Defining the Diaphragm Satellite Cell Pool using ES Cell-Derived Myogenic Progenitors

A. Specific Aims:

In Duchenne (DMD) and other types of muscular dystrophies, the diaphragm is one of the most affected skeletal muscles and its progressive weakening is the cause of the respiratory dysfunctions leading these patients to death. In normal conditions, muscle homeostasis is ensured by satellite cells, a population of resident stem cells located under the muscle basal lamina. Since the diaphragm represents an important target for the treatment of muscular dystrophies, it is fundamental to understand the molecular regulation of satellite cells isolated from this muscle in order to develop an appropriate therapeutic strategy. Only with the recent advances in genomic analysis, deciphering the molecular mechanisms that regulate each specific cell type became a reachable goal. Moreover, the development of protocols for the culture and differentiation of both mouse and human Embryonic Stem (ES) cells has provided the unprecedented opportunity to study and manipulate cellular events that before could be only partially investigated. Our laboratory has demonstrated that differentiating ES cells adopt the skeletal myogenic cell fate upon up-regulation of the transcription factors Pax3 or Pax7. Using a doxycycline (dox)-inducible expression system, we developed a protocol to generate and expand a population of ES-derived skeletal myogenic progenitors, which, upon transplantation in a mouse model of DMD, are able to produce Dystrophin+ myofibers and to seed the satellite cell pool. The ability of ES-derived myogenic progenitors to repopulate the muscle niche makes these cells the closest cell model to study satellite cells. Importantly, whereas isolation of satellite cells from muscle tissues provides only a few hundred thousand cells, in vitro differentiating ES cells have tremendous proliferative potential and are easily amenable to genetic modification, allowing for the generation of large numbers of myogenic progenitors. Taking advantage of these premises, in this proposal we aimed 1) to identify the molecular determinants characterizing diaphragm satellite cells and investigate their role by using an ES-derived myogenic population through co-expression of both Pax7 and Pax3; and 2) test the ability of our in vitro system to repair the dystrophic diaphragm upon intravenous injection in murine models.

B. Studies and Results:

B1. To dissect the Pax7+Pax3+-co-dependent transcriptional regulation

To expand our knowledge of the molecular mechanisms regulating specifically diaphragm satellite cells, we decided to perform an unbiased analysis of this cell population by employing different genome wide approaches. As shown in Fig. 1A, whole transcriptome (RNA-seq) analysis revealed a subset of genes specifically up-regulated in diaphragm satellite cells compared to their hindlimb counterparts. As expected, and in agreement with the studies mentioned above, Pax3 is one of the genes detected in the diaphragm-specific group. Importantly, based on this finding, we speculate this approach will allow the identification of new genes regulating diaphragm satellite cells. About 200 genes were found significantly up-regulated (fold induction>2) in diaphragm- vs hindlimb- satellite cells. Gene ontology analysis revealed that these genes are associated with different biological processes, including RNA processing, proliferation, tissue morphogenesis and chemotaxis (Fig. 1B). Because one of the goals of this study is dissecting the transcriptional mechanisms governing diaphragm satellite cells, we initially focused our interest on transcription factors and chromatin associated proteins. As expected, Pax7 was equally expressed in both groups while Pax3 is up-regulated in satellite cells isolated from the diaphragm (Fig. 1C). Interestingly, some of genes upregulated in the diaphragm satellite cells have been described to regulate proliferation of progenitors during development as well as in the adult. Based on these evidences, it is reasonable to hypothesize that these candidate genes may contribute to maintain the satellite cell pool in the diaphragm.

In addition to RNA-seq, to better understand the transcriptional regulation of diaphragm satellite cells, we envisioned the necessity of studying the epigenetic regulation of these cells. Because of their rarity, it has been challenging to isolate sufficient numbers of satellite cells for reliably performing Chromatin-immunoprecipitation followed sequencing (ChIP-seq) studies in satellite cells. This represents a major obstacle for dissecting the role of transcription factors in satellite cells. We believe in the importance of understanding these mechanisms and, to overcome the low cell number limitation, we decided to address this question by analyzing chromatin accessibility using ATAC-seq. Chromatin accessibility correlates with transcription factor binding and nucleosome occupancy and therefore provides information about the epigenetic status of a specific locus. The advantage of this technique is that it requires a small cell number and limited sample processing prior to

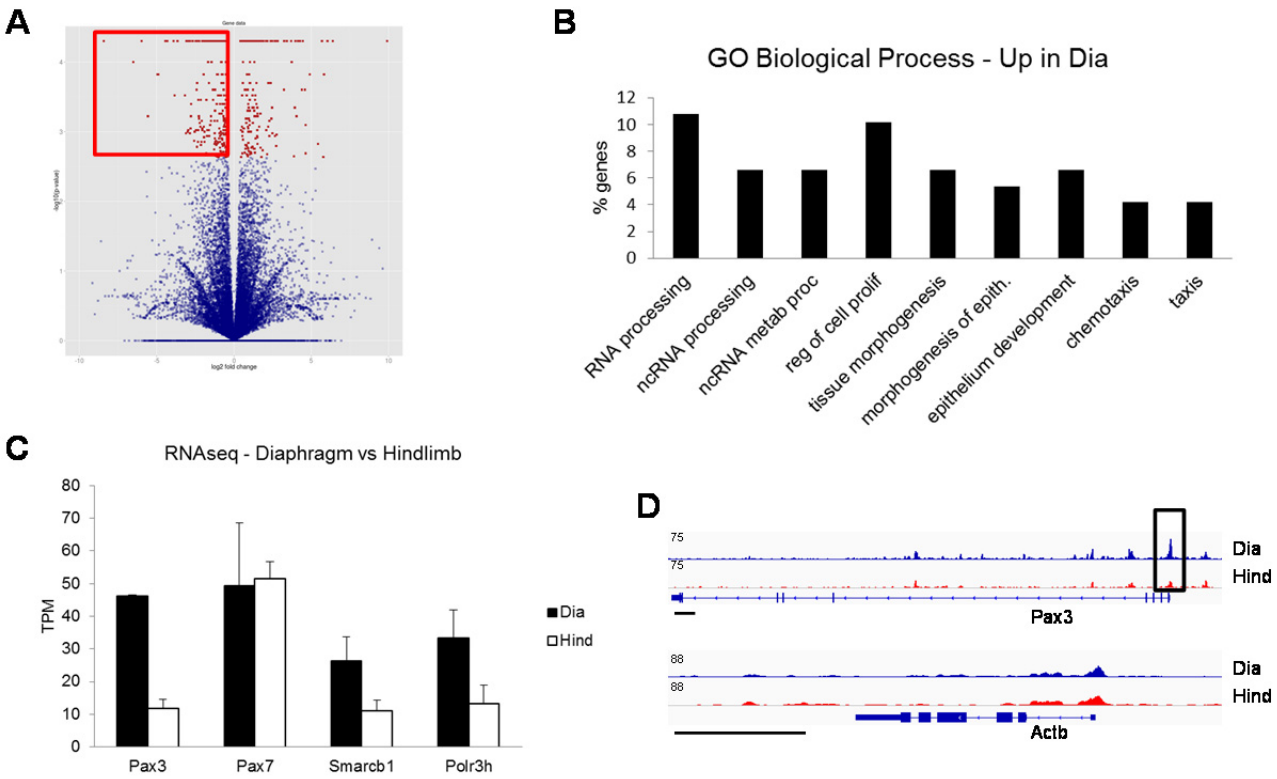

Figure 1. Transcriptional and epigenetic analysis of Diaphragm Satellite cells. A) Volcano plot showing RNA-seq data from Hindilimb vs Diaphragm satellite cells. Red square represents genes up-regulated in the Diaphragm. B) Functional classification of genes up-regulated in Diaphragm (revise this panel overhead: "Up in Dia". C) Expression levels of selected genes from RNA-seq. TPM: Transcript Per Million reads. D) Genome tracks displaying ATAC-seq data for selected genes. Bar: 5Kb.

sequencing. As shown in Fig. 1D, ATAC-seq allowed us to identify loci characterized by increased accessibility in the genes up-regulated in diaphragm satellite cells. These loci may represent important regulatory regions driving diaphragm-specific gene expression, and further analyses are required to define their role in this process.

In parallel to the studies on diaphragm satellite cells, we also developed an in vitro model based on ES cells to investigate in more detail the cooperation between Pax7 and Pax3. Using myogenic progenitors expressing Pax3-only, Pax7-only and both Pax7 and Pax3 (Pax7-ires-Pax3), we performed chromatin-immunoprecipitation followed by sequencing to study the genome wide binding of Pax3 and Pax7. As shown in Fig. 2, we successfully detected the genomic binding of these transcription factors in ES-derived myogenic progenitors, although we observed a lower binding of Pax3 when co-expressed with Pax7 (Pax7-ires-Pax3). We explain this observation with the lower protein expression of the second cDNA in the bicistronic construct. Nonetheless, at few sites Pax3 binding is still maintained (downstream of Prex1). We are currently analyzing these data to identify differences in the genomic binding due to co-expression of Pax7 and Pax3 that might account for the increased proliferation observed for this line.

B2. To determine the engraftment potential of Pax7+Pax3+ myogenic progenitors into the diaphragm

Because of its location, diaphragm is not an accessible muscle for direct intramuscular transplantation and therefore, any cell therapy approach would require cells to possess an intrinsic homing capability. Our first attempts to deliver myogenic progenitors upon intravenous transplantation evidenced a lower level of engraftment in this muscle. Because other investigators have demonstrated that extensive exercise in dystrophic mice (e.g. 30 minutes of swimming) promotes damage, thus making the mouse pathology closer in severity to the human patient, and improves cell homing to diaphragm, we envisioned that increasing the extent of tissue damage could improve the engraftment of myogenic progenitors into this muscle. For this

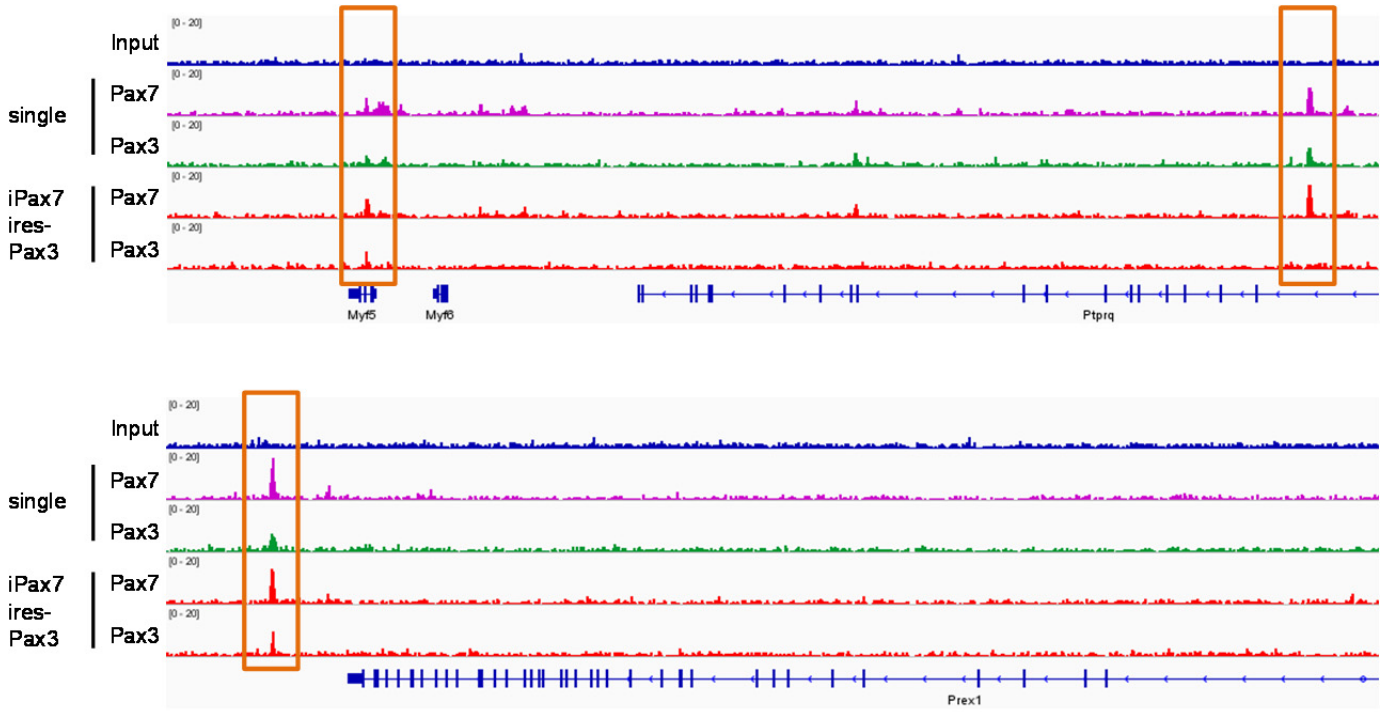

Figure 2. ChIP-seq comparison of Pax3 and Pax7 binding in myogenic progenitors. Inducible Pax3 or Pax7 (single) myogenic progenitors were compared to inducible Pax7-ires-Pax3 cells, which express both transcription factors. Selected IGV tracks show Pax3 and Pax7 binding at Myf5 (upper panel) and Prex1 (lower panel) regulatory regions.

reason, we generated a new mouse model in which tissue damage can be induced by expression of the Diphtheria Toxin (DTA). Importantly, DTA gene is encoded in the ROSA locus and is expressed only upon CRE-mediated recombination, which in our system is driven by the tissue specific Human Skeletal muscle Actin (HSA) promoter. Because a constitutive CRE-driver would be active since the embryonic development and would probably result in embryonic lethal phenotype, CRE activity is regulated by a ligand-dependent system to induce nuclear translocation of the recombinase only upon tamoxifen injection (Mer-CRE-Mer). By generating this new model, we have now a better tool to study the role of new candidate regulators of diaphragm satellite cells identified in our genome wide studies. Importantly, we began the characterization of the *HSA-MCM:ROSA-DTA:mdx^{4cv}* mice by collecting the diaphragm 3 days post tamoxifen injection (Fig. 3). As expected, compared to DTA-negative muscles, DTA+ diaphragms display tissue degeneration thus demonstrating this approach might represent an improvement of the current methods to injure this muscle.

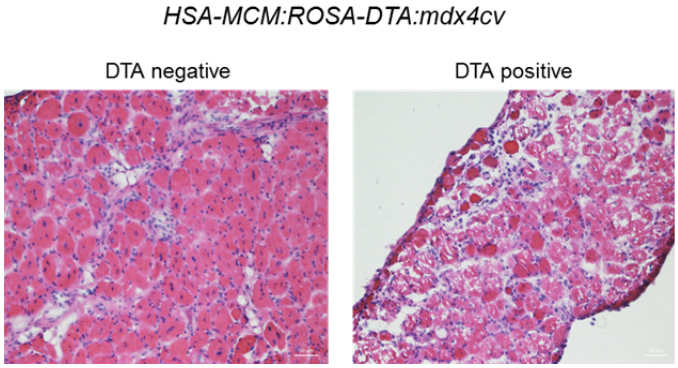

Figure 3. Muscle degeneration upon tamoxifen-injection in *HSA-MCM;ROSA-DTA;mdx^{4cv}* mice. Hematoxylin-Eosin staining of cryosections from diaphragm of DTA negative and DTA positive tamoxifen-injected animals.

C. Significance:

Our goal is to identify the molecular mechanisms distinguishing diaphragm satellite cells and apply this knowledge to improve regenerative potential of ES-derived myogenic progenitors for treatment of muscular dystrophies. These studies have also important implications for basic developmental biology.

D. Publications:

Oliveira NAJ*, Magli A*, Zhan WZ, Oliveira VKP, Ortiz-Cordero C, Selvaraj S, Ervasti JM, Mantilla CB, and Perlingeiro RCR. Rescue of Diaphragm Function upon Infusion of Pluripotent-Derived Myogenic Progenitors in a Mouse Model of Dystroglycanopathy. *Submitted*. *cofirst author.