Genome-wide Computational Analysis Reveals Cardiomyocyte-specific Transcriptional Cis-regulatory Motifs That Enable Efficient Cardiac Gene Therapy

Melvin Y Rincon1,2, Shilpita Sarcar1, Dina Danso-Abeam2, Marleen Keyaerts3, Janka Matrai1,2, Ermira Samara-Kuko1,2, Abel Acosta-Sanchez4, Takis Athanasopoulos5, George Dickson5, Tony Lahoutte3, Pieter De Bleser6, Thierry VandenDriessche1,2 and Marinee K Chuah1,2

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

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 nontarget 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, may provoke an untoward immune responses against the gene-modified target cells and/or the therapeutic transgene.

Received 16 April 2014; accepted 29 August 2014; advance online publication 7 October 2014. doi:10.1038/mt.2014.178

The two first authors contributed equally to this work.

Correspondence: Marinee Chuah, Free University of Brussels (VUB), Department of Gene Therapy & Regenerative Medicine, Faculty of Medicine & Pharmacy, Building D, room D306, Laarbeeklaan 103, Brussels B-1090, Belgium. E-mail: marinee.chuah@vub.ac.be or Thierry VandenDriessche, Free University of Brussels (VUB), Department of Gene Therapy & Regenerative Medicine, Faculty of Medicine & Pharmacy, Building D, room D306, Laarbeeklaan 103, Brussels B-1090, Belgium. E-mail: thierry.vandenDriessche@vub.ac.be
product. The use of cardiac-specific promoters may improve transcriptional targeting in cardiomyocytes and potentially overcome some of the limitations of using ubiquitously expressed promoters. 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 multipronged approach that relies not only on optimizing “transductional” targeting but also boosts expression by enhancing “transcriptional targeting.”

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. 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.

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 (CS-CRMs). 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 CS-CRMs boost transcriptional targeting after cardiac gene therapy up to 100-fold. This type of multidisciplinary approach—at the nexus of genomics, computational biology, and gene therapy—remains 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 applications 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 multistep computational approach that allowed us to identify evolutionary-conserved CS-CRMs associated with genes that are highly expressed in the heart (Figure 1). This in silico strategy was initially developed to identify CRMs associated with differential gene expression following specific in vitro stimuli. However, to our knowledge, this type of bioinformatics 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 overrepresentation of a given TFBS but also its context-dependent co-occurrence with other TFBS on a genome-wide scale. 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-CRMs ranging from size of 117 to 689 bp were identified for the heart (Table 1 and Supplementary Table S1, Supplementary Figure S1). These CS-CRMs comprised binding sites for eight different TFs including SRF, CTF/NF1, MEF2, R5R5C4, COUP-TF1, HIF1, 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), Casq2 (CS-CRM4), Casq2 (CS-CRM5), Ankr1 (CS-CRM6), Ankr1 (CS-CRM7), and Ankr1 (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 (Supplementary Table S1 and Supplementary 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.
Table 1: Transcription factor binding sites (TFBS) strongly associated with high cardiac-specific expression.

| Name   | Gene                        | Length (bp) | TFBS                                                                 |
|--------|-----------------------------|-------------|----------------------------------------------------------------------|
| CS-CRM1| Myf3                        | 150         | SRF, NF1, MEF2, RSRFC4, COUP-TF1, HFH1, HNF3α, HNF3β                 |
| CS-CRM2| Bm7                        | 689         | HNF3α, HNF3β, SRF, RSRFC4, NF1                                      |
| CS-CRM3| Mef2                       | 183         | HNF3α, HNF3β, MEF2, SRF, NF1, RSRFC4                                |
| CS-CRM4| Casq2b                     | 219         | HNF3α, MEF2, SRF, NF1, RSRFC4, HNF3β, HFH1                          |
| CS-CRM5| Ankrd1α                    | 299         | HNF3α, HNF3β, MEF2, SRF, NF1, RSRFC4                                |
| CS-CRM6| Ankrd1β                    | 277         | MEF2, HNF3α, HFH1, HNF3β, NF1                                       |
| CS-CRM7| Ankrd1α                    | 397         | HNF3α, HNF3β, MEF2, SRF, COUP-TF1, NF1                              |

Evolutionarily 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. CS-CRM, cardiac-specific cis-regulatory module.

In vivo validation of CS-CRMs

To validate the de novo-designed CS-CRM in vivo, we generated AAV that expressed the humanized green fluorescent reporter protein (GFP) from a chimeric promoter. This promoter was composed of the heart-specific αMHC promoter linked to the different CS-CRM (Figure 2a). We selected the AAV9 serotype to obtain efficient cardiac gene transfer after intravenous injection of 10\(^{11}\) viral genome (VG) in C57BL/6 mice. Seventy percentage of the de novo-designed CS-CRM (five out of eight: 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 (Figure 3a,b), consistent with the increase in GFP expression levels (Figure 2b–d). In particular, the CS-CRM4 and CS-CRM7 elements resulted in a significant 100- and 70-fold (P < 0.01) increase in GFP messenger RNA (mRNA) expression respectively, compared to the control without CS-CRM (Figure 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 expression levels in the heart (Figure 4a–d). This was confirmed at two different

![Figure 2](image-url)
Combining synthetic promoters with CS-CRM4s

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., Sp5-12) \(^{16}\) and tested its impact on gene expression in these tissues. This chimeric synthetic CS-CRM4/Sp5-12 promoter was used to express luciferase from a self-complementary AAV9 (scAAV9), that further increases transduction compared to single-stranded AAV. The AAV terminal repeat 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 mice and determined luciferase expression by bioluminescent imaging 2 and 4 weeks postinjection (Figure 7a). 

Ex vivo analysis of the organs was performed 5 weeks postinjection (Supplementary Figures S5 and S6). Comprehensive comparative analysis by bioluminescent imaging revealed that this de novo-designed chimeric synthetic CS-CRM4/Sp5-12 combination was the most potent compared to any of the other promoters tested (Figure 7b,c and Supplementary Figure S5). In particular, a robust and significant 20-fold increase in cardiac expression was obtained using the CS-CRM4/Sp5-12 promoter compared to the 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/\(\alpha\)MHC promoter was comparable to that of CMV (Figure 7b,c and Supplementary Figure S5). Moreover, CS-CRM4/Sp5-12 was also more robust than the CS-CRM4/\(\alpha\)MHC chimera promoter, yielding a significant 19-fold increase in luciferase activity (Figure 7b,c). Finally, we demonstrated that the chimeric synthetic CS-CRM4/Sp5-12 promoter resulted in a nearly fivefold increase in cardiac gene expression compared to Sp5-12, whereas only a slight or no increase in skeletal muscle expression was apparent (Figure 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 (Figure 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., \(\alpha\)MHC versus Sp5-12). Reporter gene expression was absent in most of the other organs and tissues (i.e., kidney, spleen, brain, and diaphragm) (Supplementary Figure S6). Nevertheless, we observed minor expression in the lungs of the mice injected with the vector containing the Sp5-12 promoter. Moreover, expression of the reporter gene was observed in the liver (Supplementary Figures S6 and S7). These observations challenge the prevailing assumption that the synthetic Sp5-12 promoter is only expressed in heart and skeletal muscle.

We subsequently injected the AAV9sc-CS-CRM4/Sp5-12 vector (5 \(\times\) 10\(^{10}\) vg intravenously) into 4-week-old C57BL/6 mice (male) and monitored long-term luciferase expression for...
up to 8 weeks postinjection. Similar results were now obtained in immune-competent C57BL/6 mice as described above in the severe combined immune-deficient model (Figure 8 and Supplementary 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 severe combined immune-deficient mice. Moreover, a significant twofold increase in luciferase expression was apparent compared to AAV9sc-Spc5-12.

DISCUSSION
The current study explores the use of bioinformatics 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 CS-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 αMHC 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 Calcium Upregulation by Percutaneous Administration of

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) (1,011 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 semiautomatic fashion using the Zeiss Axiovert 200 inverted microscope. Panels are show at ×5 magnification, 2-second exposure time (a, c). Representative images by confocal microscopy are shown at ×10 magnification (b, d). Six images were obtained from different fragments of each organ, n = 3 mice per group; a representative image is shown. α-MHC, α-myosin heavy chain; AAV9, adeno-associated virus serotype 9; CS-CRM, cardiac-specific cis-regulatory module; GFP, green fluorescent protein; vg, viral genome.
Gene Therapy in Cardiac Disease (CUPID) trial that is based on AAV1-SERCA2a myocardial transduction. 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. These cellular immune responses were previously shown to eliminate AAV-transduced cells, at least in liver. To facilitate transduction in the face of preexisting anti-AAV antibodies would be more challenging and would require AAV capsid decoys instead or plasma exchange. In addition to the use of an optimal CS-CRM/synthetic promoter combination, the AAV serotype itself significantly impacts on cardiac tropism and expression. 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.

The present in silico approach does not only allow for the identification of TFBS that are overrepresented in CS-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-specific 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 overrepresentation of a single TFBS element. 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 overrepresentation of a given TFBS, regardless of its context, as discussed previously. The current computational approach that defines the CS-CRMs has been further improved beyond the initial description of the algorithm, 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. 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 this study to analyze this further but it sets the stage to further refine the computational analysis. This 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 toward 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.

This 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 toward a more comprehensive approach toward 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 nonhuman 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 copy numbers per 100 ng DNA.

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 (10¹¹ vg/mouse) (a) or AAV9-CS-CRM7-αMHC-GFP (10¹¹ vg/mouse) (n = 3) (b) determined by qPCR using vector specific primers. Relative to a standard curve of known vector copy numbers (mean ± SD) are shown. (c, d) Heart sections of mice injected with AAV9-CS-CRM4-αMHC-GFP (10¹¹ 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 pericellular location of Cnx43 on cardiomyocytes consistent with its participation in intercellular gap junctions essential for Ca²⁺ handling. (e) Chromatin immunoprecipitation (ChIP) assay on heart of mice injected with AAV9-CS-CRM4-αMHC-GFP (10¹¹ 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 ± SD). α-MHC, α-myosin heavy chain; AAV9, adeno-associated virus serotype 9; CS-CRM, cardiac-specific cis-regulatory module; DAPI, 4',6-diamidino-2-phenylindole; GFP, green fluorescent protein; qPCR, quantitative PCR; TFBS, transcription factor binding site; vg, viral genome. *P ≤ 0.05.
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

MATERIALS AND METHODS

Identification of CS-CRM by computational analysis. A computational sequential approach (Figure 1) was used to identify tissue-specific CRMs: (i) identification of cardiac-specific genes that are highly expressed based on statistical analysis of microarray expression data of normal human heart tissues; (ii) extraction of the corresponding promoter sequences from publicly available databases; (iii) identification of the regulatory modules and the TFBS using a differential distance matrix/multidimensional scaling approach;18 (iv) searching the genomic context of the highly expressed genes for evolutionary-conserved clusters of TFBS (i.e., CRMs). A detailed outline of the bioinformatics approach is outlined in the Supplementary Materials and Methods.

Generation of CRM constructs. The CS-CRMs (CS-CRM1 to CS-CRM8) (Table 1) were synthesized by conventional oligonucleotide synthesis. The different CS-CRM were cloned upstream of the αMHC,9 in the context of an AAV backbone (Figure 2a). Additionally, the CS-CRM4 was cloned upstream of the novel synthetic promoter Spc5-1216 using a scAAV backbone.19 The luciferase (Luc) complementary DNA was used to substitute the GPA 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 CMV promoter 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 vg/ml) was determined by quantitative reverse transcriptase-PCR. Typically, for all vectors we achieved titers in the normal range of 2×1012 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, 2 days old C57BL/6 mice were injected with vectors containing the different CS-CRM elements (2×109 to 3×1011 vg/mouse) via periorbital vein. Mice were euthanized by cervical dislocation after inhaled anesthesia using isoflurane 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 postinjection organs were harvested and examined. For the
bioluminescent imaging analysis (see below), 10^10 vg scAAV9-luciferase vectors were injected in adult severe combined immune-deficient mice.

**mRNA analysis.** Quantitative PCR 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 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 ± SEM (n = 3 mice per group, dose: 5 × 10^10 vg/mouse) containing the different promoters (i.e., SPc5-12; CMV; CS-CRM4/SPc5-12). Images were taken 2, 4, and 8 weeks postinjection. Whole body bioluminescent images were represented based on a color scale, showing intensities ranging from 1.28 × 10^6 (blue) to 1.45 × 10^7 (red) photons/second/cm^2/sr.

**Detection of luciferase transgene expression.** Luciferase expression was analyzed using an in vivo Photon Imager ( Biospace, Paris, France) after intravenous injection of 50 μl luciferin substrate per 10 g of body weight (150 mg/kg; Promega, Madison, WI). Representative in vivo bioluminescence transgene expression images were taken at 2, 4, and 8 weeks postvector injection. Subsequently at 5 and 8 weeks postviral vector injection, the organs were removed and bioluminescence signal was quantified during 5 minutes.

**SUPPLEMENTARY MATERIAL**

**Figure S1.** Organization of the CS-CRM elements.

**Figure S2.** Dose–response analysis of cardiac-specific CRM (CS-CRM).

**Figure S3.** Tissue specificity of cardiac-specific CRM (CS-CRM).

**Figure S4.** Cardiac transduction efficiency analysis.

![Figure 8](image-url)
Figure S5. Ex vivo determination of luciferase expression in heart and skeletal muscle.

Figure S6. Ex vivo luciferase expression in liver, lung, kidney, spleen, brain, and diaphragm.

Figure S7. Quantification of ex vivo luciferase expression in liver, kidney, spleen, brain, and diaphragm.

Figure S8. Ex vivo luciferase expression and quantification in lung, liver, kidney, spleen, brain, and diaphragm.

Table S1. DNA sequence of CS-CRM.

Materials and Methods

ACKNOWLEDGMENTS

This study was supported by grants from EU FP7 (222878, PERSISt), AFM, FWO, EHA, Geconcerteerde Onderzoeksaaltje (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 M.K.C. and T.V.L. and M.K.C. are Senior Clinical Investigators of the Research Foundation—Flaniders (Belgium) (FWO).

The research at ICMI is funded by the Belgian State, National Kankerplan, Vlaamse Liga tegen Kanker, and Vlaamse Stichting tegen Kanker. M.Y.R. is a recipient of a Fundación Cardiovascular de Colombia Volcicollas Fellowship and an AFM PhD Fellowship. We thank Bing Yan for technical assistance. 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. and A.A. performed molecular biology experiments and vector production/characterization and animal experiments. P.D. conducted the bioinformatics analysis and wrote part of the manuscript. M.K.C. and T.L. contributed to the bioinformatics imaging studies. G.D. contributed to the synthetic promoter studies. T.V. and M.K.C. designed experiments, co-ordinated the work, and wrote the manuscript. T.A. and G.D. contributed the SPCs-12. T.V. and M.K.C. are joined senior corresponding authors. The authors declare no competing financial interests.

REFERENCES

1. Zsebo, K, Varoshinsky, A, Rudy, J, Wagner, K, Greenberg, B, Jessup, M, et al. (2014). Long-term effects of AAV1/SECA2a 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, Cappolla, T, Pauly, DF, Jaski, B, et al.; Calcium Upregulation by Permissive Administration of Gene Therapy in Cardiac Disease (CUPID) Investigators. (2011). Calcium upregulation by permissive administration of gene therapy in cardiac disease (CUPID): a phase 2 trial of intraocular gene therapy of sancroptilicum Ca2+-ATPase in patients with advanced heart failure. Circulation 124: 313–318.

3. Lai, Y and Duan, D (2012). Progress in gene therapy of dystrophic heart disease. Mol Ther 14: 1319–1329.

4. Manno, CS, Pierce, GF, Arruda, VR, Glader, B, Ragni, M, Rasko, JJ et al. (2004). Treatment of phenylketonuria using minicircle-based naked-DNA gene transfer to murine liver. Hepatology 39: 1035–1043.

5. Chuah, MK, Petrus, I, De Bleser, P, Le Guiner, C, Gernoux, G, Adjali, O et al. (2014). Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response. Nat Med 12: 342–347.

6. Mingozzi, F, Maus, MV, Hui, DJ, Sabatino, DE, Murphy, SL, Rasko, JE et al. (2007). CD8(+) T-cell responses to adeno-associated virus capsid in patients with advanced heart failure. Circulation 116: 3594–3600.

7. Lai, Y and Duan, D (2012). Progress in gene therapy of dystrophic heart disease. Gene Ther 19: 678–685.

8. Manno, CS, Pierce, GE, Arruda, VR, Glader, B, Ragni, M, Rasko, JE et al. (2007). CD8(+) T-cell responses to adeno-associated virus capsid in humans. Nat Med 13: 419–422.

9. Nathwani, AC, Tuddenham, EG, Ranajaran, S, Rosales, C, McIntosh, J, Linch, DC et al. (2011). Adenovirus-associated virus-mediated gene transfer in hemophilia B. N Engl J Med 365: 2357–2365.

10. Brown, BD, Cantore, A, Mancini, D, Cappolla, T, Pauly, DF, Jaski, B, et al.; Calcium Upregulation by Permissive Administration of Gene Therapy in Cardiac Disease (CUPID) Investigators. (2011). Calcium upregulation by permissive administration of gene therapy in cardiac disease (CUPID): a phase 2 trial of intraocular gene therapy of sancroptilicum Ca2+-ATPase in patients with advanced heart failure. Circulation 124: 313–318.

11. Inagaki, K, Fues, S, Storm, TA, Gibson, GA, Mietzen, CE, 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 14: 45–53.

12. Vandendriessche, T, Zhao, C, Andrutis, A, Peters, 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 5: 16–24.

13. Pesando, J, Bunge, H, Koller, DM, Bird, MJ, Giron, A, Wendtner, CM et al. (2003). In vivo selection of viral vectors with modified tropism: the adeno-associated virus display. Mol Ther 15: 151–157.