experimental section, the inventors could identify nucleic acid regulatory elements that specifically enhanced gene expression in both heart and skeletal muscle, and skeletal muscle-specific nucleic acid regulatory elements. These regulatory elements were subsequently 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:

In a first aspect, there is provided a nucleic acid regulatory element for enhancing muscle-specific gene expression, comprising a sequence selected from the group consisting of: SEQ ID NO:10, SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, a sequence having at least 95% identity to any of said sequences, or a functional fragment thereof, wherein said functional fragment comprises at least 20, preferably at least 25, contiguous nucleotides from the sequence from which it is derived, and wherein said functional fragment comprises at least 1, preferably at least 5, of the transcription factor binding sites (TFBS) that are present in the sequence from which it is derived.

In a second aspect, there is provided a nucleic acid regulatory element for enhancing muscle-specific gene expression hybridizing under stringent conditions to the nucleic acid regulatory element of the first aspect, or to its complement.

In a third aspect, there is provided use of the nucleic acid regulatory element of the first or second aspect in a nucleic acid expression cassette, or a vector.

In a fourth aspect, there is provided a nucleic acid expression cassette comprising at least one nucleic acid regulatory element of the first or second aspect, operably linked to a promoter and a transgene.

In a fifth aspect, there is provided a vector comprising the nucleic acid regulatory element of the first or second aspect, or the nucleic acid expression cassette of the fourth aspect.

In a sixth aspect, there is provided a pharmaceutical composition comprising the nucleic acid expression cassette of the fourth aspect, or the vector of the fifth aspect, and a pharmaceutically acceptable carrier.

In a seventh aspect, there is provided the nucleic acid regulatory element of the first or second aspect, the nucleic acid expression cassette of the fourth aspect, the vector of the fifth aspect, or the pharmaceutical composition of the sixth aspect for use in gene therapy.

In an eighth aspect, there is provided the nucleic acid regulatory element of the first or second aspect, the nucleic acid expression cassette of the fourth aspect, the vector of the fifth aspect, or the pharmaceutical composition of the sixth aspect for use in vaccination therapy, preferably prophylactic vaccination.

In a ninth aspect, there is provided an *in vitro* or *ex vivo* method for expressing a transgene product in muscle cells comprising:

- introducing the nucleic acid expression cassette of the fourth aspect, or the vector of the fifth aspect into the muscle cells;

- expressing the transgene product in the muscle cells.

The present disclosure also provides the following embodiments:

~~Aspect~~ Embodiment 1: a nucleic acid regulatory element for enhancing muscle-specific gene expression comprising, consisting essentially of, or consisting of a functional fragment of a sequence selected from the group consisting of: SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, 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, 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 5, more preferably at least 10 or at least 15, of the transcription factor binding sites (TFBS) that are present in the sequence from which it is derived.

Embodiment~~Aspect~~ 2: the nucleic acid regulatory element according to ~~aspect~~ embodiment 1 for enhancing cardiac and skeletal muscle-specific gene expression, comprising a functional fragment of a sequence selected from the group consisting of: SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, 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, wherein said functional fragment comprise 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 5, more preferably at least 10 or at least 15, of the transcription factor binding sites (TFBS) that are present in the sequence from which it is derived.

Embodiment~~Aspect~~ 3: the nucleic acid regulatory element according to embodiment~~aspect~~ 1 for enhancing skeletal muscle-specific gene expression comprising a functional fragment of a sequence selected from the group consisting of: SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, 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, wherein said functional fragment comprise 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 5, more preferably at least 10 or at least 15, of the transcription factor binding sites (TFBS) that are present in the sequence from which it is derived.

EmbodimentAspect 4: the nucleic acid regulatory element according to embodimentaspect 1, 2, or 3, comprising, consisting essentially of, or consisting of a sequence selected from the group consisting of: the nucleotide sequence from position 33 to 58 in SEQ ID NO:1; the nucleotide sequence from position 90 to 142 in SEQ ID NO:1; the nucleotide sequence from position 143 to 233 in SEQ ID NO:1; the nucleotide sequence from position 240 to 310 in SEQ ID NO:1; the nucleotide sequence from position 90 to 233 in SEQ ID NO:1; the nucleotide sequence from position 47 to 130 in SEQ ID NO:5; the nucleotide sequence from position 252 to 293 in SEQ ID NO:5; the nucleotide sequence from position 330 to 450 in SEQ ID NO:5; the nucleotide sequence from position 10 to 180 in SEQ ID NO:10 (i.e. SEQ ID NO:37); the nucleotide sequence from position 190 to 240 in SEQ ID NO:10 (i.e. SEQ ID NO:38); the nucleotide sequence from position 241 to 300 in SEQ ID NO:10 (i.e. SEQ ID NO:39); the nucleotide sequence from position 241 to 360 in SEQ ID NO:10 (i.e. SEQ ID NO:41); the nucleotide sequence from position 380 to 420 in SEQ ID NO:10 (i.e. SEQ ID NO:40); or a sequence having at least 95% identity to any of said sequences.

EmbodimentAspect 5: the nucleic acid regulatory element according to any one of embodimentaspects 1 to 4 comprising, consisting essentially of or consisting of a sequence selected from the group consisting of: the nucleotide sequence from position 33 to 310 in SEQ ID NO:1; the nucleotide sequence from position 47 to 450 in SEQ ID NO:5; the nucleotide sequence from position 10 to 420 in SEQ ID NO:10, or a sequence having at least 95% identity to any of said sequences.

EmbodimentAspect 6: the nucleic acid regulatory element according to any one of embodimentaspects 1 to 5, comprising, consisting essentially of, or consisting of a sequence selected from the group consisting of: SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, 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.

EmbodimentAspect 7: the nucleic acid regulatory element according to any one of embodimentaspects 1, 2, or 4 to 6 for enhancing cardiac and skeletal muscle-specific gene expression comprising, consisting essentially of, or consisting of a sequence selected from the group consisting of: SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, 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.

EmbodimentAspect 8: the nucleic acid regulatory element according to any one of embodimentaspects 1, 3 to 6 for enhancing skeletal muscle-specific gene expression comprising, consisting essentially of, or consisting of a sequence selected from the group consisting of: SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, 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.

EmbodimentAspect 9: the nucleic acid regulatory element according to any one of embodimentaspects 1 to 8, comprising, consisting essentially of, or consisting of a sequence selected from the group consisting of: SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, a functional fragment thereof comprising SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, or SEQ ID NO:13, 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.

EmbodimentAspect 10: the nucleic acid regulatory element according to any one of embodimentaspects 1, 2, 4 to 7, 9, comprising, consisting essentially of, or consisting of a sequence selected from the group consisting of: SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, a functional fragment thereof comprising SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, or SEQ ID NO:6, 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.

EmbodimentAspect 11: the nucleic acid regulatory element according to any one of embodimentaspects 1, 3 to 6, 8, or 9, comprising, consisting essentially of, or consisting of a sequence selected from the group consisting of: SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, a functional fragment thereof comprising SEQ

ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, or SEQ ID NO:13, 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.

EmbodimentAspect 12: a nucleic acid regulatory element for enhancing muscle-specific gene expression comprising, consisting essentially of, or consisting of the complement of a sequence as defined in any one of embodimentaspects 1 to 11.

EmbodimentAspect 13: a nucleic acid regulatory element for enhancing muscle-specific gene expression hybridizing under stringent conditions to the nucleic acid regulatory element according to any one of embodimentaspects 1 to 12.

EmbodimentAspect 14: the nucleic acid regulatory element according to any one of embodimentaspects 1 to 13, having a maximal length of 1000 nucleotides, preferably 800 nucleotides, more preferably 600 nucleotides, such as 550 nucleotides, 500 nucleotides, 450 nucleotides, 400 nucleotides, 350 nucleotides, or 300 nucleotides, still comprising the regulatory element defined by any one of SEQ ID Nos: 1-13, or any one of the functional fragments thereof defined in any one of embodimentaspects 3 to 5.

EmbodimentAspect 15: use of the nucleic acid regulatory element according to any one embodimentaspects 1 to 14 in a nucleic acid expression cassette, or a vector, more particularly for enhancing muscle-specific expression of said nucleic acid expression cassette or vector.

EmbodimentAspect 16: a nucleic acid expression cassette comprising at least one, such as one, two, three, four, five or more, nucleic acid regulatory element according to any one of embodimentaspects 1 to 14, operably linked to a promoter.

EmbodimentAspect 17: the nucleic acid expression cassette according to embodimentaspect 16, wherein the nucleic acid regulatory element is operably linked to a promoter and a transgene.

EmbodimentAspect 18: the nucleic acid expression cassette according any one of embodimentaspects 16 or 17, wherein the promoter is a muscle-specific promoter, preferably the promoter from the desmin (DES) gene.

EmbodimentAspect 19: the nucleic acid expression cassette according to any one of embodimentaspects 17 or 18, wherein the transgene encodes a therapeutic protein or an immunogenic protein.

EmbodimentAspect 20: the nucleic acid expression cassette according to any one of embodimentaspects 17 to 19, wherein the transgene encodes a secretable protein or a structural protein, such as dystrophin or sarcoglycan.

EmbodimentAspect 21: the nucleic acid expression cassette according to any one of embodimentaspects 16 to 20, further comprising an intron, preferably the Minute Virus of Mouse (MVM) intron.

EmbodimentAspect 22: the nucleic acid expression cassette according to any one of embodimentaspects 16 to 21, further comprising a polyadenylation signal, preferably the Simian Virus 40 (SV40) polyadenylation signal.

EmbodimentAspect 23: a vector comprising the nucleic acid regulatory element according to any one of embodimentaspects 1 to 14, or the nucleic acid expression cassette according to any one of embodimentaspects 16 to 22.

EmbodimentAspect 24: the vector according to embodimentaspect 23, which is a viral vector, preferably an adeno-associated viral (AAV) vector, more preferably an AAV9 vector.

EmbodimentAspect 25: the vector according to embodimentaspect 23, which is a non-viral vector, preferably a plasmid, a minicircle or a transposon-based vector, such as a PiggyBac-based vector or a Sleeping Beauty-based vector.

EmbodimentAspect 26: a pharmaceutical composition comprising the nucleic acid expression cassette according to any one of embodimentaspects 16 to 22, or the vector according to any one of embodimentaspects 23 to 25, and a pharmaceutically acceptable carrier.

EmbodimentAspect 27: the nucleic acid regulatory element according to any one of embodimentaspects 1 to 14, the nucleic acid expression cassette according to any one of embodimentaspects 16 to 22, the vector according to any one of embodimentaspects 23 to 25, or the pharmaceutical composition according to embodimentaspect 26 for use in medicine.

EmbodimentAspect 28: the nucleic acid regulatory element according to any one of embodimentaspects 1 to 14, the nucleic acid expression cassette according to any one of embodimentaspects 16 to 22, the vector according to any one of embodimentaspects 23 to 25, or the pharmaceutical composition according to embodimentaspect 26 for use in gene therapy, preferably muscle-directed gene therapy.

EmbodimentAspect 29: the nucleic acid regulatory element, the nucleic acid expression cassette, the vector, or the pharmaceutical composition according for use according to embodimentaspect 28, wherein the gene therapy is for muscular dystrophy (e.g. Duchenne muscular dystrophy (DMD) / Becker muscular dystrophy (BMD)), myotonic dystrophy, neuromuscular disease, motor neuron diseases (MND), such as e.g. Charcot-Marie-Tooth disease (CMT), spinal muscular atrophy (SMA), and amyotrophic lateral sclerosis (ALS), Emery-Dreifuss muscular dystrophy, facioscapulohumeral muscular dystrophy (FSHD), congenital muscular dystrophies, congenital myopathies, limb girdle muscular dystrophy, metabolic myopathies, muscle inflammatory diseases, myasthenia, mitochondrial myopathies,

anomalies of ionic channels, nuclear envelop diseases, cardiomyopathies, cardiac hypertrophy, heart failure, distal myopathies, cardiovascular diseases, hemophilia, *including hemophilia A and B, and diabetes.*

EmbodimentAspect 30: the nucleic acid regulatory element according to any one of embodimentaspects 1 to 14, the nucleic acid expression cassette according to any one of embodimentaspects 16 to 22, the vector according to any one of embodimentaspects 23 to 25, or the pharmaceutical composition according to embodimentaspect 26 for use as a vaccine, preferably a prophylactic vaccine, or for use in vaccination therapy, preferably prophylactic vaccination.

EmbodimentAspect 31: A method, preferably an *in vitro* or *ex vivo* method, for expressing a transgene product in muscle cells, preferably heart muscle cells and/or skeletal muscle cells, comprising:

- introducing the nucleic acid expression cassette according to any one of embodimentaspects 16 to 22, or the vector according to any one of embodimentaspects 23 to 25 into the muscle cells;

- expressing the transgene product in the muscle cells.

BRIEF DESCRIPTION OF DRAWINGS

FIG 1: Flow diagram of the identification and validation of nucleic acid regulatory elements. A computational approach was used to identify the nucleic acid regulatory elements involving the following steps: (1) identification of tissue-specific genes that are highly expressed e.g. based on statistical analysis of micro-array expression data of normal human tissues; (2) extraction of the corresponding promoter sequences from publicly available databases; (3) identification of the regulatory modules and the transcription factor binding sites (TFBS) they contain (4) Next, the genomic context of the highly expressed genes was searched for evolutionary conserved clusters of TFBS (i.e. nucleic acid regulatory elements). The identified nucleic acid regulatory elements were *de novo* designed and validated *in vivo* by testing whether inclusion in a construct increases expression of a reporter gene.

FIG 2: Nucleotide sequences of the cardiac and skeletal muscle-specific regulatory elements CSk-SH1 (A, SEQ ID NO: 1), CSk-SH2 (B, SEQ ID NO: 2), CSk-SH3 (C, SEQ ID NO: 3), CSk-SH4 (D, SEQ ID NO: 4), CSk-SH5 (E, SEQ ID NO: 5), and CSk-SH6 (F, SEQ ID NO: 6).

FIG 3 shows a schematic representation of the AAV9sc-CSk-SH/Sk-SH-Des-Luc2 vectors disclosed herein. The expression cassette was packaged in a self-complementary (sc) adeno-associated virus, serotype 9 (AAV9). The cardiac and skeletal muscle-specific