



(12) **DEMANDE DE BREVET CANADIEN**
CANADIAN PATENT APPLICATION
(13) **A1**

(86) Date de dépôt PCT/PCT Filing Date: 2016/12/22
(87) Date publication PCT/PCT Publication Date: 2017/06/29
(85) Entrée phase nationale/National Entry: 2018/06/20
(86) N° demande PCT/PCT Application No.: EP 2016/082314
(87) N° publication PCT/PCT Publication No.: 2017/109039
(30) Priorité/Priority: 2015/12/22 (EP PCT/EP2015/081075)

(51) Cl.Int./Int.Cl. *A61K 48/00* (2006.01),
C12N 15/86 (2006.01), *C12N 15/861* (2006.01),
C12N 15/864 (2006.01)
(71) Demandeur/Applicant:
VRIJE UNIVERSITEIT BRUSSEL, BE
(72) Inventeurs/Inventors:
CHUAH, MARINEE, BE;
VANDENDRIESSCHE, THIERRY, BE
(74) Agent: GOWLING WLG (CANADA) LLP

(54) Titre : ELEMENTS REGULATEURS D'ACIDES NUCLEIQUES SPECIFIQUES A L'ENDOTHELIUM ET METHODES
ET UTILISATION ASSOCIEES
(54) Title: ENDOTHELIUM-SPECIFIC NUCLEIC ACID REGULATORY ELEMENTS AND METHODS AND USE
THEREOF

(57) **Abrégé/Abstract:**

The present invention relates to nucleic acid regulatory elements that are able to enhance endothelial cell-specific expression of genes, methods employing these regulatory elements and uses of these elements. Expression cassettes and vectors containing these nucleic acid regulatory elements are also disclosed. The present invention is particularly useful for applications using gene therapy, more particularly endothelial cell-directed gene therapy, and for vaccination purposes.

(51) **International Patent Classification:**
A61K 48/00 (2006.01) *C12N 15/861* (2006.01)
C12N 15/86 (2006.01) *C12N 15/864* (2006.01)

(21) **International Application Number:**
PCT/EP2016/082314

(22) **International Filing Date:**
22 December 2016 (22.12.2016)

(25) **Filing Language:** English

(26) **Publication Language:** English

(30) **Priority Data:**
PCT/EP2015/081075
22 December 2015 (22.12.2015) EP

(71) **Applicant:** **VRIJE UNIVERSITEIT BRUSSEL** [BE/BE]; Pleinlaan 2, 1050 Brussel (BE).

(72) **Inventors:** **CHUAH, Marinee**; Tiensesteenweg 250, 3360 Bierbeek (BE). **VANDENDRIESSCHE, Thierry**; Tiensesteenweg 250, 3360 Bierbeek (BE).

(74) **Agents:** **VANHALST, Koen** et al.; E. Gevaertdreef 10a, 9830 Sint-Martens-Latem (BE).

(81) **Designated States** (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,

BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

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

(54) **Title:** ENDOTHELIUM-SPECIFIC NUCLEIC ACID REGULATORY ELEMENTS AND METHODS AND USE THEREOF

(57) **Abstract:** The present invention relates to nucleic acid regulatory elements that are able to enhance endothelial cell-specific expression of genes, methods employing these regulatory elements and uses of these elements. Expression cassettes and vectors containing these nucleic acid regulatory elements are also disclosed. The present invention is particularly useful for applications using gene therapy, more particularly endothelial cell-directed gene therapy, and for vaccination purposes.


WO 2017/109039 A1

ENDOTHELIUM-SPECIFIC NUCLEIC ACID REGULATORY ELEMENTS AND METHODS AND USE THEREOF

FIELD

The present invention relates to nucleic acid regulatory elements that are able to enhance
5 endothelial-specific expression of genes, methods employing these regulatory elements
and use thereof. The invention further encompasses expression cassettes, vectors and
pharmaceutical compositions comprising these regulatory elements. The present invention
is particularly useful for applications using gene therapy, more particularly endothelial-
directed gene therapy, and for vaccination purposes.

BACKGROUND

Endothelial cells form a single cell layer that lines all blood vessels and regulate
exchanges between the bloodstream and the surrounding tissues. Signals from
endothelial cells organize the growth and development of connective tissue cells that form
the surrounding layers of the blood-vessel wall. New blood vessels can develop from the
15 walls of existing small vessels by the outgrowth of endothelial cells, which have the
capacity to form hollow capillary tubes even when isolated in culture. Endothelial cells of
developing arteries and veins express different cell-surface proteins, which may control
the way in which they link up to create a capillary bed. (Molecular Biology of the Cell. 4th
Edition). A homeostatic mechanism ensures that blood vessels permeate every region of
20 the body. Cells that are short of oxygen increase their concentration of hypoxia-inducible
factor 1 (HIF-1), which stimulates the production of vascular endothelial growth factor
(VEGF). VEGF acts on endothelial cells, causing them to proliferate and invade the
hypoxic tissue to supply it with new blood vessels. (Molecular Biology of the Cell. 4th
Edition).

25 Endothelial cell phenotypes vary between different organs, between different segments of
the vascular loop within the same organ, and between neighbouring endothelial cells of
the same organ and blood vessel type. In addition to differences in structure, endothelial
cells show remarkable heterogeneity in function. For example, the endothelial cells in the
liver, called liver sinusoidal endothelial cells (LSEC), form a continuous lining of the liver
30 capillaries, or sinusoids, separating parenchymal cells and fat-storing cells from sinusoidal
blood. LSECs represent unique, highly specialized endothelial cells in the body. LSECs
differ in fine structure from endothelial cells lining larger blood vessels and from other
capillary endothelia in that they lack a distinct basement membrane and also contain open

pores, or fenestrae, in the thin cytoplasmic projections which constitute the sinusoidal wall. This distinctive morphology supports the protective role played by liver endothelium, the cells forming a general barrier against pathogenic agents and serving as a selective sieve for substances passing from the blood to parenchymal and fat-storing cells, and vice versa. Sinusoidal endothelial cells, furthermore, significantly participate in the metabolic and clearance functions of the liver. They have been shown to be involved in the endocytosis and metabolism of a wide range of macromolecules, including glycoproteins, lipoproteins, extracellular matrix components, and inert colloids, establishing endothelial cells as a vital link in the complex network of cellular interactions and cooperation in the liver.

In addition, LSECs have long been noted to contribute to liver regeneration after liver injury. In normal liver, the major cellular source of hepatocyte growth factor (HGF) is the hepatic stellate cell, but after liver injury, HGF expression has been thought to increase markedly in proliferating LSECs (DeLeve et al. Liver sinusoidal endothelial cells and liver regeneration J Clin. Invest. 2013). Another unexpected function of LSEC was recently reported (Shahani et al., J. Thromb. Hemost 2014), demonstrating that LSECs and not hepatocytes express coagulation factor VIII (FVIII). Moreover, endothelial cells, including LSECs, also express von Willebrand factor (vWF). It is known that secreted FVIII would be relatively unstable unless it is associated with vWF. Deficiency of FVIII, a co-factor in the intrinsic coagulation pathway, results in hemophilia A. Liver transplantation in both FVIII-deficient dogs and patients with hemophilia A corrects these disorders. Although the liver is known to be the main site of FVIII production, other organs are probably also important for the regulation of FVIII secretion. Recent studies have shown that lung endothelial cells can synthesize FVIII. Microvascular endothelial cells from lung, heart, intestine, and skin as well as endothelial cells from pulmonary artery constitutively secreted FVIII and released it after treatment with phorbol-myristate acetate and epinephrine. By contrast, endothelial cells from the aorta, umbilical artery and umbilical vein did not constitutively secrete FVIII or release it after treatment with agonists, probably because of a lack of FVIII synthesis. Extrahepatic endothelial cells from certain vascular beds therefore appear to be an important FVIII production and storage site with the potential to regulate FVIII secretion in chronic and acute conditions (Shahani et al. Blood. 2010 Jun 10;115(23):4902-9). In addition, LSECs have also been reported to induce immunosuppressive IL-10-producing Th1 cells via the Notch pathway (Neumann et al. Eur J Immunol. 2015 Jul;45(7):2008-16) suggesting an important immune-modulatory role. Therefore, LSEC dysfunction has been regarded as a key event in multiple liver disorders. Future studies are likely to disclose

more fully the role of LSEC in the regulation of liver hemodynamics, in liver metabolism and blood clearance, in the maintenance of hepatic structure, in the pathogenesis of various liver diseases, and in the aging process in the liver (De Leeuw et al. J Electron Microsc Tech. 1990 Mar;14(3):218-36).

5 The endothelium is involved in many disease states, either as a primary determinant of pathophysiology or as a secondary target. In particular, endothelial cell dysfunction can be caused by acquired, complex multifactorial, genetic or infectious diseases. In some cases, the underlying endothelial cell defect can be life-threatening, for which no effective cure is presently available. Dysfunction of the vascular endothelium is a hallmark of many human
10 diseases. The endothelium is directly involved in many different diseases including peripheral vascular disease, stroke, heart disease, diabetes, insulin resistance, chronic kidney failure, tumor growth, metastasis, venous thrombosis, and severe viral infectious diseases. Consequently, the endothelium has substantial untapped potential as a therapeutic target. In particular, endothelial cells are attractive target cells for gene
15 therapy to enable robust and/or sustained expression of the cognate therapeutic genes.

In particular, endothelial dysfunction is one of the major pathophysiological mechanisms that leads towards coronary artery disease and other atherosclerotic diseases. Atherosclerosis is a progressive vascular disease characterized by the accumulation of lipids, inflammatory cells, and fibrous elements. In Western societies, it is the underlying
20 cause of approximately 50% of all deaths. Dysfunction or injury of vascular endothelial cells is critical for the development of atherosclerosis. The endothelium functions as a selectively permeable barrier between blood and tissues as it can regulate transcytosis and generate effector molecules such as nitric oxide (NO) that regulate thrombosis, inflammation, vascular tone and vascular remodeling. For example, overexpression of
25 STAMP2 suppresses atherosclerosis and stabilizes plaques in diabetic mice. Similarly, it had been reported that over-expression of ABCG1 by somatic gene transfer to the atherosclerotic vessel wall results in a significant improvement of plaque morphology and composition, and of vascular function in vivo (Heart Int. 2012 Jun 5; 7(2): e12.).

Endothelial cells also play a key role in angiogenesis and vasculogenesis. Angiogenesis is
30 the physiological process through which new blood vessels form from pre-existing vessels. This is distinct from vasculogenesis, which is the de novo formation of endothelial cells from mesoderm cell precursors. Angiogenesis and vasculogenesis are normal and vital processes in growth and development. However, they also represent a fundamental step in cancer progression, justifying the use of angiogenesis or vasculogenesis inhibitors

in the treatment of cancer. Conversely, promoting angiogenesis and vasculogenesis may benefit the treatment of ischemia which is associated with decrease in blood supply to certain organs or tissues. Consequently, the delivery of genes into endothelial cells that either promote or inhibit angiogenesis and/or vasculogenesis opens new perspectives for the treatment of cardiovascular disease and cancer, respectively. This includes a plethora of therapeutic genes including VEGF, PLGF, FGF, sFLT1, antibodies directed against these factors and their cognate receptors, cytokines, chemokines, etc.

Furthermore, endothelial cells are also promising targets for gene therapy to express therapeutic proteins that are missing or defective in genetic disorders that result from mutations in the respective genes. For example, this includes FVIII, vWF or ADAMTS13 (a disintegrin and metalloproteinase with a thrombospondin type 1 motif, member 13). As mentioned above, hemophilia A is due to a deficiency in FVIII. Moreover, deficiency in vWF causes a bleeding diathesis in patients suffering from von Willebrand disease (VWD). Conversely, a deficiency in ADAMTS13 is linked to the development of microvascular thrombosis characteristic of thrombotic thrombocytopenic purpura (TTP). Consequently, to establish an effective cure for these genetic diseases, robust expression of FVIII, vWF or ADAMTS13 in the endothelial cells is required. In addition, given their proximity to the blood, endothelial cells are also ideally suited to express other therapeutic proteins that are normally not expressed by endothelial cells but that can be directly secreted in the blood. This includes, but is not limited to, other coagulation factors (e.g. factor VII, IX, XI etc.), serum proteins (α 1-antitrypsin, AAT, antibodies, growth factors etc.).

Finally, endothelial cells also play a key role in viral infection. For example, the Ebola virus is an aggressive pathogen that causes a highly lethal hemorrhagic fever syndrome in humans and nonhuman primates. The virus infects microvascular endothelial cells and compromises vascular integrity. Infection of endothelial cells also induces a cytopathic effect and damage to the endothelial barrier. Similarly, Dengue virus causes leakage of the vascular endothelium, resulting in dengue hemorrhagic fever and dengue shock syndrome. The endothelial cell lining of the vasculature regulates capillary permeability and is altered by immune and chemokine responses which affect fluid barrier functions of the endothelium. Human endothelial cells are susceptible to infection by dengue virus. Following attachment to human endothelial cell receptors, dengue virus causes a productive infection that has the potential to increase viral dissemination and viremia. This provides the potential for dengue virus-infected endothelial cells to directly alter barrier

functions of the endothelium, contribute to enhancement of immune cell activation, and serve as potential targets of immune responses which play a central role in dengue pathogenesis.

Hence, there is a need to establish effective cures by gene therapy to enable robust expression of the cognate therapeutic genes in the endothelial cells. This requires the development of potent expression cassettes containing the genes of interest. Consequently, there is a need to identify robust nucleic acid regulatory elements capable of substantially increasing transcription in the endothelium.

SUMMARY OF THE INVENTION

To achieve a robust and specific expression in endothelial cells, the inventors have developed a computational approach to identify robust nucleic acid regulatory elements such as cis-regulatory elements (CREs) that are capable of substantially increasing transcription in endothelial cells (also called EC-CREs or EC-REs) when combined with an endothelial specific promoter. Endothelial specific nucleic acid regulatory elements were identified in silico and subsequently validated in in vitro cell lines and also in vivo in mice.

These nucleic acid regulatory elements are critically important for the regulation of gene expression in an endothelial cell type-specific manner. They are typically composed of clusters of transcription factor binding site (TFBS) motifs. The types and arrangement of TFBS and epigenetic modification patterns influence gene expression levels and specificity. Conventional methods of vector design relied on haphazard trial-and-error approaches whereby transcriptional enhancers were combined with promoters to boost expression levels. Though this could sometimes be effective, it often resulted in non-productive combinations that resulted in either modest or no increased expression levels of the gene of interest and/or loss of tissue specificity. Moreover, these conventional approaches did not take into account the importance of including evolutionary conserved regulatory motifs into the expression modules. The development of nucleic acid regulatory elements that can lead to robust and specific expression in endothelial cells will be very useful for achieving safe and efficient gene delivery to endothelial cells for the treatment of disorders related to endothelial cell dysfunction.

The present inventors have relied on a computational approach (cf. Figure 1) to identify robust nucleic acid regulatory elements that boost gene expression at the transcriptional level in endothelial cells (designated herein as "EC"). This requires the following

computational steps: (1) endothelial cell-specific genes were identified that are highly and specifically expressed based on expression data from endothelial cells; (2) publicly available databases (ENSEMBL) were used for extracting the sequences upstream of the Transcriptional Start Site (TSS) of the selected genes. 3) These sequences were then submitted into UCSC Genome Browser Database for locating the transcription start site in human genome. To extract the corresponding endothelial cell nucleic acid regulatory elements, defined herein as the nucleic acid regulatory elements, the sequences were selected based on the following criteria: a) rich TFBS content, b) epigenetic signatures associated with high DNase hypersensitivity or chromatin accessibility (i.e. histone modifications), and c) evolutionary conserved clusters of TFBS associated with highly expressed endothelial cell-specific genes.

As shown in the experimental section, the inventors identified nucleic acid regulatory elements that will specifically enhance gene expression in endothelial cells.

The endothelial cell regulatory elements will subsequently be validated *in vivo*, yielding efficient and tissue-specific gene expression. This approach hence, allows for the use of lower and thus safer vector doses, while maximizing therapeutic efficacy.

The invention therefore provides the following aspects:

Aspect 1. A nucleic acid regulatory element for enhancing endothelial cell-specific gene expression comprising, consisting essentially of, or consisting of the sequence selected from the group consisting of: SEQ ID NO:1 to 33, or a sequence having at least 80%, preferably at least 85%, more preferably at least 90%, even more preferably at least 95%, such as 95%, 96%, 97%, 98%, or 99%, identity to any of these sequences. In a preferred embodiment of said aspect, said nucleic acid regulatory element for enhancing endothelial cell-specific gene expression comprises, consists essentially of, or consists of the sequence of SEQ ID NO.22, or a sequence having at least 80%, preferably at least 85%, more preferably at least 90%, even more preferably at least 95%, such as 95%, 96%, 97%, 98%, or 99%, identity to any of these sequences.

Aspect 2. The nucleic acid regulatory element according to aspect 1, having a maximal length of 1000 nucleotides, preferably 800 nucleotides, more preferably 700 nucleotides, most preferably of 610 nucleotides, still comprising the regulatory element defined by any one of SEQ ID Nos: 1 to 33.

Aspect 3. A nucleic acid regulatory element for enhancing gene expression in endothelial cells comprising, consisting essentially of, or consisting of the complement of a

sequence as defined by any one of SEQ ID Nos: 1 to 33, or hybridizing under stringent conditions to a sequence as defined by any one of SEQ ID Nos: 1 to 33.

Aspect 4. Use of the nucleic acid regulatory element according to any one aspects 1 to 3 in a nucleic acid expression cassette, or a vector, more particularly for enhancing gene expression in endothelial cells of said nucleic acid expression cassette or vector.

Aspect 5. A nucleic acid expression cassette comprising at least one, such as one, two, three, four, five or more, nucleic acid regulatory elements according to any one of aspects 1 to 3, operably linked to a promoter.

Aspect 6. The nucleic acid expression cassette according to aspect 5, wherein the nucleic acid regulatory element is operably linked to a promoter and a transgene.

Aspect 7. The nucleic acid expression cassette according any one of aspects 5 or 6, wherein the promoter is an endothelial cell-specific promoter, such as the promoter of any one of the genes selected from the group comprising: IFI27, ICAM2, VWF, EDN1, ENG, ECSCR, CDH5 (vascular endothelial cadherin promoter, cadherin 5 type 2), PECAM1, HHIP, TIE1 and HYAL2.

Aspect 8. The nucleic acid expression cassette according to any one of aspects 5 to 7, wherein the transgene encodes a therapeutic protein or an immunogenic protein.

Aspect 9. The nucleic acid expression cassette according to any one of aspects 5 to 8, wherein the transgene encodes a secretable protein or a structural protein.

Aspect 10. The nucleic acid expression cassette according to aspect 8 or 9, wherein said transgene is selected from the group comprising: hepatocyte growth factor (HGF), coagulation factor VIII (FVIII), coagulation factor VII (FVII), tissue factor (TF), tissue factor pathway inhibitor (TFPI), coagulation factor IX (FIX), coagulation factor XI (FXI), von Willebrand factor (vWF), ADAMTS13, VEGF, PLGF, FGF, sFLT1, α 1-antitrypsin, AAT, apolipoprotein A-I (apoA-I), matrix metalloproteinases including but not limited to matrix metalloproteinase-3 (TIMP-3), nitric oxide synthase (NOS), antibodies, growth factors, cytokines, chemokines and antibodies, including but not limited to antibodies directed against any one of said transgenes, factors and their cognate receptors or against any secreted protein or viral protein, small interfering RNA, guide RNA, endonuclease, and Cas9.

Aspect 11. The nucleic acid expression cassette according to any one of aspects 5 to 10, further comprising a polyadenylation signal, preferably the Simian Virus 40 (SV40)

polyadenylation signal, a synthetic polyadenylation signal or a bovine growth hormone polyadenylation signal.

Aspect 12. A vector comprising the nucleic acid regulatory element according to any one of aspects 1 to 3, or the nucleic acid expression cassette according to any one of
5 aspects 5 to 11.

Aspect 13. The vector according to aspect 12, which is a viral vector, preferably a lentiviral vector (LV), an adeno-associated viral (AAV) vector, or an adenoviral vector (AV). In specific examples of said aspect, the vector is a self- or non-self inactivating lentiviral vector, preferably a self inactivating lentiviral vector.

10 In specific examples the LV has the following components: EC-CRE-PM-TG, with EC-CRE being one of the newly identified regulatory sequences as defined in SEQ ID Nos 1-33; PM being an endothelial cell-specific promotor such as, but not limited to, those referred to in aspect 7, and TG being a transgene such as, but not limited to, the transgenes identified herein and in particular those defined in aspect 10.

15 Taking the example with endothelial cell-specific promotor ICAM2, and using the self inactivating lentiviral vector backbone pCDH, the vector can be:

pCDH-CDH5-EC-CRE1a-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

20 pCDH-CDH5-EC-CRE1b-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

pCDH-CDH5-EC-CRE1c-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

pCDH-CDH5-EC-CRE1d-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

25 pCDH-CDH5-EC-CRE1e-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

pCDH-HYAL2-EC-CRE1f-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

30 pCDH-ECSCR-EC-CRE1a-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

pCDH-ECSCR-EC-CRE1b-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

pCDH-EDN1-EC-CRE1a-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

pCDH-ENG-EC-CRE1a-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

5 pCDH-ENG-EC-CRE1b-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

pCDH-ENG-EC-CRE1c-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

10 pCDH-HHIP-EC-CRE1a-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

pCDH-HHIP-EC-CRE1b-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

pCDH-HYAL2-EC-CRE1a-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

15 pCDH-HYAL2-EC-CRE1b-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

pCDH-HYAL2-EC-CRE1c-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

20 pCDH-ICAM2-EC-CRE1a-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

pCDH-ICAM2-EC-CRE1b-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

pCDH-ICAM2-EC-CRE1c-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

25 pCDH-IFI27-EC-CRE1a-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

pCDH-IFI27-EC-CRE1b-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

30 pCDH-PECAM1-EC-CRE1a-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

pCDH-TIE1-EC-CRE1a-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

pCDH-TIE1-EC-CRE1b-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10,

5 pCDH-VWF-EC-CRE1a-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10, or

pCDH-VWF-EC-CRE1b-ICAM2-TG, such as wherein the TG is any one of those identified in aspect 10.

10 In anyone of said VL constructs, the ICAM2-promotor can be replaced by another endothelial cell specific promotor, such as, but not limited to those exemplified in aspect 7. In anyone of said VL constructs, the vector backbone can be exchanged by another suitable backbone such as those known to the person skilled in the art.

Aspect 14. The vector according to aspect 12, which is a non-viral vector, preferably a plasmid, a minicircle, an episomal vector, or a transposon-based vector, such as a
15 PiggyBac-based vector or a Sleeping Beauty-based vector.

Aspect 15. A pharmaceutical composition comprising the nucleic acid expression cassette according to any one of aspects 5 to 11, or the vector according to any one of aspects 12 to 14, and a pharmaceutically acceptable carrier.

Aspect 16. The nucleic acid regulatory element according to any one of aspects 1 to 3,
20 the nucleic acid expression cassette according to any one of aspects 5 to 11, the vector according to any one of aspects 12 to 14, or the pharmaceutical composition according to aspect 15 for use in medicine, more preferably for use in gene therapy, in particular for use in treating endothelial cell dysfunction, preferably such as any one of the diseases or disorders selected from the group comprising: liver diseases, hemophilia A, von
25 Willebrand disease, microvascular thrombosis, thrombotic thrombocytopenic purpura, peripheral vascular disease, coronary artery diseases, atherosclerotic diseases, stroke, heart disease, diabetes, insulin resistance, chronic kidney failure, tumor growth, metastasis, venous thrombosis, ischemia, tumour growth, tumour vascularisation, cancer and viral infectious diseases such as Ebola, Dengue and Dengue hemorrhagic fever.

30 Aspect 17. The nucleic acid regulatory element according to any one of aspects 1 to 3, the nucleic acid expression cassette according to any one of aspects 4 to 11, the vector according to any one of aspects 12 to 14, or the pharmaceutical composition according to aspect 15 for use as a vaccine, preferably a prophylactic vaccine, or for use in vaccination

therapy, preferably prophylactic vaccination. Alternatively, said nucleic acid regulatory element according to any one of aspects 1 to 3, the nucleic acid expression cassette according to any one of aspects 4 to 11, the vector according to any one of aspects 12 to 14, or the pharmaceutical composition according to aspect 15 can be for use in induction
5 of immunotolerance to the transgene.

Aspect 18. A method, preferably an *in vivo* method, for expressing a transgene product in endothelial cells, comprising:

- introducing the nucleic acid expression cassette according to any one of aspects 4 to 11, or the vector according to any one of aspects 12 to 14, comprising the nucleic acid regulatory element according to any one of aspects 1 to 3, into the cells; and
10

Aspect 19. A method, preferably an *in vitro* or *ex vivo* method, for expressing a transgene product in endothelial cells, comprising:

- introducing the nucleic acid expression cassette according to any one of aspects 4 to 11, or the vector according to any one of aspects 12 to 14, comprising the nucleic acid regulatory element according to any one of aspects 1 to 3, into the cells; and
15

- expressing the transgene product in the cells.

Aspect 20. A method for treating an endothelial cell-related disease or disorder comprising the administration of a therapeutically effective amount of the nucleic acid expression cassette according to any one of aspects 5 to 11, the vector according to any one of aspects 12 to 14, or the pharmaceutical composition according to aspect 15, each
20 comprising the nucleic acid regulatory element according to any one of aspects 1 to 3, to a subject in need thereof.

Aspect 21. The method according to aspect 20, wherein said endothelial cell-related disease or disorder is selected from the group comprising: endothelial cell dysfunction, preferably such as any one of the diseases or disorders selected from the group comprising: liver diseases, hemophilia A, von Willebrand disease, microvascular thrombosis, thrombotic thrombocytopenic purpura, peripheral vascular disease, coronary artery diseases, atherosclerotic diseases, stroke, heart disease, diabetes, insulin resistance, chronic kidney failure, tumor growth, metastasis, venous thrombosis, ischemia, tumour growth, tumour vascularisation, cancer and viral infectious diseases such as Ebola
30 Dengue and Dengue hemorrhagic fever.

BRIEF DESCRIPTION OF DRAWINGS

FIG. 1: The selection strategies of endothelial-specific CREs

As shown in the experimental section (Example 1), the inventors identified nucleic acid regulatory elements that will specifically enhance gene expression in endothelial cells. The endothelial specific regulatory elements were subsequently be validated *in vitro* and *in vivo* assays in mice. The details of the *in vitro* and *in vivo* validation are described in Example 2 to 6 below. The successful use of the endothelial CREs will hence, allows for the use of lower and thus safer vector doses, while maximizing therapeutic efficacy.

FIG. 2: Lentiviral vector design

Lentiviral vectors were produced as described previously (VandenDriessche et al., Blood 2002). Briefly, the lentiviral vector-containing plasmids were cotransfected with a VSV-G expression plasmid, a gag-pol and Rev expression plasmid. Lentiviruses were produced by transient co-transfection of HEK293T (293T) cells using supplemented Dulbecco modified Eagle medium (Invitrogen) with 10% heat-inactivated fetal bovine serum (Invitrogen) and 1% penicillin/streptomycin. A total of 60 µg lentiviral plasmid was used for transfection of one double-tray culture chamber: 60 µg lentiviral plasmid, 30 µg pRSV-REV, 30 µg pMDLg/pRRE and 30 µg pCMV-VSV-G. Plasmid was pre-complexed with calcium phosphate (Calcium phosphate transfection kit, Invitrogen) for 30 minutes at room temperature. Transfection media was added to the cells for 16 hours and then replaced by fresh medium containing NU-serum (Invitrogen) and Sodium Butyrate (Sigma). Viral supernatant was harvested 48 and 72 hours after transfection and concentrated using a Centricon concentrator (Millipore) (2000 rpm for 1 hours at 4°C). Aliquots of viruses were stored at - 80°C. The physical titer in nanograms per microliter of all LVs was determined using a p24 colorimetric enzyme-linked immunosorbent assay (ELISA) kit (Cell Biolabs) according to the manufacturer's instructions. This value was then used to calculate an estimated vector titer equivalent in transducing units (TU) per milliliter. Polybrene (8 µg/mL) was added to the concentrated vectors to enhance transduction.

FIG. 3: Flow cytometry analysis of HUVECs and LSECs transduced with LV CMV-GFP vector (MOI 50) at 72 hr timepoint post transduction.

FIG. 4: FVIII antigen expression expressed in ng/ml in HUVECs and LSECS transduced with different lentiviral vector designs. FVIII antigen expression levels were determined.at 24, 48 and 72 hrs after transduction.

FIG. 5: The FVIII expression levels of each EC-CRE-ICAM2-FVIII construct in HUVECs. The percentage compared to the ICAM2 construct without EC-CREs. The expression values were represented as mean with standard error of mean (SEM; n = 3 for each group).

5 **FIG. 6: FVIII expression after lentiviral in vivo transduction.** FVIII protein expression was determined using a human FVIII-specific ELISA on the plasma samples collected 5 weeks after lentiviral vector injection. The mice cohorts (n= 4 mice per cohort) include ICAM2-FVIII (no CRE control), HYAL2-EC-CRE1a-ICAM2-FVIII, HYAL2-EC-CRE1b-ICAM2-FVIII & IFI27-EC-CRE1b-ICAM2-FVIII.

10 **FIG. 7: mRNA analysis of the CD146-positive endothelial cells isolated from liver and spleen.** The human FVIII expression encoded by the lentiviral vector was normalized to endogenous mouse GAPDH expression.

FIG. 8: plasmid maps of lentiviral vectors

15 All endothelial CRE's were cloned, in a self inactivating the lentiviral vectors called pCDH as shown in the map. A: pCDH-HYAL2-EC-CRE1a-ICAM2-FVIII (SEQ ID NO. 50); B: pCDH-ICAM2-FVIII (SEQ ID NO. 49); C: pCDH-IFI27-EC-CRE1b-ICAM2-FVIII (SEQ ID NO. 52); D: pCDH-HYAL2-EC-CRE1b-ICAM2-FVIII (SEQ ID NO. 51).

DESCRIPTION

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

The terms "comprising", "comprises" and "comprised of" as used herein are synonymous with "including", "includes" or "containing", "contains", and are inclusive or open-ended and do not exclude additional, non-recited members, elements or method steps. The terms also encompass "consisting of" and "consisting essentially of", which enjoy well-
25 established meanings in patent terminology.

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

30 The terms "about" or "approximately" as used herein when referring to a measurable value such as a parameter, an amount, a temporal duration, and the like, are meant to encompass variations of and from the specified value, such as variations of +/-10% or less, preferably +/-5% or less, more preferably +/-1% or less, and still more preferably +/-0.1% or less of and from the specified value, insofar such variations are appropriate to

perform in the disclosed invention. It is to be understood that the value to which the modifier “about” refers is itself also specifically, and preferably, disclosed.

Whereas the terms “one or more” or “at least one”, such as one or more members or at least one member of a group of members, is clear per se, by means of further
5 exemplification, the term encompasses inter alia a reference to any one of said members, or to any two or more of said members, such as, e.g., any ≥ 3 , ≥ 4 , ≥ 5 , ≥ 6 or ≥ 7 etc. of said members, and up to all said members. In another example, “one or more” or “at least one” may refer to 1, 2, 3, 4, 5, 6, 7 or more.

The discussion of the background to the invention herein is included to explain the context
10 of the invention. This is not to be taken as an admission that any of the material referred to was published, known, or part of the common general knowledge in any country as of the priority date of any of the claims.

Throughout this disclosure, various publications, patents and published patent specifications are referenced by an identifying citation. All documents cited in the present
15 specification are hereby incorporated by reference in their entirety. In particular, the teachings or sections of such documents herein specifically referred to are incorporated by reference.

Unless otherwise defined, all terms used in disclosing the invention, including technical and scientific terms, have the meaning as commonly understood by one of ordinary skill in
20 the art to which this invention belongs. By means of further guidance, term definitions are included to better appreciate the teaching of the invention. When specific terms are defined in connection with a particular aspect of the invention or a particular embodiment of the invention, such connotation is meant to apply throughout this specification, i.e., also in the context of other aspects or embodiments of the invention, unless otherwise defined.

25 In the following passages, different aspects or embodiments of the invention are defined in more detail. Each aspect or embodiment so defined may be combined with any other aspect(s) or embodiment(s) unless clearly indicated to the contrary. In particular, any feature indicated as being preferred or advantageous may be combined with any other feature or features indicated as being preferred or advantageous.

30 Reference throughout this specification to “one embodiment”, “an embodiment” means that a particular feature, structure or characteristic described in connection with the embodiment is included in at least one embodiment of the present invention. Thus, appearances of the phrases “in one embodiment” or “in an embodiment” in various places throughout this specification are not necessarily all referring to the same embodiment, but

may. Furthermore, the particular features, structures or characteristics may be combined in any suitable manner, as would be apparent to a person skilled in the art from this disclosure, in one or more embodiments. Furthermore, while some embodiments described herein include some but not other features included in other embodiments, combinations of features of different embodiments are meant to be within the scope of the invention, and form different embodiments, as would be understood by those in the art. For example, in the appended claims, any of the claimed embodiments can be used in any combination.

For general methods relating to the invention, reference is made *inter alia* to well-known textbooks, including, e.g., "Molecular Cloning: A Laboratory Manual, 2nd Ed." (Sambrook *et al.*, 1989), "Current Protocols in Molecular Biology" (Ausubel *et al.*, 1987).

In one aspect, the invention relates to a nucleic acid regulatory element for enhancing gene expression in endothelial cells or tissue comprising, consisting essentially of (i.e., the regulatory element may for instance additionally comprise sequences used for cloning purposes, but the indicated sequences make up the essential part of the regulatory element, e.g. they do not form part of a larger regulatory region such as a promoter), or consisting of: a sequence selected from the group consisting of: SEQ ID NO:1 to 33, a sequence having at least 80%, preferably at least 85%, more preferably at least 90%, even more preferably at least 95%, such as 95%, 96%, 97%, 98%, or 99%, identity to any of these sequences, or a functional fragment of a sequence selected from the group consisting of: SEQ ID NO:1 to 33.

Tables 2 and 3 below depict the core nucleotide sequence of the different nucleic acid regulatory elements for enhancing gene expression in endothelial cells or tissue. Table 1 lists the corresponding genes and lengths.

A 'nucleic acid regulatory element', 'cis-acting regulatory element', 'CRE' or 'regulatory element' as used herein refers to a transcriptional control element, in particular a non-coding cis-acting transcriptional control element, capable of regulating and/or controlling transcription of a gene, in particular tissue-specific transcription of a gene. Regulatory elements comprise at least one transcription factor binding site (TFBS), more in particular at least one binding site for a tissue-specific transcription factor, most particularly at least one binding site for an endothelial cell-specific transcription factor. Typically, regulatory elements as used herein increase or enhance promoter-driven gene expression when compared to the transcription of the gene from the promoter alone, without the regulatory elements. Thus, regulatory elements particularly comprise enhancer sequences, although it is to be understood that the regulatory elements enhancing transcription are not limited

to typical far upstream enhancer sequences, but may occur at any distance of the gene they regulate. Indeed, it is known in the art that sequences regulating transcription may be situated either upstream (e.g. in the promoter region) or downstream (e.g. in the 3'UTR) of the gene they regulate *in vivo*, and may be located in the immediate vicinity of the gene or further away. Of note, although regulatory elements as disclosed herein typically comprise naturally occurring sequences, combinations of (parts of) such regulatory elements or several copies of a regulatory element, i.e. regulatory elements comprising non-naturally occurring sequences, are themselves also envisaged as regulatory element. Regulatory elements as used herein may comprise part of a larger sequence involved in transcriptional control, e.g. part of a promoter sequence. However, regulatory elements alone are typically not sufficient to initiate transcription, but require a promoter to this end. The regulatory elements disclosed herein are provided as nucleic acid molecules, i.e. isolated nucleic acids, or isolated nucleic acid molecules. Said nucleic acid regulatory elements hence have a sequence which is only a small part of the naturally occurring genomic sequence and hence is not naturally occurring as such, but is isolated therefrom.

The term "nucleic acid" as used herein typically refers to an oligomer or polymer (preferably a linear polymer) of any length composed essentially of nucleotides. A nucleotide unit commonly includes a heterocyclic base, a sugar group, and at least one, e.g. one, two, or three, phosphate groups, including modified or substituted phosphate groups. Heterocyclic bases may include *inter alia* purine and pyrimidine bases such as adenine (A), guanine (G), cytosine (C), thymine (T) and uracil (U) which are widespread in naturally-occurring nucleic acids, other naturally-occurring bases (e.g., xanthine, inosine, hypoxanthine) as well as chemically or biochemically modified (e.g., methylated), non-natural or derivatised bases. Sugar groups may include *inter alia* pentose (pentofuranose) groups such as preferably ribose and/or 2-deoxyribose common in naturally-occurring nucleic acids, or arabinose, 2-deoxyarabinose, threose or hexose sugar groups, as well as modified or substituted sugar groups. Nucleic acids as intended herein may include naturally occurring nucleotides, modified nucleotides or mixtures thereof. A modified nucleotide may include a modified heterocyclic base, a modified sugar moiety, a modified phosphate group or a combination thereof. Modifications of phosphate groups or sugars may be introduced to improve stability, resistance to enzymatic degradation, or some other useful property. The term "nucleic acid" further preferably encompasses DNA, RNA and DNA/RNA hybrid molecules, specifically including hnRNA, pre-mRNA, mRNA, cDNA, genomic DNA, amplification products, oligonucleotides, and synthetic (e.g., chemically synthesised) DNA, RNA or DNA/RNA hybrids. A nucleic acid can be naturally occurring, e.g., present in or isolated from nature; or can be non-naturally occurring, e.g.,

recombinant, i.e., produced by recombinant DNA technology, and/or partly or entirely, chemically or biochemically synthesised. A “nucleic acid” can be double-stranded, partly double stranded, or single-stranded. Where single-stranded, the nucleic acid can be the sense strand or the antisense strand. In addition, nucleic acid can be circular or linear.

5 As used herein “transcription factor binding site”, “transcription factor binding sequence” or “TFBS” refers to a sequence of a nucleic acid region to which transcription factors bind. Non-limiting examples of TFBS include binding sites for or such as: POLR2A, Pol2(b), GATA2, GATA-2, FOS, c-Fos, NR3C1, Freac-2, MAX, FOX2P, MYC, SRY, SOX9, SRF, JUN, TEAD4, EZH2, TBP, TCF7L2, SPI1, RELA, JUND, MXI1, JUNB, BHLHE40, 10 RCOR1, TCF12, TAL1, EP300, HDAC2, GTF2F1, SIN3AK20, FOSL2, ETS1, CTBP2, GATA3, CEBPB, FOXA1, YY1, RFX5, TAF1, REST, ELF1, CTCF, SMC3, FOXP2, RUNX3, NRF1, HDAC6, IRF4, PAX5, RAD21, WRNIP1, ERalpha_a, PU.1, TCF4, TAL1, HDAC2, GATA3, Mxi1, GTF2F1, ELF1, NRSF, CTCF, SMC3, Ini1, IRF4, PAX5, CTCF, Pol2-4H8, YY1, CTCF, FOXO1, FOXJ2, GATA-X, Gfi-1, Hand1/E47, MAZ, USF1, REST, 15 TFAP2A, TFAP2C, CHD2, ZNF274, BACH1, EBF, EBF1, ATF2:c-Jun, CREB1, ATF, Tax/CREB, CREB1, EGR1, NF-kappaB, c-Rel, Pax-3, FOXO4, SOX5, GR, ZNF263, Lmo2 complex, AP-4, HEN1, E2F6, PML, TRIM28, SMARCA4, RBBP5, NRF2F, TBL1XR1, STAT5A, MAFF, REST, JUND, IRF1, MAFK, eGFP-JunDATF1, ARID3A, ATF3, E2F6, GATA2, GATA-1, Brg1, TAL1, JunB, NR2F2, HDAC8, BCL3, ATF2, CBX3, 20 HNF4, FOXA2, KAP1, UBTF, GABP, GABPA, BCLAF1, SP1, FOXM1, MEF2A, ZNF143, ZBTB7A, NANOG, CTCFL, NFKB, CCNT2, EBF1, FOXA1, Max, c-Myc, STAT1, STAT2, MZF1, SMARCC1, E2F4, FOSL1, STAT3, P300, AP2gamma, MafF1, JunD, AP2alpha, FOXA2, HMGN3, ZBTB33, P300, Nkx2-2, Nkx2-5, SRF, YY1, HTF, CHX10, HNF1, OCT, Ncx, AP-2rep, Lmo2 complex, SOX5, GATA-1, CDP CR1, Cart-1, NFIV, RXRA, SREBP1, 25 MYBL2, HNF4G, HNF4A, HEY1, ZEB1, PHF8, CHD1, PU-1, RSRFC4, MEF-2, and/or Lyf-1. The nucleic acid regulatory elements described herein can comprise any one or more of said TFBS, or combinations thereof. Transcription factor binding sites may be found in databases such as Transfac®.

Sequences disclosed herein may be part of sequences of regulatory elements capable of 30 controlling transcription of endothelial cell-specific genes *in vivo*. Particular examples for endothelial-specific regulatory elements may in particular be controlling the following genes: IFI27, ICAM2, VWF, EDN1, ENG, ECSCR, CDH5, PECAM1, HHIP, TIE1 or HYAL2.

Accordingly, in embodiments, the nucleic acid regulatory elements disclosed herein 35 comprise a sequence from CDH5 regulatory elements, i.e. regulatory elements that

control expression of the CDH5 gene (Cadherin 5 or VE-cadherin gene) *in vivo*, e.g. regulatory elements comprising any one or more of SEQ ID NOs: 1 to 6, or functional fragments thereof. In other embodiments, the nucleic acid regulatory elements disclosed herein comprise a sequence from ECSCR regulatory elements, i.e. regulatory elements

5 that control expression of the ECSCR gene (Endothelial Cell-Specific Chemotaxis Regulator gene) *in vivo*, e.g. regulatory elements comprising any one or more of SEQ ID NOs: 7 or 8, or functional fragments thereof. In other embodiments, the nucleic acid regulatory elements disclosed herein comprise a sequence from EDN1 regulatory elements, i.e. regulatory elements that control expression of the EDN1 gene (Endothelin 1

10 gene) *in vivo*, e.g. regulatory elements comprising SEQ ID NO: 9, or functional fragments thereof. In other embodiments, the nucleic acid regulatory elements disclosed herein comprise a sequence from ENG regulatory elements, i.e. regulatory elements that control expression of the ENG gene (Endoglin gene) *in vivo*, e.g. regulatory elements comprising any one or more of SEQ ID NOs: 10 to 12, or functional fragments thereof. In other

15 embodiments, the nucleic acid regulatory elements disclosed herein comprise a sequence from HHIP regulatory elements, i.e. regulatory elements that control expression of the HHIP gene (Hedgehog Interacting Protein gene) *in vivo*, e.g. regulatory elements comprising any one or more of SEQ ID NOs: 13 or 14, or functional fragments thereof. In other embodiments, the nucleic acid regulatory elements disclosed herein comprise a

20 sequence from HYAL2 regulatory elements, i.e. regulatory elements that control expression of the HYAL2 gene (Hyaluronoglucosaminidase 2 gene) *in vivo*, e.g. regulatory elements comprising any one or more of SEQ ID NOs: 15 to 17, or functional fragments thereof. In other embodiments, the nucleic acid regulatory elements disclosed herein comprise a sequence from ICAM2 regulatory elements, i.e. regulatory elements that

25 control expression of the ICAM2 gene (Intercellular Adhesion Molecule 2 gene) *in vivo*, e.g. regulatory elements comprising any one or more of SEQ ID NOs: 18 to 20, or functional fragments thereof. In other embodiments, the nucleic acid regulatory elements disclosed herein comprise a sequence from IFI27 regulatory elements, i.e. regulatory elements that control expression of the IFI27 gene (Interferon, Alpha-Inducible Protein 27

30 gene) *in vivo*, e.g. regulatory elements comprising any one or more of SEQ ID NOs: 21 to 23, or functional fragments thereof. In other embodiments, the nucleic acid regulatory elements disclosed herein comprise a sequence from PECAM1 regulatory elements, i.e. regulatory elements that control expression of the PECAM1 gene (Platelet/Endothelial Cell Adhesion Molecule 1 gene) *in vivo*, e.g. regulatory elements comprising SEQ ID NO: 24,

35 or functional fragments thereof. In other embodiments, the nucleic acid regulatory elements disclosed herein comprise a sequence from TIE1 regulatory elements, i.e.

regulatory elements that control expression of the TIE1 gene (Tyrosine Kinase With Immunoglobulin-Like And EGF-Like Domains 1 gene) *in vivo*, e.g. regulatory elements comprising any one or more of SEQ ID NOs: 25 or 26, or functional fragments thereof. In other embodiments, the nucleic acid regulatory elements disclosed herein comprise a sequence from VWF regulatory elements, i.e. regulatory elements that control expression of the VWF gene (Von Willebrand Factor gene) *in vivo*, e.g. regulatory elements comprising any one or more of SEQ ID NOs: 27 or 28, or functional fragments thereof. In other embodiments, the nucleic acid regulatory elements disclosed herein comprise a sequence comprising any one or more of SEQ ID NOs: 29 to 33, or functional fragments thereof.

As used herein, the terms "identity" and "identical" and the like refer to the sequence similarity between two polymeric molecules, e.g., between two nucleic acid molecules, e.g., two DNA molecules. Sequence alignments and determination of sequence identity can be done, e.g., using the Basic Local Alignment Search Tool (BLAST) originally described by Altschul et al. 1990 (J Mol Biol 215: 403-10), such as the "Blast 2 sequences" algorithm described by Tatusova and Madden 1999 (FEMS Microbiol Lett 174: 247-250). Typically, the percentage sequence identity is calculated over the entire length of the sequence. As used herein, the term "substantially identical" denotes at least 90%, preferably at least 95%, such as 95%, 96%, 97%, 98% or 99%, sequence identity.

The term 'functional fragment' as used in the application refers to fragments of the regulatory element sequences disclosed herein that retain the capability of regulating endothelial cell-specific expression, i.e. they can still confer tissue specificity and they are capable of regulating expression of a (trans)gene in the same way (although possibly not to the same extent) as the sequence from which they are derived. Functional fragments may preferably comprise at least 20, at least 25, at least 30, at least 35, at least 40, at least 45, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 120, at least 150, at least 200, at least 250, at least 300, at least 350, or at least 400 contiguous nucleotides from the sequence from which they are derived. Also preferably, functional fragments may comprise at least 1, more preferably at least 2, at least 3, or at least 4, even more preferably at least 5, at least 10, or at least 15, of the transcription factor binding sites (TFBS) that are present in the sequence from which they are derived.

"endothelial cell-specific expression" as used in the application, refers to the preferential or predominant expression of a (trans)gene (as RNA and/or polypeptide) in endothelial cells and tissue comprising or built from endothelial cells, as compared to other (i.e. non-endothelial) cells or tissues. According to particular embodiments, at least 50%, more

particularly at least 60%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, at least 99% or 100% of the (trans)gene expression occurs within endothelial cells or endothelial tissue.

The term "endothelial cell" as used herein encompasses all endothelial cell types, such as the cells forming a single cell layer that lines all blood vessels and regulates exchanges between the bloodstream and the surrounding tissues. Many endothelial cell types exist and their phenotypes vary between different organs, between different segments of the vascular loop within the same organ, and between neighbouring endothelial cells of the same organ and blood vessel type. Non-limiting examples of such endothelial cells are: liver sinusoidal endothelial cells (LSEC), (micro)vascular endothelial cells from e.g. lung, heart, intestine, skin, retina, arterial endothelial cells, such as endothelial cells from pulmonary artery, the aorta, umbilical artery and umbilical vein, extrahepatic endothelial cells from certain vascular beds, blood-brain barrier ECs, bone marrow ECs, and high endothelial venule cells (HEVs).

According to a particular embodiment, endothelial cell specific expression entails that there is less than 10%, less than 5%, less than 2% or even less than 1% 'leakage' of expressed gene product to other organs or tissue than those comprising or built by endothelial cells, such as muscle, heart, lung, liver, brain, kidney and/or spleen.

The same applies *mutatis mutandis* for endothelial progenitor cell (EPC)-specific expression, which may be considered as a particular form of endothelial cell-specific expression. Hence, throughout the application, where endothelial cell-specific is mentioned in the context of expression, endothelial progenitor cell (EPC)-specific expression is also explicitly envisaged.

In embodiments, the invention relates to a nucleic acid regulatory element for enhancing gene expression in endothelial cells or tissue derived therefrom comprising, consisting essentially of, or consisting of a sequence selected from the group consisting of: SEQ ID NO:1 to 33; a sequence having at least 80%, preferably at least 85%, more preferably at least 90%, even more preferably at least 95%, such as 95%, 96%, 97%, 98%, or 99%, identity to any of the sequences selected from the group consisting of SEQ ID NO: 1 to 33; or a functional fragment thereof, wherein said functional fragment comprises at least 20, preferably at least 25, more preferably at least 50, at least 100, at least 200 or at least 250, contiguous nucleotides from the sequence from which it is derived, and wherein said functional fragment comprises at least 1, preferably at least 2, 3, 4, or 5, more preferably at least 10 or at least 15 transcription factor binding sites (TFBS) such as those TFBS that are present in the sequence from which it is derived.

It is also possible to make nucleic acid regulatory elements that comprise an artificial sequence by combining two or more identical or different sequences disclosed herein or functional fragments thereof. Accordingly, in certain embodiments a nucleic acid regulatory element for enhancing gene expression in endothelial cells is provided comprising at least two sequences selected from the group consisting of: SEQ ID NO:1-33.

For example, disclosed herein is a nucleic acid regulatory element comprising, consisting essentially of, or consisting of 2, 3, 4, or 5 repeats, e.g. tandem repeats, of any one of SEQ ID NOs:1 to 33, or combinations thereof.

Particular examples of nucleic acid regulatory elements that comprise an artificial sequence include the regulatory elements that are obtained by rearranging the transcription factor binding sites (TFBS) that are present in the sequences disclosed herein. Said rearrangement may encompass changing the order of the TFBSs and/or changing the position of one or more TFBSs relative to the other TFBSs and/or changing the copy number of one or more of the TFBSs. For example, also disclosed herein is a nucleic acid regulatory element for enhancing endothelial cell-specific gene expression, in particular endothelial cell-specific gene expression, comprising binding sites for e.g. Sp1, EGR-1, ETS and GATA. Further for example, also disclosed herein is a nucleic acid regulatory element for enhancing endothelial cell-specific gene expression, in particular comprising binding sites for one or more of: Sp1, EGR-1, ETS and GATA and combinations thereof. In some embodiments, these nucleic acid regulatory elements comprise at least two, such as 2, 3, 4, or more copies of any one or more of the recited TFBSs.

In some embodiments, the vector used is a lentiviral vector. In other embodiments, the vector used is an adeno-associated viral vector. In yet other embodiment, the vector used is an adenoviral vector. In case a lentiviral vector is used, it can be a self inactivating or a non self-inactivating lentiviral vector. A self inactivating lentiviral vector is sometimes preferred for clinical use since it is considered safer.

In case the regulatory element is provided as a single stranded nucleic acid, e.g. when using a single-stranded AAV vector, the complement strand is considered equivalent to the disclosed sequences. Hence, also disclosed herein is a nucleic acid regulatory element for enhancing endothelial cell-specific gene expression comprising, consisting essentially of, or consisting of the complement of a sequence described herein, in particular a sequence selected from the group consisting of: SEQ ID NOs:1 to 33; a sequence having at least 80%, preferably at least 85%, more preferably at least 90%,

even more preferably at least 95%, such as 95%, 96%, 97%, 98%, or 99%, identity to any of these sequences; or a functional fragment thereof as defined herein.

Also disclosed herein is a nucleic acid regulatory element for enhancing endothelial cell-specific gene expression hybridizing under stringent conditions to a nucleic acid regulatory
5 element described herein, in particular to the nucleic acid regulatory element comprising, consisting essentially of, or consisting of a sequence selected from the group consisting of: SEQ ID NOs:1 to 33; a sequence having at least 90%, preferably at least 95%, such as 95%, 96%, 97%, 98%, or 99%, identity to any of these sequences; a functional fragment thereof as defined herein; or to its complement. Said nucleic acid regulatory elements do
10 not need to be of equal length as the sequence they hybridize to. In preferred embodiments, the size of said hybridizing nucleic acid regulatory element does not differ more than 25% in length, in particular 20% in length, more in particular 15% in length, most in particular 10% in length from the sequence it hybridizes to.

The expression 'hybridize under stringent conditions' refers to the ability of a nucleic acid
15 molecule to hybridize to a target nucleic acid molecule under defined conditions of temperature and salt concentration. Typically, stringent hybridization conditions are no more than 25°C to 30°C (for example, 20°C, 15°C, 10°C or 5°C) below the melting temperature (T_m) of the native duplex. Methods of calculating T_m are well known in the art. By way of non-limiting example, representative salt and temperature conditions for
20 achieving stringent hybridization are: 1x SSC, 0.5% SDS at 65°C. The abbreviation SSC refers to a buffer used in nucleic acid hybridization solutions. One liter of the 20x (twenty times concentrate) stock SSC buffer solution (pH 7.0) contains 175.3 g sodium chloride and 88.2 g sodium citrate. A representative time period for achieving hybridization is 12 hours.

25 Preferably the regulatory elements as described herein are fully functional while being only of limited length. This allows their use in vectors or nucleic acid expression cassettes without unduly restricting their payload capacity. Accordingly, in embodiments, the regulatory element disclosed herein is a nucleic acid of 1500 nucleotides or less, 1000 nucleotides or less, 900 nucleotides or less, 800 nucleotides or less, 700 nucleotides or
30 less, more preferably 610 nucleotides or less, such as 550 nucleotides or less, 500 nucleotides or less, 450 nucleotides or less, 400 nucleotides or less, 350 nucleotides or less, or 300 nucleotides or less (i.e. the nucleic acid regulatory element has a maximal length of 1500 nucleotides, 1000 nucleotides, 900 nucleotides, 800 nucleotides, 700 nucleotides, preferably 610 nucleotides, such as 550 nucleotides, 500 nucleotides, 450
35 nucleotides, 400 nucleotides, 350 nucleotides, or 300 nucleotides).

However, it is to be understood that the disclosed nucleic acid regulatory elements retain regulatory activity (i.e. with regard to specificity and/or activity of transcription) and thus they particularly have a minimum length of 20 nucleotides, 25 nucleotides, 30 nucleotides, 35 nucleotides, 40 nucleotides, 45 nucleotides, 50 nucleotides, 100 nucleotides, 150 nucleotides, 200 nucleotides, 250 nucleotides, 300 nucleotides, 350 nucleotides or 400 nucleotides.

In certain embodiments, the invention provides for a nucleic acid regulatory element of 1000 nucleotides or less, preferably 900 nucleotides or less, preferably 800 nucleotides or less, preferably 700 nucleotides or less of a sequence selected from the group consisting of: SEQ ID NOs:1 to 33; a sequence having at least 80%, preferably at least 85%, more preferably at least 90%, even more preferably at least 95%, such as 95%, 96%, 97%, 98%, or 99%, identity to any of said sequences; or a functional fragment thereof as defined herein.

The nucleic acid regulatory elements disclosed herein may be used in a nucleic acid expression cassette. Accordingly, in an aspect the invention provides for the use of the nucleic acid regulatory elements as described herein in a nucleic acid expression cassette.

In an aspect the invention provides a nucleic acid expression cassette comprising a nucleic acid regulatory element as described herein, operably linked to a promoter. In embodiments, the nucleic acid expression cassette does not contain a transgene. Such nucleic acid expression cassette may be used to drive expression of an endogenous gene. In preferred embodiments, the nucleic acid expression cassette comprises a nucleic acid regulatory element as described herein, operably linked to a promoter and a transgene.

As used herein, the term 'nucleic acid expression cassette' refers to nucleic acid molecules that include one or more transcriptional control elements (such as, but not limited to promoters, enhancers and/or regulatory elements, polyadenylation sequences, and introns) that direct (trans)gene expression in one or more desired cell types, tissues or organs. Typically, they will also contain a transgene, although it is also envisaged that a nucleic acid expression cassette directs expression of an endogenous gene in a cell into which the nucleic acid cassette is inserted.

The term 'operably linked' as used herein refers to the arrangement of various nucleic acid molecule elements relative to each other such that the elements are functionally connected and are able to interact with each other. Such elements may include, without

limitation, a promoter, an enhancer and/or a regulatory element, a polyadenylation sequence, one or more introns and/or exons, and a coding sequence of a gene of interest to be expressed (i.e., the transgene). The nucleic acid sequence elements, when properly oriented or operably linked, act together to modulate the activity of one another, and ultimately may affect the level of expression of the transgene. By “modulate” is meant increasing, decreasing, or maintaining the level of activity of a particular element. The position of each element relative to other elements may be expressed in terms of the 5' terminus and the 3' terminus of each element, and the distance between any particular elements may be referenced by the number of intervening nucleotides, or base pairs, between the elements. As understood by the skilled person, operably linked implies functional activity, and is not necessarily related to a natural positional link. Indeed, when used in nucleic acid expression cassettes, the regulatory elements will typically be located immediately upstream of the promoter (although this is generally the case, it should definitely not be interpreted as a limitation or exclusion of positions within the nucleic acid expression cassette), but this need not be the case *in vivo*. E.g., a regulatory element sequence naturally occurring downstream of a gene whose transcription it affects is able to function in the same way when located upstream of the promoter. Hence, according to a specific embodiment, the regulatory or enhancing effect of the regulatory element is position-independent.

In particular embodiments, the nucleic acid expression cassette comprises one nucleic acid regulatory element as described herein. In alternative embodiments, the nucleic acid expression cassette comprises two or more, such as, e.g., 2, 3, 4, 5, 6, 7, 8, 9, or 10, nucleic acid regulatory elements as described herein, i.e. they are combined modularly to enhance their regulatory (and/or enhancing) effect. In further embodiments, at least two of the two or more nucleic acid regulatory elements are identical or substantially identical. In yet further embodiments, all of the two or more regulatory elements are identical or substantially identical. The copies of the identical or substantially identical nucleic acid regulatory elements may be provided as tandem repeats in the nucleic acid expression cassette. In alternative further embodiments, at least two of the two or more nucleic acid regulatory elements are different from each other, that is to say, are defined by a different SEQ ID NO:.. The nucleic acid expression cassette may also comprise a combination of identical and substantially identical nucleic acid regulatory elements and non-identical nucleic acid regulatory elements.

For example, the nucleic acid expression cassette may comprise a nucleic acid regulatory element comprising SEQ ID NO:1, and a nucleic acid regulatory element comprising any

one or more of SEQ ID Nos: 2 to 33. Alternatively, this can be done for remaining regulatory elements defined by SEQ ID NOs:2 to 33, which can be combine with any one or more of the other regulatory elements.

As used in the application, the term 'promoter' refers to nucleic acid sequences that regulate, either directly or indirectly, the transcription of corresponding nucleic acid coding sequences to which they are operably linked (e.g. a transgene or endogenous gene). A promoter may function alone to regulate transcription or may act in concert with one or more other regulatory sequences (e.g. enhancers or silencers, or regulatory elements). In the context of the present application, a promoter is typically operably linked to a regulatory element as disclosed herein to regulate transcription of a (trans)gene. When a regulatory element as described herein is operably linked to both a promoter and a transgene, the regulatory element can (1) confer a significant degree of endothelial cell-specific expression *in vivo* (and/or *in vitro* in cell lines derived from endothelial cell- or tissue) of the transgene, and/or (2) can increase the level of expression of the transgene in endothelial cells (and/or *in vitro* in cell lines derived from endothelial cells or tissue).

The promoter may be homologous (i.e. from the same species as the animal, in particular mammal, to be transfected with the nucleic acid expression cassette) or heterologous (i.e. from a source other than the species of the animal, in particular mammal, to be transfected with the expression cassette). As such, the source of the promoter may be any virus, any unicellular prokaryotic or eukaryotic organism, any vertebrate or invertebrate organism, or any plant, or may even be a synthetic promoter (i.e. having a non-naturally occurring sequence), provided that the promoter is functional in combination with the regulatory elements described herein. In preferred embodiments, the promoter is a mammalian promoter, in particular a murine or human promoter.

The promoter may be an inducible or constitutive promoter.

Non-limiting exemplary endothelial cell-specific promoters are: the promoters of the genes depicted in Table 1 below, more preferably the

In particularly preferred embodiments, the promoter is a mammalian promoter, in particular a murine or human promoter.

In preferred embodiments, the promoter is from the vascular-endothelial cadherin gene, in particular the murine or human cadherin-5 gene, such as the promoter as defined in SEQ ID NO: 34 (cf. Table 4).

In preferred embodiments, the promoter is from the endothelin-1 gene, in particular the murine or human endothelin-1 gene, such as the promoter as defined in SEQ ID NO: 35 (cf. Table 4).

5 In preferred embodiments, the promoter is from the endoglin gene, in particular the murine or human endoglin gene, such as the promoter as defined in SEQ ID NO: 36 (cf. Table 4).

In preferred embodiments, the promoter is from the Fms-Related Tyrosine Kinase 1 gene, in particular the murine or human Fms-Related Tyrosine Kinase 1 gene, such as the promoter as defined in SEQ ID NO: 37 (cf. Table 4).

10 In preferred embodiments, the promoter is from the Intercellular Adhesion Molecule 2 gene, in particular the murine or Intercellular Adhesion Molecule 2 gene (ICAM2), such as the promoter as defined in SEQ ID NO: 38 (cf. Table 4).

15 Furthermore, the promoter does not need to be the promoter of the transgene in the nucleic acid expression cassette, although it is possible that the transgene is transcribed from its own promoter.

To minimize the length of the nucleic acid expression cassette, the regulatory elements may be linked to minimal promoters, or shortened versions of the promoters described herein. A 'minimal promoter' (also referred to as basal promoter or core promoter) as used herein is part of a full-size promoter still capable of driving expression, but lacking at least
20 part of the sequence that contributes to regulating (e.g. tissue-specific) expression. This definition covers both promoters from which (tissue-specific) regulatory elements have been deleted- that are capable of driving expression of a gene but have lost their ability to express that gene in a tissue-specific fashion and promoters from which (tissue-specific) regulatory elements have been deleted that are capable of driving (possibly decreased)
25 expression of a gene but have not necessarily lost their ability to express that gene in a tissue-specific fashion. Preferably, the promoter contained in the nucleic acid expression cassette disclosed herein is 1000 nucleotides or less in length, 900 nucleotides or less, 800 nucleotides or less, 700 nucleotides or less, 600 nucleotides or less, 500 nucleotides or less, 400 nucleotides or less, 300 nucleotides or less, or 250 nucleotides or less. One
30 particular non-limiting example of such a minimal promoter is the EDN1mini promoter (cf. Table 4).

The term 'transgene' as used herein refers to particular nucleic acid sequences encoding a polypeptide or a portion of a polypeptide to be expressed in a cell into which the nucleic acid sequence is introduced. However, it is also possible that transgenes are expressed

as RNA, typically to control (e.g. lower) the amount of a particular polypeptide in a cell into which the nucleic acid sequence is inserted. These RNA molecules include but are not limited to molecules that exert their function through RNA interference (shRNA, RNAi), micro-RNA regulation (miR) (which can be used to control expression of specific genes), catalytic RNA, antisense RNA, RNA aptamers, ZFN, TALEN, CRISPR/Cas9 or similar DNA or RNA cutters, etc..

How the nucleic acid sequence is introduced into a cell is not essential to the invention, it may for instance be through integration in the genome or as an episomal plasmid. Of note, expression of the transgene may be restricted to a subset of the cells into which the nucleic acid sequence is introduced. The term 'transgene' is meant to include (1) a nucleic acid sequence that is not naturally found in the cell (i.e., a heterologous nucleic acid sequence); (2) a nucleic acid sequence that is a mutant form of a nucleic acid sequence naturally found in the cell into which it has been introduced; (3) a nucleic acid sequence that serves to add additional copies of the same (i.e., homologous) or a similar nucleic acid sequence naturally occurring in the cell into which it has been introduced; or (4) a silent naturally occurring or homologous nucleic acid sequence whose expression is induced in the cell into which it has been introduced.

The transgene may be homologous or heterologous to the promoter (and/or to the animal, in particular mammal, in which it is introduced, e.g. in cases where the nucleic acid expression cassette is used for gene therapy).

The transgene may be a full length cDNA or genomic DNA sequence, or any fragment, subunit or mutant thereof that has at least some biological activity. In particular, the transgene may be a minigene, i.e. a gene sequence lacking part, most or all of its intronic sequences. The transgene thus optionally may contain intron sequences. Optionally, the transgene may be a hybrid nucleic acid sequence, i.e., one constructed from homologous and/or heterologous cDNA and/or genomic DNA fragments. By 'mutant form' is meant a nucleic acid sequence that contains one or more nucleotides that are different from the wild-type or naturally occurring sequence, i.e., the mutant nucleic acid sequence contains one or more nucleotide substitutions, deletions, and/or insertions. The nucleotide substitution, deletion, and/or insertion can give rise to a gene product (i.e. e., protein or nucleic acid) that is different in its amino acid/nucleic acid sequence from the wild type amino acid/nucleic acid sequence. Preparation of such mutants is well known in the art. In some cases, the transgene may also include a sequence encoding a leader peptide or signal sequence such that the transgene product will be secreted from the cell.

The transgene that may be contained in the nucleic acid expression cassettes described herein typically encodes a gene product such as RNA or a polypeptide (protein).

In embodiments, the transgene encodes a therapeutic protein. The therapeutic protein may be a secretable protein. Non-limiting examples of secretable proteins, in particular
5 secretable therapeutic proteins, include hepatocyte growth factor (HGF), coagulation factor VIII (FVIII), coagulation factor VII (FVII), coagulation factor IX (FIX), coagulation factor XI (FXI), tissue factor (TF), tissue factor pathway inhibitor (TFPI), von Willebrand factor (vWF), ADAMTS13, VEGF, PLGF, FGF, sFLT1, α 1-antitrypsin, AAT,, apolipoprotein A-I (apoA-I), matrix metalloproteinases including but not limited to matrix
10 metalloproteinase-3 (TIMP-3), insulin, erythropoietin, lipoprotein lipase, nitric oxide synthase (NOS), antibodies or nanobodies, including but not limited to antibodies directed against any one of said transgenes, factors and their cognate receptors or against any secreted protein or viral protein, small interfering RNA, guide RNA, endonuclease, and Cas9, growth factors, cytokines, chemokines, plasma factors etc.. The therapeutic protein
15 may also be a structural protein. Non-limiting examples of structural proteins, in particular structural therapeutic proteins, include proteins modulating vascular relaxation and vasoconstriction, atherosclerosis. In preferred embodiments, the transgene comprises the nitric oxide synthase (NOS).

In embodiments, the transgene encodes an immunogenic protein. Non-limiting examples
20 of immunogenic proteins include epitopes and antigens derived from a pathogen.

As used herein, the term "immunogenic" refers to a substance or composition capable of eliciting an immune response.

Other sequences may be incorporated in the nucleic acid expression cassette disclosed herein as well, typically to further increase or stabilize the expression of the transgene
25 product (e.g. introns and/or polyadenylation sequences).

Any intron can be utilized in the expression cassettes described herein, but may not be necessary. The term "intron" encompasses any portion of a whole intron that is large enough to be recognized and spliced by the nuclear splicing apparatus. Typically, short, functional, intron sequences are preferred in order to keep the size of the expression
30 cassette as small as possible which facilitates the construction and manipulation of the expression cassette. In some embodiments, the intron is obtained from a gene that encodes the protein that is encoded by the coding sequence within the expression cassette. The intron can be located 5' to the coding sequence, 3' to the coding sequence, or within the coding sequence. An advantage of locating the intron 5' to the coding

sequence is to minimize the chance of the intron interfering with the function of the polyadenylation signal. In embodiments, the nucleic acid expression cassette disclosed herein further comprises an intron. Non-limiting examples of suitable introns are Minute Virus of Mice (MVM) intron, beta-globin intron (betaVS-II), factor IX (FIX) intron A, Simian virus 40 (SV40) small-t intron, and beta-actin intron. Preferably, the intron is MVM intron.

Any polyadenylation signal that directs the synthesis of a polyA tail is useful in the expression cassettes described herein, examples of those are well known to one of skill in the art. Exemplary polyadenylation signals include, but are not limited to, polyA sequences derived from the Simian virus 40 (SV40) late gene, the bovine growth hormone (BGH) polyadenylation signal, the minimal rabbit β -globin (mRBG) gene, and the synthetic polyA s(SPA) site as described in Levitt et al. (1989, Genes Dev 3:1019-1025). Preferably, the polyadenylation signal is derived from SV40 (i.e. SV40 pA).

In particular embodiments, the invention provides a nucleic acid expression cassette comprising, consisting essentially of, or consisting of a nucleic acid regulatory element selected from the group consisting of SEQ ID NO: 1 to 33 or a sequence having 95% identity to said sequence, operably linked to a promoter, preferably a promoter selected from the group consisting of the promoter from the cadherin-5, endothelin-1, endoglin, Fms-Related Tyrosine Kinase 1, or Intercellular Adhesion Molecule 1 gene or the promoter, and a transgene, preferably a transgene encoding a luciferase.. In yet further embodiments the nucleic acid expression cassette further comprises a polyadenylation signal, preferably a polyadenylation signal derived from SV40.

In particular embodiments, the invention provides a nucleic acid expression cassette comprising, consisting essentially of, or consisting of a nucleic acid regulatory element selected from the group consisting of SEQ ID NO: 1 to 33 or a sequence having 95% identity to said sequence, operably linked to a promoter, preferably the promoter from the cadherin-5, endothelin-1, endoglin, Fms-Related Tyrosine Kinase 1, or Intercellular Adhesion Molecule 1 gene, and a transgene, preferably a transgene encoding a therapeutic or structural protein as defined herein. In yet further embodiments, the nucleic acid expression cassette further comprises a polyadenylation signal. In particular embodiments, any one of the following transgenes can be introduced: secretable proteins, in particular secretable therapeutic proteins, including hepatocyte growth factor (HGF), coagulation factor VIII (FVIII), coagulation factor VII (FVII), coagulation factor IX (FIX), coagulation factor XI (FXI), tissue factor (TF), tissue factor pathway inhibitor (TFPI), von Willebrand factor (vWF), ADAMTS13, VEGF, PLGF, FGF, sFLT1, α 1-antitrypsin (AAT), matrix metalloproteinases including but not limited to matrix metalloproteinase-3 (TIMP-3)

(TIMP-3), insulin, erythropoietin, lipoprotein lipase, antibodies or nanobodies, growth factors, cytokines, chemokines, plasma factors etc.. The therapeutic protein may also be a structural protein. Non-limiting examples of structural proteins, in particular structural therapeutic proteins, modulating vascular relaxation and vasoconstriction, atherosclerosis.

5 In preferred embodiments, the transgene comprises the nitric oxide synthase (NOS).

The nucleic acid regulatory element and the nucleic acid expression cassette disclosed herein may be used as such, or typically, they may be part of a nucleic acid vector. Accordingly, a further aspect relates to the use of a nucleic acid regulatory element as described herein or a nucleic acid expression cassette as described herein in a vector, in
10 particular a nucleic acid vector.

In an aspect, the invention also provides a vector comprising a nucleic acid regulatory element as disclosed herein. In further embodiments, the vector comprises a nucleic acid expression cassette as disclosed herein.

The term 'vector' as used in the application refers to nucleic acid molecules, e.g. double-
15 stranded DNA, which may have inserted into it another nucleic acid molecule (the insert nucleic acid molecule) such as, but not limited to, a cDNA molecule. The vector is used to transport the insert nucleic acid molecule into a suitable host cell. A vector may contain the necessary elements that permit transcribing the insert nucleic acid molecule, and, optionally, translating the transcript into a polypeptide. The insert nucleic acid molecule
20 may be derived from the host cell, or may be derived from a different cell or organism. Once in the host cell, the vector can replicate independently of, or coincidental with, the host chromosomal DNA, and several copies of the vector and its inserted nucleic acid molecule may be generated. The vectors can be episomal vectors (i.e., that do not integrate into the genome of a host cell), or can be vectors that integrate into the host cell
25 genome. The term 'vector' may thus also be defined as a gene delivery vehicle that facilitates gene transfer into a target cell. This definition includes both non-viral and viral vectors. Non-viral vectors include but are not limited to cationic lipids, liposomes, nanoparticles, PEG, PEI, plasmid vectors (e.g. pUC vectors, bluescript vectors (pBS) and pBR322 or derivatives thereof that are devoid of bacterial sequences (minicircles))
30 transposons-based vectors (e.g. PiggyBac (PB) vectors or Sleeping Beauty (SB) vectors), etc. Viral vectors are derived from viruses and include but are not limited to retroviral, lentiviral, adeno-associated viral, adenoviral, herpes viral, hepatitis viral vectors or the like. Typically, but not necessarily, viral vectors are replication-deficient as they have lost the ability to propagate in a given cell since viral genes essential for replication have been
35 eliminated from the viral vector. However, some viral vectors can also be adapted to

replicate specifically in a given cell, such as e.g. a cancer cell, and are typically used to trigger the (cancer) cell-specific (onco)lysis. Virosomes are a non-limiting example of a vector that comprises both viral and non-viral elements, in particular they combine liposomes with an inactivated HIV or influenza virus (Yamada et al., 2003). Another
5 example encompasses viral vectors mixed with cationic lipids.

In preferred embodiments, the vector is a viral vector, such as a retroviral, lentiviral, adenoviral, or adeno-associated viral (AAV) vector, more preferably a lentiviral vector. Lentiviral vectors are preferably derived from the human immune deficiency virus (HIV) though other lentiviral vectors based on other lentiviruses could also be used (including
10 but not limited to Equine infectious anemia virus). Lentiviral vectors can transduce endothelial cells. Production of lentiviral vectors can be achieved by (VandenDriessche et al. J. Thromb Hemostasis, 2007) transient co-transfection of lentiviral vector plasmids encoding the gene of interest with the gag-pol, rev and env-encoding helper constructs. Typically, a heterologous envelope is used such as the vesicular stomatitis virus G
15 glycoprotein (VSV-G) or an endotheliotropic envelope including but not limited to envelopes that confer antibody or nanobody (i.e. single chain antibody)-mediated endothelial retargeting targeting specific endothelial cell surface markers (VandenDriessche & Chuah, Blood. 2013 Sep 19;122(12):1993-4; Abel et al., Blood. 2013 Sep 19;122(12):2030-8; Buchholz et al. Trends Biotechnol. 2015 Dec;33(12):777-90;
20 Münch et al., Mol Ther. 2011 Apr;19(4):686-93; Anliker et al., Nat Methods. 2010 Nov;7(11):929-35).

In another embodiment the vector is an adeno-associated viral (AAV) vector. AAV vectors are preferably used as self-complementary, double-stranded AAV vectors (scAAV) in order to overcome one of the limiting steps in AAV transduction (i.e. single-stranded to
25 double-stranded AAV conversion) (McCarty, 2001, 2003; Nathwani et al, 2002, 2006, 2011; Wu et al., 2008), although the use of single-stranded AAV vectors (ssAAV) are also encompassed herein.

Production of AAV vector particles can e.g. be achieved by transient co-transfection of AAV-vector and AAV helper constructs, encoding AAV capsids into HEK293 cells,
30 followed by a purification step based on cesium chloride (CsCl) density gradient ultracentrifugation, as described (Vanden Driessche et al., 2007). Capsids can also be derived from different serotypes or are specifically modified to enhance endothelial cell transduction either by evolution or selection, antibody (nanobody engineering) or the use of DARPin (Work et al., Mol Ther. 2006 Apr;13(4):683-93; Münch et al., Nat Commun.

2015 Feb 10;6:6246; Buchholz et al., Trends Biotechnol. 2015 Dec;33(12):777-90; White et al. Circulation. 2004 Feb 3;109(4):513-9.)

In yet another embodiment the vector is an adenoviral vector. Adenoviral vectors are preferably derived from the human adenovirus 5 serotype or from other serotypes that display increased tropism to endothelial cells, including but not limited to Ad5T*F35++ (White et al., J Cardiothorac Surg. 2013 Aug 9;8:183. doi: 10.1186/1749-8090-8-183). Alternatively, the capsid can be engineered to enhance the endotheliotropic properties of the adenoviral vectors including but not limited to the references below (Nicol et al., FEBS Lett. 2009 Jun 18;583(12):2100-7; Nicklin and Baker, Mol Ther. 2008 Dec;16(12):1904-5; Work et al., Methods Mol Med. 2005;108:395-413; Work et al., Genet Vaccines Ther. 2004 Oct 8;2(1):14). They can be derived from either early-generation or helper-dependent adenoviral vectors (Mol Ther. 2010 Dec;18(12):2121-9). Production of these vectors has after transfection of adenoviral vector and helper constructs in HEK293T cells has been described previously (Mol Ther. 2010 Dec;18(12):2121-9).

Since the nucleic acid regulatory elements are de facto modular, also combinations of the best endothelial cell-specific nucleic acid regulatory elements with any other endothelial cell-specific nucleic acid regulatory elements to maximize expression in the desired target tissue are tested. Consequently, this can lead to the generation of a versatile endothelial cell-specific nucleic acid regulatory element platform tailor-made for diseases that affect endothelial cells and tissues encompassing those. Furthermore, the endothelial cell-specific nucleic acid regulatory elements can also be combined with other promoters or nucleic acid regulatory elements active in other target tissues.

In other embodiments, the vector is a non-viral vector, preferably a plasmid, a minicircle, or a transposon-based vector, such as a Sleeping Beauty(SB)-based vector or piggyBac(PB)-based vector.

In yet other embodiments, the vector comprises viral and non-viral elements.

In particular embodiments, the invention provides a vector comprising a nucleic acid expression cassette comprising a nucleic acid regulatory element comprising, consisting essentially of, or consisting of a nucleic acid regulatory element selected from the group consisting of SEQ ID NO:1 to 33, a promoter, preferably the promoter from the cadherin-5, endothelin-1, endoglin, Fms-Related Tyrosine Kinase 1, or Intercellular Adhesion Molecule 1 gene, a transgene, preferably a transgene encoding a therapeutic structural or secretable protein, and a polyadenylation signal. In particular, any one of the following transgenes can be introduced: secretable proteins, in particular secretable therapeutic

proteins, including hepatocyte growth factor (HGF), coagulation factor VIII (FVIII), coagulation factor VII (FVII), coagulation factor IX (FIX), coagulation factor XI (FXI), tissue factor (TF), tissue factor pathway inhibitor (TFPI), von Willebrand factor (vWF), ADAMTS13, VEGF, PLGF, FGF, sFLT1, α 1-antitrypsin (AAT), apolipoprotein A-I (apoA-I),
 5 matrix metalloproteinases including but not limited to matrix metalloproteinase-3 (TIMP-3) (TIMP-3), insulin, erythropoietin, lipoprotein lipase, antibodies or nanobodies, growth factors, cytokines, chemokines, plasma factors etc.. The therapeutic protein may also be a structural protein. Non-limiting examples of structural proteins, in particular structural therapeutic proteins, including proteins modulating vascular relaxation, vasoconstriction or
 10 atherosclerosis. In preferred embodiments, the transgene comprises the nitric oxide synthase (NOS).

In particular embodiments, the invention provides a vector comprising a nucleic acid expression cassette comprising a nucleic acid regulatory element comprising, consisting essentially of, or consisting of a nucleic acid regulatory element selected from the group
 15 consisting of SEQ ID NO:1 to 33, a promoter, preferably the promoter from the cadherin-5, endothelin-1, endoglin, Fms-Related Tyrosine Kinase 1, or Intercellular Adhesion Molecule 1 gene, , a transgene, preferably a transgene encoding secretable proteins, in particular secretable therapeutic proteins, including hepatocyte growth factor (HGF), coagulation factor VIII (FVIII), coagulation factor VII (FVII), coagulation factor IX (FIX),
 20 coagulation factor XI (FXI), tissue factor (TF), tissue factor pathway inhibitor (TFPI), von Willebrand factor (vWF), ADAMTS13, VEGF, PLGF, FGF, sFLT1, α 1-antitrypsin (AAT), apolipoprotein A-I (apoA-I), matrix metalloproteinases including but not limited to matrix metalloproteinase-3 (TIMP-3) (TIMP-3), insulin, erythropoietin, lipoprotein lipase, antibodies or nanobodies, growth factors, cytokines, chemokines, plasma factors etc.. The
 25 therapeutic protein may also be a structural protein. Non-limiting examples of structural proteins, in particular structural therapeutic proteins, including proteins modulating vascular relaxation, vasoconstriction or atherosclerosis. In preferred embodiments, the transgene comprises the nitric oxide synthase (NOS).

30 The nucleic acid expression cassettes and vectors disclosed herein may be used, for example, to express proteins that are normally expressed and utilized in endothelial cells (i.e. structural proteins), or to express proteins that are expressed in endothelial cells and that are then exported to the blood stream for transport to other portions of the body (i.e. secretable proteins). For example, the expression cassettes and vectors disclosed herein
 35 may be used to express a therapeutic amount of a gene product (such as a polypeptide, in

particular a therapeutic protein, or RNA) for therapeutic purposes, in particular for gene therapy. Typically, the gene product is encoded by the transgene within the expression cassette or vector, although in principle it is also possible to increase expression of an endogenous gene for therapeutic purposes. In an alternative example, the expression
5 cassettes and vectors disclosed herein may be used to express an immunological amount of a gene product (such as a polypeptide, in particular an immunogenic protein, or RNA) for vaccination purposes.

The nucleic acid expression cassettes and vectors as taught herein may be formulated in a pharmaceutical composition with a pharmaceutically acceptable excipient, i.e., one or
10 more pharmaceutically acceptable carrier substances and/or additives, e.g., buffers, carriers, excipients, stabilisers, etc. The pharmaceutical composition may be provided in the form of a kit.

The term “pharmaceutically acceptable” as used herein is consistent with the art and means compatible with the other ingredients of the pharmaceutical composition and not
15 deleterious to the recipient thereof.

Accordingly, a further aspect of the invention relates to a pharmaceutical composition comprising a nucleic acid expression cassette or a vector described herein.

The use of nucleic acid regulatory elements described herein for the manufacture of these pharmaceutical compositions is also disclosed herein.

20 In embodiments, the pharmaceutical composition may be a vaccine. The vaccine may further comprise one or more adjuvants for enhancing the immune response. Suitable adjuvants include, for example, but without limitation, saponin, mineral gels such as aluminium hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil or hydrocarbon emulsions, bacilli Calmette-Guerin (BCG),
25 Corynebacterium parvum, and the synthetic adjuvant QS-21. Optionally, the vaccine may further comprise one or more immunostimulatory molecules. Non-limiting examples of immunostimulatory molecules include various cytokines, lymphokines and chemokines with immunostimulatory, immunopotentiating, and pro-inflammatory activities, such as interleukins (e.g., IL-1, IL-2, IL-3, IL-4, IL-12, IL-13); growth factors (e.g., granulocyte-
30 macrophage (GM)-colony stimulating factor (CSF)); and other immunostimulatory molecules, such as macrophage inflammatory factor, Flt3 ligand, B7.1; B7.2, etc.

In a further aspect, the invention relates to the nucleic acid regulatory elements, the nucleic acid expression cassettes, the vectors, or the pharmaceutical compositions described herein for use in medicine.

As used herein, the terms “treat” or “treatment” refer to both therapeutic treatment and prophylactic or preventative measures. Beneficial or desired clinical results include, but are not limited to, prevention of an undesired clinical state or disorder, reducing the incidence of a disorder, alleviation of symptoms associated with a disorder, diminishment
5 of extent of a disorder, stabilized (i.e., not worsening) state of a disorder, delay or slowing of progression of a disorder, amelioration or palliation of the state of a disorder, remission (whether partial or total), whether detectable or undetectable, or combinations thereof. “Treatment” can also mean prolonging survival as compared to expected survival if not receiving treatment.

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

As used herein the terms “prevention”, “preventive treatment” or “prophylactic treatment”
20 and the like encompass preventing the onset of a disease or disorder, including reducing the severity of a disease or disorder or symptoms associated therewith prior to affliction with said disease or disorder. Such prevention or reduction prior to affliction refers to administration of the nucleic acid regulatory elements, the nucleic acid expression cassettes, the vectors, or the pharmaceutical compositions described herein to a patient
25 that is not at the time of administration afflicted with clear symptoms of the disease or disorder. “Preventing” also encompasses preventing the recurrence or relapse-prevention of a disease or disorder for instance after a period of improvement. In embodiments, the nucleic acid regulatory elements according to any one of SEQ ID Nos: 1-33, the nucleic acid expression cassettes, the vectors, or the pharmaceutical compositions described
30 herein may be for use in gene therapy, in particular endothelial cell-directed gene therapy.

Also disclosed herein is the use of the nucleic acid regulatory elements according to any one of SEQ ID Nos: 1-33, the nucleic acid expression cassettes, the vectors, or the pharmaceutical compositions described herein for the manufacture of a medicament for gene therapy, in particular endothelial cell-directed gene therapy.

Also disclosed herein is a method for gene therapy, in particular endothelial cell-directed gene therapy in a subject in need of said gene therapy comprising:

- introducing in the subject, in particular in endothelial cells or tissue of the subject, a nucleic acid expression cassette, a vector or a pharmaceutical composition described herein, wherein the nucleic acid expression cassette, the vector or the pharmaceutical composition comprises a nucleic acid regulatory element according to any one of SEQ ID Nos: 1-33, operably linked to a promoter and a transgene; and
- expressing a therapeutically effective amount of the transgene product in the subject, in particular in endothelial tissue or cells of the subject.

The transgene product may be any one of the following transgenes can introduced: secretable proteins, in particular secretable therapeutic proteins, including hepatocyte growth factor (HGF), coagulation factor VIII (FVIII), coagulation factor VII (FVII), coagulation factor IX (FIX), coagulation factor XI (FXI), tissue factor (TF), tissue factor pathway inhibitor (TFPI), von Willebrand factor (vWF), ADAMTS13, VEGF, PLGF, FGF, sFLT1, α 1-antitrypsin (AAT), apolipoprotein A-I (apoA-I), insulin, erythropoietin, lipoprotein lipase, antibodies or nanobodies, growth factors, cytokines, chemokines, plasma factors etc.. The therapeutic protein may also be a structural protein. Non-limiting examples of structural proteins, in particular structural therapeutic proteins, including proteins modulating vascular relaxation, vasoconstriction or atherosclerosis, In preferred embodiments, the transgene comprises the nitric oxide synthase (NOS).

Alternatively, the transgene product may be RNA, such as siRNA., or a nuclease such as ZFN, TALEN, CRISPR/Cas9 or similar DNA or RNA editing systems.

Exemplary diseases and disorders that may benefit from gene therapy using the nucleic acid regulatory elements, the nucleic acid expression cassettes, the vectors, or the pharmaceutical compositions described herein include: liver diseases, hemophilia A, von Willebrand disease, microvascular thrombosis, thrombotic thrombocytopenic purpura, peripheral vascular disease, coronary artery diseases, atherosclerotic diseases, stroke, heart disease, diabetes, insulin resistance, chronic kidney failure, tumor growth, metastasis, venous thrombosis, ischemia, tumour growth, tumour vascularisation, cancer and viral infectious diseases such as Ebola, Dengue fever and dengue hemorrhagic fever.

Gene therapy protocols have been extensively described in the art. These include, but are not limited to, intramuscular injection of plasmid (naked or in liposomes), hydrodynamic gene delivery in various tissues, including muscle, interstitial injection, instillation in

airways, application to endothelium, intra-hepatic parenchyme, and intravenous or intra-arterial administration. Various devices have been developed for enhancing the availability of DNA to the target cell. A simple approach is to contact the target cell physically with catheters or implantable materials containing DNA. Another approach is to utilize needle-free, jet injection devices which project a column of liquid directly into the target tissue under high pressure. These delivery paradigms can also be used to deliver vectors. Another approach to targeted gene delivery is the use of molecular conjugates, which consist of protein or synthetic ligands to which a nucleic acid-or DNA-binding agent has been attached for the specific targeting of nucleic acids to cells (Cristiano et al., 1993). In embodiments, the nucleic acid regulatory elements, the nucleic acid expression cassettes, the vectors, or the pharmaceutical compositions described herein may be for use as a vaccine, more particularly for use as a prophylactic vaccine.

Also disclosed herein is the use of the nucleic acid regulatory elements, the nucleic acid expression cassettes, the vectors, or the pharmaceutical compositions described herein for the manufacture of medicament or a vaccine, in particular for the manufacture of a prophylactic vaccine.

Also disclosed herein is a method of vaccination, in particular prophylactic vaccination, of a subject in need of said vaccination comprising:

- introducing in the subject, in particular in endothelial tissue or cells of the subject, a nucleic acid expression cassette, a vector or a pharmaceutical composition described herein, wherein the nucleic acid expression cassette, the vector or the pharmaceutical composition comprises a nucleic acid regulatory element according to any one of SEQ ID Nos: 1 to 33, operably linked to a promoter and a transgene; and
- expressing an immunologically effective amount of the transgene product in the subject, in particular in endothelial cells or tissue of the subject.

As used herein, a phrase such as "a subject in need of treatment" includes subjects that would benefit from treatment of a recited disease or disorder. Such subjects may include, without limitation, those that have been diagnosed with said disease or disorder, those prone to contract or develop said disease or disorder and/or those in whom said disease or disorder is to be prevented.

The terms "subject" and "patient" are used interchangeably herein and refer to animals, preferably vertebrates, more preferably mammals, and specifically include human patients and non-human mammals. "Mammalian" subjects include, but are not limited to, humans,

domestic animals, commercial animals, farm animals, zoo animals, sport animals, pet and experimental animals such as dogs, cats, guinea pigs, rabbits, rats, mice, horses, cattle, cows; primates such as apes, monkeys, orang-utans, and chimpanzees; canids such as dogs and wolves; felids such as cats, lions, and tigers; equids such as horses, donkeys, and zebras; food animals such as cows, pigs, and sheep; ungulates such as deer and giraffes; rodents such as mice, rats, hamsters and guinea pigs; and so on. Preferred patients or subjects are human subjects.

A 'therapeutic amount' or 'therapeutically effective amount' as used herein refers to the amount of gene product effective to treat a disease or disorder in a subject, i.e., to obtain a desired local or systemic effect. The term thus refers to the quantity of gene product that elicits the biological or medicinal response in a tissue, system, animal, or human that is being sought by a researcher, veterinarian, medical doctor or other clinician. Such amount will typically depend on the gene product and the severity of the disease, but can be decided by the skilled person, possibly through routine experimentation.

An "immunologically effective amount" as used herein refers to the amount of (trans)gene product effective to enhance the immune response of a subject against a subsequent exposure to the immunogen encoded by the (trans)gene. Levels of induced immunity can be determined, e.g. by measuring amounts of neutralizing secretory and/or serum antibodies, e.g., by plaque neutralization, complement fixation, enzyme-linked immunosorbent, or microneutralization assay.

Typically, the amount of (trans)gene product expressed when using an expression cassette or vector as described herein (i.e., with at least one nucleic acid regulatory element) are higher than when an identical expression cassette or vector is used but without a nucleic acid regulatory element therein. More particularly, the expression is at least double as high, at least five times as high, at least ten times as high, at least 20 times as high, at least 30 times as high, at least 40 times as high, at least 50 times as high, or even at least 60 times as high as when compared to the same nucleic acid expression cassette or vector without nucleic acid regulatory element. Preferably, the higher expression remains specific to endothelial tissues or cells. Furthermore, the expression cassettes and vectors described herein direct the expression of a therapeutic amount of the gene product for an extended period. Typically, therapeutic expression is envisaged to last at least 20 days, at least 50 days, at least 100 days, at least 200 days, and in some instances 300 days or more. Expression of the gene product (e.g. polypeptide) can be measured by any art-recognized means, such as by antibody-based assays, e.g. a Western Blot or an ELISA assay, for instance to evaluate whether

therapeutic expression of the gene product is achieved. Expression of the gene product may also be measured in a bioassay that detects an enzymatic or biological activity of the gene product.

Also disclosed herein is the use of the nucleic acid regulatory elements according to SEQ ID Nos: 1 to 33, or the nucleic acid expression cassettes, or the vectors disclosed herein comprising said nucleic acid regulatory elements, for transfecting or transducing endothelial cells.

Further disclosed herein is the use of the nucleic acid expression cassettes or the vectors disclosed herein comprising the nucleic acid regulatory elements according to SEQ ID Nos: 1 to 33, for expressing a transgene product in endothelial cells, wherein the nucleic acid expression cassette or the vector comprises said nucleic acid regulatory element disclosed herein operably linked to a promoter and a transgene.

Further disclosed herein is a method for expressing a transgene product in endothelial cells, comprising:

- transfecting or transducing said cells with a nucleic acid expression cassette or a vector disclosed herein, wherein the nucleic acid expression cassette or the vector comprises a nucleic acid regulatory element according to any one of SEQ ID Nos: 1 to 33, operably linked to a promoter and a transgene; and
- expressing the transgene product in said cells.

Non-viral transfection or viral vector-mediated transduction of endothelial cells may be performed by *in vitro*, *ex vivo* or *in vivo* procedures. The *in vitro* approach requires the *in vitro* transfection or transduction of endothelial cells, e.g. cells previously harvested from a subject, cell lines or cells differentiated from e.g. induced pluripotent stem cells or embryonic cells. The *ex vivo* approach requires harvesting of the endothelial cells from a subject, *in vitro* transfection or transduction, and optionally re-introduction of the transfected cells into the subject. The *in vivo* approach requires the administration of the nucleic acid expression cassette or the vector disclosed herein into a subject. In preferred embodiments, the transfection of the endothelial cells is performed *in vitro* or *ex vivo*.

It is understood by the skilled person that the use of the nucleic acid regulatory elements, the nucleic acid expression cassettes and vectors disclosed herein has implications beyond gene therapy, e.g. coaxed differentiation of stem cells into endothelial cell precursors or endothelial cells, transgenic models for over-expression of proteins in endothelial cells or their precursors, etc..

The invention is further explained by the following non-limiting examples

EXAMPLES

Example 1: Identification of endothelial cell-specific nucleic acid regulatory elements

To identify the endothelial cell genes that are highly expressed, we followed several steps. First, we obtained the list of genes that are highly expressed in endothelial cells from the publication of Bhasin et al., 2010 (Genomics 2010, 11:342) showing 104 genes that were identified as endothelial-restricted genes. Subsequently, the specificity and robustness of expression of the endothelial-restricted genes was compared to that of 6 types of endothelial cells (i.e. LSEC: Liver Sinusoid Endothelial cells, HCAEC: Coronary Artery Endothelial cells, HMVEC: Dermal Microvascular Endothelial cells, HUVEC: Human Umbilical Vein Endothelial cells, IEn: Iliac Artery Endothelial cells, RE: Retinal Endothelial cells) based on the Reference Database of Gene Expression Analysis (RefExA). We identified 11 genes (Table 1) from this endothelial-restricted gene list that are highly expressed and specific among these quintessential endothelial cell types. Consequently, these 11 genes were then used for designing the endothelial-specific cis-regulatory elements (CREs: Table 2)

Table 1: Highly expressed genes in endothelial cells

Highly expressed genes in endothelial cells	ENSEMBL Acc. N°
IFI27	ENSG00000165949
ICAM2	ENSG00000108622
VWF	ENSG00000110799
EDN1	ENSG00000078401
ENG	ENSG00000106991
ECSCR	ENSG00000279686
CDH5	ENSG00000179776
PECAM1	ENSG00000261371
HHIP	ENSG00000164161
TIE1	ENSG00000066056
HYAL2	ENSG00000068001

Candidate CREs were selected using the University of California Santa Cruz (UCSC) Genome Browser database based on i) high DNase hypersensitivity sites; ii) high content of epigenetic markers associated with open chromatin (acetylation, methylation); iii) high content of transcription factor binding sites; iv) strong evolutionary conservation. The ideal CREs were expected to exhibit co-existence of predicted motifs together with DNase clusters, high conservation level in vertebrates, and explicit histone modification patterns. Therefore, 28 potential CRE sequences were selected based on those criteria (Figure 1 and Table 2).

Table 2 The EC-CREs sequence for highly expression in endothelial cells.

Gene	EC-CRE	Size (bp)	Sequence
CDH5	EC-CRE1a	76	GCCCTCACAAAGGAACAATAACAGGAAACCATCCCA GGGGGAAGTGGGCCAGGGCCAGCTGGAAAACCTGA AGGGG (SEQ ID NO: 1)
CDH5	EC-CRE1b	277	GGCCGAGGCAGCCGCCACCGCAGGGCCTGCCTAT CTGCAGCCAGCCCAGCCCTCACAAAGGAACAATAAC AGGAAACCATCCCAGGGGGGAAGTGGGCCAGGGCCA GCTGGAAAACCTGAAGGGGAGGCAGCCAGGCCTCC CTCGCCAGCGGGGTGTGGCTCCCCTCCAAAGACGGT CGGCTGACAGGCTCCACAGAGCTCCACTCACGCTCA GCCCTGGACGGACAGGCAGTCCAACGGAACAGAAAC ATCCCTCAGCCCACAGGCACGGTGAGTG (SEQ ID NO: 2)
CDH5	EC-CRE1c	408	GCTTCCTCCTCTGCTACTAATCTGGTCTCACAGACCA TCCCATTTCCTGCTAGCCCACCAGCCGCCTTCCTTGC TCCAATGACACTTCCTGGCCTTGTGCCCTCCTGTTA CCTCCTTTGCCTCCAGAGAGGTTGGAGCAGAGGCTG GGCAGTGCCAGAAATCAGGCATGAAATCCTCAGGGG GACCAAGGAGGCACCAGCCTCCCTCCCACAGTCTCA GCTACCTCTGCTACGGTGACCCCCAGCCCCACCCCT GGGGCCCACAGCTCATGCCTGGCTCACCATTTCCTTT GTTTATGGACCACAGGAACAGTCGTTTTTCAGGGCAGA GTCAACTTCCTCATGGACTGGGAGTACAAAGGGAATT GGCAGATGGTGCCAGGACAGGCCCTGTCCCCATCTG CCACAGC (SEQ ID NO: 3)
CDH5	EC-CRE1d	173	GCTTCCTCCTCTGCTACTAATCTGGTCTCACAGACCA TCCCATTTCCTGCTAGCCCACCAGCCGCCTTCCTTGC TCCAATGACACTTCCTGGCCTTGTGCCCTCCTGTTA CCTCCTTTGCCTCCAGAGAGGTTGGAGCAGAGGCTG GGCAGTGCCAGAAATCAGGCATGAAA (SEQ ID NO: 4)
CDH5	EC-CRE1e	307	CCCAGCTGAGGGCTGGTGCCAGAGCCGTGTCTGCTT GCCCCATCAAGAGGTGGGAGGGATTGATCCACCTTC CTGCCCCACAGATGGTGCAGCCTCCAACCTATTGTTT TCCAGGACGCTTCGGTGGAGAGCACAAGGAATGTAG GGTCTAGAAACAGGAAGCCCTGGCTTCCGCTGGACA AGGTTTCCTCCAGACTCAGGCCTGCCCTCCAGACAA CAAGGCAGGGCCCTTGGTCCCACCCTGCCCTGCCTG

Gene	EC-CRE	Size (bp)	Sequence
			GCTCACTGGGCCACCCCAAGGAAGGCCTTGCCCTCTCTGGGCTTCTGCATGTGA (SEQ ID NO: 5)
CDH5	EC-CRE1f	450	AGGTATCCACCAAGGGGGCCCAAGAGCTGCTGAGCCCCTGAGCAGCCCTACCAATTTTCAGCTTATGGTGGTGAGGGGTAGGGGAGGTGTATACTGGCCTGGAAGGGGTTAAGCTGCCCCGCCTGCAGCCTCAGCCTGAGCTATTGTGTTGCCAAACAAGGGCCGGACATGAGGGCAGGAAGCCAGCAGGGGGCCACACATTTTCTGCAAAGTTGGATGATTCACTGCTGACTTGGGGACACCCAGGGGGACAGAGGGGACACCATCCCAGGAAGAATCTTAGGCTCATTTTGCCCACATGGACCCATGACTGTTCCCTGTATCCTCTCTCTGCACCCCCTCAGTCACACTGAAGCAACTATGAGAATCCCATTGACAGATGGGACCATCGAGGCTGAGGGGAAGCTGTGCAGCCAGTCCAAGGTCACACAACCAACACAAGGTAGAAGCAGGG (SEQ ID NO: 6)
ECSCR	EC-CRE1a	123	AGGGCCCCTGGAGCTGGTCCCAATGTGTTTCCTTCTATTCTTTTGACAGGAAGCTCCTGGAGAGCCAGTCCCCACCCCATCCCGCCCCAGCACTCCCTCTCTCTTCTCCACTATGGACAGAG (SEQ ID NO: 7)
ECSCR	EC-CRE1b	499	TAGGGCCTCTCTGAAAGATGTGGGGAGTCCTATCTGCATTGGGATCCCTGAGGAGGGAGAGGAATGTGGAGATTTCAGGGTCCAGGGAGCATGGGTGACTGGTGGGCTGGGCTTCCAGGCTGAATCATGGGAAAGGAGAACCTGTCTGAAACAGTACTGGGCGGGATTGGTGTTAGATTCAGGAAAACCCCCAGGCGGTCTGTGGTGGAACCTGATGGACCCTCAGAAGGGAAGAGAATGGGGATGGGGCCAGGTTGCCATGGTTGGTCATTGTGCATAGGCACTAGAGGCCATGCTGGGTGGGCACAGTCGCTGCTGCAGCCTCACATCCTCATCTGGACATGGCTGAGCAGGGGCCCTGGAGCTGGTCCCAATGTGTTTCCTTCTATTCTTTTGACAGGAAGCTCCTGGAGAGCCAGTCCCCACCCCATCCCGCCCCAGCACTCCCTCTCTCTTCTCCACTATGGACAGAGCCTCCACTGAGCTGCTGCCTGCC (SEQ ID NO: 8)
EDN1	EC-CRE1a	455	GAGACATAAAAGGAAAATGAAGCGAGCAACAATTAAAAAAAATTCCCCGCACACAACAATACAATCTATTTAAAC TGTGGCTCATACTTTTCATACCAATGGTATGACTTTTTTTCTGGAGTCCCCTCTTCTGATTCTTGAACCTCCGGGGCTGGCAGCTTGCAAAGGGGAAGCGGACTCCAGCACTGCACGGGCAGGTTTAGCAAAGGTCTCTAATGGGTATT TCTTTTTTCTTAGCCCTGCCCCCGAATTGTCAGACGGCGGGCGTCTGCCTCTGAAGTTAGCAGTGATTTCTTTT CGGGCTGGCCTTATCTCCGGCTGCACGTTGCCTGT TGGTGACTAATAACACAATAACATTGTCTGGGGCTGG AATAAAGTCGGAGCTGTTTACCCCCACTCTAATAGGG GTTCAATATAAAAAGCCGGCAGAGAGCTGTCCAAGTCAGACGCGCCTC (SEQ ID NO: 9)

Gene	EC-CRE	Size (bp)	Sequence
ENG	EC-CRE1a	153	TGTCCACTTCTCCTGACCCCTCGGCCGCCACCCAG AAGGCTGGAGCAGGGACGCCGTGCTCCGGCCGCC TGCTCCCCTCGGGTCCCCGTGCGAGCCCACGCCGG CCCCGGTGCCCCGCCCGCAGCCCTGCCACTGGACAC AGGATAAGGCC (SEQ ID NO: 10)
ENG	EC-CRE1b	125	GGGCCCCCACCAGTGACAAAGCCCGTGGCACTTC CTCTACCCGGTTGGCAGGCGGCCTGGCCCAGCCCCT TCTCTAAGGAAGCGCATTTCCTGCCTCCCTGGGCCG GCCGGGCTGGATGAGCC (SEQ ID NO: 11)
ENG	EC-CRE1c	498	GGGATGGGAGGGTGGGGTGCTTGGGGAGACAAGCC TAGAGCCTGGGCCCTCCCACCCCACTGCCTCCCCC ATCCCAGGGCCCCCACCAGTGACAAAGCCCGTGG CACTTCCTCTACCCGGTTGGCAGGCGGCCTGGCCCA GCCCTTCTCTAAGGAAGCGCATTTCCTGCCTCCCTG GGCCGGCCGGGCTGGATGAGCCAGGAGCTCCCTGC TGCCGGTCATACCACAGCCTTCATCTGCGCCCTGGG GCCAGGACTGCTGCTGTCACTGCCATCCATTGGAGC CCAGCACCCCCTCCCCGCCCATCCTTCGGACAGCAA CTCCAGCCCAGCCCCGCGTCCCTGTGTCCACTTCTC CTGACCCCTCGGCCGCCACCCAGAAAGGCTGGAGC AGGGACGCCGTCGCTCCGGCCGCTGCTCCCCTCG GGTCCCCGTGCGAGCCCACGCCGGCCCCGGTGCCC GCCCGCAGCCCTGCCACTGGACACAGGATAAGGC (SEQ ID NO: 12)
HHIP	EC-CRE1a	136	AGCGGTGACGTCAAGGGGCGCGCTGTGGCAGCACC TCCCCGCGCGCTAGTTAAAAAGAAGAAGAAAGAGG GAACGAAACATGAGAGGCTGTGTGAGAAGCTGCAGC CGCCGGCAGAGGAGACCTCAGCATCATCT (SEQ ID NO: 13)
HHIP	EC-CRE1b	574	CTGGGCGGGGGCGCGCGAGAAGCGGTGACGTCAAG GGGCGCGCTGTGGCAGCACCTCCCCGCGCGCTAGT TAAAAAGAAGAAGAAAAGAGGGAACGAAACATGAGA GGCTGTGTGAGAAGCTGCAGCCGCCGGCAGAGGAG ACCTCAGCATCATCTAGAGCCCAGCGCTGGCCCTGC CTCCGCCTGCCCCGCCGCCGCGTCGCCGTTTCTGT TCCTGCTACTGTCCCACCTAAACAACCTCCCGTTACAC GGACAAGTGAACATCTGTGGCTGTCCTCTCCTTTTCT TCCTCCTCTTCCAACCTCCTTCTCCTCCTCCCACTTCC CAGCCGCAGCAGAAAGCCCCCAACCCAACTGACACT GGCACAACCTGCAAACGGGTGTCATCCGCACAACCTTAT CTCGCTCCTCGGGCTCCCCTAAGGCATTGGACCCAT CGCCGCGTCTTTTATTTTTTGTCAAAGTTGCATCGCTG TACATATTTTGTCCCCGCCACCTCCCTCTGTCTCTG GAGTGCCCTACAGCCCCGCAAACCTCCTCCTGGAGCT GCGCCCTAGTGCCCCTGCTGGGCAGTGGCGT (SEQ ID NO: 14)
HYAL2	EC-CRE1a	170	AGGGAACCTCCCTGTGCTGGGCCTACCCAGCTGACCC CATCGCTGGAAACAATGGGGGTGAGGCAACACTTCC CACTCTCTCCCGCCGGGCTGTGCTCACTTCCTTCCT GCTGGCTGCCTGAGGAAGTGTCCCTGCCCTGGGACA

Gene	EC-CRE	Size (bp)	Sequence
			GTCTGGCCTAGCCTTTGTTTCCCCG (SEQ ID NO: 15)
HYAL2	EC-CRE1b	470	GACAGGCTTCTGAGTGTAGGGAGCTGGTCTGCCAGT CTTTCGGAGGTTTGAAGTTGTCAAGGCTAGGGCAGG ATCACCATATCCAGCCTGGACTTGCAGTTCTGTGGGG TGCCTCCCCATACCCCCATAAGATGCCAAACATGAGG CCCTGTCATCCTCCATGGTCCCCCTCTACTGGCTGTT CAAGGCCAGGGCTCTCCCATGCCAGATAGCATCCT GTCTCCTACCACCACTGTCCCAGCCTGAGGGAACTC CCTGTGCTGGGCCTACCCAGCTGACCCCATCGCTGG AAACAATGGGGGTGAGGCAACACTTCCCCACTCTCTC CCGCCGGGCTGTGCTCACTTCCTTCCTGCTGGCTGC CTGAGGAAGTGTCCCTGCCCTGGGACAGTCTGGCCT AGCCTTTGTTTCCCCGGGGGTCCCCACCCATGGAGC TTTCAAGGCTTCTGGCCCCTGTGAAGCCAGCACA (SEQ ID NO: 16)
HYAL2	EC-CRE1c	602	CAGTGGAAAAAACGGACTCAGCTACTGGAAGTCCC CCCGACCCTCCCCCAAGGCTAGTTCCCTTCTTGGG CACCTGCTCTGGGGGACCATCAGCTGAACGACCCCC AAGTATTTTGACTCCCAAAAGCACCAACACCTGACCC CATCCTCTCACACCCTACTGGATTTGAGGATGGGCCC CAATCCTAGGGAAGGAGTGAAGAGGTTCCCTAGTGT TGGAAGCTGTGGGTGTGGGGGAGATTGGCACCTGAT CCTGAGCCCATAGCCTTCCTGTCACCTGGCGCAGCT GGCGGGGCCAGATCCTACTCGGGAAGGGTGGGGAG GGCAGCCAGCCAGCAGGGCATTCTGGAGGGAAACA GGGTCAAGGCGATCTCCTCCCCCACGCCTGTTCTG GCCCTTTCCTCTCAGGGGGGCAGCAGGAAGTGAGGAG AAAGGGCTGGGATGGGAGGCGGGAGCGGATGGGAG GGAATGGGGTTTATCAAGTCCTCGGCGAGCTGCCCA ACGGGCAGCAGCTGGCGCAAGTAGCCTAGCTGGAG AGGCTCACCCCAGGAAGGAGGGAGGCCACCGACCT ACTGGGCCGACGGAAGTCCCACACAGGTGA (SEQ ID NO: 17)
ICAM2	EC-CRE1a	265	CTTGCATAGATGGCCAGCGTTCATACTTTCTGCTTGT TTGTACAAAGTCATTCTTCTAGAGTAATTGTTGTAAAA TTGCTAGGCAAGGTGGCAGGTCTGATAAGATTTGATG ACGTAATGGCTCTTAGTCGCTAATAAGAGGCTTTTGT GGAGTGGCGTGTACAGCCAGCGAAGGCTCAGCTCT GTGATCTTCGCCTGCCTCACTTGGGGGACCAGAAGG CAGCTTGTCTTGGAAGTGCCTCATTCACAGAAGACCC CATTGAG (SEQ ID NO: 18)

Gene	EC-CRE	Size (bp)	Sequence
ICAM2	EC-CRE1b	545	TTTACCTAGCATGATCTTGGCAGTTCAAAGAGGAATG TGCCAGAAACCAAGCAAAGAAGAAAAAGAAATAAATG GAAATGGAAAGTGATCTGCTCAGAGGCCACAAAGTT GAGGGAGGAGGTTTCCAGAGTGGGATTTGGCCCAA TGTTGCCTGAGGAAGTACGTAAAGGGGTCTCAACTCT GGCTACACAACAGAACAGCAGGACTGTGTGTGCAGC TCACGAAGTGGGTACACAGGGTAATCTGCAAGTTCTG GGTGGATCTAGCACCTGGATTGTTAAACTTGCATAG ATGGCCAGCGTTCATACTTTCTGCTTGTTTGTACAAA GTCATTCTTCTAGAGTAATTGTTGTAAAATTGCTAGGC AAGGTGGCAGGTCTGATAAGATTTGATGACGTAATGG CTCTTAGTCGCTAATAAGAGGCTTTTGTGGAGTGGCG TGTCACAGCCAGCGAAGGCTCAGCTCTGTGATCTTC GCCTGCCTCACTTGGGGGACCAGAAGGCAGCTTGTC TTGGAAGTGCCTCATTACAGAAGACCCCAT (SEQ ID NO: 19)
ICAM2	EC-CRE1c	554	ATGGCAGCTGGCAGGTGCCTTCACGTCCAGGGTTTC CAGAGAGAAAGCATCTCTCCTCCGCAGAGACCCTCC CACGCTCTCCCTCCCTCAAATTAGTGCATCTACATAG ACCGCCCTCCTTATAAACAGTCTCTCAGGGGATCCTA GCCCATTCCAAATCTACCTGTGATTGCAGAATCGCAA GGAATGTGATTTACCGCAGATCGCGGGGGCGTCGTGT CTTTTAGGGGACCTGCTCACTTTGGCCACTAGGTGG CGGGCAGTGCAGCCCCTGCTCCTGTGACCCCTGAGC GTTCAGCGTTTCCGCCGCTCCGCCCCACTCCGTAG GGGGAGCTGATGAGATGAGGTTGAGGTCCAGGAAGA CGTCAAGGGCTTGGTTTTGTAAACAACTCCATTCTC GCTCGCTGATAAGTTTTCTAAGTGATGCATATTCACAA CCTTGTCCCATCCAAGGACCCAAGAATTAACACATTA CATAATATGGACAGCCCCCTCCTGTCCAACGGGCAT GATTTTGGGGTCTGATATTCTGTGGATCTGTGCAATA GTCAAC (SEQ ID NO: 20)
IFI27	EC-CRE1a	450	AGACTTTTTTTGAAAAACGGAACATCTGCCTATCGCA AGGACTACTATTATTCTGAAAATCACCTTCTTCATTAG AAAGTAATATTTATCATTTTATTATAGAACTTTGATCTT ACTTCTTGTGACTTCATTCTGCGTAGAGCACACTCCC ATCCTTGAATTAAATGACAAAGCATTTTATATTAAGTG ACAATGACTGATGCCATGGGCAAATCCTATTTCTGTA AATAACTGAATTTTCTTCTGGACTGCGCATGAGGGGA GAAAGATGTCTGCAGTTTCGGTTTCCTGGAAAATGAA ACCTATCTCATTTGTTGCCTGTGTCAAGGGGCAGTGC TTCAGTCGGGGTGGAGCTGCTTAAAAGGCCTGGGAT CACACCCTTTGGGAACACATCCAAGCTTAAGACGGTG AGGTCAGCTTCACATTCTCAGGAACTCTCCTTCTTTG GGT (SEQ ID NO: 21)

Gene	EC-CRE	Size (bp)	Sequence
IFI27	EC-CRE1b	570	GAGACTTTTTTTTGA AAAACGGAACATCTGCCTATCGC AAGGACTACTATTATTCTGAAAATCACCTTCTTCATTA GAAAGTAATATTTATCATTTTATTATAGAACTTTGATCT TACTTCTTGTGACTTCATTCTGCGTAGAGCACACTCC CATCCTTGAATTAAATGACAAAGCATT TTTATATTA ACT GACAATGACTGATGCCATGGGCAAATCCTATTTCTGT AAATAACTGAATTTTCTTCTGGACTGCGCATGAGGGG AGAAAGATGTCTGCAGTTTCGGTTTCCTGGAAAATGA AACCTATCTCATTTGTTGCCTGTGTCAAGGGGCGAGTG CTTCAGTCGGGGTGGAGCTGCTTAAAAGGCCTGGGA TCACACCCTTTGGGAACACATCCAAGCTTAAGACGGT GAGGTCAGCTTTCACATTCTCAGGAACTCTCCTTCTTT GGGTAAGACTGGGAGGGTGGGCAGGAGCTACCCTT CCCGTGGCCCCGGACCTTGGGTGGGCTGTGGGCTC AGGGAGCGGAGGGGAGGCCTTAAGCATCCACTCTCT GCCCGGTGTTTTTGTTC (SEQ ID NO: 22)
IFI27	EC-CRE1c	513	AGGTGGGGATGAGGGGCTAAGTATGAACCAAGGAGC TAGAAATACAGCACTGGAAGCTGGAAGCAGGGGGCT TGGAGACTGGGAGCTGGAGTGCGTGTGGGCAGGGT GTGGCAGCAGCCGGCAGAGGCCATTTCCCCTTGGCA GAACATTCACCATGTGACCCTGAGCATGTCTTTGAAC TCCTCTGAGCTCCTGTTTCCTCTCCAGAGAAAAGGCT GGTAATGCCCATTCAGGGTTATGGTCAGGATTGCATA GGGTGAAACAATAGAGATTGAACACAGTAGACATGAA AGAGATGCCAGGGCTCAGCTCCCTTTGGTTTAGTTGC TTCCAGTGTGCTCTGTGGCAACACCACGGAGCCCTA GAGCTGTCTCTTTGAGCCGCTCTGAATGTGCCTCTTA CATAATCTCCTGGGCAACATCTGCTCCCCTAATGAGA TTTGCTCCCCAGCAAAGATAAGAACTTGCCAACCAC TCCCCTGGTCCAGCATTTGGCCAAGGCAGACACTGA GG (SEQ ID NO: 23)
PECA M1	EC-CRE1a	217	ACCTCACTCAATGCATGGAAGTTGACACAATGGCTCA ACATTAGCGTTGGGCTGATTCATCATTTGGCTGTTGA CACCAGCCTCTGGCCCAGCCAGGACAGAAAAAGGGC CCCTGAGGAACTTCTGGCTCTGTTCCCTCTATGGGG GAGGGGCGAGTGGACTTGTGATAAGACAGGGTGTTAG GGTGAGGTGGACTTGGGGAAACAGGATATTTCTAA (SEQ ID NO: 24)
TIE1	EC-CRE1a	97	GGGGGGAGGGGAGACCCCAGAACAAATGTCCCCCAC CCCACCCCCCTCCTCAATAGGCGGAAGCCACTGGCT TCCTCCCTTTCCTGCCTCCTGCCTCC (SEQ ID NO: 25)
TIE1	EC-CRE1b	427	GTGTGTTTGTGCCGGGGGGAGGGGAGACCCCAGAA CAATGTCCCCCACCCACCCCCCTCCTCAATAGGCG GAAGCCACTGGCTTCCTCCCTTTCTGCCTCCTGCCT CCTTTGTGCCAGCAAGACTGAGTACTGGAGAGAGAC AGGGGATGGGAAAAATCAGTCCAGCTGTCCCCAGGT CTGCCCTTACCATAACCTTCCCCCACCTCAAGTGAC TCCTCCCAGGCCACACCCATCCCCAGCCTTGTGGGG GCCAGATTGGGGGGCCTAGAGGCTCAAAGGCAGAAT GAGTCCTCCCACCCCCTACCCTGCCACCCCTCCCAC CCAAGCCACCTCATTTCTCCTCCTCCCCAGCACCGA

Gene	EC-CRE	Size (bp)	Sequence
			CCCACACTGACCAACACAGGCTGAGCAGTCAGGCCC ACAGCATCTGACCCCAGGCCCAGCTCGTC (SEQ ID NO: 26)
VWF	EC-CRE1a	119	CTACAAAGCTTTATCAGCTTGGAGGTACTTCTAATAC CATTTCTTTTCATTGTTTCCTTTTGGTAATTAAAAGGA GGCCAATCCCCTGTTGTGGCAGCTCACAGCTATTGT GGTGGGAA (SEQ ID NO: 27)
VWF	EC-CRE1b	385	CTGCCAGGAGGTCTCCCTCCAACTCTACAAAGCTTT ATCAGCTTGGAGGTACTTCTAATACCATTTCTTTTCAT TGTTTCCTTTTGGTAATTAAAAGGAGGCCAATCCCCT GTTGTGGCAGCTCACAGCTATTGTGGTGGGAAAGGG AGGGTGGTTGGTGGATGTCACAGCTTGGGCTTTATCT CCCCCAGCAGTGGGGACTCCACAGCCCCTGGGCTAC ATAACAGCAAGACAGTCCGGAGCTGTAGCAGACCTG ATTGAGCCTTTGCAGCAGCTGAGAGCATGGCCTAGG GTGGGCGGCACCATTGTCCAGCAGCTGAGTTTCCCA GGGACCTTGGAGATAGCCGCAGCCCTCATTTGCAGG GGAAGGTATGGCCTTTGGAA (SEQ ID NO: 28)

Alternatively, the VISTA enhancer browser (<http://enhancer.lbl.gov>) was also applied, a central resource for experimentally validated human and mouse non-coding fragments with gene enhancer activity. This also provided the predicted DNA elements associated with high expression in blood vessels. The predicted sequences from VISTA were selected based on the validated data using mouse embryonic staining. Up to 3 VISTA sequences were selected from these validated data. However, since the DNA fragment sizes were too large to be accommodated into a viral vector, the selected sequences were further trimmed down or separated into sub-fragments using the UCSC genome browser using the aforementioned criteria. This resulted in 5 CREs derived from the 3 selected VISTA sequences (Table 3). All of the endothelial-specific CREs sequences were further validated both *in vitro* and *in vivo* to investigate their specificity and robustness in endothelial cells.

Table 3: The EC-CREs sequence for highly expression in endothelial cells from VISTA Enhancer browser.

VISTA code	EC-CRE	Size (bp)	Sequence
------------	--------	-----------	----------

Hs185 9	EC-CRE- V1	517	GCCATTGGCTGGTCCTTCACTGACAGCAGAAACT TGGCCAATGGCAATCAATCAGGGGGGCCCCGCGCT GCCTTAAATACCAGCAGAGCAAACAGCCTCAGAC AAAGCTGCGCCGTGTATCAATTACCGAGAGGCT GCGTGCTCCTCTGGGCGGAGGGAGCCGGAGCG AGCGGCCAGGGCTGCTGCCCCAGCTGATAAGG GCCCCGATTGTTCGGGGACAGCTGGCAGCCCCGA TAAGGGCCTGCTCGCCCCGAGATAATGGCAGTGG GCAGGCGCCTCGCGGCAGTTTAGAATTTCTTGG GTCTCCAAGAAAGGTTCTATTAAGCCCACTGACC CCAATTGAATATTAATTAGCTAATTAACGGATTTA TTGTTCCACGCCATTTCTGGAGAGGGCCATTTTTTT TCGAGTGCCATTATTTTTGTAAATGATTTTTTCGCA TTGTTCATAATTGAATCTTTGCAGCTGCCAGCATC TTCTGCATGATTTGGCAAAAAAAGGAAGCAGAA GCACTTAGGGTT (SEQ ID NO: 29)
Hs217 9	EC-CRE- V2	511	CCCCTAATAAACAGGAAGGCATCCGCGCCATTA GTATCCATCCTTTTCAGAGCATCTGAGACCTGTC TGGACCATCAAAGCCATCCCCAGCCCCCAGGAG CCTACTGGAGGAGACACCAGCCTCGCCAAAACA ATTCTCCATTGTGTTCTTCCCCTTAGAAATCATGG GTTTGTAAACAGGCCCTTACATTTCAGCAGGTCC TGCCCTGGCTTTGTGCTGGTGTGTTGTTTTTTCTT CCCTGAACAATGTCCTTTCCAGTAGGGCCAGCC GTTACACCATTTGTCTGAGACCCTTGGACTACAG GAAACATCACCAGATTCTTATCAGTTGGGGGCAG GAGTGGGGGGGTGAACAGATGAGATCATGTCCA CAGAGCAAGTGGCTGGTGTGCCACGTCATTCCC CATGCCTTCATCTGTGAGAGCAGAGCCCGCTCG CCCTCCATCAATCTGGGCTTCATGTGTCCAGAGT CCAGTCCCCTCTATTTGGTGGTAGACACCTGGAC TCTGTT (SEQ ID NO: 30)
Hs217 9	EC-CRE- V3	390	TCCTACAGTTATGGGTCTGTCTTTCCTGCTAGAT CAGAAGCTCCAGGGACCTGTCTGTCTTATTAACC TTTCTTCCCCTCATGATGCCTGGCCCAGGACTCC ACGTTCCAGAGGCAGTTTAATGTCTACAGAACTGA TGGATGCTCCATACCCTGTATTCATAAGCCTGTG TTTTGCTGCCAAACACCAGAGGGCACTGTTAGCA TGTCGATGAAGATTATAAACCCTCAGACCTGGAA GGCTGGGAAAGGCTTATGAAAATCTTGCTTCTGT TTTGGGATTACATAAGACTTATTGCGCCTCAATTG TTCTAAACACATCTGTTTCGAGTTTATTCATGAGGC ACGTTCCCTGTTGGGGTTAGAGATGAGTTTGAAAG CTCCCCGTCACAGG (SEQ ID NO: 31)
Hs188 2	EC-CRE- V4	573	AGGGACTGTTGGGCAGCCCCAGACTGGCACAGG TGGATCGGGTGCCTAGGCAGGGGGTGGTGAGTT ATGGCGCAGCTGTCTTGGTGGCTGGGGGGGAGCA GGGATAAGGGTGGACTTCTTAGTGACCGCTCTCT GCCCCAGGAGGTAGAGTCCTGGGGGGCTGGGCT GGCCTGAGAGACGCCCCCTCATCCTTTCCAGGG TGAGGTACGAGGGCTCCGCCCCCTCCTGATATC ACCAGGCCTAGGGCAGCATCCTGATGGGGGAG GGGCAAGTGACCCGGGGCCCTGGACTGCAGGAA CAGCCCCTCCTCCACTGGTGGAGTTCCCACTTC

			CTGCGGAAGGAACTATGTTAGAAGTTGTGTATAT GGGGTGGGGGTTGGGTGTGGGTGGCGGGGGG CCTGGGTGGGGTCCACTGAGTCGCCTCCCCTGT CTCCCTGCACTTCCTCCTGGAGGAAATGGGGAC AACAGGATGAAGTGAGGGCCTGCTGAGCCCAGG GCTGCCACCTGGGAGTGAAGCCGGGGCAGGCT GCAGGGTCCGGGGCCCTTCTGTGTGGGCAGGTG GAAGTGGTGGGGATGCA (SEQ ID NO: 32)
Hs188 2	EC-CRE- V5	441	GGGGAGAGGGTGTGGGGTGGGGTGGGGAGAG GGTGTGGGGTGGGGTGGGGAGAGGGGATGGGA TGGCATGGGGGGATGTGGCAGTGAGGAGGCTG GGCCCTTGGAGCTGCCGAGTGCAGGGGGCCTGG AGGACTCCGGGAAGGCGTCCTAGTGCATCAAGC GTGGGCTTGGCCTGCTTGGGTCTCCCCTCCTGG CCCCCCTAGCAATGGGCGGACTTGGGCCCGCTC TGGGAGGATTCCAGGAACGGCTCCTGCCTGGTT ATAAATAGACTTCTCCGAAAGGCCTGGGGCTGTG CCAGCTGCAGCAGGTGCCTCCCAGGCCCGGCC AGAGGGCCCCAGGCAAGGGGGTGGAGCCCGGG TGGGGGTGATGAGGATGCTGGGGTCCACTTTTG TAGCGCCAGAGGCGACGGGCTCTGTCTGGTTGT AGCATCACAGAGCTTGAT (SEQ ID NO: 33)

Example 2: Lentiviral transduction of human endothelial cells *in vitro*

To validate the potential endothelial cell-specific CREs, we first validated a robust
5 endothelial-specific promoter. The selected endothelial cell-specific CREs were cloned
upstream of this promoter. We identified several human endothelial-specific promoter
such as the human CDH5 promoter (1,303 bp) and human EDN1 promoter (455 bp) (The
sequence of the CDH5 was obtained from Genecopoeia (<http://www.genecopoeia.com>)
and for EDN1mini promoter, we selected the promoter sequence using the same concept
10 as we select the CREs from UCSC). In addition, we also identified several endothelial
promoters that are commercially available (Invivogen, USA) such as the as ENG promoter
(888 bp), FLT1 promoter (1,037 bp), and ICAM2 promoter (399 bp). The sequences of
these promoters are provided in Table 4. The endothelial-specific promoters were cloned
into the lentiviral vector plasmid upstream of the FVIII or GFP reporter gene (Figure 2).

15 Nine different lentiviral constructs were generated (designated as:

- 1) pLVX-CMV-GFP (SEQ ID NO:40),
- 2) pLVX-CMV-FVIII (SEQ ID NO:41),
- 3) pLVX-hEDN1mini-FVIII (SEQ ID NO:42),
- 4) pLVX-hEDN1mini-Kozak-FVIII (SEQ ID NO:43),
- 20 5) pLVX-hICAM2-FVIII (SEQ ID NO:44),
- 6) pLVX-hICAM2-Kozak-FVIII (SEQ ID NO:45)
- 7) pLVX-hICAM2-Kozak-Luc2 (SEQ ID NO:46)

8) pLVX-EC-CRE-h(uman)ICAM2-Kozak-Luc2 (SEQ ID NO:47)
9) pLVX-EC-CRE-h(uman)ICAM2-Kozak-FVIII (SEQ ID NO:48), by conventional cloning.
pLVX can also be a pCDH backbone or another (lenti)viral backbone.

5 **Table 4** Endothelial-specific promoter sequences

Promoter	Host Species	Size (bp)	Sequence	Source
CDH5	Human	1303	GCTTGCCCAGCTATATAATAAAACAAGTTTGGGACTTCC CAACCATTACCCCATGGAAAAACAGAAGCAACTCTTCAA AGGACAGATTCCCAGGATCTGCCCTGGGAGATTCCAAA TCAGTTGATCTGGGGTGAGCCCAGTCCTCTGTAGTTTTT AGAAGCTCCTCCTATGTCTCTCCTGGTCAGCAGAATCTT GGCCCCCTCCCTTCCCCCCAGCCTCTTGGTTCTTCTGGG CTCTGATCCAGCCTCAGCGTCACTGTCTTCCACGCCCC TCTTTGATTCTCGTTTATGTCAAAAGCCTTGTGAGGATG AGGCTGTGATTATCCCCATTTTACAGATGAGGAACTGT GGCTCCAGGATGACACAACTGGCCAGAGGTCACATCAG AAGCAGAGCTGGGTCACTTGACTCCACCCAATATCCCT AAATGCAAACATCCCCTACAGACCGAGGCTGGCACCTT AGAGCTGGAGTCCATGCCCGCTCTGACCAGGAGAAGC CAACCTGGTCCTCCAGAGCCAAGAGCTTCTGTCCCTTT CCCATCTCCTGAAGCCTCCCTGTACCTTTAAAGTCCAT TCCCACAAAGACATCATGGGATCACCACAGAAAATCAA GCTCTGGGGCTAGGCTGACCCCAGCTAGATTTTTGGCT CTTTTATACCCCAGCTGGGTGGACAAGCACCTTAAACC CGCTGAGCCTCAGCTTCCCGGGCTATAAAATGGGGGTG ATGACACCTGCCTGTAGCATTCCAAGGAGGGTTAAATG TGATGCTGCAGCCAAGGGTCCCCACAGCCAGGCTCTTT GCAGGTGCTGGGTTCAAGTCCCAGAGCTGAGGCCGG GAGTAGGGGTTCAAGTGGGGTGCCCCAGGCAGGGTCC AGTGCCAGCCCTCTGTGGAGACAGCCATCCGGGGCCG AGGCAGCCGCCACCGCAGGGCCTGCCTATCTGCAGC CAGCCCAGCCCTCACAAAGGAACAATAACAGGAAACCA TCCCAGGGGGGAAGTGGGCCAGGGCCAGCTGGAAAACC TGAAGGGGAGGCAGCCAGGCCTCCCTCGCCAGCGGG GTGTGGCTCCCCTCCAAAGACGGTCGGCTGACAGGCT CCACAGAGCTCCACTCACGCTCAGCCCTGGACGGACA GGCAGTCCAACGGAACAGAAACATCCCTCAGCCCACAG GCACGGTGAGTGGGGGCTCCCACACTCCCCTCCACCC CAAACCCGCCACCCTGCGCCCAAGATGGGAGGGTCCT CAGCTTCCCCATCTGTAGAATGGGCATCGTCCCCTCC CATGACAGAGAGGCTC (SEQ ID NO: 34)	Genecopoeia
EDN1	Human	455	GAGACATAAAAGGAAAATGAAGCGAGCAACAATTAAAAA AAATTCCCCGCACACAACAATAATCTATTTAACTGT GGCTCATACTTTTCATACCAATGGTATGACTTTTTTCTG GAGTCCCCTCTTCTGATTCTTGAAGTCCGGGGCTGGCA GCTTGCAAAGGGGAAGCGGACTCCAGCACTGCACGGG CAGGTTTAGCAAAGGTCTCTAATGGGTATTTTCTTTTCT TAGCCCTGCCCCCGAATTGTCAGACGGCGGGCGTCTG CCTCTGAAGTTAGCAGTGATTTCTTTTCGGGCCTGGCC TTATCTCCGGCTGCACGTTGCCTGTTGGTGACTAATAAC ACAATAACATTGTCTGGGGCTGGAATAAAGTCGGAGCT GTTTACCCCCACTCTAATAGGGGTTCAATATAAAAAGCC GGCAGAGAGCTGTCCAAGTCAGACGCGCCTC (SEQ ID NO: 35)	UCSC

Promoter	Host Species	Size (bp)	Sequence	Source
ENG	Human	888	CGCCTTGCTGTGCCACTTTGGGACTTCCCTCCCTAGCC TGAGCTTCAGTTTTCTGCCTGTTAGGCAGCCCCATGT CAACTGCACTTAGTAGGCCGGGTTTGATGCCCCGACAAG ACGTGAAGTGGTGGAGGTGGGCAGGATCCCAGCGCTA CCATCTTCTTGAACCACTGATCTCAACACATCGGATTTT TGTTTCCTCATCTGCAAAATGGGATCAGTGAGCTCAGGT GGGTCACAAATTCTACAGGAACACTTTAGCCAAGCCC GGCCCCCTGAAAGTTCCCTCGGTGGGCTGTTAGGGT GATTGTTTTTCTATCTGTGGGGCTCCCTGATGCGTCCCAC CCACCAGCCTTGAGAGGGTGGGATGGGAGGGTGGG GTGCTTGGGGAGACAAGCCTAGAGCCTGGGCCCTCCC ACCCCACTGCCTCCCCCATCCCAGGGCCCCCACCC AGTGACAAAGCCCGTGGCACTTCTCTACCCGGTTGGC AGGCGGCTGGCCCAGCCCCTTCTCTAAGGAAGCGCA TTTCCTGCCTCCCTGGGCCGGCCGGGCTGGATGAGCC GGGAGCTCCCTGCTGCCGGTCATACCACAGCCTTCATC TGCGCCCTGGGGCCAGGACTGCTGCTGTCACTGCCAT CCATTGGAGCCCAGCACCCCCCTCCCCGCCCATCCTTCG GACAGCAACTCCAGCCCAGCCCCGCGTCCCTGTGTCC ACTTCTCCTGACCCCTCGGCCGCCACCCAGAAAGGCTG GAGCAGGGACGCCGTGCTCCGGCCGCTGCTCCCCT CGGGTCCCCGTGCGAGCCCACGCCGGCCCCGGTGCC CGCCCGCAGCCCTGCCACTGGACACAGGATAAGGCC AGCGCACAGGCCCCCCACGTGGACACC (SEQ ID NO: 36)	Invivogen Inc.
FLT1	Human	1037	TTTGCTTCTAGGAAGCAGAAGACTGAGGAAATGACTTG GGCGGGTGCAATGCGGCCAAAAAAGACACGGACA CGTCCCCTGGGACCTGAGCTGGTTCGCAGTCTTCCCA AAGGTGCCAAGCAAGCGTCAGTTCCCTCAGGCGCTCC AGGTTCACTGCCTTGTGCCGAGGGTCTCCGGTGCCTTC CTAGACTTCTCGGGACAGTCTGAAGGGGTCAGGAGCG GCGGGACAGCGCGGGAAGAGCAGGCAAGGGGAGACA GCCGGACTGCGCCTCAGTCCTCCGTGCCAAGAACACC GTCGCGGAGGCGCGGCCAGCTTCCCTTGGATCGGACT TTCCGCCCTAGGGCCAGGCGGCGGAGCTTCAGCCTT GTCCCTTCCCCAGTTTTCGGGCGGCCCCAGAGCTGAG TAAGCCGGGTGGAGGGAGTCTGCAAGGATTTCTGAG CGCGATGGGCAGGAGGAGGGGCAAGGGCAAGAGGGC GCGGAGCAAAGACCCTGAACCTGCCGGGGCCGCGCTC CCGGGCCCCGCGTCGCCAGCACCTCCCCACGCGCGCTC GGCCCCGGGCCACCCGCCCTCGTCGGCCCCCGCCCCCT CTCCGTAGCCGCAGGGAAGCGAGCCTGGGAGGAAGAA GAGGGTAGGTGGGGAGGCGGATGAGGGGTGGGGGAC CCCTTGACGTACCCAGAAGGAGGTGCCGGGGTAGGAA GTGGGCTGGGGAAAGGTTATAAATCGCCCCCGCCCTC GGCTGCTCTTCATCGAGGTCCGCGGGAGGCTCGGAGC GCGCCAGGCGGACACTCCTCTCGGCTCCTCCCCGGCA GCGGCGGCGGCTCGGAGCGGGCTCCGGGGCTCGGGT GCAGCGGCCAGCGGGCGCCTGGCGGCGAGGATTACC CGGGGAAGTGTTGTCTCCTGGCTGGAGCCGCGAGAC GGGCGCTCAGGGCGCGGGGCCGGCGGCGGCGAACAA GAGGACGGACTCTGGCGGCCGGGTGCTTGGCCGCGG GGAGCGCGGGCACCGGGCGAGCAGGCCGCGTCCGCGC TCACC (SEQ ID NO: 37)	Invivogen Inc.
ICAM2	Human	399	GTCTCCCAGGCATGACTCCAACAATGCATCCCATGGGA TTTGGGGTTCCCCAGATCTGGGGCTTGTAGGCCTGACT CTCCCCTGTGCACACGTCTCATACACGCATGCGTGCAC CCATTGCCTGCCCCGCCCTTGACACAGGGAGTCAGCA GGGAGGACTGGGTTATGCCCTGCTTATCAGCAGCTTCC CAGCTTCCTCTGCCTGGATTCTTAGAGGCCTGGGGTCC TAGAACGAGCTGGTGCACGTGGCTTCCCAAAGATCTCT CAGATAATGAGAGGAAATGCAGTCATCAGTTTGCAGAA GGCTAGGGATTCTGGGCCATAGCTCAGACCTGCGCCC ACCATCTCCCTCCAGGCAGCCCTTGGCTGGTCCCTGCG AGCCCGTGGAGACTGCCAGTC (SEQ ID NO: 38)	Invivogen Inc.

The Kozak consensus sequence is present in eukaryotic mRNA and is known to improve expression by enhancing translation initiation. Consequently, we introduce the Kozak consensus sequence (i.e. GCCACC, SEQ ID NO:39) upstream of the FVIII or LUC2 gene within the lentiviral vector plasmids.

HUVECs or LSECs were transduced at a multiplicity of infection (MOI)= 50. Culture medium was collected at 24, 48, and 72 hrs and FVIII levels were subsequently measured in the conditioned medium using a human FVIII-specific ELISA, according to the manufacturer's instructions (Asserachrome). Using flow cytometry, the results showed that more than nearly 90 % of HUVEC and LSEC cells were transduced compared to non-transduced HUVECs (Figure 3). Relatively robust FVIII expression could be achieved in transduced HUVECs and LSECs with CMV, ICAM2 and EDN1mini. The Kozak translational consensus sequence significantly enhanced FVIII expression levels (Figure 4). In particular, the Kozak consensus optimized ICAM2 construct yielded the highest FVIII expression in both HUVECs and LSECs. These results confirmed the robustness of the selected promoters in endothelial cells.

Example 3: *In vitro* validation of endothelial-specific (EC) CREs in transfected HUVECs

To validate whether the different EC-CREs identified by genome-wide computational analysis, led to enhanced FVIII expression when coupled to an EC-specific human *ICAM2* promoter, human umbilical vein endothelial cells (HUVECs) were transfected *in vitro* with the corresponding lentiviral vector constructs: pCDH-EC-CRE-ICAM2-FVIII, with EC-CRE representing the respective regulatory elements named in Figure 5, which were cloned upstream of the *ICAM2* promoter driving the human FVIII gene in a pCDH lentiviral self-inactivating backbone. Figure 8 shows 4 examples of such expression cassettes comprising HYAL2-EC-CRE1a (Figure 8a – SEQ ID NO. 50), HYAL2-EC-CRE1b (Figure 8d – SEQ ID NO. 51) and IFI27-EC-CRE1b (Figure 8c – SEQ ID NO. 52). The control vector (without EC-CRE) is depicted in Figure 8b (SEQ ID NO. 49). Analogously, the person skilled in the art would be capable of cloning the other EC-CRE's in a similar manner in the expression vector backbone.

HUVECs were seeded at 1.5×10^5 cells/well of 6-well plate transfected 24 hr later with cationic lipid-based Lipofectamine 3000 (Invitrogen, USA). For HUVECs, 2.5 microgram of each plasmid were mixed with 3.75 microliter of Lipofectamine reagents. The P3000 reagent was mixed with each plasmid at 2 microliter per 1 microgram of plasmid and

incubated at room temperature for 5 mins before adding to the HUVECs. Sixteen hours after transfection, the cell culture medium was removed and then replaced with fresh medium. 72 hrs later 100 microliter of the culture medium was collected and stored in -80°C for FVIII quantification using a human FVIII-specific ELISA (Asserachrome). The results showed that about 70% of the CRMs resulted in increased FVIII expression in transfected HUVECs in vitro, relative to the control lentiviral vector without CRE (i.e. 23 out of 32). (Figure 5).

Example 4: *In vivo* validation of endothelial-specific (EC) CREs following lentiviral transduction in mice

Next, it was validated whether the EC-CREs identified by genome-wide computational analysis, led to enhanced FVIII expression in mice in vivo. Self-inactivating lentiviral vectors were used to express the human codon usage optimized B-domain deleted FVIII from an EC-specific human *ICAM2* promoter. To test the impact of the EC-CRE on FVIII expression, the HYAL2-EC-CRE1a (SEQ ID NO. 50), HYAL2-EC-CRE1b (SEQ ID NO. 51) and IFI27-EC-CRE1b (SEQ ID NO. 52) elements were cloned upstream of the *ICAM2* promoter driving the human FVIII gene. Said expression cassettes comprise respectively the EC-CRE's HYAL2-EC-CRE1a, HYAL2-EC-CRE1b and IFI27-EC-CRE1b. Analogously, the person skilled in the art would be capable of cloning the other EC-CRE's in a similar manner in an expression vector backbone. In these examples the pCDH-*ICAM2*-FVIII backbone is used. A lentiviral vector identical in design but without any upstream EC-CRE was used as control (SEQ ID NO. 49) to compare FVIII expression levels. Lentiviral vector particles were manufactured by transient cotransfection of HEK293 packaging cells with lentiviral vector and helper plasmids (Cyagen, USA). Vector titer was determined and expressed in Transducing Units per ml (TU/ml). Lentiviral vectors were retro-orbitally injected in 2 day-old neonatal CB17-SCID mice (Taconic). The vector preparation was supplemented with 40 microgram/ml polybrene in a total volume of 80 microliter. A total vector dose of 1×10^8 TU was used. Plasma was collected 5 weeks post-injection and FVIII was measured using a human FVIII-specific ELISA.

A significant increase was detected in FVIII expression *in vivo* when the EC-CRE were present compared to a control lentiviral vector without EC-CRE (Figure 6). In particular, a 27-fold increase in FVIII could be detected following in vivo transduction with lentiviral vectors containing the IFI27-EC-CRE1b compared to a control lentiviral vector without EC-CRE (*ICAM2*-FVIII). Similarly, the HYAL2-EC-CRE1a also boosted FVIII expression, but to a lesser extent (5-fold). This is consistent with the increased FVIII expression following

in vitro HUVEC transfection with the IFI27-EC-CRE1b and HYAL2-EC-CRE1a vectors compared to controls without EC-CRE (ICAM2-FVIII) (Figure 5). HYAL2-EC-CRE1b did not increase FVIII expression in vivo, consistent with the lower levels of FVIII expression in transfected HUVECs in vitro (Figure 5).

5

Example 5: Confirmation of increased gene expression by EC-CRE in organ-derived endothelial cells isolated from lentivirally transduced mice

The mice were injected with the lentiviral vectors containing the ICAM2-FVIII (no CRE control – SEQ ID NO. 49) or IFI27-EC-CRE1b-ICAM2-FVIII (SEQ ID NO. 52) expression cassette. After euthanization, the liver and spleen were processed to obtain their respective endothelial populations (i.e. liver sinusoidal endothelial cells, splenic endothelial cells). First, a single cell suspension was obtained from the liver and spleen tissue using the GentleMACS dissociator (Product no- 130-093-235, Miltenyi Biotec) according to the manufacturer's protocol (Liver Dissociation Kit: Product code: 130-105-807; Miltenyi Biotec. [-http://www.miltenyibiotec.com/en/products-and-services/macs-sample-preparation/sample-dissociation/tissue-dissociation-kits/liver-dissociation-kit-mouse.aspx](http://www.miltenyibiotec.com/en/products-and-services/macs-sample-preparation/sample-dissociation/tissue-dissociation-kits/liver-dissociation-kit-mouse.aspx)), spleen (*Spleen Dissociation Kit: Product code: 130-095-926; Miltenyi Biotec.* [-http://www.miltenyibiotec.com/en/products-and-services/macs-sample-preparation/sample-dissociation/tissue-dissociation-kits/spleen-dissociation-kit-mouse.aspx](http://www.miltenyibiotec.com/en/products-and-services/macs-sample-preparation/sample-dissociation/tissue-dissociation-kits/spleen-dissociation-kit-mouse.aspx)).

Subsequently, the single cell suspension from each organ was subjected to MACS cell separation technology (Miltenyi Biotec) to sort out the respective endothelial populations. The single cell suspension obtained from liver and spleen were therefore tagged with CD146 microbeads, according to the manufacturer's instructions (*Product code: 130-092-007; Miltenyi Biotec.* [-http://www.miltenyibiotec.com/en/products-and-services/macs-cell-separation/cell-separation-reagents/endothelial-cells/cd146-lsec-microbeads-mouse.aspx](http://www.miltenyibiotec.com/en/products-and-services/macs-cell-separation/cell-separation-reagents/endothelial-cells/cd146-lsec-microbeads-mouse.aspx)), allowing positive selection of the respective liver-derived and splenic endothelial cells. One to 1.6×10^6 CD146-positive endothelial cells were obtained from the liver and $6.9-8.1 \times 10^5$ CD146-positive endothelial cells were obtained from the spleen. The isolated cells were plated at a density of 25000 cells/well of 48 well plates in 200 microliter of Endothelial Basal Medium supplemented with growth factors.

Total RNA was isolated from the cells using RNeasy Micro Kit (Qiagen) according to manufacturer's instruction. Isolated RNA concentrations were measured using Nanodrop 1000 (Thermo scientific, MA, USA). Complementary DNA (cDNA) was synthesized from

35

75 ng-35 ng isolated RNA using Superscript III First-Strand synthesis system (Invitrogen) according to manufacturer's instructions. The qRT-PCR was performed using SYBR Green qPCR mix (Life technology) in a qPCR ABI Prism 7900HT (Applied Biosystems, Foster City/CA, USA) using FVIII specific primers 5'- AACGGCTACGTGAACAGAAG-3' (forward – SEQ ID NO. 53) and 5'- GATAGGGCTGATTTCCAGGC-3' (reverse – SEQ ID NO. 54). The expression levels were normalized to GAPDH (glyceraldehyde-3-phosphate dehydrogenase) mRNA expression, obtained by using the forward primer 5'- GAAGGTGAAGGTCGGAGTC-3' (SEQ ID NO. 55) and reverse primer 5'- GAAGATGGTGATGGGATTTC-3' (SEQ ID NO. 56).

- 10 The results showed that the IFI27-EC-CRE1b element enhanced FVIII expression in CD146-positive endothelial cells obtained from liver or spleen, as reflected by increased FVIII mRNA levels in mice injected with the lentiviral vector containing the IFI27-EC-CRE1b-ICAM2-FVIII cassette compared to the ICAM2-FVIII control (Figure 7).

CLAIMS

1. A nucleic acid regulatory element for enhancing endothelial cell-specific gene expression comprising, consisting essentially of, or consisting of the sequence
5 selected from the group consisting of: SEQ ID NO:1 to 33, preferably SEQ ID NO:22, or a sequence having at least 80%, preferably at least 85%, more preferably at least 90%, even more preferably at least 95%, such as 95%, 96%, 97%, 98%, or 99%, identity to any of these sequences.
2. The nucleic acid regulatory element according to claim 1, having a maximal length
10 of 1000 nucleotides, preferably 800 nucleotides, more preferably 700 nucleotides, most preferably of 610 nucleotides, still comprising said regulatory element.
3. A nucleic acid regulatory element for enhancing gene expression in endothelial cells comprising, consisting essentially of, or consisting of the complement of a sequence as defined by any one of SEQ ID Nos: 1 to 33, preferably SEQ ID
15 NO:22, or hybridizing under stringent conditions to a sequence as defined by any one of SEQ ID Nos: 1 to 33, preferably SEQ ID NO:22.
4. Use of the nucleic acid regulatory element according to any one claims 1 to 3 in a nucleic acid expression cassette, or a vector, more particularly for enhancing gene expression in endothelial cells of said nucleic acid expression cassette or vector.
- 20 5. A nucleic acid expression cassette comprising at least one, such as one, two, three, four, five or more, nucleic acid regulatory elements according to any one of claims 1 to 3, operably linked to a promoter.
6. The nucleic acid expression cassette according to claim 5, wherein the nucleic acid regulatory element is operably linked to a promoter and a transgene.
- 25 7. The nucleic acid expression cassette according any one of claims 5 or 6, wherein the promoter is a endothelial cells-specific promoter, such as the promoter of any one of the genes selected from the group comprising: IFI27, ICAM2, VWF, EDN1, ENG, ECSCR, CDH5, PECAM1, HHIP, TIE1 and HYAL2.
8. The nucleic acid expression cassette according to any one of claims 5 to 7,
30 wherein the transgene encodes a therapeutic protein or an immunogenic protein.
9. The nucleic acid expression cassette according to any one of claims 5 to 8, wherein the transgene encodes a secretable protein or a structural protein, such as hepatocyte growth factor (HGF), coagulation factor VIII (FVIII), coagulation factor VII (FVII), coagulation factor IX (FIX), coagulation factor XI (FXI), tissue factor (TF), tissue factor pathway inhibitor (TFPI), von Willebrand factor (vWF),
35

ADAMTS13, VEGF, PLGF, FGF, sFLT1, α 1-antitrypsin (AAT), apolipoprotein A-I (apoA-I), matrix metalloproteinases including but not limited to matrix metalloproteinase-3 (TIMP-3) (TIMP-3), insulin, nitric oxide synthase (NOS), antibodies, growth factors, and including but not limited to antibodies directed against any one of said transgenes, factors and their cognate receptors or against any secreted protein or viral protein, small interfering RNA, guide RNA, endonuclease, and Cas9.

10. The nucleic acid expression cassette according to any one of claims 5 to 9, further comprising a polyadenylation signal, preferably the Simian Virus 40 (SV40) polyadenylation signal.

11. A vector comprising the nucleic acid regulatory element according to any one of claims 1 to 3, or the nucleic acid expression cassette according to any one of claims 5 to 10.

12. The vector according to claim 11, which is a viral vector, preferably a lentiviral vector, an adeno-associated viral (AAV) vector, or a adenoviral vector.

13. The vector according to claim 11, which is a non-viral vector, preferably a plasmid, a minicircle, an episomal vector, or a transposon-based vector, such as a PiggyBac-based vector or a Sleeping Beauty-based vector.

14. A pharmaceutical composition comprising the nucleic acid expression cassette according to any one of claims 5 to 10, or the vector according to any one of claims 11 to 13, and a pharmaceutically acceptable carrier.

15. The nucleic acid regulatory element according to any one of claims 1 to 3, the nucleic acid expression cassette according to any one of claims 5 to 10, the vector according to any one of claims 11 to 13, or the pharmaceutical composition according to claim 14 for use in medicine, more preferably for use in gene therapy, in particular for use in treating endothelial cell dysfunction, preferably such as any one of the diseases or disorders selected from the group comprising: liver diseases, hemophilia A, von Willebrand disease, microvascular thrombosis, thrombotic thrombocytopenic purpura, peripheral vascular disease, coronary artery diseases, atherosclerotic diseases, stroke, heart disease, diabetes, insulin resistance, chronic kidney failure, tumor growth, metastasis, venous thrombosis, ischemia, tumour growth, tumour vascularisation, cancer and viral infectious diseases such as Ebola, Dengue fever and Dengue hemorrhagic fever.

16. The nucleic acid regulatory element according to any one of claims 1 to 3, the nucleic acid expression cassette according to any one of claims 4 to 10, the vector according to any one of claims 11 to 13, or the pharmaceutical composition

according to claim 14 for use as a vaccine, preferably a prophylactic vaccine, or for use in immune tolerance induction, vaccination therapy, preferably prophylactic vaccination.

17. A method, preferably an *in vitro* or *ex vivo* method, for expressing a transgene product in endothelial cells, comprising:

- introducing the nucleic acid expression cassette according to any one of claims 4 to 10, or the vector according to any one of claims 11 to 13, comprising the nucleic acid regulatory element according to any one of claims 1 to 3, into the cells; and
- expressing the transgene product in the cells.

18. A method for treating an endothelial cell-related disease or disorder comprising the administration of a therapeutically effective amount of the nucleic acid expression cassette according to any one of claims 4 to 10, the vector according to any one of claims 11 to 13, or the pharmaceutical composition according to claim 14, each comprising the nucleic acid regulatory element according to any one of claims 1 to 3, to a subject in need thereof.

19. The method according to claim 18, wherein said endothelial cell-related disease or disorder is selected from the group comprising: endothelial cell dysfunction, preferably such as any one of the diseases or disorders selected from the group comprising: liver diseases, hemophilia A, von Willebrand disease, microvascular thrombosis, thrombotic thrombocytopenic purpura, peripheral vascular disease, coronary artery diseases, atherosclerotic diseases, stroke, heart disease, diabetes, insulin resistance, chronic kidney failure, tumor growth, metastasis, venous thrombosis, ischemia, tumour growth, tumour vascularisation, cancer and viral infectious diseases such as Ebola, Dengue and Dengue hemorrhagic fever.

Overview of Endothelial CREs Search Strategy

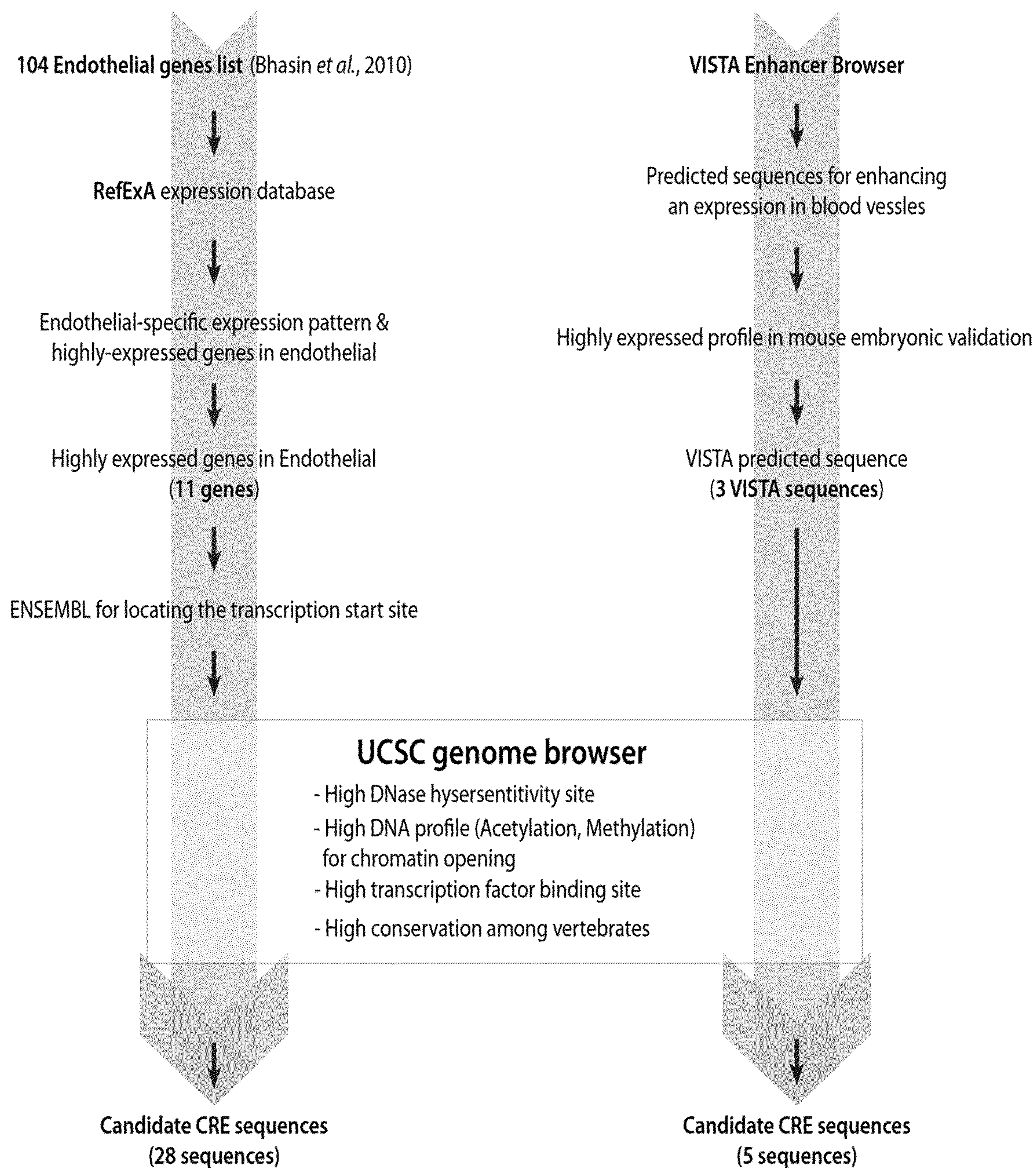


Figure 1

Lentiviral plasmid constructs for endothelial experiments

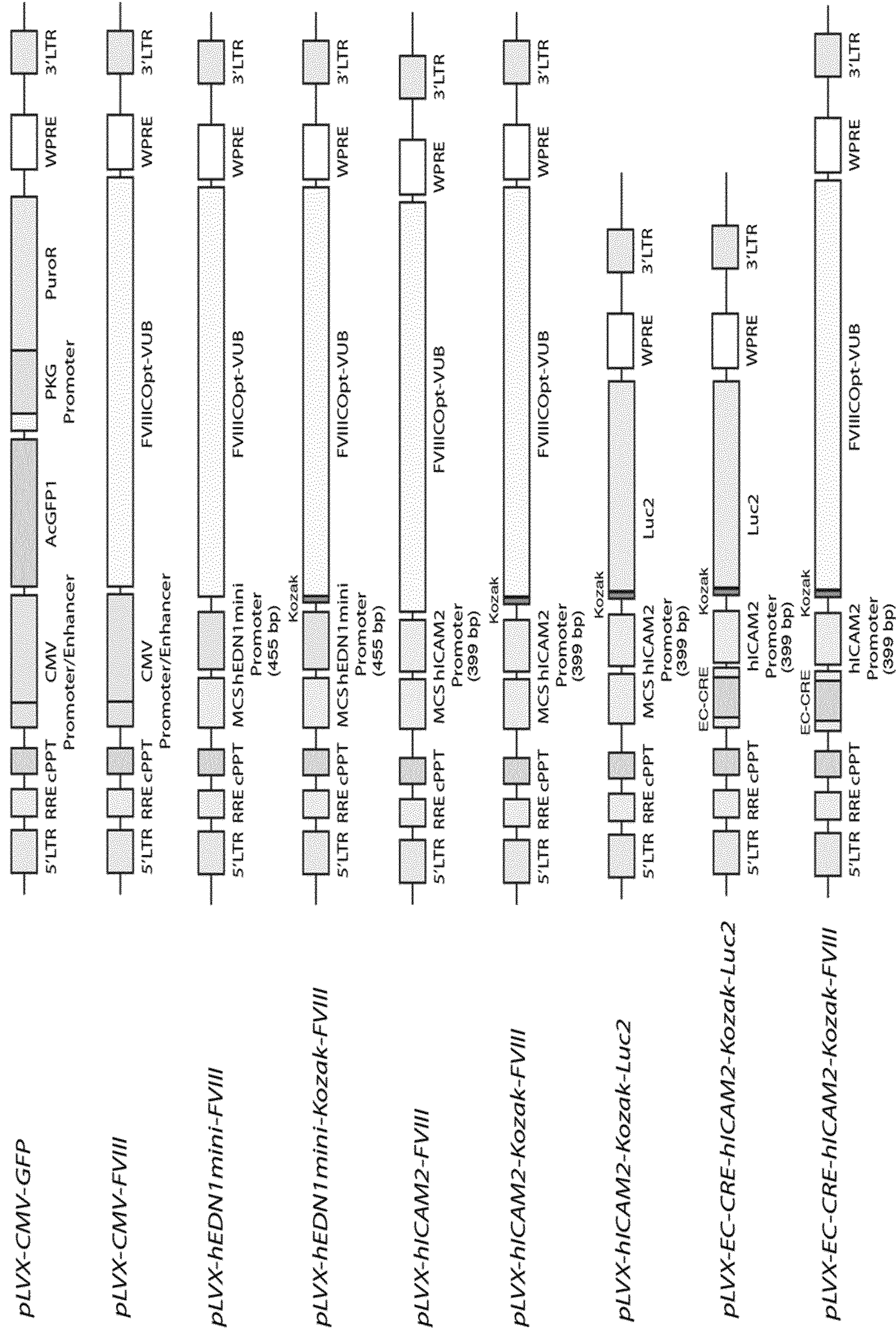


Figure 2

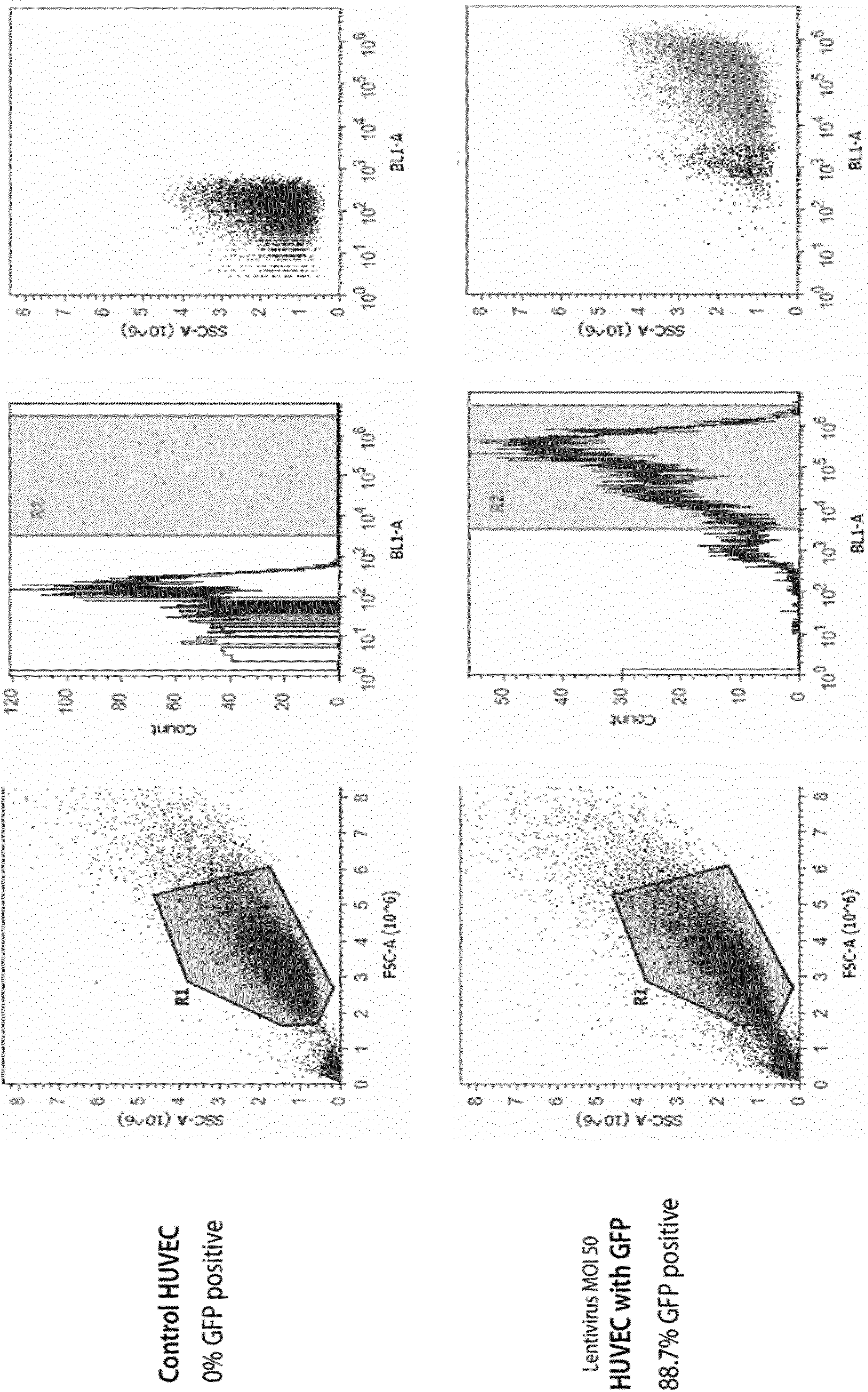


Figure 3A

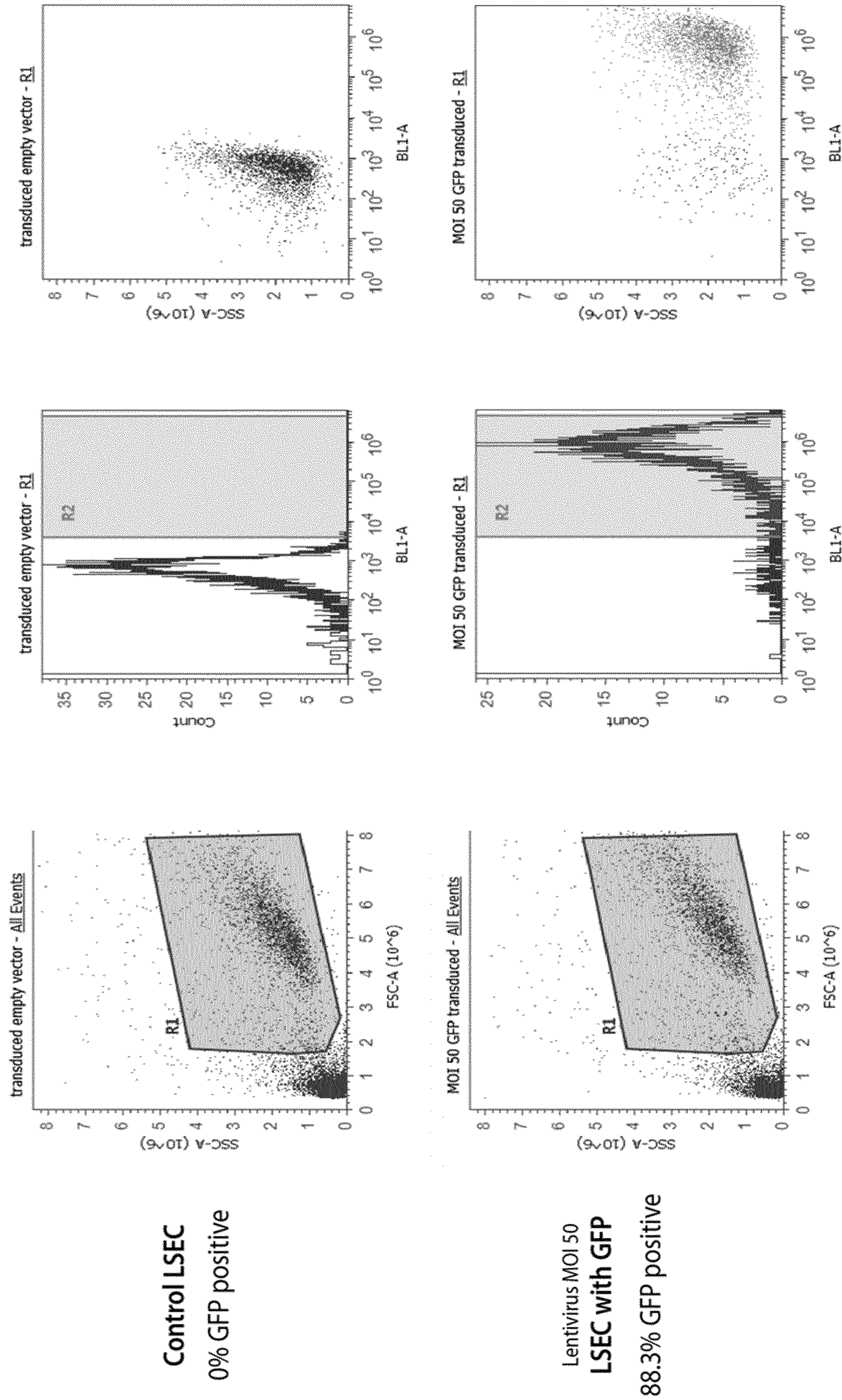


Figure 3B

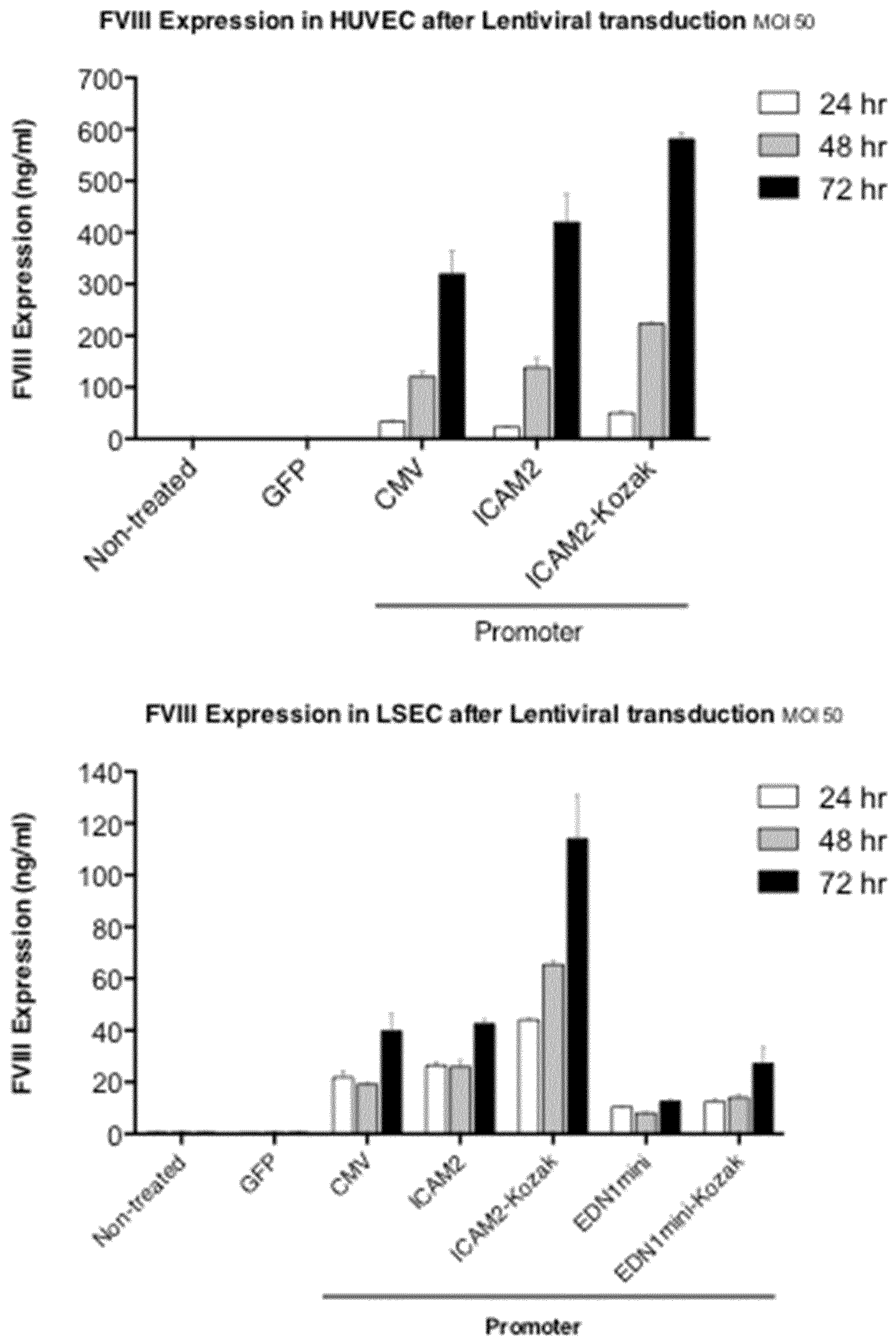


Figure. 4

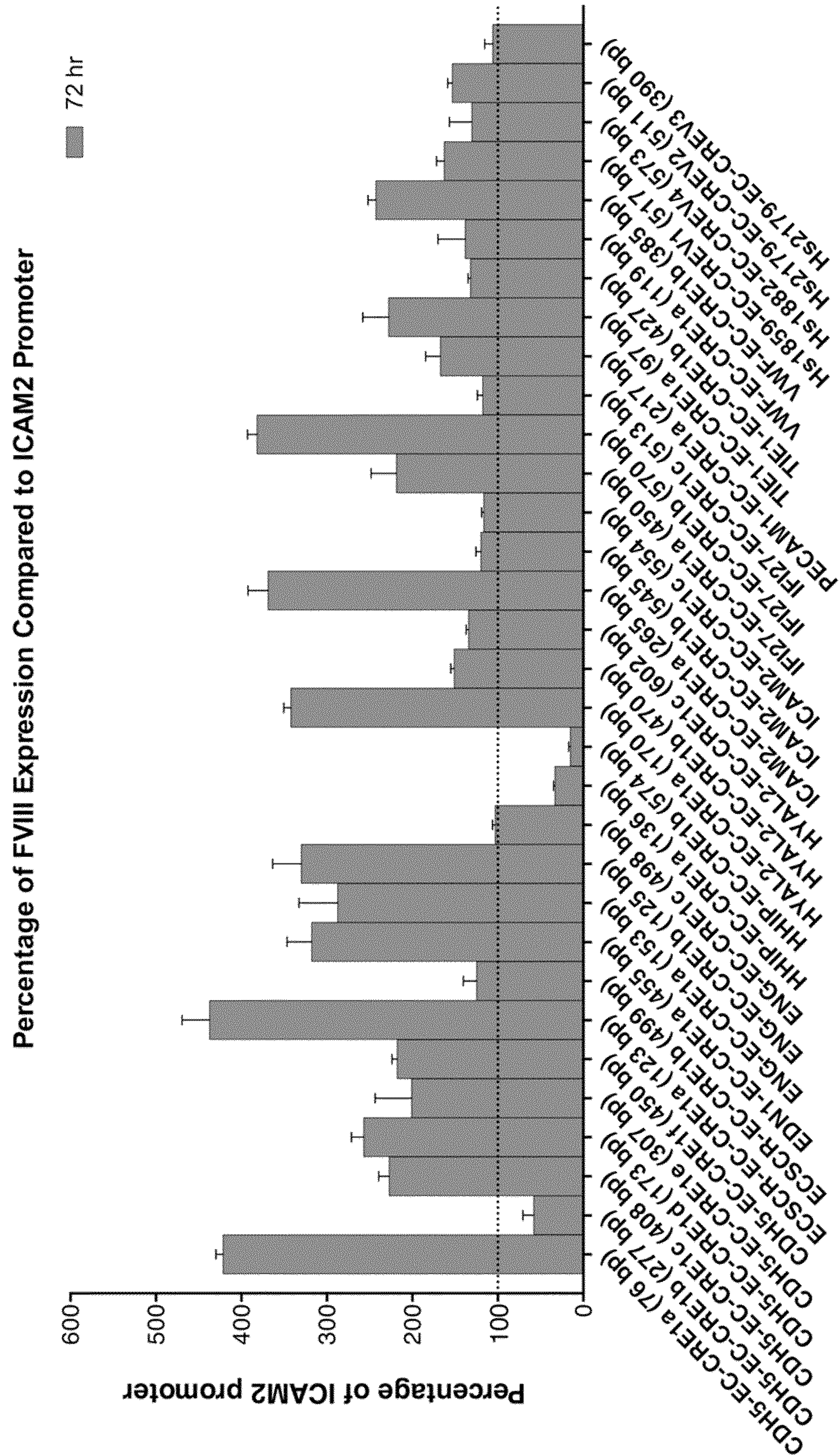


Figure 5

7/11

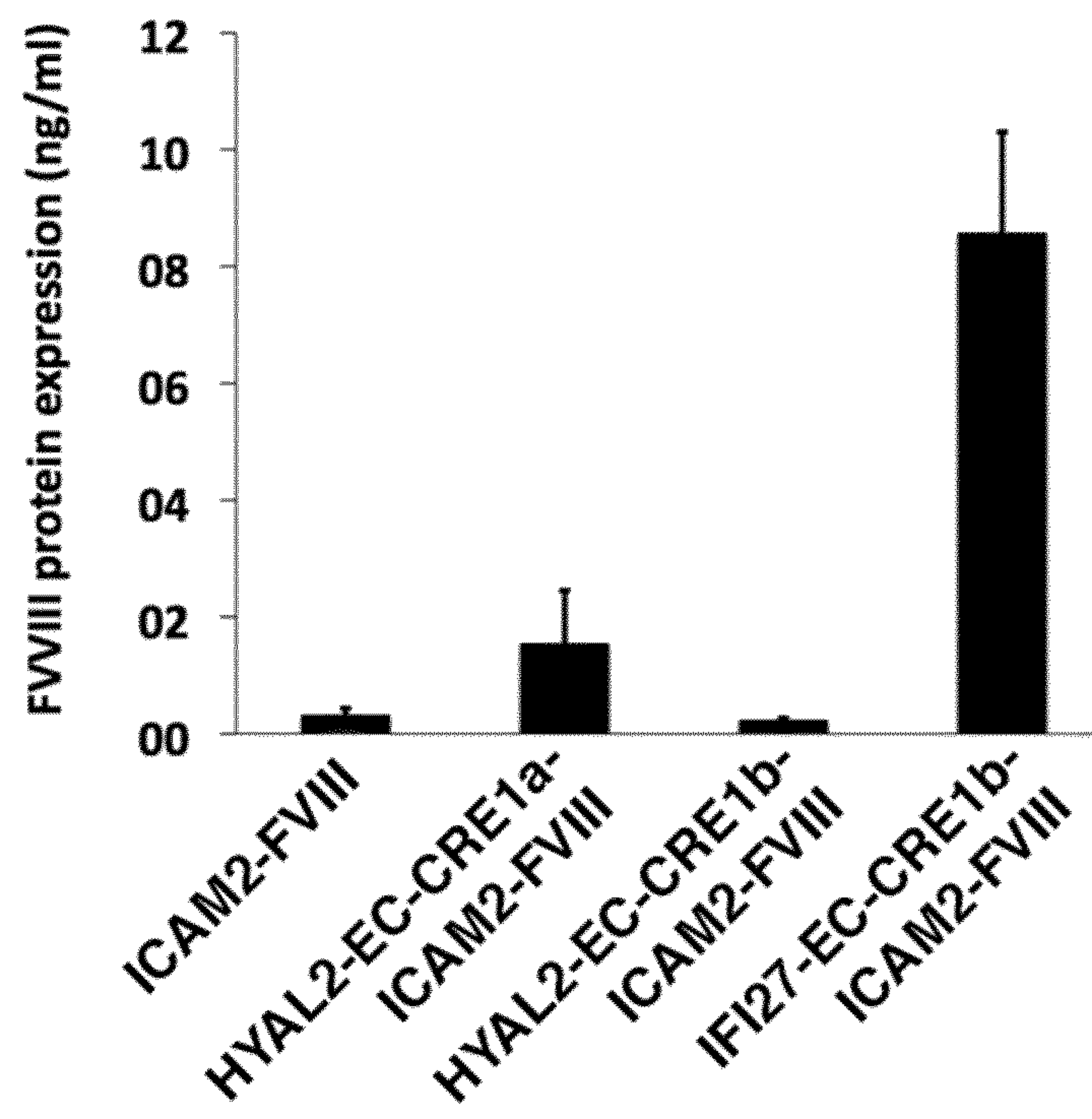


Figure 6

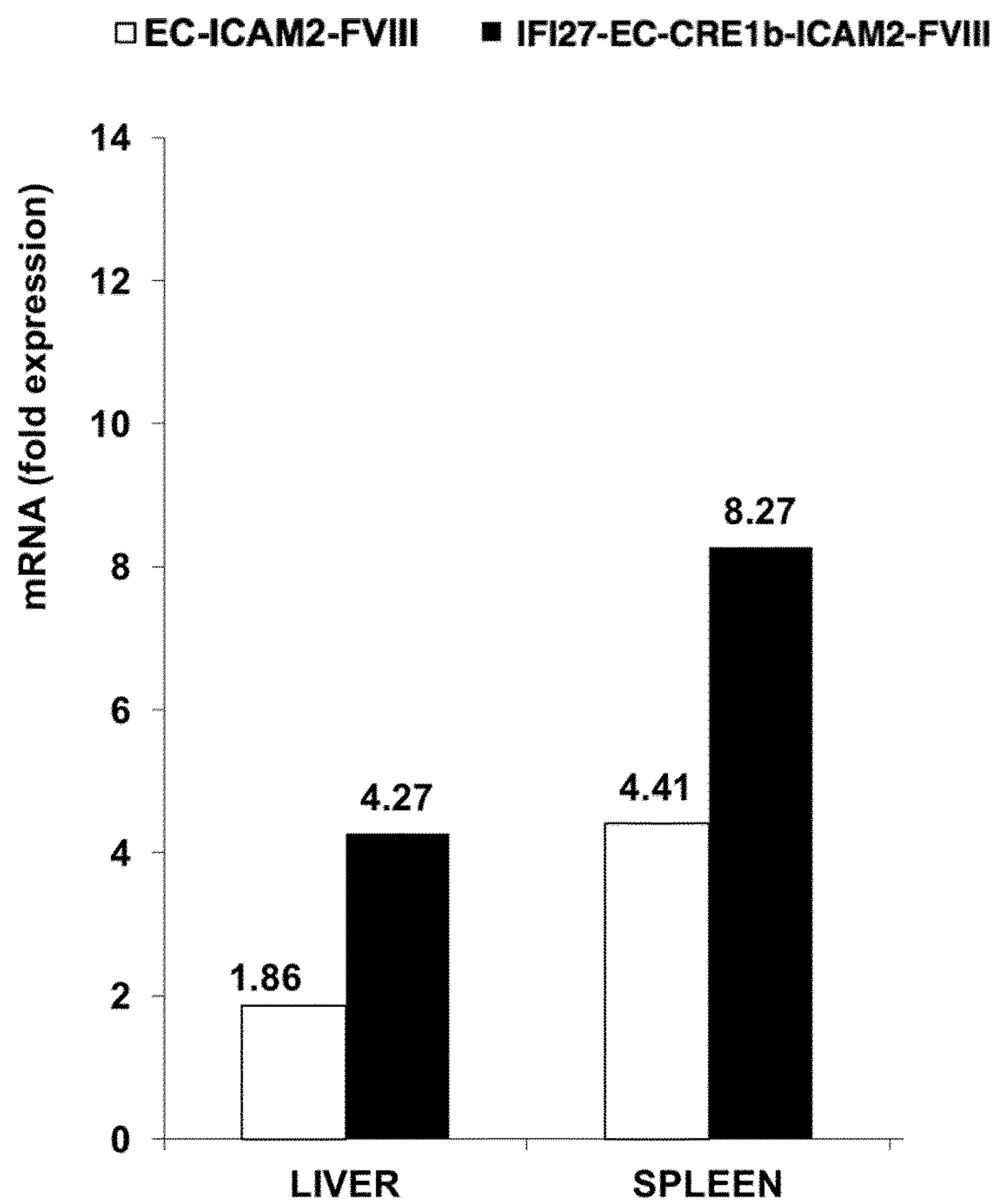


Figure 7

Figure 8a

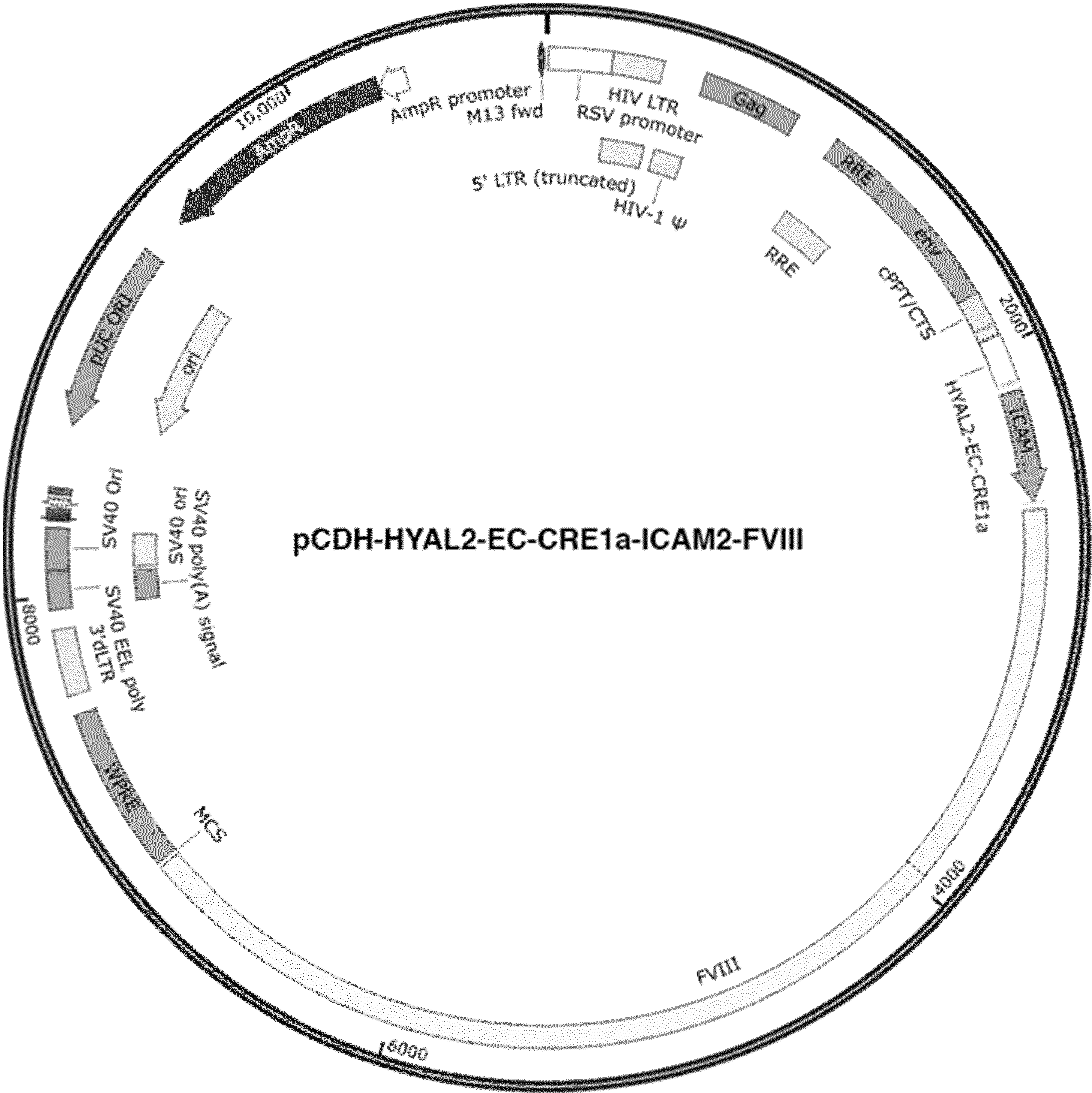


Figure 8b

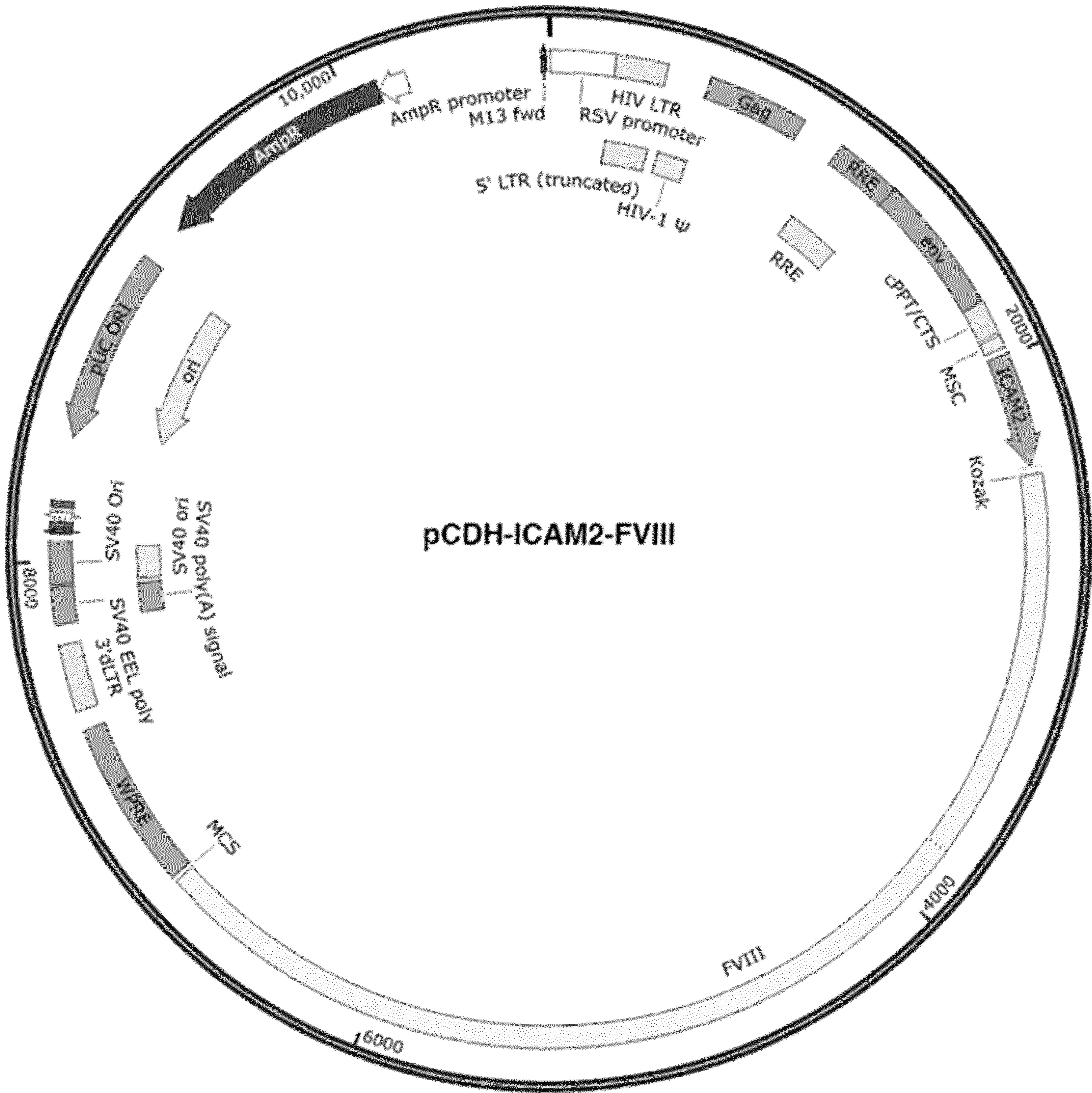


Figure 8c

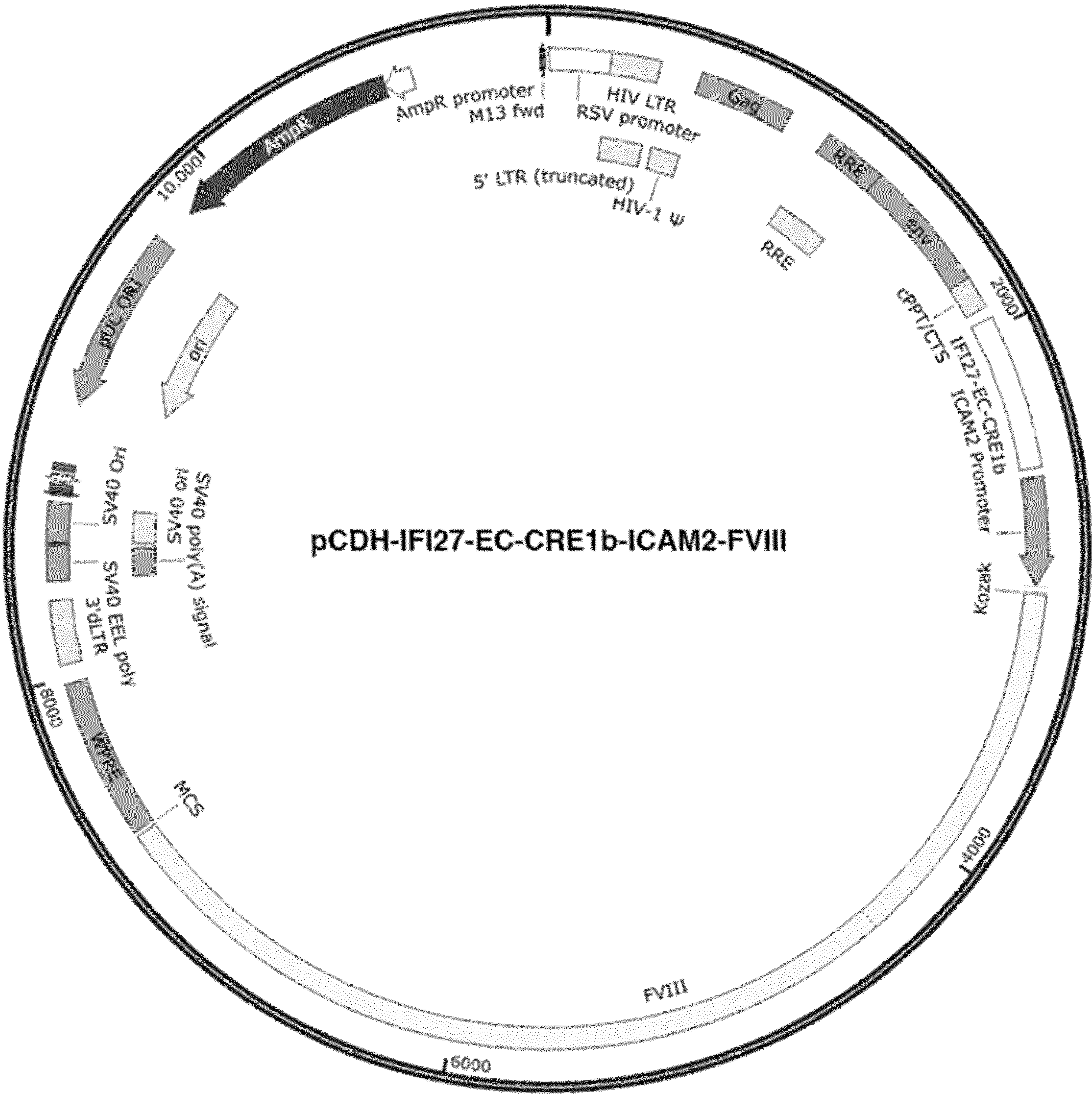


Figure 8d

