



(51) International Patent Classification:

C12N 7/00 (2006.01) A61K 48/00 (2006.01)
C12N 15/09 (2006.01) A01K 67/027 (2006.01)
C12N 15/864 (2006.01) C12N 15/86 (2006.01)
C12N 5/10 (2006.01)

(21) International Application Number:

PCT/GB2013/051960

(22) International Filing Date:

23 July 2013 (23.07.2013)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

1213117.3 24 July 2012 (24.07.2012) GB

(71) Applicant: UCL BUSINESS PLC [GB/GB]; The Network Building, 97 Tottenham Court Road, London, W1T 4TP (GB).

(72) Inventor: NATHWANI, Amit; UCL Cancer Institute, Paul O'Gorman Building, 72 Huntley Street, London, WC1E 6BT (GB).

(74) Agents: HUTTER, Anton et al.; Venner Shipley LLP, 200 Aldersgate, London, EC1A 4HD (GB).

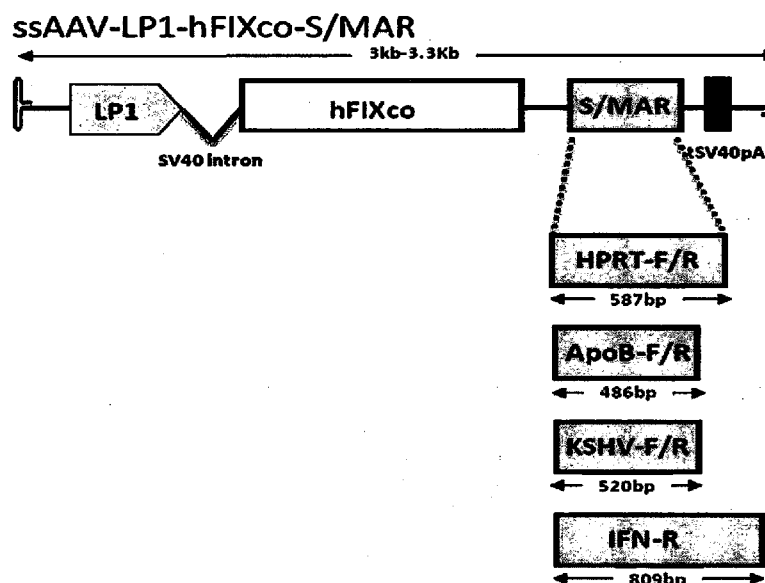
(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, 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, 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.

(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, 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).

Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: TRANSGENE EXPRESSION

Figure 9A

(57) Abstract: The invention relates to transgene expression in host cells using recombinant expression cassettes and expression vectors harbouring such cassettes. A Parvoviridae expression cassette comprises a transgene and a matrix attachment region (MAR) sequence or a fragment thereof from a hypoxanthine phosphoribosyltransferase (HPRT) gene. The invention extends to the therapeutic uses of such cassettes and vectors for expressing transgenes for treating a range of different medical conditions.

- 1 -

TRANSGENE EXPRESSION

The present invention relates to transgene expression in host cells using recombinant expression cassettes and expression vectors harbouring such cassettes. The invention extends to the therapeutic uses of such cassettes and vectors for
5 expressing transgenes for treating a range of different medical conditions.

The clinical application of gene therapy with adeno-associated virus (AAV) vectors shows great promise. Recent studies have established safe and successful treatment of patients with Leber's congenital amaurosis and haemophilia B following a single
10 administration of recombinant adeno-associated virus (rAAV) vectors. However, the need for high titres of vector for therapeutic transgene expression and its associated toxicity suggests that further optimization of this technology is essential prior to its broader application in the clinic.

The inventors therefore set out to develop a novel, more potent AAV expression cassette, and investigated the inclusion of a matrix attachment region (MAR). MARs are AT-rich sequences that attach to the nuclear matrix thereby forming anchor points that serve to coordinate nuclear activities such as initiation of DNA replication or regulation of transcription by bringing genes in close proximity to key nuclear
20 regulatory elements. Furthermore, MARs are thought to maintain genes in an "unwound" open chromatin state, which is an absolute requirement for their role as origins of replication, because it allows the initiation of strand synthesis, a function that is supported by the presence of topoisomerase II binding sites within MAR sequences. The "unwound" euchromatinised DNA state is also important for
25 enhancing gene expression.

The inventors have previously demonstrated in unpublished work that, following gene transfer, a proportion of the AAV genome becomes heterochromatinised and is then maintained in a transcriptionally inactive state within the target cell. They
30 therefore considered if the inclusion of MARs within an AAV vector would prevent proviral heterochromatinisation from occurring, thereby increasing transgene expression and the potency of the AAV vector system. They constructed and tested the role of the four different matrix associated regions (MARs) in the context of rAAV-mediated gene transfer in murine models. These MARs included domains
35 from the interferon β (IFN- β) gene, the hypoxanthine phosphoribosyltransferase (HPRT) gene, the apolipoprotein B (ApoB) gene, and the Kaposi sarcoma herpes

- 2 -

virus (KSHV) genome. These four MARs were selected on the basis of size (<1000bp) and were then graded according to their *in silico* theoretical approximation of the “energy (GX)” of unwinding.

5 The IFN- β MAR has previously been shown to enhance expression of transgenes in the context of plasmid or retroviral vectors (1, 2). Furthermore, the ApoB MAR has been extensively studied for its ability to insulate transgene expression from chromosomal position effects (3, 4). In contrast, the interest in the HPRT and the KSHV MARs, to date, has merely been in their ability to serve as the origin of
10 replication and support of autonomous replication respectively (5-7). The *in silico* studies suggested that the IFN- β and ApoB MARs were likely to be most efficient at enhancing gene expression with GX values of 1.55 and 0.8, respectively. The higher GX (2.5) value for the HPRT MAR suggested that it was least likely to maintain the rAAV genome in an unwound state and, therefore, was most unlikely to enhance
15 transgene expression. Indeed, the HPRT MAR has previously been shown to exhibit all the properties of MARs except the ability to enhance transgene expression (5-7).

However, as described in the Examples, a comparison of the four MARs in the context of the LP1-hFIXco, a Factor IX (i.e. FIX) expression cassette that is being
20 tested in the clinic, revealed highly surprising data for the vector containing the HPRT MAR. Indeed, inclusion of the HPRT MAR in the vector resulted in a significant and totally unexpected boost in the expression of Factor IX.

Thus, according to a first aspect of the invention, there is provided a Parvoviridae
25 expression cassette comprising a transgene and a matrix attachment region (MAR) sequence or a fragment thereof from a hypoxanthine phosphoribosyltransferase (HPRT) gene.

The inventors were surprised to observe that the HPRT matrix attachment region or
30 fragment thereof present in the expression cassette unexpectedly enhanced transgene expression in a host cell. As shown in Figure 1, the inclusion, in an expression cassette, of the HPRT MAR downstream of the transgene, FIX, (in a recombinant vector - rAAV-LP1-hFIXco) surprisingly resulted in the highest level of transgene expression. The expression was at least 10- fold higher than the level of
35 transgene expression achieved with an identical construct containing the IFN- β MAR, and over 100-fold higher than that observed with the same construct but without any MAR sequences. This was totally unexpected. These data suggest that

- 3 -

different classes of MARs have divergent biological properties in relation to the upregulation of transgene expression. The published literature suggests that the HPRT MAR does not have enhancer activity. However, the results described herein suggest that the HPRT MAR element or a fragment thereof is a surprisingly potent enhancer of transgene expression. This has substantial implications for gene therapy of many different conditions, and will undoubtedly impact the design of future generations of Parvoviridae expression cassettes for use in the clinic. Moreover, as shown in Figure 2, substantially lower doses of vector containing HPRT MAR are required to achieve therapeutic levels of transgene expression, and this has important safety implications as well as easing the burden on vector production.

It will be appreciated that HPRT is a transferase, which catalyzes the conversion of hypoxanthine to inosine monophosphate, and guanine to guanosine monophosphate via transfer of the 5-phosphoribosyl group from 5-phosphoribosyl 1-pyrophosphate. This enzyme plays a central role in the generation of purine nucleotides through the purine salvage pathway. Human HPRT gene is 40.5kb long and the HPRT protein has 218 amino acid residues. The accession number of human HPRT is NG_012329.1.

The HPRT sequence appears to function optimally in the forward orientation in the AAV-LP1-hFIXco HPRT vector as described herein. However, it is recognised that the context of the MARs, such as HPRT, within the expression cassette (or transgene itself) will have an impact on which orientation functions best. The MAR elements, such as HPRT, can improve transgene expression when cloned into the 3' position as shown in the examples, but it is highly likely that MARs can exert enhancer effect when cloned in the 5' region of promoters, introns and within intragenic regions. Therefore, the MAR sequence or fragment thereof may be provided anywhere within the expression cassette, including at the 5' end of the promoter and/or within an intron of the transgene. Preferably, however, the MAR sequence or fragment thereof is provided 3' of the transgene. The MAR sequence or fragment thereof may be operatively linked to the transgene, i.e. the MAR sequence is capable of enhancing transgene expression.

The MAR sequence or fragment thereof is preferably a Scaffold/Matrix Attachment Region (S/MAR) element. The MAR region of the human HPRT gene is an AT-rich sequence and may be substantially set out as SEQ ID No:1. The HPRT MAR maps to the first intron of the wild-type HPRT gene between +3858 to +4430bp.

- 4 -

5 GTCAAGAATGCTGATTTTAAATTCTTCTAACAGCTTTATCCCTCAGAAGGGAAGGCAAGCA
 AGTTATATATGTAGTTTATTTGTAAGACTGATATGAAATTGGAAGATGAATCTACTATTA
 GCTTTAATTATTTTACATTTAGGAATATTGCATCAGTAACTCATAATTTTGGTTTTCTG
 10 TTATCCTGAGTTAACACAAATTATCCAAGGAGATGGCGGATCATCTGCTTTGAGGTGTTT
 TTTTTTGAGAATTTTAATGTATCTGAATATAAAAGGTAAAAATATGCCAACTAGCAATTT
 CTGCCCATTCCAGAAGTTTGGAAATATTACTCATTACTAGGAATTAAATAAAATATGGTT
 TATCTATTGTTATACCTCTTTTAATTCACATAGCTCATTTTTATCTTTTATTTTGTGTTG
 15 TTTTTTTGAGATGGAGTCTTGCTCTGTCACCAGGCAGGAGTGCAGTGATGCAATCTCGG
 CTCACTCTAGCCACCGACTCCCTGGTTCAAGCGATTCTCCTGCCTGAGCCTTCTGAGTAG
 CTGGGATTACAGGCAGGCACCAACACGCCACG

[SEQ ID No.1]

15 Thus, the MAR sequence or fragment thereof provided in the expression cassette of
 the first aspect may comprise a nucleic acid sequence substantially as set out in SEQ
 ID No.1, or a functional fragment or variant thereof. It will be appreciated that
 suitable fragments or variants of SEQ ID No.1 may include truncations of this
 sequence while still being substantially AT-rich. Preferably, the fragment or variant
 comprises at least 20, 25, 30, 35, 40, 45, 50 or 55 ATs.

20

As shown in Figures 5 and 6, the inventors have created fragments of the MAR
 sequence of SEQ ID No.1, and determined which sequences mediated enhancer
 activity. Thus, a suitable fragment of the MAR sequence may comprise nucleotides
 205 to 450 of SEQ ID No.1, which is referred to as Fragment 2. However,
 25 surprisingly, the inventors have determined that the enhancer activity of the MAR
 sequence is predominantly mediated by nucleotides 320 to 450 of SEQ ID No.1,
 which is referred to in the Figures as Fragment 2b. Thus, it is preferred that a
 fragment of the MAR sequence comprises nucleotides 205 to 450 of SEQ ID No.1.

30

The expression cassette may comprise DNA from a member of the Dependovirus
 genus. It is preferred that the expression cassette comprises DNA from an adeno-
 associated virus (AAV) vector. Thus, preferably the expression cassette is provided in
 the form of a recombinant expression vector.

35

Thus, in a second aspect, there is provided a recombinant vector comprising an
 expression cassette according to the first aspect.

40

The recombinant vector may be a plasmid, cosmid or phage and/or be a viral vector.
 The vector may be single- or double-stranded. Such vectors are useful for
 transforming host cells, and for replicating the transgene therein. The vector may

- 5 -

comprise DNA from, or is a, member of the Dependovirus genus. It is preferred that the vector comprises DNA from, or is, an adeno-associated virus (AAV) vector.

The skilled technician will appreciate that the transgene containing HPRT MAR or S/MAR may be combined with many types of backbone vector for expression purposes. Examples of suitable Parvoviridae backbone vectors include: single-stranded AAV vectors, self-complementary AAV vectors pseudotyped with serotype 2, 5, 8 or 9 capsid or any other capsid protein arising from wild type AAV or engineered forms of AAV capsid.

In one embodiment, the nucleic acid sequence of the internal terminal repeats (ITR) in the expression cassette is substantially set out as SEQ ID No:2. The sequence shows the two ITRs and identifies the area where expression cassettes may be inserted.

```

ggccactccc tctctgcgcg ctgcctcgtt cactgaggcc gggcgaccaa aggtcgcccc
>>.....S'ITR.....>

acgcccgggc ttgccccggg cggcctcagt gagcgagcga gcgcgcagag agggagtggc
>.....S'ITR.....>

caactccatc actagggggt cctggagggg tggagtcgtg ac-Expression cassette-
>.....S'ITR... ..>

aggaaccctt agtgatggag ttggccactc cctctctgcg cgtctcgtcg ctactgagg
>>.....S'ITR.....>

ccgggcgacc aaaggtcgcc cgacgcccgg gctttgcccg ggcggcctca gtgagcgagc
>.....S'ITR.....>

gagcgcgcgc
>...ITR.>>

```

[SEQ ID No.2]

Thus, the expression cassette or vector may comprise a nucleic acid sequence substantially as set out in SEQ ID No.2, or a functional fragment or variant thereof.

The cassette or vector may further comprise a variety of other functional elements including a suitable promoter for initiating transgene expression upon introduction of the vector in a host cell. For instance, the vector is preferably capable of autonomously replicating in the nucleus of the host cell. In this case, elements which induce or regulate DNA replication may be required in the recombinant vector. Alternatively, the recombinant vector may be designed such that it integrates into the

- 6 -

genome of a host cell. In this case, DNA sequences which favour targeted integration (e.g. by homologous recombination) are envisaged. Suitable promoters may include the SV40 promoter, CMV, EF1a, PGK, viral long terminal repeats, as well as inducible promoters, such as the Tetracycline inducible system, as examples. The
 5 cassette or vector may also comprise a terminator, such as the Beta globin, SV40 polyadenylation sequences or synthetic polyadenylation sequences.

The vector may also comprise DNA coding for a gene that may be used as a selectable marker in the cloning process, i.e. to enable selection of cells that have been
 10 transfected or transformed, and to enable the selection of cells harbouring vectors incorporating heterologous DNA. For example, ampicillin, neomycin, puromycin or chloramphenicol resistance is envisaged. Alternatively, the selectable marker gene may be in a different vector to be used simultaneously with the vector containing the transgene. The cassette or vector may also comprise DNA involved with regulating
 15 expression of the transgene, or for targeting the expressed polypeptide to a certain part of the host cell.

Purified vector may be inserted directly into a host cell by suitable means, e.g. direct endocytotic uptake. The vector may be introduced directly into cells of a host subject
 20 (e.g. a eukaryotic or prokaryotic cell) by transfection, infection, electroporation, microinjection, cell fusion, protoplast fusion or ballistic bombardment. Alternatively, vectors of the invention may be introduced directly into a host cell using a particle gun.

25 In one embodiment, the nucleic acid sequence of the expression cassette of the first aspect (comprising a promoter, a S/MAR and a PolyA (without transgene) with ITRs) is substantially set out as SEQ ID No: 3.

```

30  ggccactccc tctctgcgcg ctcgtcgcgt cactgaggcc gggcgaccaa aggtcgccccg
    >>.....ITR.....>

    acgcccgggc ttgcccggg cggcctcagt gagcgagcga gcgcgcagag agggagtggc
    >.....ITR.....>

35  caactccatc actagggggt cctggagggg tggagtcgtg acccctaaaa tgggcaaaca
    >.....ITR.....>

                                     >>.....LP1.....>

40  ttgcaagcag caaacagcaa acacacagcc ctccctgcct gctgaccttg gagctggggc
    >.....LP1.....>

    agaggtcaga gacctctctg ggcccatgcc acctccaaca tccactcgac ccttgggaat
  
```

- 7 -

```

>.....LP1.....>
ttcgggtggag aggagcagag gttgtcctgg cgtgggttag gtagtgtgag aggggaatga
>.....LP1.....>
5  ctcctttcgg taagtgcagt ggaagctgta cactgccag gcaaagcgtc cgggcagcgt
>.....LP1.....>
10  aggcgggcga ctcagatccc agccagtga cttagcccct gtttgctcct ccgataactg
>.....LP1.....>
gggtgacctt ggtaaatatt caccagcagc ctccccggtt gccctctggt atccactgct
>.....LP1.....>
15  taaatacggg cgaggacagg gccctgtctc ctcagcttca ggcaccacca ctgacctggg
>.....LP1.....>
acagtgaatc cggactctaa ggtaaatata aaatttttaa gtgtataatg tgtaaacta
>..LP1..>>
20  ctgattctaa ttgtttctct cttttagatt ccaacctttg gaactgaatt ctagaccacc
agctgacctg aacgcgtgat atcggtaccg ctagcgcggc cgcaagcttg gtcaagaatg
                                     HPRT SMAR >>.....>
25  ctgattttaa ttcttctaac agctttatcc ctcagaaggg aaggcaagca agttatatat
>.....HPRT SMAR .....>
gtagttttatt tgtaagactg atatgaaatt ggaagatgaa tctactatta gctttaatta
30  >.....HPRT SMAR .....>
tttttacatt taggaatatt gcatcagtaa ctcataattt tggttttctg ttatcctgag
>.....HPRT SMAR .....>
35  ttaacacaaa ttatccaagg agatggcgga tcatctgctt tgagggtgtt ttttttgaga
>.....HPRT SMAR .....>
attttaatgt atctgaatat aaaaggtaaa aatatgccaa ctagcaattt ctgcccattc
>.....HPRT SMAR .....>
40  cagaagtttg gaaatattac tcattactag gaattaaata aaatatggtt tatctattgt
>.....HPRT SMAR .....>
tatacctctt ttaattcaca tagctcattt ttatctttta tttttgttg tttttttga
45  >.....HPRT SMAR .....>
gatggagtct tgcctgtgca ccaggcagga gtgcagtgat gcaatctcgg ctactctag
>.....HPRT SMAR .....>
50  ccaccgactc cctggttcaa gcgattctcc tgctgagcc ttctgagtag ctgggattac
>.....HPRT SMAR .....>
aggcaggcac caccacgcc agcgcgcccg cactagttac gtactcgatg ctttatttgt
>.....HPRT SMAR .....>
55  >>....SV40pA.....>
gaaatttgtg atgtattgc tttatttgta accattataa gctgcaataa acaagttaac
>.....SV40pA.....>
60  aacaacaatt gcattcattt tatgtttcag gttcaggggg aggtgtggga ggttttttaa

```

- 8 -

```

>.....SV40pA.....>
actagagcat ggctacgtag ataagtagca tggcgggtta atcattaact acaaggaacc
> SV40pA
5                                     >>ITR.>

cctagtgatg gagttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgggcg
>.....ITR.....>
10 accaaaggte gcccgacgcc cgggctttgc ccgggcggcc tcagtgagcg agcgagcgcg
>.....ITR.....>

cc
>> ITR
15                                     [SEQ ID No.3]

In another embodiment, the nucleic acid sequence of the expression cassette of the
first aspect (comprising a promoter, a transgene (i.e. Factor IX), a S/MAR and a
PolyA tail (without transgene) with ITRs) is substantially set out as SEQ ID No: 4.

20 ggccactccc tctctgcgcg ctgcctcgtt cactgaggcc gggcgaccaa aggtcgccccg
>>.....ITR.....>

acgcccgggc tttgccggg cggcctcagt gagcgagcga gcgcgcagag agggagtggc
25 >.....ITR.....>

caactccatc actaggggtt cctggagggg tggagtcgtg acccctaaaa tgggcaaaca
>.....ITR.....>>
>>.....LP1.....>
30 ttgcaagcag caaacagcaa acacacagcc ctccctgcct gctgaccttg gagctggggc
>.....LP1.....>

agaggtcaga gacctctctg ggcccatgcc acctccaaca tccactcgac cccttggaa
35 >.....LP1.....>

ttcgggtggag aggagcagag gttgtcctgg cgtgggttag gtagtgtgag aggggaatga
>.....LP1.....>

40 ctcccttcgg taagtgcagt ggaagctgta cactgccag gcaaagcgtc cgggcagcgt
>.....LP1.....>

aggcggggcg ctcatatccc agccagtgga cttagcccct gtttgctcct ccgataactg
>.....LP1.....>
45 gggtgacctt ggttaatat caccagcagc ctccccggtt gccctctgg atccactgct
>.....LP1.....>

taaatacggg cgaggacagg gccctgtctc ctgagcttca ggcaccacca ctgacctggg
50 >.....LP1.....>

acagtgaatc cggactctaa ggtaaataa aaatttttaa gtgtataatg tgtaaacta
>...LP1...>>

55 ctgattctaa ttgtttctct ctttttagatt ccaacctttg gaactgaatt ctgaccacc
atgcagaggg tgaacatgat catggctgag agccctggcc tgatcaccat ctgcctgctg

```

- 9 -

```

>>.....hFIXco'.....>
ggctacctgc tgtctgctga gtgcactgtg ttcttggacc atgagaatgc caacaagatc
>.....hFIXco'.....>
5 ctgaacaggc ccaagagata caactctggc aagctggagg agtttgtgca gggcaacctg
>.....hFIXco'.....>
gagagggagt gcatggagga gaagtgcagc tttagaggagg ccaggagggt gtttgagaac
10 >.....hFIXco'.....>
actgagagga ccactgagtt ctggaagcag tatgtggatg gggaccagtg tgagagcaac
>.....hFIXco'.....>
15 ccttgcttga atgggggcag ctgcaaggat gacatcaaca gctatgagtg ctggtgcccc
>.....hFIXco'.....>
tttggttttg agggcaagaa ctgtgagctg gatgtgacct gcaacatcaa gaatggcaga
>.....hFIXco'.....>
20 tgtgagcagt tctgcaagaa ctctgctgac aacaaggtgg tgtgcagctg cactgagggc
>.....hFIXco'.....>
25 tacaggctgg ctgagaacca gaagagctgt gagcctgctg tgccattccc atgtggcaga
>.....hFIXco'.....>
gtgtctgtga gccagaccag caagctgacc agggctgagg ctgtgttccc tgatgtggac
>.....hFIXco'.....>
30 tatgtgaaca gcaactgaggc tgaaccatc ctggacaaca tcaccagag caccagagc
>.....hFIXco'.....>
ttcaatgact tcaccagggt ggtggggggg gaggatgcca agcctggcca gttcccttg
35 >.....hFIXco'.....>
caagtgggtgc tgaatggcaa ggtggatgcc ttctgtgggg gcagcattgt gaatgagaag
>.....hFIXco'.....>
40 tggattgtga ctgtgcccc ctgtgtggag actgggggtga agatcactgt ggtggctggg
>.....hFIXco'.....>
gagcacaaca ttgaggagac tgagcacact gagcagaaga ggaatgtgat caggatcatc
>.....hFIXco'.....>
45 cccaccaca actacaatgc tgccatcaac aagtacaacc atgacattgc cctgctggag
>.....hFIXco'.....>
ctggatgagc cctgggtgct gaacagctat gtgaccccca tctgcattgc tgacaaggag
50 >.....hFIXco'.....>
tacaccaaca tottctgaa gtttggtctt ggctatgtgt ctggctgggg cagggtgttc
>.....hFIXco'.....>
55 cacaagggca ggtctgccct ggtgctgcag tacctgaggg tgccctggt ggacagggcc
>.....hFIXco'.....>
acctgcttga ggagcaccaa gttcaccatc tacaacaaca tgttctgtgc tggcttccat
>.....hFIXco'.....>
60

```

- 10 -

```

gagggggggca gggacagctg ccaggggggac tctggggggcc cccatgtgac tgaggtggag
>.....hFIXco'.....>

ggcaccagct tctgactgg catcatcagc tggggggagg agtgtgcat gaagggcaag
5 >.....hFIXco'.....>

tatggcatct acaccaaagt ctccagatat gtgaactgga tcaaggagaa gaccaagctg
>.....hFIXco'.....>>

10 acctgaacgc gtgatatcgg taccgctagc gcgcccgcaa gcttgggtcaa gaatgctgat
>>.HPRT SMAR ...>

tttaattott ctaacagctt tatccctcag aagggaaggc aagcaagtta tatatgtagt
>.....HPRT SMAR .....>
15 ttatttgtaa gactgatatg aaattggaag atgaatctac tattagcttt aattattttt
>.....HPRT SMAR .....>

acatttagga atattgcac agtaactcat aattttgggtt ttctgttata ctgagttaac
20 >.....HPRT SMAR .....>

acaaattata caaggagatg gcggatcac tgctttgagg tgtttttttt tgagaatttt
>.....HPRT SMAR .....>

25 aatgtatctg aatataaaag gtaaaaatat gccaaactagc aatttctgcc cattccagaa
>.....HPRT SMAR .....>

gtttggaaat attactcatt actaggaatt aaataaaata tggtttatct attgttatac
>.....HPRT SMAR .....>
30 ctottttaat tcacatagct catttttata ttttattttt gtttgttttt ttgagatgg
>.....HPRT SMAR .....>

agtcttgctc tgtcaccagg caggagtga gtgatgcaat ctgggtcac ttagccacc
35 >.....HPRT SMAR .....>

gactccctgg ttcaagcag tctcctgcct gagccttctg agtagctggg attacaggca
>.....HPRT SMAR .....>

40 ggcaccacca cgcacagcgc ggccgcacta gttacgtact cgatgcttta ttgtgaaat
>...HPRT SMAR ...>>
>>.....SV40pA.....>

ttgtgatgct attgctttat ttgtaaccat tataagctgc aataaacaag ttaacaacaa
45 >.....SV40pA.....>

caattgcatt cattttatgt ttcaggttca gggggagggtg tgggagggtt tttaaactag
>.....SV40pA.....>>

50 agcatggcta cgtagataag tagcatggcg ggtaaatcat taactacaag gaaccctag
>>...ITR...>

tgatggagtt ggccactccc tctctgcgcg ctgcctcgt cactgaggcc gggcgacaa
>.....ITR.....>
55 aggtgcgccg acgcccgggc tttgcccggg cgccctcagt gagcgagcga gcgcgcc
>.....ITR.....>>

```

[SEQ ID No.4]

Thus, the expression cassette or vector of the invention may comprise a nucleic acid sequence substantially as set out in either SEQ ID No.3 or SEQ ID No.4, or a functional fragment or variant thereof.

5

The transgene may be any gene encoding a protein, which may have therapeutic or industrial utility. For example, the transgene may encode cystic fibrosis transmembrane conductance regulator (CFTR), a blood clotting factor, such as Factor VIII, Factor VII, Factor IX, or alpha-galactosidase A or glucocerebrosidase. Factor IX was used as the transgene in the examples.

10

It will be appreciated that expression cassettes and vectors of the invention may be used to treat a wide variety of conditions.

15

Thus, in a third aspect, the invention provides an expression cassette according to the first aspect or a recombinant vector according to the second aspect, for use in therapy.

20

In a fourth aspect, there is provided an expression cassette according to the first aspect or a recombinant vector according to the second aspect, for treating cystic fibrosis, a haematological disorder, a blood clotting disorder, haemophilia A, haemophilia B, congenital FVII and FX deficiency, a urea cycle disorder, a lysosomal disorder, a disorder affecting the liver, any form of malignant disorder or Leber's congenital amaurosis.

25

In a fifth aspect, there is provided a method of treating, ameliorating or preventing cystic fibrosis, a haematological disorder, a blood clotting disorder, haemophilia A, haemophilia B, congenital FVII and FX deficiency, a urea cycle disorder, a lysosomal disorder, a disorder affecting the liver, any form of malignant disorder or Leber's congenital amaurosis, the method comprising administering, to a subject in need of such treatment, an expression cassette according to the first aspect or a recombinant vector according to the second aspect.

30

Preferably, the cassette or vector may be used to treat haemophilia B.

35

In a sixth aspect, there is provided a host cell comprising an expression cassette according to the first aspect or a vector according to the second aspect.

The host cell may be a bacterial cell. Alternatively, the host cell may be an animal cell, for example a mouse or rat or human cell. The host cell may be transformed with cassettes or vectors according to the invention, using known techniques, which
5 will depend on the type of cell.

In a seventh aspect, there is provided a transgenic host organism comprising at least one host cell according to the sixth aspect.

10 The host may be a bacterium. The host organism may be a multicellular organism, which is preferably non-human. For example, the host organism may be a mouse or rat.

It will be appreciated that vectors and medicaments according to the invention may
15 be used in a monotherapy (i.e. the sole use of a recombinant vector), for use in gene therapy techniques. Alternatively, vectors and medicaments according to the invention may be used as an adjunct to, or in combination with, known therapies or drugs.

20 The vectors and medicaments according to the invention may be combined in compositions having a number of different forms depending, in particular, on the manner in which the composition is to be used. Thus, for example, the composition may be in the form of a powder, tablet, capsule, liquid, ointment, cream, gel, hydrogel, aerosol, spray, micellar solution, transdermal patch, liposome suspension
25 or any other suitable form that may be administered to a person or animal in need of treatment. It will be appreciated that the vehicle of medicaments according to the invention should be one which is well-tolerated by the subject to whom it is given, and preferably enables delivery of the agents across the blood-brain barrier.

30 Medicaments comprising vectors and medicaments of the invention may be used in a number of ways. For instance, oral administration may be required, in which case the agents may be contained within a composition that may, for example, be ingested orally in the form of a tablet, capsule or liquid. Compositions comprising vectors and medicaments of the invention may be administered by inhalation (e.g. intranasally).

35 Compositions may also be formulated for topical use. For instance, creams or ointments may be applied to the skin.

- 13 -

Vectors and medicaments according to the invention may also be incorporated within a slow- or delayed-release device. Such devices may, for example, be inserted on or under the skin, and the medicament may be released over weeks or even months. The device may be located at least adjacent the treatment site. Such devices
5 may be particularly advantageous when long-term treatment with vectors and medicaments according to the invention is required and which would normally require frequent administration (e.g. at least daily injection).

In a preferred embodiment, vectors and medicaments according to the invention
10 may be administered to a subject by injection into the blood stream or directly into a site requiring treatment. Injections may be intravenous (bolus or infusion) or subcutaneous (bolus or infusion), or intradermal (bolus or infusion).

It will be appreciated that the amount of the vector and medicament that is required
15 is determined by its biological activity and bioavailability, which in turn depends on the mode of administration, the physiochemical properties of the vector and medicament, and whether it is being used as a monotherapy or in a combined therapy. The frequency of administration will also be influenced by the half-life of the vector within the subject being treated. Optimal dosages to be administered may
20 be determined by those skilled in the art, and will vary with the particular vector in use, the strength of the pharmaceutical composition, the mode of administration, and the advancement of the disease being treated. Additional factors depending on the particular subject being treated will result in a need to adjust dosages, including subject age, weight, gender, diet, and time of administration.

25 Typically, a dose of between 2×10^9 and 2×10^{14} of vector genome per kg of patient may be used. Preferably, a dose of between 2×10^{10} and 2×10^{13} of vector genome per kg of patient or between 2×10^{11} and 2×10^{12} of vector genome per kg of patient may be used.

30 The vector or medicament may be administered before, during or after onset of the disease condition. Daily doses may be given as a single administration (e.g. a single daily injection). Alternatively, the vector or medicament may require administration twice or more times during a day. As an example, vectors and
35 medicaments may be administered as two (or more depending upon the severity of the bacterial infection being treated) daily doses of between between 2×10^{11} and 2×10^{12} of vector genome per kg of patient (i.e. assuming a body weight of 70kg). A

- 14 -

patient receiving treatment may take a first dose upon waking and then a second dose in the evening (if on a two dose regime) or at 3- or 4-hourly intervals thereafter. Alternatively, a slow release device may be used to provide optimal doses of vectors and medicaments according to the invention to a patient without the need to
5 administer repeated doses. Known procedures, such as those conventionally employed by the pharmaceutical industry (e.g. *in vivo* experimentation, clinical trials, etc.), may be used to form specific formulations of the vectors and medicaments according to the invention and precise therapeutic regimes (such as daily doses of the vectors and the frequency of administration).

10

In an eighth aspect of the invention, there is provided a pharmaceutical composition comprising an expression cassette of the first aspect or a vector according to the second aspect, and a pharmaceutically acceptable vehicle.

15

The invention also provides in a ninth aspect, a process for making the composition according to the eighth aspect, the process comprising combining a therapeutically effective amount of an expression cassette of the first aspect or a vector according to the second aspect, and a pharmaceutically acceptable vehicle.

20

The composition may be used for treating cystic fibrosis, a haematological disorder, a blood clotting disorder, haemophilia A, haemophilia B, congenital FVII and FX deficiency, a urea cycle disorder, a lysosomal disorder, a disorder affecting the liver, a malignant disorder or Leber's congenital amaurosis. Preferably, the composition may be used to treat haemophilia B, i.e. it is a haemophilia B treatment composition.

25

A "subject" may be a vertebrate, mammal, or domestic animal. Hence, medicaments according to the invention may be used to treat any mammal, for example livestock (e.g. a horse), pets, or may be used in other veterinary applications. Most preferably, the subject is a human being.

30

A "therapeutically effective amount" of vector is any amount which, when administered to a subject, is the amount of drug that is needed to treat the infection, or produce the desired effect.

35

For example, the therapeutically effective amount of vector used may be from about between 2×10^{11} and 2×10^{12} of vector genome per kg of patient.

- 15 -

A "pharmaceutically acceptable vehicle" as referred to herein, is any known compound or combination of known compounds that are known to those skilled in the art to be useful in formulating pharmaceutical compositions.

5 In one embodiment, the pharmaceutically acceptable vehicle may be a solid, and the composition may be in the form of a powder or tablet. A solid pharmaceutically acceptable vehicle may include one or more substances which may also act as flavouring agents, lubricants, solubilisers, suspending agents, dyes, fillers, glidants, compression aids, inert binders, sweeteners, preservatives, dyes, coatings, or tablet-
10 disintegrating agents. The vehicle may also be an encapsulating material. In powders, the vehicle is a finely divided solid that is in admixture with the finely divided active agents according to the invention. In tablets, the active agent (i.e. the vector) may be mixed with a vehicle having the necessary compression properties in suitable proportions and compacted in the shape and size desired. The powders and
15 tablets preferably contain up to 99% of the active agents. Suitable solid vehicles include, for example calcium phosphate, magnesium stearate, talc, sugars, lactose, dextrin, starch, gelatin, cellulose, polyvinylpyrrolidone, low melting waxes and ion exchange resins. In another embodiment, the pharmaceutical vehicle may be a gel and the composition may be in the form of a cream or the like.

20

However, the pharmaceutical vehicle may be a liquid, and the pharmaceutical composition is in the form of a solution. Liquid vehicles are used in preparing solutions, suspensions, emulsions, syrups, elixirs and pressurized compositions. The vector according to the invention may be dissolved or suspended in a
25 pharmaceutically acceptable liquid vehicle such as water, an organic solvent, a mixture of both or pharmaceutically acceptable oils or fats. The liquid vehicle can contain other suitable pharmaceutical additives such as solubilisers, emulsifiers, buffers, preservatives, sweeteners, flavouring agents, suspending agents, thickening agents, colours, viscosity regulators, stabilizers or osmo-regulators. Suitable
30 examples of liquid vehicles for oral and parenteral administration include water (partially containing additives as above, e.g. cellulose derivatives, preferably sodium carboxymethyl cellulose solution), alcohols (including monohydric alcohols and polyhydric alcohols, e.g. glycols) and their derivatives, and oils (e.g. fractionated coconut oil and arachis oil). For parenteral administration, the vehicle can also be an
35 oily ester such as ethyl oleate and isopropyl myristate. Sterile liquid vehicles are useful in sterile liquid form compositions for parenteral administration. The liquid

- 16 -

vehicle for pressurized compositions can be a halogenated hydrocarbon or other pharmaceutically acceptable propellant.

Liquid pharmaceutical compositions, which are sterile solutions or suspensions, can be utilized by, for example, intramuscular, intrathecal, epidural, intraperitoneal, intravenous and particularly subcutaneous injection. The vector may be prepared as a sterile solid composition that may be dissolved or suspended at the time of administration using sterile water, saline, or other appropriate sterile injectable medium.

The vectors and compositions of the invention may be administered orally in the form of a sterile solution or suspension containing other solutes or suspending agents (for example, enough saline or glucose to make the solution isotonic), bile salts, acacia, gelatin, sorbitan monoleate, polysorbate 80 (oleate esters of sorbitol and its anhydrides copolymerized with ethylene oxide) and the like. The vectors used according to the invention can also be administered orally either in liquid or solid composition form. Compositions suitable for oral administration include solid forms, such as pills, capsules, granules, tablets, and powders, and liquid forms, such as solutions, syrups, elixirs, and suspensions. Forms useful for parenteral administration include sterile solutions, emulsions, and suspensions.

It will be appreciated that the invention extends to any nucleic acid or peptide or variant, derivative or analogue thereof, which comprises substantially the amino acid or nucleic acid sequences of any of the sequences referred to herein, including functional variants or functional fragments thereof. The terms "substantially the amino acid/nucleotide/peptide sequence", "functional variant" and "functional fragment", can be a sequence that has at least 40% sequence identity with the amino acid/nucleotide/peptide sequences of any one of the sequences referred to herein, for example 40% identity with the nucleotide sequence identified as SEQ ID No:1, 2, 3, or 4, and so on.

Amino acid/polynucleotide/polypeptide sequences with a sequence identity which is greater than 50%, more preferably greater than 65%, 70%, 75%, and still more preferably greater than 80% sequence identity to any of the sequences referred to are also envisaged. Preferably, the amino acid/polynucleotide/polypeptide sequence has at least 85% identity with any of the sequences referred to, more preferably at least

- 17 -

90%, 92%, 95%, 97%, 98%, and most preferably at least 99% identity with any of the sequences referred to herein.

The skilled technician will appreciate how to calculate the percentage identity
5 between two amino acid/polynucleotide/polypeptide sequences. In order to
calculate the percentage identity between two amino
acid/polynucleotide/polypeptide sequences, an alignment of the two sequences must
first be prepared, followed by calculation of the sequence identity value. The
percentage identity for two sequences may take different values depending on:- (i)
10 the method used to align the sequences, for example, ClustalW, BLAST, FASTA,
Smith-Waterman (implemented in different programs), or structural alignment from
3D comparison; and (ii) the parameters used by the alignment method, for example,
local vs global alignment, the pair-score matrix used (e.g. BLOSUM62, PAM250,
Gonnet etc.), and gap-penalty, e.g. functional form and constants.

15 Having made the alignment, there are many different ways of calculating percentage
identity between the two sequences. For example, one may divide the number of
identities by: (i) the length of shortest sequence; (ii) the length of alignment; (iii) the
mean length of sequence; (iv) the number of non-gap positions; or (iv) the number of
20 equivalenced positions excluding overhangs. Furthermore, it will be appreciated that
percentage identity is also strongly length dependent. Therefore, the shorter a pair of
sequences is, the higher the sequence identity one may expect to occur by chance.

Hence, it will be appreciated that the accurate alignment of protein or DNA
25 sequences is a complex process. The popular multiple alignment program ClustalW
(Thompson *et al.*, 1994, Nucleic Acids Research, 22, 4673-4680; Thompson *et al.*,
1997, Nucleic Acids Research, 24, 4876-4882) is a preferred way for generating
multiple alignments of proteins or DNA in accordance with the invention. Suitable
parameters for ClustalW may be as follows: For DNA alignments: Gap Open Penalty
30 = 15.0, Gap Extension Penalty = 6.66, and Matrix = Identity. For protein alignments:
Gap Open Penalty = 10.0, Gap Extension Penalty = 0.2, and Matrix = Gonnet. For
DNA and Protein alignments: ENDGAP = -1, and GAPDIST = 4. Those skilled in the
art will be aware that it may be necessary to vary these and other parameters for
optimal sequence alignment.

35 Preferably, calculation of percentage identities between two amino
acid/polynucleotide/polypeptide sequences may then be calculated from such an

- 18 -

alignment as $(N/T)*100$, where N is the number of positions at which the sequences share an identical residue, and T is the total number of positions compared including gaps but excluding overhangs. Hence, a most preferred method for calculating percentage identity between two sequences comprises (i) preparing a sequence
5 alignment using the ClustalW program using a suitable set of parameters, for example, as set out above; and (ii) inserting the values of N and T into the following formula:- Sequence Identity = $(N/T)*100$.

Alternative methods for identifying similar sequences will be known to those skilled
10 in the art. For example, a substantially similar nucleotide sequence will be encoded by a sequence which hybridizes to the sequences shown in SEQ ID No's: 1-15, or their complements under stringent conditions. By stringent conditions, we mean the nucleotide hybridises to filter-bound DNA or RNA in 3x sodium chloride/sodium citrate (SSC) at approximately 45°C followed by at least one wash in 0.2x SSC/0.1%
15 SDS at approximately 20-65°C. Alternatively, a substantially similar polypeptide may differ by at least 1, but less than 5, 10, 20, 50 or 100 amino acids from the sequences referred to herein or encoded by the described nucleic acid molecules.

Due to the degeneracy of the genetic code, it is clear that any nucleic acid sequence
20 described herein could be varied or changed without substantially affecting the sequence of the protein encoded thereby, to provide a functional variant thereof. Suitable nucleotide variants are those having a sequence altered by the substitution of different codons that encode the same amino acid within the sequence, thus producing a silent change. Other suitable variants are those having homologous
25 nucleotide sequences but comprising all, or portions of, sequence, which are altered by the substitution of different codons that encode an amino acid with a side chain of similar biophysical properties to the amino acid it substitutes, to produce a conservative change. For example small non-polar, hydrophobic amino acids include glycine, alanine, leucine, isoleucine, valine, proline, and methionine. Large non-
30 polar, hydrophobic amino acids include phenylalanine, tryptophan and tyrosine. The polar neutral amino acids include serine, threonine, cysteine, asparagine and glutamine. The positively charged (basic) amino acids include lysine, arginine and histidine. The negatively charged (acidic) amino acids include aspartic acid and glutamic acid. It will therefore be appreciated which amino acids may be replaced
35 with an amino acid having similar biophysical properties, and the skilled technician will know the nucleotide sequences encoding these amino acids.

- 19 -

All of the features described herein (including any accompanying claims, abstract and drawings), and/or all of the steps of any method or process so disclosed, may be combined with any of the above aspects in any combination, except combinations where at least some of such features and/or steps are mutually exclusive.

5

For a better understanding of the invention, and to show how embodiments of the same may be carried into effect, reference will now be made, by way of example, to the accompanying diagrammatic drawings, in which:-

10 **Figure 1** shows, in the top panel, a schematic of various embodiments of AAV vectors that were tested. A control vector (above) is shown having a Factor IX expression cassette (LP1-hFIXco) with a SV40 polyadenylation signal sequence, as well as test vectors (below) containing a MAR sequence which is positioned 3' of a hFIX transgene. The MAR was from either the interferon B gene, the HPRT gene, the
15 Apolipoprotein B gene or the Kaposi sarcoma herpes virus. The bottom panel shows the hFIX levels in murine plasma for each of the vectors tested. For those vectors including a MAR, the MAR was tested in forward (F) and reverse (R) orientations. The hFIX levels have been corrected for transgene copy number in the liver at 4 weeks after tail vein administration of 1×10^{11} vg/mouse (n=4/group);

20 **Figure 2** shows hFIX levels in murine plasma at 2 weeks after a single tail vein administration of the control or MAR (N=3/group) containing vector at a dose of either 1×10^9 vg/mouse (1×10^9) or 1×10^{10} vg/mouse;

Figure 3 shows that AAV vectors containing HPRT MAR mediate higher levels of hFIX expression in macaques. A dose of 4×10^{12} vg/kg of ssAAV8-LP1-hFIXco (blue
25 bar) or ssAAV8-LP1-hFIXco-HPRT (red bar) was injected into the peripheral vein of male rhesus macaques as a bolus infusion. Human FIX expression was assessed at 7 days after gene transfer. Shown are the results of repeated measurements expressed as mean $\pm$ SEM;

Figure 4 shows that single stranded AAV containing HPRT MAR have the same
30 potency as self-complementary AAV vectors. Equivalent alkaline gel based titres (1×10^9 vg/mouse) of ss-rAAV-LP1-hFIXco-HPRT (red line with closed circles) and scAAV-LP1-hFIXco (black line with closed squares) pseudotyped with serotype 8 capsid were administered into the tail vein of cohorts (n=3) of male C57Bl/6 mice. FIX levels were assessed at various time points and expressed as mean percentage of
35 normal values $\pm$ SEM;

Figure 5 is a schematic of the 573bp HPRT MAR showing the wild type HPRT gene co-ordinates (+3858-+4430) at the top. Smaller HPRT fragments were created by

- 20 -

PCR and their size and co-ordinates with reference to the 573bp HPRT MAR are shown ranging from 0-573bp;

Figure 6 shows results of gene transfer studies in cohorts of male C57Bl/6 mice (n=3) with serotype 8 capsid pseudotyped single stranded AAV vector containing the LP1-hFIXco expression cassette including either the full length, 573bp MAR or its sub-fragments. The data shows that the level of FIX expression observed with full length 573bp MAR at 4 weeks after gene transfer was comparable to that achieved with HPRT fragments 2 and 2b suggesting that the enhancer activity of HPRT MAR is mediated predominantly by sequences in its 131bp sub-fragment;

Figure 7 shows quantitative chromatin immunoprecipitation (qChIP). Analysis of the pattern of (active) histone lysine acetylation (acH3), dimethylation of lysine 4 (2MeK4) and (inactive) histone dimethylation of lysine 9 (2meK9H3) at the promoter region of the viral genome in mice injected with ssAAV8-LP1-hFIXco with and without HPRT SMA/R. There is a decrease in the 2meK9H3 mark in the AAV promoter in the mice injected with the construct containing the HPRT SMA/R, whilst the acH3 and 2MeK4 marks increase or remain constant. Enrichment values for the modifications studied were normalized to the Input;

Figure 8 shows the construction and characterisation of single stranded AAV-hFIXco containing IFN β S/MAR (A) Schematic of vectors. Each vector is represented schematically as it is packaged inside the virion. Common features of all vectors include the hybrid liver specific promoter (LP1), SV40 intron, codon-optimised human FIX cDNA (hFIXco) and a truncated SV40 late polyA (tSV40pA) flanked by AAV inverted terminal repeats (ITRs represented as hairpin loops). ssAAV-LP1-hFIXco-Control and vector genomes containing the IFN β S/MAR in the 3' region of the expression cassette (IFN-F and IFN-R) are 2.4Kb and 3.3Kb in size respectively. (B) Characterisation of ssAAV-FIXco proviral DNA. Proviral DNA was extracted and electrophoresed on a 1% alkaline agarose gel. The ssAAV-LP1-hFIXco-Control genome migrated as a 2.4Kb monomer with a proportion of DNA observed as self-complementary dimers at approximately 5Kb. The ssAAV-LP1-hFIXco-IFN-R genome migrated in a 3.3Kb monomeric form only. The control consisted of scAAV vector particles: with vector genomes observed as 5Kb dimers. (C) Transgene expression in C57BL/6 mice. hFIX expression profiles over a period of 8 weeks after tail-vein administration of 4×10^{12} vg/kg (n=6) of ssAAV-hFIXco-IFN-F/R and ssAAV-hFIXco-Control vectors into 4 to 6 week old male C57BL/6 mice. All results are shown as an average together with the standard error of the mean ($\pm$ SEM). (D) Proviral copy numbers in liver. Proviral copy number was determined by a Q-PCR assay on liver samples collected 6 weeks after tail vein administration of ssAAV-

- 21 -

hFIXco-IFN-F/R and ssAAV-hFIXco-Control vectors. Standards consisted of serial dilutions of ssAAV-LP1-hFIXco vector plasmid diluted in murine liver genomic DNA derived from a naïve animal. Results are presented as mean proviral copies/cell $\pm$ SEM;

5 **Figure 9** shows construction of ssAAV-LP1-hFIXco containing alternative S/MARs. (A) Schematic of vectors. Each vector is represented schematically as it is packaged in the virion. Common features of each vector are as previously mentioned in Figure 8A. The HPRT and ApoB S/MAR elements and KSHV control element were cloned 3' of the hFIXco cDNA in forward and reverse orientation (HPRT-F/R, ApoB-F/R and
10 KSHV- F/R respectively). All vectors depicted are approximately 3-3.3Kb in size. (B) Transgene expression in C57BL/6 mice. hFIX levels at 4 weeks after tail-vein administration of 4×10^{12} vg/kg (n=3) of all vectors depicted are represented graphically as averages $\pm$ SEM. (C) Proviral copy numbers in liver. Proviral copy number was determined by a Q-PCR assay on liver samples collected 4 weeks after
15 transduction and depicted as mean proviral copies/cell $\pm$ SEM. (D) HFIX mRNA levels in liver. mRNA levels were determined by a Q-PCR assay on cDNA generated from total liver RNA collected 4 weeks after tail-vein administration of 4×10^{12} vg/kg of ssAAV-hFIXco-IFN-R, ssAAV-hFIXco-HPRTF/R and ssAAV-hFIXco-Control into C57BL/6 mice. Results are presented as mean mRNA levels corrected for the amount
20 of housekeeping cDNA (GAPDH) present in each sample $\pm$ SEM and then corrected for proviral copy number to take account of variation in transduction;

Figure 10 shows lower doses of ssAAV-LP1-hFIXco-HPRT-F mediate therapeutic transgene expression in male C57BL/6 mice. (A) Transgene expression. hFIX expression profiles over a period of 4 weeks after tail-vein administration of 4×10^{10}
25 vg/kg (n=3) of ssAAV-hFIXco-HPRT-F and ssAAV-hFIXco-Control vectors into 6 to 8 week old male C57BL/6 mice. All results are shown as an average $\pm$ SEM. (B) Proviral copy numbers in liver. Proviral copy number was determined by a Q-PCR assay on liver samples collected 4 weeks after tail-vein administration of 4×10^{10} vg/kg of ssAAV-hFIXco-HPRT-F and ssAAV-hFIXco-Control vectors. Results are
30 depicted as mean proviral copies/cell $\pm$ SEM;

Figure 11 shows higher potency of HPRT S/MAR containing vector in rhesus macaques. Human FIX levels in rhesus plasma were determined at the indicated time points after peripheral vein administration of 4×10^{12} vg/kg of ssAAV-LP1-hFIXco-HPRT-F and ssAAV-LP1-hFIXco-Control vectors. Results are depicted as
35 averages of repeat ELISA results $\pm$ SEM;

Figure 12 shows deletion analysis of HPRT S/MAR. (A) Schematic: Smaller HPRT S/MAR fragments were generated by a standard PCR method and cloned 3' of the

- 22 -

hFIXco cDNA in forward orientation (FR1, FR2, FR2a and FR2b). (B) Transgene expression in C57BL/6 mice. hFIX levels at 2 weeks after tail-vein administration of 4×10^{10} vg/kg (n=3) of ssAAV-hFIXco-HPRT-F and ssAAV-hFIXco-HPRT-FR2b vectors into 6 to 8 week old male C57BL/6 mice are represented graphically as averages $\pm$ SEM;

Figure 13 shows HPRT S/MAR improves potency of self complementary LP1-hFIXco vector. hFIX levels at 4 weeks after tail-vein administration of 4×10^{10} vg/kg (n=3) of scAAV-hFIXco-HPRT-FR2b and scAAV-hFIXco-Control vectors into 6 to 8 week old male C57BL/6 mice are represented graphically as averages $\pm$ SEM;

Figure 14 shows epigenetic modification of ssAAV-LP1-hFIXco proviral DNA (A) IFN β S/MAR ChIP. ChIP analysis was carried out on chromatin extracted from liver samples collected 6 weeks after tail-vein administration of 4×10^{12} vg/kg of ssAAV-LP1-hFIXco-IFN-R and ssAAV-LP1-hFIXco-Control vectors into 4 to 6 week old male C57BL/6 mice. Vector genome regions amplified by Q-PCR included proximal and distal regions of the LP1 promoter (152 and 366 respectively). Antibodies used for ChIP included anti H3K9me2, anti HP1 α , anti H3K4me2 and anti H3Ac. (B) HPRT S/MAR ChIP. Similar analysis was performed on liver samples collected 4 weeks after administration of 4×10^{12} vg/kg of ssAAV-LP1-hFIXco-HPRT-F and ssAAV-LP1-hFIXco-Control vectors. Antibodies used for ChIP included anti H3K9me2 anti H3K4me2 and anti H3Ac. All results are presented as a percentage of total input chromatin (calculated as an average from 2 Ct values) $\pm$ SEM; and **Figure 15** shows cytosine methylation profiles of the AAV LP1 promoter revealed by bisulphite sequencing. Bisulphite sequencing was carried out on DNA extracted from rhesus liver samples 1 week, 2 months and 1 year after peripheral vein administration of scAAV8-LP1-hFIXco vector at a dose of 2×10^{12} vg/kg. The vector genome is represented schematically (top). Boxes represent CpG regions of the LP1 promoter (bottom). Unfilled boxes depict non-methylated CpG regions.

Examples

Example 1

Materials & Methods

Molecular Cloning of MAR elements into pAV-LP1-hFIXco-SV40pa plasmid and generation of Short HPRT MAR fragments by PCR

All other MAR sequences were synthesised in a PUC57 plasmid vector with flanking Not1 restriction sites (Genescript), allowing for all new MARs to be cloned into Not1

- 23 -

digested pAV-LP1-hFIXco-mcs-SV4opA plasmid DNA in both forward and reverse orientation. pAV-LP1-hFIXco-MCS-SV4opA is a single stranded AAV expression cassette in which the codon optimised hFIX gene is under the control of the LP1 liver specific promoter. Smaller HPRT MAR fragments were next generated by PCR from
5 1µg of plasmid DNA containing the HPRT MAR (pAV-LP1-hFIXco-HPRT MAR-SV4opA) using the Hot Start Taq PCR kit (Qiagen). To the template 1µl of each primer, 1µl DNTPs, 10µl 10X PCR buffer, 0.5µl Hot start Taq polymerase and water was added to a final volume of 100µl per reaction.

10 Using a thermal cycler, reactions were initiated by heating samples at 95°C for 15 minutes. The following cycling conditions were then repeated 35 times: 1 min 94°C denaturation, 1 min annealing (temperature set to 5°C below the melting temperature of primer pairs), 1 min 72°C extension. After the 35 cycles a final 10 min extension step at 72°C was carried out. All reactions were done in conjunction with a
15 no template negative control. PCR products were run on a 1% agarose gel at 120V. The appropriate bands were cut from the gel and DNA was purified using the Gel Purification Kit (Qiagen) as per the manufacturer's instructions.

Restriction Enzyme digests

20 Primers for the generation of short HPRT MAR fragments were designed to incorporated Kpn1 (forward primers) and Not1 (reverse primers) restriction sites for cloning into the pAV-LP1-hFIXco-mcs-SV4opA plasmid in forward orientation. PCR products and pAV-LP1-hFIXco-mcs-SV4opA were first digested with Kpn1 followed by purification using the PCR purification kit (Qiagen) and subsequent Not1
25 digestion.

All restriction digests were carried out using the following conditions. 5µg DNA (or the remaining amount of PCR product following gel purification) was added to 10µl of the appropriate reaction buffer, 1µl 100X BSA, 8µl of the required enzyme (all
30 NEB) and molecular biology water up to 100µl. Reactions were carried out at 37°C for 4 hrs. All digested Vector backbone and MAR insert DNA samples were run on a 1% agarose gel at 120V. Required bands were excised and purified using the Gel Purification kit (Qiagen) as per the manufacturer's instructions.

Alkaline Phosphatase Treatment of Backbone DNA

35

- 24 -

To remove 5' phosphate groups from vector backbone DNA (to prevent self ligation) remaining DNA was incubated at 37°C for 1hr with 1X NEB buffer 3 and 0.5units/ μ g DNA calf intestinal phosphatase (CIP) (NEB). DNA was purified by using the gel extraction kit (Qiagen) as per the manufacturer's instructions.

5

Ligations

All ligations for the cloning of MAR elements into the pAV-LP1-hFIXco-mcs-SV40pA backbone were carried out using T4 DNA ligase (NEB). On ice, vector backbone DNA was mixed with an approximate 3 fold excess (as determined by the appearance of
10 band intensity on a 1% agarose gel) of MAR fragment DNA. To the DNA 2 μ l of T4 DNA ligase buffer, 2 μ l of T4 DNA ligase and molecular biology water was added to a final volume of 20 μ l. As negative control vector backbone DNA alone was used under the same reaction conditions. Control reactions were later used to determine relative background levels of self-ligated backbone DNA. All reactions were carried out by
15 overnight incubation at 4°C for transformation into DH5 α competent E.Coli (NEB) the following day.

Transformation

For transformation of pAV-LP1-hFIXco-MAR constructs, DH5 α competent cells
20 were thawed briefly on ice and divided into 25 μ l aliquots. 5 μ l of each ligation mixture including negative control was then added to one individual aliquot of competent cells and left for 30 min on ice. For positive control 1 μ l of PUC19 plasmid DNA (NEB) was also added to 25 μ l of DH5 α cells and incubated on ice for the same period of time. Next, to allow for the uptake of DNA, all samples and controls were heat
25 shocked at 42°C for 20 sec and immediately placed on ice for a further 5 minutes. 425 μ l of SOC medium (NEB) was then added to each sample followed by a 1 hr incubation at 37°C with shaking. As all constructs contained an ampicillin resistance gene, samples were spread onto agar plates containing 100 μ g/ml ampicillin under aseptic technique. Plates were incubated at 37°C overnight for the selection of
30 colonies the following day.

Screening colonies

To determine whether ligations had been successful plates were first checked to ensure that higher numbers of colonies were present for each construct in
35 comparison to negative control. The positive control plate was used to verify that transformation had occurred. Following these checks colonies were picked using

- 25 -

pipette tips and placed in 5ml of LB medium. Samples were left in a 37°C shaking incubator overnight. The following day, DNA was extracted from 2ml of the cultures using the NucleoSpin® Plasmid kit (Machery Nagel) as per the manufacturer's instructions. 3 ml of each culture was reserved and stored at 4°C.

5

To ascertain whether the correct clones had been generated the extracted DNA was used in a diagnostic restriction digest whereby the enzymes selected would generate fragments of sizes that enabled correct constructs to be differentiated from the incorrect. All diagnostic digests were carried out at 37°C for 3 hrs with the following components per sample: 5µl DNA, 1µl buffer, 0.1µl BSA, 0.25µl enzyme and 3.65µl molecular biology water. Samples were run on a 1% agarose gel at 120V.

10

Where positive digests were identified the corresponding 3ml of reserved bacterial culture was used to inoculate a further litre of LB medium. These cultures were incubated at 37°C overnight with shaking. The next day DNA was extracted from the cultures using the Mega Prep kit (Qiagen) as per the manufacturer's instructions. The presence of the correct construct was again verified by restriction digest as above. The collected DNA was then used for the preparation of AAV 2/8 viruses.

15

AAV2/8 Production

20

For each virus 40× 15cm plates of HEK293T cells were used at approximately 70% confluence on the day of transfection. Cells were cultured at 37 °C, 5% CO₂ in DMEM (PAA) with 10% FBS (GIBCO) and enzymatically passaged every 2-3 days. For the transfection mixture 6ml of Polyethyleneimine (PEI) (Polysciences) was added to 54ml serum free DMEM. In a separate tube 1.8mg of helper plasmid (containing essential genes from the adenoviral genome to support rescue and replication of the viral genome), 0.6mg 2/8 plasmid (containing rep and cap genes), 0.6mg of transgene containing plasmid, and 40µg of PCL10.1-EF1α-GFP plasmid was added to serum free DMEM to a final volume of 62ml. The mixture containing DNA was then filtered into the PEI containing mixture through a 0.2µm syringe filter. The solution was mixed and incubated at room temperature for 15min. 3ml of transfection solution was then added to each 15cm dish. After 72hrs of incubation, cells were harvested using cell scrapers and transferred to 250ml conical tubes (a small fraction of each prep was reserved for flow cytometry to determine transfection efficiency as a percentage of GFP positive cells). All tubes were centrifuged at 1500rpm for 10min. Supernatants were discarded and cell pellets resuspended in 25ml 1×TD buffer (140mM NaCl, 5mM KCl, 0.7mM K₂HPO₄, 3.5mM MgCl₂, 25mM Tris pH7.5). Cells

25

30

35

- 26 -

were again spun for 10min at 1500rpm and supernatants were discarded. Cell pellets were then collectively resuspended in 40ml 1×TD buffer. To lyse the cells and release viral particles freeze thaw cycles of 30min at -80°C followed by 10 min at 37°C were repeated 5 times.

5

Prior to purification cell were incubated at 37°C for 30min with 0.5% deoxycholic acid and 50units/ml Benzonase. Lysates were then spun at 4000×g for 30minutes at 18°C, passed through a 0.45µm filter and diluted 1/3 in filtered PBS.

10 Virus purification by affinity chromatography (Äkta explorer)

Purification of AAV2/8 was carried out using a 5ml AVB Sepharose (GE Healthcare) packed column. Briefly, following line washes with the appropriate buffers, the affinity medium was equilibrated with filtered PBS pH 7.5 for 10-20min (or until pH and absorbance readings were stable) at a flow rate of 2ml/min. The cell lysates
15 obtained from 40 plate preparations was then loaded onto the column at 5ml/min (traceable as an increase in absorbance at 260 and 280nm as unbound protein and DNA exits the column). Residual unbound protein and DNA was washed off the column in filtered PBS allowing for viral particles exclusively to be eluted in 50mM Glycine pH 2.7 (5ml/min). The eluate was collected in 1ml fractions into tubes
20 containing 30µl Tris pH 8.8 to neutralise the glycine. The virus containing fractions were identified by a peak in absorbance at 260 and 280nm. To remove glycine and restore pH, fractions were pooled together and dialysed overnight at 4°C in PBS using a 10KDa cut-off dialysis cassette (Slide-A-Lyzer Thermo Scientific).

25 Titration of AAV8 (QPCR based method)

Using primers against the LP1 promoter, the titre of all AAV2/8 viruses was determined from a standard generated from LP1 containing plasmid DNA. Each virus was quantified in three different dilutions (1:100, 1:1000 and 1:10000 in molecular biology water). 10µl of virus samples, negative control samples (Water)
30 and standards were added to a 96 well plate in triplicate. To each well the following master mix was added: 12.5µl SYBR Green PCR master mix (Qiagen), 1µl forward primer, 1µl reverse primer, 0.5µl molecular biology water (15µl master mix/well). Plates were then sealed, briefly spun by centrifugation and placed in a thermal cycler (Eppendorf). After an initial heating period of 5 min at 95°C the following steps were
35 carried out for 40 cycles: 95°C 10 sec and 60°C 30 sec. Copy number of viral genomes could then be obtained using Realplex software (Eppendorf). To calculate viral

- 27 -

genome copies per ml, copy number recorded per 10µl reaction was multiplied by the dilution factor and 100. As all standards used were double stranded DNA, values derived for single stranded viruses were multiplied by 2.

5 AAV8 capsid evaluation (acrylamide gel based method)

The amount of AAV8 capsid was assessed by relative band intensity of capsid proteins (VP1, VP2 and VP3). In brief samples of purified vector were heated at 95°C in an equal volume of Laemmli buffer (SIGMA) for 5 min to denature proteins.

10 Samples, protein ladder and a control virus sample of known titre were then run on a pre-made polyacrylamide gel (Thermo Scientific) in 1×HEPES running buffer (121g Tris, 238g HEPES, 10g SDS and water up to 1L) at 150V for 30 min. Gels were then stained for 1hr using Coomassie blue dye. After destaining Gels were imaged using Genesnap software.

15 Titration of AAV8 (alkaline gel based method)

Viruses were run on denaturing alkaline gels and band intensity was quantified by comparison against a standard (DNA ladder). To make the gels 2ml of 10X alkaline electrophoresis buffer (50X=30g 1.5M NaOH, 50ml 0.5M EDTA pH 8.0, up to 500ml deionised water) was added to 98mls of 0.8% agarose (pre cooled below 20 50°C). 8.5µl of alkaline sample loading buffer (200µl glycerol, 80µl 50X alkaline running buffer, 60µl 20% SDS, Xylene cyanol, up to 1ml deionised water) was added to each 25µl aliquot of virus to be titred. Samples were mixed well and placed on ice. Samples were then run at 20V on the gel overnight at 4°C with 5µl Hyperladder I (Bioline) loaded in a separate lane. The next day gels were washed for one hour in 25 300ml Tris pH 8.0 with agitation. Gels were then transferred to 100ml of 4x Gelred solution (prepared in 0.1M NaCl) and left for two hours in darkness with agitation. After rinsing twice with tap water gels were imaged using Genesnap (Syngene) software. Images were then used for quantification of viral DNA using Gene tools software (Syngene). Output raw volume and background values for the intensity of 30 each band were exported into an excel file. To calculate viral titre, background was subtracted from raw volume to obtain net volume values and a standard curve was generated by plotting ng of DNA in the Hyperladder against corresponding net volumes. The equation from the chart was used to calculate ng of DNA in each sample and these values were multiplied by 40 to derive DNA concentration in 35 ng/ml. Viral titre could then be calculated by determining the number of AAV

- 28 -

genomes in the given amount of DNA (based on the length of the viral genome in question).

Animal work

5 All mice used in the study were male C57BL/6 aged between 4 and 8 weeks. AAV2/8 viruses were prepared in X-VIVO 10 media and administered via tail vein injection with 3-4 mice per group. Mice were bled at regular intervals and plasma was collected by centrifugation (4000×g for 15 min at 4°C) of blood samples in tri-sodium citrate. 4 weeks after the injection date mice were sacrificed and livers were
10 taken for further investigations.

Human Factor IX ELISA

Levels of hFIX in mouse plasma and supernatants from transduced cell cultures were determined by FIX ELISA. Standards for analysis on cell cultures and mouse plasma
15 were prepared by serial dilutions of 1000% recombinant hFIX coagulation reference (Technoclone) in X-Vivo 10 (GIBCO) or C57BL/6 plasma respectively. Huh7 supernatants and corresponding standards were diluted 1:2 in PBST +2% bovine serum albumin (dilution buffer). Mouse plasma samples and corresponding standards were diluted 1:100 in dilution buffer (where required, plasma samples
20 were initially diluted in naive C57BL/6 plasma to ensure values fell within the range of the standard curve). All Samples and standards were prepared on ice and stored long term at -80°C where required. 50µl Coating antibody (9µl rhesus anti-human antisera diluted in 5ml 0.1M NaHCO₃ pH 9.2) was added to each well of 96 well ELISA plates and incubated at 4°C overnight. All wells were washed 5 times in with
25 200µl PBST and then blocked in PBST + 6% BSA for 1hr at 37°C. Plates were washed 5 times in PBST (200µl/well) and 50µl of diluted standards and samples were then added to the wells in duplicate. Plates were incubated at 37°C for 2hrs. Following incubation plates were washed 5 times in PBST (200µl/well). Secondary antibody was prepared by adding 5µl Goat antihuman FIX-HRP peroxidase labelled antibody
30 (Affinity Biological Inc) to 10ml of dilution buffer. 100µl of diluted secondary antibody was added to each well. Plates were incubated with secondary antibody for 1hr at 37°C followed by 5 washes with PBST. 200µl SIGMA FAST™ OPD peroxidase substrate was then added to each well and colour was allowed to develop for 3-4 minutes. To stop the reaction 50µl/well of 3M HCL was added. Plates were read at
35 490nm on an ELISA plate reader.

Viral Genome copy number analysis

To determine viral genome copy number per cell, genomic DNA was extracted from mouse liver samples using the DNeasy® Blood and Tissue kit (Qiagen) as per the manufacturer's instructions. For each sample DNA was diluted to an approximate
5 concentration of 7.5ng/μl and loaded in 5μl aliquots in duplicate to a 96 well QPCR plate. 20μl of QPCR master mix (12.5μl SYBR Green master mix, 1μl forward LP1 primer, 1μl reverse LP1 primer, 5.5μl molecular biology water per well) was then added to each well and cycling conditions were carried out as described previously. Copy number was determined from an LP1 plasmid DNA standard (using LP1
10 primers) and DNA concentration for each sample was determined from levels of housekeeping gene expression (derived from a mouse genomic DNA standard using Gapdh primers in a separate reaction). To calculate copies per cell LP1 copy number results were divided by the amount of DNA in each sample (ng); these figures were then multiplied by 37.5 (the theoretical amount of DNA loaded in each well) and
15 divided by 6528.5 (the approximate number of cells from which 37.5ng of DNA is derived). To normalise FIX levels for viral genome copy number (FIX/copy/cell) values of % normal FIX were divided by copy number from the corresponding liver DNA sample.

FIX RNA copy number analysis

To determine RNA levels per viral genome copy per cell, RNAlater® (Qiagen) stabilised mouse liver samples were placed in liquid nitrogen and disrupted with a mortar and pestle. Samples were homogenised using Qiashredder® (Qiagen) columns and RNA was extracted using the RNeasy Plus® kit (Qiagen) as per the
25 manufacturer's instructions. RNA was then converted to cDNA by using the cDNA synthesis kit (Bioline). 2μl of cDNA samples, cDNA standards and negative controls (-RT and water) were then loaded onto a 96well QPCR plate in duplicate. 23μl of master mix (12.5μl SYBR Green master mix, 1μl forward LP1 primer, 1μl reverse LP1 primer and 8.5μl molecular biology water per well) was then added to each well and
30 QPCR cycling conditions were carried out as previously described. To normalise the data for the amount of cDNA loaded for each sample, a separate QPCR to quantify levels of housekeeping gene cDNA was performed using murine Gapdh primers (1X) suitable for cDNA samples (Qiagen). To calculate RNA/copy/cell values of LP1 copies were divided by results for cDNA concentration (ng). Results were then divided by
35 viral genome copy number per cell to give an approximation of transcription levels per copy.

- 30 -

Chromatin Immunoprecipitation

Liver tissue (about 200mg) was chopped into small pieces (sesame seed size), transferred to a 40 um nylon cell strainer and washed in PBS with protease inhibitors. Livers were fixed with 1% formaldehyde at RT for 10 min, lysed and sonicated using a Bioruptor (Diagenode) for 45 minutes. The sonicated extract was immunoprecipitated with specific antibodies (4ug, O/N): 2MeK9H3 (ab12220); 2MeK4H3 (Millipore 07-030); acH3 (Millipore 06-599). After washing the beads, the immunoprecipitated DNA/protein complexes were eluted, and reverse cross-linked. DNA was purified and subjected to qPCR with specific primers for the LP1 promoter region. (TTCGGTAAGTGCAGTGGGAAG – SEQ ID No: 5).

Experiments conducted

The inventors have tested the role of the four different matrix associated regions (MARs) in the context of rAAV-mediated gene transfer in murine models. These MARs included domains from the interferon β (IFN- β), hypoxanthine-guanine phosphoribosyltransferase (HPRT) and apolipoprotein B (ApoB) genes. These MARs were selected on the basis of size (<1000bp) and were then graded according to their *in silico* theoretical approximation of the “energy (GX)” of unwinding. As additional control sequence of similar length from the Kaposi sarcoma herpes virus (KSHV) genome was cloned downstream of the codon optimised FIX cDNA.

The *in silico* studies suggested that the IFN- β and ApoB MARs were likely to be most efficient at enhancing gene expression with GX values of 1.55 and 0.8, respectively. The higher GX (2.5) value for the HPRT MAR suggested that it was least likely to maintain the rAAV genome in an unwound state and, therefore, may not enhance transgene expression. Head-to-head comparisons of the four MARs in the context of the LP1-hFIXco, an expression cassette that is being tested in the clinic, revealed highly surprising, but consistent data, as described below.

Results

Referring to Figure 1, the inclusion of the 587bp HPRT MAR downstream of the rAAV-LP1-hFIXco resulted in the highest level of human FIX in the plasma of mice following a single tail vein administration of 1×10^{11} single-stranded LP1-hFIXco vector particles/mouse. This is at least 10x fold higher than the level of expression achieved with an identical construct containing the IFN β MAR and over 100x fold higher than observed with the same construct without MAR.

With reference to Figure 2, reduction of vector dose by two logs to 1×10^9 vg/mouse resulted in hFIX expression at supra physiologic human FIX levels ($508.22 \pm 83.362\%$ of normal) therapeutic levels in mice following administration of rAAV-LP1-hFIXco containing the HPRT MAR, but not when the same expression cassette was used in the context of other MAR elements. This means that when the HPRT MAR is included in a FIX expression cassette, substantially lower doses of vector are required to achieve therapeutic levels of transgene expression. This has important safety implications as well as easing the burden on vector production.

Referring to Figure 3, there is shown that AAV vectors containing HPRT MAR mediate higher levels of hFIX expression in macaques. Figure 4 shows that single stranded AAV containing HPRT MAR have the same potency as self-complementary AAV vectors.

Figure 5 is a schematic of the 573bp HPRT MAR showing the wild type HPRT gene co-ordinates (+3858-+4430) at the top, and Figure 6 shows the results of gene transfer studies in cohorts of male C57Bl/6 mice (n=3) with serotype 8 capsid pseudotyped single stranded AAV vector containing the LP1-hFIXco expression cassette including either the full length, 573bp MAR or its sub-fragments. The data shows that the level of FIX expression observed with full length 573bp MAR at 4 weeks after gene transfer was comparable to that achieved with HPRT fragments 2 and 2b suggesting that the enhancer activity of HPRT MAR is mediated predominantly by sequences in its 131bp sub-fragment.

With reference to Figure 7, the inclusion of HPRT MAR reduces the interaction of histones with inactive marks with the LP1 promoter whilst facilitating an association with histones containing active marks thus enhancing transcription from the promoter.

Conclusions

These data suggest that different classes of MARs have divergent biological properties in relation to the upregulation of rAAV transgene expression. The published literature suggests that the HPRT MAR does not have enhancer activity. However, the results described herein suggest that, in the context of rAAV mediated gene transfer, the HPRT MAR element is a surprisingly potent enhancer of transgene

- 32 -

expression. This has substantial implications for gene therapy of haemophilia B, as well as other disorders which are amenable to AAV-mediated gene transfer, such as those affecting the liver including Factor VIII, Factor VII, alpha-galactosidase A, and glucocerebrosidase, and will undoubtedly impact the design of future generations of AAV expression cassettes for use in the clinic.

Example 2

Materials and Methods

AAV-hFIX vector production and purification

The LP1-hFIXco AAV expression cassette encoding hFIX has been previously described (Nathwani et al., 2006). Key elements of this cassette include a hybrid liver-specific promoter (LP1) driving the expression of a codon optimized hFIX cDNA. In order to accommodate S/MAR elements, this expression cassette was modified by inserting a BbsI-HpaI oligonucleotide encoding a multiple cloning site (MCS) followed by a truncated SV40 polyadenylation (tSV40pA) sequence downstream of the hFIX cDNA (AAV-LP1-hFIXco-Control) (Hart et al., 1985). The IFN β S/MAR was derived from MP4253.SFGmSR-preMSV.eGFP plasmid (a gift from Dr Martin Pule, UCL Cancer Institute, UK) and ligated as a blunt fragment into the *EcoR*-V site of the MCS in LP1-hFIXco-Control plasmid. HPRT, ApoB (Apolipoprotein B) and KSHV (Kaposi sarcoma herpes virus) elements were synthesized based on published sequences; (Namciu et al., 1998; Stedman et al., 2008; Sykes et al., 1988; Nathwani et al., 2001) and inserted into the MCS at the *Not*-I site. Truncated S/MAR fragments were generated using standard polymerase chain reaction (PCR) methods and inserted into the MCS as *Kpn*-I/*Not*-I fragments. All vectors were made by the adenovirus-free transient transfection method previously described, using a chimeric AAV2 Rep-8Cap packaging plasmid (pAAV2-8) and an adenoviral helper plasmid (Nathwani et al., 2001). Vectors were purified as described before (Davidoff et al., 2004) and vector genome (vg) titers determined by standard alkaline gel or quantitative real-time PCR (Q-PCR) based methods (Fagone et al., 2011).

Animal studies

All procedures were performed in accordance with institutional guidelines under protocols approved by the Institutional Biosafety Committee and the Institutional Animal Care and Use Committee at St Jude Children's Research Hospital, Memphis,

- 33 -

USA. All animal work carried out in the United Kingdom was performed under the authority of the UK Home Office Project and Personal Licenses regulations and was compliant with the guidelines of the University College London ethical review committee. Tail-vein administration of rAAV vector particles was performed in 6 to 8 week old immunocompetent male C57BL/6 mice (Charles River Laboratories, UK) as described before (Nathwani et al., 2001). Mice were killed between 2 to 12 weeks after administration of AAV8-LP1-hFIXco vectors to harvest the liver for subsequent molecular analysis. Captive-bred male *Macaca mulatta* approximately 4 years of age and weighing between 4 and 5 kg were purchased from Charles River Laboratories (Sierra, NV). Single stranded AAV8-LP1-hFIXco vectors containing S/MAR or control sequences were administered into the peripheral vein as a bolus infusion as previously described (Nathwani et al., 2007). Human FIX antigen levels in murine and rhesus samples were determined by enzyme-linked immunosorbent assay (ELISA) as previously described (Nathwani et al., 2002).

Molecular studies

AAV transgene copy number in liver was quantified using quantitative real-time PCR (Q-PCR) with the following primers designed to amplify a 283bp region of the LP1 promoter: 5' primer, 5'-GGA GAG GAG CAG AGG TTG TC-3' [SEQ ID No: 6]; 3' primer, 5' - TGG TGG TGC CTG AAG CTG AG -3' [SEQ ID No: 7]. The Q-PCR reaction was performed using the QuantiFast SYBR Green kit as per the manufacturer's instructions (Qiagen, Crawley, UK). Primers against murine GAPDH used to normalize for variation in DNA loading were: 5' primer, 5'-TGGAGAGCCCGCTCAGACCC-3' [SEQ ID No: 8]; 3' primer, 5'-GGATTGGGTGTCCCTGCGCC-3' [SEQ ID No: 9].

To determine transgene expression levels, approximately 1 µg total liver RNA from each sample was subjected to reverse transcription using the cDNA Synthesis Kit (Bioline, London, UK). 5µL cDNA was then used in a Q-PCR reaction as described previously with primers designed to amplify a 127bp region of the hFIXco cDNA: 5' primer, 5'- GGG CAA GTA TGG CAT CTA CA -3' [SEQ ID No: 10]; 3' primer, 5'- AAA GCA TCG AGT ACGTAACT-3' [SEQ ID No: 11]. GAPDH primers used to establish equivalent cDNA loading were sourced from the Mm_Gapdh_3_SG QuantiTect primer assay (Qiagen, Crawley, UK).

- 34 -

Chromatin Immunoprecipitation

200mg of transduced fresh liver was pushed through a 40µm nylon cell strainer (BD Biosciences, Oxford, UK) in 2ml ice cold PBS containing complete protease inhibitor cocktail (Roche, Burgess Hill, UK) and 2% foetal bovine serum (FBS). Cell pellets

5 were washed twice in 10ml of the above mentioned PBS solution; then Protein-DNA crosslinks were formed by adding formaldehyde to final concentration of 1% followed by incubation at room temperature for 10min. 0.125M Glycine (final concentration) was added to quench formaldehyde for 5min at room temperature followed by 2 washes in ice cold PBS with protease inhibitor and 2% FBS. To fragment DNA,

10 chromatin sonication parameters used were 15 s pulse on and 15 s pulse off for 45 min using a Bioruptor sonicator (Diagenode, Liège, Belgium). Chromatin immunoprecipitation was carried out using the EZ-Magna ChIP A-Chromatin Immunoprecipitation Kit (Millipore, Feltham, UK) as per the manufacturer's instructions. Antibodies used included anti-H3K4me2, anti-H3ac, anti-HP1α (from

15 Millipore, Feltham, UK) anti H3K9me2 and Rabbit control IgG (from Abcam, Cambridge, UK). IPs were carried out at 4°C overnight with either 4µg of specific antibody or with no antibody as control. DNA pull down was quantified by Q-PCR using the following primer pairs: 152 (proximal LP1 promoter) 5' primer, 5'-GGAGTCGTGACCCCTAAAATG-3' [SEQ ID No: 12]; 3' primer, 5'-CTCTGACCTCTGCCCCAGCTC-3' [SEQ ID No: 13]. 366 (distal LP1 promoter) 5' primer, 5'-TTCGGTAAGTGCAAGTGAAG-3' [SEQ ID No: 14]; 3' primer, 5'-GTTATCGGAGGAGCAAACAG-3' [SEQ ID No: 15].

Results

25 Incorporation of interferon-β S/MAR within AAV expression cassettes improves transgene expression from the liver

In initial studies, an 809bp S/MAR element derived from the 5' region of the human IFN-β gene was cloned downstream of the inventors' previously described LP1-hFIXco expression cassette in forward and reverse orientations in the context of

30 single stranded AAV vectors (see Figure 8A, ssAAV-LP1-hFIXco-IFN-F and ssAAV-LP1-hFIXco-IFN-R respectively). Single stranded AAV vectors were used for this initial analysis instead of a self-complementary vector because of its limited packaging capacity of <2.3kb. The control vector consisted of ssAAV-LP1-hFIXco without the S/MAR sequences. These vectors were packaged with serotype 8 capsid

35 and generated using the HEK293T transient transfection method at similar yields ranging from 2-6x10⁴ ssAAV genome particles/293T cell. Assessment of the viral genome extracted from 5x10¹⁰ particles on alkaline agarose gels showed that the

- 35 -

ssAAV-LP1-hFIXco-IFN-R was packaged as a 3.3kb genome in a monomeric form. In contrast, a proportion of the control vector (ssAAV-LP1-hFIXco-Control) was packaged as self-complementary dimer DNA that ran at approximately twice the length of its 2.4kb monomeric genome size (see Figure 8B).

5

Each of the three vectors was injected into the tail vein of 6-8 week old, male, C57Bl/6 immunocompetent mice (N=6) at a dose of 4×10^{12} vector genome (vg)/kg. Following vector administration, transgene expression was assessed by determining plasma hFIX levels at varying time points over a period of 8 weeks. The kinetics of hFIX expression with or without the S/MAR element was similar, with hFIX being detectable at 2 weeks and reaching steady state levels by 4 weeks (see Figure 8C), consistent with previous observations. Highest levels of hFIX expression were observed in mice transduced with ssAAV-LP1-hFIXco-IFN-R with steady state hFIX levels of $188 \pm 18\%$ of normal. This was almost 10-fold higher ($p < 0.001$, One-way ANOVA) than levels observed in cohorts of mice transduced with ssAAV-LP1-hFIXco-Control ($17 \pm 1.2\%$ of normal) and ssAAV-LP1-hFIXco-IFN-F ($20 \pm 4\%$ of normal). This difference in hFIX expression was not accounted for by variation in transduction efficiency as the proviral copy number in liver at 12 weeks following gene transfer was comparable at approximately 0.3 copies/hepatocyte (see Figure 8D).

20

Incorporation of HPRT and ApoB S/MARs in AAV Expression Cassettes Enhances Transgene Expression

To assess if other previously described S/MARs enhance AAV transgene expression, the inventors cloned a 587bp S/MAR element derived from the human HPRT gene and a 486bp S/MAR from the human ApoB gene into the ssAAV-LP1-hFIXco vector plasmid downstream of the hFIXco cDNA in forward and reverse orientation (ssAAV8-LP1-hFIXco-HPRT-F, ssAAV8-LP1-hFIXco-HPRT-R, ssAAV8-LP1-hFIXco-ApoB-F and ssAAV8-LP1-hFIXco-ApoB-R, respectively). To eliminate the possibility that enhanced transgene expression is merely an artifact of cloning sequences 3' of the hFIXco cDNA, a 520bp non-S/MAR control element derived from a region of the Kaposi sarcoma herpes virus (KSHV) genome known to promote autonomous replication (a property attributed to some S/MARs) was cloned into the same position of the vector genome in forward and reverse orientation (ssAAV8-LP1-hFIXco-KSHV-F and ssAAV8-LP1-hFIXco-KSHV-R respectively) (see Figure 9A).

35

- 36 -

Each of the new vectors, pseudotyped with serotype 8 capsid protein, was administered at a dose of 4×10^{12} vg/kg as a single bolus injection into the tail vein of 6-8 week old male C57BL/6 mice (N=3) and plasma hFIX levels were determined at 4 weeks after gene transfer (see Figure 9B). For comparison a group of mice was also transduced with the same dose of ssAAV8-LP1-hFIXco-IFN-R. The highest levels of transgene expression were observed in the ssAAV8-LP1-hFIXco-HPRT-F and ssAAV8-LP1-hFIXco-ApoB-R transduced cohorts with hFIX levels of $1401 \pm 87\%$ and $1540 \pm 149\%$ of normal, respectively. This was over 3-fold higher than observed in the cohort transduced with ssAAV8-LP1-hFIXco-IFN-R ($398 \pm 79\%$) and ssAAV8-LP1-hFIXco-HPRT-R ($403 \pm 75\%$) and approximately 28-fold higher than levels observed with ssAAV8-LP1-hFIXco-Control transduced animals ($52 \pm 3\%$) ($p < 0.0001$, One-way ANOVA). Animals transduced with ssAAV8-LP1-hFIXco-ApoB-F, ssAAV8-LP1-hFIXco-KSHV-F and ssAAV8-LP1-hFIXco-KSHV-R had hFIX levels that were comparable to those observed in the ssAAV8-LP1-hFIXco-Control cohort at $47 \pm 7\%$, $22 \pm 6\%$ and $62 \pm 6\%$, respectively.

After accounting for variation in transgene copy number (see Figure 9C) in the liver at 4 weeks after gene transfer, animals transduced with ssAAV8-LP1-hFIXco-HPRT-F had at least four times higher levels of hFIX/transgene copy when compared to the cohorts transduced with other S/MAR containing AAV vectors including ssAAV8-LP1-hFIXco-ApoB-R (see Figure 1, lower panel).

The LP1 promoter has previously been shown to restrict transcription to hepatocytes. The inventors, therefore, assessed hFIX mRNA levels in the liver at 4 weeks after gene transfer in a selected number of cohorts. After normalizing the data to the GAPDH housekeeping gene transcript levels and proviral DNA copy number, the mRNA levels/copy of the transgene were greatest in the ssAAV8-LP1-hFIXco-HPRT-F cohort at 433 ± 132 RU (see Figure 9D), which was approximately 2-fold higher than mRNA levels observed in the ssAAV8-LP1-hFIXco-HPRT-R (264 ± 57 RU) and ssAAV8-LP1-hFIXco-IFN-R (176 ± 81 RU) cohorts but an over 20-fold increase in transcript levels when compared to the ssAAV8-LP1-hFIXco-Control animals (18 ± 6 RU). This trend in mRNA levels mirrors the hFIX protein expression profile observed in the respective groups of animals suggesting that the increase in hFIX protein observed following gene transfer with S/MAR containing vectors was due, at least in part, to an increase in transcription or mRNA stability.

- 37 -

The ssAAV8-LP1-hFIXco-HPRT-F vector mediated higher levels of transgene expression even at a lower vector dose. At 4 weeks after tail vein administration of a two log lower dose of vector (4×10^{10} vg/kg), hFIX levels of $508 \pm 83\%$ of normal were observed in the ssAAV8-LP1-hFIXco-HPRT-F cohort, which was approximately 80-fold higher than that observed in the ssAAV8-LP1-hFIXco-Control mice ($6 \pm 1\%$, $P=0.0038$, T test, see Figure 10A) even though the transgene copy number in the liver of each cohort was roughly comparable at ~ 0.005 - 0.008 copies/cell (see Figure 10B).

10 Higher transgene expression with ssAAV8-LP1-hFIXco-HPRT-F in non-human primates

The non-human primate model provides an opportunity to validate observation in murine models in a context relevant to humans (Herzog et al., 2011). Peripheral vein administration of 4×10^{12} vg/kg of ssAAV8-LP1-hFIXco-Control or ssAAV8-LP1-hFIXco-HPRT-F AAV vector into male adolescent rhesus monkeys was well-tolerated with no perturbation of vital signs (pulse, respiration and temperature) or liver transaminases (alanine aminotransferase < 45 U/L) over a period of 37 weeks after gene transfer.

20 Human FIX was detectable in monkey plasma within 72 hours of vector administration, reaching peak levels of $53.9 \pm 10\%$ at 5 days in the ssAAV8-LP1-hFIXco-HPRT-F transduced monkey. In contrast the animal transduced with ssAAV8-LP1-hFIXco-Control had peak hFIX levels of 5% (see Figure 11). Similarly, steady state hFIX levels were 4-fold higher in the monkey transduced with ssAAV8-LP1-hFIXco-HPRT-F ($17 \pm 4\%$ of normal) when compared to the animal that received ssAAV8-LP1-hFIXco-Control ($4 \pm 2\%$ of normal). The mean transgene copy number in the liver by Q-PCR was 32 ± 9 and 13 ± 7 proviral copies/cells respectively for the monkeys transduced with ssAAV8-LP1-hFIXco-Control and ssAAV8-LP1-hFIXco-HPRT-F. Thus, single stranded AAV vector containing HPRT-S/MAR mediated higher FIX expression/copy of proviral DNA when compared to control vector in NHPs.

30 Identification of the minimal HPRT sequences required to enhance AAV transgene expression

35 The HPRT S/MAR was divided into smaller fragments as outlined in Figure 12A and then cloned into ssAAV-LP1-hFIXco in order to identify the minimum sequences required for enhancement of AAV transgene expression. A dose of 4×10^{10} vg/kg of

- 38 -

serotype 8 capsid protein pseudotyped ssAAV vector containing truncated HPRT S/MAR fragments was administered via the tail vein into 6-8 week old male C57Bl/6 mice (N=3). Plasma hFIX levels and proviral copy number in the liver were assessed 2 weeks after vector administration. For comparison an equivalent dose of ssAAV8-LP1-hFIXco-HPRT-F vector containing the full 587bp HPRT S/MAR was administered in another cohort. Highest levels of transgene expression were observed in the ssAAV8-LP1-hFIXco-HPRT-F transduced cohort with hFIX levels of $46 \pm 9\%$ of normal (see Figure 12B). hFIX levels observed in the ssAAV8-LP1-hFIXco-HPRT-FR2b cohort were similar at $35 \pm 7\%$ of normal ($P > 0.05$). In comparison, animals transduced with ssAAV8-LP1-hFIXco-HPRT-FR2 ($24 \pm 5\%$), ssAAV8-LP1-hFIXco-HPRT-FR1 ($4 \pm 1\%$) and ssAAV8-LP1-hFIXco-HPRT-FR2a ($3 \pm 2\%$) had significantly lower plasma hFIX levels despite similar levels (between ~ 0.3 - 1.4 copies/hepatocyte) of gene transfer as assessed by proviral copy number in the liver. These data suggests that the 130bp HPRT fragment 2b may contain the core elements required for enhancing the AAV transgene expression in-vivo.

This fragment was, therefore, cloned downstream of the codon optimized FIX cDNA in our scAAV-LP1-hFIXco expression cassette (scAAV8-LP1-hFIXco-HPRT-FR2b), containing truncated SV40 polyadenylation sequences. AAV8 capsid pseudotyped vector was next administered into the tail vein of 6 to 8 week old male C57BL/6 mice at a dose of 4×10^{10} vg/kg and plasma hFIX levels and proviral copy number in the liver were assessed at 4 weeks after gene transfer. Mice transduced with scAAV8-LP1-hFIXco-HPRT-FR2b expressed hFIX at approximately $105 \pm 6\%$ (see Figure 13), which was 35-fold higher ($P = 0.0038$, student T-test) than that observed in scAAV8-LP1-hFIXco transduced animals ($3 \pm 1\%$) suggesting that S/MAR elements are also effective in a self-complementary format. The proviral copy number in the liver of transduced animals was comparable in the two groups at between 2.1-3.1 copies/hepatocyte.

S/MARs confer higher transgene expression through epigenetic modification of the ssAAV-LP1-hFIXco genome

Bisulfite sequencing of the LP1 promoter, which contains several small CpG rich regions, did not show any CpG methylation in proviral DNA extracted from liver tissue; even at 1 year after gene transfer (see Figure 15). Next, the inventors evaluated *in-vivo* chromatinization of ssAAV-LP1-hFIXco-Control in comparison to ssAAV-LP1-hFIXco-IFN-R or ssAAV-LP1-hFIXco-HPRT-F genomes extracted from murine liver at 4-8 weeks after tail vein administration of 4×10^{12} vg/kg using

- 39 -

chromatin immunoprecipitation (ChIP) experiments. When compared to the ssAAV-LP1-hFIXco-IFN-R cohort, mice transduced with ssAAV-LP1-hFIXco-Control consistently showed an enrichment of approximately 2 to 4-fold of the heterochromatin markers (dimethylation of lysine 9 on histone 3 [H3K9me2] and heterochromatic adaptor proteins HP1 α) in the LP1 promoter region (see Figure 14A). Both H3K9me2 and HP1 α marks are associated with inactive promoters (Lienert et al., 2011; Wreggett et al., 1994). Additionally, there was an increase in association of histone H3 with dimethylation on lysine 4 (H3K4me2) as well as acetylation of H3 lysines (H3Ac) with the distal region of the LP1 promoter of ssAAV-LP1-hFIXco-IFN-R when compared to the LP1 promoter of ssAAV-LP1-hFIXco-Control. Both H3K4me2 and H3Ac marks correlate with transcriptionally active genes (Yan et al., 2006). ChIP analysis of ssAAV-LP1-hFIXco-IFN-F showed the same in-vivo chromatinization pattern as the control vector (data not shown).

In comparison to the ssAAV-LP1-hFIXco-HPRT-F cohort, ssAAV-LP1-hFIXco-Control genomes again showed an approximate 2 to 3-fold enrichment of the H3K9me2 heterochromatin marker (see Figure 14B). Further to this, a 2 to 4-fold enrichment of H3K4me2 and 2 to 6-fold enrichment of H3Ac was observed in association with ssAAV-LP1-hFIXco-HPRT-F genomes in both promoter regions under analysis. Collectively, these data suggest that the higher levels of transgene expression observed with ssAAV-LP1-hFIXco-IFN-R and ssAAV-LP1-hFIXco-HPRT-F was due to the maintenance of these expression cassettes in an open chromatin status.

Discussion

Clinical proof of concept of AAV gene therapy has been realized in several diseases including hemophilia B and Leber's congenital amaurosis using relatively high vector doses. In this study, the inventors show that the potency of a AAV-LP1-hFIXco expression cassette that was recently used in the hemophilia B clinical trial can be significantly improved by incorporating S/MAR elements. In common with episomal plasmids and onco-retroviral vector systems, the data show that S/MARs surprisingly enhance AAV mediated expression of FIX in mice and nonhuman primates through an increase in transgene mRNA levels.

Surprisingly, the orientation of the S/MAR sequences in AAV vectors appears to be important for all three elements examined, a feature that has also been observed in the context of retroviral and lentiviral vectors. Furthermore, the AAV-LP1-hFIXco

- 40 -

provirus does not undergo DNA methylation in the liver of mice following systemic administration. However, LP1 promoter sequences within the proviral genome become associated with histones containing repressive marks. The inclusion of S/MARs appears to decrease the level of proviral DNA heterchromatinization.

5

The data suggest that S/MARs improve AAV transgene expression by maintaining AAV concatamers in an open chromatin, transcriptionally active, configuration. However, in addition to chromatin remodeling function, S/MARs have other biological properties that could potentially influence AAV transgene expression. For instance, some S/MARs have enhancer properties capable of upregulating the expression of heterologous reporters (Stief et al., 1989). Additionally, they have the ability to bring the AAV proviral DNA in close proximity to key nuclear regulatory elements within the nucleus by forming “anchor points” with nuclear matrix.

10

Alternatively S/MARs may provide significant enhancement of transcription by providing topological separation of individual expression cassettes within large AAV concatamers, thus reducing read-through transcription and/or “insulating” the proviral DNA from the inhibitory influences of other host genes. Previous studies show that some or all of these mechanisms are in play in the context of gene transfer. The human IFN β gene S/MAR element has been shown to prevent epigenetic silencing of episomally maintained plasmid vectors as well as retroviral vectors. Similarly the ApoB S/MAR used in this study was reported to insulate associated genes from positional silencing effects. In contrast, the HPRT S/MAR is primarily documented to have a role in supporting autonomous replication. As such, the observation that highest enhancement of AAV transgene expression was achieved with the HPRT vector was totally unexpected.

15

20

25

S/MARs do not have an obvious consensus sequence although prototypic elements are AT-rich. The structure of S/MAR in relation to the transgene appears to have a greater influence on activity than the overall nucleotide base composition. This may explain there may be an orientation-dependent effect in the context of rAAV mediated gene transfer. Deletion analysis of the HPRT S/MAR showed that a 130bp region (Fragment 2b) was able to enhance FIX expression in the context of self-complementary vectors. The only unique nucleotide consensus sequence found in HPRT fragment 2b was TYRTTT, which occurs twice in this sequence and may represent a core enhancer. Alternatively, the potency of fragment 2b may be purely structural due to the large quantity of poly-pyrimidine and poly-purine tracts present in the sequence. These enable the formation of secondary non-B DNA structures,

30

35

- 41 -

which facilitate base-unpairing and strand-separation, thereby catalyzing transcription.

Several agents including androgens and proteasome inhibitors have been found to
5 enhance rAAV mediated gene transfer. These drugs have to be administered prior to
or at the time of gene transfer, thus making scheduling of vector administration
more complex. In addition, they have significant side-effects which could adversely
influence the toxicity profile of rAAV vector. In contrast, S/MAR containing rAAV
vectors provides a novel, safe, inbuilt molecular strategy for augmenting rAAV
10 transgene expression by overcoming some of the limitation of AAV biology.
Therefore, the incorporation of S/MARs into the rAAV expression cassettes
represents an important step in improving the potency and safety profile of AAV
vectors, a prerequisite to the wider clinical use of this gene transfer platform, which
already shows significant promise.

15

Summary

Adeno-associated virus vector encoding codon optimized human factor IX (AAV-
LP1-hFIXco) shows great promise in severe hemophilia B patients. However,
hepatocellular toxicity was observed in some patients treated at the high dose levels.
20 In the inventor's desire to improve vector potency, various scaffold/matrix
attachment regions (S/MARs) were cloned at the 3' end of a modified single-
stranded (ss) AAV-LP1-hFIXco expression cassette. Amongst these ssAAV-LP1-
hFIXco-HPRT-F, which contained the hypoxanthine-guanine phosphoribosyl-
transferase S/MAR in the forward orientation, enhanced FIX expression by 28-fold
25 in mice. In monkeys, AAV-LP1-hFIXco-HPRT-F mediated 4-fold higher steady-state
plasma FIX levels than achieved with an equivalent dose of ssAAV-LP1-hFIXco.
When cloned a into self-complementary AAV-LP1-hFIXco expression cassette, a
130bp region (fragment-2b) within the HPRT S/MAR was sufficient to enhance FIX
expression in mice by 35-fold over that observed with an identical vector without
30 S/MAR. Mechanistic studies showed that HPRT S/MAR enhanced AAV transgene
expression by reducing heterochromatin marks (H3K9me2) in the promoter region,
resulting in an increase in FIX mRNA levels by 20-fold. S/MARs therefore provide a
novel inbuilt process for enhancing AAV transgene expression through epigenetic
modification of the viral genome; thus improving the prospects of therapeutic gene
35 transfer in humans using lower; potentially safer doses of AAV vector. These data
suggest that S/MARs can enhance AAV mediated transgene expression in murine

and non-human primate models by preventing heterchromatinization of the proviral DNA.

References

- 5 1. Dang Q, Auten J, Plavec I. Human beta interferon scaffold attachment region inhibits de novo methylation and confers long-term, copy number-dependent expression to a retroviral vector. *J Virol* 2000;74(6):2671-2678.
- 10 2. Kurre P, Morris J, Thomasson B, Kohn DB, Kiem HP. Scaffold attachment region-containing retrovirus vectors improve long-term proviral expression after transplantation of GFP-modified CD34+ baboon repopulating cells. *Blood* 2003;102(9):3117-3119.
- 15 3. Namciu SJ, Blochlinger KB, Fournier RE. Human matrix attachment regions insulate transgene expression from chromosomal position effects in *Drosophila melanogaster*. *Mol Cell Biol* 1998;18(4):2382-2391.
4. Namciu SJ, Fournier RE. Human matrix attachment regions are necessary for the establishment but not the maintenance of transgene insulation in *Drosophila melanogaster*. *Mol Cell Biol* 2004;24(23):10236-10245.
- 20 5. Sykes RC, Lin D, Hwang SJ, Framson PE, Chinault AC. Yeast ARS function and nuclear matrix association coincide in a short sequence from the human HPRT locus. *Mol Gen Genet* 1988;212(2):301-309.
6. Koina E, Piper A. An inactive X specific replication origin associated with a matrix attachment region in the human X linked HPRT gene. *J Cell Biochem* 2005;95(2):391-402.
- 25 7. Debatisse M, Toledo F, Anglana M. Replication initiation in mammalian cells: changing preferences. *Cell Cycle* 2004;3(1):19-21.
8. Nathwani AC, Gray JT, Ng CY, et al. Self-complementary adeno-associated virus vectors containing a novel liver-specific human factor IX expression cassette enable highly efficient transduction of murine and nonhuman primate liver. *Blood*. 2006;107(7):2653-2661.
- 30 9. Hart RP, McDevitt MA, Ali H, Nevins JR. Definition of essential sequences and functional equivalence of elements downstream of the adenovirus E2A and the early simian virus 40 polyadenylation sites. *Mol Cell Biol*. 1985;5(11):2975-2983.
- 35 10. Namciu SJ, Blochlinger KB, Fournier RE. Human matrix attachment regions insulate transgene expression from chromosomal position effects in *Drosophila melanogaster*. *Mol Cell Biol*. 1998;18(4):2382-2391.
- 40 11. Stedman W, Kang H, Lin S, Kissil JL, Bartolomei MS, Lieberman PM. Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. *EMBO J*. 2008;27(4):654-666.
- 45

- 43 -

12. Nathwani AC, Davidoff A, Hanawa H, Zhou JF, Vanin EF, Nienhuis AW. Factors influencing in vivo transduction by recombinant adeno-associated viral vectors expressing the human factor IX cDNA. *Blood*. 2001;97(5):1258-1265.
- 5 13. Davidoff AM, Ng CY, Sleep S, et al. Purification of recombinant adeno-associated virus type 8 vectors by ion exchange chromatography generates clinical grade vector stock. *J Virol Methods*. 2004;121(2):209-215.
- 10 14. Fagone P, Wright JF, Nathwani AC, Nienhuis AW, Davidoff AM, Gray JT. Systemic Errors in Quantitative Polymerase Chain Reaction Titration of Self-Complementary Adeno-Associated Viral Vectors and Improved Alternative Methods. *Hum Gene Ther*. 2011.
- 15 15. Nathwani AC, Gray JT, McIntosh J, et al. Safe and efficient transduction of the liver after peripheral vein infusion of self-complementary AAV vector results in stable therapeutic expression of human FIX in nonhuman primates. *Blood*. 2007;109(4):1414-1421.
- 20 16. Nathwani AC, Davidoff AM, Hanawa H, et al. Sustained high-level expression of human factor IX (hFIX) after liver-targeted delivery of recombinant adeno-associated virus encoding the hFIX gene in rhesus macaques. *Blood*. 2002;100(5):1662-1669.
- 25 17. Herzog RW, Davidoff AM, Markusic DM, Nathwani AC. AAV vector biology in primates: finding the missing link? *Mol Ther*. 2011;19(11):1923-1924.
18. Lienert F, Mohn F discriminate pluripotent from terminally differentiated cells. *PLoS Genet*. 2011;7(6):e1002090.
- 30 19. Wreggett KA, Hill F, James PS, Hutchings A, Butcher GW, Singh PB. A Mammalian Homolog of Drosophila Heterochromatin Protein-1 (Hp1) Is a Component of Constitutive Heterochromatin. *Cytogenetics and Cell Genetics*. 1994;66(2):99-103.
- 35 20. Yan C, Boyd DD. Histone H3 acetylation and H3 K4 methylation define distinct chromatin regions permissive for transgene expression. *Mol Cell Biol*. 2006;26(17):6357-6371.

- 44 -

Claims

1. A Parvoviridae expression cassette comprising a transgene and a matrix attachment region (MAR) sequence or a fragment thereof from a hypoxanthine phosphoribosyltransferase (HPRT) gene.
2. An expression cassette according to claim 1, wherein the cassette comprises a promoter which initiates transgene expression upon introduction in a host cell.
3. An expression cassette according to claim 2, wherein the MAR sequence or fragment thereof is provided at the 5' end of a promoter and/or within an intron of the transgene.
4. An expression cassette according to either claim 1 or claim 2, wherein the MAR sequence or fragment thereof is provided 3' of the transgene.
5. An expression cassette according to any preceding claim, wherein the MAR sequence or fragment thereof comprises a nucleic acid sequence substantially as set out in SEQ ID No.1, or a functional fragment or variant thereof.
6. An expression cassette according to any preceding claim, wherein the MAR sequence or fragment thereof comprises nucleotides 205 to 450 of SEQ ID No.1.
7. An expression cassette according to any preceding claim, wherein the MAR sequence or fragment thereof comprises nucleotides 320 to 450 of SEQ ID No.1.
8. An expression cassette according to any preceding claim, wherein the cassette comprises DNA from a member of the Dependovirus genus.
9. An expression cassette according to any preceding claim, wherein the cassette comprises DNA from an adeno-associated virus (AAV) vector.
10. An expression cassette according to any preceding claim, wherein the cassette comprises a nucleic acid sequence substantially as set out in SEQ ID No.2, or a functional fragment or variant thereof.

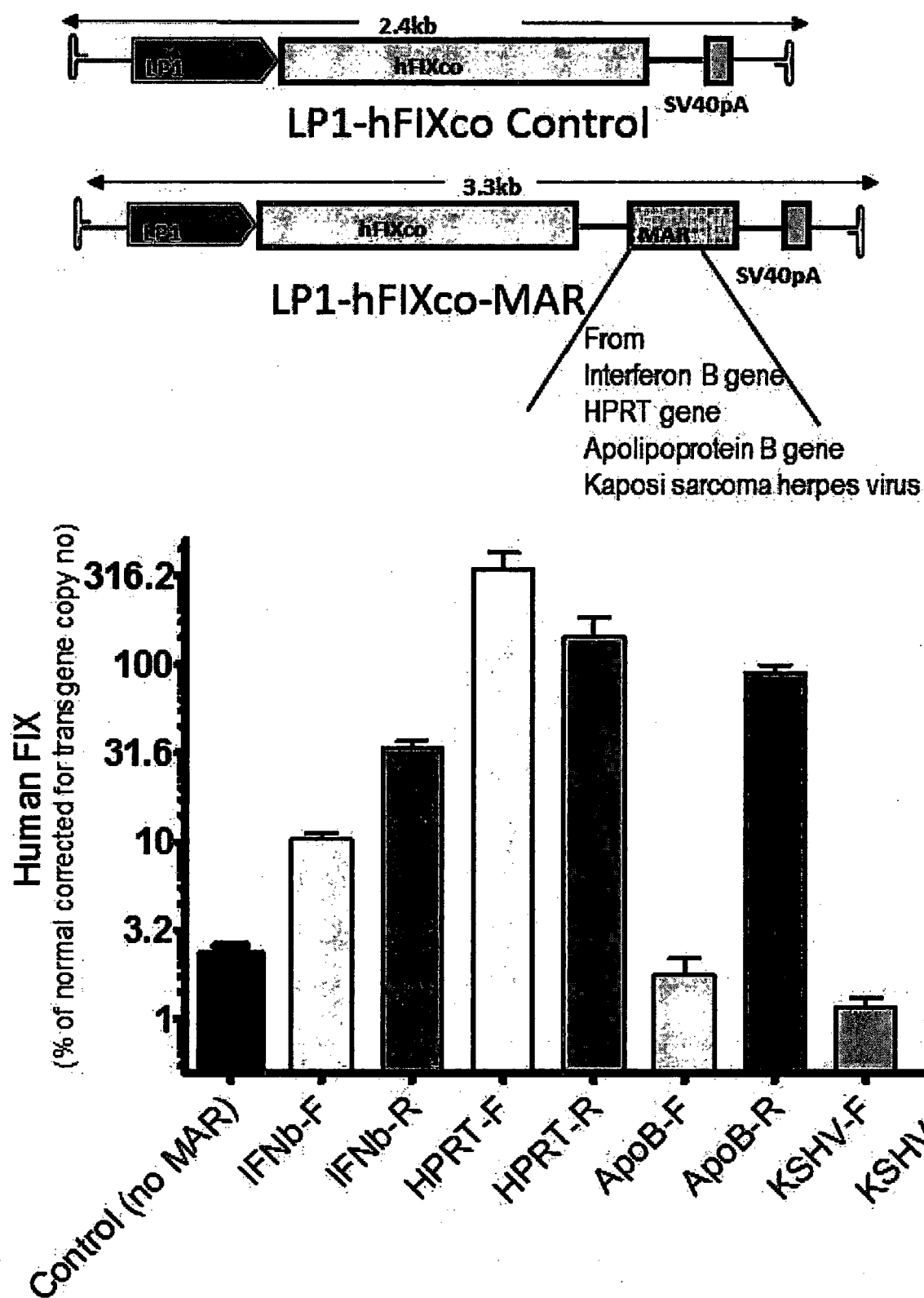
- 45 -

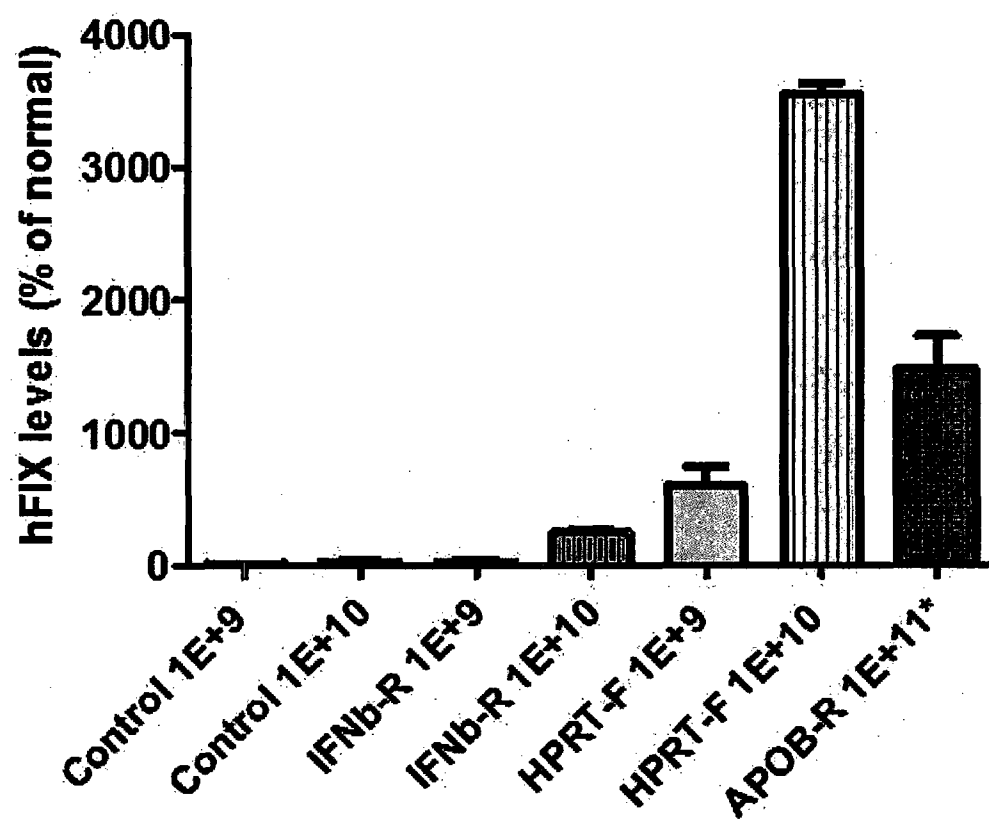
11. An expression cassette according to any preceding claim, wherein the cassette comprises a nucleic acid sequence substantially as set out in either SEQ ID No.3 or SEQ ID No.4, or a functional fragment or variant thereof.
- 5 12. An expression cassette according to any preceding claim, wherein the transgene encodes cystic fibrosis transmembrane conductance regulator (CFTR), a blood clotting factor, such as Factor VIII, Factor VII, Factor IX, or alpha-galactosidase A or glucocerebrosidase.
- 10 13. A recombinant vector comprising an expression cassette according to any one of claims 1-12.
14. An expression cassette according to any one of claims 1-12, or a vector according to claim 13, for use in therapy.
- 15 15. An expression cassette according to any one of claims 1-12, or a vector according to claim 13, for treating cystic fibrosis, a haematological disorder, a blood clotting disorder, haemophilia A, haemophilia B, congenital FVII and FX deficiency, a urea cycle disorder, a lysosomal disorder, a disorder affecting the liver, any form of malignant disorder or Leber's congenital amaurosis.
- 20 16. An expression cassette or a vector according to claim 15, for treating haemophilia B.
- 25 17. A host cell comprising an expression cassette according to any one of claims 1-12, or a vector according to claim 13.
18. A host cell according to claim 17, wherein the host cell is a bacterial cell, or an animal cell.
- 30 19. A transgenic host organism comprising at least one host cell according to either claim 17 or claim 18.
20. A pharmaceutical composition comprising an expression cassette according to any one of claims 1-12, or a vector according to claim 13, and a pharmaceutically acceptable vehicle.
- 35

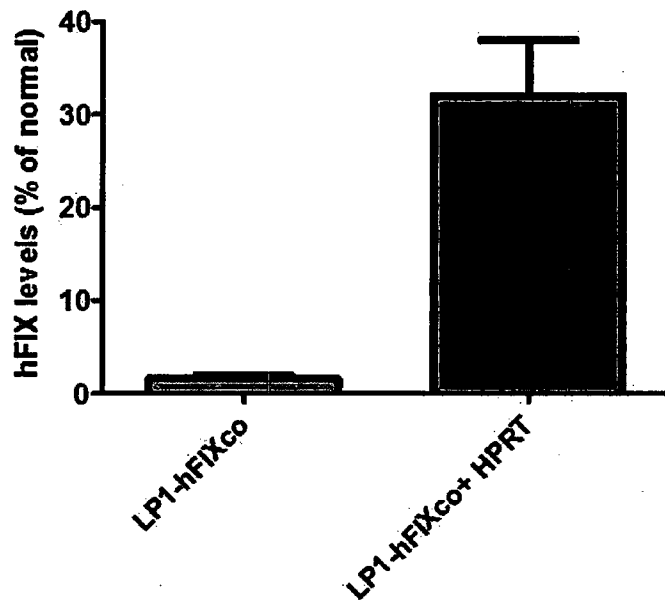
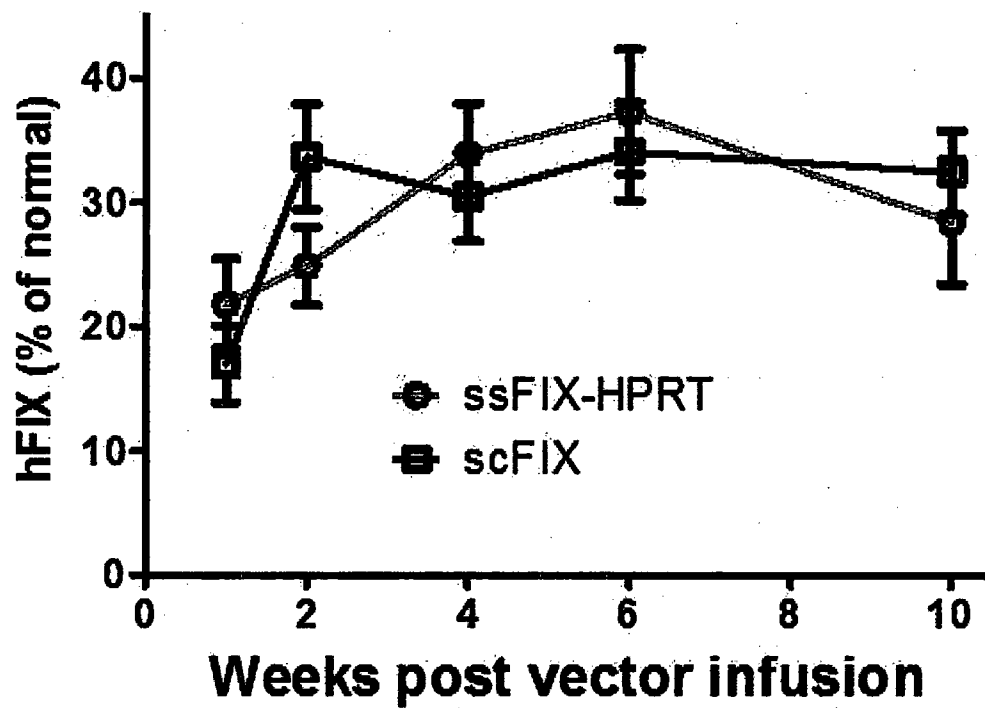
- 46 -

21. A composition according to claim 20, wherein the composition is for treating cystic fibrosis, a haematological disorder, a blood clotting disorder, haemophilia A, haemophilia B, congenital FVII and FX deficiency, a urea cycle disorder, a lysosomal disorder, a disorder affecting the liver, any form of malignant disorder or Leber's
5 congenital amaurosis.
22. A process for making the composition according to either claim 20 or claim 21, the process comprising combining a therapeutically effective amount of an expression cassette according to any one of claims 1-12, or a vector according to
10 claim 13, and a pharmaceutically acceptable vehicle.

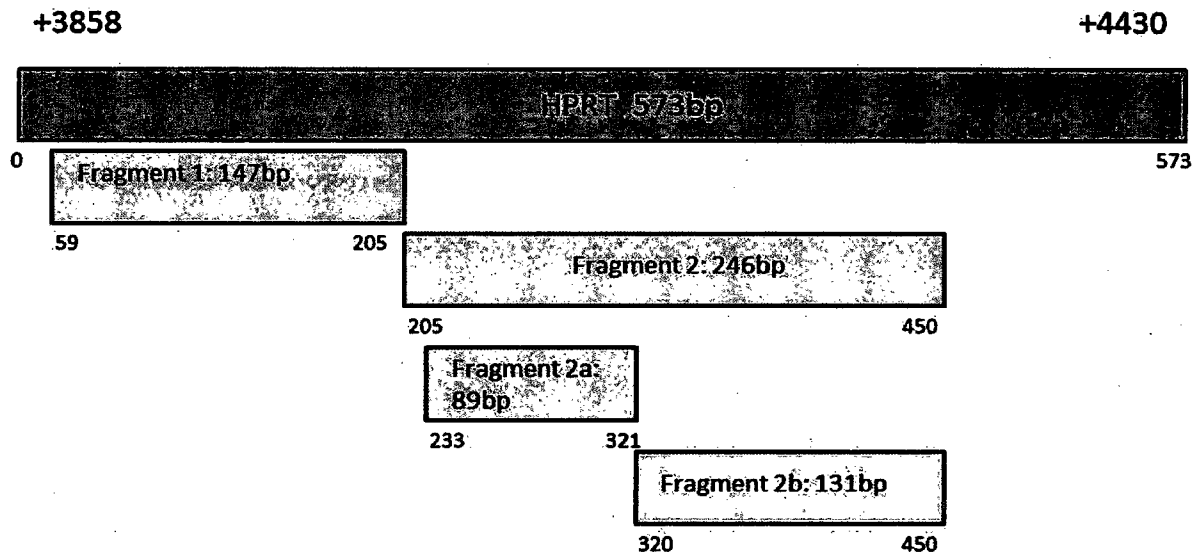
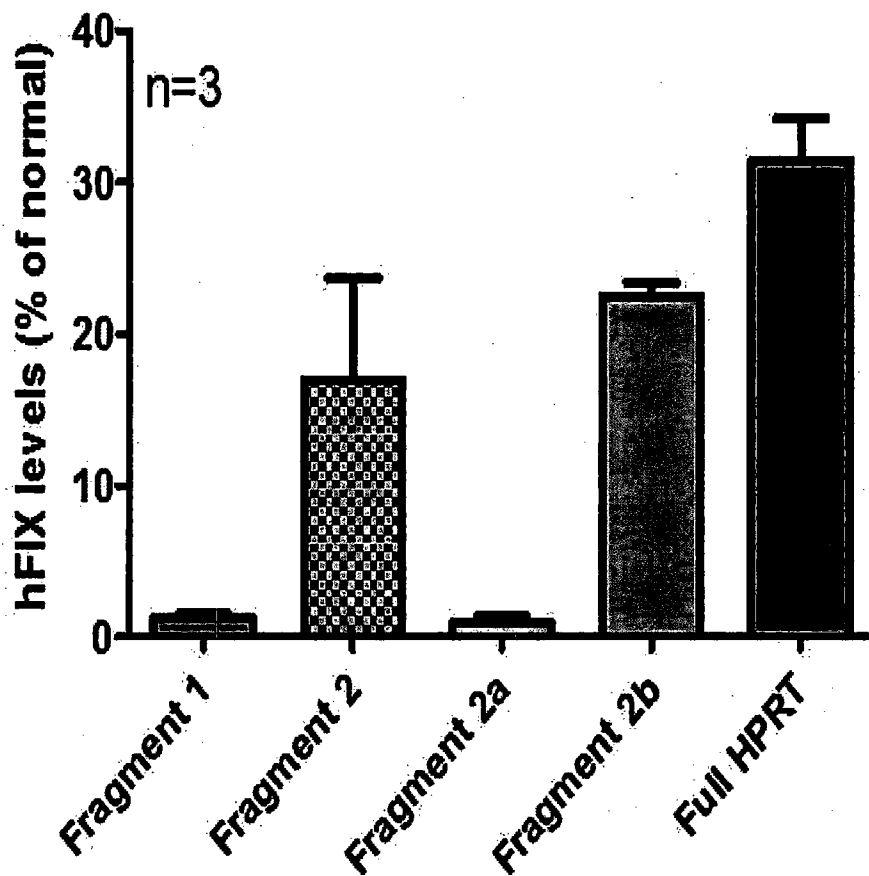
1/15

Figure 1

2/15**Figure 2**

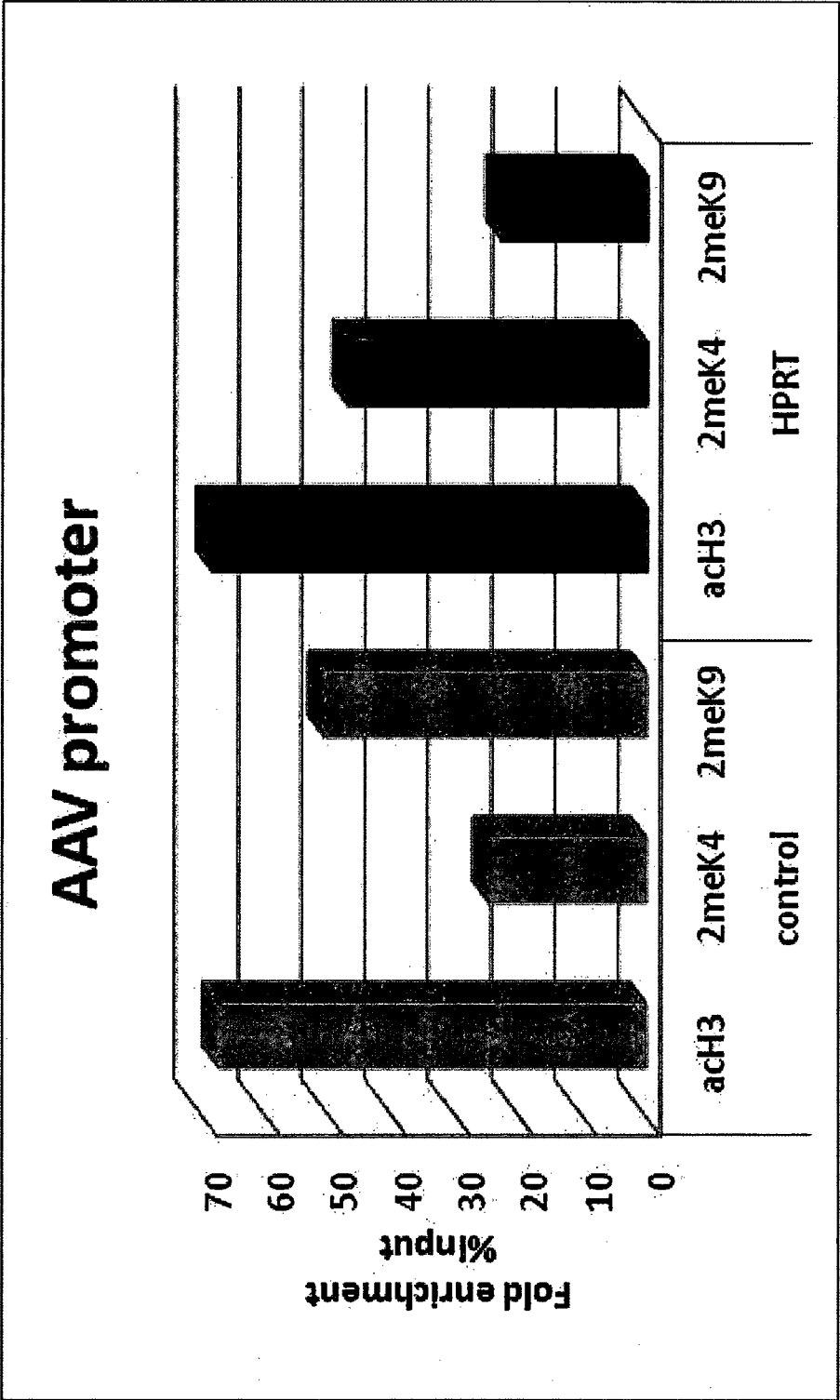
3/15**Figure 3****Figure 4**

4/15

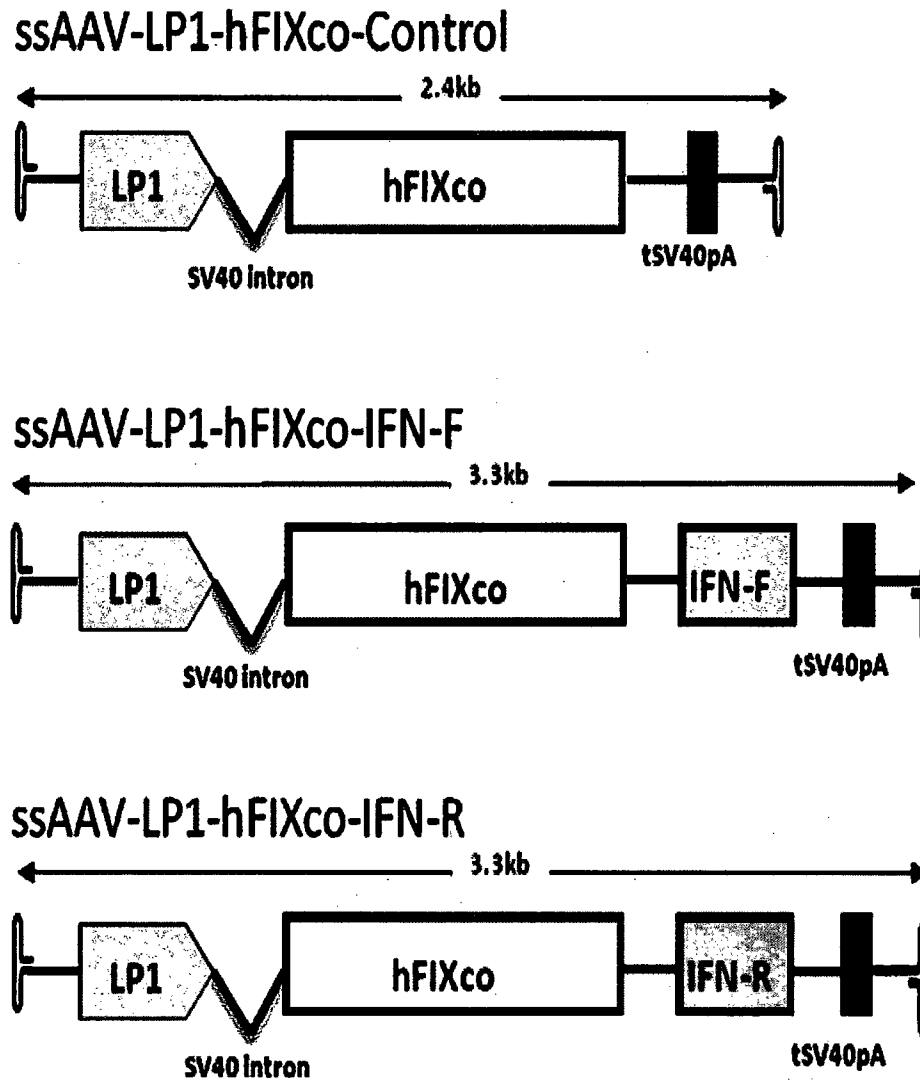
Figure 5**Figure 6**

5/15

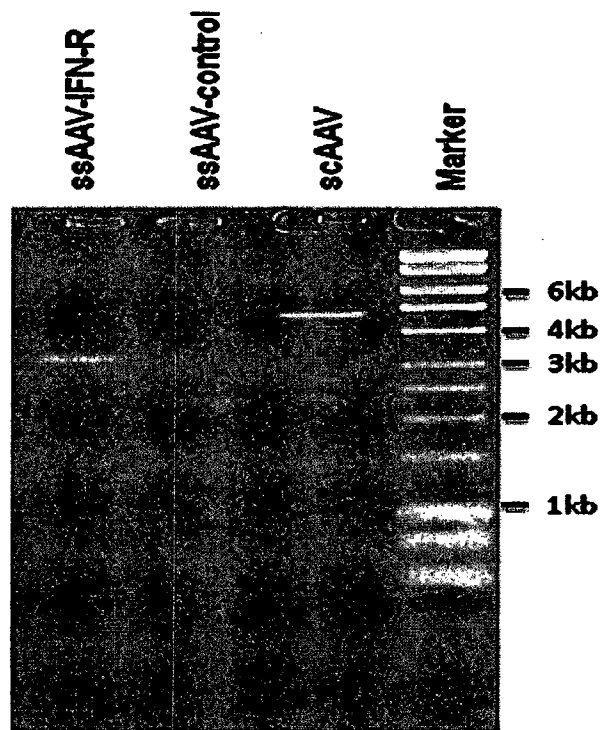
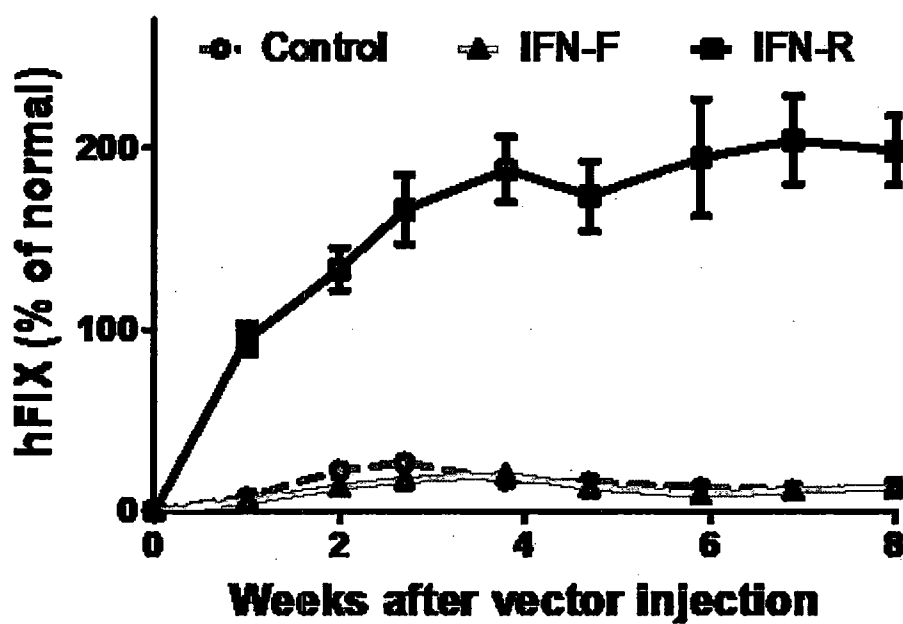
Figure 7



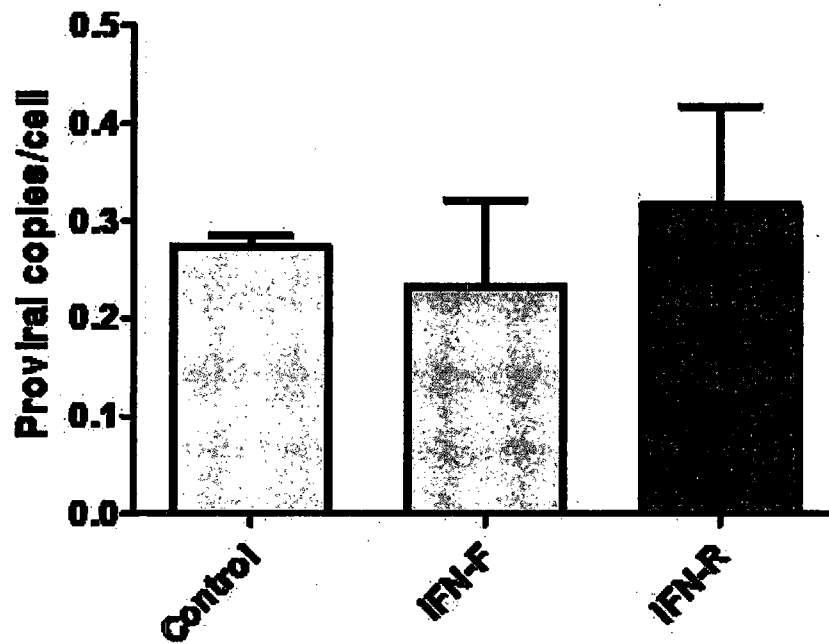
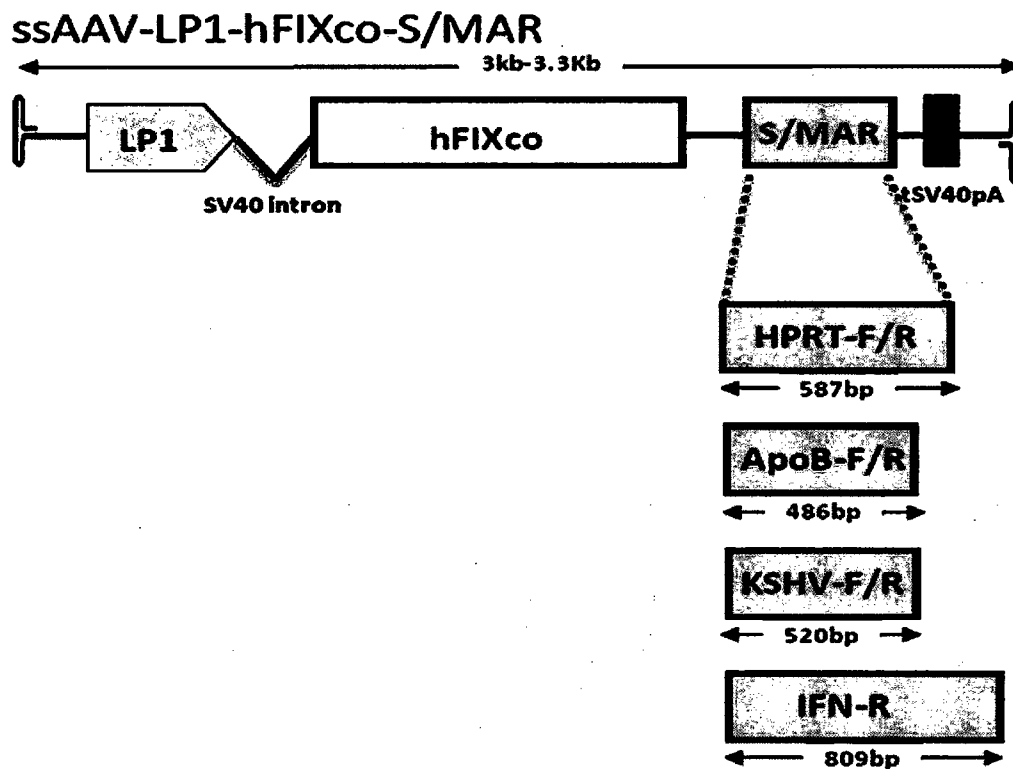
6/15

Figure 8A

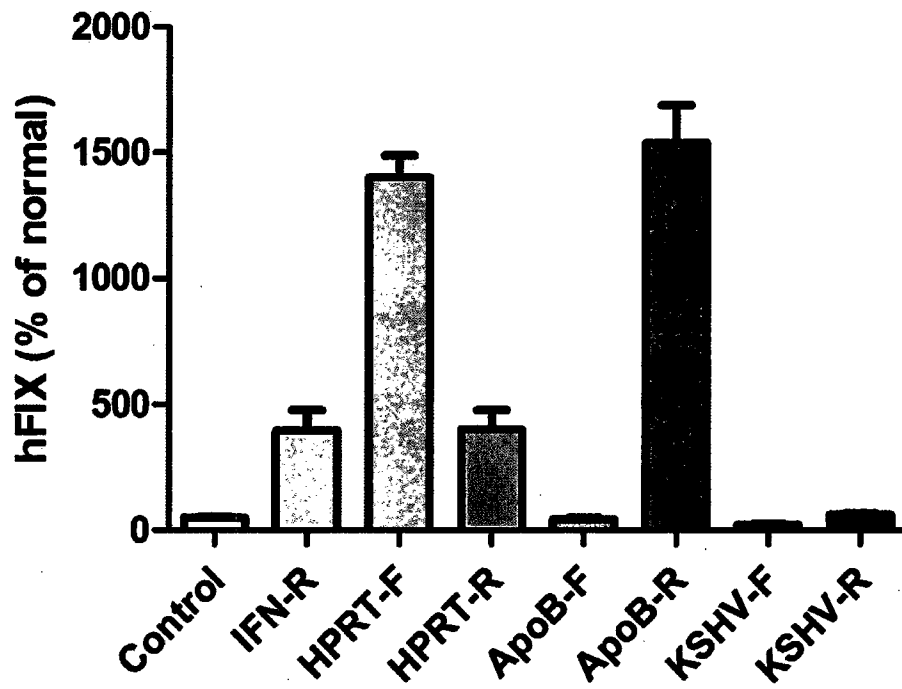
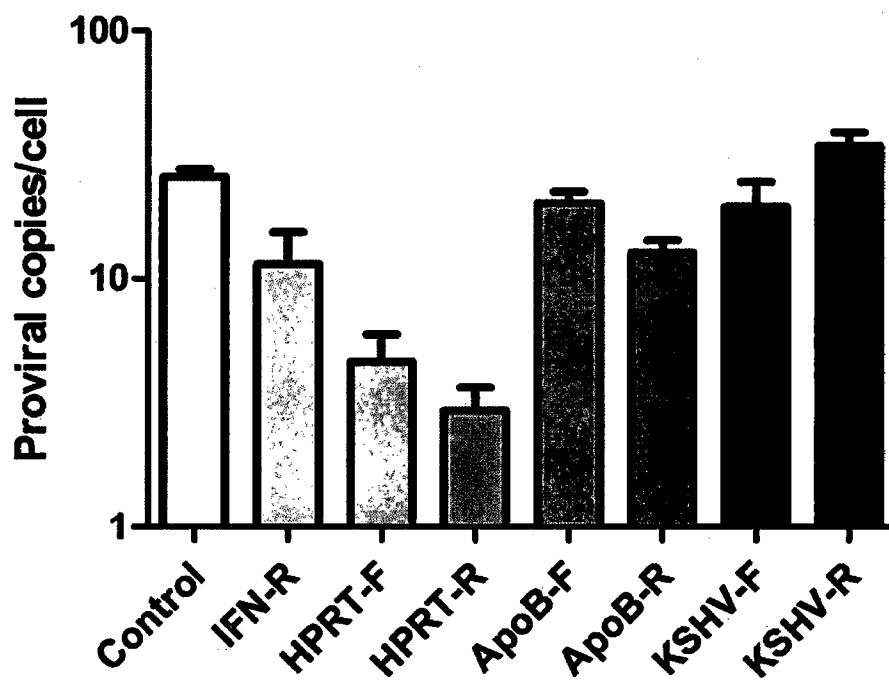
7/15

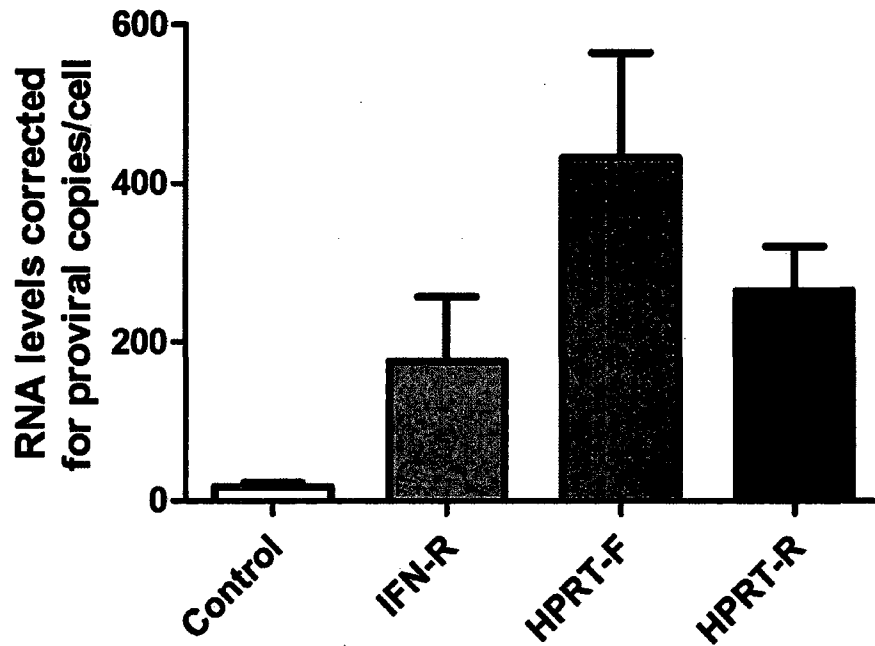
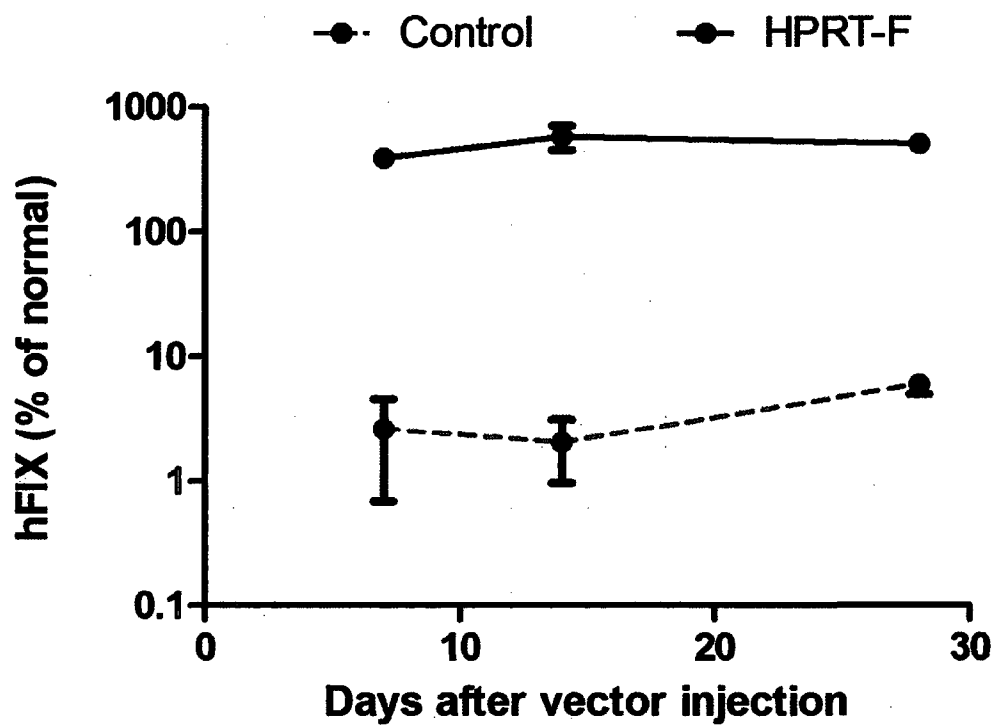
Figure 8B**Figure 8C**

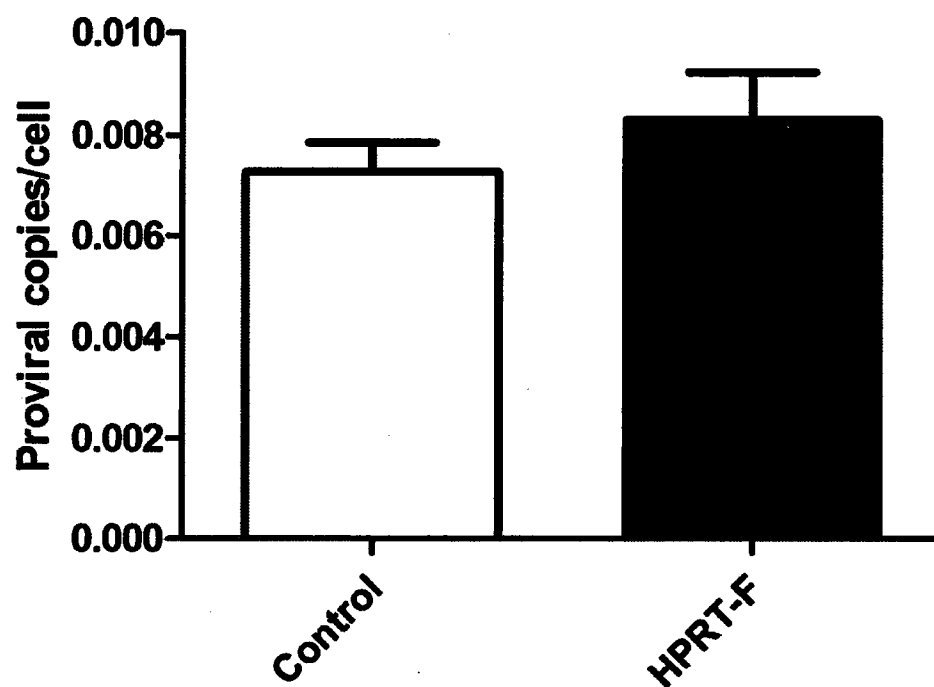
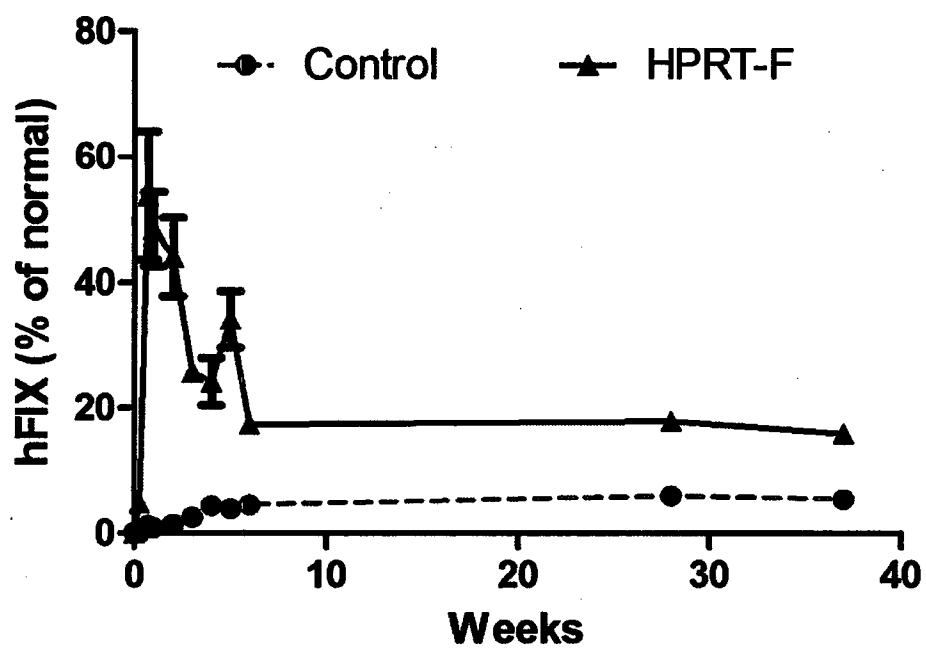
8/15

Figure 8D**Figure 9A**

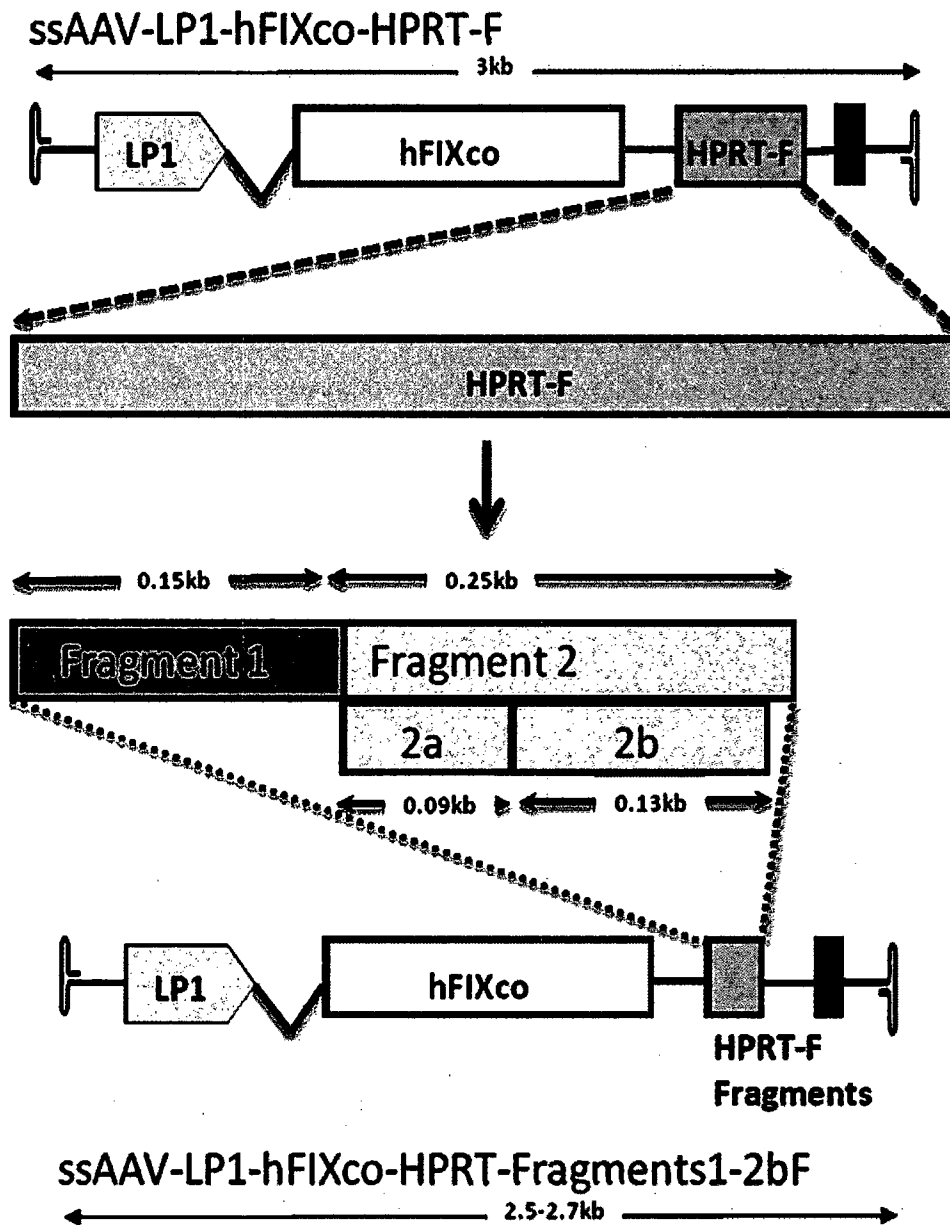
9/15

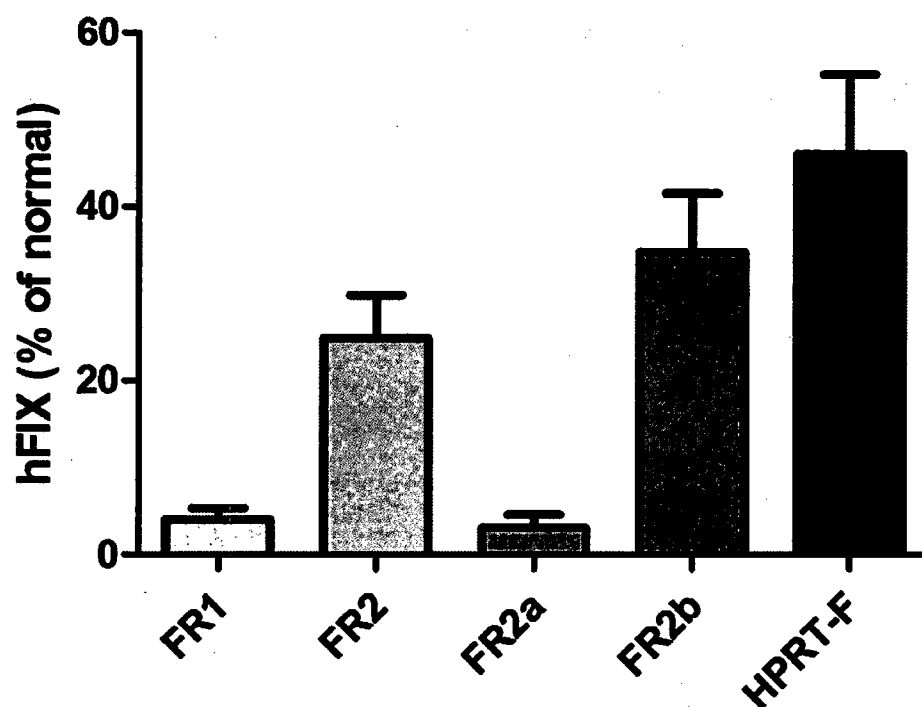
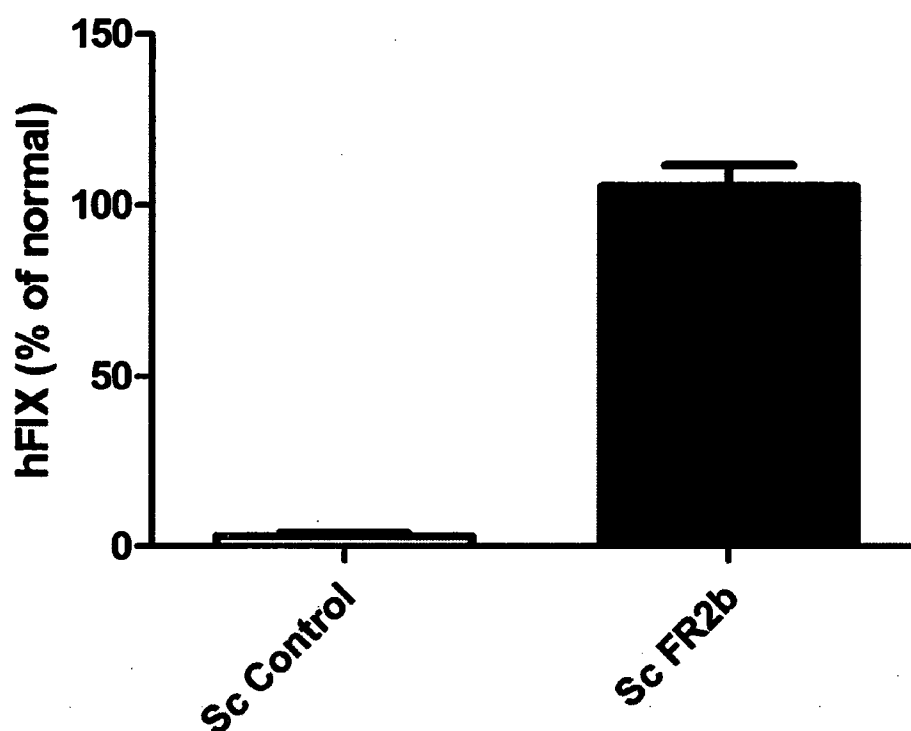
Figure 9B**Figure 9C****Figure 9D**

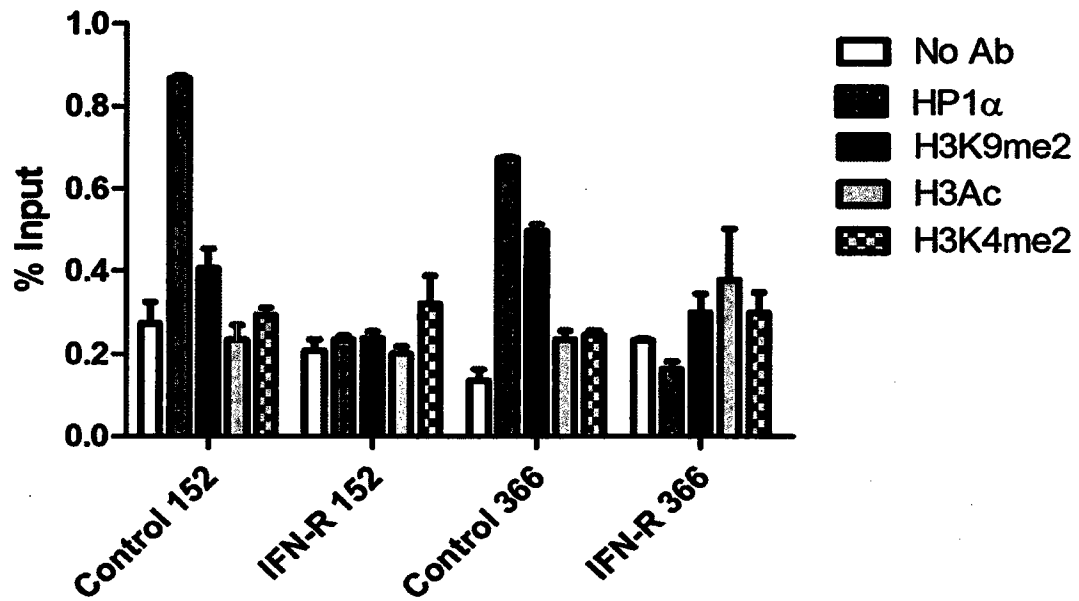
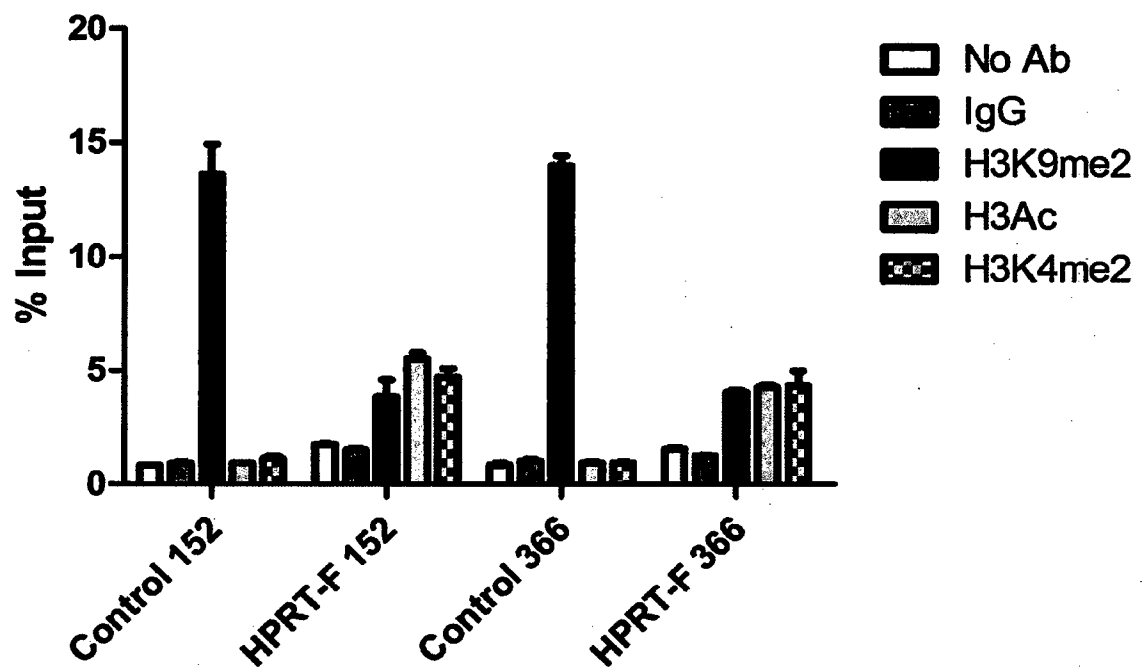
10/15Figure 10AFigure 10B

11/15Figure 11

12/15

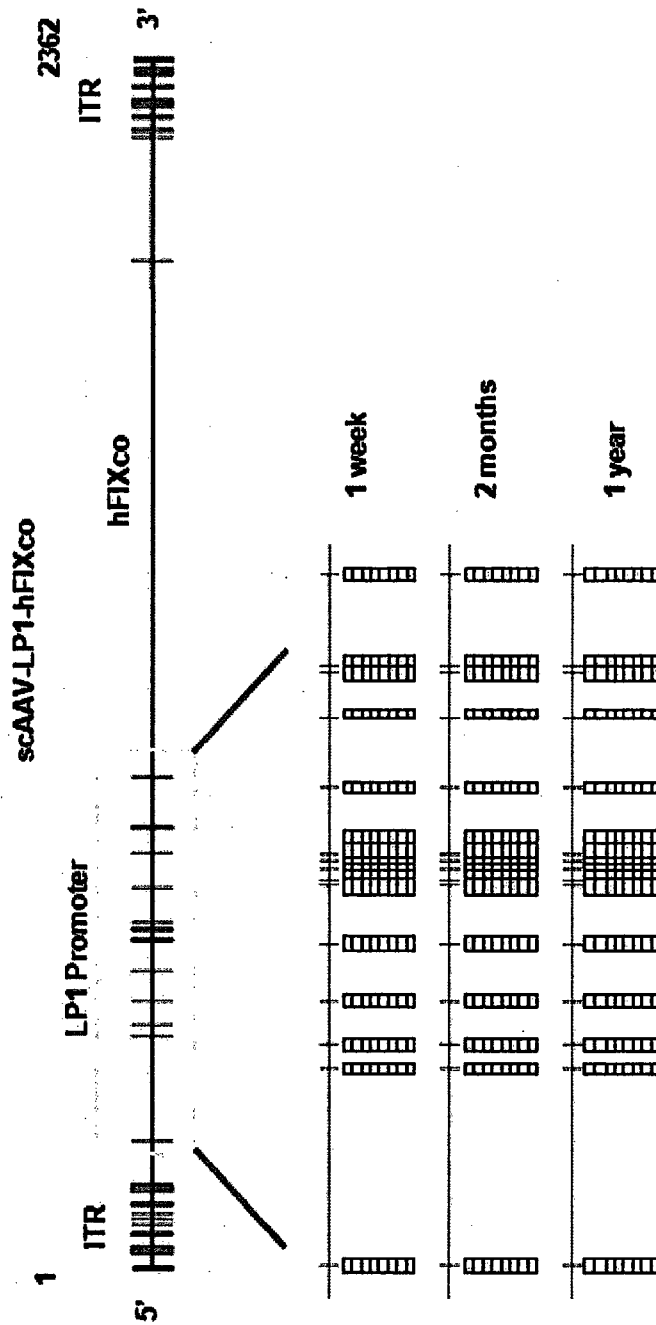
Figure 12A

13/15Figure 12BFigure 13

14/15**Figure 14A****Figure 14B**

15/15

Figure 15



INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2013/051960

A. CLASSIFICATION OF SUBJECT MATTER

INV. C12N7/00 C12N15/09 C12N15/864 C12N5/10 A61K48/00
A01K67/027 C12N15/86

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)

C12N A61K

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, Sequence Search, 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.
A	MAYUYO MORI-UCHINO ET AL: "Enhanced transgene expression in the mouse skeletal muscle infected by the adeno-associated viral vector with the human elongation factor 1[alpha] promoter and a human chromatin insulator", THE JOURNAL OF GENE MEDICINE, vol. 11, no. 7, 1 July 2009 (2009-07-01), pages 598-604, XP055079593, ISSN: 1099-498X, DOI: 10.1002/jgm.1337 the whole document ----- -/--	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

20 September 2013

Date of mailing of the international search report

01/10/2013

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

Brenz Verca, Stefano

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2013/051960

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	NATHWANI AMIT C ET AL: "Self-complementary adeno-associated virus vectors containing a novel liver-specific human factor IX expression cassette enable highly efficient transduction of murine and nonhuman primate liver", BLOOD, AMERICAN SOCIETY OF HEMATOLOGY, US, vol. 107, no. 7, 1 April 2006 (2006-04-01), pages 2653-2661, XP002534783, ISSN: 0006-4971, DOI: 10.1182/BLOOD-2005-10-4035 [retrieved on 2005-12-01] cited in the application the whole document -----	1-22
A	WO 2011/005968 A1 (UCL BUSINESS PLC [GB]; THROMBOSIS RES INST [GB]; ST JUDE CHILDRENS RES) 13 January 2011 (2011-01-13) the whole document -----	1-22
A	L. H. REID ET AL: "Regulatory elements in the introns of the human HPRT gene are necessary for its expression in embryonic stem cells.", PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, vol. 87, no. 11, 1 June 1990 (1990-06-01), pages 4299-4303, XP055079882, ISSN: 0027-8424, DOI: 10.1073/pnas.87.11.4299 the whole document -----	1-22
A	WO 2006/066066 A2 (UNIV NORTH CAROLINA [US]; UNIV FLORIDA [US]; BOWLES DAWN E [US]; LI CH) 22 June 2006 (2006-06-22) the whole document -----	1-22
A	WO 2010/018444 A2 (AVESTHAGEN LTD [IN]; PATELL VILLOO MORAWALA [IN]; MAITYA SUNIT [IN]; V) 18 February 2010 (2010-02-18) the whole document -----	1-22

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/GB2013/051960

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2011005968	A1	13-01-2011	CA 2804602 A1 13-01-2011
		EP 2451474 A1 16-05-2012	
		US 2013024960 A1 24-01-2013	
		WO 2011005968 A1 13-01-2011	

WO 2006066066	A2	22-06-2006	AU 2005316476 A1 22-06-2006
		CA 2591544 A1 22-06-2006	
		CN 101124328 A 13-02-2008	
		EP 1828390 A2 05-09-2007	
		ES 2385837 T3 01-08-2012	
		JP 5136766 B2 06-02-2013	
		JP 2008523813 A 10-07-2008	
		NZ 555830 A 31-01-2009	
		US 2008269149 A1 30-10-2008	
		US 2011104119 A1 05-05-2011	
		WO 2006066066 A2 22-06-2006	

WO 2010018444	A2	18-02-2010	AU 2009280913 A1 18-02-2010
		CA 2736580 A1 18-02-2010	
		CN 102177240 A 07-09-2011	
		EP 2331693 A2 15-06-2011	
		JP 2012506694 A 22-03-2012	
		KR 20110044769 A 29-04-2011	
		WO 2010018444 A2 18-02-2010	
