Accepted Article Preview: Published ahead of advance online publication

GENOME-WIDE COMPUTATIONAL ANALYSIS REVEALS CARDIOMYOCYTE-SPECIFIC TRANSCRIPTIONAL CIS-REGULATORY MOTIFS THAT ENABLE EFFICIENT CARDIAC GENE THERAPY

Melvin Y. Rincon, Shilipita Sarcar, Dina Danso-Abeam, Marleen Keyaerts, Janka Matrai', Ermira Samara-Kuko, Abel Acosta-Sanchez, Takis Athanasopoulos, George Dickson, Tony Lahoutte, Pieter De Bleser, Thierry VandenDriessche', Marinee K. Chuah

Cite this article as: Melvin Y. Rincon, Shilipita Sarcar, Dina Danso-Abeam, Marleen Keyaerts, Janka Matrai', Ermira Samara-Kuko, Abel Acosta-Sanchez, Takis Athanasopoulos, George Dickson, Tony Lahoutte, Pieter De Bleser, GENOME-WIDE COMPUTATIONAL ANALYSIS REVEALS CARDIOMYOCYTE-SPECIFIC TRANSCRIPTIONAL CIS-REGULATORY MOTIFS THAT ENABLE EFFICIENT CARDIAC GENE THERAPY, *Molecular Therapy* accepted article preview online 08 September 2014; doi:10.1038/mt.2014.178

Received 16 April 2014; accepted 29 August 2014; Accepted article preview online 08 September 2014

MOLECULAR THERAPY – RESEARCH ARTICLE

GENOME-WIDE COMPUTATIONAL ANALYSIS REVEALS CARDIOMYOCYTE-SPECIFIC TRANSCRIPTIONAL CIS-REGULATORY MOTIFS THAT ENABLE EFFICIENT CARDIAC GENE THERAPY

Melvin Y. Rincon[1,2#], Shilipita Sarcar[1,#], Dina Danso-Abeam[2], Marleen Keyaerts[3], Janka Matrai[1,2,], Ermira Samara-Kuko[1,2], Abel Acosta-Sanchez[4], Takis Athanasopoulos[5], George Dickson[5], Tony Lahoutte[3], Pieter De Bleser[6], Thierry VandenDriessche[1,2]*, Marinee K. Chuah[1,2]*.

[1]Department of Gene Therapy & Regenerative Medicine, Free University of Brussels (VUB), Brussels, Belgium; [2]Center for Molecular & Vascular Biology, Department of Cardiovascular Sciences, University of Leuven, Leuven, Belgium; [3]Nuclear Medicine Department, UZ Brussel & In vivo Cellular and Molecular Imaging Lab, Free University of Brussels (VUB); [4]Vesalius Research Center, Flanders Institute of Biotechnology (VIB) & University of Leuven, Leuven, Belgium; [5]School of Biological Sciences, Royal Holloway - University of London, Egham, Surrey, UK; [6]Inflammation Research Center, Flanders Institute of Biotechnology and Department of Biomedical Molecular Biology, Ghent, University, Ghent, Belgium.

[*]Joined corresponding authors; [#]equal contributions.

Short title: Rational *in silico* design of cardiac-specific *CRMs*

CONTACT INFORMATION

Marinee Chuah & Thierry VandenDriessche

Free University of Brussels (VUB)

Department of Gene Therapy & Regenerative Medicine

Faculty of Medicine & Pharmacy

Building D, room D306

Laarbeeklaan 103 B-1090 Brussels Belgium

E-mail: marinee.chuah@vub.ac.be & thierry.vandendriessche@vub.ac.be

Phone: +32 477 529653 Fax: +32 2 477 4159

ABSTRACT

Gene therapy is a promising emerging therapeutic modality for the treatment of cardiovascular diseases and hereditary diseases that afflict the heart. Hence, there is a need to develop robust cardiac-specific expression modules that allow for stable expression of the gene of interest in cardiomyocytes. We therefore explored a new approach based on a genome-wide bio-informatics strategy that revealed novel cardiac-specific *cis*-acting regulatory modules (*CRMs*). These transcriptional modules contained evolutionary conserved clusters of putative transcription factor binding sites that correspond to a 'molecular signature' associated with robust gene expression in the heart. We then validated these cardiac-specific *CRMs in vivo* using an adeno-associated viral vector serotype 9 (AAV9) that drives a reporter gene from a quintessential cardiac-specific α myosin heavy chain (α*MHC)* promoter. Most *de novo* designed cardiac-specific *CRMs* resulted in a >10-fold increase in cardiac gene expression. The most robust *CRMs* enhanced cardiac-speciific transcription 70 to 100-fold. Expression was sustained and restricted to cardiomyocytes. We then combined the most potent *CS-CRM4* with a synthetic heart and muscle-specific promoter (*SPc5-12*) and obtained a significant 20-fold increase in cardiac gene expression compared to the cytomegalovirus promoter. This study underscores the potential of rational vector design to improve the robustness of cardiac gene therapy.

LIST OF ABBREVIATIONS

AAV: adeno-associated virus; BLI: bioluminescent imaging; *CRM*: *cis*-regulatory module; Casq: calsequestrin; *CS-CRM*: cardiac-specific *cis*-regulatory module; ChIP: chromatin immunoprecipitation; *CMV*: cytomegalovirus; GAPDH: glyceraldehyde-3-phosphate dehydrogenase; GFP: green fluorescent protein; DDM: differential distance matrix; HFH1: Hepatocyte nuclear factor 1 homologue 1; HNF: hepatic nuclear factor; *ITR*: *inverted terminal repeats;* αMHC: alpha myosin heavy chain; MDS: multidimensional scaling; mRNA: messenger RNA; MEF2: myocyte enhancer factor-2; qPCR: quantitative polymerase chain reaction; qRT-PCR: quantitative reverse transcriptase polymerase chain reaction; SRF: serum response factor; *TFBS*: transcription factor binding site; vg: vector genomes

KEYWORDS

Gene therapy, adeno-associated virus, heart, cardiovascular disease.

INTRODUCTION

Gene therapy offers promising prospects for the treatment of acquired cardiovascular diseases and hereditary disorders afflicting the heart. Long-term expression and sustained therapeutic effects have been reported in animal models and the first clinical trial results targeting heart failure are encouraging [1–3].

Nevertheless, there is a need to further improve the efficacy and safety of cardiac gene therapy applications. Increasing the vector dose typically boosts the therapeutic efficacy. However, this concomitantly increases the risk of possible undesirable side-effects. In particular, higher vector doses may trigger untoward adaptive or innate immune responses or result in direct cellular toxicity, depending on the vector type. For instance, one of the major risks associated with the use of adeno-associated viral vectors (AAV), relates to the induction of T cell-mediated immune responses against the vector capsid antigens displayed on major histocompatibility class I antigens of the transduced cells[4–6]. This likely triggers the elimination of the gene-modified cells by the immune effector cells which in turn contributes to cellular toxicity and short-term gene expression. Most vectors used in clinical cardiac gene therapy rely on ubiquitously expressed promoter, such as the cytomegalovirus (*CMV*) promoter to drive the therapeutic gene[1]. However, this may not only result in cardiac gene expression but also provoke unwanted expression of the gene of interest in non-target cells, depending on the intrinsic tropism of the vector. This may in turn provoke undesirable cellular consequences and possibly also influence the immune response against the transgene product.

Indeed, it has been demonstrated that the inadvertent expression of the gene of interest in antigen-presenting cells (APCs), may provoke an untoward immune responses against the gene-modified target cells and/or the therapeutic transgene product[7,8]. The use of cardiac-specific promoters may improve transcriptional targeting in cardiomyocytes and potentially overcome some of the limitations of using ubiquitously expressed promoters[9] Importantly, these cardiac-specific promoters provide an added layer of control over transgene activity following systemic gene delivery. However, most cardiac-specific promoters (e.g. α myosin heavy chain, *αMHC* promoter) typically yield lower levels of expression of the gene of interest. Hence, there is a need to further improve cardiac-specific gene therapy vectors using a multi-pronged approach that relies not only on optimizing 'transductional' targeting[10–13] but also boosts expression by enhancing 'transcriptional targeting'[14]. The development of more robust cardiac-specific vectors may allow for the use of lower and thus potentially safer vector doses for cardiac gene therapy applications. Most conventional methods of improving vector design, rely on haphazard *ad hoc* trial-and-error approaches whereby transcriptional enhancers are combined with promoters to increase the levels of expression of the gene of interest and/or overcome transcriptional repression[14,15]. Moreover, the design of a given gene therapy vector is often based on the *in vitro* characteristics of its regulatory elements in cell lines. However, this approach is not always predictive as *in vitro* and *in vivo* vector performances do not always correlate[16,17]

In the current study, we validated an alternative strategy of improving transcriptional targeting to cardiomyocytes by '*de novo* computational design'. We therefore employed a comprehensive *in silico* strategy that relies on the genome-wide identification of transcriptional cardiac-specific *cis*-regulatory modules (*CRM*). These heart-specific *CRMs* contain a 'molecular signature' composed of clusters of transcription factor binding site (*TFBS*) motifs that are characteristic of highly expressed heart-specific genes. Moreover, this comprehensive computational analysis takes into consideration evolutionary conserved transcriptional regulatory motifs, which is particularly relevant in anticipation of clinical translation. Most importantly, these cardiac-specific *CRMs* boost transcriptional targeting after cardiac gene therapy up to 100-fold. This type of multi-disciplinary approach -at the nexus of genomics, computational biology and gene therapy- remain largely unexplored, which underscores the novelty of the current study. Consequently, this approach offers unique opportunities to generate more robust cardiac-specific gene therapy vectors with potentially broad implications for the field. Furthermore, the validation of these heart-specific *CRMs* provides new insights into the molecular determinants underlying transcriptional control in cardiomyocytes.

RESULTS

Computational de novo design of heart-specific CRMs

To design robust cardiac-specific gene therapy vectors, we relied on a multi-step computational approach that allowed us to identify evolutionary conserved cardiac-specific *CRMs* (*CS-CRMs*) associated with genes that are highly expressed in the heart (Fig. 1). This *in silico* strategy was initially developed to identify *CRMs* associated with differential gene expression

following specific *in vitro* stimuli[18]. However, to our knowledge, this type of bio-informatics analysis had not yet been explored in the context of gene therapy and had not yet been validated *in vivo*. One of the unique features of this computational strategy is that it takes into account not only the over-representation of a given *TFBS* but also its context-dependent co-occurrence with other *TFBS* on a genome-wide scale[18]. This comprehensive genome-wide *in silico* analysis allowed us to take into account the actual context of the TFBS that are part of these transcriptional modules.

Eight different *CS-CRM*s ranging from size of 117 bp to 689 bp were identified for the heart (Table 1 & Table S1, Figure S1). These *CS-CRMs* comprised binding sites for 8 different TFs including SRF, CTF/NF1, MEF2, RSRFC4, COUP-TF1, HFH1, HNF3α and HNF3β (Table 1). The *CS-CRM* (i.e. *CS-CRM1* to *CS-CRM8*) correspond to *TFBS* clusters in the promoters of the following heart-specific genes: *Myl3 (CS-CRM1), Brd7 (CS-CRM2) Myl2 (CS-CRM3), Casq2e1 (CS-CRM4), Casq2e2 (CS-CRM5), Ankrd1e1 (CS-CRM6), Ankrd1e2 (CS-CRM7)* and *Ankrd1e3 (CS-CRM8)* (Table 1). These distinct *CS-CRMs* contain a 'molecular signature' that are characteristic of genes that are highly expressed in the heart. Most *CRMs* contain identical *TFBS* but each *CRM* is unique with respect to their specific arrangement. The *CS-CRMs* were evolutionary conserved among 44 divergent species, suggesting strong selection pressure to maintain these particular *TFBS* combinations for high cardiac-specific expression. We have shown the corresponding *CRM* sequences from a few selected species (Table S1 & Figure S1). This evolutionary conservation increases the likelihood that the performance of the

CRMs is preserved following gene therapy in humans. This may ultimately reduce attrition rate in gene therapy clinical trials.

In vivo validation of cardiac-specific CRMs

To validate the *de novo* designed *CS-CRM in vivo*, we generated adeno-associated vectors (AAV) that expressed the humanized green fluorescent reporter protein (GFP) from a chimeric promoter. This promoter was composed of the heart-specific α-myosin heavy chain (αMHC) promoter linked to the different *CS-CRM* (Fig. 2a). We selected the AAV9 serotype to obtain efficient cardiac gene transfer after intravenous injection of 10^{11} vg in C57Bl/6 mice. Seventy percent of the *de novo* designed *CS-CRM* (5 out of 8: *i.e. CS-CRM1, CS-CRM4, CS-CRM6, CS-CRM7* and *CS-CRM8*) resulted in a significant >10-fold increase ($p<0.05$) in transcription compared to the control without *CS-CRM* (Fig 3a & b), consistent with the increase in GFP protein expression levels (Fig 2b-2d). In particular, the *CS-CRM4* and *CS-CRM7* elements resulted in a significant 100 and 70-fold ($p< 0.01$) increase in *GFP mRNA* expression respectively, compared to the control without *CS-CRM* (Fig 3a & b). These two *CS-CRMs* share very similar types of *TFBS*, such as MEF2, RSRFC4, HFH1, NF1, HNF3α and HNF3β but differ in their specific arrangement. Consequently, these selected *CS-CRM* yielded the highest GFP protein expression levels in the heart (Fig 4 a-d). This was confirmed at two different vector doses (Fig 2b & Figure S2). Overall, the mRNA levels correlated strongly with the GFP fluorescence. Cardiac specificity was maintained since *GFP mRNA* and protein expression was absent or limited in any other organ or tissue, (Fig 4, 5a-b & Figure S3 a-h). All the AAV9-CS-

CRM-αMHC-GFP constructs resulted in stable cardiac gene transfer with comparable efficiencies (Figure S4 a-b), confirming that the differences in reporter gene expression directly reflect the differences in potency of the respective *CS-CRM*. Furthermore, though the AAV9 vectors resulted in significant number of vector genomes in different tissues besides heart, including liver, lungs, skeletal muscle and to a lesser extent in spleen, kidney and brain (Fig 6a & b), expression was restricted to the heart and more specifically to cardiomyocytes as confirmed by immunohistochemistry for troponin-T and connexin-43 (Fig 6c & d). These *in vivo* data validate the bio-informatics algorithm and establish proof-of-concept that the *de novo* design of cardiac-specific *CS-CRM* resulted in robust cardiomyocyte-specific expression following gene therapy. Finally, we demonstrated the binding of MEF2, SRF and HNF3β on the most potent *CS-CRM4* element by chromatin immunoprecipitation (ChIP) using heart from mice that were injected with AAV vectors containing *CS-CRM4* (Fig. 2a & Figure S2). The ChIP assays revealed a specific enrichment of the MEF2, SRF and HNF3β TFs on *CS-CRM*4. In particular, in the case of MEF2, a robust 14-fold enrichment of *CS-CRM*4 over the negative control was apparent. In addition, a 9.5 and 4.5-fold enrichment was observed for SRF and NHF3, respectively (Fig. 6e). This provided independent experimental confirmation of the *in silico* predicted putative TFBS that are located in the *CS-CRM*4 element.

Combining synthetic promoters with cardiac-specific CRMs

Next, we combined the most potent *CS-CRM4* element with a synthetic promoter that is known to confer high cardiac and skeletal muscle-specific

expression (i.e. *SPc5-12*)[16] and tested its impact on gene expression in these tissues. This chimeric synthetic *CS-CRM4/SPc5-12* promoter was used to express luciferase from a self-complementary AAV9 (scAAV9), that further increases transduction compared to ssAAV. The AAV terminal repeat (TR) mutant generates self-complementary vectors to overcome the rate-limiting step to transduction *in vivo*[19–21]. We injected 10^{10} vg scAAV9-luciferase vectors in adult severe combined immune deficient (SCID) mice and determined luciferase expression by bioluminescent imaging (BLI) 2 and 4 weeks post-injection (Fig. 7a). *Ex vivo* analysis of the organs was performed 5 weeks post-injection (Figure S5 & S6). Comprehensive comparative analysis by BLI revealed that this *de novo* designed chimeric synthetic *CS-CRM4/SPc5-12* combination was the most potent compared to any of the other promoters tested (Fig 7b, 7c & Figure S5). In particular, a robust and significant 20-fold increase in cardiac expression was obtained using the *CS-CRM4/SPc5-12* promoter compared to the cytomegalovirus (*CMV*) promoter. *CMV* is commonly used in cardiac gene therapy clinical applications and serves as the 'gold standard' given its robust expression in cardiomyocytes. The level of cardiac-specific luciferase gene expression driven from the *CS-CRM4/αMHC* promoter was comparable to that of *CMV* (Fig. 7b, Fig. 7c & Figure S5). Moreover, *CS-CRM4/SPc5-12* was also more robust than the *CS-CRM4/αMHC* chimeric promoter, yielding a significant 19-fold increase in luciferase activity (Fig. 7b, Fig. 7c). Finally, we demonstrated that the chimeric synthetic *CS-CRM4/SPc5-12* promoter resulted in a nearly 5-fold increase in cardiac gene expression compared to *SPc5-12* whereas only a slight or no increase in skeletal muscle expression was apparent (Fig. 7b & c). The

preferential increase in cardiac expression as opposed to muscle by the *CS-CRM4* element is consistent with the preferential binding of cardiac-specific TFs on *CS-CRM4* (Fig. 6e). This implies that the extent of the increase by the *CS-CRM* is depending, at least in part, on the promoter used to drive the gene of interest (i.e. αMHC *vs. SPc5-12*). Reporter gene expression was absent in most of the other organs and tissues (i.e. kidney, spleen, brain and diaphragm) (Figure S6). Nevertheless, we observed minor expression in the lungs of the mice injected with the vector containing the SPc5-12 promoter. Moreover, expression of the reporter gene was observed in the liver (Figure S6 & S7). These observations challenge the prevailing assumption that the synthetic SPc5-12 promoter is only expressed in heart and skeletal muscle.

We subsequently injected the AAV9sc-CS-CRM4/SPc5-12 vector ($5x10^{10}$ vg i.v.) into 4 week old C57BL/6 mice (male) and monitored long-term luciferase expression for up to 8 weeks post-injection. Similar results were now obtained in immune-competent C57BL/6 mice as described above in the SCID model (Fig. 8 & Figure S8). This confirms that luciferase expression does not trigger any untoward immune reactions and/or that the expression is not silenced, consistent with the long-term persistence of the AAV transduced luciferase-expressing cells. The highest level of cardiac gene expression was obtained with the AAV9sc-CS-CRM4/Spc5-12 vector, yielding a significant 19-fold increase (at 8 weeks) compared to AAV9sc-CMV consistent with the results obtained in SCID mice. Moreover, a significant 2-fold increase in luciferase expression was apparent compared to AAV9sc-Spc5-12.

DISCUSSION

The current study explores the use of bio-informatics to improve the efficiency of cardiac gene therapy. Based on a comprehensive genome-wide *in silico* analysis, we discovered evolutionary conserved clusters of *TFBS* motifs that correspond to cardiac-specific *CRMs*. The corresponding genes linked to these *CS-CRMs* were all highly and specifically expressed in the heart. This indicates that a specific '*TFBS* signature' characterizes robust cardiac-specific expression. Using cardiotropic AAV9 vectors, we validated experimentally that the majority of these *CRMs* resulted in a significant >10-fold increase in gene expression in the heart, while maintaining a high degree of cardiac selectivity. The increased protein expression levels were consistent with an increased transcriptional activity. In particular, the most robust *CS-CRM* elements (i.e. *CS-CRM7* and *CS-CRM4*) resulted in a 70-to 100-fold increase in cardiac transcription from the quintessential cardiac α*MHC* promoter. It is therefore particularly encouraging that these *CRMs* effectively boost the performance of the αMHC promoter, since it has been used previously to coax gene expression in the heart following gene therapy[9]. Moreover, we demonstrated that the combination of the *CS-CRM4* element with a synthetic heart/muscle-specific promoter (i.e. *SPc5-12*) resulted in a 20-fold increase in cardiac expression levels compared to when the *CMV* promoter was used. This has important translational implications since the *CMV* promoter is widely used in gene therapy clinical trials for heart failure, including the promising CUPID trial that is based on AAV1-SERCA2a myocardial transduction[2,22]. Consequently, this novel synthetic chimeric *CS-CRM4/SPc5-12* promoter constitutes an

attractive alternative to *CMV*. It may be well suited to boost the performance of AAV vectors in future gene therapy clinical trials for heart failure and hereditary disorders that affect heart and skeletal muscle (e.g. Duchenne muscular dystrophy).

Consequently, higher expression levels can be achieved with the same amount of vector. This also implies that a therapeutic effect could potentially be achieved at lower vector doses, which would ease the manufacturing constraints and minimize the risk of possible immune complications. In particular, the use of lower vector doses may minimize the risk of triggering AAV-specific cytotoxic T cell immune responses[5,23,24]. These cellular immune responses were previously shown to eliminate AAV-transduced cells, at least in liver[4–6,25,26]. To facilitate transduction in the face of pre-existing anti-AAV antibodies[2,22] would be more challenging and would require AAV capsid decoys instead or plasma-exchange[27,28]. In addition to the use of an optimal *CS-CRM*/synthetic promoter combination, the AAV serotype itself significantly impacts on cardiac tropism and expression[10–13]. The optimization of cardiac gene therapy therefore requires optimal cardiotropic AAV serotypes and the use of optimal cardiac-specific regulatory elements, specifically tailored to maximize expression of the gene of interest in cardiomyocytes. In future studies, we will explore and validate the use of these optimized AAV vectors for delivery and expression of therapeutic genes in preclinical models of cardiovascular disease.

Another limitation of conventional vector design is that the typical regulatory elements are often quite large, limiting the size of the transgene that can be packaged into a viral vector. Typically, trimming down promoter sizes results in loss of potency and/or specificity. In contrast, our study shows that potency and specificity was maintained even when selecting small-size *CS-CRMs* (<300bp). Given their small size, these *CS-CRMs* could be used in conjunction with relatively large therapeutic transgenes and/or can be readily accommodated in scAAV vector designs, which have an intrinsic limited packaging capacity of 2.5 kb, to overcome the rate-limiting transduction steps *in vivo* [19–21].

The present *in silico* approach does not only allow for the identification of TFBS that are over-represented in cardiac-specific *CRMs* but it also takes into consideration the specific context of these TFBS. Consequently, the current analysis is more comprehensive and allows for the identification of cardiac-specifc *TFBS* elements that tend to cluster together in *CS-CRMs* of genes that are highly expressed in the heart. It therefore takes into account the actual context-dependent *TFBS* interactions from a broad genome-wide perspective, instead of just relying on the over-representation of a single *TFBS* element[18]. Our previous studies had shown that the present *in silico* analysis is more reliable, compared to other data-mining strategies that typically rely solely on the over-representation of a given *TFBS*, regardless of its context, as discussed previously[18]. The current computational approach that defines the *CS-CRMs* has been further improved beyond the initial description of the algorithm[18], by taking into account phylogenetic

conservation of the *TFBS* clusters. This increases the likelihood that the superior performance of the *CS-CRMs* is maintained upon clinical translation. Furthermore, the current study corroborates the algorithm and establishes a causal relationship between these evolutionary conserved *TFBS* clusters in the *CS-CRMs* and high cardiac-specific gene expression. In contrast, the previously published bioinformatics approach was applied in a different context and merely established correlations between TFBS clusters and differentially expressed genes following a given physiologic stimulus *in vitro* [18]. Moreover, direct proof of the impact of such TFBS clusters on gene expression levels *in vivo* was lacking. The current study overcomes this limitation by *in vivo* gene transfer, hereby providing direct experimental *in vivo* validation of the *genome-wide* computational analysis. Though the large majority of the selected *CS-CRMs* resulted in a significant increase in expression of the gene of interest, a few elements had a more limited impact. It is therefore possible that other unknown factors that were not taken into account in the computational analysis, may have influenced gene expression levels, such as epigenetic factors, chromatin remodeling and/or the spacing between a given *CS-CRM* and the *promoter/TATA box* and their respective orientation on the DNA helix, It is beyond the scope of the present study to analyze this further but it sets the stage to further refine the computational analysis. The present study demonstrates that the *CS-CRM* elements result in a significant increase in cardiac gene expression when used in conjunction with either the *αMHC* promoter or the synthetic *SPc5-12* promoter. This suggests that these *CS-CRM* elements may also effectively boost cardiac gene expression in combination with other heart-specific promoters but this

would need to be verified experimentally on a case-by-case basis. *αMHC* is widely expressed in normal rodent hearts, justifying the use of the corresponding *αMHC* promoter to achieve cardiomyocyte-specific gene expression in mouse models. However, failing rodent hearts have a propensity for *βMHC* instead of *αMHC*. In humans, MHC expression in the failing heart is skewed towards βMHC instead of *αMHC.* Other promoters that are highly expressed in the failing human heart would therefore need to be considered for clinical gene therapy applications. The *SPc5-12* promoter by itself and its *CS-CRM4/SPc5-12* derivative constitute attractive alternatives.

The present study has broad implications for the design of better gene therapy vectors that target different diseases. Our results indicate that *in silico* 'rational vector design' may therefore pave the way towards a more comprehensive approach towards vector optimization with broad implications to improve the efficacy and safety of gene therapy. Indeed, we have recently confirmed that the same computational approach can be employed to identify liver-specific and skeletal muscle-specific *CRMs* that boost the performance of gene therapy vectors in mouse disease models (i.e. hemophilia, phenylketonuria) or non-human primates, which is consistent with our present findings[29–31]. In particular, in these studies, we observed a significant 10 to 100-fold increase in liver-specific gene expression when the corresponding *de novo* designed tissue-specific *CRMs* were employed. Our findings may ultimately impact not only on the clinical translation of gene therapy but also on fundamental biological and transgenic studies that rely on the use of robust tissue-specific gene expression[32,33].

METHODS

Identification of cardiac-specific *CRM*s by computational analysis

A computational sequential approach (Fig. 1) was used to identify tissue-specific *CRMs*: (1) identification of cardiac-specific genes that are highly expressed based on statistical analysis of micro-array expression data of normal human heart tissues; (2) extraction of the corresponding promoter sequences from publicly available databases (3) identification of the regulatory modules and the *TFBS* using a differential distance matrix (DDM)/ multidimensional scaling (MDS) approach[18] (4) searching the genomic context of the highly expressed genes for evolutionary conserved clusters of *TFBS* (i.e. *CRMs*). A detailed outline of the bio-informatics approach is outlined in the Supplementary Materials & Methods.

Generation of *CRM* constructs

The cardiac-specific *CRMs* (*CS-CRM1* to *CS-CRM8*) (Table 1) were synthesized by conventional oligonucleotide synthesis. The different *CS-CRM* were cloned upstream of the α-myosin heavy chain (α*MHC*)[9], in the context of an adeno-associated viral vector (AAV) backbone (Fig. 2a). Additionally, the *CS-CRM4* was cloned upstream of the novel synthetic promoter SPc5-12[16] using a self-complementary AAV (scAAV) backbone[19]. The luciferase (Luc) cDNA was used to substitute the *GFP* as reporter gene. Transgene expression was compared with two other designed vectors, one containing the *CMV* promoter instead of the SPc5-12 and a reference group,

in which the *CS-CRM* was removed. Cloning details are available upon request.

AAV vector production and purification

AAV vectors were prepared as previously described[12]. Genome-containing vectors and empty AAV capsid particles were purified by cesium chloride gradient. For all vector, the titer (in viral genomes (vg)/ml) was determined by qRT-PCR. Typically, for all vectors we achieved titers in the normal range of 2-5x10^{12} vg/ml.

Animal procedures

All animal procedures were approved by the institutional animal ethics committee of the University of Leuven and Vrije Universiteit Brussels. For the *in vivo* screening of the *CS-CRMs*, two-days old C57BL/6 mice were injected with vectors containing the different *CS-CRM* elements (2X10^{10} to 3X10^{11} vg/mouse) via periorbital vein. Mice were euthanized by cervical dislocation after inhaled anesthesia using Isofluorane at 4% in a chamber (depth of anesthesia was evaluated by loss of purposeful voluntary movement and loss of response to reflex stimulation) 5 weeks post-injection and organs were harvested and examined. For the BLI analysis (see below), 10^{10} vg scAAV9-luciferase vectors were injected in adult severe combined immune deficient (SCID) mice.

mRNA analysis

qPCR analysis from mRNA extracted from different organs collected were performed. All results were normalized to mRNA levels of the endogenous murine glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) gene.

Transduction efficiency and vector biodistribution

Transduction efficiency was evaluated by quantifying the *GFP* transgene copy numbers in the different tissues, as described previously[34]. The results were expressed as mean AAV copy number/100 ng of genomic DNA. The qPCR standard consisted of serially diluted plasmids of known quantity.

Chromatin immunoprecipitation assays (ChIP assays)

Genomic DNA regions of were isolated by using specific ChIP grade antibodies[35],[36]. Quantitative PCRs (qPCRs) were carried out on specific genomic region, the resulting signals were normalized for primer efficiency by carrying out qPCR for each primer pair using input DNA. Primer sequences include for positive control and negative control are available upon request.

Immunohistochemistry

Following paraffin embedding protocol, the heart samples were analyzed using anti-troponin-T (1:100, Thermo Scientific, Waltham, MA, USA)[37] as primary antibody and biotinylated secondary antibody. After PBS washings and incubation with DyLight 549-Streptavidin the slides were mounted with fluorescent mounting medium (Without counterstain). Immunostaining analysis for connexin 43 and GFP from heart slides was conducted using

primary connexin 43 rabbit polyclonal antibody (Santa Cruz Biotechnology, Santa Cruz, USA) or anti-GFP goat polyclonal antibody (Abcam, Cambridge, USA)[38,39]. Nuclei were stained with DAPI (Invitrogen Co., Carlsbad, USA)

Detection of luciferase transgene expression

Luciferase expression was analyzed using an *In vivo* Photon Imager (Biospace, France) after intravenous injection of 50 µl luciferin substrate (150 mg/kg, Promega). Representative *in vivo* bioluminescence transgene expression images were taken at 2, 4 and 8 weeks post vector injection. Subsequently at 5 and 8 weeks post viral vector injection, the organs were removed and bioluminescence signal was quantified during 5 min

ACKNOWLEDGEMENTS

This study was supported by grants from EU FP7 (222878, PERSIST), AFM, FWO, EHA, Geconcerteerde Onderzoeksactie (GOA) (EPIGEN, VUB), STK, Willy Gepts Fund (VUB), VUB Strategic Research Program 'Groeier', VUB Industrieel Onderzoeksfonds (Groups of Expertise in Applied Research (GEAR), VUB Proof of Concept Grant (Industrieel Onderzoeksfonds) to Marinee Chuah and Thierry VandenDriessche. Tony Lahoutte and Marleen Keyaerts are Senior Clinical Investigators of the Research Foundation – Flanders (Belgium) (FWO). The research at ICMI is funded by the Belgian State, Nationaal Kankerplan, Vlaamse Liga tegen Kanker and Vlaamse Stichting tegen Kanker. Melvin Y. Rincon is a recipient of a Fundación

Cardiovascular de Colombia/Colciencias Fellowship and an AFM PhD Fellowship. We thank Bing Yan for technical assistance.

AUTHOR CONTRIBUTIONS

M.Y.R. and S.S. designed and performed the *in vivo* experiments, microscopy analysis and the molecular analyses and analyzed the data. J.M. performed experiments, microscopy analysis and analyzed data. E.S., A.A. performed molecular biology experiments and vector production/characterization and animal experiments. P.D. conducted the bio-informatics analysis and wrote part of the paper. M.K. and T.L. contributed to the BLI studies. G.D. contributed to the synthetic promoter studies. T.V. and M.C. designed experiments, coordinated the work, and wrote the manuscript. T.A. and G.D. contributed the SPc5-12. T.V. and M.C. are joined senior corresponding authors. The authors declare no competing financial interests.

REFERENCES

1. Zsebo, K, Yaroshinsky, A, Rudy, JJ, Wagner, K, Greenberg, B, Jessup, M, *et al.* (2014). Long-term effects of AAV1/SERCA2a gene transfer in patients with severe heart failure: analysis of recurrent cardiovascular events and mortality. *Circ. Res.* **114**: 101–108.
2. Jessup, M, Greenberg, B, Mancini, D, Cappola, T, Pauly, DF, Jaski, B, *et al.* (2011). Calcium Upregulation by Percutaneous Administration of Gene Therapy in Cardiac Disease (CUPID): a phase 2 trial of intracoronary gene therapy of sarcoplasmic reticulum Ca2+-ATPase in patients with advanced heart failure. *Circulation* **124**: 304–313.
3. Lai, Y and Duan, D (2012). Progress in gene therapy of dystrophic heart disease. *Gene Ther.* **19**: 678–685.
4. Manno, CS, Pierce, GF, Arruda, VR, Glader, B, Ragni, M, Rasko, JJ, *et al.* (2006). Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response. *Nat. Med.* **12**: 342–347.
5. Mingozzi, F, Maus, MV, Hui, DJ, Sabatino, DE, Murphy, SL, Rasko, JEJ, *et al.* (2007). CD8(+) T-cell responses to adeno-associated virus capsid in humans. *Nat. Med.* **13**: 419–422.
6. Nathwani, AC, Tuddenham, EGD, Rangarajan, S, Rosales, C, McIntosh, J, Linch, DC, *et al.* (2011). Adenovirus-associated virus vector-mediated gene transfer in hemophilia B. *N. Engl. J. Med.* **365**: 2357–2365.
7. Brown, BD, Cantore, A, Annoni, A, Sergi, LS, Lombardo, A, Della Valle, P, *et al.* (2007). A microRNA-regulated lentiviral vector mediates stable correction of hemophilia B mice. *Blood* **110**: 4144–4152.
8. Mátrai, J, Cantore, A, Bartholomae, CC, Annoni, A, Wang, W, Acosta-Sanchez, A, *et al.* (2011). Hepatocyte-targeted expression by integrase-defective lentiviral vectors induces antigen-specific tolerance in mice with low genotoxic risk. *Hepatol. Baltim. Md* **53**: 1696–1707.
9. Pacak, CA, Sakai, Y, Thattaliyath, BD, Mah, CS and Byrne, BJ (2008). Tissue specific promoters improve specificity of AAV9 mediated transgene expression following intra-vascular gene delivery in neonatal mice. *Genet. Vaccines Ther.* **6**: 13.
10. Pacak, CA, Mah, CS, Thattaliyath, BD, Conlon, TJ, Lewis, MA, Cloutier, DE, *et al.* (2006). Recombinant adeno-associated virus serotype 9 leads to preferential cardiac transduction in vivo. *Circ. Res.* **99**: e3–9.
11. Inagaki, K, Fuess, S, Storm, TA, Gibson, GA, Mctiernan, CF, Kay, MA, *et al.* (2006). Robust systemic transduction with AAV9 vectors in mice: efficient global cardiac gene transfer superior to that of AAV8. *Mol. Ther. J. Am. Soc. Gene Ther.* **14**: 45–53.
12. Vandendriessche, T, Thorrez, L, Acosta-Sanchez, A, Petrus, I, Wang, L, Ma, L, *et al.* (2007). Efficacy and safety of adeno-associated viral vectors based on serotype 8 and 9 vs. lentiviral vectors for hemophilia B gene therapy. *J. Thromb. Haemost. JTH* **5**: 16–24.
13. Perabo, L, Büning, H, Kofler, DM, Ried, MU, Girod, A, Wendtner, CM, *et al.* (2003). In vitro selection of viral vectors with modified tropism: the adeno-associated virus display. *Mol. Ther. J. Am. Soc. Gene Ther.* **8**: 151–157.

14. Müller, OJ, Leuchs, B, Pleger, ST, Grimm, D, Franz, W-M, Katus, HA, *et al.* (2006). Improved cardiac gene transfer by transcriptional and transductional targeting of adeno-associated viral vectors. *Cardiovasc. Res.* **70**: 70–78.
15. Dai, Y, Roman, M, Naviaux, RK and Verma, IM (1992). Gene therapy via primary myoblasts: long-term expression of factor IX protein following transplantation in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **89**: 10892–10895.
16. Li, X, Eastman, EM, Schwartz, RJ and Draghia-Akli, R (1999). Synthetic muscle promoters: activities exceeding naturally occurring regulatory sequences. *Nat. Biotechnol.* **17**: 241–245.
17. Miao, CH, Ohashi, K, Patijn, GA, Meuse, L, Ye, X, Thompson, AR, *et al.* (2000). Inclusion of the hepatic locus control region, an intron, and untranslated region increases and stabilizes hepatic factor IX gene expression in vivo but not in vitro. *Mol. Ther. J. Am. Soc. Gene Ther.* **1**: 522–532.
18. De Bleser, P, Hooghe, B, Vlieghe, D and van Roy, F (2007). A distance difference matrix approach to identifying transcription factors that regulate differential gene expression. *Genome Biol.* **8**: R83.
19. McCarty, DM, Fu, H, Monahan, PE, Toulson, CE, Naik, P and Samulski, RJ (2003). Adeno-associated virus terminal repeat (TR) mutant generates self-complementary vectors to overcome the rate-limiting step to transduction in vivo. *Gene Ther.* **10**: 2112–2118.
20. Bish, LT, Sleeper, MM, Brainard, B, Cole, S, Russell, N, Withnall, E, *et al.* (2008). Percutaneous transendocardial delivery of self-complementary adeno-associated virus 6 achieves global cardiac gene transfer in canines. *Mol. Ther. J. Am. Soc. Gene Ther.* **16**: 1953–1959.
21. Sipo, I, Fechner, H, Pinkert, S, Suckau, L, Wang, X, Weger, S, *et al.* (2007). Differential internalization and nuclear uncoating of self-complementary adeno-associated virus pseudotype vectors as determinants of cardiac cell transduction. *Gene Ther.* **14**: 1319–1329.
22. Jaski, BE, Jessup, ML, Mancini, DM, Cappola, TP, Pauly, DF, Greenberg, B, *et al.* (2009). Calcium upregulation by percutaneous administration of gene therapy in cardiac disease (CUPID Trial), a first-in-human phase 1/2 clinical trial. *J. Card. Fail.* **15**: 171–181.
23. Mingozzi, F, Meulenberg, JJ, Hui, DJ, Basner-Tschakarjan, E, Hasbrouck, NC, Edmonson, SA, *et al.* (2009). AAV-1-mediated gene transfer to skeletal muscle in humans results in dose-dependent activation of capsid-specific T cells. *Blood* **114**: 2077–2086.
24. VandenDriessche, T (2009). Muscling through AAV immunity. *Blood* **114**: 2009–2010.
25. Kay, MA (2011). State-of-the-art gene-based therapies: the road ahead. *Nat. Rev. Genet.* **12**: 316–328.
26. Seymour, LW and Thrasher, AJ (2012). Gene therapy matures in the clinic. *Nat. Biotechnol.* **30**: 588–593.
27. Mingozzi, F, Anguela, XM, Pavani, G, Chen, Y, Davidson, RJ, Hui, DJ, *et al.* (2013). Overcoming preexisting humoral immunity to AAV using capsid decoys. *Sci. Transl. Med.* **5**: 194ra92.
28. VandenDriessche, T and Chuah, MK (2013). Vector Decoys Trick the Immune Response. *Sci. Transl. Med.* **5**: 194fs28–194fs28.
29. Nair, N, Rincon, MY, Evens, H, Sarcar, S, Dastidar, S, Samara-Kuko, E, *et al.* (2014). Computationally designed liver-specific transcriptional cis-

regulatory modules and hyper-functional factor IX improve liver-targeted gene therapy for hemophilia B. *Blood*doi:10.1182/blood-2013-10-534032.
30. Markusic, DM and Herzog, RW (2014). A David promoter with Goliath strength. *Blood* **123**: 3068–3069.
31. Viecelli, HM, Harbottle, RP, Wong, SP, Schlegel, A, Chuah, MK, Vandendriessche, T, *et al.* (2014). Treatment of phenylketonuria using minicircle-based naked-DNA gene transfer to murine liver. *Hepatol. Baltim. Md*doi:10.1002/hep.27104.
32. Chuah, MK, Petrus, I, De Bleser, P, Le Guiner, C, Gernoux, G, Adjali, O, *et al.* (2014). Liver-Specific Transcriptional Modules Identified by Genome-Wide In Silico Analysis Enable Efficient Gene Therapy in Mice and Non-Human Primates. *Mol. Ther. J. Am. Soc. Gene Ther.*doi:10.1038/mt.2014.114.
33. Di Matteo, M, Samara-Kuko, E, Ward, N, Waddington, S, McVey, JH, Chuah, MK, *et al.* (2014). Hyperactive PIGGYBAC Transposons for Sustained and Robust Liver-targeted Gene Therapy. *Mol. Ther. J. Am. Soc. Gene Ther.*doi:10.1038/mt.2014.131.
34. Valouev, A, Johnson, DS, Sundquist, A, Medina, C, Anton, E, Batzoglou, S, *et al.* (2008). Genome-wide analysis of transcription factor binding sites based on ChIP-Seq data. *Nat. Methods* **5**: 829–834.
35. Cunha, PMF, Sandmann, T, Gustafson, EH, Ciglar, L, Eichenlaub, MP and Furlong, EEM (2010). Combinatorial binding leads to diverse regulatory responses: Lmd is a tissue-specific modulator of Mef2 activity. *PLoS Genet.* **6**: e1001014.
36. Naderi, A, Meyer, M and Dowhan, DH (2012). Cross-regulation between FOXA1 and ErbB2 signaling in estrogen receptor-negative breast cancer. *Neoplasia N. Y. N* **14**: 283–296.
37. Bedelbaeva, K, Gourevitch, D, Clark, L, Chen, P, Leferovich, JM and Heber-Katz, E (2004). The MRL mouse heart healing response shows donor dominance in allogeneic fetal liver chimeric mice. *Cloning Stem Cells* **6**: 352–363.
38. Jovanova-Nesic, K, Koruga, D, Kojic, D, Kostic, V, Rakic, L and Shoenfeld, Y (2009). Choroid plexus connexin 43 expression and gap junction flexibility are associated with clinical features of acute EAE. *Ann. N. Y. Acad. Sci.* **1173**: 75–82.
39. Towne, C, Raoul, C, Schneider, BL and Aebischer, P (2008). Systemic AAV6 delivery mediating RNA interference against SOD1: neuromuscular transduction does not alter disease progression in fALS mice. *Mol. Ther. J. Am. Soc. Gene Ther.* **16**: 1018–1025.

TABLE LEGENDS

Table 1. Transcription factor binding sites (TFBS) strongly associated with high cardiac-specific expression. Evolutionary conserved CS-CRM enriched in TFBS associated with high cardiac-specific expression. The *CS-CRM* designation, corresponding gene, size (in bp) and TFBS clusters are shown. Subscripts a, b or c refer to *CS-CRM* present at different locations within the same promoter of a given gene

TABLE 1

Name	Gene	Length (bp)	TFBS
CS-CRM1	*Myl3*	150	SRF, NF1, MEF2, RSRFC4, COUP-TF1, HFH1, HNF3α, HNF3β
CS-CRM2	*Brd7*	689	HFH1, HNF3α, HNF3β, SRF, RSRFC4, NF1
CS-CRM3	*Myl2*	183	HFH1, HNF3α, HNF1, MEF2, SRF, NF1, RSRFC4
CS-CRM4	*Casq2e*[a]	219	HNF3α, MEF2, SRF, NF1, RSRFC4, HNF3β, HFH1.
CS-CRM5	*Casq2e*[b]	117	HNF3α, HNF3β, MEF2, NF1, HFH1
CS-CRM6	*Ankrd1e*[a]	299	HNF3α, HNF3β, HNF1, HFH1
CS-CRM7	*Ankrd1e*[b]	277	MEF2, HNF3α, HFH1, HNF3β, NF1
CS-CRM8	*Ankrd1e*[c]	397	HFH1, HNF3α, MEF2, SRF, COUP-TF1, NF1, HNF3β.

FIGURE LEGENDS

Figure 1. Multi-step *in silico* strategy. A computational approach was used to identify cardiac-specific *CRMs*. DDM/MDS: Distance difference matrix/Multidimensional scaling; TFBS: transcription factor binding site

Figure 2. Validation of cardiac-specific *cis*-regulatory module (*CS-CRM*). (a) Schematic representation of AAV9-*CS-CRM*-αMHC-GFP vector used in this study. The expression cassette was packaged in a single-stranded (ss) adeno-associated virus serotype 9 (AAV9), flanked by the 5' and 3' AAV2 inverted terminal repeats (ITR). The α-myosin heavy chain (MHC) promoter drives the humanized recombinant green fluorescent protein (GFP) reporter transgene. The cardiac-specific *CS-CRM* (i.e. *CS-CRM1* to *CS-CRM8*) were cloned upstream of the α-MHC promoter. The β-globin intron and bovine growth hormone polyadenylation site (pA) are also indicated. The control vector AAV9-αMHC-GFP is identical to AAV9-*CS-CRM*-αMHC-GFP but does not contain any *CS-CRM* elements. Representative results at one vector dose (10^{11} vg/mouse) were shown (see below). The expression pattern was confirmed when the experiment was repeated at different vector doses (i.e. $3x10^{11}$ vg/mouse and $2x10^{10}$ vg/mouse) (data not shown) (b) Fluorescent imaging of intact hearts of mice injected with 10^{11} vg/mouse of AAV9-*CS-CRM*-αMHC-GFP containing the different *CS-CRM* (i.e. *CS-CRM1* to *CS-CRM8*) or the AAV9-αMHC-GFP vector without *CS-CRM* as a control (indicated as no *CS-CRM*). Panels are shown at 12.8x magnification, 10 s constant exposure time. (c) Longitudinal tissue fragments of the heart were

excised using a scalpel and put into a 96-well plate. This allowed us to acquire fluorescent composite images of the heart tissue by scanning the entire tissue fragment in a semi-automatic fashion using the Zeiss Axiovert 200 inverted microscope. Mice were injected with the different *CRM* vectors shown in (b).Panels are shown at 5x magnification, 2 s exposure time. (d) Confocal microscopy of intact myocardium of mice injected with 10^{11} vg/mouse of AAV9-*CS-CRM*-αMHC-GFP containing the different *CS-CRM* (i.e. *CS-CRM1* to *CS-CRM8*) or the AAV9-αMHC-GFP vector without *CS-CRM* as a control (indicated as no *CS-CRM*). A representative confocal scan is shown. There was no background fluorescence in PBS-injected mice (Fig. S2). Pictures were taken at 10x magnification (n=3 mice per group)

Figure 3. Validation of cardiac-specific *cis*-regulatory module (*CS-CRM*). (a) Semi-quantitative RT-PCR and (b) quantitative real-time RT-PCR of GFP mRNA expression levels in the heart (mRNA was purified from each mice, 3 mice per group and 3 samples of each organ were analyzed per qPCR in triplicate, averages are presented) 6 weeks after intravenous injection of the AAV9-CS-CRM-αMHC-GFP vectors (10^{11} vg/mouse) containing the different *CS-CRM* elements (i.e. *CS-CRM1* to *CS-CRM8*) compared to AAV9-αMHC-GFP control. Relative cardiac mRNA expression levels (mean ± s.e.m.) are shown. MW: molecular weight marker; H_2O: control sample. GAPDH is shown for comparison in the semi-quantitative RT-PCR and it was used for normalization in the quantitative real-time RT-PCR .

Figure 4. Cardiac specificity of the *CS-CRM*. Fluorescent (a, c) and confocal microscopy (b, d) images of representative sections from selected tissues of the mice injected with AAV9-CS-CRM4-αMHC-GFP (a, b) or AAV9-CS-CRM7-αMHC-GFP (c, d) (1011 vg/mouse) (n=3). A tissue fragment of each organ was excised using a scalpel and put into a 96-well plate. This allowed us to acquire fluorescent composite images of the heart tissue by scanning the entire tissue fragment in a semi-automatic fashion using the Zeiss Axiovert 200 inverted microscope. Panels are show at 5x magnification, 2 s exposure time (a, c). Representative images by confocal microscopy are shown at 10x magnification (b, d). Six images were obtained from different fragments of each organ, n = 3 mice per group; a representative image is shown.

Figure 5. Cardiac specificity of the *CS-CRM*. (a) Semi-quantitative RT-PCR and (b) quantitative real-time RT-PCR of *GFP* mRNA expression levels in different organs 6 weeks after intravenous injection of the AAV9-CS-CRM4-αMHC-GFP vectors (10^{11} vg/mouse) (n = 3). Expression levels (mean + s.d.) relative to the corresponding organ of the mice injected with the control group are shown.

Figure 6. Biodistribution and transduction efficiency. Biodistribution and transduction efficiency (a,b) analysis in different organs of mice (n = 3) injected with AAV9-CS-CRM4-αMHC-GFP (10^{11} vg/mouse) (a) or AAV9-CS-CRM7-αMHC-GFP (10^{11} vg/mouse) (n= 3) (b) determined by qPCR using vector specific primers. Relative to a standard curve of known vector copy

numbers (mean ± s.d.) are shown. (c, d) Heart sections of mice injected with AAV9-CS-CRM4-αMHC-GFP (10^{11} vg/mouse) were subjected to immunostaining for troponin-T (TNNT) (red), connexin-43 (Cnx43) (red) and GFP (green). Nuclei were stained with DAPI (blue). Note the peri-cellular location of Cnx43 on cardiomyocytes consistent with its participation in inter-cellular gap junctions essential for Ca^{2+} handling. (e) Chromatin immuno-precipitation (ChIP) assay on heart of mice injected with AAV9-*CS-CRM4*-αMHC-GFP (10^{11} vg/mouse). Antibodies specific to HNF3α, MEF2 or SRF1 and PCR primers specific for the corresponding TFBS were used. In particular, PCR primers were designed to amplify a region within the vector corresponding to *CS-CRM4* (that binds HNF3α, MEF2 and SRF1), a positive control region in the genome corresponding to the cognate TFBS elements (+) or and untranscribed region on chromosome 6 as negative control (-). Binding events per 10^3 cells were determined for each of the corresponding primers pairs. Significant differences compared to the negative control were indicated (t-test, *p ≤0.05, mean ± s.d.).

Figure 7. Whole body bioluminescence imaging and quantitation of luciferase expression in heart and skeletal muscle in SCID mice. (a) *In vivo* bioluminescence imaging of transgene expression in SCID mice injected intravenously with scAAV9-luciferase vectors (n=4 mice per group) containing the different promoters (i.e. *CS-CRM4/αMHC; SPc5-12; CMV; CS-CRM4/SPc 5-12*). Images were taken 2 and 4 weeks post injection. Whole body bioluminescence imaging were represented based on a color scale, showing intensities ranging from $5.22X10^3$ (blue) – $6.82X10^5$ (red) ph/s/cm^2/sr**.** (b) *Ex vivo*

determination of luciferase expression in heart and skeletal muscles. Values were expressed as total flux (photons/sec/cm^2/sr) and (c) fold-difference in total flux with respect to AAV9sc-CS-CRM4/αMHC vector. The fold-difference was indicated above the graph for the AAV9sc-CS-CRM4/SPc5-12 vector. Values were expressed as mean + s.e.m. (n=3 mice per group). Gastroc.: gastrocnemius; Quads: quadriceps

Figure 8. Whole body bioluminescence imaging and quantitation of luciferase expression in heart and skeletal muscle in C57Bl/6 mice. (a) *In vivo* bioluminescence imaging of luciferase expression in C57BL/6 mice injected intravenously with scAAV9-luciferase vectors (n=3 mice per group, 5x10^{10} vg/mouse) containing the different promoters (i.e. *SPc5-12; CMV; CS-CRM4/SPc 5-12*). Images were taken 2, 4 and 8 weeks post injection. Whole body bioluminescent images were represented based on a color scale, showing intensities ranging from 1.26X10^6 (blue) – 1.45X10^7 (red) ph/s/cm^2/sr. (b) *Ex vivo* bioluminescence imaging of harvested heart and skeletal muscles from C57BL/6 mice injected intravenously with the scAAV9-luciferase vector (n=3 mice per group, dose: 5x10^{10} vg/mouse) containing the different promoters (i.e. *SPc5-12; CMV; CS-CRM4/SPc 5-12*). Images of the harvested tissues and organs were taken 8 weeks post injection. Bioluminescent images of all respective tissues and organs were represented based on a color scale, showing intensities ranging from 8.34X10^4 (blue) – 1.28X10^6 (red) ph/s/cm^2/sr. (c) Quantification of luciferase expression *ex vivo* in harvested heart and skeletal muscles at week 5 and week 8 from C57BL6 mice injected intravenously with scAAV9-luciferase vectors (n=3 mice per group, dose:

$5x10^{10}$ vg/mouse) containing the different promoters (i.e. *SPc5-12; CMV; CS-CRM4/SPc 5-12)*. Values were expressed as mean + s.e.m. (n=3 mice per group). Gastroc.: gastrocnemius; Quads: quadriceps

Fig. 1

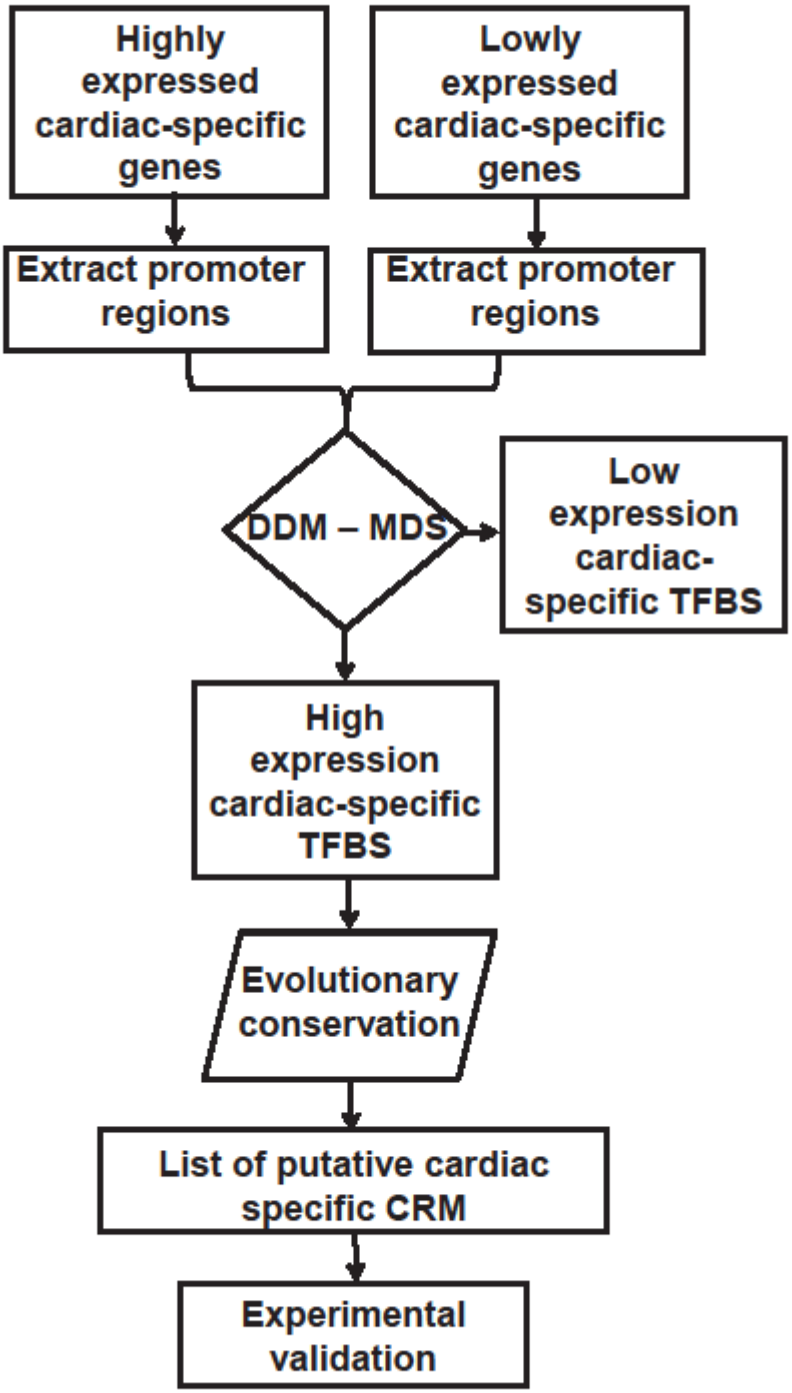

Fig. 2

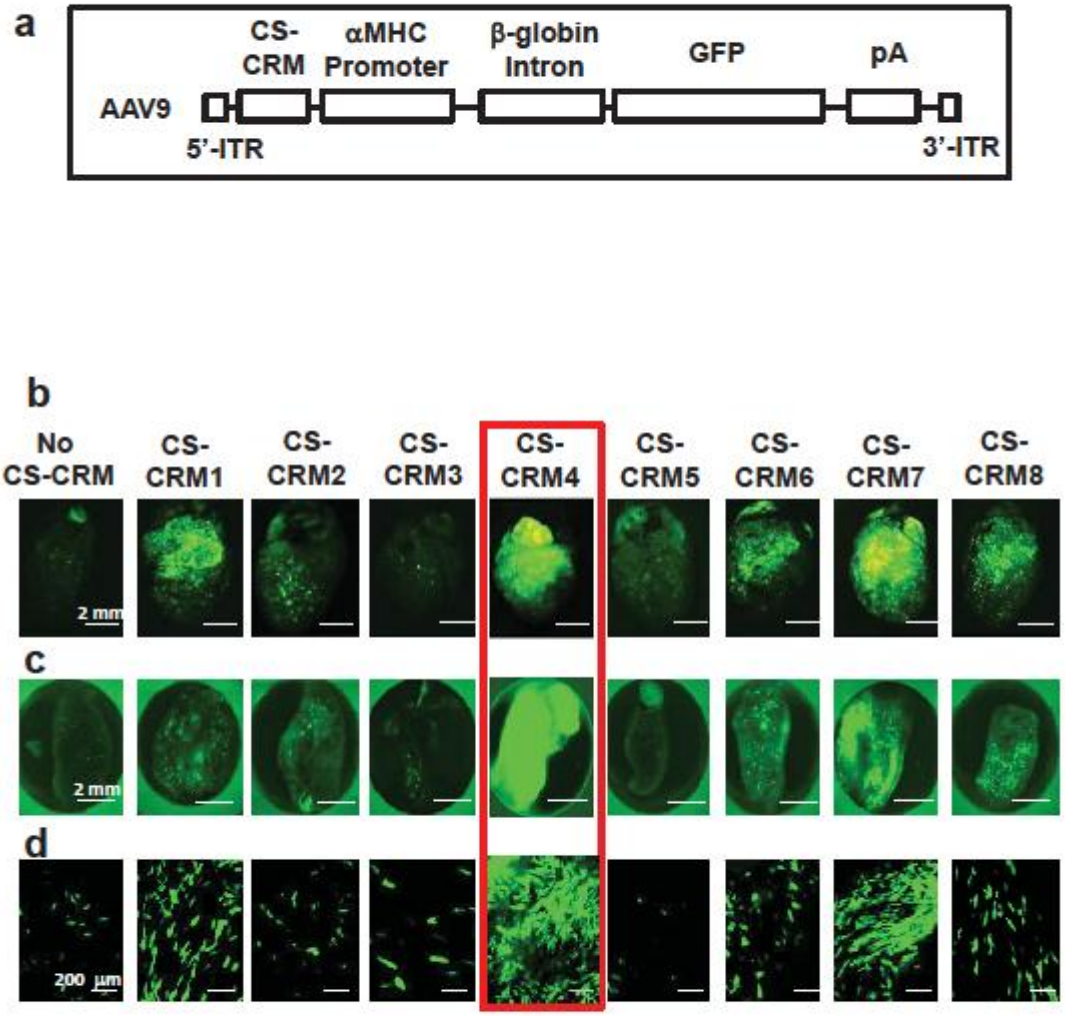

Fig. 3

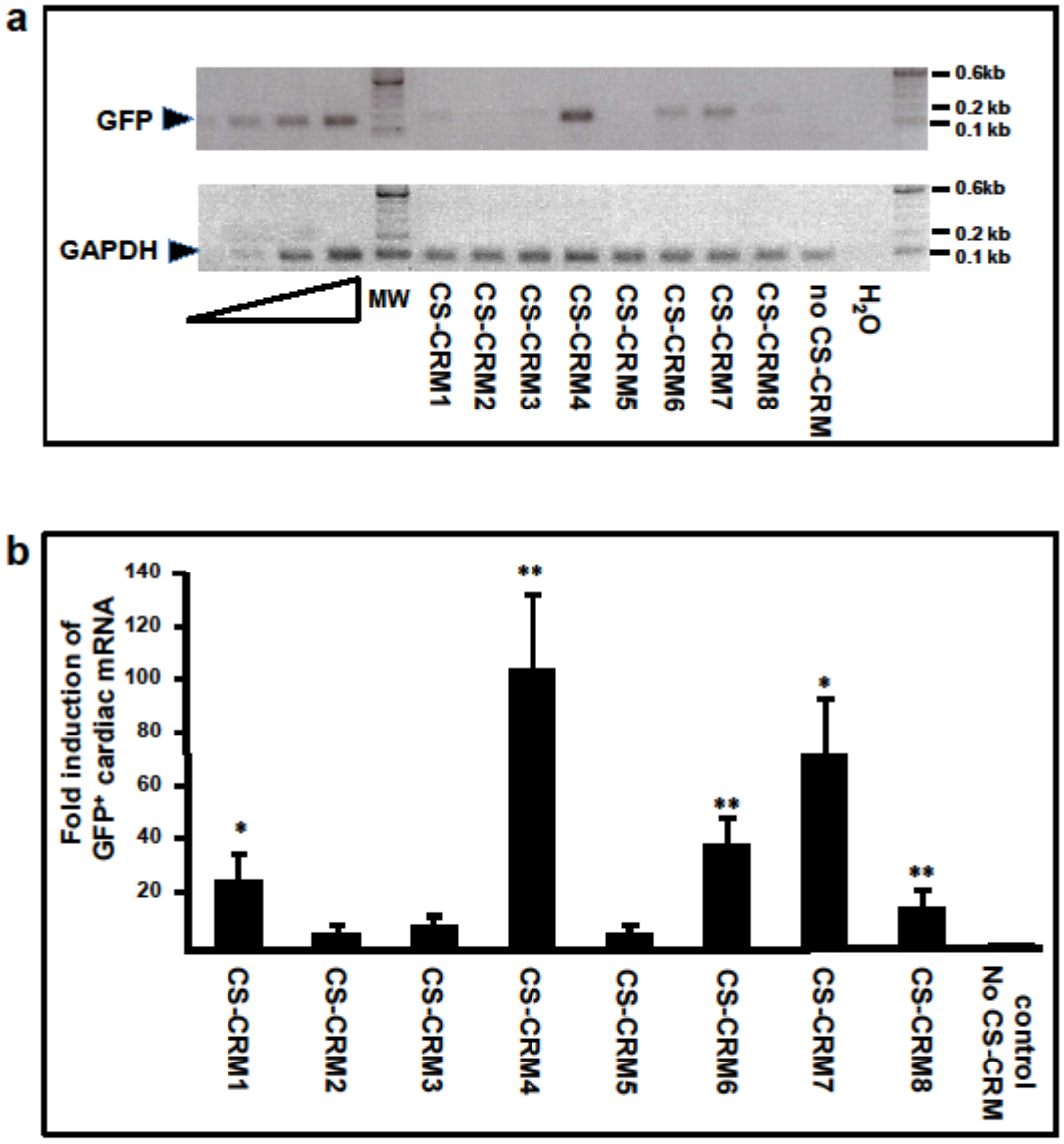

Fig. 4

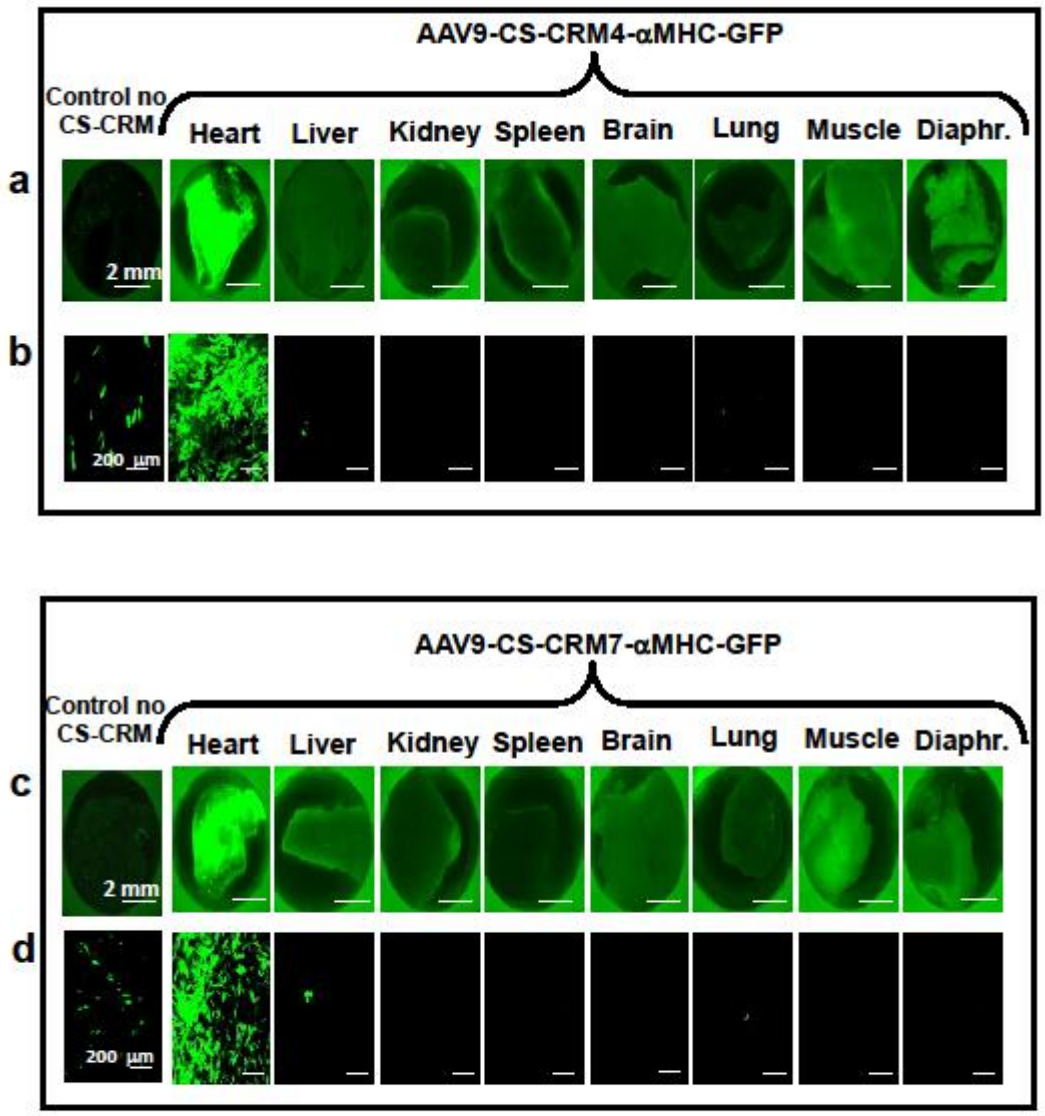

Fig. 5

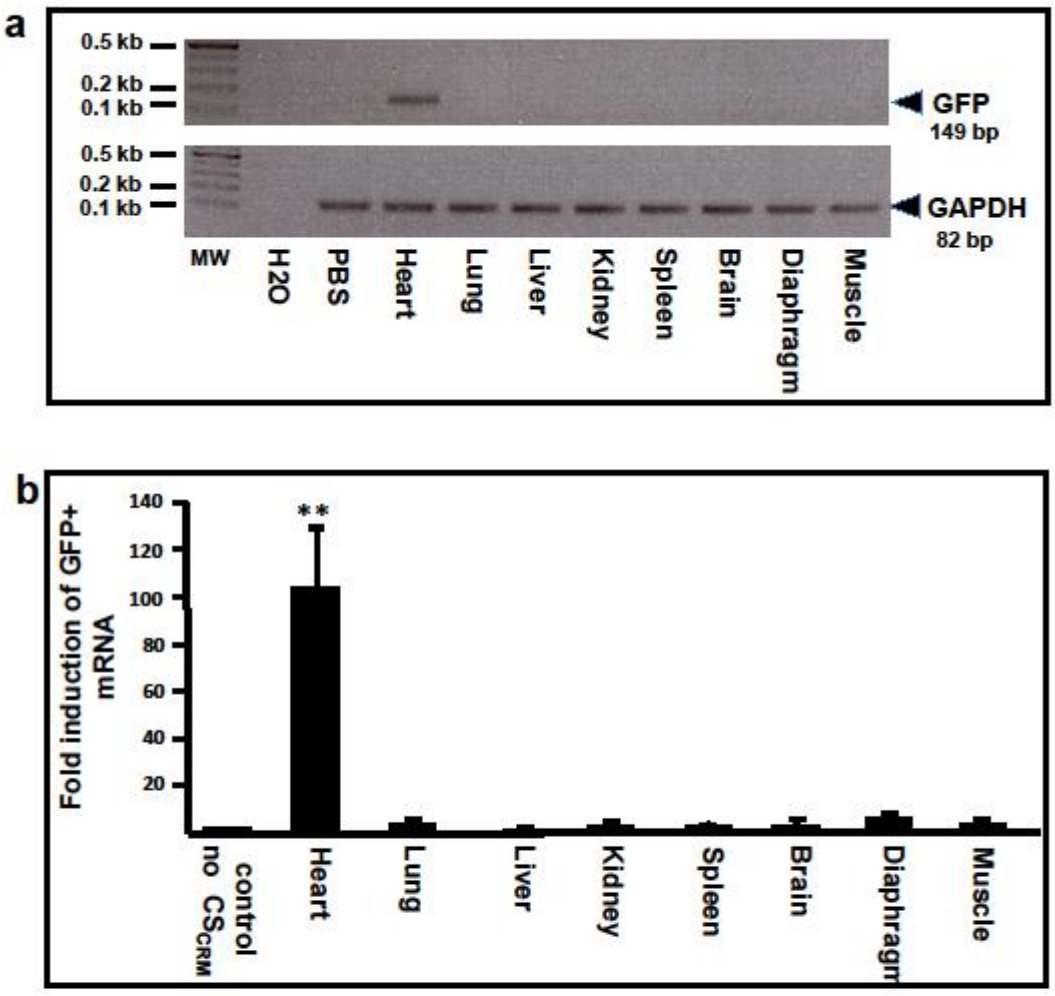

Fig. 6

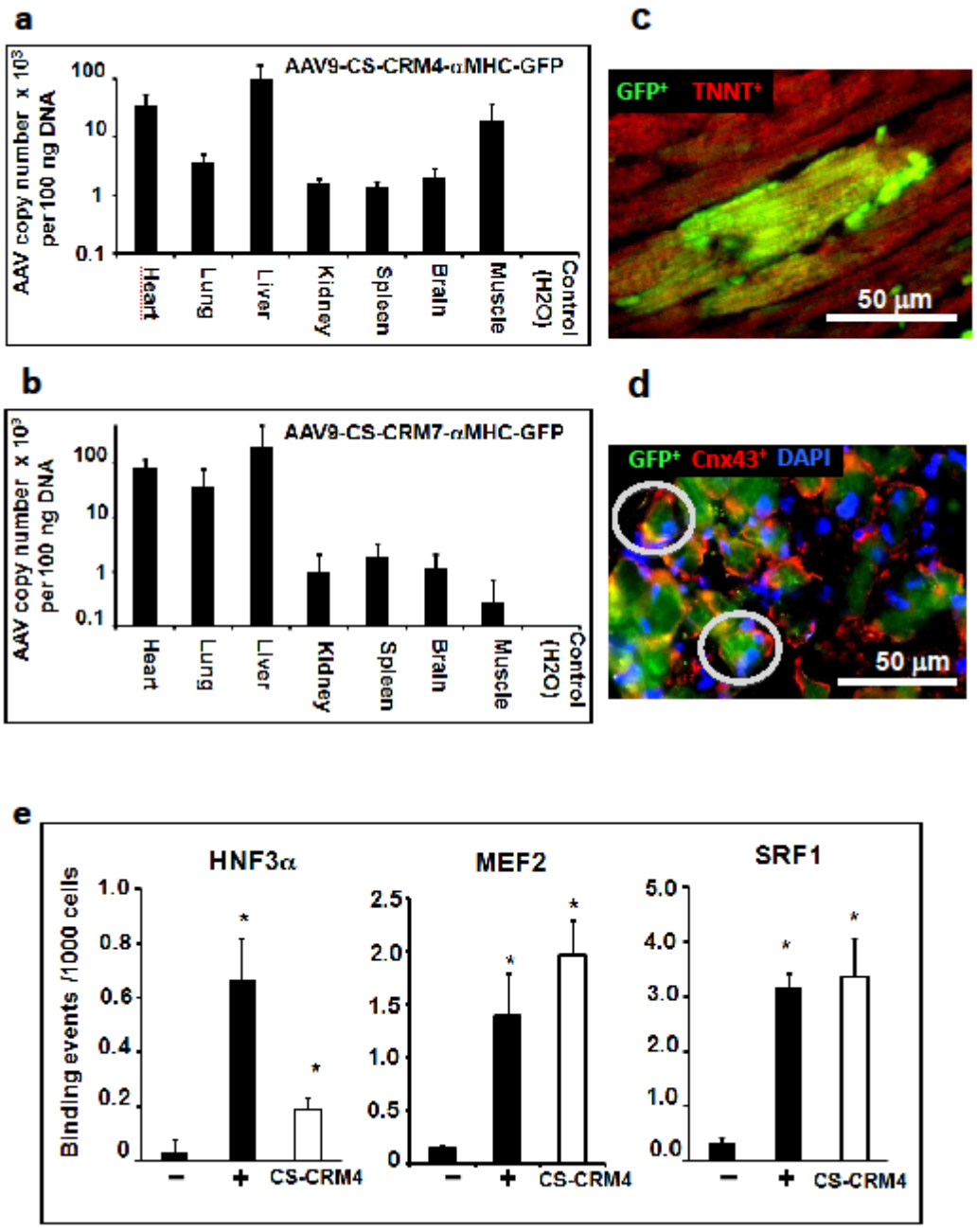

Fig. 7

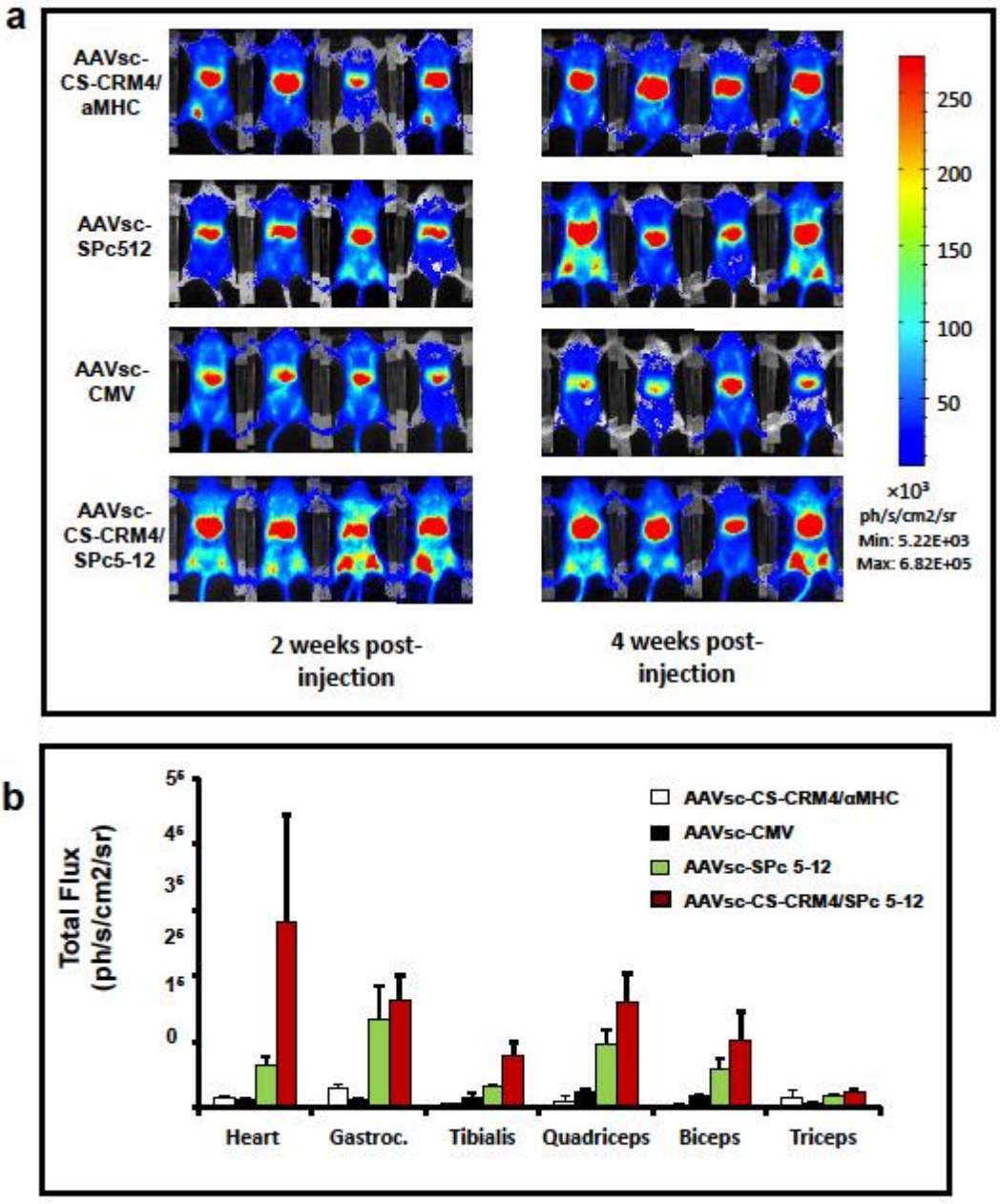

Figure 8