

# miRNA Lentiviral Vector Integration and Gene Targeting Efficacy in Cardiac Progenitors

Mariana Loperfido<sup>1</sup>, Stefania Crippa<sup>1</sup> and Maurilio Sampaolesi<sup>1,2\*</sup>

<sup>1</sup>Translational Cardiology, Stem Cell Research Institute, KULeuven, Herestraat 49, 3000 Leuven, Belgium

<sup>2</sup>Human Anatomy, University of Pavia, Via Forlanini 8, 27100 Pavia, Italy

## Abstract

Stem and progenitor cardiac cells are challenging for possible cell therapy application. Several research laboratories are exploiting the feasibility of autologous cell therapy approach to get rid of immunosuppressive treatments responsible for undesirable side effects. Recently, we showed that cardiac progenitors isolated from *Sgcb* null mice, animal model of limb-girdle muscular dystrophy type 2E, undergo an aberrant differentiation *in vitro* and *in vivo* due to the dysregulation of miR669. This miRNA family is able to inhibit the skeletal myogenic program directly targeting MyoD 3' UTR. Using lentiviral technology we provided evidence that it is possible to rescue the dystrophic aberrant phenotype by miRNA669 overexpression without gene correction. However, how the viruses carrying the miRNAs were positioned in the genome upon transduction and how their localization site could influence the rescue potential was not analysed. Here we investigate the integration profile of lentiviral vector carrying the pre-miR669 in infected polyclonal and clonal populations derived from *Sgcb* cardiac progenitors. Our study reveals that the retroviral insertion sites (RIS) are largely restricted to coding genes (65%). Although with the limitation of our analysis, we found no hits for cancer-related genes and several sequenced RIS brought to light genes mainly involved in muscle function. Thus our data show that lentiviral vector insertional profile is cell-specific, however, the chromatin state of target cells positively influences the viral integrations.

**Keywords:** Muscular dystrophy; Cardiomyopathy; Cardiac progenitors; miRNAs; Lentiviral vectors; Integration profile analysis; Gene and cell therapy

## Introduction

Muscular dystrophies are heterogeneous genetic diseases caused by progressive degeneration of skeletal muscle tissue [1]. Mutations in genes encoding for skeletal muscle specific proteins lead to the lack of one of several proteins either at the plasma membrane or, less frequently, within internal membranes. This increases the probability of damage during contraction and eventually leads to fiber degeneration [2], although the molecular mechanisms are not yet understood in detail. Fiber degeneration is counterbalanced by the regeneration of new myocytes at the expense of resident myogenic. Skeletal muscle regeneration is mainly sustained by satellite cells [3,4], local myogenic progenitors localized underneath the basal lamina of muscle fibers. However, after repeated injuries, satellite cells become exhausted, losing their regenerative ability. Differently, cardiac muscle is less efficient to regenerate and tends to develop scar tissue after injuries [5]. In the last years several groups have reported the presence of local stem/progenitor cells able to differentiate into cardiac [6-10] and skeletal muscle lineages [11-17]. This is relevant for muscular dystrophy pathologies, since cardiac tissue is frequently affected in several forms of muscular dystrophy, including limb-girdle muscular dystrophies (LGMD), in which sarcoglycan protein mutations are responsible for the muscle degeneration.

LGMD type 2E is due to mutation in gene encoding for beta sarcoglycan protein and affected patients develop severe cardiomyopathy and mild muscle wasting. The group of Campbell has developed *Sgcb*-null mice that similarly to LGMD2E patients present both muscular dystrophy and severe cardiomyopathy [18].

*Sgcb*-null mice represent an excellent animal model for stem cell therapy strategies to promote regeneration of tissues characterized by

slow cellular turnover [7,10,19]. In *Sgcb*-null micemyogenic progenitors fail to repair damaged muscle fibers and large areas of necrosis were observed in skeletal muscles at all ages, more severe muscle pathology in comparison to *Sgca*-null mouse model [20]. In addition, the *Sgcb*-null heart shows the presence of degenerative foci already at 4 weeks leading to extensive alteration in 20 week old mice. Consequently at 30 week old all the ischemic lesions are replaced by diffuse areas of fibrosis.

In the last decade several evidences led to the identification of cardiaclocal stem cells able to migrate into heart ischemic regions and participate to cardiac regeneration [6,21]. Several groups identified cardiac stem/progenitor cells with similar characteristics, however, on the basis of differential marker expression and their multipotent characteristic were named differently [22-25]. Our group in collaboration with Cossu et al., recently showed that vessel associated stem cells termed mesoangioblasts (MABs), are present in skeletal muscle and heart, sharing pericyte markers, limited self-renewal, and undergo skeletal [11,12,26] and cardiac [8,9] myogenesis.

Although this topic remains quite controversial, the majority of researchers believe that the number of cardiac progenitors strongly increases in acute and chronic diseases. However, they are not able to

\*Corresponding author: Maurilio Sampaolesi, Ph.D, Stem Cell Research Institute, KULeuven, Herestraat 49, 3000 Leuven, Belgium, Tel: +32-(0)163-30295; Fax: +32-(0)163-30294; E-mail: [maurilio.sampaolesi@med.kuleuven.be](mailto:maurilio.sampaolesi@med.kuleuven.be)

Received May 08, 2012; Accepted June 27, 2012; Published June 29, 2012

**Citation:** Loperfido M, Crippa S, Sampaolesi M (2012) miRNA Lentiviral Vector Integration and Gene Targeting Efficacy in Cardiac Progenitors. J Stem Cell Res Ther S9:003. doi:10.4172/2157-7633.S9-003

**Copyright:** © 2012 Loperfido M, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

counteract extensive cardiac degeneration, likely because they may be responsible for repairing limited damaged areas or get exhausted in repeated attempts to regenerate the failing heart [27,28].

We recently isolated and characterized cardiac progenitors from *Sgcb*-null mice and found miR669 responsible for differentiation impairment of dystrophic cardiac progenitors [29]. However, *Sgcb*-null cardiac progenitors differentiated into skeletal muscle fibers both *in vitro* and *in vivo*. It is noteworthy that those dystrophic cardiac progenitors were not influenced by stem cell niche, and maintain their skeletal muscle commitment when transplanted into damaged skeletal muscles or infarcted hearts. Although cardiac differentiation in adult progenitor cells is still largely unknown, several recent papers have pointed out the role of microRNAs (miRNAs), small non-coding RNA, involved in pathophysiological aspects of myogenic progenitor cells and myocytes [30-32]. In our attempts to identify the molecular mechanisms responsible to this aberrant differentiation ability, we found that *Sgcb*-null cardiac progenitors do not express members of miR669 family. miR669 family encoded as a cluster in *Sfmbt2* gene, mainly involved in epigenetic silencing of myogenic genes [33,34]. We showed that miR669 directly inhibits the MyoD 3' untranslated region (UTR) and, consequently, skeletal muscle differentiation. To date, all the identified muscle miRNAs indirectly promote myogenesis, rather than acting directly on key regulatory factors for muscle differentiation. Gain and loss of function studies clearly show that miR669 act within a control system to switch skeletal/cardiac muscle fate. We then rescued the aberrant differentiation phenotype by targeting the dystrophic cardiac progenitor with lentiviral vectors carrying pre-miR669a. These data indicate that *ex vivo* gene therapy for cardiac disease can be supported by miRNA regulative elements to specifically drive stem cell differentiation. However, further experiments are necessary to understand the role of integration sites and the miRNA expression efficacy in miRNA-based gene therapy protocols.

The aim of this study was to develop a transduction protocol for *ex vivo* cell therapy using dystrophic cardiac progenitors and explore the integration profile of miRNA669 lentiviral vectors in clonal and polyclonal settings. Lentiviral vectors have attractive properties as gene-delivery vehicles. They allow permanent integration and stable expression of a therapeutic gene or miRNA into a target cells, and potentially give rise to significantly safer integration site profile. Results from this study provide important information for miRNA transduction protocols to be eventually adopted on human *ex vivo* cell therapy systems.

## Material and Methods

### Isolation of *Sgcb*-null cardiac progenitors and muscle differentiation

Cardiac progenitors were isolated from heart of 2 weeks *Sgcb*-null mice, as previously described [29]. Briefly, hearts were kept in DMEM without FCS with antibiotics and divided in aorta, ventricle, and atrium. Each piece was rinsed in PBS, sharply dissected in ~ 2 mm pieces with a scalpel, and transferred to a 3.5 cm 1% gelatin-coated Petri dish in presence of 20% FBS-DMEM plus 5 mM glutamine and antibiotics. Fragments were cultured for at least 8 days, until the outgrowth of fibroblast-like cells, small round, and refractile cells appeared. This cell population was detached by trypsin and collected by gently pipetting, counted and cloned by limited dilution on 1% gelatin-coated p96 well plates. Different clones were selected by phase contrast morphology and

expanded. Skeletal muscle differentiation was induced spontaneously in differentiation medium (DMEM 2% HS).

### Immunofluorescence analysis and microscope image acquisition

*Sgcb*-null cardiac progenitors were cultivated as previously reported [29] or exposed to differentiation medium (DMEM 2% HS) and fixed 3 and 7 days with 4% paraformaldehyde for 10 min at 4°C. Then, they were stained with antibody against MyHC (MF20, DSHB) over night at 4°C and incubated with Alexa Fluor® conjugated secondary antibody (Invitrogen) for 1 hour at room temperature in the dark. DAPI (Molecular Probes, Invitrogen) was used for nuclear staining.

Images were taken with a fluorescent inverted microscope (Eclipse Ti-U; Nikon) equipped with a QICAM Fast 1354 camera and using Image-Pro Plus software. The number of nuclei stained with DAPI inside myosin-positive cells was counted. The fusion index (FI) was calculated against the total number of nuclei.

### Lentiviral vector generation and *Sgcb*-null cardiac progenitor transduction

The vector used in this study is a 3rd generation lentiviral vector, previously described as LV-CMV-eGFP-miR669a2x (Supplementary Figure S1). *Sgcb*-null clone (H4 Ven KO) was transduced with LV-CMV-eGFP-miR669a2x in multiple infections (LV1a, LV1b, LV1c) or sorted for GFP expression by FACS, in order to derive monoclonal infected clones (A8, C12, D1, F5 and G8). All the infections were performed following the same protocol and with a multiplicity of infection (MOI) of 50.

### RNA extraction and qRT-PCR analysis for MyoD expression

Total RNA was extracted from proliferating and differentiated wt and *Sgcb*-null cardiac clones using the RNeasy Micro Kit (Qiagen). Incubation with DNase I was performed on column and cDNA was synthesized from 1µg of total RNA using SuperScript® III Reverse Transcriptase (Invitrogen). qRT-PCR was performed using Platinum® SYBR® Green qPCRSuperMix-UDG (Life Technologies) and primers for MyoD (*Forward* 5'-CCACTCCGGGACATAGACTTGAC-3'; *Reverse* 5'-TCTGGTGAGTCGAAACACGGATC-3').

### MiRNA extraction and TaqMan assay analysis for miR669a

Total miRNA population was extracted from proliferating and differentiated wt and *Sgcb*-null cardiac clones using the mirVana™ miRNA Isolation Kit (Ambion). Quantification of miRNA was performed in two-step RT-PCR according to the TaqMan MicroRNA Assays protocol. Briefly, cDNA is reverse transcribed from 5ng of miRNA using specific provided miRNA primers from the TaqMan® MicroRNA Reverse Transcription Kit. PCR products were amplified from cDNA samples using the Taq Man® Universal PCR Master Mix.

Primers for miR669a were purchased from Applied Biosystems (miR-669a, # 001683).

### DNA extraction and integration profile analyses by LM-PCR

Genomic DNA was extracted with QIAamp DNA Blood Mini-Kit (Qiagen) from infected *Sgcb*-null cardiac clones and used for LM-PCR analysis (Linker-Mediated Polymerase Chain Reaction) in order to identify the genomic loci where the vector landed.

Two different LM-PCR protocols specific for LV-CMV-eGFP-miR669a2x vector were set up (see Supplementary Figure S2) i) by a single digestion of the genomic DNA with *MseI* enzyme; ii) by a double

digestion of the genomic DNA with *NcoI* enzyme (present inside the vector), and with *MseI* enzyme. This strategy was designed to enrich the population of fragments containing the 3'LTR in the first exponential PCR of the protocol.

Digestion products were ligated with an *MseI*-complementary linker cassette (Quick Ligation™ Kit, New England Biolabs) and underwent to two sequential rounds of exponential PCR steps performed by using primers specific for 3'LTR and linker cassette.

AmpliTaq Gold® DNA Polymerase (Applied Biosystems) and Taq Extender PCR Additive (Agilent Technologies) were used for the PCR steps.

LM-PCR products were cloned using TOPO TA Cloning Kit (Invitrogen) and sequenced. Results with the correct LTR and linker primer sequences were selected for analysis, then mapped in the mouse genome using Ensemble database (NCBI m37, www.ensembl.org) in order to identify unique Retroviral Insertion Sites (RIS).

For the integration profile analysis was used Quick Map program (www.gtsg.org), whereas FASTA sequences were automatically aligned to the mouse genome version 37 (ENS52).

#### Functions of genes hit by LV-CMV-eGFP-miR669a2x vector

To analyse the gene functions directly hit or the single gene closest to each insertion site, the hit gene list for all insertion sites from infected cardiac progenitors was uploaded on the Ingenuity Pathways Analysis software (IPA, www.ingenuity.com).

The contribution of hit genes with respect to the functional categories listed in the IPA database were analysed. Molecular and cellular functions, pathways, physiological system development and functions, disease and disorder have been considered.

## Results

### MyoD expression and lentiviral integration of miR669a-transduced *Sgcb*-null cardiac progenitors subjected up to 4 rounds of infections

In order to evaluate the integration sites we followed two alternative routes based on the analysis of vector integrations in polyclonal and monoclonal infected cardiac progenitors as reported in Figure 1A. We analysed the vector integrations in polyclonal *Sgcb*-null cells subjected to multiple injections (red dashed line) and in monoclonal infected clones (green dashed line), derived from GFP+ *Sgcb*-null cardiac progenitors. TaqMan assay analysis revealed that miR669a expression (Figure 1B, left panel) is proportional to the round of infections and reach a plateau after the third round of infection. Consequently qPCR analysis for MyoD expression (Figure 1B, right panel) in *Sgcb*-null clones decrease proportionally to the rounds of infection, suggesting a dose-dependent effect of miR669a. The wt cardiac progenitors, used as control, showed high expression of miR669a and the absence of MyoD. In order to associate the therapeutic effect of miR669a and the reduction of skeletal muscle differentiation we performed an immunofluorescence analysis after 3 or 5 days of serum starvation of transduced clones to induce myotube formation. The strongest inhibitory effect on myotube formations have been observed on LV1c clone, subjected to four rounds of infections (Figure 1C right last panels).

### MyoD expression and myogenic differentiation in miR669a-transduced *Sgcb*-null cardiac progenitor clones

To evaluate the effect of single infection GFP+ cells were sorted by Flow Cytometry and MyoD expression in comparison the miR669a presence was determined. Although we were able to increase miR669a expression faintly (Figure 2A, left panel), this was enough to partially reduce MyoD expression (Figure 2A, right panel). The inhibitory effect was confirmed by immunofluorescence analysis showing that transduced *Sgcb*-null cardiac progenitors slowed down their myogenic differentiation compare to untransduced cells (Figure 2B).

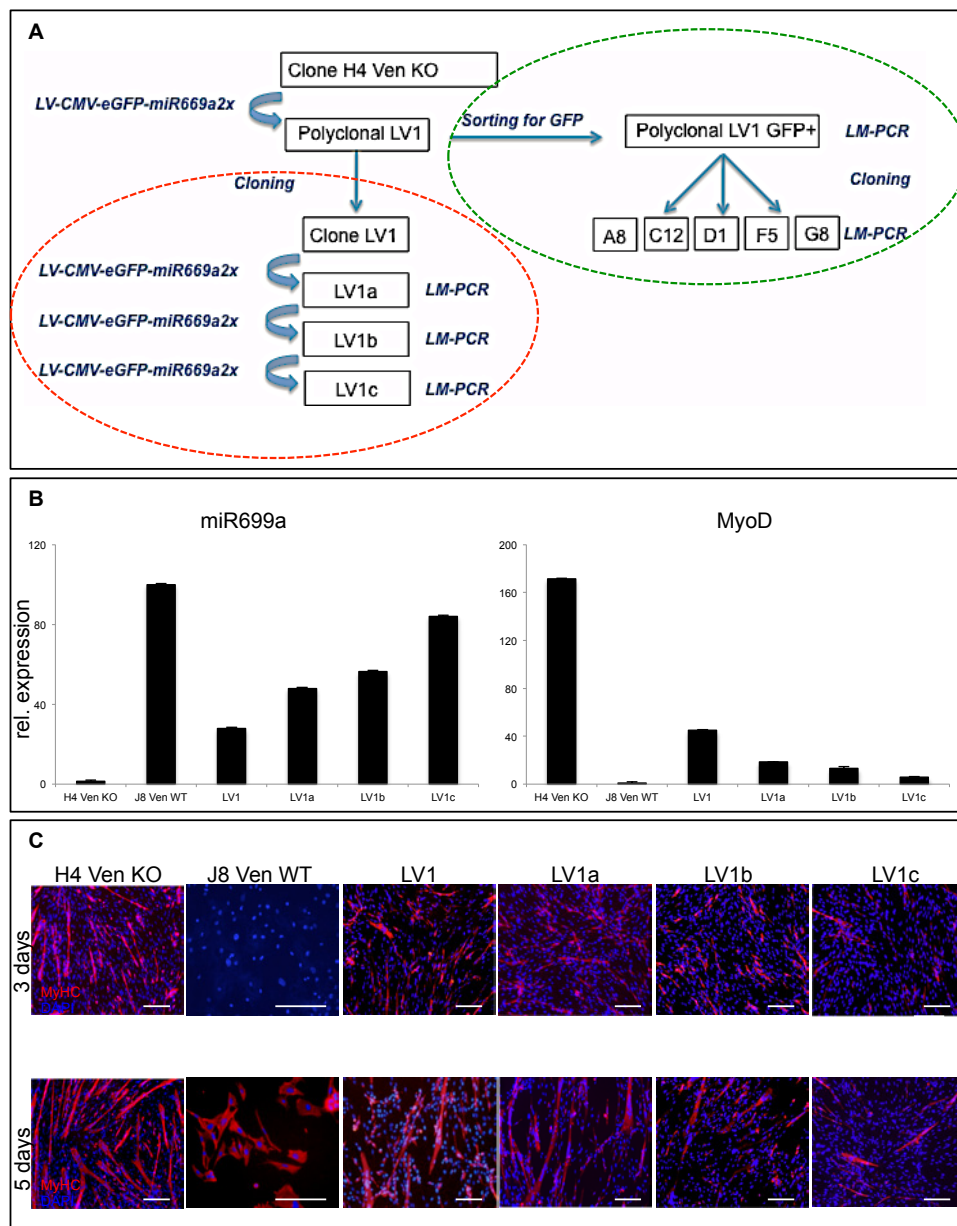
We then cloned GFP+ cells in order to investigate the biological inhibitory effect of miRNA based therapeutic protocol in relation to the lentiviral vector integration sites. A8, F5 and G8 clones were isolated from miR669a-transduced *Sgcb*-null cardiac progenitors and processed for further analysis. In each clone the expression of miR669a caused a proportionally reduction in MyoD expression (Figure 2A) that affect myotube formation (Figure 2B), when the transduced dystrophic progenitors were cultivated under differentiation conditions. As expected clones expressing higher level of miR669a (Figure 2B, A8 and G8 clones) resulted in lower fusion index compared to the uninfected or lower miR669a transduced cardiac progenitors (Figure 2B, F5 clone).

### Analysis of vector integration sites in *Sgcb*-null cardiac progenitors

In order to study the integration site profile of the lentiviral vector LV-CMV-eGFP-miR669a2x in *Sgcb*-null cardiac progenitor cells, we collected insertion sites by shotgun cloning-generated insertion libraries of LM-PCR products. We used two different protocols of digestion of the genomic DNA from infected clones, before the ligation with the linker cassette: i) a single digestion with *MseI*; ii) a double digestion with *NcoI* and *MseI* enzymes. The purpose was to determine if the extra-digestion with an enzyme that cuts into the vector (*NcoI*) might increase the sensibility of the LM-PCR and then the number of RIS to detect. The distribution of the vector was analysed following two different approaches (as shown in Figure 1A): i) on LV1 cardiac progenitor polyclonal population after multiple infections (LV1a, LV1b, LV1c clones); ii) on LV1 polyclonal population sorted for GFP+ and derived clones (A8, C12, D1, F5 and G8), obtained by limiting dilutions.

The insertion site distribution of LV1 polyclonal population and after multiple infections had shown the typical polyclonal integration pattern of the lentiviral vector, with significant preference for transcription units (Table 1). Within the genes hit by the vector, particularly interesting were *ATPase* gene (*Atp2b1*) and *Protein Phosphatase 1* gene (*Ppp1r12a*) (Supplementary Figure S3) since they are involved in muscle function. *ATPase* gene encodes the plasma membrane calcium ATPase isoform 1, an enzyme that plays a critical role in intracellular calcium homeostasis and when altered it is responsible for neurological and muscular disorders. The protein encoded by *Protein Phosphatase 1* gene is one of the myosin-binding subunits of myosin phosphatase, which regulates the extent of phosphorylation of myosin light chain (MLC). It's involved in different pathways, such as the cardiac  $\beta$ -adrenergic signalling.

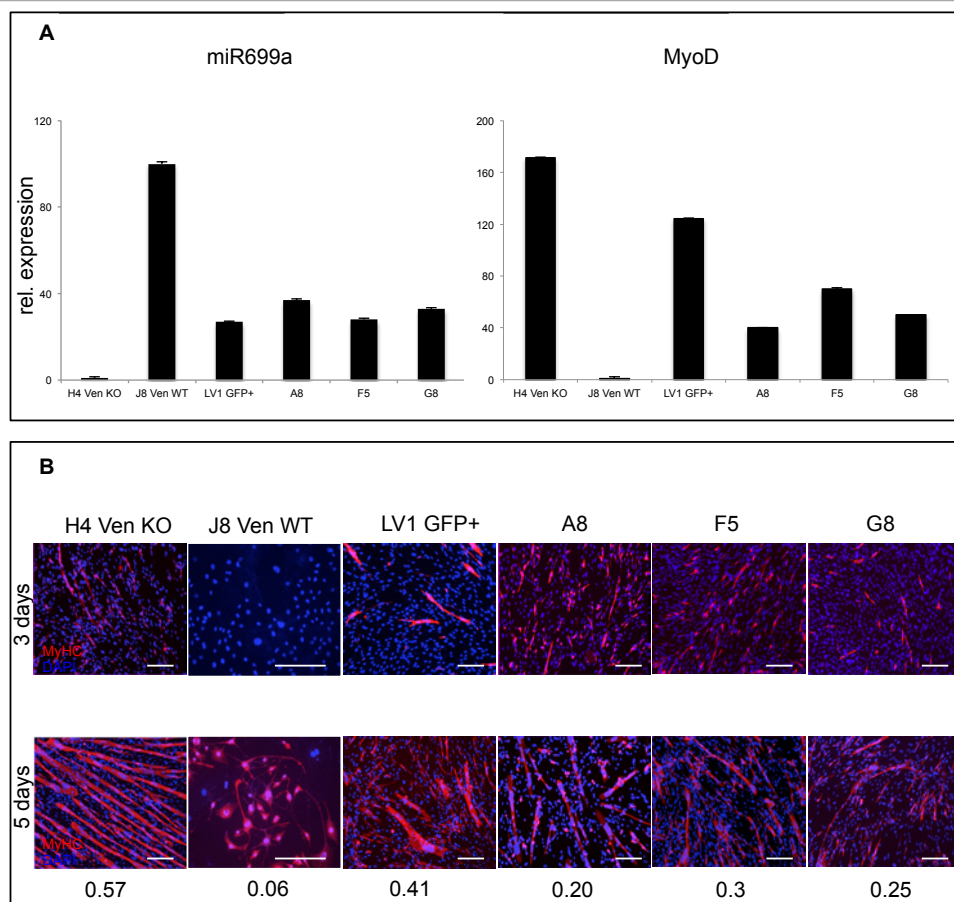
Also the analysis of RIS in GFP+ sorted *Sgcb*-null cardiac progenitors revealed a polyclonal repertoire: unique insertions were retrieved in different genomic loci. While in GFP+ clones derived by limiting dilutions from the same population (A8, C12, D1, F5 and G8) only one or two unique integrations per clone have been found (Table 2).



**Figure 1: Dose-response studies on MyoD expression and myogenic differentiation, in miR669a-transfected *Sgcb*-null cardiac progenitors. (A)** Flowchart of experimental plan. We adopted two different approaches to evaluate the integration site profile of LV-CMV-eGFP-miR669a2x vector in infected cardiac progenitors by LM-PCR. 1) Analysis of vector integrations in polyclonal *Sgcb*-null cells upon multiple infections (red dashed circle). 2) Analysis of vector integrations in monoclonal infected clones (A8, C12, D1, F5 and G8), derived from GFP+ *Sgcb*-null cardiac progenitors previously transduced with the same viral vector (green dashed circle). Note that the protocol of infection and MOI have been standardized for all experiments. **(B)** TaqMan assay analysis for miR669a expression (left panel) and qPCR analysis for MyoD expression (right panel) in wt and *Sgcb*-null clone after a single infection (LV1) with LV-CMV-eGFP-miR669a2x or multiple infections (LV1a, after double infection; LV1b, after three infections and LV1c after four infections) using the same protocol. MyoD expression decreases in a dose-response manner according to miR669a increment. **(C)** Immunofluorescence analysis of differentiating *Sgcb*-null (H4 Ven KO, LV1, LV1a, LV1b and LV1c) and wt (J8 Ven WT) cardiac clones, at 3 and 5 d after serum starvation (2% HS). Skeletal muscle differentiation was impaired by miR669a transduction in dose dependent manner. Note that only few myocytes are present in multiple infected clones compared to controls, after 3 and 5 days from serum starvation; scale bar = 100 μm.

Of particular note is the identification of the vector insertion site in *Utrn* gene, encoding the utrophin protein (Supplementary Figure S4). In mice, this protein is located at the neuromuscular synapse and myotendinous junctions, and plays a fundamental role for normal membrane maintenance. Another interesting gene where the lentiviral vector landed is *Tardbp*, whose mutations are associated with neurodegenerative disorders including frontotemporal lobar degeneration and amyotrophic lateral sclerosis (ALS).

The analysis on the total number of RIS retrieved from infected *Sgcb*-null cardiac progenitor cells shows a preference of the lentiviral vector to integrate inside gene (65%) than outside gene (35%) (Table 3). Most of the insertions were obtained by using the single digestion protocol (MseI enzyme), even though all the integrations identified in exons have been detected by the double digestion protocol (NcoI+MseI enzymes), likely due to a higher sensibility of the technique. When we considered the region around the transcription start site (TSS) of the



**Figure 2:** MyoD expression and myogenic differentiation in miR669a-transduced Sgcb-null cardiac progenitors after single infection and FACS. The slightly increase in miR669a expression in transduced cells(A, left panel) was able to reduce MyoD expression(A, right panel), and consequently Sgcb-null cardiac progenitors slowed down their myogenic differentiation (B), as showed by a reduction of MyHC positive myotubes. A8, F5, G8 clones were isolated from miR669a-transduced Sgcb-null cardiac progenitors, previously sorted for GFP expression (LV1 GFP+). The fusion index (FI) analysis confirmed the myogenic impairments in transduced clones. Indeed, clones (A8 and G8) that expressed higher level of miR669a resulted in lower number of myotubes and FI (as reported below the panels) compared to the clone (F5) with low miR669a expression. Scale bar = 100 μm.AL

NcoI+MseI			MseI		
Chr	Gene	Position	Chr	Gene	Position
<b>4</b>	<b>Gm12649 (LincRNA)</b>	-	<b>3</b>	<b>2700069118Rik (LincRNA)</b>	-
10	Atp2b1	Exon 15	5	Kntc1	Intron 24
18	Rbm27	Intron 11	6	Glcci1	Intron 2
			10	Ppp1r12a	Intron 4
			13	Rhobtb3	Intron 12
			16	2310008H04Rik	Intron 11

**Table 1: Vector integration profile in LV1 clone upon multiple infections.** The analysis was performed after genomic DNA digestions with NcoI+MseI or MseI restriction enzymes. The integration profile of LV1 has been acquired by LM-PCR and depicted a typical polyclonal integration pattern, as showed by multiple integration sites in open chromatin regions. Insertion sites detected outside genes are shown in bold, and the closest gene identified is reported in the table.

genes hit by the vector, most of integrations were distributed in a 50-100 kb window downstream the TSS, confirming the tendency of the lentiviral vector to integrate in intronic/exonic regions, rather than closer to the proximity of the TSS (Figure 3A). Interesting, none of the RIS was identified in “Cancer” genes, obtained from the list of the RTCGD [35].

To analyse the insertion distribution on chromosomes, we calculated the frequency of the vector insertions versus a reference

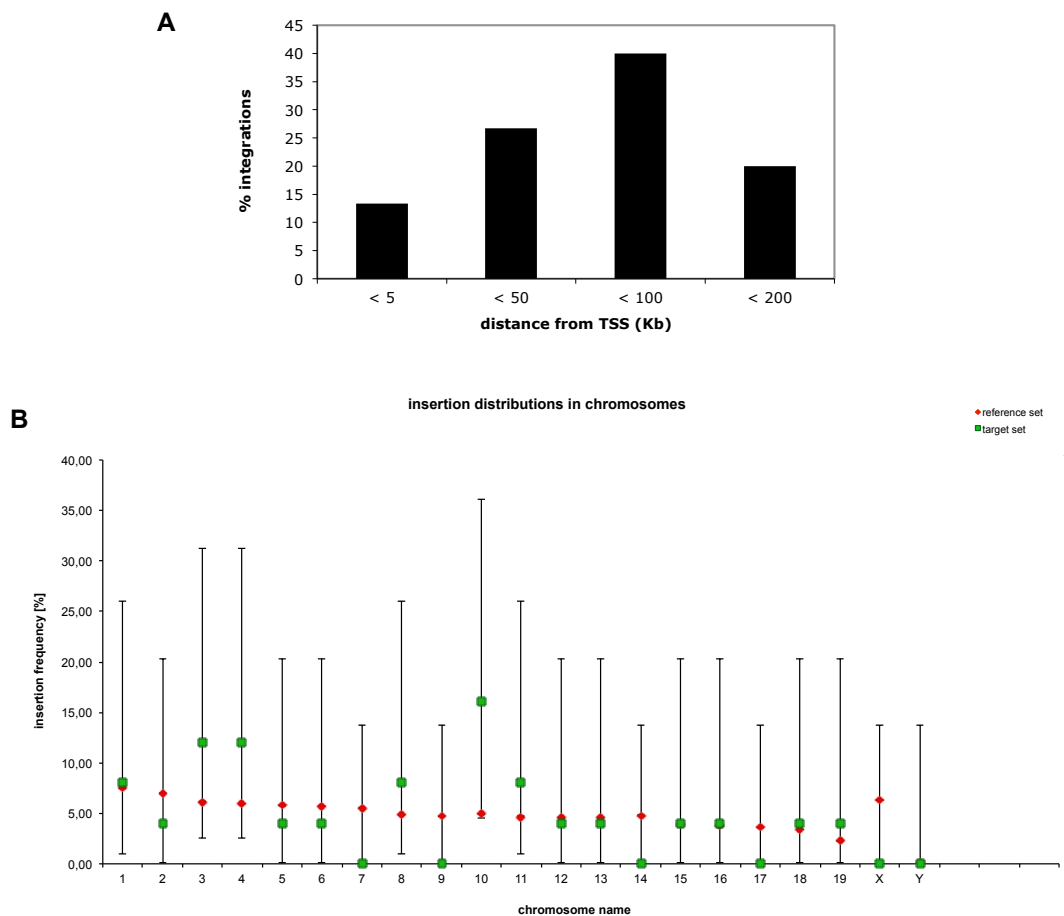
set of 1.000.000 random insertions detected in the mouse genome, using GTSG-QuickMap software. As shown in Figure 3B, both sets of insertions (target versus reference) match quite completely and there is not a particular preference of the lentiviral vector to integrate on a specific chromosome.

#### Functional analysis of genes hit by the vector

To analyse the function of genomic loci involved by RIS, genes

		LV1 GFP+			
<i>NcoI</i> + <i>MseI</i>				<i>MseI</i>	
Chr	Gene	Position	Chr	Gene	Position
<b>1</b>	<b>Fcgr4</b>	-	<b>4</b>	<b>Gm11925 (LincRNA)</b>	-
<b>8</b>	<b>Foxl1</b>	-	<b>10</b>	<b>Trhde</b>	-
1	Usp37	Intron 12	2	B230339M05Rik	Intron 22
4	Tardbp	Intron 3	12	Prkar2b	Intron 2
10	Utrn	Intron 23			
19	Nt5c2	Intron 3			
A8					
-	-	-	11	Acox1	Intron 3
C12					
10	Utrn	Intron 23	-	-	-
D1					
-	-	-	10	Ppp1r12a	Intron 4
F5					
3	Tbl1xr1	Intron 10	3	Tbl1xr1	Intron 10
G8					
-	-	-	<b>8</b>	<b>Efnb2</b>	-
-	-	-	<b>15</b>	<b>Angpt1</b>	-

**Table 2: Identification of vector insertions in transduced GFP+ Sgcb-null cardiac progenitors, and in derived clones.** The analysis was performed on the genomic DNA by LM-PCR, after digestions with *NcoI*+*MseI* or *MseI* restriction enzymes. The overall distribution of the vector integration sites in the GFP+ sorted population confirmed the polyclonal profile of lentiviral integrations, while in the GFP+ clones (A8, C12, D1, F5 and G8), obtained by limiting dilutions, only one or two unique insertions were retrieved. Insertion sites detected outside genes are shown in bold, and the closest gene identified is also reported in bold in the table.



**Figure 3: (A)** Insertion sites distribution was analyzed in four distance categories inside genes. The histogram confirms the classical profile of the lentiviral vector to integrate inside genes, rather than in the close proximity of the TSS. **(B)** Frequency of the vector insertions on chromosomes was analyzed by GTSG- QuickMap software. Integration distribution of the LV-CMV-eGFP-miR669a2x vector is represented in green (target set) and 1.000.000 random insertions (reference set) detected in mouse genome (Random Set Mus Musculus ENS52) are shown in red.

directly hit by the vector or the single gene closest to each insertion site were uploaded on the ingenuity pathways analysis (IPA) to look for related biological functions.

We considered the contribution of hit genes with respect to 4 different functional categories listed in the IPA database (Figure 4): molecular and cellular functions, pathways, physiological system development and functions, diseases and disorders.

The cardiac progenitor dataset showed that genes involved by RIS are related to some functions as cellular assembly and organization, cell-to-cell signalling and interaction, and molecular transport. Considering the nature of the cell type infected with the lentiviral vector, pathways in which hit genes might play an important role are cardiac  $\beta$ -adrenergic and calcium signalling, agrin interactions at neuromuscular junctions, nitric oxide signalling in the cardiovascular system.

With regard to the physiological system and disease/disorder categories, hit genes are involved in nervous system architecture and muscular development and function.

## Discussion

Unlike some organs, the heart has limited capacity for repair after injury and until the last decade it has been considered an organ composed by terminally differentiated somatic cells, explaining at least in part its poor regenerative capacity. Several researchers, however, may be influenced by the extraordinary regenerative ability of zebrafish cardiac tissue [36], identified cardiac progenitors able to participate to cardiomyogenesis [6,9,37]. The fact that ischemic events in mammalian heart lead to the fibrosis scar formation highlights the poor regeneration capacity of those cardiac committed adult stem cells. Nevertheless, understanding their differentiation molecular circuit could help in the treatment of acute and chronic heart diseases. Recently, we identified a miRNA dysfunction in cardiac progenitors isolated from *Sgcb* null mice, animal model for limb-girdle muscular dystrophy type 2E. We proved that correcting the miRNA669 expression in adult cardiac stem cells was sufficient to rescue the cardiomyogenic ability of dystrophic cells. At least two member of miR669 family are capable to directly decrease the expression of MyoD in dystrophic cardiac progenitors, in which high level of intracellular  $Ca^{2+}$  [38] activates calpain proteases responsible for YinYang1 (*Yy1*) proteolytic degradation [39,40]. *Yy1* is a transcription factor responsible for Scm-like with four *mbt domains* 2 (*Sfmbt2*) expression, functionally involved in epigenetic silencing of skeletal muscle genes [33,34]. Thus, the complexity of myogenic process require a fine-tune regulation of miR669 family expression in consideration of potential ex-vivo cell therapy protocols. Lentiviral vectors have been largely used in gene therapy applications [41-45] and in the last decade many studies focused on the identification of retroviral integration sites (RIS) to monitor vector-host interaction in long-term follow-up of patients [46]. However, integration site analysis has now considered critical in gene therapy protocols considering the adverse effects occurred in SCID clinical trials during the early 2000s.

In this study, we investigated the insertion site distribution of the lentiviral vector LV-CMV-eGFP-miR669a2x on polyclonal and monoclonal *Sgcb*-null progenitor cells. Transduced cells were able to negatively regulate MyoD expression, as expected, and shown a typical lentiviral integration profile in both polyclonal and monoclonal populations. Currently, it is well accepted that lentiviral vectors show a non-random integration profile possibly influencing the fate of transduced cells [47-49]. On the other hand, insertion site selection

during in vitro transduction could be driven by cellular mechanisms such as tethering of transcription factors (TF) to specific genomic regions according to the presence of TF binding sites [50] and seems dependent on cellular determinants as well as on vector design [51]. According to the literature we found that the lentiviral insertion in polyclonal progenitor population before and after multiple infections displayed the typical polyclonal integration pattern, with significant preference for introns of coding genes, often associated with myogenic development and function.

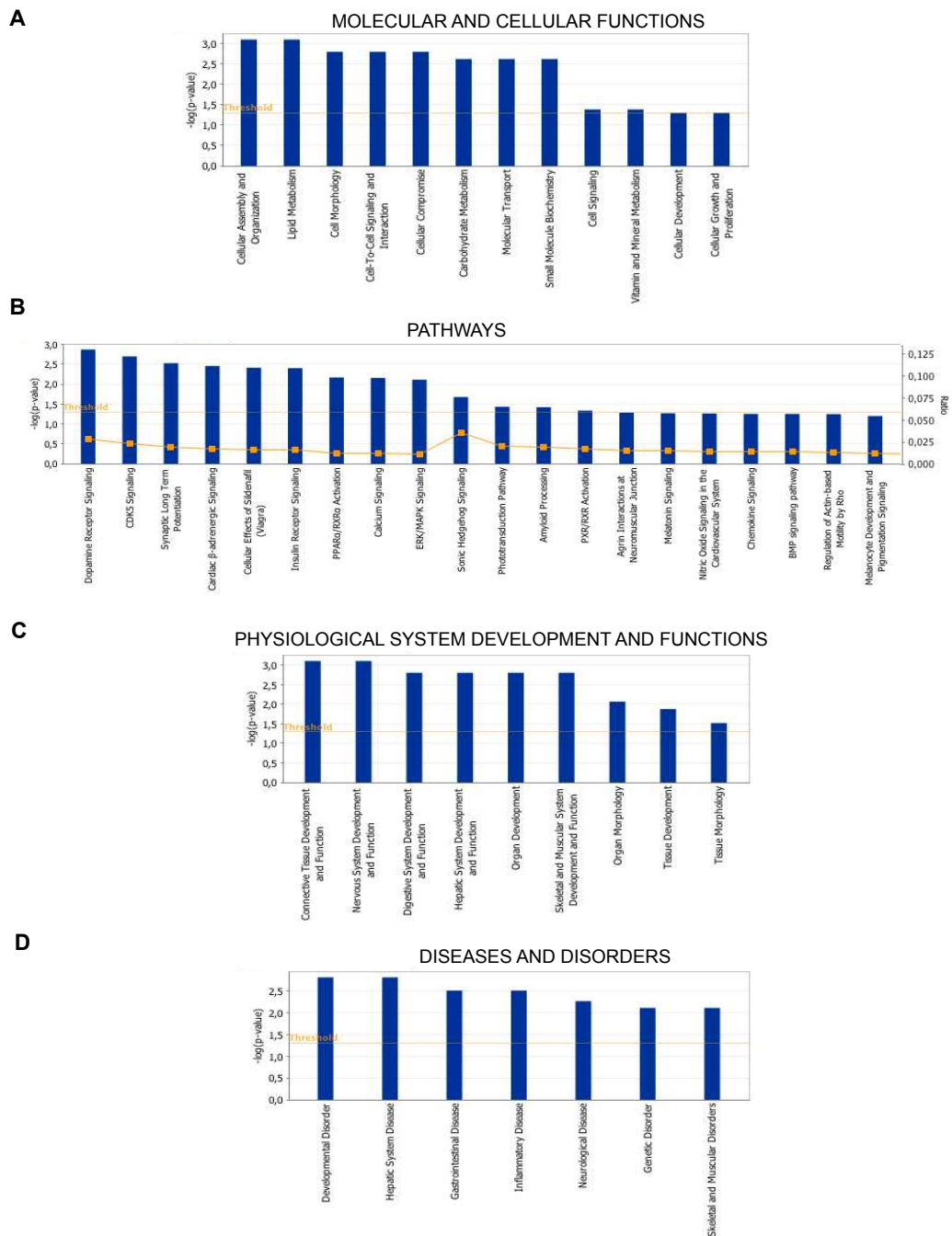
Together our data unequivocally show that a higher number of RIS has been identified inside coding regions compare to the number of integrations landed into intergenic regions. It is noteworthy that the integration sites are in gene directly correlated with muscle function, such as *utrophin*, *ATPase* and *Protein Phosphatase 1* gene. The latter is a functional subunit of key enzyme namely myosin phosphatase, which regulates the interaction of myosin and actin, necessary to generate the force needed for muscle contraction. Similarly, ATPase activity is positively correlated with muscle contraction while utrophin, a cytoskeleton protein mainly localized in the neuromuscular junction of skeletal muscles and in the Purkinje fibers, transverse tubules, and intercalated disks of heart. Utrophin upregulation observed in dystrophic muscles of mdx mice, animal model for Duchenne muscular dystrophy, can be obtained with some drugs, which are considered potential treatments for DMD patients.

A8 transduced dystrophic clone showed the lowest MyoD expression and consequently the lowest fusion index. The unique integration site found in A8 clone is in the intron 3 of *Acox1* gene encoding peroxisomal acyl-CoA oxidase involved in the beta-oxidation of fatty acids. It is interesting that this unique insertion in the *Acox1* gene intron was able to guarantee a strong MyoD inhibition, much more efficiently than the double intergenically insertions observed in the G8 clone (Table 2).

Although cell-specific insertion preferences are due to the expression profile and epigenetic status of the cardiac dystrophic progenitor cells at the time of infection we observed a broad variability of hit genes, as shown in the gene categories listed in Figure 4. In addition, due the limitation of our analysis we cannot exclude the presence of other insertion sites. We have extensively used two DNA digestion protocols, however, only in one case (clone F5, Table 2) we were able to detect the same unique insertion site with both protocols, suggesting the two methods are necessary to identify a larger number of different insertions.

*Sgcb*-null cardiac progenitors aberrantly differentiate in skeletal myocytes because they lack key microRNAs regulating MyoD expression. This can explain the inefficient cardiac regeneration in *Sgcb*-null mice and maybe the unusual cardiac MyoD expression in human oncocyctic myocarditis [52]. Several small non coding RNAs are dysregulated in cardiac diseases, as shown in a recent study in which 43 miRNAs were found differentially expressed in cardiac samples from patients affected by ischemic cardiomyopathy, dilated cardiomyopathy or aortic stenosis [53].

In addition, the different heart pathologies are associated with a distinct expression repertoire of miRNAs, suggesting that miRNA expression profiles correctly matched samples by their clinical diagnosis. Further investigations are necessary to understand if miRNA gain and loss of function could interfere with the specific cardiac degeneration. In this respect, our cell model system might prove useful



**Figure 4:** Biological functions of genes hit by lentiviral vector integrations. The histograms show the contribution of hit genes with respect to four different functional categories listed in the IPA database: Molecular and cellular functions (A), Pathways (B), Physiological system development and functions (C), Diseases and disorders (D).

integrations	inside gene	outside gene
(NcoI+MseI)	26% (33% intron, 7% exon)	13%
(MseI)	39% (60% intron, 0% exon)	22%
total	65% (93% intron, 7% exon)	35%

**Table 3: Genomic distribution of lentiviral integration sites in infected Sgcb-null cardiac progenitors.** Percentages of vector integrations obtained by two different digestion protocols (NcoI+MseI and MseI) in infected cardiac progenitors landing inside or outside genes are shown. Note that LV-CMV-eGFP-miR669a2x vector preferentially integrated inside genes as expected.



to test different miRNAs involved in cardiac function. Thus, lentiviral vector technology, supported by a safe integration profile, provides a useful tool to manipulate miRNAs, which in turn affect specific biological processes and maybe directly or indirectly could improve cardiac function. Indeed, a specific set of miRNAs (miR-1, mir 133, mir 208, and mir 499) was able to induce the expression of cardiac markers in fibroblasts that *in vitro* efficiently showed the presence of L-type channel, spontaneous calcium oscillations, and contractility [54]. The Authors reported similar results when lentiviral microRNAs were injected directly into the infarcted murine heart at the time of injury. It could be interesting to evaluate if the polyclonal profile of lentiviral integrations in the infected fibroblasts contain genes involved in cardiomyogenesis, and thus their positive alteration could facilitate the fibroblast transdifferentiation.

Although the expression of microRNAs can be finely tuned through different approaches, we provide evidence that lentiviral vectors could nicely rescue the miR669 expression in *Scgb* null cardiac progenitors. We also give novel information on how the general chromatin state of myogenic progenitors drives lentiviral vector integrations, since all the hit genes are directly related to the cellular biochemistry of skeletal muscle function. Finally, with the limitation of our analysis, we found that none of the RIS was identified in cancer-related genes, although it could be considered an arbitrary group, since most of stemness markers belong to this gene category.

#### Acknowledgment

This work was supported in part by: FWO- Odysseus Program n. G.0907.08; Research Council of the University of Leuven n. OT/09/053; CARE-MI n.242038 FP7-EC grant; the Italian Ministry of University and Scientific Research grant n. 2005067555\_003, PRIN 2006-08), CARIPO 2005-2008 to MS. We are grateful to Catherine Verfaillie for continuous support; Danny Huylebroeck, and An Zwijsen for helpful discussion; ML is and SC was supported by FWO. We thank K.P. Campbell (Howard Hughes Medical Institute, University of Iowa) for providing *Sgcb-null* mice, Christina Vochten and Luigi Vercesi for the professional secretarial service and Paolo Luban for a kind donation.

#### References

- Emery AE (2002) The muscular dystrophies. *Lancet* 359: 687-695.
- Blake DJ, Weir A, Newey SE, Davies KE (2002) Function and genetics of dystrophin and dystrophin-related proteins in muscle. *Physiol Rev* 82: 291-329.
- Mauro A (1961) Satellite cell of skeletal muscle fibers. *J Biophys Biochem Cytol* 9: 493-495.
- Quattrocelli M, Cassano M, Crippa S, Perini I, Sampaolesi M (2010) Cell therapy strategies and improvements for muscular dystrophy. *Cell Death Differ* 17: 1222-1229.
- Crippa S, Cassano M, Sampaolesi M (2012) Role of miRNAs in Muscle Stem Cell Biology: Proliferation, Differentiation and Death. *Curr Pharm Des* 18: 1718-1729.
- Quaini F, Urbanek K, Beltrami AP, Finato N, Beltrami CA, et al. (2002) Chimerism of the transplanted heart. *N Engl J Med* 346: 5-15.
- Anversa P, Leri A, Kajstura J (2006) Cardiac regeneration. *J Am Coll Cardiol* 47: 1769-1776.
- Galvez BG, Sampaolesi M, Barbuti A, Crespi A, Covarello D, et al. (2008) Cardiac mesoangioblasts are committed, self-renewable progenitors, associated with small vessels of juvenile mouse ventricle. *Cell Death Differ* 15: 1417-1428.
- Galvez BG, Covarello D, Tolorenzi R, Brunelli S, Dellavalle A, et al. (2009) Human cardiac mesoangioblasts isolated from hypertrophic cardiomyopathies are greatly reduced in proliferation and differentiation potency. *Cardiovasc Res* 83: 707-716.
- Yi BA, Wernet O, Chien KR (2010) Pregenerative medicine: developmental paradigms in the biology of cardiovascular regeneration. *J Clin Invest* 120: 20-28.

- Sampaolesi M, Torrente Y, Innocenzi A, Tonlorenzi R, D'Antona G, et al. (2003) Cell therapy of alpha-sarcoglycan null dystrophic mice through intra-arterial delivery of mesoangioblasts. *Science* 301: 487-492.
- Sampaolesi M, Blot S, D'Antona G, Granger N, Tonlorenzi R, et al. (2006) Mesoangioblast stem cells ameliorate muscle function in dystrophic dogs. *Nature* 444: 574-579.
- Montarras D, Morgan J, Collins C, Relaix F, Zaffran S, et al. (2005) Direct isolation of satellite cells for skeletal muscle regeneration. *Science* 309: 2064-2067.
- Zheng B, Cao B, Crisan M, Sun B, Li G, et al. (2007) Prospective identification of myogenic endothelial cells in human skeletal muscle. *Nat Biotechnol* 25: 1025-1034.
- Cerletti M, Jurga S, Witczak CA, Hirshman MF, Shadrach JL, et al. (2008) Highly efficient, functional engraftment of skeletal muscle stem cells in dystrophic muscles. *Cell* 134: 37-47.
- Crisan M, Yap S, Casteilla L, Chen CW, Corselli M, et al. (2008) A perivascular origin for mesenchymal stem cells in multiple human organs. *Cell Stem Cell* 3: 301-313.
- Quattrocelli M, Cassano M, Crippa S, Perini I, Sampaolesi M (2010) Cell therapy strategies and improvements for muscular dystrophy. *Cell Death Differ* 17: 1222-1229.
- Durbeej M, Cohn RD, Hrstka RF, Moore SA, Allamand V, et al. (2000) Disruption of the beta-sarcoglycan gene reveals pathogenetic complexity of limb-girdle muscular dystrophy type 2E. *Mol Cell* 5: 141-151.
- Nadal-Ginard B, Anversa P, Kajstura J, Leri A (2005) Cardiac stem cells and myocardial regeneration. *Novartis Found Symp* 265: 142-154.
- Duclos F, Straub V, Moore SA, Venzke DP, Hrstka RF, et al. (1998) Progressive muscular dystrophy in alpha-sarcoglycan-deficient mice. *J Cell Biol* 142: 1461-1471.
- Beltrami AP, Barlucchi L, Torella D, Baker M, Limana F, et al. (2003) Adult cardiac stem cells are multipotent and support myocardial regeneration. *Cell* 114: 763-776.
- Oh H, Bradfute SB, Gallardo TD, Nakamura T, Gaussin V, et al. (2003) Cardiac progenitor cells from adult myocardium: homing, differentiation, and fusion after infarction. *Proc Natl Acad Sci USA* 100: 12313-12318.
- Oh H, Chi X, Bradfute SB, Mishina Y, Pocius J, et al. (2004) Cardiac muscle plasticity in adult and embryo by heart-derived progenitor cells. *Ann NY Acad Sci* 1015: 182-189.
- Beltrami AP, Cesselli D, Bergamin N, Marcon P, Rigo S, et al. (2005) Investigation on possible cell sources to be utilized for cardiac cell therapy. *Pathologica* 97: 185.
- Laugwitz KL, Moretti A, Lam J, Gruber P, Chen Y, et al. (2005) Postnatal isl1+ cardioblasts enter fully differentiated cardiomyocyte lineages. *Nature* 433: 647-653.
- Dellavalle A, Sampaolesi M, Tonlorenzi R, Tagliafico E, Sacchetti B, et al. (2007) Pericytes of human skeletal muscle are myogenic precursors distinct from satellite cells. *Nat Cell Biol* 9: 255-267.
- Beltrami AP, Urbanek K, Kajstura J, Yan SM, Finato N, et al. (2001) Evidence that human cardiac myocytes divide after myocardial infarction. *N Engl J Med* 344: 1750-1757.
- Urbanek K, Torella D, Sheikh F, De Angelis A, Nurzynska D, et al. (2005) Myocardial regeneration by activation of multipotent cardiac stem cells in ischemic heart failure. *Proc Natl Acad Sci USA* 102: 8692-8697.
- Crippa S, Cassano M, Messina G, Galli D, Galvez BG, et al. (2011) miR669a and miR669q prevent skeletal muscle differentiation in postnatal cardiac progenitors. *J Cell Biol* 193: 1197-1212.
- Bi F, Liu N, Fan D (2003) Small interfering RNA: a new tool for gene therapy. *Curr Gene Ther* 3: 411-417.
- Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 116: 281-297.
- van Rooij E, Quiat D, Johnson BA, Sutherland LB, Qi X, et al. (2009) A family of microRNAs encoded by myosin genes governs myosin expression and muscle performance. *Dev Cell* 17: 662-673.

33. Wu S, Trievel RC, Rice JC (2007) Human SFMBT is a transcriptional repressor protein that selectively binds the N-terminal tail of histone H3. *FEBS Lett* 581: 3289-3296.
34. Liang R, Bates DJ, Wang E (2009) Epigenetic Control of MicroRNA Expression and Aging. *Curr Genomics* 10: 184-193.
35. Akagi K, Suzuki T, Stephens RM, Jenkins NA, Copeland NG (2004) RTCGD: retroviral tagged cancer gene database. *Nucleic Acids Res* 32: D523-D527.
36. Gupta V, Poss KD (2012) Clonally dominant cardiomyocytes direct heart morphogenesis. *Nature* 484: 479-484.
37. Porrello ER, Mahmoud AI, Simpson E, Hill JA, Richardson JA, et al. (2011) Transient regenerative potential of the neonatal mouse heart. *Science* 331: 1078-1080.
38. Sampaolesi M, Yoshida T, Iwata Y, Hanada H, Shigekawa M (2001) Stretch-induced cell damage in sarcoglycan-deficient myotubes. *Pflugers Arch* 442: 161-170.
39. Galvagni F, Cartocci E, Oliviero S (1998) The dystrophin promoter is negatively regulated by YY1 in undifferentiated muscle cells. *J Biol Chem* 273: 33708-33713.
40. Walowitz JL, Bradley ME, Chen S, Lee T (1998) Proteolytic regulation of the zinc finger transcription factor YY1, a repressor of muscle-restricted gene expression. *J Biol Chem* 273: 6656-6661.
41. Montini E, Cesana D, Schmidt M, Sanvito F, Ponzoni M, et al. (2006) Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. *Nat Biotechnol* 24: 687-696.
42. Cartier N, Hacein-Bey-Abina S, Bartholomae CC, Veres G, Schmidt M, et al. (2009) Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. *Science* 326: 818-823.
43. Cavazzana-Calvo M, Payen E, Negre O, Wang G, Hehir K, et al. (2010) Transfusion independence and HMG2A activation after gene therapy of human  $\beta$ -thalassaemia. *Nature* 467: 318-322.
44. Biffi A, Bartolomae CC, Cesana D, Cartier N, Aubourg P, et al. (2011) Lentiviral vector common integration sites in preclinical models and a clinical trial reflect a benign integration bias and not oncogenic selection *Blood* 117: 5332-5339.
45. Scaramuzza S, Biasco L, Ripamonti A, Castiello MC, Loperfido M, et al. (2012) Preclinical Safety and Efficacy of Human CD34(+) Cells Transduced With Lentiviral Vector for the Treatment of Wiskott-Aldrich Syndrome. *Mol Ther*.
46. Biasco L, Baricordi C, Aiuti A (2012) Retroviral integrations in gene therapy trials. *Mol Ther* 20: 709-716.
47. Bushman F, Lewinski M, Ciuffi A, Barr S, Leipzig J, et al. (2005) Genome-wide analysis of retroviral DNA integration. *Nat Rev Microbiol* 3: 848-858.
48. Cattoglio C, Facchini G, Sartori D, Antonelli A, Miccio A, et al. (2007) Hot spots of retroviral integration in human CD34+ hematopoietic cells. *Blood* 110: 1770-1778.
49. Brady T, Agosto LM, Malani N, Berry CC, O'Doherty U, et al. (2009) HIV integration site distributions in resting and activated CD4+ T cells infected in culture. *AIDS* 23: 1461-1471.
50. Felice B, Cattoglio C, Cittaro D, Testa A, Miccio A, et al. (2009) Transcription factor binding sites are genetic determinants of retroviral integration in the human genome. *PLoS one* 4: e4571.
51. Lewinski MK, Yamashita M, Emerman M, Ciuffi A, Marshall H, et al. (2006) Retroviral DNA integration: viral and cellular determinants of target-site selection. *PLoS Pathog* 2: e60.
52. Hotarkova S, Hermanova M, Povysilova V, Dvorak K, Feit J, et al. (2004) Demonstration of MyoD1 expression in oncocytic cardiomyopathy: report of two cases and review of the literature. *Pathol Res Pract* 200: 59-65.
53. Ikeda S, Kong SW, Lu J, Bisping E, Zhang H, et al. (2007) Altered microRNA expression in human heart disease. *Physiol Genomics* 31: 367-373.
54. Jayawardena TM, Egemnazarov B, Finch EA, Zhang L, Payne JA, et al. (2012) MicroRNA-Mediated In vitro and In vivo Direct Reprogramming of Cardiac Fibroblasts to Cardiomyocytes. *Circ Res* 110: 1465-1473.

This article was originally published in a special issue, **Cardiac Resident Stem Cells & Progenitors** handled by Editor(s). Dr. Vincenzo Lionetti, Scuola Superiore Sant Anna University, Italy