 ***	10.36 ***	32.99 ***
MLV _{IN_W380A-LANA}	9174	6.62 *** ***	7.33 *** ***	27.84 *** ***
MRC	30234	2.27	2.27	14.65

B

Vector	Total Sites	%TSS within 2kb	%CpG within 2kb	%DHS within 2kb
MLV _{IN_WT}	9963	20.65 *** ***	20.87 *** ***	45.96 *** ***
MLV _{IN_1-380}	5744	10.08 ***	10.24 ***	32.54 ***
MLV _{IN_1-380-CBX}	3738	6.45 *** ***	6.37 *** ***	26.51 *** ***
MLV _{IN_1-380-CDYL}	2598	9.95 ***	10.61 ***	33.8 ***
MLV _{IN_1-380-E2}	4140	10.14 ***	10.72 ***	35.92 ***
MLV _{IN_1-380-LANA}	2424	5.98 *** ***	6.85 *** ***	27.31 *** ***
MRC	17226	2.49	2.43	15.19

In a next step, we analysed integration preferences relative to a wider set of genomic features to evaluate overall vector integration (Figure 3B). Tile colour depicts the correlation for an integration dataset with the respective genomic feature (left) relative to matched random control (MRC), as quantified by the area under the receiver operating characteristic (ROC) curve. MLVIN_W390A-CDYL and MLVIN_W390A-E2 integration profiles were similar to that of MLVIN_W390A, while MLVIN_W390A-CBX and MLVIN_W390A-LANA showed significant differences. Addition of the CBX chromodomain (MLVIN_W390A-CBX) significantly shifted integration frequencies for most of the selected genomic features, redistributing integration towards a more random pattern. On the other hand, MLVIN_W390A-LANA only showed significant effects for the typical determinants of MLV-vector integration, such as DHS and CpG islands in the smaller window sizes.

Since some of the peptides specifically recognize chromatin marks, we also analysed integration preferences near a collection of 41 epigenetic features (Figure 3C). In line with previous data, uncoupling of BET interaction (MLVIN_WT vs. MLVIN_W390A) yields a integration pattern that moves towards MRC (Ashkar et al. cited above). Albeit to a lesser extent, MLVIN_W390A retroviral integration still correlates with histone marks associated with open and transcriptionally active chromatin and disfavours transcriptionally silent regions or heterochromatin (De Ravin et al. 2014). Fusion of the CBX1 chromodomain and the LANA peptide to INW390A results in a more randomized distribution for markers correlating with transcriptionally active open chromatin (shifting towards MRC; blue tiles overall shift towards black), whereas addition of the other peptides had no effect (Figure 3C). CBX1 is known to bind H3K9me2/3 epigenetic marks via its chromodomain (Kaustov et al. 2011). In-depth analysis revealed that MLVIN_W390A-CBX integration shifts more into transcriptionally silent heterochromatin regions, generally disfavoured for integration, marked by di- and/or trimethylation of H3K9 and H3K27.

Together, these data show that fusion of peptide-tethers to the C-terminal end of MLV INW390A can effectively retarget integration. As integration is shifted away from traditional MLV integration markers, known to associate with insertional mutagenesis in gene therapeutic trials, a potentially safer integration site profile might be obtained.

Addition of small peptides to MLVIN_1-380 rescues its integration defects

Aside from W390A substitution, deletion of the C-terminal tail of MLV IN (Δ 23 aa, IN1-380) similarly detargeted MLV integration. However, deletion resulted in 2-3 fold

lower integration efficiency, possibly due to IN destabilization (Ashkar et al. cited above). Therefore we asked whether this defect could be rescued by the addition of the respective peptide fragments to IN1-380 (MLVIN_1-380-CBX, MLVIN_1-380-CDYL, MLVIN_1-380-E2 and MLVIN_1-380-LANA; Figure 4A).

5 In line with previous results (Ashkar et al. cited above), MLVIN_1-380 transduction efficiency decreased ~3 fold compared to MLVIN_WT. Interestingly, addition of any of the peptides to MLVIN_1-380 improved transduction efficiency at different vector dilutions (Figure 4B) and resulted in comparable expression levels (MFI at day 3; Figure 4C). These data were corroborated at 10 days post transduction (Figure 7C-
10 D). Similar results were obtained in HeLa cells (data not shown). Next, we asked whether these IN-chimeras also redistributed the respective MLV vector integration profiles. Integration site amplification yielded 28607 unique integration sites and their computationally generated MRC sites (Table 3B). As in previous work (Ashkar et al. cited above), MLVIN_1-380 integration decreased near TSSs, CpG islands and
15 DHS (Figure 8A and Table 3B). Fusion of the CBX and LANA peptides to MLVIN_1-380 resulted in comparable integration site distributions as observed for MLVIN_W390A-CBX and MLVIN_W390A-LANA (compare Table 3A and 3B). Additionally, more detailed analysis using genomic and epigenetic heat maps corroborated the results observed for MLVIN_W390A: integration of MLVIN_1-380-
20 CBX shifted more towards random compared to MLVIN_WT and MLVIN_1-380 for a range of genomic features (Figure 8B), and histone modifications (Figure 8C). Together, these data highlight the specificity of the detargeting effects achieved by the fusion of CBX and LANA peptides to MLVIN_W390A and MLVIN_1-380.

25 **Addition of peptide tethers does not alter the local MLV integration site sequence**

Retroviral INs show weak but discernible target sequence preferences surrounding the site of integration. This local integration site sequence is mainly determined by IN contacts with the (nucleosomal) DNA template (Wu et al. 2005; Holman et al.
30 2005). To assess whether addition of alternative peptide tethers to MLV INW390A influenced the local integration site sequence, we constructed sequence logos (Figure 9). As expected, results indicate that the local integration site sequence preferences remained unaffected. Similar results were obtained for MLVIN_1-380 peptide fusions.

35 **Evaluation of the gene therapeutic potential of next generation BinMLV vectors**

Comparative integrome analysis let us single out MLVIN_W390A-CBX and MLVIN_W390A-LANA as BinMLV vectors that demonstrated substantial detargeting from traditional MLV markers without compromising transduction efficiency. In a next step, we determined the safety profile for the respective next generation BinMLV vector constructs by determining integration frequencies near a set of previously defined genomic regions that should be avoided to prevent transformation (Sadelain et al. 2012; Papapetrou et al. 2011), such as those proximal to transcription start sites (<50kb TSS), to cancer-related oncogenes (<300kb AllOnco), or to any miRNA coding regions (<300kb miRNA), or within transcription units (TU) and ultraconserved regions (UCR). For each integrome data set the percentage of unsafe integrations near these features was determined (Figure 5A). In addition, we calculated the integration events that were not captured by any of the above-mentioned criteria (fraction of safe integrations; Figure 5A). All BinMLV vector designs resulted in a larger fraction of safe integrations compared to MLVIN_WT (12.4%, 15.5% and 14.1% for MLVIN_W390A, MLVIN_W390A-CBX and MLVIN_W390A-LANA, respectively, compared to 10.6% for MLVIN_WT). Moreover, when assessing the individual safe harbour criteria, MLVIN_W390A-CBX integrated less frequent near each of these features when compared to both MLVIN_WT and MLVIN_W390A, suggesting an overall safer integration profile that is less likely to disturb nearby genes. Of note, evaluation of the integromes of the respective MLVIN_1-380 vectors corroborated these safety profiles (Table 3).

Next, we validated the potential of next generation BinMLV vectors by (i) transducing more relevant primary cells and (ii) by assessing the integration profile safety relative to MLVIN_WT and MLVIN_W390A. To judge the therapeutic potential of next generation BinMLV vectors, we challenged primary CD4+ T-cells with MLVIN_WT, MLVIN_W390A, MLVIN_W390A-CBX and MLVIN_W390A-LANA. In line with SupT1 cells, MLVIN_W390A-CBX and MLVIN_W390A-LANA showed transduction efficiencies (percentage of GFP+ cells) and mean fluorescence intensities (MFI) that were comparable to MLVIN_WT and MLVIN_W390A at 3 days post transduction (Figure 5B). Re-analysis at 10 days post transduction confirmed sustained expression levels, even for the vector constructs that display retargeted integration preferences (Figure 5B, compare lower and upper panels). Likewise, transduction at different multiplicities of infection (MOI) of mouse hematopoietic lineage depleted (lin-) bone marrow cells resulted in transduction efficiencies for MLVIN_W390A-CBX and MLVIN_W390A-LANA that were comparable to those of MLVIN_WT and MLVIN_W390A (Figure 5C). In accordance, efficient integration was corroborated by

comparable integrated vector copy numbers (VCN) for the respective MOIs (Figure 5D). Collectively, these data demonstrate that the next generation BinMLV vectors perform in line with MLVIN_WT, but display an improved integration pattern in potentially safer regions.

5

Integration site safety is improved for MLV.SIN.SFIN_W390A-CBX vector in IVIM assay

An essential component in the validation of new viral vectors includes the evaluation of the genotoxicity. To estimate whether the altered integration profile of our MLV vectors carrying IN peptide-fusions is potentially safer, we determined the incidence of immortalizing events in a clonal dominance assay (in vitro immortalization assay (IVIM); (Modlich et al. 2006 cited above). To this end, we packaged a self-inactivating (SIN) gammaretroviral vector genome (MLV.SIN) that contains the spleen focus-forming virus enhancer/promoter (SF) as an internal promoter, known to trigger insertional transformation events (Modlich, U. et al., (2008) *Leukemia*, **22(8)**, 1519–1528; Zychlinski, D. et al. (2008) *Molecular Therapy* **16(4)** 718–725) in MLV-particles that contain INWT, INW390A, INW390A-CBX or INW390A-LANA (MLV.SIN.SFIN_WT, MLV.SIN.SFIN_W390A, MLV.SIN.SFIN_W390A-CBX and MLV.SIN.SFIN_W390A-LANA; Figure 6A).

20 Mouse hematopoietic lineage depleted (lin⁻) bone marrow stem cells were transduced with each vector at different MOIs in independent IVIM assays. To ensure clonal outgrowth, at least 55% of cells (transduction efficiency, TE) are to be transduced, corresponding to a mean copy number of >2 (Kustikova 2003) (Figure 6B). The MLV.SIN.SF genome architecture resulted in transformed lin⁻ cells for all MLV vector contexts tested (Figure 6B-C). The number of clones in an assay is a measure for the incidence of clonal transformation events. MLV.SIN.SFIN_W390A and MLV.SIN.SFIN_W390A-LANA resulted in replicating clone numbers in line with MLV.SIN.SFIN_WT, whereas a reduction in clone numbers was observed for MLV.SIN.SFIN_W390A-CBX (Figure 6B and 6C). The mean replating frequency of IVIM clones was comparable for MLV.SIN.SFIN_WT, MLV.SIN.SFIN_W390A and MLV.SIN.SFIN_W390A-LANA (4.2×10^{-3} , 4.8×10^{-3} and 4.1×10^{-3} , respectively; Figure 6B), while the replating frequency for MLV.SIN.SFIN_W390A-CBX was reduced (2.7×10^{-3}). Besides the incidence of clonal transformation, the fitness of the immortalized cells is estimated by correcting the replating frequency for the mean number of integrated vector copies (mVCN). MLV.SIN.SFIN_W390A-CBX displayed a

35

reduced replating frequency/copy number (repl. freq/copy no.), when compared to any of the other vector contexts (Figure 6B and 6D).

Taken together, these data indicate that tinkering with the physical MLV vector configuration by modifying the MLV IN protein allows MLV.SIN.SFIN_W390A-CBX to
5 integrate in regions that are potentially safer, and demonstrated a reduced genotoxic outcome for a gammaretroviral vector genome design known to elicit insertional mutagenesis (MLV.SIN.SF).

Claims

1. A retroviral plasmid comprising a nucleotide sequence which codes for a gamma-retroviral-Integrase (IN), which is independent from bromodomain and extra-terminal (BET) family of proteins, wherein said nucleotide sequence is operably linked to a nucleotide sequence which codes for a chromatin binding domain (CBD) or a chromatin binding fragment thereof.
2. The plasmid according to claim 1, wherein said gamma-retroviral-IN is a Murine Leukemia Virus (MLV)-IN.
3. The plasmid according to claim 1 or 2, wherein said gamma-retroviral-IN codes for a C-terminal modified retroviral-IN, wherein at least one amino acid in the BET interaction domain is deleted or mutated, wherein the BET interaction domain corresponds to amino acids 389-405 of MLV-IN.
4. The plasmid according to claim 3, whereby the conserved W of said BET interaction domain, corresponding to W390 of MLV-IN, is mutated or deleted.
5. The plasmid according to claim 3 or 4, comprising a W390A mutation.
6. The plasmid according to claim 5, wherein said integrase is a C-terminal modified integrase, wherein the region corresponding to amino acids 382 to 408 of MLV-integrase is deleted.
7. The plasmid according to any one of claims 1 to 6, wherein the chromatin binding domain or chromatin binding fragment thereof comprises a chromodomain.
8. The plasmid according to any one of claims 1 to 6, wherein the chromatin binding fragment is the peptide from latency associated nuclear antigen (LANA) peptide with SEQ ID NO: 1, the peptide of heterochromatin-binding protein 1 β (CBX1) with SEQ ID NO: 2, or a chromatin binding variant thereof having at least 80% sequence identity to SEQ ID NO: 1 or SEQ ID NO: 2.

9. A host cell transfected with a retroviral plasmid according to any one of claims 1 to 8.
10. The host cell of claim 9, wherein said retroviral plasmid is stably integrated in the cellular genome.
11. A method for obtaining an infectious viral vector particle, comprising the steps of :
a) transfecting a helper cell with :
-a retroviral plasmid according to any one of claims 1 to 8,
-a transfer plasmid comprising self-inactivating (SIN) LTR sequences, 5' LTR and 3'LTR sequences and a heterologous nucleotide sequence, and
- an envelope plasmid comprising an ENV nucleotide sequence and
b) culturing said transfected in a suitable medium under conditions for the release of infectious viral vector particles.
12. An infectious viral vector particle comprising a nucleotide sequence which codes for a gamma-retroviral-Integrase, which is independent from bromodomain and extra-terminal (BET) family of proteins, operably linked to a nucleotide sequence which codes for a chromatin binding domain, or chromatin binding fragment thereof, and comprising a heterologous nucleotide sequence.
13. The infectious viral vector particle according to claim 11, wherein said heterologous nucleotide sequence is a transgene under the control of a promoter sequence, suitable for expression in a mammalian cell.
14. A host cell that has been transduced with an infectious viral vector particle according to claim 12 or 13.
15. In vitro use of the infectious retroviral vector particle according to claim 12 or 13 for the delivery of a heterologous sequence or transgene to a host cell.
16. The in vitro use according to claim 15 in gene therapy.

17. The use according to claim 15 or 16 wherein host cells are selected in which said heterologous sequence or transgene has been integrated in a "safe" genomic location.
- 5 18. The use according to any one of claims 15 to 17, wherein the host cell is a hematopoietic cell or a stem cell.
19. The use according to any one of claims 15 to 18, wherein said host cell is hematopoietic stem cell or a progenitor cell.
- 10 20. The infectious retroviral vector particle according to claim 12 or 13, wherein said infectious retroviral vector particle is present in a form suitable for being introduced into a target cell, for use as a medicament in gene therapy.
- 15 21. A method of gene therapy, which method comprises administering the infectious retroviral vector particle according to claim 12 or 13 to said subject.
22. The method according to claim 21, wherein a heterologous sequence or transgene is introduced in a subject in need thereof and wherein said subject is a mammal or human.
- 20

	390	400	
MLV	GGGPSSRLT	<u>WRVQ</u> RSQNPLKIRLTRE	-----AP [SEQ ID NO: 5]
AKR-MLV	PP---IKPS	<u>WRVQ</u> RSQNPLKIRLT <u>RG</u>	-----AP [SEQ ID NO: 6]
BaEV	TPGPTSSGT	<u>WRLR</u> SEDPLKIRLSR	-----T [SEQ ID NO: 7]
Cas-BR-E	SP---ARTA	<u>WKVQ</u> RSQNPLKIRLSRE	-----PS [SEQ ID NO: 8]
en-FeLV	TPGPETPKT	<u>WKLH</u> RSENPLKIRLSR	-----V [SEQ ID NO: 9]
GaLV	SAP---DES	<u>WELEK</u> TDHPLKLRIRRRRDE	-----SAK [SEQ ID NO:10]
KoRV	GAP---DES	<u>WELEK</u> TDHPLKLRVRRRRNE	-----STA [SEQ ID NO:11]
PERV	--P---DSG	<u>WRAEK</u> TENPLKLRLHRLVPYSNNNSPGQX	[SEQ ID NO:12]
		* ** ***** *	

Figure 1

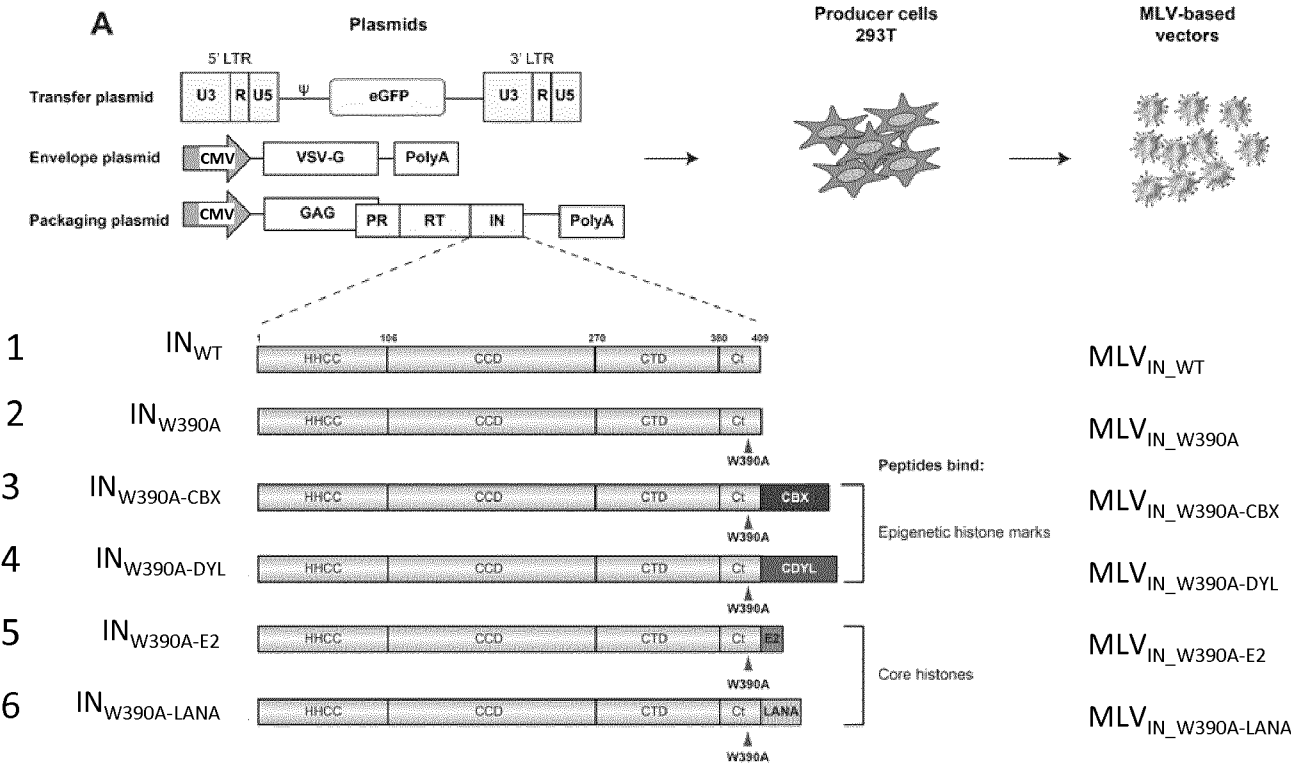
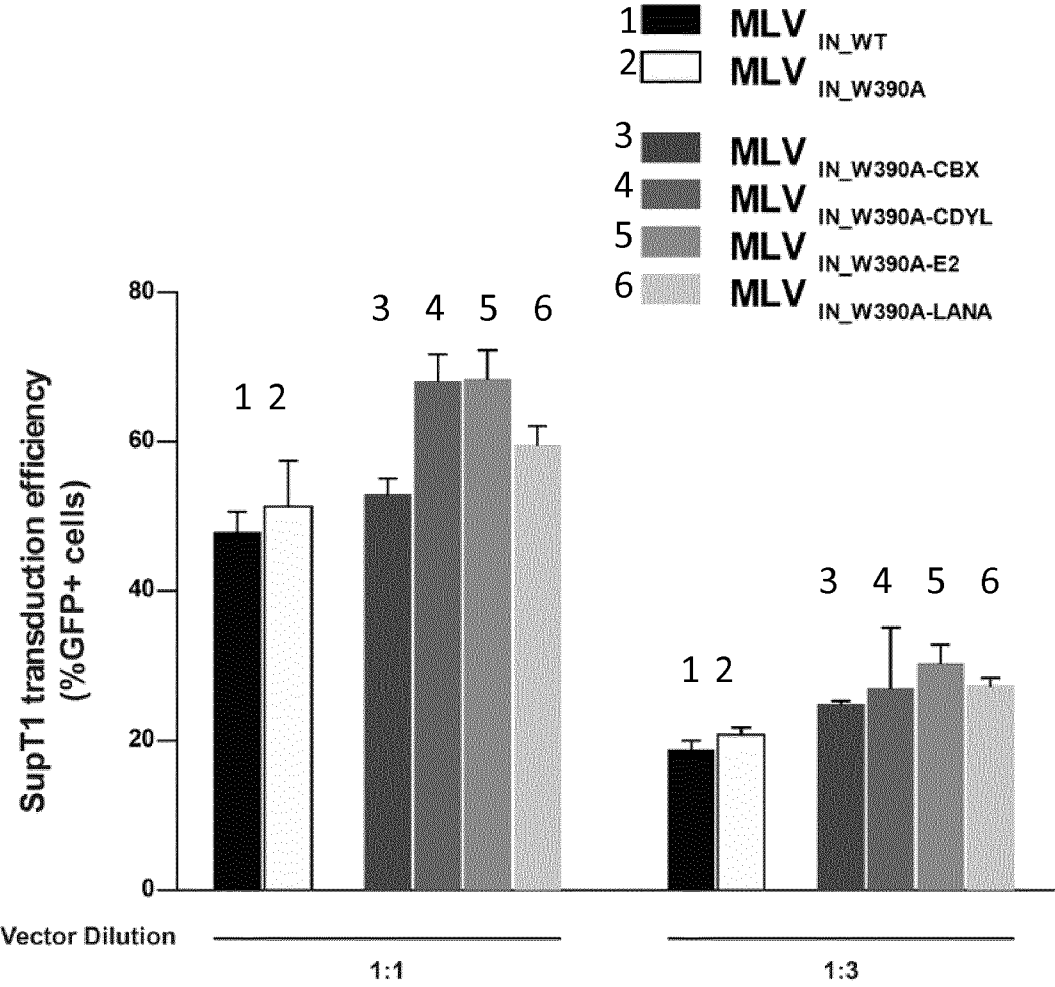
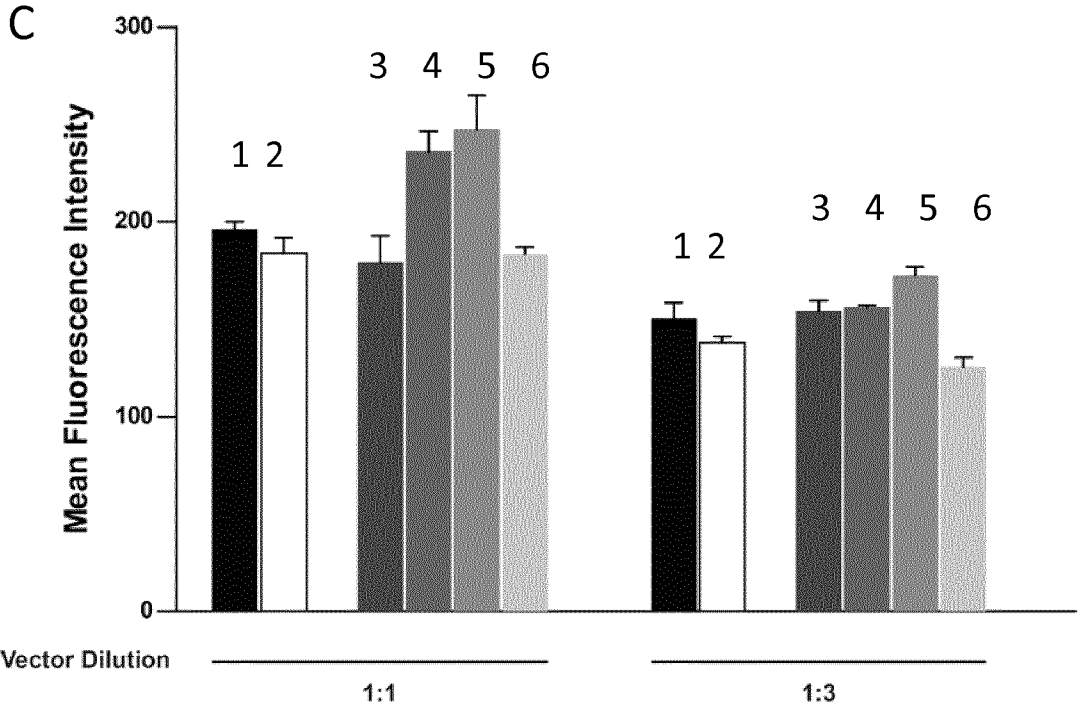


Figure2

B



C



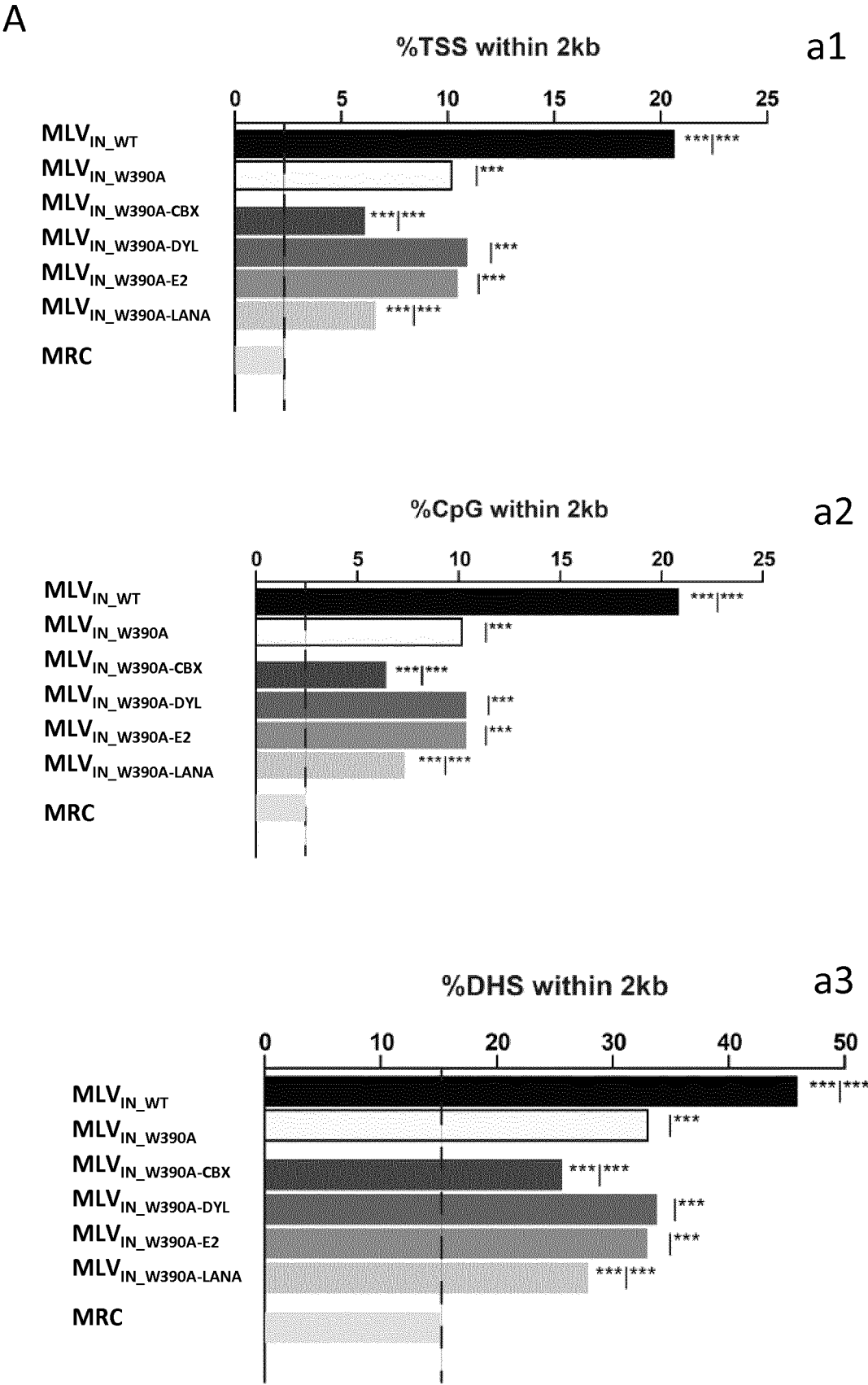


Figure 3

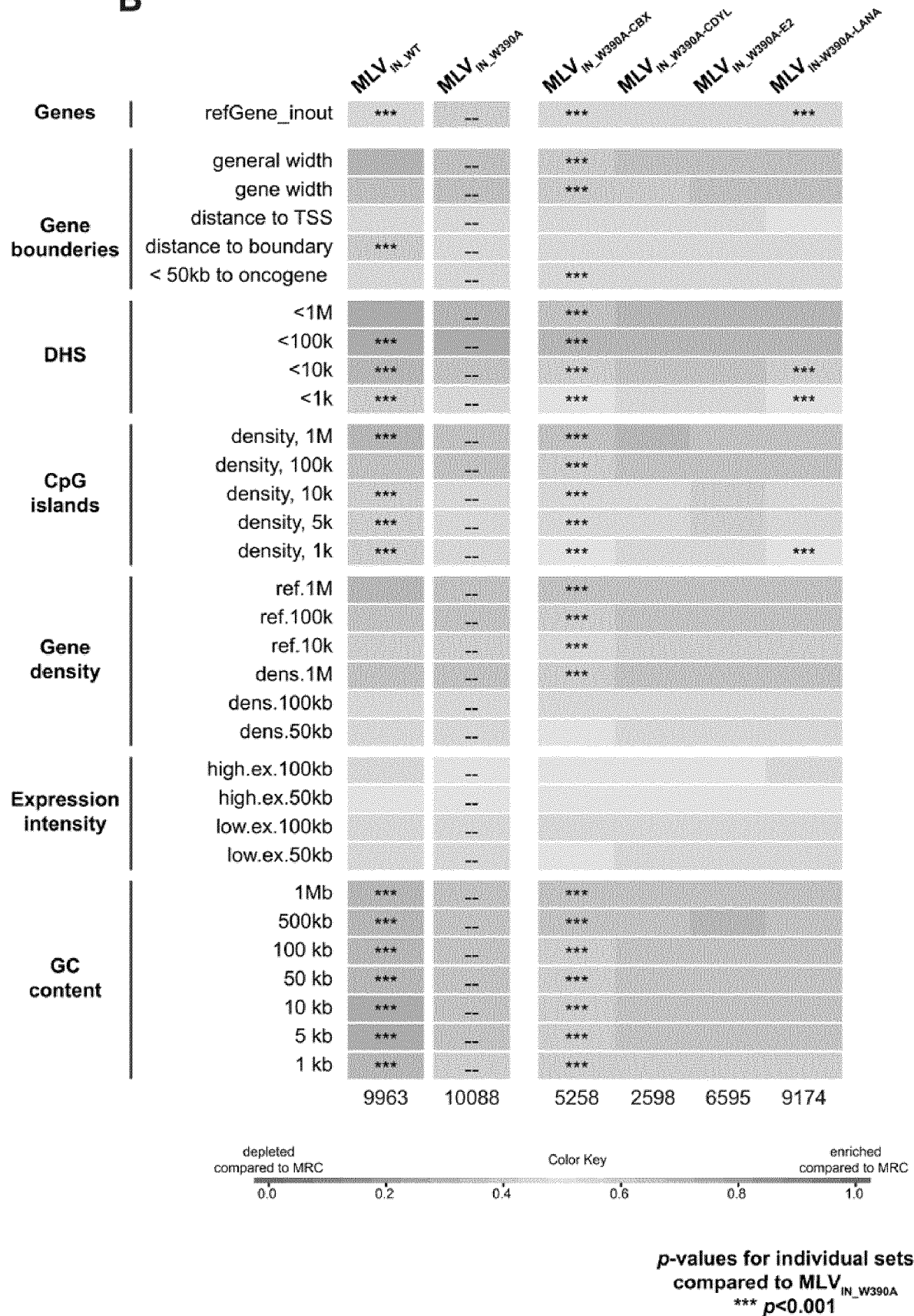
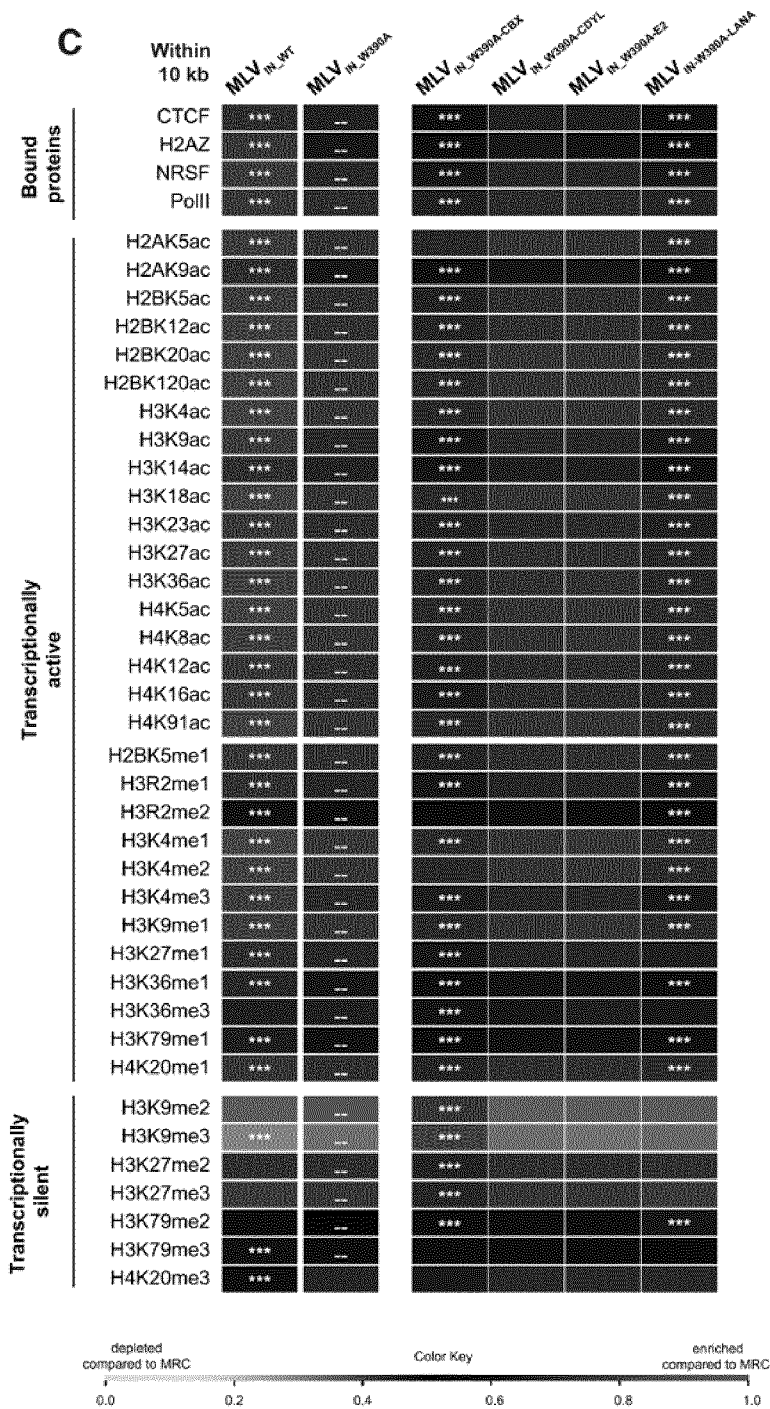
B

Figure 3 (continued)



p-values for individual sets compared to MLV_{IN_W390A}
*** *p*<0.001

Figure 3 (continued)

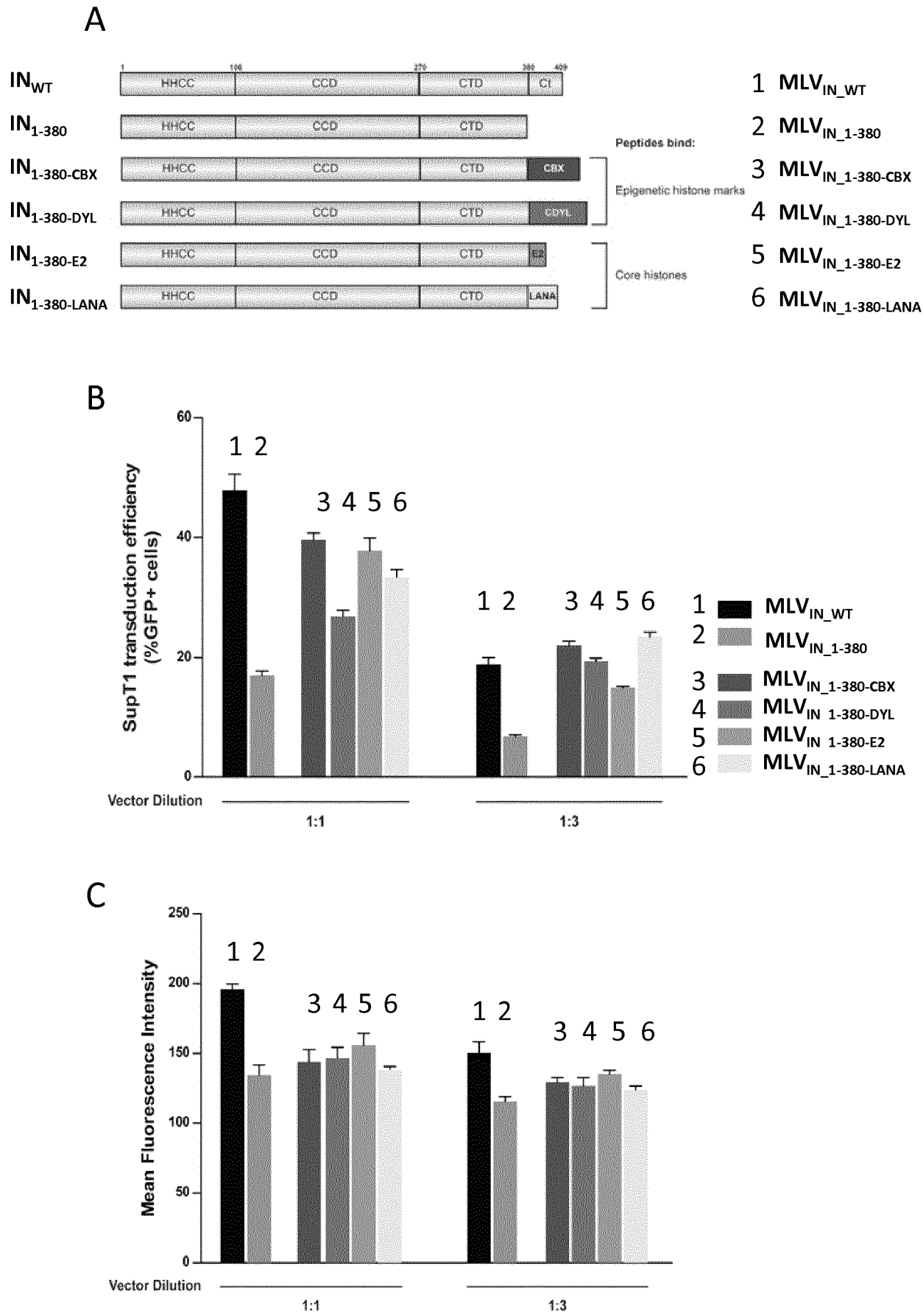


Figure 4

A

Vector	< 50kb TSS	< 300kb Allonco.	< 300kb miRNA	TU	In UCR	Fraction of safe integrations
MLV _{IN_WT}	42.0% ***	38.6% ***	23.4%	58.0% ***	6.5%	10.6% ***
MLV _{IN_W390A}	36.8%***	36.2%***	22.5%	61.4%***	6.1%	12.4%***
MLV _{IN_W390A-CBX}	34.1%*** ***	33.9%***	20.9%***	56.1% ***	5.0%***	15.5%*** ***
MLV _{IN_W390A-LANA}	37.8%***	38.3%	22.8%	58.4% ***	6.1%	14.1%*** ***
MRC	24.1%	23.1%	14.2%	46.5%	5.4%	28.4%

B

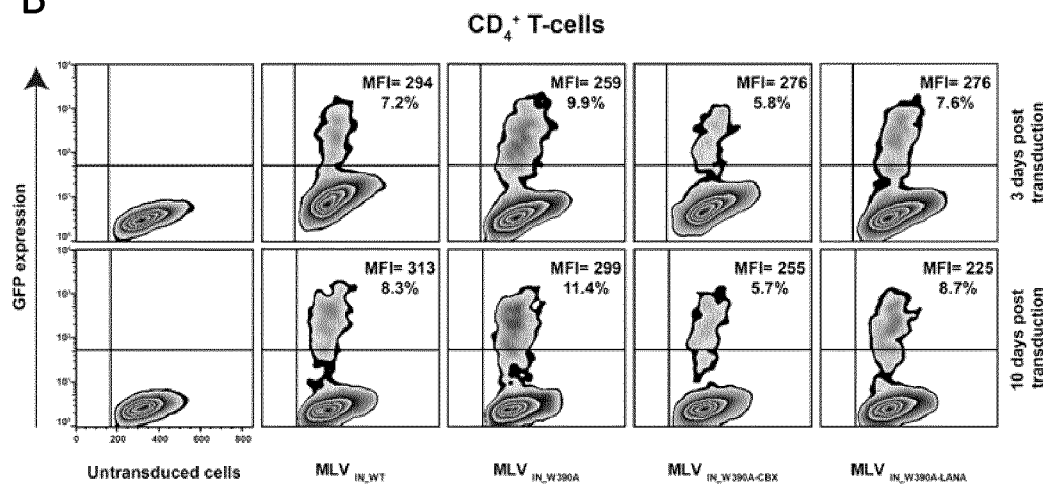


Figure 5

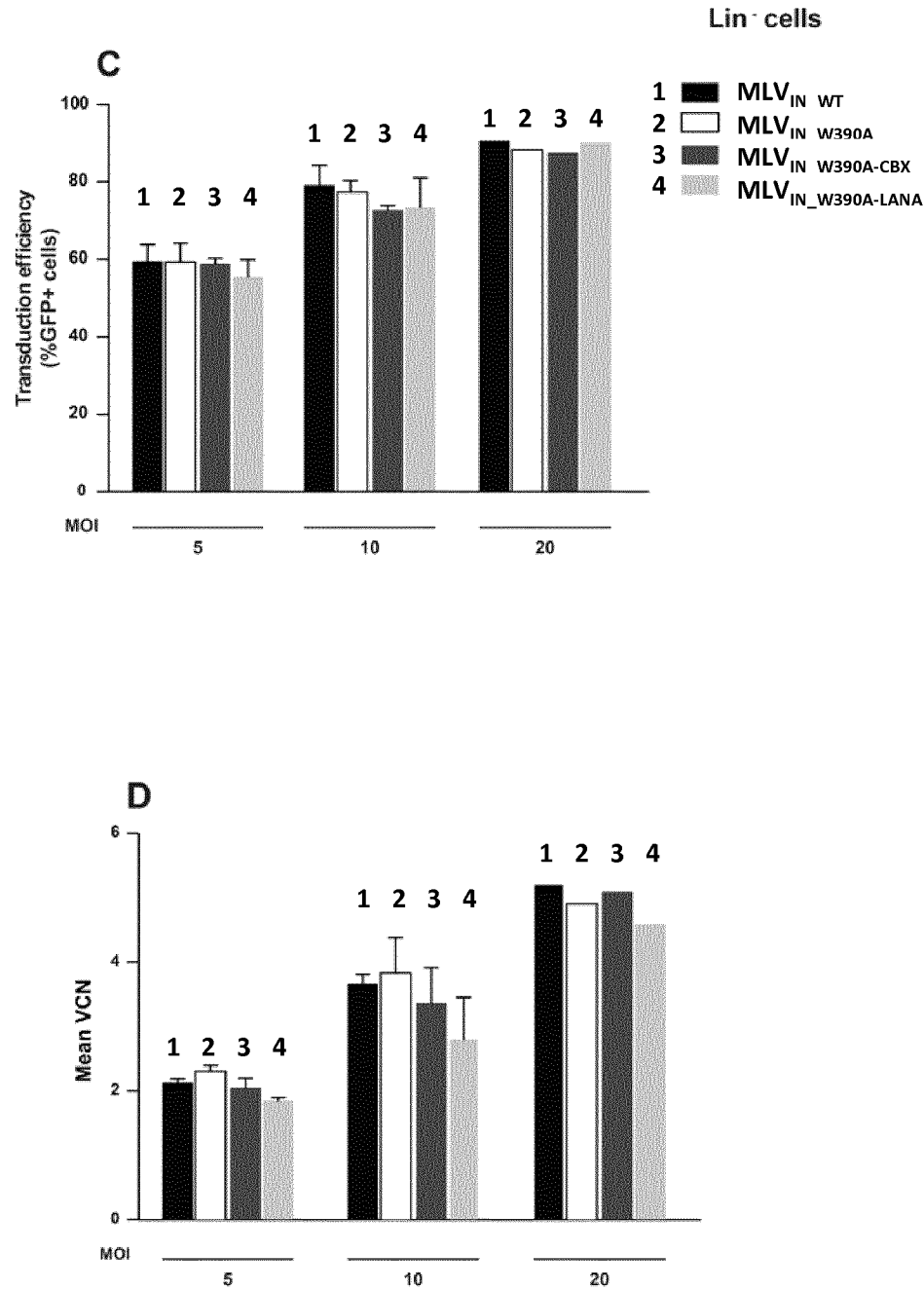
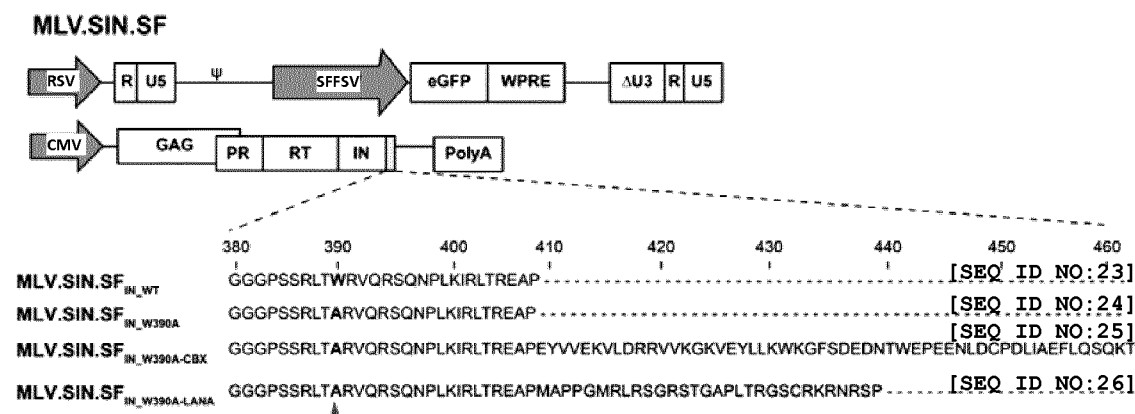


Figure 5 (continued)

A



B

	Vector	MOI	VCN	No.of pos.wells	Repl. freq.	Repl.freq./ copy no.
MLV.SIN.SF _{IN_WT}	5	5	2.1	0	0	0
	5	5	2.2	18	4.7x10 ⁻³	2.1x10 ⁻³
	10	10	3.5	11	2.6x10 ⁻³	7.5x10 ⁻⁴
	10	10	3.7	16	4.1x10 ⁻³	1.1x10 ⁻³
	10	10	3.8	17	4.4x10 ⁻³	1.1x10 ⁻³
	20	20	5.2	29	9.3x10 ⁻³	1.8x10 ⁻³
	Mean			91/288	4.2x10 ⁻³	1.1x10 ⁻³
MLV.SIN.SF _{IN_W390A}	5	5	2.2	11	2.6x10 ⁻³	1.2x10 ⁻³
	5	5	2.4	18	4.7x10 ⁻³	2.0x10 ⁻³
	10	10	3.8	8	1.8x10 ⁻³	4.8x10 ⁻⁴
	10	10	4.4	20	5.4x10 ⁻³	1.2x10 ⁻³
	10	10	3.3	19	5.0x10 ⁻³	1.5x10 ⁻³
	20	20	4.9	20	9.3x10 ⁻³	1.9x10 ⁻³
	Mean			96/288	4.8x10 ⁻³	1.4x10 ⁻³
MLV.SIN.SF _{IN_W390A-CBX}	5	5	1.9	0	0	0
	5	5	2.2	8	1.8x10 ⁻³	8.3x10 ⁻⁴
	10	10	3.0	1	2.1x10 ⁻⁴	7.0x10 ⁻⁵
	10	10	4.0	10	2.3x10 ⁻³	5.8x10 ⁻⁴
	10	10	3.1	23	6.5x10 ⁻³	2.1x10 ⁻³
	20	20	5.1	20	5.4x10 ⁻³	1.1x10 ⁻³
	Mean			62/288	2.7x10 ⁻³	7.8x10 ⁻⁴
MLV.SIN.SF _{IN_W390A-LANA}	5	5	1.9	8	1.8x10 ⁻³	9.6x10 ⁻⁴
	5	5	1.8	19	5.0x10 ⁻³	2.8x10 ⁻³
	10	10	3.5	26	7.8x10 ⁻³	2.2x10 ⁻³
	10	10	2.7	17	4.4x10 ⁻³	1.6x10 ⁻³
	10	10	2.2	14	3.5x10 ⁻³	1.6x10 ⁻³
	20	20	4.6	9	2.1x10 ⁻³	4.5x10 ⁻⁴
	Mean			93/288	4.1x10 ⁻³	1.6x10 ⁻³

Figure 6

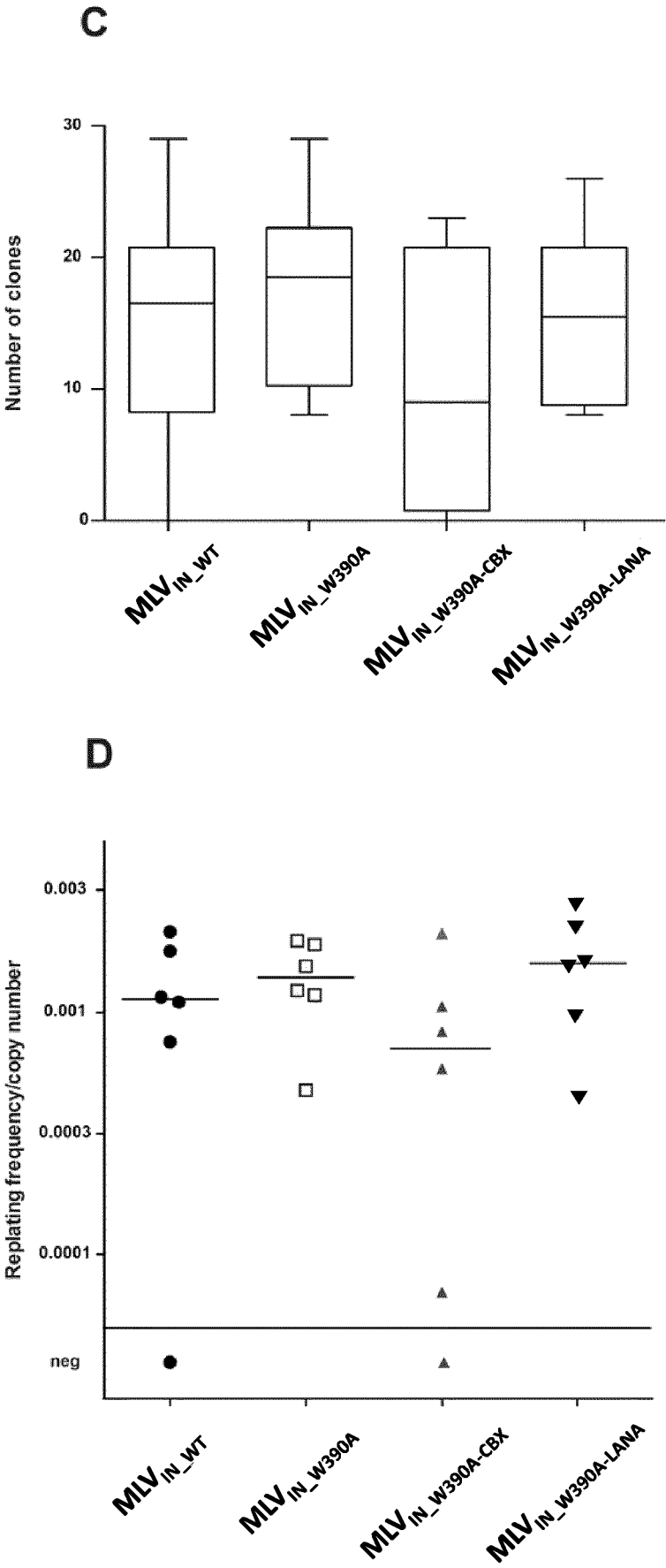


Figure 6 (continued)

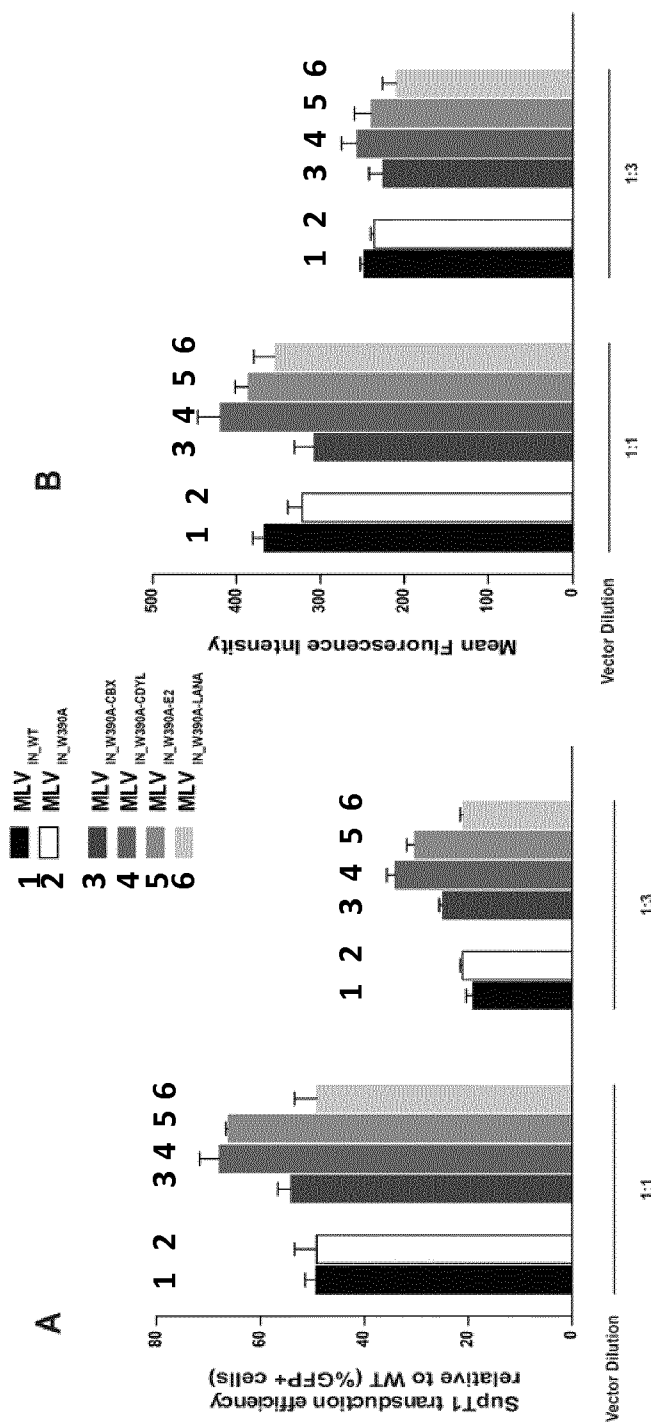


Figure 7

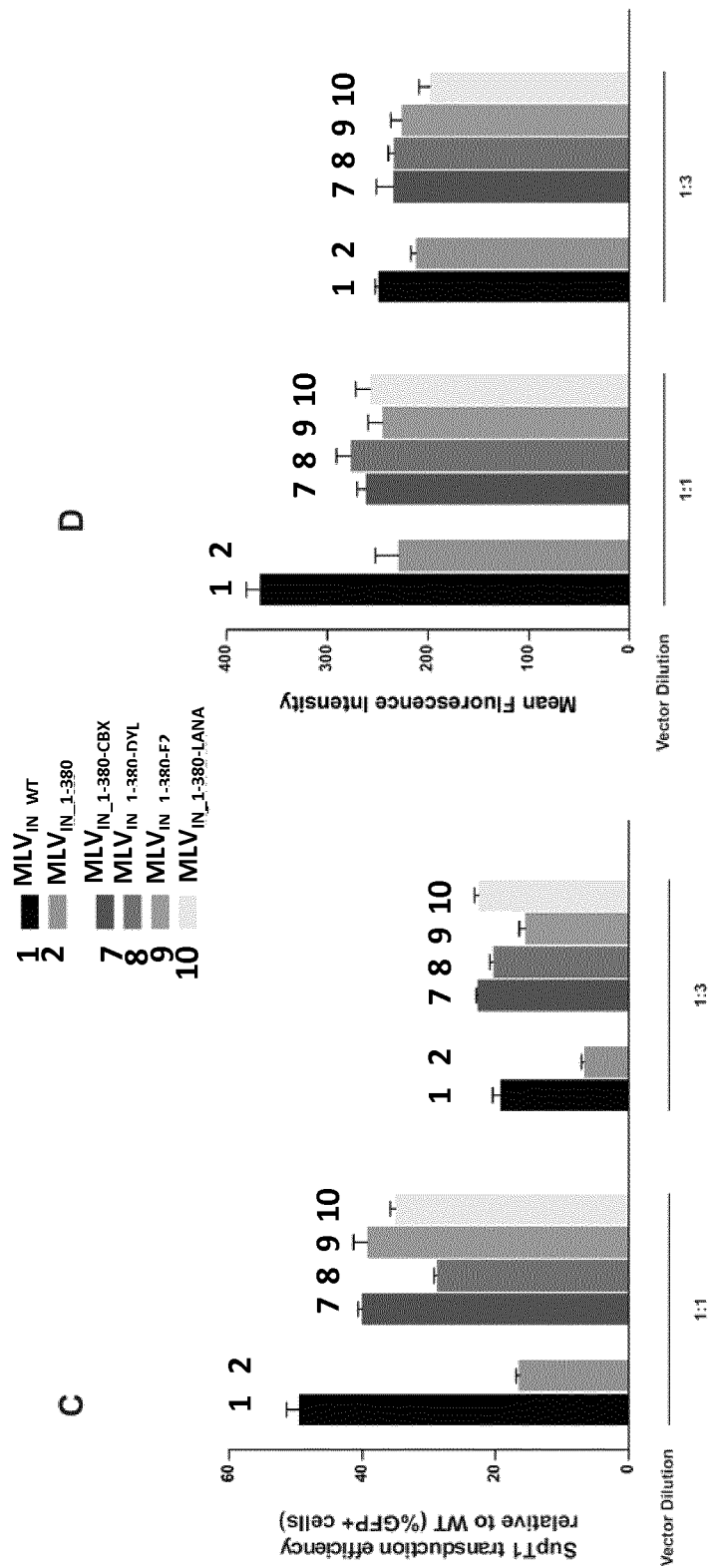


Figure 7 (continued)

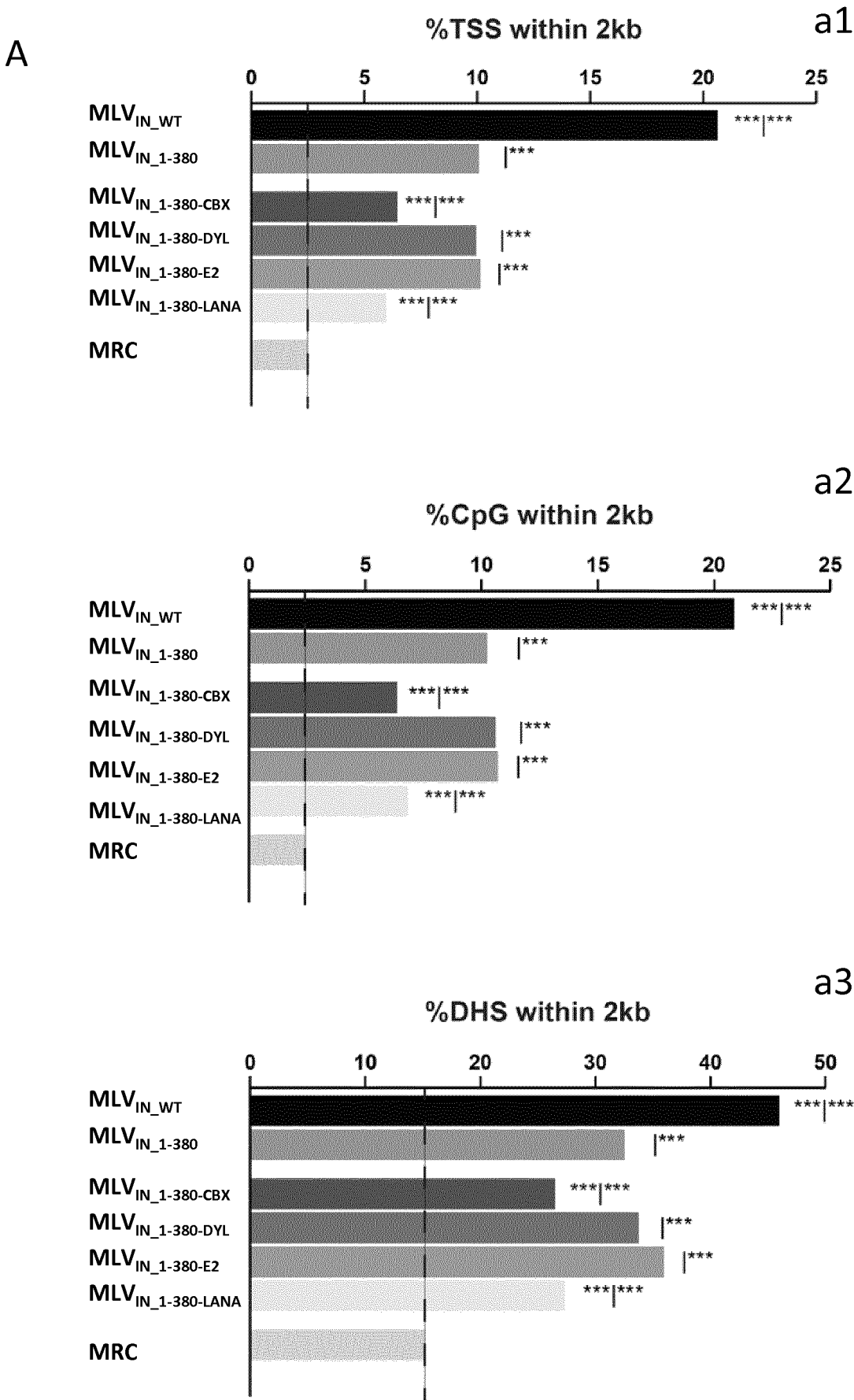
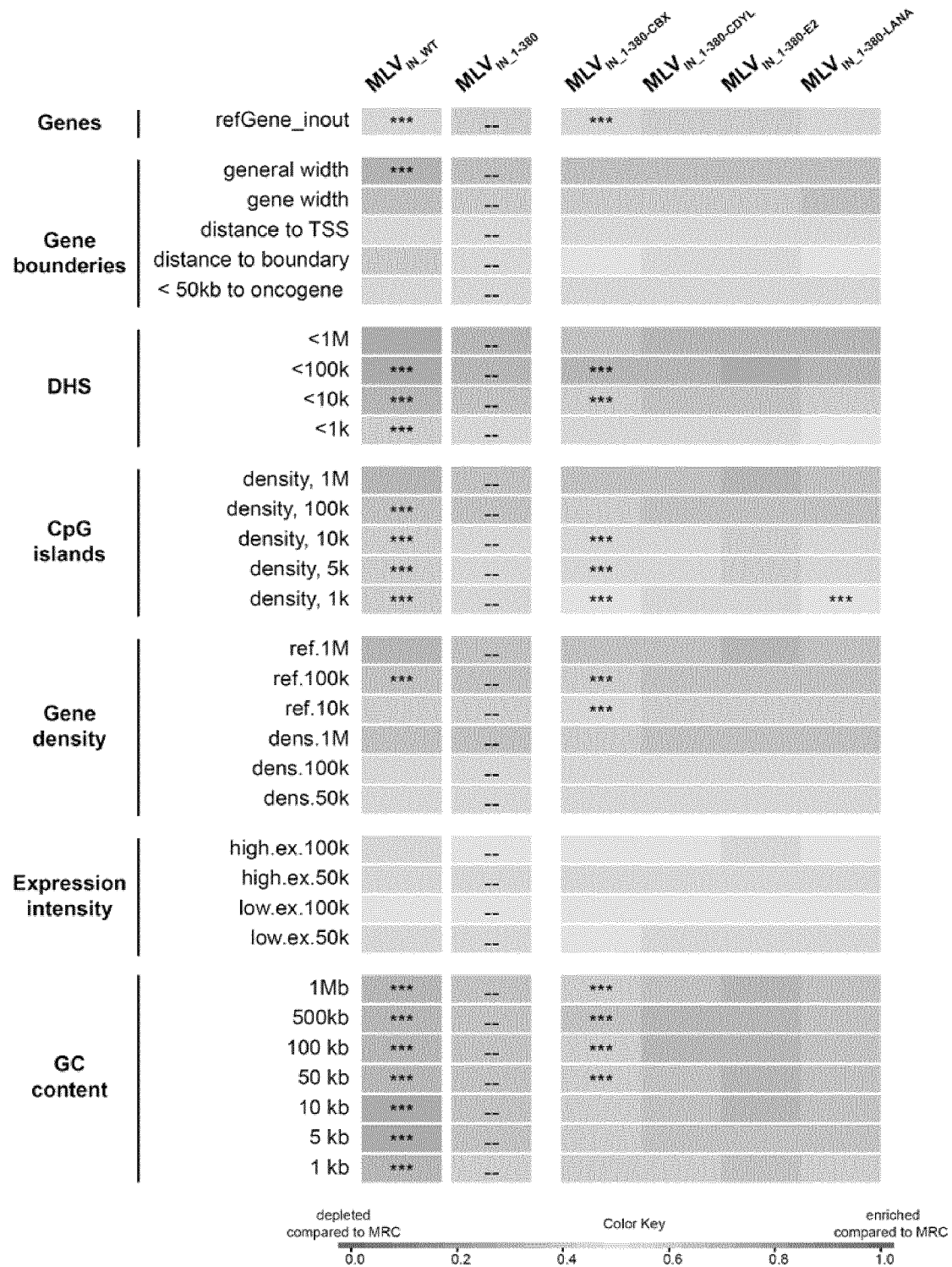


Figure 8



p-values for individual sets
compared to MLV_{IN_1-380}
*** *p*<0.001

Figure 8 (continued)

C

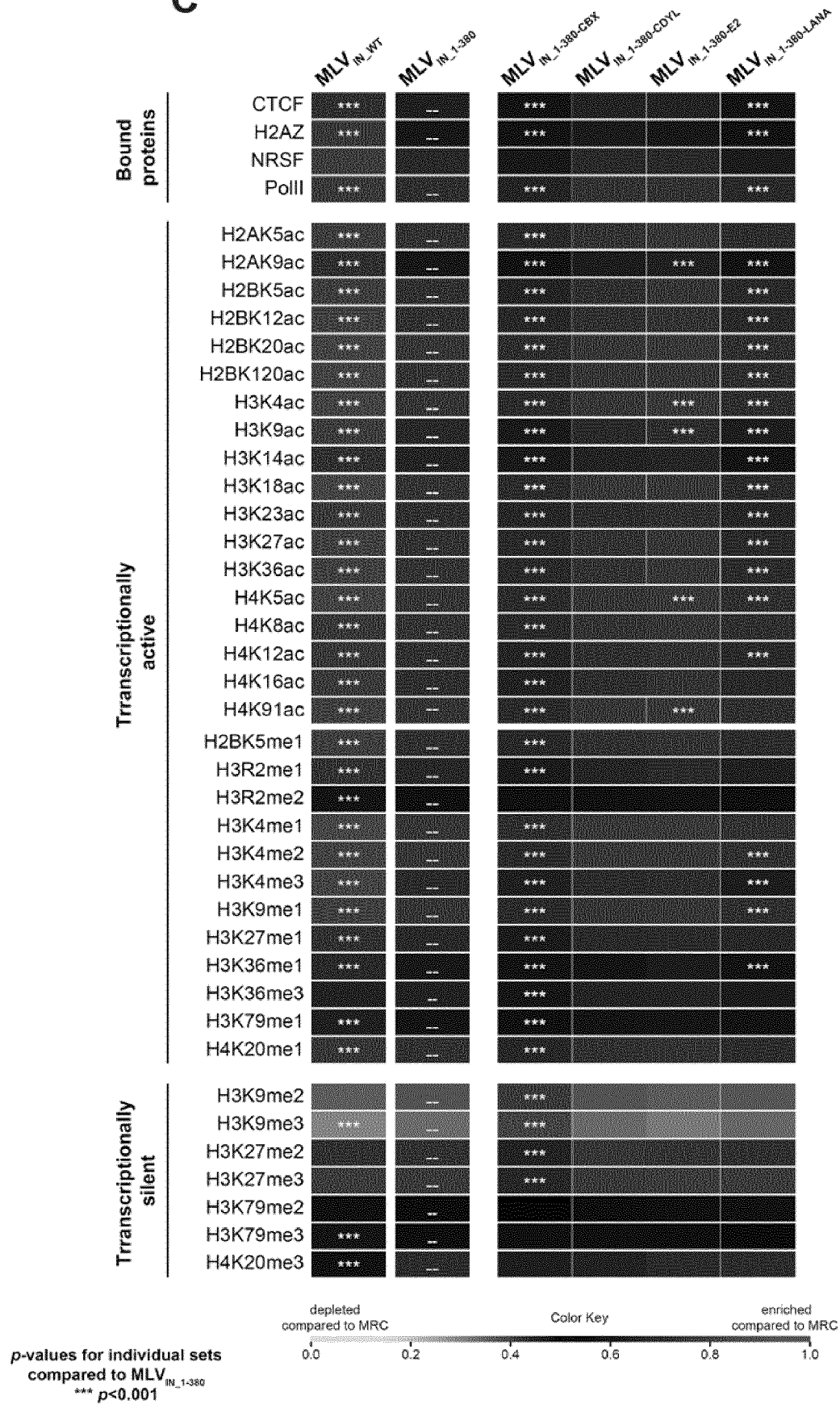


Figure 8 (continued)

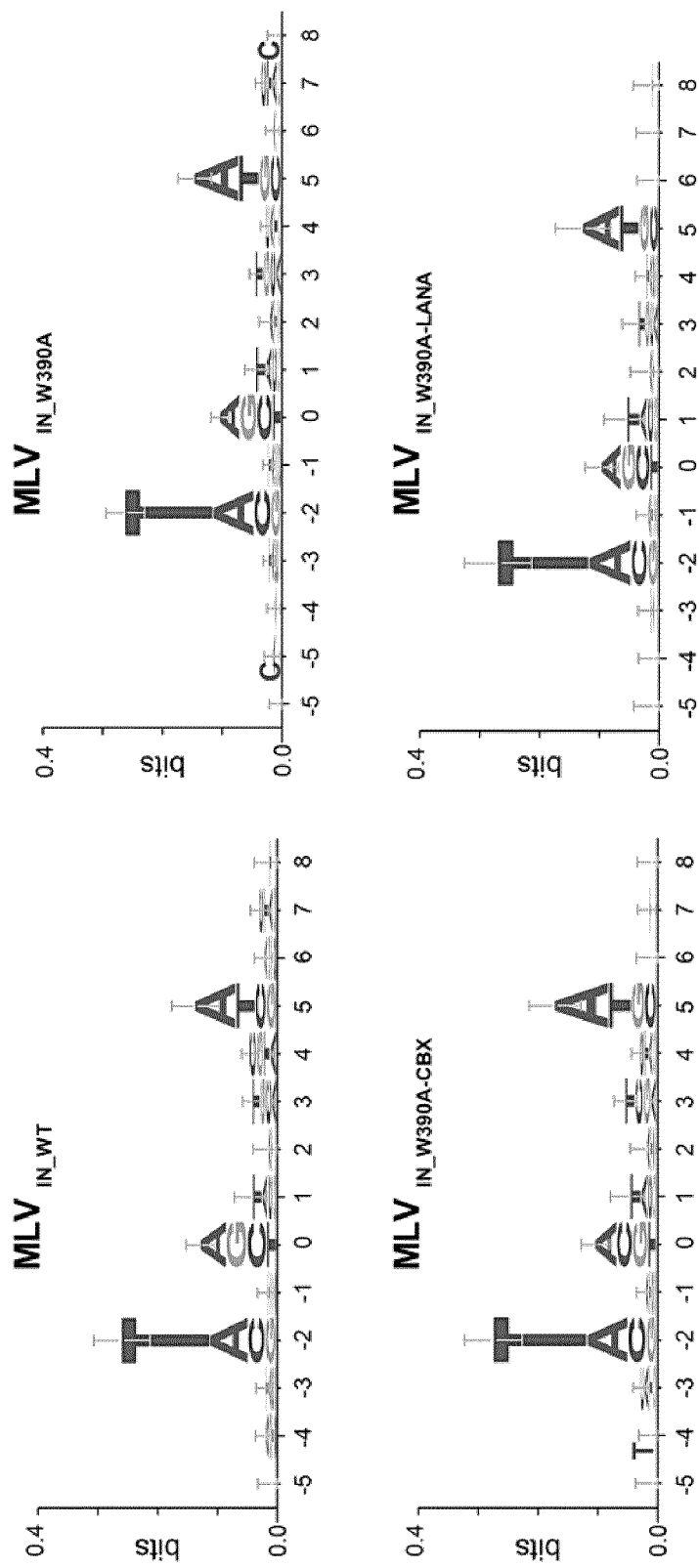


Figure 9

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2017/076615

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N15/86 C07K14/005
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C07K A61K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, BIOSIS, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	SARA EL ASHKAR ET AL: "BET-independent MLV-based Vectors Target Away From Promoters and Regulatory Elements", MOLECULAR THERAPY - NUCLEIC ACIDS, vol. 3, 1 January 2014 (2014-01-01), pages e179-2531, XP055450922, GB ISSN: 2162-2531, DOI: 10.1038/mtna.2014.33 cited in the application abstract page 6, column 2 - page 8, column 1; figures 1-4 ----- -/--	1-22



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

15 February 2018

Date of mailing of the international search report

02/03/2018

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Mossier, Birgit

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2017/076615

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ZEGER DEBYSER ET AL: "Host factors for retroviral integration site selection", TRENDS IN BIOCHEMICAL SCIENCES, vol. 40, no. 2, 1 February 2015 (2015-02-01), pages 108-116, XP055450928, AMSTERDAM, NL ISSN: 0968-0004, DOI: 10.1016/j.tibs.2014.12.001 abstract; figure 4 -----	1-22
A	AXEL SCHAMBACH: "Toward a Safer Integration Profile of MLV-based Retroviral Vectors", MOLECULAR THERAPY, vol. 22, no. 8, 1 August 2014 (2014-08-01), pages 1405-1406, XP055450925, & 15TH ANNUAL MEETING OF THE AMERICAN-SOCIETY-OF-GENE-AND-CELL-THERAPY (ASGCT); PHILADELPHIA, PA, USA; MAY 16 -19, 2012 ISSN: 1525-0016, DOI: 10.1038/mt.2014.124 abstract; figure 1; table 1 -----	1-22
A	RIK GIJSBERS ET AL: "LEDGF Hybrids Efficiently Retarget Lentiviral Integration Into Heterochromatin", MOLECULAR THERAPY, vol. 18, no. 3, 1 March 2010 (2010-03-01), pages 552-560, XP055285803, & 15TH ANNUAL MEETING OF THE AMERICAN-SOCIETY-OF-GENE-AND-CELL-THERAPY (ASGCT); PHILADELPHIA, PA, USA; MAY 16 -19, 2012 ISSN: 1525-0016, DOI: 10.1038/mt.2010.36 abstract page 558, column 1, paragraph 2 - column 2, paragraph 3; figure 1; table 1 -----	1-22
A	JAN DE RIJCK ET AL: "The BET Family of Proteins Targets Moloney Murine Leukemia Virus Integration near Transcription Start Sites", CELL REPORTS, vol. 5, no. 4, 1 November 2013 (2013-11-01), pages 886-894, XP055451029, US ISSN: 2211-1247, DOI: 10.1016/j.celrep.2013.09.040 abstract page 891, column 2, paragraph 2 - page 892; figures 1-4 ----- -/--	1-22

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2017/076615

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MEEHAN A M ET AL: "Chromatin tethering and retroviral integration: Recent discoveries and parallels with DNA viruses", BIOCHIMICA ET BIOPHYSICA ACTA. GENE REGULATORY MECHANISMS, ELSEVIER, AMSTERDAM, NL, vol. 1799, no. 3-4, 1 March 2010 (2010-03-01), pages 182-191, XP026965615, ISSN: 1874-9399 [retrieved on 2009-10-15] abstract; figures 1-3; table 1 -----	1-22
A	WO 01/25484 A2 (UNIV MICHIGAN [US]; ROBERTSON ERLE S [US]; COTTER MURRAY A [US]) 12 April 2001 (2001-04-12) abstract page 12, line 11 - page 13, line 31 -----	1-22
Y,P	WO 2016/176152 A2 (UNIV WASHINGTON STATE [US]) 3 November 2016 (2016-11-03) abstract page 8, line 9 - line 21 page 9, line 1 - line 9; claims 1,6,7,62; figures 1,2; examples 1,3,4 -----	1-22
X,P	SARA EL ASHKAR ET AL: "Engineering Next-Generation BET-Independent MLV Vectors for Safer Gene Therapy", MOLECULAR THERAPY - NUCLEIC ACIDS, vol. 7, 1 June 2017 (2017-06-01), pages 231-245, XP055450923, GB ISSN: 2162-2531, DOI: 10.1016/j.omtn.2017.04.002 the whole document -----	1-22

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2017/076615

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0125484	A2	12-04-2001	
		AU 7623700 A	10-05-2001
		US 6482587 B1	19-11-2002
		US 2003133948 A1	17-07-2003
		US 2004248081 A1	09-12-2004
		WO 0125484 A2	12-04-2001

WO 2016176152	A2	03-11-2016	NONE
