University of São Paulo
"Luiz de Queiroz" College of Agriculture

Functional analysis of the *SAP30* gene and its relation to muscle development

Bruna Petry

Thesis presented to obtain the degree of Doctor in Science. Area: Animal Science and Pastures

Piracicaba
2021

Bruna Petry
Biologist

Functional analysis of the *SAP30* gene and its relation to muscle development

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Advisor:
Prof. Dr. **LUIZ LEHMANN COUTINHO**

Thesis presented to obtain the degree of Doctor in Science. Area: Animal Science and Pastures

Piracicaba
2021

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP

Petry, Bruna

Functional analysis of *SAP30* gene and its relation to muscle development / Bruna Petry. – – versão revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2021.

63 p.

Tese (Doutorado) - - USP / Escola Superior de Agricultura “Luiz de Queiroz”.

1.Hipertrofia 2. C2C12 3. CRISPR/Cas9. 4. siRNA. 5. RNA-seq. I. Título

DEDICATION

I dedicate this thesis to the loves of my whole life: my father and my mother.

ACKNOWLEDGMENTS

First, my parents - Ivanier and Malizi Petry, for all love and support during my whole life. They always went to great lengths to support me in the craziest decisions. You are everything for me. It was and always will be for you. Also, to my life mate Amora (dog), for always welcoming me with joy, even on difficult days.

My grandparents: Alma F. Massing, Silário Massing (*in memoriam*), Benno A. Petry (*in memoriam*), Ilse S. Petry (*in memoriam*) for all the love and care. Love you!

My advisor Dr Luiz Lehmann Coutinho for trusting me with the challenge of executing this project, in addition to all the teachings over these 4 years. Thank you!

Coutinho's Lab friends: Aline, Francisco, Thais, Raquel, Simone, Fernanda, Nirlei, Jorge, Ricardo, Giovana, Carol (Purcell), Carol (Fernandes), Marcela, Pilar, Mirele, Barbara, Vinicius, Fabio, Luan, Graciela, Marianne, Simara, for all support.

Best and amazing friends: Gabriel (Gabu), Mayara (Guria), Priscila (Pri), tia Rita, tio Luis, tia Marlene (*in memoriam*), Suelen, Marinara, Aline Viancelli, Fernanda Rigon, Fabricio, Izally, Morgana, Aline Pelizzer, Francieli, Ju Matté, Fernanda D'agostini, Gisele, Andressa, Tai, José Luis, Shilton, Evelin, for always being with me, having a coffee, drinking a beer, talking in long calls, or just sending good vibes, even geographically distant. Love you!

Friends from Dra Erika Jorge's Lab: Aline (Copolinha), Rayan, Julia, Iago, Clara, Ricardo, Chico, and a special thanks for Dra Erika and Dra Gerluza for all the support during my time in Belo Horizonte, MG.

Friends from Koltes' lab: James, Dawn, Ellie, Mary Sue, Kayla, Cori, Abbey, Kyu-Sang, Luke. Thanks for all the learning, exchange of experiences, enthusiasm, friendship and all the welcome during my stay in Ames. I had a great time with you, guys!

My Amazing American Mom, Judy, thanks for everything! Love you!

My Ames' family: Raquel, Thamires, Sonia, Oscar, Sebastian, Esneider, Juber, Tomas, Marcela, David, Athena, Megan, Thales, Julia, Letícia, Ithalo, Guilherme, Debora, Iris, Camila, Ana Paula, Anelise, Raissa.

Finally, I wish to express my gratitude to University of São Paulo – ESALQ for the privilege to studied in this college, aggregating knowledge in my academic life. A special acknowledgment to CAPES and FAPESP (process number 2017/25119-1) for the funding support.

SUMMARY

DEDICATION 3
ACKNOWLEDGMENTS 4
1. Introduction 9
1.1 Main Objective 12
1.2 Specific objectives 12
1.3 Systematization text 13
1.4 References 14
2. SAP30 GENE IS A PROBABLE REGULATOR OF MUSCLE HYPERTROPHY IN CHICKENS ... 17
2.1 Abstract 17
2.2 Introduction 17
2.3 Methods 19
2.4 Results 21
2.5 Discussion 23
2.6 Conclusion 25
2.7 References 25
3. UNRAVELING THE SAP30 GENE FUNCTION IN MUSCLE CELLS TRHOUGH CRISPR/CAS9 34
3.1 Abstract 34
3.2 Introduction 34
3.3 Methods 35
3.4 Results 38
3.5 Discussion 57
3.6 Conclusion 58
3.7 References 60
APPENDIX 63

RESUMO

Análise funcional do gene *SAP30* e sua relação com o desenvolvimento muscular

Entender a maquinaria molecular envolvida em determinadas características de interesse econômico é importante para que informações genéticas possam ser aplicadas junto aos programas de melhoramento animal, visando estabelecer maior acurácia na seleção genômica. Estudos de GWAS identificaram regiões de QTL e CNV localizadas no cromossomo 4 na espécie *Gallus gallus,* em uma população brasileira de frangos de corte, associadas com características de desenvolvimento muscular, como peso e porcentagem de peito. Buscas por genes candidatos presentes nessas regiões e que pudessem estar participando do desenvolvimento muscular foram realizadas. O gene *SAP30* foi identificado como um possível candidato ao desenvolvimento muscular. Diante disso, e com o objetivo de validar sua função, *SAP30* foi editado usando CRISPR/Cas9 e teve sua expressão reduzida por meio de siRNA, em cultura de mioblastos *in vitro*. Após essas análises, análises morfométricas foram realizadas na cultura celular para avaliar as alterações ocorridas no tamanho de área dessas células. Posteriormente, as análises de sequenciamento de RNA, genes diferencialmente expressos e enriquecimento funcional foram realizadas para investigar redes gênicas e processos biológicos que as mutações e/ou redução da expressão do *SAP30* estariam ocasionando. Os resultados para ambos os trabalhos mostram o envolvimento do *SAP30* na hipertrofia de células musculares. Tanto a redução na expressão do gene, quanto a indução de mutações em sua sequência, fizeram com que os miotubos tivessem maior área nos grupos tratados, quando comparado aos controles. As análises de sequenciamento de RNA revelaram, para ambos, uma importante diferença na expressão de genes relacionados ao desenvolvimento muscular, bem como, genes que atuam em vias de formação e regeneração do músculo. Assim, pela primeira vez, apresentamos o gene *SAP30* como um possível regulador de hipertrofia muscular.

Palavras-chave: Hipertrofia. Crescimento muscular. C2C12. siRNA. CRISPR/Cas9.

ABSTRACT

Functional analysis of *SAP30* gene and its relation to muscle development

Understanding the molecular machinery involved in certain traits of economic interest is important because the genetic information can be applied to animal breeding programs, defining greater accuracy in genomic selection. GWAS studies have identified general QTL and CNV regions on chromosome 4 in *Gallus gallus* species, in a Brazilian broiler population, associated with muscle development traits such as weight and percentage of breast. Searches for candidate genes present in these regions and that could be participating in muscle development were performed. The *SAP30* gene was identified as a possible candidate for muscle development. Therefore, and to validate its function, *SAP30* was edited using CRISPR/Cas9 and had its expression reduced by means of siRNA, in myoblast culture in vitro. After these analyses, morphometric analyzes were performed in cell culture to evaluate changes in the area size of these cells. Subsequently, RNA sequencing, differentially expressed genes and functional enrichment analyzes were performed to investigate gene networks and biological processes that mutations and/or reduced expression of *SAP30* would be causing. The results for both papers presented the involvement of *SAP30* in muscle cell hypertrophy. Both the reduction in gene expression and the induction of mutations in its sequence made the myotubes have a larger area in the treated groups when compared to controls. RNA sequencing analyzes revealed, for both, an important difference in the expression of genes related to muscle development, as well as genes that act in muscle formation and regeneration pathways. So, for the first time, we present the *SAP30* gene as a possible regulator of muscle hypertrophy.

Key-words: Hypertrophy. Muscle growth. C2C12. siRNA. CRISPR/Cas9.

1. INTRODUCTION

In the world ranking, Brazil occupies the place of largest exporter and third largest producer of chicken meat (ABPA, 2020) and, among the agricultural activities, poultry is one of the most prominent. The increase in chicken meat exports was associated with good production practices, such as improvements in the poultry, in the nutritional quality of the feed, and genetic improvement for selection of production traits (Avila et al., 2007).

The intense genetic improvement that has occurred in recent years has brought advances in poultry production, as muscle development. The advances in molecular biology technologies, the discovery of candidate genes, and the identification of mutations related to the traits of zootechnical interest in chickens, have become part of the genetic improvement programs, providing greater accuracy in the results of genomic selection (Dekkers, 2012).

The chicken meat quality and improvement are important to the poultry industry, also for the human nutrition. Chicken meat has a high nutritional value, with high protein and low-fat contents (Kim et al., 2020). Understanding the muscle development is an important strategy to improve muscle deposition. Several quantitative trait loci (QTL) associated with multiple traits of economic interest have been mapped, including growth and body composition, egg production, resistance to diseases, and responses to antibodies. As an example, in a study to identify SNPs (single nucleotide polymorphisms) candidates for important characteristics in the *Gallus gallus* species, genetic variations were found close to important disease resistance genes such as *TGF-β17*, cytokines, histocompatibility complex genes (MHC), in addition to genes related to antibody responses, such as *IL4* and *IL13* (Wong et al., 2004).

Other studies have identified genes and genetic variations related to the important developmental traits of chickens, as muscle fiber and carcass quality characteristics (Zhang et al., 2008), body composition and meat quality (Liu et al., 2013), and abdominal fat, skin fat and body fat distribution (Ikeobi et al., 2002). Boschiero et al. (2013) studied SNP-type mutations in the *IGF1* and *KDM5A* genes and observed that in the *IGF1* gene this genetic variation is associated with characteristics of body growth, feed conversion, and percentage of abdominal fat, while a SNP in the *KDM5A* gene was associated with fat deposition abdominal in chickens. Felício et al. (2013) worked with mutations associated with chicken meat quality, as thigh size, loss from cooking, and body weight. Also, candidate genes located in QTLs involved with muscle growth (Pértille et al., 2015), fat deposition (Moreira et al., 2015), and muscle deposition (Godoy et al., 2015) were studied. Recently our group studied QTLs associated with muscle traits, by Genome-Wide Association Studies (GWAS), and one region in the chromosome 4 (43Mb) was associated with breast muscle yield in the TT Reference population (Trevisoli et al., 2021), also a study using the same chicken population detected Copy Number Variations (CNV) associated with performance traits, including birth weight, body weight, feed intake, feed conversion ratio, and body weight gain (Fernandes et al., 2021).

Studies discovering candidate genes associated with economical traits are important to characterize the metabolic and genetic pathways involved in determine the phenotype. As previously mentioned, there are many studies that have already identified mutations and genes associated with important traits, however, validation studies using cell culture model of the genes are still scarce. With this aim, Petry et al. (2021) used QTL and CNV information from TT Reference Population (Marchesi et al., 2018) to select a candidate gene to perform its validation in muscle cell culture, *in vitro*. The *SAP30* gene was identified overlapping in a QTL and CNV region associated with breast meat and percentage in the same broiler chicken population (TT Reference population). This gene has a gene ontology term associated with skeletal muscle cell development and was chosen as a candidate for validation in our study.

Sin3A-Associated protein 30 (*SAP30*) gene is a member of the Sin3A corepressor complex and was demonstrated important in the cell growth affecting gene expression (Zhang et al., 1998), and participating in N-CoR-mediated repression by specific transcription factors (Laherty et al., 1998). As an example, a study has shown that *SAP30* is a transcription factor participating in the *NETO2* gene expression, in a renal cell carcinoma in humans (Snezhkina et al., 2018). The SAP30 protein is conserved between humans and yeast, acting as a histone deacetylase complex, regulating gene expression (Zhang et al., 1998), moreover, the same protein is conserved between human, chicken and mouse, where chicken and mouse genomes have a similar and conserved GALNT-SAP30-HAND organization, showing the evolutionary similarity between these species (Viiri et al., 2009).

Even with some evidence of the *SAP30* gene and protein participation as a transcription factor, and the fact that it has been found as a possible candidate for muscle development in chickens (QTL and CNV analyzes), the role of this gene in muscle development is still unknown. In view of this, the use of techniques as gene editing (such as CRISPR/Cas9, Clustered Regulatory Interspaced Short Palindromic Repeats) and small interfering RNAs (siRNA) has enabled to validate the effect of the candidate genetic changes in the control of the phenotype by the activation or repression of parts of the genome. Cellular reprogramming and modeling of the organism for adaptations against degenerative diseases or genetic alterations that cause damage and have brought benefits in the medical, human health and zootechnical areas (Du and Qi, 2016; Singh et al., 2017), where there is great importance as a tool to understand the genetic machinery behind the characteristics of interest.

The CRISPR technique has been providing discussions of possible applications in genetic engineering in several areas of knowledge. Firstly described as an alternative to adaptive immunity in prokaryotes, this mechanism is based on the incorporation of small fragments of invading DNA into the host genome, through the enzyme Cas, making the organism - previously pathogenic to the prokaryote - be recognized as similar (Burmistrz and Pyrc, 2015). Knowing about the sequence of the adaptive immunity process of these microorganisms, the model started to be adopted in the biological field as an alternative technique for genome editing (Makarova et al., 2011).

The editing or knockout of genes has been widely used in several animal models. In mice, for example, CRISPR/Cas9 was used to change SNP-like bases in embryonic cells, and the efficiency of changing the Thymine base to Cytosine (T>C) was 100% successful (Liang et al., 2017). Also, the use of the CRISPR/Cas9 system obtained results of 90 to 100% efficiency in silencing complete genes in cells of mice and monkeys (Zuo et al., 2017). Although further studies are needed, research has recently been published that addresses the use of the CRISPR/Cas9 system in the removal of the HIV virus from mouse cells, obtaining good results in eliminating the virus from host cells (Yin et al., 2017) and demonstrating a significant step towards human clinical trials. Another study related to the well-being and control of zoonoses was carried out in cattle (Gao et al., 2017), where bases were inserted in the animal's genome to make it resistant to tuberculosis. In this study, the authors achieved success in the formation of transgenic fibroblast cells by inserting a DNA sequence responsible for the formation of macrophage proteins (NRAMP1) which is responsible for increasing resistance to this pathogenesis in cattle.

In chickens, studies also demonstrate the importance of the technique for editing genes responsible for characteristics of zootechnical interest. An example of this was a study in which the differentiation of embryonic stem cells (ECSs) into sperm cells (SSCs) was investigated by knocking out the *Stra8* gene - a gene typically expressed in mitosis and meiosis events in mammalian cells (Zhang et al., 2017). As a result, the knockout of the gene was confirmed and the generation of SSCs was blocked in these cells, indicating that the CRISPR/Cas9 system was effective to perform the knockout of the *Stra8* gene in chicken cells, inhibiting the differentiation of ECSs into SSCs (Zhang et al., 2017). The positive results that the studies with CRISPR/Cas9 have been obtaining suggest that this technique can be used in the edition of candidate genes to regulate specific traits, to characterize the function of the genes, since most of the characteristics of interest are polygenic control, and the gene knockout aiming at the loss of function of a gene is important to know its participation in the regulation of the phenotype.

Small Interfering RNAs (siRNAs) are 20-30 nucleotides noncoding RNAs that regulate genes and genomes, in different genomic levels such as chromatin structure, RNA processing and stability, transcription, and translation (Carthew and Sontheimer, 2009). The use of siRNA to silencing a gene to investigate its function is well stablished in mammalian cells and may help to identify the importance of the specific genes in determinate traits (Sakurai et al. 2010). Figeac et al., (2019) studied the influence of pathways involved in the myogenesis in skeletal muscle, silencing a VGLL3 protein by siRNA. Another study investigated the participation of the *CSRP3* gene in the muscular dystrophy using chickens' primary myoblasts, here the gene knockdown was performed by siRNA to evaluate its function (Cui et al., 2020). As CRISPR/Cas9, the siRNA is also useful to modulate the gene expression, making it possible to understand the molecular machinery involved behind the traits, as muscle development.

Important processes of development, growth, and regeneration of skeletal muscle occur throughout the life of vertebrates and all these phases are regulated by the same myogenic markers. C2C12 cells, for example, reproduce in a controlled and fast manner all the steps of myogenesis, which include the determination, proliferation, and differentiation of skeletal muscle cells (Lee et al., 2005). Therefore, the *in vitro* system is a good model for studies of characterization and function of genes or mutations that are associated with this cell type. Understanding the cellular and molecular programming that determines the formation of skeletal musculature is the basis for it to be possible to characterize the main role of candidate genes in the development, proliferation, and differentiation of muscle tissue. Therefore, our hypothesis is that the silencing of a candidate gene associated with muscle development can modulates the differentiation of muscle cells.

1.1 Main Objective

The general objective was to identify a candidate gene associated with muscle development in chickens from previously study of our research group and validate its function in muscle cell culture applying siRNA and CRISPR/Cas9 tools.

1.2 Specific objectives

a) Identify a candidate gene localized in quantitative trait loci (QTL) and copy number variation (CNV) regions that were previously associated with muscle development in a broiler chicken population.

b) Apply the gene knockdown on muscle cells cultured by using siRNA technique and evaluate the differential gene expression between the control and treated groups.

c) Apply the gene editing on muscle cells cultured by using CRISPR/Cas9 technique, evaluate the mutations occurred in the DNA sequence and the differential gene expression between the groups of cells.

1.3 Systematization text

Many studies have identified variants and genes that are candidates for certain characteristics of economic interest using genetic information from farm animals. However, studies aiming to validate the specific function of these variants and genes are scarce. Therefore, our work has contributed to many of these validation studies, bringing the idea of using the information previously obtained through bioinformatics and statistical analysis and applying this information in a cell culture model. Thus, two studies were performed to validate the candidate *SAP30* gene previously identified in the QTL and CNV regions associated with muscle characteristics, by Genome-wide Association Study (GWAS). The first study used the siRNA technique to reduce the expression of the gene (SAP30-knockdown) and observe what would be the consequences of this knockdown in the culture of muscle cells. In the second study, also using *SAP30* as a candidate gene, the technique known as CRISPR/Cas9 was performed to edit the DNA sequence of this gene. Then, the changes that each type of mutation would also cause in muscle cell culture was evaluated. As a result, for both studies, we observed that the changes - both in the expression of the *SAP30* gene, and in the mutations caused by CRISPR/Cas9 - conferred a hypertrophy phenotype in muscle cells. The analysis of RNA sequencing for SAP30-knockdown showed a difference in the expression level of genes related to muscle growth, development, and contraction; while the analysis of differentially expressed genes for cells that received CRISPR/Cas9 showed genes related to important pathways of embryogenic muscle development, such as the NOTCH and WNT pathways. Thus, the two studies presented below showed through morphometric and molecular analysis of cell culture, the possible participation of the *SAP30* gene as a regulator of muscle hypertrophy.

1.4 References

ABPA (2020). Associação Brasileira de Proteína Animal, relatório anual 2020.

Avila, V., Kunz, A., Bellaver, C., Jaenisch, F.R.F., Mazzuco, H., Trevisol, I.M., et al. (2007). Boas práticas na produção de frangos de corte. Available at: https://www.infoteca.cnptia.embrapa.br/infoteca/bitstream/doc/433206/1/publicacaos8t285e.pdf [Accessed May 20, 2021].

Boschiero, C., Jorge, E. C., Ninov, K., Nones, K., do Rosário, M. F., Coutinho, L. L., et al. (2013). Association of IGF1 and KDM5A polymorphisms with performance, fatness and carcass traits in chickens. *J. Appl. Genet.* 54, 103–112. doi:10.1007/s13353-012-0129-6.

Burmistrz, M. Pyrć, K (2015). CRISPR-Cas Systems in Prokaryotes. Available at: http://www.pjmonline.org/wp-content/uploads/archive/vol6432015193.pdf [Accessed May 20, 2021].

Carthew, R. W., and Sontheimer, E. J. (2009). Origins and Mechanisms of miRNAs and siRNAs. *Cell* 136, 642–655. doi:10.1016/j.cell.2009.01.035.

Cui, C., Han, S., Tang, S., He, H., Shen, X., Zhao, J., et al. (2020). The autophagy regulatory molecule CSRP3 interacts with LC3 and protects against muscular dystrophy. *Int. J. Mol. Sci.* 21, 749. doi:10.3390/ijms21030749.

Dekkers, J. (2012). Application of Genomics Tools to Animal Breeding. *Curr. Genomics* 13, 207–212. doi:10.2174/138920212800543057.

Du, D., and Qi, L. S. (2016). An introduction to CRISPR technology for genome activation and repression in mammalian cells. *Cold Spring Harb. Protoc.* 2016, 1–5. doi:10.1101/pdb.top086835.

Felício, A. M., Boschiero, C., Balieiro, J. C. C., Ledur, M. C., Ferraz, J. B. S., Michelan Filho, T., et al. (2013). Identification and association of polymorphisms in CAPN1 and CAPN3 candidate genes related to performance and meat quality traits in chickens. *Genet. Mol. Res.* 12, 472–482. doi:10.4238/2013.February.8.12.

Fernandes, A. C., da Silva, V. H., Goes, C. P., Moreira, G. C. M., Godoy, T. F., Ibelli, A. M. G., et al. (2021). Genome-wide detection of CNVs and their association with performance traits in broilers. *BMC Genomics* 22, 354. doi:10.1186/s12864-021-07676-1.

Figeac, N., Mohamed, A. D., Sun, C., Schönfelder, M., Matallanas, D., Garcia-Munoz, A., et al. (2019). VGLL3 operates via TEAD1, TEAD3 and TEAD4 to influence myogenesis in skeletal muscle. *J. Cell Sci.* 132. doi:10.1242/jcs.225946.

Gao, Y., Wu, H., Wang, Y., Liu, X., Chen, L., Li, Q., et al. (2017). Single Cas9 nickase induced generation of NRAMP1 knockin cattle with reduced off-target effects. *Genome Biol.* 18, 1–15. doi:10.1186/s13059-016-1144-4.

Godoy, T. F., Moreira, G. C. M., Boschiero, C., Gheyas, A. A., Gasparin, G., Paduan, M., et al. (2015). SNP and INDEL detection in a QTL region on chicken chromosome 2 associated with muscle deposition. *Anim. Genet.* 46, 158–163. doi:10.1111/age.12271.

Ikeobi, C. O. N., Woolliams, J. A., Morrice, D. R., Law, A., Windsor, D., Burt, D. W., et al. (2002). Quantitative trait loci affecting fatness in the chicken. *Anim. Genet.* 33, 428–435. doi:10.1046/j.1365-2052.2002.00911.x.

Kim, Hye-Jin Kim, JinJoo Jeon, Ki-Chang Nam, Kwan-Seob Shim, Jong-Hyun Jung, et al. (2020). Comparison of the quality characteristics of chicken breast meat from conventional and animal welfare farms under refrigerated storage | Elsevier Enhanced Reader. Available at: https://reader.elsevier.com/reader/sd/pii/S0032579119579202?token=27D282A847D410CFD0D335D086805CC6C2DE4DF85853477028F6E0C325C685A45A7A263807E85E865BD6B6E1B855946C&originRegion=us-east-1&originCreation=20210522111748 [Accessed May 22, 2021].

Laherty, C. D., Billin, A. N., Lavinsky, R. M., Yochum, G. S., Bush, A. C., Sun, J. M., et al. (1998). SAP30, a component of the mSin3 corepressor complex involved in N-CoR-mediated repression by specific transcription factors. *Mol. Cell* 2, 33–42. doi:10.1016/S1097-2765(00)80111-2.

Lee, M. H., Jang, M. H., Kim, E. K., Han, S. W., Cho, S. Y., and Kim, C. J. (2005). Nitric oxide induces apoptosis in mouse C2C12 myoblast cells. *J. Pharmacol. Sci.* 97, 369–376. doi:10.1254/jphs.FPJ04017X.

Liang, P., Sun, H., Sun, Y., Zhang, X., Xie, X., Zhang, J., et al. (2017). Effective gene editing by high-fidelity base editor 2 in mouse zygotes. *Protein Cell* 8, 601–611. doi:10.1007/s13238-017-0418-2.

Liu, R., Sun, Y., Zhao, G., Wang, F., Wu, D., Zheng, M., et al. (2013). Genome-Wide Association Study Identifies Loci and Candidate Genes for Body Composition and Meat Quality Traits in Beijing-You Chickens. *PLoS One* 8, 61172. doi:10.1371/journal.pone.0061172.

Makarova, K. S., Aravind, L., Wolf, Y. I., and Koonin, E. V (2011). Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. doi:10.1186/1745-6150-6-38.

Marchesi, J. A. P. et al. Relationship of runs of homozygosity with adaptive and production traits in a paternal broiler line. Animal, 12, 1126–1134 (2018)

Moreira, G. C. M., Godoy, T. F., Boschiero, C., Gheyas, A., Gasparin, G., Andrade, S. C. S., et al. (2015). Variant discovery in a QTL region on chromosome 3 associated with fatness in chickens. *Anim. Genet.* 46, 141–147. doi:10.1111/age.12263.

Pértille, F., Zanella, R., Felício, A. M., Ledur, M. C., Peixoto, J. O., and Coutinho, L. L. (2015). Identification of

polymorphisms associated with production traits on chicken (Gallus gallus) chromosome 4. *Genet. Mol. Res.* 14, 10717–10728. doi:10.4238/2015.September.9.11.

Trevisoli, P.A., Moreira, G.C.M., C, Cesar, A.S.M., Petrini, J., Ledur, M.C., Mourão, G.B., Garrick, D. and Coutinho, L.L. (2021). A Missense Mutation in the MYBPH Gene Is Associated With Abdominal Fat Traits in Meat-Type Chickens. Front. Genet. 12:698163. doi: 10.3389/fgene.2021.698163

Sakurai, K., Chomchan, P., and Rossi, J. J. Silencing of Gene Expression in Cultured Cells Using Small Interfering RNAs. *Curr. Protoc. Cell Biol* 47. doi:10.1002/0471143030.cb2701s47.

Singh, V., Braddick, D., and Dhar, P. K. (2017). Exploring the potential of genome editing CRISPR-Cas9 technology. *Gene* 599, 1–18. doi:10.1016/j.gene.2016.11.008.

Snezhkina, A. V., Nyushko, K. M., Zaretsky, A. R., Shagin, D. A., Sadritdinova, A. F., Fedorova, M. S., et al. (2018). Transcription Factor SAP30 Is Involved in the Activation of NETO2 Gene Expression in Clear Cell Renal Cell Carcinoma. *Mol. Biol.* 52, 385–392. doi:10.1134/S0026893318020152.

Viiri, K. M., Heinonen, T. Y., Mäki, M., and Lohi, O. (2009). Phylogenetic analysis of the SAP30 family of transcriptional regulators reveals functional divergence in the domain that binds the nuclear matrix. *BMC Evol. Biol.* 9, 1–15. doi:10.1186/1471-2148-9-149.

Wong, G. K. S., Liu, B., Wang, J., Zhang, Y., Yang, X., Zhang, Z., et al. (2004). A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. *Nature* 432, 717–722. doi:10.1038/nature03156.

Yin, C., Zhang, T., Qu, X., Zhang, Y., Putatunda, R., Xiao, X., et al. (2017). In Vivo Excision of HIV-1 Provirus by saCas9 and Multiplex Single-Guide RNAs in Animal Models. *Mol. Ther.* 25, 1168–1186. doi:10.1016/j.ymthe.2017.03.012.

Zhang, Y., Sun, Z. W., Iratni, R., Erdjument-Bromage, H., Tempst, P., Hampsey, M., et al. (1998). SAP30, a novel protein conserved between human and yeast, is a component of a histone deacetylase complex. *Mol. Cell* 1, 1021–1031. doi:10.1016/S1097-2765(00)80102-1.

Zhang, Y., Wang, Y., Zuo, Q., Li, D., Zhang, W., Wang, F., et al. (2017). CRISPR/Cas9 mediated chicken Stra8 gene knockout and inhibition of male germ cell differentiation. *PLoS One* 12, e0172207. doi:10.1371/journal.pone.0172207.

Zhang, Z. R., Liu, Y. P., Jiang, X., Du, H. R., and Zhu, Q. (2008). Study on association of single nucleotide polymorphism of CAPN1 gene with muscle fibre and carcass traits in quality chicken populations. *J. Anim. Breed. Genet.* 125, 258–264. doi:10.1111/j.1439-0388.2008.00723.x.

Zuo, E., Cai, Y. J., Li, K., Wei, Y., Wang, B. A., Sun, Y., et al. (2017). One-step generation of complete gene knockout mice and monkeys by CRISPR/Cas9-mediated gene editing with multiple sgRNAs. *Cell Res.* 27, 933–945. doi:10.1038/cr.2017.81.

2. SAP30 GENE IS A PROBABLE REGULATOR OF MUSCLE HYPERTROPHY IN CHICKENS

Bruna Petry[1], Gabriel Costa Monteiro Moreira[1], Aline Gonçalves Lio Copola[2], Marcela Maria de Souza[3], Fernanda Cristina da Veiga[1], Erika Cristina Jorge[2], Jane de Oliveira Peixoto[4], Mônica Corrêa Ledur[4], James E. Koltes[3], Luiz Lehmann Coutinho[1]

[1] Animal Science Department, University of São Paulo (USP) / Luiz de Queiroz College of Agriculture (ESALQ), Piracicaba, SP, Brazil
[2] Department of Morphology, Institute of Biological Sciences, Federal University of Minas Gerais (UFMG), Belo Horizonte, MG, Brazil
[3] Animal Science Department, Iowa State University (ISU), Ames, IA, United States of America
[4] Embrapa Suínos e Aves, Concórdia, SC, Brazil

* Correspondence:
Luiz Lehmann Coutinho[1*]
E-mail: llcoutinho@usp.br

Keywords: Hypertrophy, siRNA, Knockdown, skeletal muscle, C2C12, muscle cell growth

2.1 Abstract

Despite recent advances in QTL mapping and genome selection, it is still difficult to precisely establish the effects of candidate genes on the phenotype. Understanding the role of molecular gene networks in the regulation of the phenotype helps to establish strategies to improve chicken production, increase product value and reduce production costs. The aim of this study was to evaluate *SAP30* as a candidate gene for muscle development and to validate its function in cell culture *in vitro*. The *SAP30* gene was down regulated in C2C12 muscle cell culture using siRNA technology to evaluate its impact on morphometric traits and gene expression by RNA-seq analysis. Modulation of *SAP30* expression increased C2C12 myotube area, indicating a role in muscle hypertrophy. RNA-seq analysis identified several upregulated genes annotated in muscle development in treated cells (SAP30-knockdown), corroborating the role of *SAP30* gene in muscle development regulation. Here, we provide experimental evidence of the involvement of *SAP30* gene as a probable regulator of muscle cell hypertrophy.

2.2 Introduction

The intense genetic improvement that has occurred in recent years has increased poultry production, remodeling the industry over time (Saxena and Kolluri, 2018). Initially, chickens were selected for greater meat production by crossings lines previously selected for fast growth, carcass yield, and better feed conversion (Thiruvenkadan et al., 2011; Zuidhof et al., 2014). Despite the great advances in productive indexes, some undesirable effects have been observed affecting primarily the musculoskeletal (Julian, 2005) and cardiovascular systems (Olkowski, 2007). Ongoing selection for growth rate, carcass and prime cuts yields are correlated with pathophysiological changes in muscle tissue in modern lines, mainly the increase of breast myopathies (Soglia et al., 2019; Maharjan et al., 2020).

Advances in genomic technologies have led the discovery of single polymorphism nucleotide (SNP), development of SNP panels, and applications in animal breeding (Dekkers, 2012). Aiming to identify genomic regions and putative candidate genes, several Genome-Wide Association Studies (GWAS) were conducted. Quantitative trait loci (QTL) have been identified in chickens for many traits of economic interest, including growth and body composition (Hu et al., 2019, Animal QTL database, 2021), albumin height, Haugh score, shell shape, final body weight, gain, and feed efficiency (Hansen et al., 2005, Animal QTL database, 2021). Our research group has characterized the genetic architecture for growth and body composition traits in an experimental population originated from a Brazilian broiler line (TT Reference Population). Candidate genes participating in the performance, organs, carcass traits (Venturini et al., 2015) and bone integrity (Fornari et al., 2014; Grupioni et al., 2017), were already discovery in this broiler population. Also, QTL associated with production traits, including: carcass fat deposition (Moreira et al., 2018), feed efficiency and body weight (Moreira et al., 2019). Using the same broiler population, a study searching for Copy Number Variation Regions (CNVR) was performed to find regions possibly associated with performance traits (Fernandes et al., 2021) and one of the CNVR was associated with breast yield (unpublished results).

Results from previous QTL (Trevisoli et al., 2021) and CNVR (unpublish results) analysis, found a genomic region that was associated with breast meat yield using information about TT Reference population (Marchesi et al., 2018). Considering the possible importance of this region in the muscle architecture, we performed the data mining to search for genes that are annotated in this region and may presented a role in muscle development. *SAP30* gene was found to be overlapping with a QTL and CNV region, both associated with musculature characteristics in chicken's population.

The *SAP30* gene is an important component of histone deacetylase complex in eukaryotic organisms (Zhang et al., 1998) and it has already been identified as a factor of transcriptional regulation in the gene expression in mouse skeletal muscle (Magnusson et al., 2005). This gene is related to the Sin3 proteins, acting as a binding protein important to repression and co-repression of gene transcription (Zhang and Iratni, 1997). *SAP30* gene has already been found to belong to a conserved region between chickens and mouse when comparing the chromosome organization (Viiri et al., 2009). The ancestral protein SAP30 was found to be conserved in many species of chloroplasts, fungi, plants, and animals, also an important conserved region called *GALNT-SAP30-HAND* according to the chromosome position between chicken and mouse (Viiri et al., 2009). Conservation across species is important because it allows studies to be carried out across different species that are expected to recapitulate the same effects. Thus, well developed experimental tools such as cell lines in the mouse could be reasonably used to evaluate the function of a gene in the chicken or other livestock species.

To explore gene function, the small interfering RNAs (siRNA) are a versatile tool, particularly when used in cell lines that can be phenotype across multiple experimental replicates in the lab. This technique has been widely used to inhibit mRNA transcription using a small sequence of RNA with mismatches in the target gene sequence under investigation (Bartel, 2004), in this way, siRNAs can silence a gene, or simply reduce its expression (Morris et al., 2004), depending on the degree of sequence similarity to the target sequence. To understand the potential functional roles of mutations or genes in phenotype variation, functional experiments such as gene silencing or knockdown experiments must be conducted to eventually identify causal genes.

Given the economic importance of the genomic region containing QTL and CNVR to produce breast meat in chickens and the fact that validation studies of gene function in cell culture are scarce, the objective of this study was to select a candidate gene for muscle development and to validate its function in cell culture system (*in vitro*). Since the *SAP30* gene is well conserved across mouse and chicken, in *vitro* morphometric and gene expression analyses were performed on mouse C2C12 myoblast cells.

2.3 Methods

2. Population and identification of candidate gene

The candidate gene selection was performed mining QTLs previously detected in a meat-type chicken population, called TT Reference Population. This is an experimental broiler population developed by the EMBRAPA Swine and Poultry National Research Center, for genomics studies. The population was generated from the expansion of a paternal pure line (TT) and 20 males were mated with 92 females, generating approximately 1,500 chickens (Marchesi et al., 2018). In this population, QTLs were identified for breast muscle and yield. More specifically, one QTL on GGA4 (43Mb) explained 0.61% of the genetic variance for breast muscle yield (Trevisoli et al., 2021), which overlapped with an inherited CNVR (GGA4: 43,406,643 - 44,413,344) associated with breast muscle weight and yield (unpublished data). These two previous studies show that this region may be participating in muscle development in chickens. Therefore, we used this information to search for the genes annotated in these regions.

Searches for deleterious and high-impact mutations that could change the amino acids sequence were carried out using the Variant Effect Predictor (VEP, McLaren et al., 2016) but no gene showed relevant mutations. Then, the next analysis for choosing the candidate gene was based on gene ontology terms for biological process, that could support the potential role of the candidate gene in muscle growth and/or development.

2.3.2 C2C12 cell culture and transfection

The cell culture and transfection were performed to characterize the *SAP30* knockdown in the muscle cells. The C2C12 myoblasts (ATCC® CRL 1772™) were plated in a density of $2x10^4$ cell/well in growth medium [GM, Dulbecco's modified Eagle's medium (DMEM) with high glucose and L -glutamine (Gibco/ThermoFisher; Waltham, MA, USA)], supplemented with 10% FBS (Fetal Bovine Serum, Gibco) and 1% penicillin/streptomycin/amphotericin B solution (Gibco/ThermoFisher; Waltham, MA, USA), in 24-well plates coated with 3% gelatin (Sigma-Aldrich). Transfection was performed when cells reached approximately 70% confluence, using a custom *SAP30* siRNA (ThermoFisher; Waltham, MA, USA) designed five target transcripts (NM_021788AF075136.1, AK010928.1, AK088745.1, BC132087.1, BC132081.1) for the treated group (knocked down cells). The Silencer negative control pre-designed siRNA (ThermoFisher; Waltham, MA, USA) was used as control group. The transfections were performed using Lipofectamine 3000 (ThermoFisher; Waltham, MA, USA) following the manufacture recommendations. After 48h at 37°C and 5% at CO_2, the myogenic differentiation was induced by culturing cells at low serum conditions [DM: DMEM high-glucose, supplemented with 2% horse serum (Gibco/ThermoFisher;

Waltham, MA, USA) and 1% penicillin/streptomycin/amphotericin B (Gibco/ThermoFisher; Waltham, MA, USA) for 3 days.

2.3.3 MTT assay, immunocytochemistry, fusion index, and morphometric analysis

Cell viability was checked based on mitochondrial activity using MTT ((3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) assay. This method is capable of measuring cell viability, by means of colorimetric absorbance, based on mitochondrial cell activity. The amount of formazan blue crystals produced is proportional to the number of viable cells in the culture, thus, active cells produce more formazan crystals and thus more blue color than cells with low metabolic activity (Mosmann, 1983). Cell viability index was determined after 3 days in GM. The absorbance was determined at 595 nm, in triplicate. Immunocytochemistry was performed to allow measuring myotube area only in differentiating cells. For that, the C2C12 cells were fixed in 4% PFA, for 15 min, permeabilized with 0.1% Tween 20, blocked with 1% BSA and incubated with mouse anti-myosin antibody (MF20 monoclonal; Hybridoma Bank, diluted 1:200), at 4°C overnight. After washing, cells were incubated with secondary antibody peroxidase-conjugated (EnVision® Systems Dual Link, anti-mouse and anti-rabbit, Agilent Technologies, Santa Clara, CA, USA) and signals were detected using DAB (Dako, Santa Clara, CA, USA) and the nuclei were counterstained with hematoxylin. Immunocytochemistry images were obtained using an inverted optical microscopy with capture system (BX41 Olympus and MoticAE31). Myotubes' length and width were obtained using ImageJ software. A stage micrometer was used for standardized analysis. Myotube area was determined by length x width and fusion index was determined following Martins et al. (2014) and Costa et al. (2021).

2.3.4 RNA isolation, library preparation and sequencing

Total RNA was isolated from each cell group (treated and control) in triplicates, using Trizol (ThermoFisher; Waltham, MA, USA) reagent following the manufacture's recommendations. RNA integrity of the samples was determined using the Bionalyzer (Agilent Technologies, Santa Clara, CA, USA) and samples with RNA Integrity Number (RIN) score ≥ 7 were selected for library preparation. Libraries were prepared using kit Illumina TruSeq Stranded mRNA Sample Prep LT Protocol (Illumina, San Diego, CA, USA) and the sequencing was performed using the HiSeq SBS v4 kit, to generate 100 bp (2x101) paired end sequence reads following the Illumina protocol. Samples were sequenced using HiSeq 2500 ultra-high throughput systems technology (Illumina, San Diego, CA, USA). The RNA sequencing was performed at the Genomic Center at ESALQ/USP (Piracicaba, São Paulo, Brazil).

2.3.5 RNA sequencing bioinformatics: quality control and read mapping

Sequence quality control (QC) was evaluated using FastQC software (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/, version 0.11.9). Library adaptors and low quality sequences were trimmed away with TRIMgalore software

(https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/, version 0.6.5), which removed the reads with Phred quality score lower than 20. Sequences reads were aligned against *Mus musculus* genome assembly (GRCh38.p6) was performed with STAR software (Dobin et al., 2013). Gene counts were performed using FeatureCounts software (Liao et al., 2014, release 2.0.1) and the reads were normalized using TPM method with RSEM software (Li and Dewey, 2011). All subsequent statistical analysis were performed using R environment (R Team Core, 2019).

2.3.6 Statistical Analysis and Differentially expressed genes

Differentially expressed (DE) genes were identified comparing control to *SAP30* knockdown cells using DESeq2 software in R (Love et al., 2014). Reads were filtered as follows to exclude: i. genes with no expression, i.e., gene annotations with zero reads, ii. Genes that may have aligned sequence due to mapping error, i.e., genes very lowly expressed with less than 1 read per sample on average, iii. genes rarely expressed across samples, i.e., genes with read counts that were not present in at least three (25% of all samples). To remove potentially batch effects, we used the Surrogate Variable Analysis (SVA) package (Leek et al., 2012). After these filters, a total number of 19,489 genes remained for DE analysis. The DE analysis included variables from the SVA analysis to remove batch effects as well as the treatment (control or SAP30-knockdown) which was the primary focus in identifying DE genes. Ontology enrichment analysis was performed using MetaCore software (MetaCore, 2021) to help interpret and summarize the cumulative biological process and pathways represented by DE genes.

2.4 Results

2.4.1 Identification of *SAP30* as a candidate gene for muscle development

A previous study from our group identified a QTL on GGA4 (43Mb) associated with breast muscle yield (Trevisoli et al., 2021). In the same region an inherited CNVR was identified (GGA4: 43,406,643 - 44,413,344) associated with breast meat weight and yield (unpublished results). These results provided us with two lines of evidence of the possible importance of this region for muscular development in chicken. We next searched for genes in this region and classified them by Biological Terms (Gene Ontology - GO terms). As showed in Table 1, *SAP30* was the only gene with a biological term related to skeletal muscle cell differentiation and, therefore, was the gene chosen as a candidate for the knockdown study.

2.4.2 *SAP30* gene knockdown promotes muscle cell hypertrophy phenotype

SAP30 role during muscle cell proliferation and differentiation was investigated using the siRNA method. C2C12 cells were transfected with siRNA control or *SAP30* siRNA when cells reached 70% confluence. All analysis were performed after 3 days of cell culture in DM.

Before assessing the effects of *SAP30* knockdown on cell morphology, we evaluated cell viability index via MTT assay after 3 days of cell culture, to ensure that transfections were not toxic to the cells. We could observe no significant differences ($p<0.05$) between control and *SAP30* knocked down cells (Figure 1a), showing that the methodology used was not toxic for the cells and the morphological differences could be taken as exclusively due to the treatment.

Morphometric analysis showed an increase in myotube area of C2C12 cells that were transfected with siRNA for *SAP30* (SAP30-knockdown), compared to the control (Figure 1b). Cell fusion index was also assessed in order to determine if *SAP30* knockdown could also be associated with skeletal muscle hyperplasia. However, our results indicated no significant difference ($p <0.05$) between the control and *SAP30* knocked down (Figure 1c). The myotube area was larger in the SAP30-knockdown group that in compared to the control group (Figure 1d).

2.4.3 DE genes analysis revealed 1,190 genes being differentially expressed between control and treated groups

To investigate the mechanism by which *SAP30* gene was modulating myotube hypertrophy, we performed RNA sequencing and differential expression analysis. Initially, the differences in gene expression between the control and SAP30-knockdown groups were checked using a Principal Component Analysis (PCA) plot. Figure 2 presents a PCA plot showing the triplicates from each group clustering together and it is possible to observe a difference between both groups. The PC1 explain about 55% of the variance between the groups and the PC2 explain about 37% of the variance.

DE gene analysis revealed a total of 1,190 (629 downregulated and 561 upregulated) genes were differentially expressed (p-adjusted<0.05) when the control group was compared with the treated group. Some differentially expressed genes with Log2 fold-change higher then 1.0 for both sides (up and downregulated) were identified (Figure 3). Among downregulated genes we found *Ranbp13* (Ran-binding protein 13), *Elmo1* (Engulfment and cell motility 1), *Catsperg2* (Cation channel sperm-associated protein subunit gamma 2), *Ccl7* (C-C Motif Chemokine Ligand 7), Meox1 (Mesenchyme Homeobox 1), *Rgs2* (Regulator of G Protein Signaling 2), *Sybu* (Syntabulin), *NOS1 (Nitric Oxid Synthase 1), ABAT* (4-Aminobutyrate Aminotransferase), *Sectm1a* (Secreted and transmembrane 1A), *Col14a1* (Collagen Type XIV Alpha 1 Chain), *Vldlr* (Very low-density lipoprotein receptor). We observed a small and not significant reduction in *SAP30* gene expression (Log2-fold change = -0.17, p-adjusted = 0.53).

Among the upregulated genes we found several genes related with muscle and actin process as *Actn2* (Actin alpha 2), *Ablim3* (Actin binding LIM protein family member 3), *Cbfa2t3* (CBFA2/RUNX1 Partner Transcriptional Co-Repressor 3), *Mef2c* (Myocyte enhancer factor 2C), *Myh6* (Myosin heavy chain 6), *Arhgef39* (Rho guanine nucleotide exchange factor 39), *Ablim2* (Actin binding LIM protein family member 2), *Mpp3* (Membrane palmitoylated protein 3), *Myh7* (Myosin heavy chain 7), *Sdcbp2* (Syndecan binding protein 2), *Clcnkb* (Chloride voltage-gated channel kb), *Igf2os* (Insulin-like growth factor 2). These findings suggest the participation of the *SAP30* gene in the hypertrophy phenotype observed, since the reduction in the expression of the *SAP30*, altered the expression of many muscle development markers.

2.4.4 Enrichment analysis reveals several terms related with muscle development

To identify the biological and cellular processes associated with *SAP30*, we performed an enrichment analysis. The top 10 GO terms were related with "muscle system process", "muscle contraction", "muscle structure development", "muscle filament sliding", "actin-myosin filament sliding", "regulation of muscle system process", "muscle organ development", and "striated muscle cell development" (Figure 4).

Figure 5 shows the developmental skeletal muscle network presenting genes that were up and downregulated in our study. Several genes associated with muscle skeletal development were upregulated (red circle) and among them were the transcription factors *MyoG*, *Mef2c*, *TEF-3* and *MEF2D*, and the generic binding proteins *Actin*, *Tropomyosin-2*, *MyHC*, *Desmin*, *HYH4*, *Colagen 5*, *Dystrophin* and *Troponin*. Among the downregulated genes (blue circle) we identified the transcription factor *MYF5*, a receptor ligand *IGF-1*, and the generic binding proteins *Colagen IV*, *Delta-sarcoglycan*, *MYRL2*, *Smooth muscle myosin* and *Tropomyosin-4*, also a generic enzyme called *nNOS* and members of Histone Deacetylase Complex (*HDCA1, HDCA4, HDCA7*).

2.5 Discussion

2.5.1 *SAP30* as a promising candidate gene regulator of muscle hypertrophy

To investigate the function of *SAP30* in muscle development, we performed gene knockdown in mouse myoblast cells (C2C12). We did not observe differences in myoblast proliferation, but morphometric analysis results revealed a potential role of the *SAP30* gene in the muscle cell hypertrophy. The cells that received the *SAP30* gene knockdown treatment developed a larger myotube area when compared with the control group. Myotubes are muscle differentiated cells, multinucleated and formed by the myoblasts' fusion to generate the myofibers and play a central role during the skeletal muscle formation and growth (Chal and Pourquié, 2017). The myoblasts' proliferation and differentiation were already associated with important economical traits in broiler chicken, including meat quality (Nihashi et al., 2019). A study measuring the muscle fibers diameters observed a positive correlation and participation of this phenotype in the breast weight and percentage in chickens (Koomkrong et al., 2015), also the fiber size, number and composition are important traits related with meat quality (Lefaucheur, 2010; Wicks et al., 2019).

Our candidate gene is a Sin3A complex component, and some members of this family were already studied in C2C12 mouse myoblasts revealing essential functions for Sin3 proteins, in different stages of muscle development, where the Sin3A-null embryos for this protein do not survive the early stages of embryogenesis (van Oevelen et al., 2010). Studies with *SAP30* gene have shown its relationship with histone deacetylase process in eukaryotes that induces Sin3, the histone deacetylases HDAC1 and HDAC2 (Zhang et al., 1998), and acting as a transcription factor for *NETO2* gene in kidney cancer (Snezhkina et al., 2018). Besides that, *SAP30* has biological terms related with histone deacetylation, molecular function terms enriched for histone and protein deacetylase activity, cellular component for histone deacetylase complex and histone deacetylase domain. In addition, one of those enriched terms is skeletal muscle cell differentiation leading to a possible relation between this gene and muscle cell development.

RNA-seq analyses reveled the expression of 19,489 genes and among those, 1,190 genes were differentially expressed between control and treated groups. Enrichment analysis of the DE genes revealed terms related to muscle system process, muscle structure development, muscle filament sliding, and actin-myosin filament sliding. The muscle system process term is responsible for muscle adaptation, contraction, and hypertrophy (Bult et al., 2019).

Muscle hypertrophy is an increase in the muscle mass occasioned by many factors, among them, the level of expression of some genes and proteins, activating or inhibiting metabolic pathways (Glass, 2005). In our study we

observed upregulation of many muscle contraction genes as *Actn3*, *DMD*, *Myh3*, *Myh7* and some Tropomyosin (*Tnnt2* and *Tnnt4)*. This observation agrees and helps explain the increase in myofiber diameter observed when *SAP30* was down regulated. Although the *SAP30* was not differentially expressed in the analysis, our results clearly demonstrates that cells treated with siRNA for *SAP30* during the proliferating phase results in upregulation of muscle contraction genes and hypertrophy.

To further explore how *SAP30* could modulate myoblast hypertrophy, it is important to investigate the expression of transcription factors associated with muscle cell determination and differentiation. Transcription factors such as *Myf5, MyoD, Mef2c, MRF4, Myf6* and *Myogenin*, are involved in the myogenic program (Zammit, 2017). The Myocyte enhancer factor 2c (*Mef2c*) is a member of the Mef2 family of transcription factors and cooperates with MyoD family of basic-loop-helix (bHLH) transcription factors to play a role during skeletal muscle development (Molkentin et al., 1995; Potthoff et al., 2007). *Mef2c* knockout mouse specifically in skeletal muscle abnormalities in the myofiber perinatally structure and the adults that survived had some alterations in the skeletal myofiber types (Potthoff et al., 2007). In our work, we observed upregulation of *MefF2c* transcripts in *SAP30* knocked down cells.

In chicken and mouse embryos, activation of the myogenic factor *Myf5* is an important signal for skeletal muscle development in the somites (Ott et al., 1991; Pownall and Emerson, 1992). Other studies reported that downregulation of *Myf5*, induced by *MyoD* gene, results in an increase in expression of Myogenin, promoting cell differentiation (Deato et al., 2008; Liu et al., 2012; Hernández-Hernández et al., 2017). These results corroborate with our study where Myogenin was upregulated and the *Myf5* was downregulated in *SAP30* knocked down C2C12 cells. Also, when we are representing the skeletal muscle development pathway it is possible to observe a positive *Myf5* gene relation with *Desmin* and *Myogenin* (*MyoG*).

During muscle development, *MyoG* gene is expressed after *Myf5* (Francetic and Li, 2011). In studies with mice MyoG-null mice, it was observed that these animals had severe deficiency in the formation of skeletal muscle, where the formation of myoblasts occurred, but the myoblasts fusion to form myotubes does not occur, leading these animals to an early death (Venuti et al., 1995). *Desmin* is a gene responsible to intermediate filament that integrates the Z-disk, sarcomers, nuclear membrane and sarcolemma, being a muscle-specific protein (Brodehl et al., 2018). A study with a Desmin-knockout mice show a rigidity of the musculature, with a large amount of collagen and the expression of genes related to the renewal of the extracellular matrix and inflammatory proteins, suggesting that the imbalance in the expression of this gene can cause cellular injuries and fibrosis (Meyer and Lieber, 2012). Missense mutations in this gene were also associated with cardioskeletal myopathy (Goldfarb et al., 1998). These are important evidence of the *Myogenin* and *Desmin* participation in the phenotype, which corroborates with our findings where both genes were upregulated, when there is cellular muscle hypertrophy.

Not only *Myogenin* and *Desmin*, but many other genes already described to be related to different stages of skeletal muscle development, such as *MyHC* (Agarwal et al., 2020), *Dystrophin* (Stone et al., 2005), *Actin* (Bertola et al., 2008), *Mef2* (Taylor and Hughes, 2017), different Tropomyosin (Gunning et al., 2008) and some collagen proteins (Moore et al., 2005) were found to be upregulated when the *SAP30* was knockdown in skeletal muscle cells.

All together, these findings suggested *SAP30* as a possible regulator of muscle growth and development, since its knocking down induces genes and transcription factors responsible for increasing myotube differentiation and the formation of muscle myofibers. Here, we show important evidence of the participation of the *SAP30* gene in cellular muscle hypertrophy. Possibly, there may be an explanation for the fact that this gene is in a CNV region, however, further studies looking for functional mutations must be carried out to understand the role of the *SAP30* gene in the regulation of the muscle hypertrophy.

2.6 Conclusion

For the first time we showed morphometric and molecular evidence of the possible participation of *SAP30* gene in the muscle hypertrophy regulation. Even though there is a slight change in its expression, many genes and muscle regulators had their expression enhanced while the myotubes were also increased. Therefore, we have shown that *SAP30* gene is a candidate for muscle development and cell hypertrophy regulation.

2.7 References

Agarwal, M., Sharma, A., Kumar, P., Kumar, A., Bharadwaj, A., Saini, M., et al. (2020). Myosin heavy chain-embryonic regulates skeletal muscle differentiation during mammalian development. *Dev.* 147. doi:10.1242/dev.184507.

Animal QTL Database Available at: https://www.animalgenome.org/cgi-bin/QTLdb/index [Accessed May 11, 2021].

Bartel, D. P. (2004). MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell* 116, 281–297. doi:10.1016/S0092-8674(04)00045-5.

Bertola, L. D., Ott, E. B., Griepsma, S., Vonk, F. J., and Bagowski, C. P. (2008). Developmental expression of the alpha-skeletal actin gene. *BMC Evol. Biol.* 8. doi:10.1186/1471-2148-8-166.

Brodehl, A., Gaertner-Rommel, A., and Milting, H. (2018). Molecular insights into cardiomyopathies associated with desmin (DES) mutations. *Biophys. Rev.* 10, 983–1006. doi:10.1007/s12551-018-0429-0.

Bult, C. J., Blake, J. A., Smith, C. L., Kadin, J. A., Richardson, J. E., Anagnostopoulos, A., et al. (2019). Mouse Genome Database (MGD) 2019. *Nucleic Acids Res.* 47, D801–D806. doi:10.1093/nar/gky1056.

Chal, J., and Pourquié, O. (2017). Making muscle: Skeletal myogenesis in vivo and in vitro. *Dev.* 144, 2104–2122. doi:10.1242/dev.151035.

Costa, A. do C., Gonçalves Lio Copola, A., Carvalho Souza, C., Meireles Nogueira, J., Aparecida Borges Silva, G., and Cristina Jorge, E. (2021). RGMa can induce skeletal muscle cell hyperplasia via association with neogenin signalling pathway. doi:10.1007/s11626-021-00555-9.

Deato, M. D. E., Marr, M. T., Sottero, T., Inouye, C., Hu, P., and Tjian, R. (2008). MyoD Targets TAF3/TRF3 to Activate Myogenin Transcription. *Mol. Cell* 32, 96–105. doi:10.1016/j.molcel.2008.09.009.

Dekkers, J. (2012). Application of Genomics Tools to Animal Breeding. *Curr. Genomics* 13, 207–212. doi:10.2174/138920212800543057.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi:10.1093/bioinformatics/bts635.

Fernandes, A. C., da Silva, V. H., Goes, C. P., Moreira, G. C. M., Godoy, T. F., Ibelli, A. M. G., et al. (2021). Genome-wide detection of CNVs and their association with performance traits in broilers. BMC Genomics 22, 354. doi:10.1186/s12864-021-07676-1.

Fornari, M. B., Zanella, R., Ibelli, A. M., Fernandes, L. T., Cantão, M. E., Thomaz-Soccol, V., et al. (2014). Unraveling the associations of osteoprotegerin gene with production traits in a paternal broiler line. doi:10.1186/2193-1801-3-682.

Francetic, T., and Li, Q. (2011). Skeletal myogenesis and Myf5 activation. *Transcription* 2, 109–114. doi:10.4161/trns.2.3.15829.

Glass, D. J. (2005). Skeletal muscle hypertrophy and atrophy signaling pathways. *Int. J. Biochem. Cell Biol.* 37, 1974–1984. doi:10.1016/j.biocel.2005.04.018.

Goldfarb, L. G., Park, K. Y., Cervenákova, L., Gorokhova, S., Lee, H. S., Vasconcelos, O., et al. (1998). Missense mutations in desmin associated with familial cardiac and skeletal myopathy. *Nat. Genet.* 19, 402–403. doi:10.1038/1300.

Grupioni, N., Stafuzza, N., Carvajal, A., Ibelli, A., Peixoto, J., Ledur, M., et al. (2017). Association of RUNX2 and TNFSF11 genes with production traits in a paternal broiler line. *Genet. Mol. Res.* 16, 16019443. doi:10.4238/gmr16019443.

Gunning, P., O'Neill, G., and Hardeman, E. (2008). Tropomyosin-based regulation of the actin cytoskeleton in time and space. *Physiol. Rev.* 88, 1–35. doi:10.1152/physrev.00001.2007.

Hansen, C., Yi, N., Zhang, Y. M., Xu, S., Gavora, J., and Cheng, H. H. (2005). Identification of QTL for production traits in chickens. *Anim. Biotechnol.* 16, 67–79. doi:10.1081/ABIO-200055016.

Hernández-Hernández, J. M., García-González, E. G., Brun, C. E., and Rudnicki, M. A. (2017). The myogenic regulatory factors, determinants of muscle development, cell identity and regeneration. *Semin. Cell Dev. Biol.* 72, 10–18. doi:10.1016/j.semcdb.2017.11.010.

Hu, Z. L., Park, C. A., and Reecy, J. M. (2019). Building a livestock genetic and genomic information knowledgebase through integrative developments of Animal QTLdb and CorrDB. *Nucleic Acids Res.* 47, D701–D710. doi:10.1093/nar/gky1084.

Julian, R. J. (2005). Production and growth related disorders and other metabolic diseases of poultry - A review. *Vet. J.* 169, 350–369. doi:10.1016/j.tvjl.2004.04.015.

Koomkrong, N., Theerawatanasirikul, S., Boonkaewwan, C., Jaturasitha, S., and Kayan, A. (2015). Breed-related number and size of muscle fibres and their response to carcass quality in chickens. *Ital. J. Anim. Sci.* 14, 638–642. doi:10.4081/ijas.2015.4145.

Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., and Storey, J. D. (2012). The SVA package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 28, 882–883. doi:10.1093/bioinformatics/bts034.

Lefaucheur, L. (2010). A second look into fibre typing - Relation to meat quality. *Meat Sci.* 84, 257–270. doi:10.1016/j.meatsci.2009.05.004.

Li, B., and Dewey, C. N. (2011). RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12, 1–16. doi:10.1186/1471-2105-12-323.

Liao, Y., Smyth, G. K., and Shi, W. (2014). Sequence analysis featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. 30, 923–930. doi:10.1093/bioinformatics/btt656.

Liu, Q. C., Zha, X. H., Faralli, H., Yin, H., Louis-Jeune, C., Perdiguero, E., et al. (2012). Comparative expression profiling identifies differential roles for Myogenin and p38α MAPK signaling in myogenesis. *J. Mol. Cell Biol.* 4, 386–397. doi:10.1093/jmcb/mjs045.

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15. doi:10.1186/s13059-014-0550-8.

Magnusson, C., Svensson, A., Christerson, U., and Tågerud, S. (2005). Denervation-induced alterations in gene expression in mouse skeletal muscle. *Eur. J. Neurosci.* 21, 577–580. doi:10.1111/j.1460-9568.2005.03855.x.

Maharjan, P., Hilton, K., Weil, J., Suesuttajit, N., Beitia, A., Owens, C. M., et al. (2020). Characterizing Woody Breast Myopathy in a Meat Broiler Line by Heat Production, Microbiota, and Plasma Metabolites. *Front. Vet. Sci.* 6. doi:10.3389/fvets.2019.00497.

Marchesi, J. A. P., Buzanskas, M. E., Cantão, M. E., Ibelli, A. M. G., Peixoto, J. O., Joaquim, L. B., et al. (2018). Relationship of runs of homozygosity with adaptive and production traits in a paternal broiler line. *Animal* 12, 1126–1134. doi:10.1017/S1751731117002671.

Martins, AF. José Xavier Neto, Ana Azambuja, Maria Lorena Sereno, Antonio Figueira, Paulo Henrique Campos-Junior, et al. (2014). Repulsive Guidance Molecules a, b and c Are Skeletal Muscle Proteins, and Repulsive Guidance Molecule a Promotes Cellular Hypertrophy and Is Necessary for Myotube Fusion. *RGMb RGMc Cells Tissues Organs* 200, 326–338. doi:10.1159/000433491.

McLaren, W.; Gil, L.; Hunt, S.E.; Riat, H.S.; Ritchie, G. R.S.; Thormann, A.; Flicek, P.; Cunningham, F. The Ensembl Variant Effect Predictor. Genome Biology, 2016.

MetaCore MetaCore Login | Clarivate. Available at: https://portal.genego.com/ [Accessed May 11, 2021].

Meyer, G. A., and Lieber, R. L. (2012). Skeletal muscle fibrosis develops in response to desmin deletion. *Am. J. Physiol. - Cell Physiol.* 302, 1609–1620. doi:10.1152/ajpcell.00441.2011.

Molkentin, J. D., Black, B. L., Martin, J. F., and Olson, E. N. (1995). Cooperative activation of muscle gene expression by MEF2 and myogenic bHLH proteins. *Cell* 83, 1125–1136. doi:10.1016/0092-8674(95)90139-6.

Moore, D. R., Phillips, S. M., Babraj, J. A., Smith, K., and Rennie, M. J. (2005). Myofibrillar and collagen protein synthesis in human skeletal muscle in young men after maximal shortening and lengthening contractions. *Am. J. Physiol. - Endocrinol. Metab.* 288. doi:10.1152/ajpendo.00387.2004.

Moreira, GCM., Boschiero, C., Mello Cesar, A. S., Reecy, J. M., Godoy, T. F., et al. A genome-wide association study reveals novel genomic regions and positional candidate genes for fat deposition in broiler chickens. doi:10.1186/s12864-018-4779-6.

Moreira, G. C. M., Poleti, M. D., Pértille, F., Boschiero, C., Cesar, A. S. M., Godoy, T. F., et al. (2019). Unraveling genomic associations with feed efficiency and body weight traits in chickens through an integrative approach. *BMC Genet.* 20. doi:10.1186/s12863-019-0783-3.

Morris, K. V., Chan, S. W. L., Jacobsen, S. E., and Looney, D. J. (2004). Small interfering RNA-induced transcriptional gene silencing in human cells. *Science (80-.).* 305, 1289–1292. doi:10.1126/science.1101372.

Mosmann, T. (1983). Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays. *J. Immunol. Methods* 65, 55–63. doi:10.1016/0022-1759(83)90303-4.

Nihashi, Y., Umezawa, K., Shinji, S., Hamaguchi, Y., Kobayashi, H., Kono, T., et al. (2019). Distinct cell proliferation, myogenic differentiation, and gene expression in skeletal muscle myoblasts of layer and broiler chickens. *Sci. Rep.* 9. doi:10.1038/s41598-019-52946-4.

Olkowski, A. A. (2007). Pathophysiology of heart failure in broiler chickens: Structural, biochemical, and molecular characteristics. in *Poultry Science* (Poultry Science Association), 999–1005. doi:10.1093/ps/86.5.999.

Ott, M.-O., Bober, E., Lyons, G., Arnold, H., and Buckingham, M. (1991). Early expression of the myogenic regulatory gene, myf-5, in precursor cells of skeletal muscle in the mouse embryo.

Petry, B.P.; Moreira, G.C.M.; Copola, A.L.; Souza, M.M.; Veiga, F.C.; Jorge, E.C.; Peixoto, J.O.; Ledur, M.C.; Koltes,

J.E.; Coutinho, L.L. SAP30 gene is a probable regulator of muscle hypertrophy in chickens. Frontiers in Genetics (2021).

Potthoff, M. J., Arnold, M. A., McAnally, J., Richardson, J. A., Bassel-Duby, R., and Olson, E. N. (2007). Regulation of Skeletal Muscle Sarcomere Integrity and Postnatal Muscle Function by Mef2c. *Mol. Cell. Biol.* 27, 8143–8151. doi:10.1128/mcb.01187-07.

Pownall, M. E., and Emerson, C. P. (1992). Sequential activation of three myogenic regulatory genes during somite morphogenesis in quail embryos. *Dev. Biol.* 151, 67–79. doi:10.1016/0012-1606(92)90214-2.

R Core Team (2019). R: The R Project for Statistical Computing. Available at: https://www.r-project.org/ [Accessed May 11, 2021].

Saxena, V. K., and Kolluri, G. (2018). "Selection Methods in Poultry Breeding: From Genetics to Genomics," in *Application of Genetics and Genomics in Poultry Science* (InTech). doi:10.5772/intechopen.77966.

Snezhkina, A. V., Nyushko, K. M., Zaretsky, A. R., Shagin, D. A., Sadritdinova, A. F., Fedorova, M. S., et al. (2018). Transcription Factor SAP30 Is Involved in the Activation of NETO2 Gene Expression in Clear Cell Renal Cell Carcinoma. *Mol. Biol.* 52, 385–392. doi:10.1134/S0026893318020152.

Soglia, F., Mazzoni, M., and Petracci, M. (2019). Spotlight on avian pathology: current growth-related breast meat abnormalities in broilers. *Avian Pathol.* 48, 1–3. doi:10.1080/03079457.2018.1508821.

Stone, M. R., O'Neill, A., Catino, D., and Bloch, R. J. (2005). Specific interaction of the actin-binding domain of dystrophin with intermediate filaments containing keratin 19. *Mol. Biol. Cell* 16, 4280–4293. doi:10.1091/mbc.E05-02-0112.

Taylor, M. V., and Hughes, S. M. (2017). Mef2 and the skeletal muscle differentiation program. *Semin. Cell Dev. Biol.* 72, 33–44. doi:10.1016/j.semcdb.2017.11.020.

Thiruvenkadan, A. K., Prabakaran, R., and Panneerselvam, S. (2011). Broiler breeding strategies over the decades: An overview. *Worlds. Poult. Sci. J.* 67, 309–336. doi:10.1017/S0043933911000328.

Trevisoli, P.A., Moreira, G.C.M., C, Cesar, A.S.M., Petrini, J., Ledur, M.C., Mourão, G.B., Garrick, D. and Coutinho, L.L. (2021). A Missense Mutation in the MYBPH Gene Is Associated With Abdominal Fat Traits in Meat-Type Chickens. Front. Genet. 12:698163. doi: 10.3389/fgene.2021.698163

van Oevelen, C., Bowman, C., Pellegrino, J., Asp, P., Cheng, J., Parisi, F., et al. (2010). The Mammalian Sin3 Proteins Are Required for Muscle Development and Sarcomere Specification. *Mol. Cell. Biol.* 30, 5686–5697. doi:10.1128/mcb.00975-10.

Venturini, G. C., Stafuzza, N. B., Cardoso, F. F., Baldi, F., Ledur, M. C., Peixoto, J. O., et al. (2015). Association between ACTA1 candidate gene and performance, organs and carcass traits in broilers | Elsevier Enhanced Reader. Available at: https://reader.elsevier.com/reader/sd/pii/S0032579119321170?token=F90F263A154EFDB88E538DF7307CE488C5B21FB4C9C52C90B485BCA7D9034708D86348F6CA23C2EF4646E15747BFDC70&originRegion=us-east-1&originCreation=20210513125514 [Accessed May 13, 2021].

Venuti, J. M., Morris, J. H., Vivian, J. L., Olson, E. N., and Klein, W. H. (1995). Myogenin is required for late but not early aspects of myogenesis during mouse development. *J. Cell Biol.* 128, 563–576. doi:10.1083/jcb.128.4.563.

Viiri, K. M., Heinonen, T. Y., Mäki, M., and Lohi, O. (2009). Phylogenetic analysis of the SAP30 family of transcriptional regulators reveals functional divergence in the domain that binds the nuclear matrix. doi:10.1186/1471-2148-9-149.

Wicks, J., Beline, M., Gomez, J. F. M., Luzardo, S., Silva, S. L., and Gerrard, D. (2019). Muscle energy metabolism, growth, and meat quality in beef cattle. *Agric.* 9, 195. doi:10.3390/agriculture9090195.

Zammit, P. S. (2017). Function of the myogenic regulatory factors Myf5, MyoD, Myogenin and MRF4 in skeletal muscle, satellite cells and regenerative myogenesis. *Semin. Cell Dev. Biol.* 72, 19–32. doi:10.1016/j.semcdb.2017.11.011.

Zhang, Y., and Iratni, R. (1997). Histone Deacetylases and SAP18, a Novel Polypeptide, Are Components of a Human Sin3 Complex.

Zhang, Y., Sun, Z. W., Iratni, R., Erdjument-Bromage, H., Tempst, P., Hampsey, M., et al. (1998). SAP30, a novel protein conserved between human and yeast, is a component of a histone deacetylase complex. *Mol. Cell* 1, 1021–1031. doi:10.1016/S1097-2765(00)80102-1.

Zuidhof, M. J., Schneider, B. L., Carney, V. L., Korver, D. R., and Robinson, F. E. (2014). Growth, efficiency, and yield of commercial broilers from 1957, 1978, and 20051. *Poult. Sci.* 93, 2970–2982. doi:10.3382/ps.2014-04291.

Table 1. List of genes found within Quantitative Trait Loci (QTL) for breast yield in GWAS by Trevisoli et al. (2021) within the TT Reference Population at EMBRAPA. The list shows the genes found on chromosome 4 (43Mb, 43001322 - 43998491) from the *Gallus gallus (Gallus_gallus*5.0, NCBI) genome build.

Gene ID	Gene name	GO term name
ENSGALG00000010764	*FBXO8*	ARF guanyl-nucleotide exchange factor activity
ENSGALG00000020201	*CEP44*	centrosome. spindle pole
ENSGALG00000010744	*GALNT7*	transferase activity. Membrane. Golgi apparatus. Golgi membrane
ENSGALG00000010745	*HMGB2*	regulation of stem cell proliferation
ENSGALG00000036252	***SAP30***	**skeletal muscle cell differentiation**
ENSGALG00000010755	*SCRG1*	mesenchymal stem cell proliferation
ENSGALG00000030530	*HAND2*	positive regulation of cardiac muscle hypertrophy
ENSGALG00000025263	*gga-mir-1776*	
ENSGALG00000032355	*novel gene*	golgi apparatus
ENSGALG00000030588	*GALNTL6*	carbohydrate binding and polypeptide N-acetylgalactosaminyltransferase activity
ENSGALG00000033732	*novel gene*	
ENSGALG00000037279	*novel gene*	
ENSGALG00000032195	*novel gene*	
ENSGALG00000033311	*novel gene*	

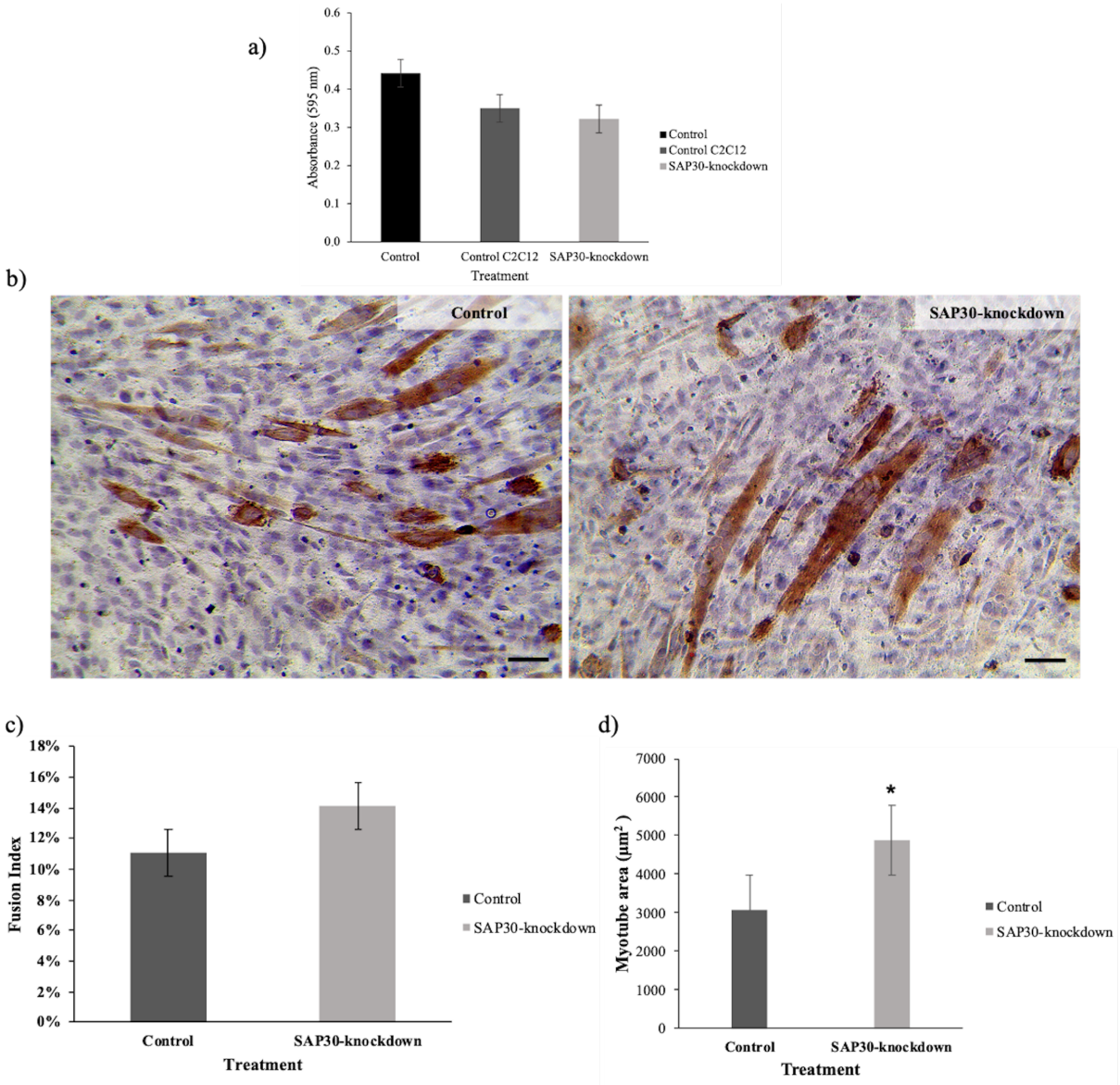


Figure 1. Cell viability index and morphological effects of C2C12 cells transfected with *SAP30* siRNA and control. a) Cell viability index determined by the MTT assay performed in C2C12 cells after 3 days of transfection with *SAP30* siRNA or siRNA control. b) Representative image of MF20 immunocytochemistry result in C2C12 culture after 3 days of transfection with siRNA control (A) and SAP30-knockdown (B). Scale bar = 100µm. c) Fusion index of MF20-positive cells transfected with siRNA control and *SAP30* siRNA after 3 days in DM. d) Myotube area of C2C12 cells transfected with the siRNA control and SAP30-knockdown. Significant values *$p<0.05$.

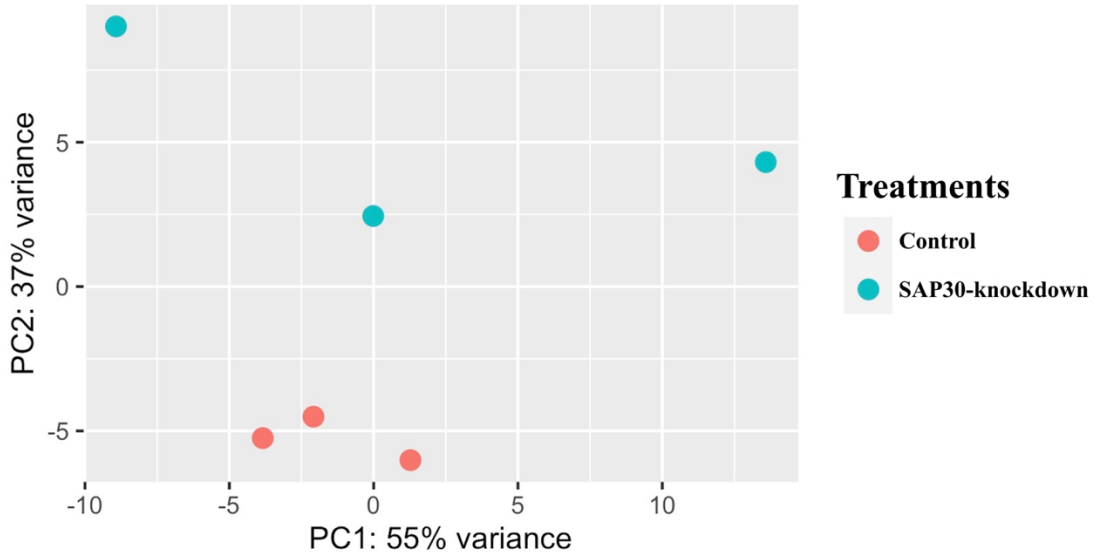


Figure 2. Principal Component Analysis (PCA) clearly showing a variation between genes expressed in the two groups (SAP30-knockdown and control) in C2C12 cells. The red dots are representing the triplicates from control group, while the blue dots are triplicates from SAP30-knockdown group.

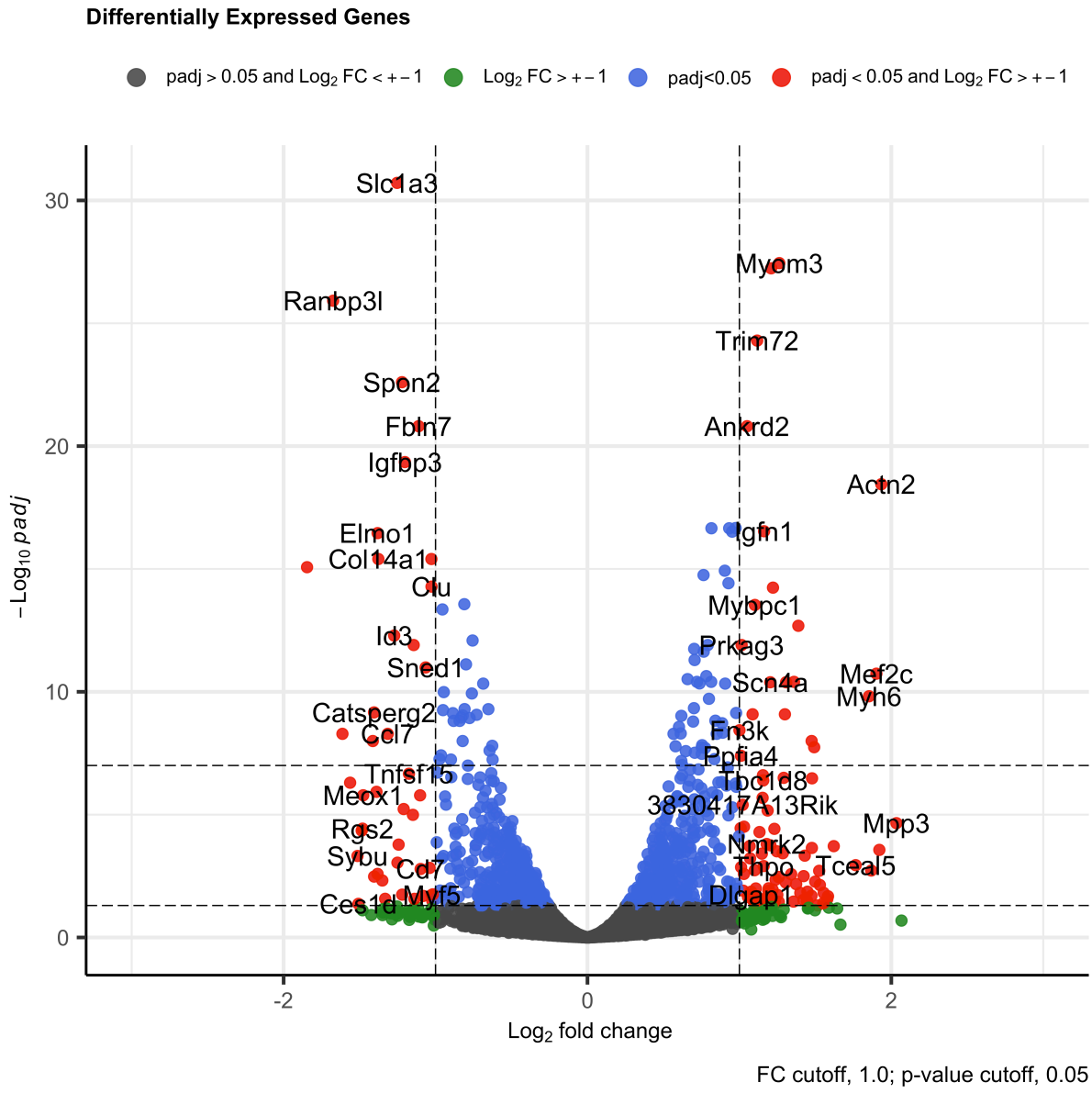


Figure 3. Volcano plot representing all differentially expressed genes with Log2Folg Change bigger then 1.0 and significant ($p<0.05$) in the differentially expressed (DE) gene analysis in C2C12 cells after 3 days in differentiation medium. The red dots are the DE genes with Log2Fold Change higher than 1.0 in both sides and with adjusted p-value < 0.05.

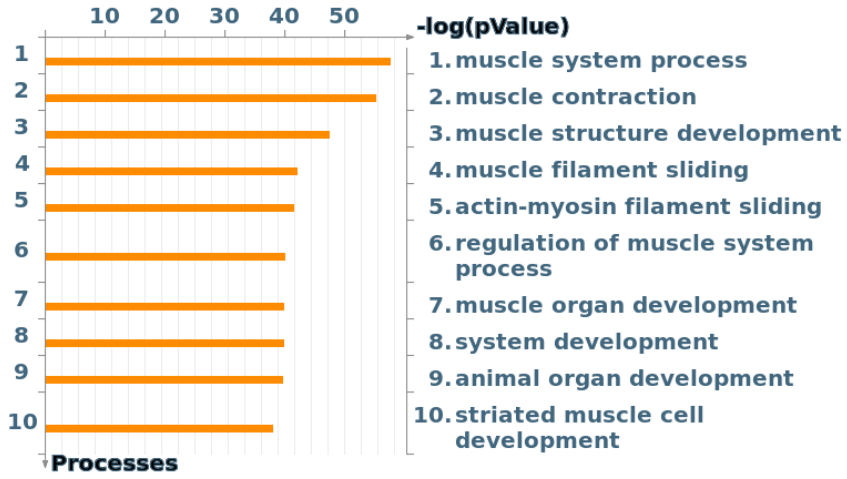


Figure 4. Gene Ontology terms showing the top 10 significantly ($p<0.05$) processes that the DE genes are related. Image created by MetaCore software.

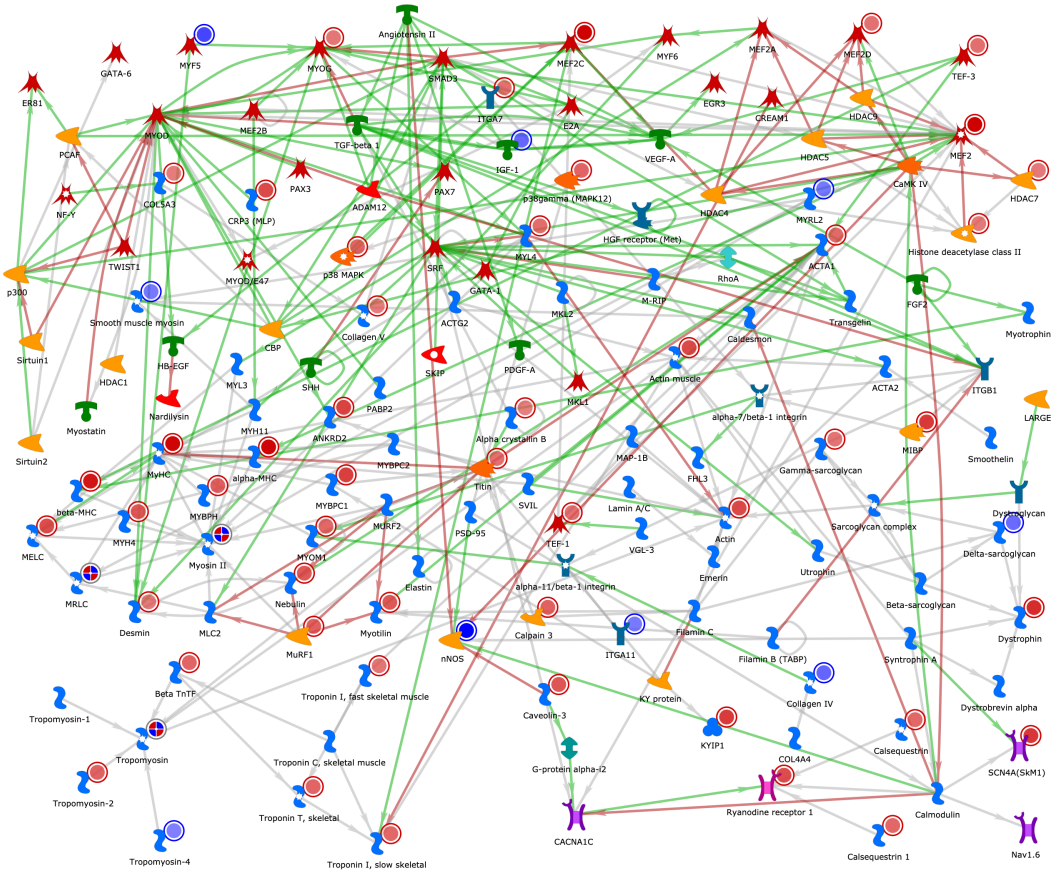


Figure 5. Developmental skeletal muscle network representing genes related with this process, based on the differentially expressed (DE) genes analysis. The genes with red circles are genes that were upregulated in the SAP30-knockdown group after the DE analysis. Genes marked with blue circle were downregulated in the DE gene analysis. The red lines between the objects means a positive/activation effects, the green lines a negative/inhibition effects and the grey lines an unspecified effect. Image created by MetaCore software.

3. UNRAVELING THE *SAP30* GENE FUNCTION IN MUSCLE CELLS TRHOUGH CRISPR/CAS9

3.1 Abstract

Understanding the molecular mechanisms involved in determining muscle development can be an interesting strategy for identifying genes and mutations involved in muscle development. Here, we used the *SAP30* gene as a candidate for muscle hypertrophy and tested the effect of mutations caused by CRISPR/Cas9 on the C2C12 cell phenotype. As a result, three different types of mutations were induced in the clones treated with the CRISPR/Cas9 system to edit the *SAP30* gene. All three clones had a hypertrophy phenotype when compared to the control clone. RNA-seq analyzes revealed several genes being up and downregulated in the mutated clones. Functional enrichment analyzes revealed the participation of the *SAP30*, regulating the expression of several genes, which have been reported to be involved in muscle development. In addition, genes that participate in muscle growth and regeneration pathways, such as NOTCH and WNT, had their expression altered when there were mutations in the *SAP30* gene. This result shows the possible participation of *SAP30* in the regulation, growth, regeneration, and hypertrophy of muscle cells.

3.2 Introduction

Understanding the molecular processes involved in muscle development is important to identify which genes are involved in the characterization of the economic traits. Studies characterizing genes and candidate mutations in the genome of farm animals have been carried out using different tools to predict the variants participation in the phenotype. However, studies characterizing the function of these genes and candidate mutations using a cell culture system - *in vitro*, are still scarce.

The use of genetic engineering tools has been contributed to understand which genes are involved with the molecular mechanisms in different traits. Gene edition, mediated by Clustered Regulatory Interspaced Palindromic Repeats associated with a Cas9 enzyme (CRISPR/Cas9) (Doudna and Charpentier, 2014) is a powerful methodology that can be used to study gene function. The CRISPR system was first described in 1987 as a series of short repeats interspaced with short sequences in the *E. coli* genome (Ishino et al., 1987). It has now been harnessed as a gene editing technique and has been used to regulate, modify, or mark genomic loci in the diverse organisms, allowing manipulation of DNA sequences according to the interest of the researchers (Doudna and Charpentier, 2014).

In animal breeding, this technique was used to study the importance of genes during the embryogenic process in chickens, where the function of the *Stra8* gene, responsible to differentiate embryogenic stem cells in spermatogenic stem cells, was characterized (Zhang et al., 2017). Also, the CRISPR/Cas9 was used to delete a single amino acid in the *NHE1* gene, conferring resistance of chickens to an important pathogen called avian leukosis virus subgroup J (Koslová et al., 2020). These positive results indicate that CRISPR/Cas9 can be used in the editing of other genes to characterize these functions. Since most of the characteristics of interest are controlled by more than one gene (polygenic characteristics), studying the knockout of an individual gene allows for precise characterization of that gene's participation in the regulation of the phenotype.

Studies using the most varied cellular models, used the technique of RNA sequencing to validate and quantify changes in gene expression in cell culture (Doynova et al., 2017; Sorelle et al., 2021). Here, the gene knockout through CRISPR/Cas9 technique was used to investigate the function of a single gene, followed by RNA sequencing of these cells after treatment to see how gene pathways were impacted by the reduction or even exclusion of the candidate gene. RNA sequencing methods provide a way to analyze the differential expression of genes, and gene networks in response to gene perturbation to characterize the molecular pathways involved in the processes of muscle development and differentiation.

As mentioned in Petry et al. (2021), *SAP30* was selected as a candidate gene because it was associated with muscle traits in a Brazilian broiler population (TT Reference population) by GWAS analysis. This gene is located on GGA4 (*Gallus gallus*) and was selected as a candidate gene for two reasons: it is in a QTL window (43Mb) associated with breast percentage (Trevisoli et al., 2021) and it is overlapping with a CNV region associated with breast meat weight and percentage (unpublished data). The *SAP30* (Sin3A-associated protein) gene acts in the process of deacetylation of histones in eukaryotic organisms (Zhang et al., 1998), and is related to Sin3 proteins, which act as a binding protein important in the processes of repression and co-repression of gene transcription (Zhang and Iratni, 1997). Petry et al. (2021) already described this gene as a possible regulator of muscle growth in myotubes' culture (*in vitro),* they observed that cells SAP30-knocked down had muscle hypertrophy phenotype. Thus, the objective of this work was to generate SAP30-mutated cell clones by CRISPR/Cas9 and observe the changes of each mutation in the muscle cell phenotype, improving knowledge about this gene function.

3.3 Methods

3.3.1 Candidate gene: *SAP30*

We initially looked for candidate genes, integrating previous results from our research group. The analyzes were conducted using data from the TT Reference Population (Marchesi et al., 2018). The candidate gene selected to be edited using CRISPR/Cas9 was *SAP30*, which has GO terms related to skeletal muscle cell differentiation. This gene is a component of the histone deacetylase complex and has already been identified as a transcriptional regulatory factor in gene expression in the skeletal muscle of mice (Magnusson et al., 2005). All the process to choose the candidate gene was already described in detail by Petry et al. (2021).

3.3.2 RNA-guide selection, vector construction, cell culture and transfection

To generate the guide RNA capable of recognizing a region of the *SAP30* gene and cleaving the double strand of DNA, we used the MIT CRISPR online design program (Massachusetts Institute of Technology, http://crispr.mit.edu/). After studying the exon, intron, and coding regions of the *SAP30* gene, we choose 250 base pairs in the first exon aiming to silencing the gene, interrupting the translation and consequently, the traduction. The

MIT CRISPR tool provided us the best guides (20 nucleotides), with the best quality score for connection and recognition to the target sequence and the PAM (Polindromic Adjacent Motif) sequence.

The selected guide RNA (sgRNA) with the PAM sequence was: 5' caccgCCGCCGCCTCGGCGGGGAAT GGG 3'and 3' aaacATTCCCCGCCGAGGCGGCGGc 5'. This guide was inserted into plasmid pSpCas9 (BB)-2A-pure, using the sgRNA construction and vector transformation protocol described in Ran et al. (2013). We used the empty plasmid as a negative control.

After the transformation, we performed the plasmidial DNA extraction, through a Miniprep reaction following the kit manufacturer's protocol (Qiagen Plasmid Mini kit, Qiagen). We did a conventional PCR analysis to amplify the region of interest of the *SAP30* gene (data not shown) and check if the cloning of the vector worked. We managed to amplify the fragment of the gene and proceeded to the maxiprep reaction.

To obtain a large amount of plasmid, a maxiprep reaction was performed using a super inoculum following the manufacturer's recommendations (Qiagen Plasmid Maxi kit, Qiagen). After extraction of plasmid DNA, we quantified the vector samples with *SAP30* and the empty vector (negative control) also, to make sure that the sgRNA designed for the *SAP30* gene had linked to the vector, a Sanger sequencing was performed (ABI 3130 Genetic Analyzer, Applied Biosystems).

As soon as we confirmed that the guide sequence that recognizes the *SAP30* gene had been inserted into the vector, we proceeded to the next step: culture of immortalized myoblasts and transfection of the vector in C2C12 cells. The mouse myoblast muscle cells (line C2C12, ATCC® CRL 1772™) were plated in a $2x10^4$ cell/well with growth medium [GM, Dulbecco's modified Eagle's medium (DMEM) with high glucose and L-glutamine (Gibco)], supplemented with 10% BFS (Bovine Fetal Serum, Gibco) and 1% penicillin/streptomycin/amphotericin B solution (Gibco) with 3% gelatin (Sigma-Aldrich)-coated plastic, as mentioned in Petry et al. (2021). When the cells reached approximately 70% confluence, the transfection was performed using Lipofectamine 3000 (ThermoFisher) following the manufacture recommendations. The cells were maintained at 37°C with 5% at CO^2 for 48 hours. After those 48 hours, the myogenic differentiation was induced by culturing cells at low serum conditions DMEM high-glucose and 2% horse serum (Gibco) for 3 days. The mutated and control cells were selected by antibiotic Puromycin (2μg/μL).

3.3.3 Cell phenotype analysis

First, to check if the treatment was not toxic to the cells, analysis based on MTT test were performed. The MTT ((3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) test measure the cell viability based on mitochondrial cell activity, by formazan blue crystals, where cells with greater activity produce a greater amount of formazan crystals (Mosmann, 1983). Also, the MTT tests was used to characterize the cell proliferation. The test was performed with the cells in 3 days of growth medium (10% BFS) and the absorbance was measuring in triplicate for each clone by ELISA equipment (Epson LX-300+II, Biotek) with optical density at 595nm.

The fusion index was calculated by the percentage of the sum of nuclei of all MF20-positive cells by the total number of nuclei in that field after 3 days in the myogenic differentiation medium, as used by Martins et al. (2015) and Costa et al. (2021).

To measure the myotube area and calculate the fusion index, the immunocytochemistry assay was performed. For that, the C2C12 cells were fixed with paraformaldehyde (PFA) 4%, for 15 minutes, permeabilized with 0.1%

Tween 20, blocked with 1% BSA and incubated with mouse anti-myosin antibody, 1:200 (MF20 monoclonal; Hybridoma Bank) at 4°C overnight. After that, the cells were incubated with secondary antibody peroxidase-conjugated (EnVision® Systems Dual Link, anti-mouse and anti-habbit, Dako). The effects of the *SAP30* knockout were evaluated thought morphometric analysis in differentiated cells labeled with Myosin heavy chain (MHC) (MF20-positive cells) and the nuclei were colored by hematoxylin to perform the fusion index analysis. The images used for these analyzes were obtained by inverted optical microscope, coupled to a Moticam 2300 capture system and the morphometric analysis were performed measuring the myotubes' length and width to calculate the area (Martins et al., 2015; Costa et al., 2021).

To check the differences in the myotube area the following model was used for the comparison among control and treated groups:

$$y_{ij} = \mu + t_i + e_{ij}$$

in which: y_{ij} is the mean area for *j*-th cell culture (*j* = 1, 2,3) associated to the treatment *i* (*i* = 1 to 4, CRISPRcontrol, CRISPRmutated_1, CRISPRmutated_2, CRISPRmutated_3), μ is the intercept, t_i is the treatment effect and e_{ij} is the residual associated to the measurement obtained of the *j*-th cell culture from the *i*-th treatment (CRISPRmutated cells). It was assumed $e_{ij} \sim N(0, \sigma_i^2)$, with σ_i^2 as the residual variance associated to each treatment. Fixed effects were estimated by generalized least squares method and variance components obtained through restricted maximum likelihood method. This model was fitted in *gls* function of *nlme* package (Douglas and Pinheiro, 1998) of R statistical software (R Core Team, 2019, version 3.5.3). Being the treatment effect significant based on F test (*p-value* < 0.05), Dunnett's pairwise comparison test was applied to identify the differences among treatments by using *multcomp* R package (Hothorn et al., 2008).

3.3.4 Sanger sequencing of a *SAP30* fragment

To inspect the mutations induced by CRISPR/Cas9 in each clone, the *SAP30* gene fragment was sequenced using Sanger sequencing (ABI 3130 Genetic Analyzer, Applied Biosystems). For that, primers designed to recognize the CRISPR region in the *Mus musculus* genome were used. The primers used for that were, forward: 5' GAACTGGCAGACATGAACGG 3', and reverse: 3' CCTCCCGCAAACAACAGAGT 5'.

3.3.5 Library preparation, RNA sequencing, Quality control and Reads alignment

RNA was extracted from 12 samples, one sample from control cells in triplicate (that received only the empty CRISPR/Cas9 plasmid) and 3 treated clones from CRISPR/Cas9 modified cells, also in triplicate (each one of the 3 cell lines received the plasmid with the *SAP30* guide-RNA). Total RNA from each sample was extracted using Trizol reagent (ThermoFisher, Massachusetts, EUA). RNA integrity of the samples was determined using the Bionalyzer (Agilent Technologies, Santa Clara, CA, USA) and samples with RNA Integrity Number (RIN) score ≥ 9.0 were selected for library preparation. RNA sequencing data were generated by HiSeq2500 ultra-high-throughput sequencing system (Illumina, San Diego, CA, USA). For that, the libraries were prepared using kit Illumina TruSeq Stranded mRNA Sample Prep LT Protocol (Illumina, San Diego, CA, USA) and the sequencing was performed using the HiSeq

SBS v4 kit, to generate 100 bp (2x101) paired end sequence reads following the Illumina protocol. The RNA-seq was performed at the Genomic Center (ESALQ/USP), Piracicaba, SP, Brazil.

The quality control for all the samples was performed using FastQC version 0.11.5 (Andrews, 2010), with a Phred score greater than or equal to 20. The adaptors were removed using TRIMgalore software (Version 0.6.5) (Krueger, 2019) and the alignment and mapping were performed against *Mus musculus* genome assembly (GRCh38.p6) using STAR software (Dobin et al., 2013). FeatureCounts software (Liao et al., 2014, release 2.0.1) was used to perform the gene counts and the reads were normalized using the TPM method with RSEM software (Li and Dewey, 2011).

3.3.6 Differentially expressed genes and Functional enrichment analysis

Statistical analyses to identify DE genes were performed using the DESeq2 software (Love et al., 2014), following the Negative Binomial Distribution from the same software. Enrichment analysis with the DE genes was performed using the Metacore software (MetaCore, 2021).

3.4 Results

3.4.1 Sanger sequencing confirming the *SAP30*-guide-RNA insertion into the pSpCas9 vector

To confirm if the guide was inserted in the pSpCas9, the region containing the U6 promoter in the vector was sequenced by Sanger, and the alignment was performed using Clustal Omega online software (Sievers et al., 2011). Figure 1 shows the results, and the sgRNA for *SAP30* gene was inserted in the region, located between 219bp and 238bp as expected.

```
U6promoter_seq      AGGCATTCCGATACAGGCTGTTAGAGAGATAATTGGAATTAATTTGACTGTAAACACAAA        60
guide_SAP30         ------------------------------------------------------------        0

U6promoter_seq      GATATTAGTACAAAATACGTGACGTAGAAAGTAATAATTTCTTGGGTAGTTTGCAGTTTT        120
guide_SAP30         ------------------------------------------------------------        0

U6promoter_seq      AAAATTATGTTTTAAAATGGACTATCATATGCTTACCGTAACTTGAAAGTATTTCGATTT        180
guide_SAP30         ------------------------------------------------------------        0

U6promoter_seq      CTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGCCGCCGCCTCGGCGGGGAATGT        240
guide_SAP30         ----------------------------------------CCGCCGCCTCGGCGGGGAAT--        20
                                                             ********************

U6promoter_seq      TTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGG        300
guide_SAP30         ------------------------------------------------------------        20

U6promoter_seq      CACCGAGTCGGTGCTTTTTTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTC        360
guide_SAP30         ------------------------------------------------------------        20

U6promoter_seq      CGTTTTTAGCGCGTGCGCCAATTCTGCAGACAAATGGCTCTAGAGGTACCCGTTACATAA        420
guide_SAP30         ------------------------------------------------------------        20

U6promoter_seq      CTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATA        480
guide_SAP30         ------------------------------------------------------------        20

U6promoter_seq      GTAACGCCAATAGGGACTTTCCATTGACGTCAATGG        516
guide_SAP30         ----------------------------------        20
```

Figure 1. Alignment of the U6 promoter region sequences of the pSpCas9 (BB)-2A-pure vector containing the SAP30 sgRNA (U6promoter_seq), and the sgRNA for the gene SAP30 (guide_SAP30). Alignment performed using Clustal Omega online software (Sievers et al., 2011).

3.4.2 Morphometric analysis showed differences between the clones

MTT assay, immunocytochemistry assay and morphometric analysis were performed to check the myotube area, after 3 days in differentiation medium. To check the cell viability and proliferation, the MTT was performed measuring the absorbance of the formazan blue crystals, as mentioned before. For this analysis, there were not significantly differences between CRISPRcontrol and CRISPRmutated clones, meaning that there were not differences in the cell proliferation and, the transfection treatment was not toxic for the cells, so, probably all the results observed are results from the SAP30-knockout.

To measure the myotube size and see the differences between CRISPRmutated cells and CRISPRcontrol, the immunocytochemistry was performed, and the photos obtained were analyzed. Figure 2 shows cells after the immunocytochemistry assay. The fusion index was calculated to check if the cells had hyperplasia phenotype. After the statistical analysis comparing all the CRISPRmutated cells with CRISPRcontrol, there was not significant difference in the fusion index between the treatments (Figure 3).

The morphometric analysis (Figure 4) reveals a difference between CRISPRmutated cells and CRISPRcontrol. All the mutated clones had an increase in the myotube area (hypertrophy), when compared with the CRISPRcontrol.

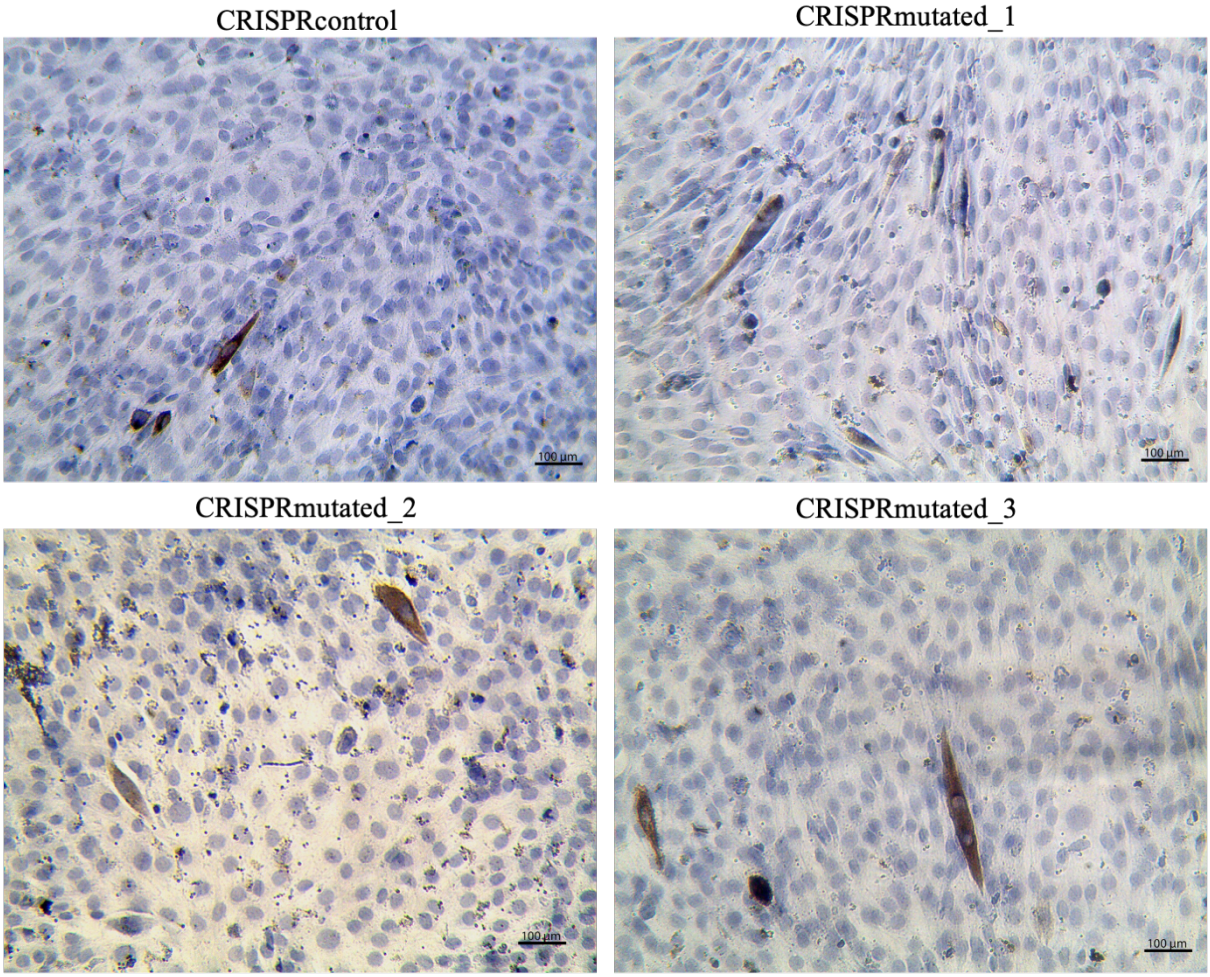


Figure 2. Images of the CRISPR/Cas9 control and treated groups obtained after immunochemistry assay with MF20 marker. Marked cells (MF20-positive) are those that differentiate into myotubes (200x). CRISPRcontrol is the C2C12 cells group that received the empty plasmid. CRISPRmutated (1, 2 and 3) are C2C12 cells that received the SAP30 guide-RNA plasmid. Scale bar = 100 µm.

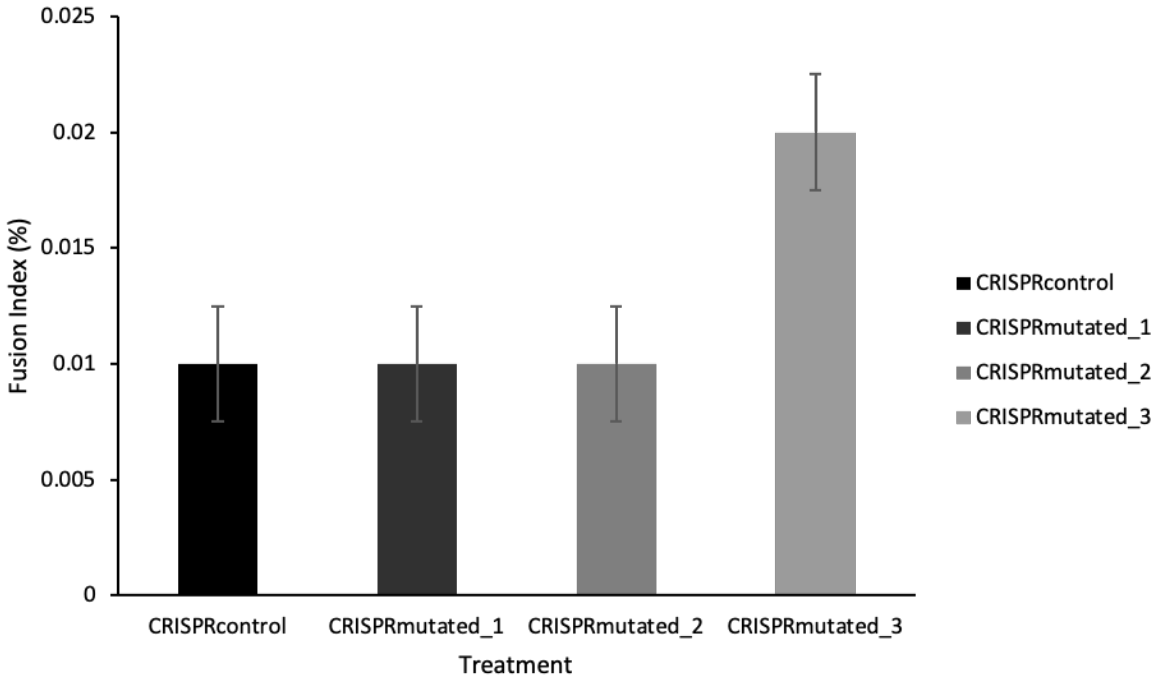


Figure 3: Fusion index results representing all clones in the x-axis and the fusion index percentage % in the y-axis. After statistical test, there was not differences between CRISPRcontrol and CRISPRmutated cells (1, 2 and 3). CRISPRcontrol is the C2C12 cells group that received the empty plasmid. CRISPRmutated (1, 2 and 3) are C2C12 cells that received the SAP30 guide-RNA plasmid.

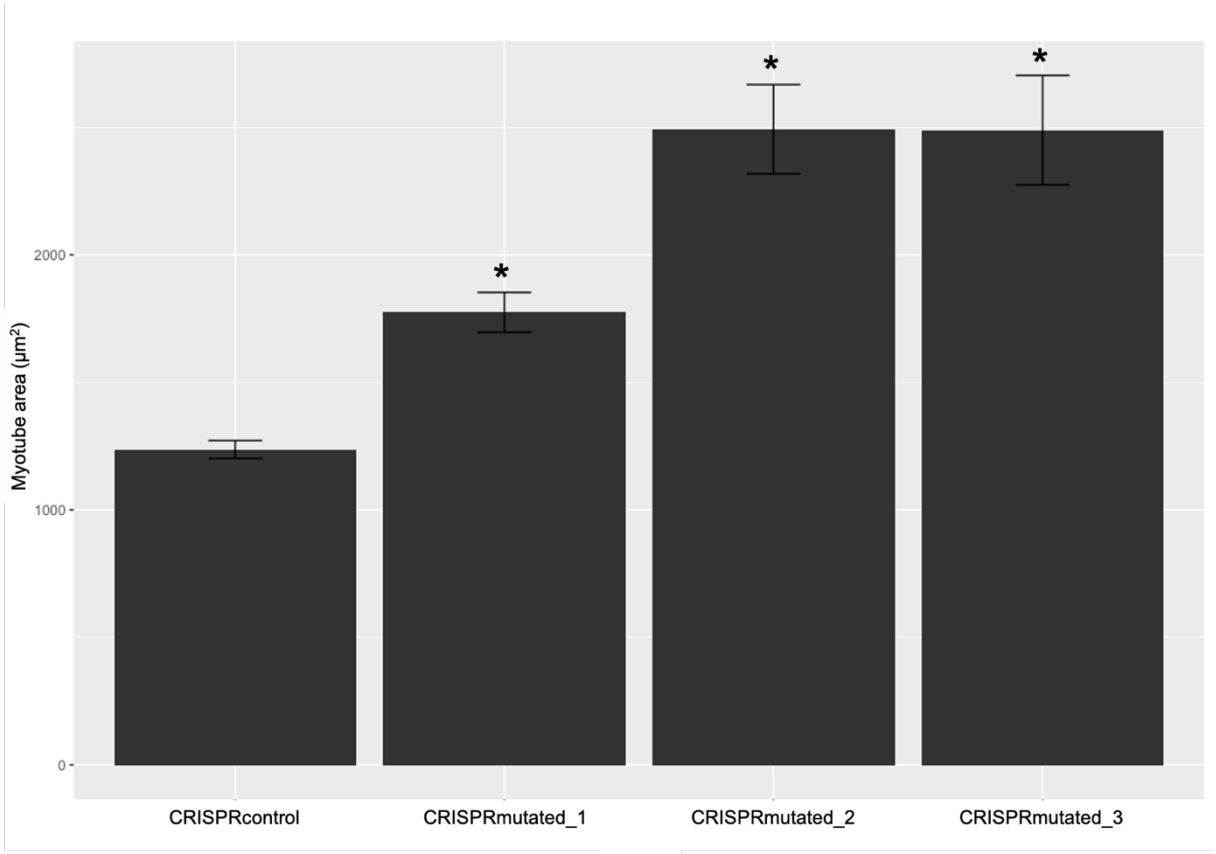


Figure 4: The impact of CRISPR induced *SAP30* mutations on myotube area compared to wildtype *SAP30* control cells. The bar plot shows an increase in myotube area in all the three treatment clones (CRISPRmutated_1, CRISPRmutated_2 and CRISPRmutated_3) when the cell lines compared to the control by Dunnett corrected $p<0.05$. CRISPRcontrol are the cells that received only the empty plasmid (no guide-RNA) and CRISPRmutated_1, CRISPRmutated_2 and CRISPRmutated_3 are the groups that received the plasmid with SAP30 guide-RNA. Different bar colors are representing different clones. p-value < 0.001 (*).

3.4.3 Sanger sequencing reveals different *SAP30* mutations

To observe which types of mutations were generated in the clones that received the CRISPR/Cas9 vector with sgRNA for the candidate gene, the sequence of the fragment containing the *SAP30* mutation was performed. Figure 5 shows the alignment using the original *SAP30 Mus musculus* sequence, from NCBI (SAP30_NCBI), CRISPRcontrol (as a control clone) and CRISPRmutated clones (CRISPRmutated_1, CRISPRmutated_2 and CRISPRmutated_3). Comparing all the results, the CRISPRmutated_1 clone had only one nucleotide (nt) deleted in the position 394 (represented by red circle), while CRISPRmutated_2 and CRISPRmutated_3 had 66nt and 96nt deleted, respectively.

```
SAP30_NCBI          ACCGTGCCGCCACCAGACTTGGGGTGCGGAGTGAACGGCCGCCCAGGGGAAGGAGGCGGA       240
CIRSPRcontrol       ACCGTGCCGCCACCAGACTTGGGGTGCGGAGTGAACGGCCGCCCAGGGGAAGGAGGCGGA       240
CRISPRmutated_1     ACCGTGCCGCCACCAGACTTGGGGTGCGGAGTGAACGGCCGCCCAGGGGAAGGAGGCGGA       240
CRISPRmutated_2     ACCGTGCCGCCACCAGACTTGGGGTGCGGAGTGAACGGCCGCCCAGGGGAAGGAGGCGGA       240
CRISPRmutated_3     ACCGTGCCGCCACCAGACTTGGGGTGCGGAGTGAACGGCCGCCCAGGGGAAGGAGGCGGA       240
                    ************************************************************

SAP30_NCBI          GCGGCGCGGCGAGCGGGGCTCTGTAGCGGCGCCGGCGCCCCCGAACTGGCAGACATGAAC       300
CIRSPRcontrol       GCGGCGCGGCGAGCGGGGCTCTGTAGCGGCGCCGGCGCCCCCGAACTGGCAGACATGAAC       300
CRISPRmutated_1     GCGGCGCGGCGAGCGGGGCTCTGTAGCGGCGCCGGCGCCCCCGAACTGGCAGACATGAAC       300
CRISPRmutated_2     GCGGCGCGGCGAGCGGGGCTCTGTAGCGGCGCCGGCGCCCCCGAACTGGCAGACATGAAC       300
CRISPRmutated_3     GCGGCGCGGCGAGCGGGGCTCTGTAGCGGCGCCGGCGCCCCCGAACTGGCAGACATGAAC       300
                    ************************************************************

SAP30_NCBI          GGCTTCACTCCGGAGGAGATGAGCCGAGGCGGGGACGCGGCCGCCGCCGTGGCCGCGGTG       360
CIRSPRcontrol       GGCTTCACTCCGGAGGAGATGAGCCGAGGCGGGGACGCGGCCGCCGCCGTGGCCGCGGTG       360
CRISPRmutated_1     GGCTTCACTCCGGAGGAGATGAGCCGAGGCGGGGACGCGGCCGCCGCCGTGGCCGCGGTG       360
CRISPRmutated_2     GGCTTCACTCCGGAGGAGATGAGCCGAGGCGGGGACGC----------------------       338
CRISPRmutated_3     ------------------------------------------------------------       300

SAP30_NCBI          GTCGCTGCTGCGGCTGCCGCCGCCTCGGCGGGGAATGGGAACGCAGCCGGCGGTGGGGCA       420
CIRSPRcontrol       GTCGCTGCTGCGGCTGCCGCCGCCTCGGCGGGGAATGGGAACGCAGCCGGCGGTGGGGCA       420
CRISPRmutated_1     GTCGCTGCTGCGGCTGCCGCCGCCTCGGCGGGG●ATGGGAACGCAGCCGGCGGTGGGGCA       419
CRISPRmutated_2     ----------------------------------------AGCCGGCGGTGGGGCA       354
CRISPRmutated_3     ----------------------------------GGGAACGCAGCCGGCGGTGGGGCA       324
                                                       ****************

SAP30_NCBI          GAGGTACCCGGTGCCGGTGCAGTGTCAGCTTCTGGGCCACCTGGAGCGGCCGGTCCCGGC       480
CIRSPRcontrol       GAGGTACCCGGTGCCGGTGCAGTGTCAGCTTCTGGGCCACCTGGAGCGGCCGGTCCCGGC       480
CRISPRmutated_1     GAGGTACCCGGTGCCGGTGCAGTGTCAGCTTCTGGGCCACCTGGAGCGGCCGGTCCCGGC       479
CRISPRmutated_2     GAGGTACCCGGTGCCGGTGCAGTGTCAGCTTCTGGGCCACCTGGAGCGGCCGGTCCCGGC       414
CRISPRmutated_3     GAGGTACCCGGTGCCGGTGCAGTGTCAGCTTCTGGGCCACCTGGAGCGGCCGGTCCCGGC       384
                    ************************************************************

SAP30_NCBI          CCTGGGCAACTCTGTTGTTTGCGGGAGGACGGTGAGCGGTGCGGGCGCGCGGCAGGCAAC       540
CIRSPRcontrol       CCTGGGCAACTCTGTTGTTTGCGGGAGGACGGTGAGCGGTGCGGGCGCGCGGCAGGCAAC       540
CRISPRmutated_1     CCTGGGCAACTCTGTTGTTTGCGGGAGGACGGTGAGCGGTGCGGGCGCGCGGCAGGCAAC       539
CRISPRmutated_2     CCTGGGCAACTCTGTTGTTTGCGGGAGGACGGTGAGCGGTGCGGGCGCGCGGCAGGCAAC       474
CRISPRmutated_3     CCTGGGCAACTCTGTTGTTTGCGGGAGGACGGTGAGCGGTGCGGGCGCGCGGCAGGCAAC       444
                    ************************************************************

SAP30_NCBI          GCCAGCTTCAGCAAAAGGATCCAGAAGAGCATCTCTCAGAAGAAGGTGAAGATCGAGCTG       600
CIRSPRcontrol       GCCAGCTTCAGCAAAAGGATCCAGAAGAGCATCTCTCAGAAGAAGGTGAAGATCGAGCTG       600
CRISPRmutated_1     GCCAGCTTCAGCAAAAGGATCCAGAAGAGCATCTCTCAGAAGAAGGTGAAGATCGAGCTG       599
CRISPRmutated_2     GCCAGCTTCAGCAAAAGGATCCAGAAGAGCATCTCTCAGAAGAAGGTGAAGATCGAGCTG       534
CRISPRmutated_3     GCCAGCTTCAGCAAAAGGATCCAGAAGAGCATCTCTCAGAAGAAGGTGAAGATCGAGCTG       504
                    ************************************************************
```

Figure 5: CRISPR/Cas9 clones selected for morphometric and RNA sequencing analyses. The nucleotide alignment was performed by ClustalW Omega software, using the gene fragment that was sequenced. The SAP30_NCBI sample is the original wildtype *SAP30* Mus musculus sequence from the National Center for Biotechnology Information (NCBI). CRISPRcontrol is the sample that received only the empty plasmid (no guide-RNA), CRISPRmutated_1 (1 nt deleted, represented by a red circle), CRISPRmutated_2 (66 nt deleted) and CRISPRmutated_3 (96 nt deleted) are the treated samples that received the plasmid with SAP30 guide-RNA.

3.4.4 Only one nucleotide deleted caused a difference in the protein structure

To compare the sequencing results from each clone and aiming to see if the mutations are disrupting the protein, we performed the amino acids prediction (Figure 6) using UniProt align tool (https://www.uniprot.org/align/) and protein structure 3D-model using Swiss-model online software

(https://swissmodel.expasy.org/). SAP30 protein has 220 amino acids, and the figure 6 shows the amino acids sequence alignment, and it is possible to observe a change in the open read frame (ORF) from the CRISPRmutated_1 around the 123-position. This is the start of the first SAP30 domain (zinc-finger domain) that is in the position 65 to 133 (NCBI, 2021), probably that is the reason why the other clones did not have changes in the amino acids structure, because the changes are not interfering in the protein domains. Figure 7 shows the results comparing the SAP30 amino acids sequence (M*us musculus*, from NCBI), CRISPRcontrol sequence and the 3 CRISPRmutated clones. The CRISPRmutated_1, the one that had only one nucleotide deleted in the sequence, have a larger difference in the protein structure, when compared with other CRISPRmutated clones (2 and 3). That is happening because even that was only one nucleotide deleted, this deletion is changing the amino acids sequence and the ORF, generating a truncated protein, with loss of function.

```
SAP30_NCBI          MNGFTPEEMSRGGDAAAAVAAVVAAAAAAAASAGNGNAAGGGAEVPGAGAVSASGPPGAAG      60
CIRSPRcontrol       MNGFTPEEMSRGGDAAAAVAAVVAAAAAAAASAGNGNAAGGGAEVPGAGAVSASGPPGAAG      60
CRISPRmutated_1     MNGFTPEEMSRGGDAAAAVAAVVAAAAAAAASAGMGTQPAVGQRYPVPVQCQLLGHLERPV      60
CRISPRmutated_2     MNGFTPEEMSRGGD----------------------AAGGGAEVPGAGAVSASGPPGAAG      38
CRISPRmutated_3     ------------------------------MNGNAAGGGAEVPGAGAVSASGPPGAAG      28
                                                    . * . *     .  *

SAP30_NCBI          PGPGQLCCLREDGERCGRAAGNASFSKRIQKSISQKKVKIELDKSARHLYICDYHKNLIQ      120
CIRSPRcontrol       PGPGQLCCLREDGERCGRAAGNASFSKRIQKSISQKKVKIELDKSARHLYICDYHKNLIQ      120
CRISPRmutated_1     PALG-------NSVVCGRTVSGAGARQATPASAK-GSRRASLRRR---------------      97
CRISPRmutated_2     PGPGQLCCLREDGERCGRAAGNASFSKRIQKSISQKKVKIELDKSARHLYICDYHKNLIQ      98
CRISPRmutated_3     PGPGQLCCLREDGERCGRAAGNASFSKRIQKSISQKKVKIELDKSARHLYICDYHKNLIQ      88
                    *. *       :.  ***:...*.  :    * .  . : .* :

SAP30_NCBI          SVRNRRKRKGSDDDGGDSPVQDIDTPEVDLYQLQVNTLRRYKRHFKLPTRPGLNKAQLVE      180
CIRSPRcontrol       SVRNRRKRKGSDDDGGDSPVQDIDTPEVDLYQLQVNTLRRYKRHFKLPTRPGLNKAQLVE      180
CRISPRmutated_1     ------------------------------------------------------------      97
CRISPRmutated_2     SVRNRRKRKGSDDDGGDSPVQDIDTPEVDLYQLQVNTLRRYKRHFKLPTRPGLNKAQLVE      158
CRISPRmutated_3     SVRNRRKRKGSDDDGGDSPVQDIDTPEVDLYQLQVNTLRRYKRHFKLPTRPGLNKAQLVE      148

SAP30_NCBI          IVGCHFKSIPVNEKDTLTCFIYSVRNDKNKSDLKADSGVH    220
CIRSPRcontrol       IVGCHFKSIPVNEKDTLTCFIYSVRNDKNKSDLKADSGVH    220
CRISPRmutated_1     ----------------------------------------    97
CRISPRmutated_2     IVGCHFKSIPVNEKDTLTCFIYSVRNDKNKSDLKADSGVH    198
CRISPRmutated_3     IVGCHFKSIPVNEKDTLTCFIYSVRNDKNKSDLKADSGVH    188
```

Figure 6: Amino acids sequences alignment showing the differences between SAP30 protein from *Mus musculus*, CRISPRcontrol and each CRISPRmutated. SAP30_NCBI is the SAP30 original amino acid sequence from *Mus musculus* (NCBI database). CRISPRcontrol is the sequence from C2C12 cells that received only the empty vector. CRISPRmutated (1, 2 and 3) are C2C12 cells that received a vector with the SAP30 guide-RNA and CRISPR/Cas9 system. Different colors represent different amino acids.

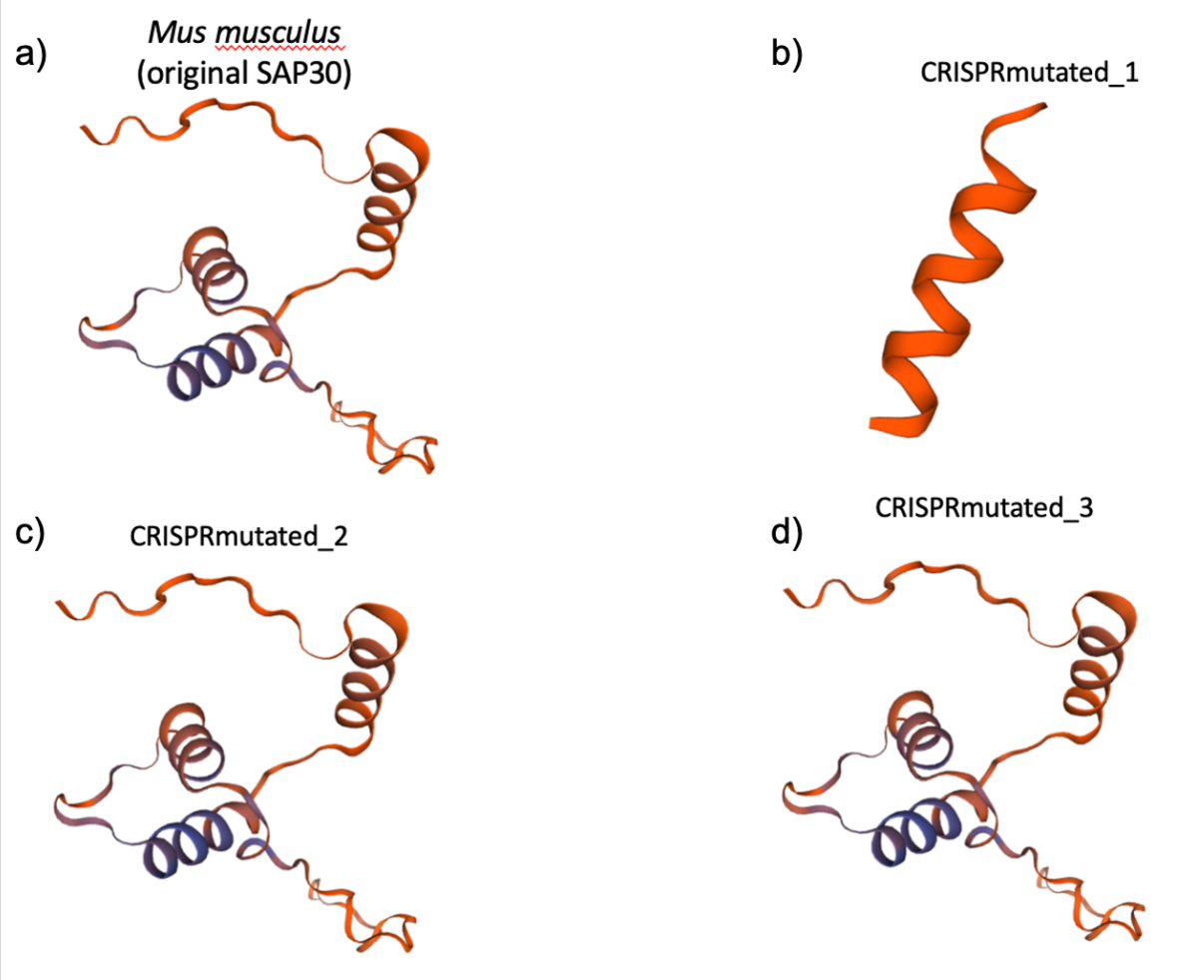


Figure 7: Images from Swiss-model software showing the differences between all the mutated clones compared with the original sequence from NCBI database. a) SAP30 protein structure, original from *Mus musculus*. b) CRISPRmutated_1 protein structure after 1 nucleotide deleted by CRISPR/Cas9. c) CRISPRmutated_2 protein structure after 66 nucleotides deleted by CRISPR/Cas9. d) CRISPRmutated_3 protein structure after 96 nucleotides deleted by CRISPR/Cas9.

3.4.5 Differentially expressed (DE) genes analysis is showing many DE genes in each CRISPRmutated clone comparing with CRISPRcontrol

To understand the mechanism by which *SAP30* increase muscle hypertrophy we compared the mRNA expression profile by differentially expressed (DE) gene analysis. The DE gene analysis compared the CRISPRcontrol (empty CRISPR/Cas9 plasmid) with each treated group/clone (CRISPR/Cas9 plasmid with the guide-RNA for *SAP30*: CRISPRmutated_1, CRISPRmutated_2 and CRISPRmutated_3) to understand the molecular mechanisms that were differing between the groups with different mutations caused by CRISPR/Cas9 in the *SAP30* gene.

The CRISPR treated samples were compared separately because, as observed in figure 5, each CRISPR/Cas9 guide RNA treated *SAP30* cell line had a different kind of mutation and these different mutations are altering the

protein reading frame and, consequently, the protein structure. For this reason, each treated clone was compared to the control clone to understand which genes were being differentially expressed between each sample.

3.4.5.1 DE analysis comparing Control x Treated group 1 (CRISPRmutated_1)

To understand which genes were involved in the hypertrophy phenotype in CRISPRmutated_1, DE analysis was performed comparing the mRNA data from CRISPRcontrol with CRISPRmutated_1, shows a total of 8,323 DE genes. Among the DE genes, 4,443 were downregulated and 4,279 were upregulated when compared to the control. Figure 8 presents the DE results as a volcano plot based on Log2 Fold Change and -log(10) adjusted p-values comparing CRISPRcontrol vs CRISPRmutated_1 cells. Our candidate gene, *SAP30* was DE in the CRISPRmutated_1 (Log2 Fold Change = -2.12; p-adjusted<0.05).

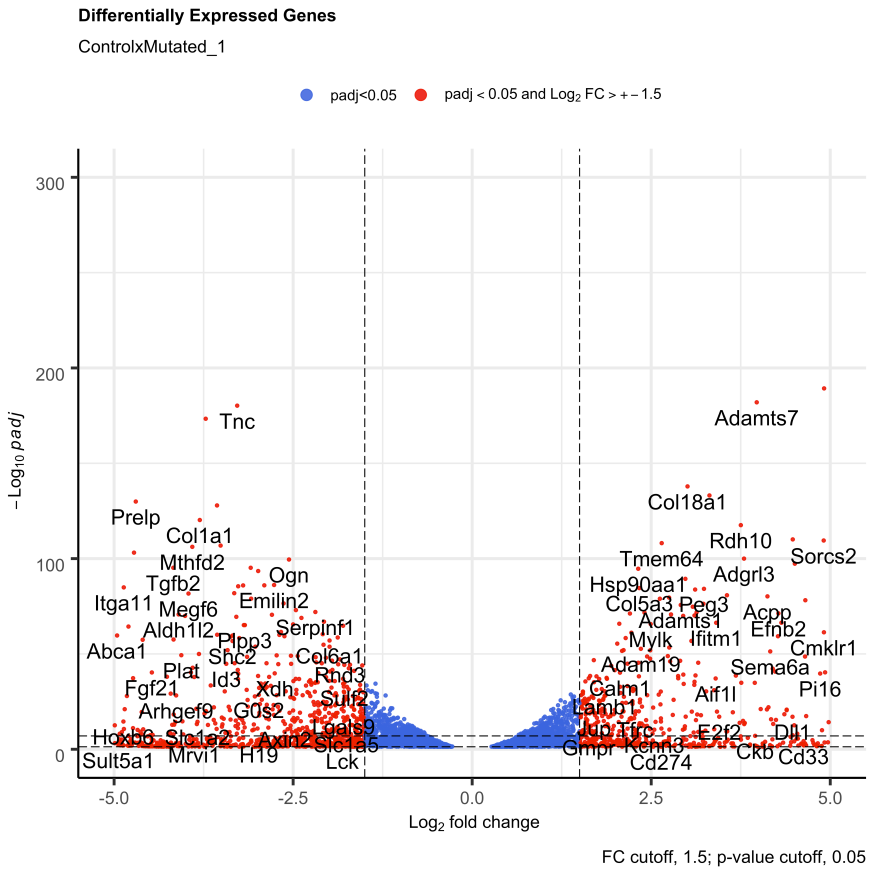


Figure 8. Volcano plot showing all the differentially expressed (DE) genes when comparing CRISPRcontrol x CRISPRmutated_1 from CRISPR/Cas9 experiment. The red dots are the DE genes with Log2 Fold Change larger than 1.5 in both sides (positive and negative) and with adjusted p-value < 0.05.

3.4.5.2 DE analysis comparing Control x Treated group 2 (CRISPRmutated_2)

The comparison between Control and CRISPRmutated_2 was performed as previously described, also to understand the relation between DE genes with the hypertrophy phenotype in the CRISPRmutated_2 cells. Here, a total of 5,565 DE genes were identified when comparing CRISPRcontrol vs. CRISPRmutated_2 cell group (66 nt deleted). A total of 2,976 genes were downregulated and 2,588 were upregulated. Again, *SAP30* was DE in the CRISPRmutated_2 and exhibited a Log2 Fold Change = -1.318 (P-adj = 0.0018). Figure 9 shows the DE genes in a volcano plot, based on their Log2 Fold Change and -log(10) adjusted p-values.

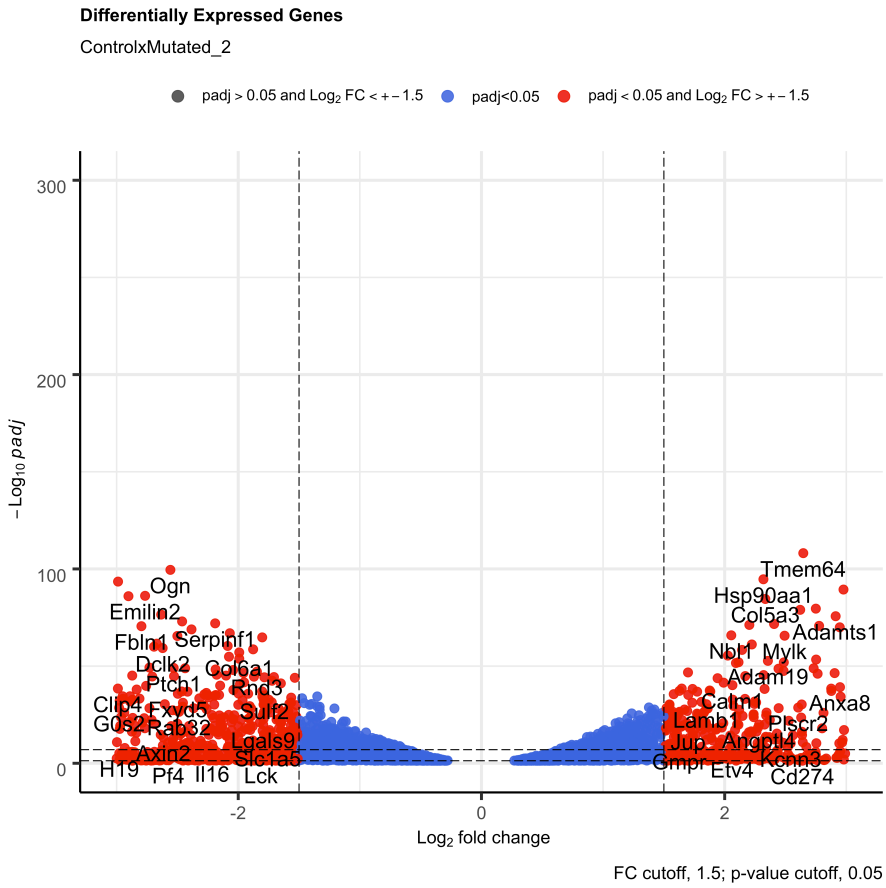


Figure 9. Volcano plot showing some differentially expressed DE genes comparing CRISPRcontrol x CRISPRmutated_2 from CRISPR/Cas9 experiment. The red dots are the differentially expressed (DE) genes with Log2 Fold Change bigger than 1.5 in both sides (positive and negative) and with p-value adjusted < 0.05.

3.4.5.3 DE analysis comparing Control x CRISPRmutated_3 cells

The RNA-sequencing data from CRISPRontrol and CRISPRmutated_3 was compared to check genes that were differentially expressed between the groups and could explain the CRISPRmutated_3 hypertrophy. A total of 7,304 DE genes were identified between the control vs. CRISPRmutated_3 cells, with 3,553 downregulated and 3,751 upregulated. Again, *SAP30* was DE with log2 Fold Change = -2.76 (P-adj =4.47e-13). The DE genes are presented in a volcano plot in Figure 10.

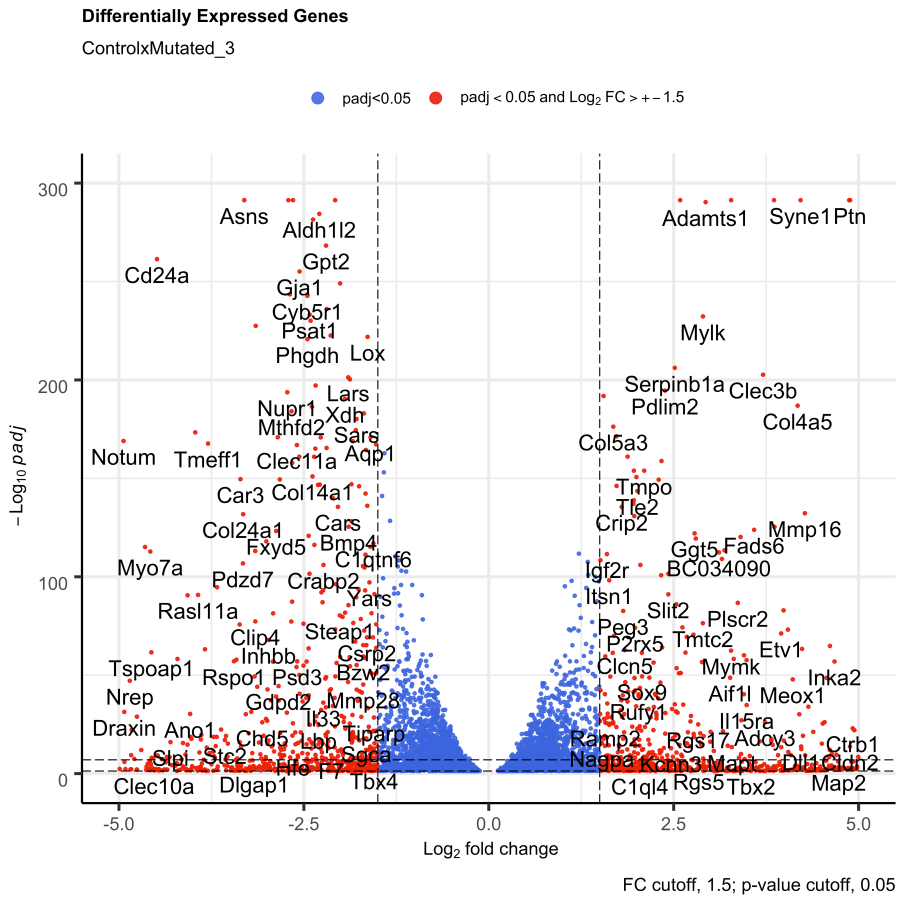


Figure 10. Volcano plot comparing CRISPRontrol vs. CRISPRmutated_3 from the CRISPR/Cas9 experiment. The red dots are the differentially expressed genes with Log2 Fold Change > 1.5 (positive and negative) and p-value adjusted < 0.05.

3.4.6 Overlap analysis for DE genes in the CRISPRmutated clones

To check how many genes were overlapping between the different mutations caused by SAP30-CRISPR/Cas9 we performed a Venn Diagram analysis (Figure 11). A total of 3,337 common genes were identified when all the DE genes were compared. After that, the functional enrichment analysis was performed separately for each clone compared with control cells using MetaCore (MetaCore, 2021).

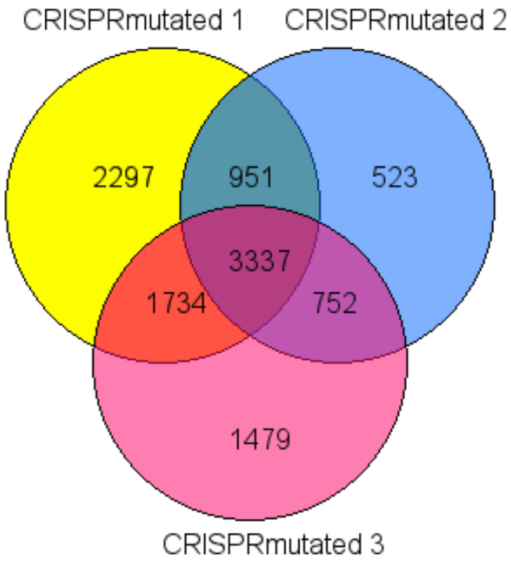


Figure 11. Venn diagram showing the DE genes overlapping across the three mutated clones (CRISPRmutated_1, CRISPRmutated_2 and CRISPRmutated_3). A total of 3337 genes were identified in each of the three independent comparisons of mutated cells compared to the control. Overlap and differences among the different clones may represent biological differences due to different types of *SAP30* mutations with different biological consequences. Venn Diagram performed using the online tool Gene List Venn Diagram (http://genevenn.sourceforge.net/).

3.4.7 Functional Enrichment analysis comparing CRISPRcontrol x CRISPRmutated_1

The functional enrichment analysis (Gene Ontology, GO terms) shows genes related with several cellular terms, as cellular metabolic process, cellular component organization or biogenesis, cellular component organization, cellular process (Figure 12). According to the networks, the DE genes are involved in important pathways related with muscle development: NOCTH and WNT. In the NOTCH signaling pathway, it is possible to see the participation of the candidate gene *SAP30* (Figure 13). The *SAP30*, as already mentioned, was downregulated, and participates in the developmental NOTCH signaling pathway. Other significant canonical pathway related to muscle development found was the Development_positive regulation of WNT/B-catenin signaling in the cytoplasm (Figure 14). In this pathway, DE genes already described related with muscle process were observed presenting changes in their expression, as *WNT*, *SMAD3*, *SMAD4* and *TGF-Beta* being upregulated and *DACT1* downregulated.

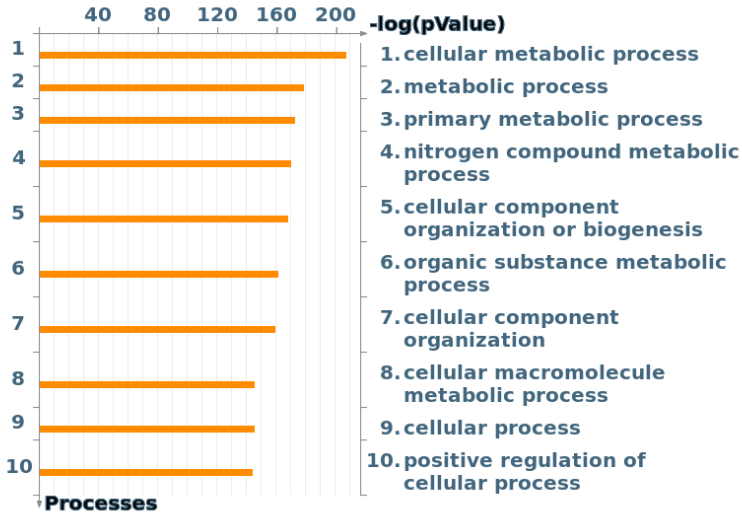


Figure 12: Gene Ontology terms showing the top 10 significantly ($p<0.05$) processes that the DE genes from CRISPRmutated_1 clone are related. Image created by MetaCore software

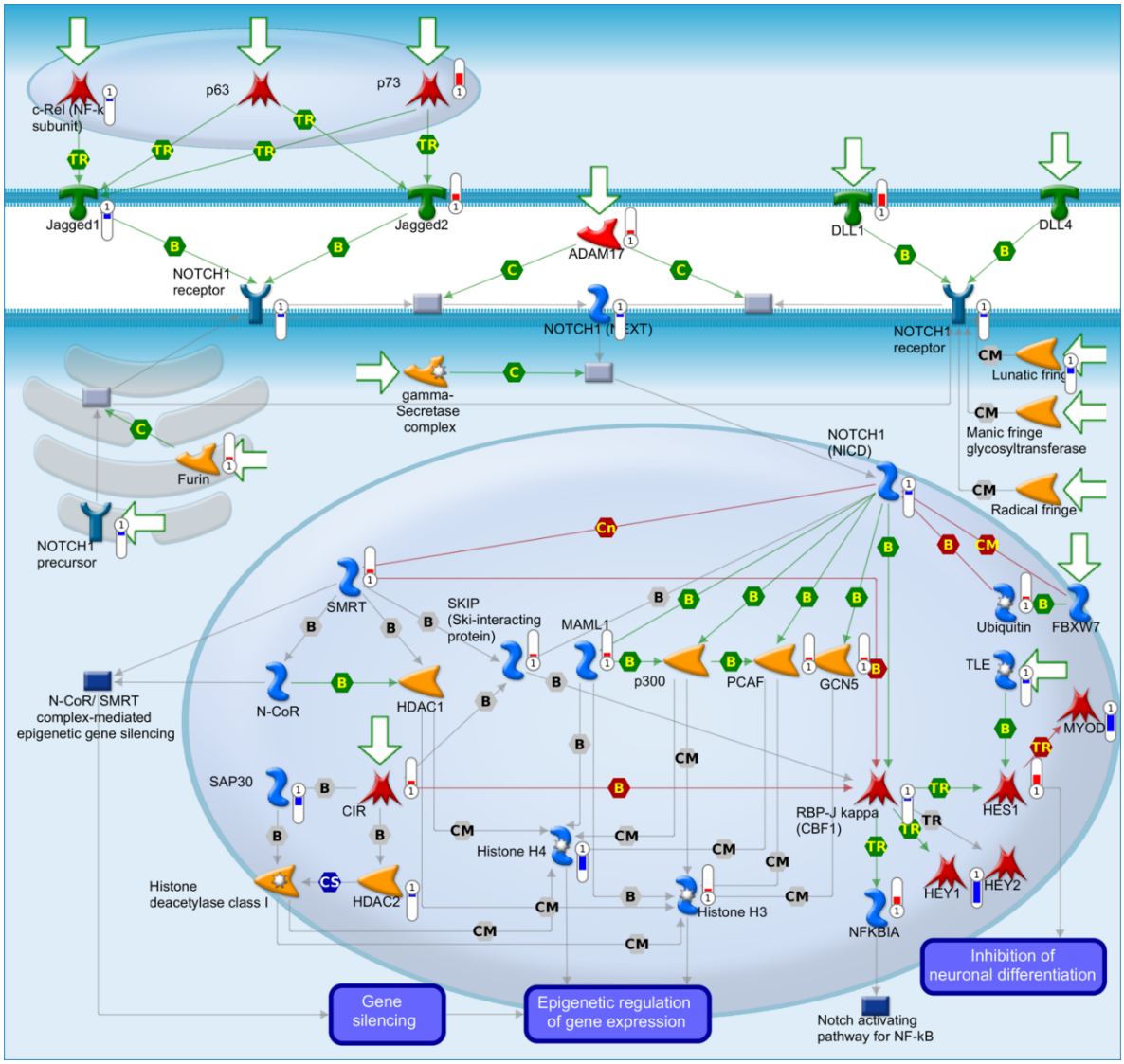


Figure 13: Developmental NOTCH signaling pathway showing the differentially expressed genes for the CRISPRmutated_1. The blue thermometer is genes that were downregulated in CRISPRmutated_1, while the red thermometer is representing upregulated genes. Red lines are representing genes inactivating other genes, green lines are representing genes activation and grey lines are representing unspecific relationship between genes. Analysis performed by MetaCore software.

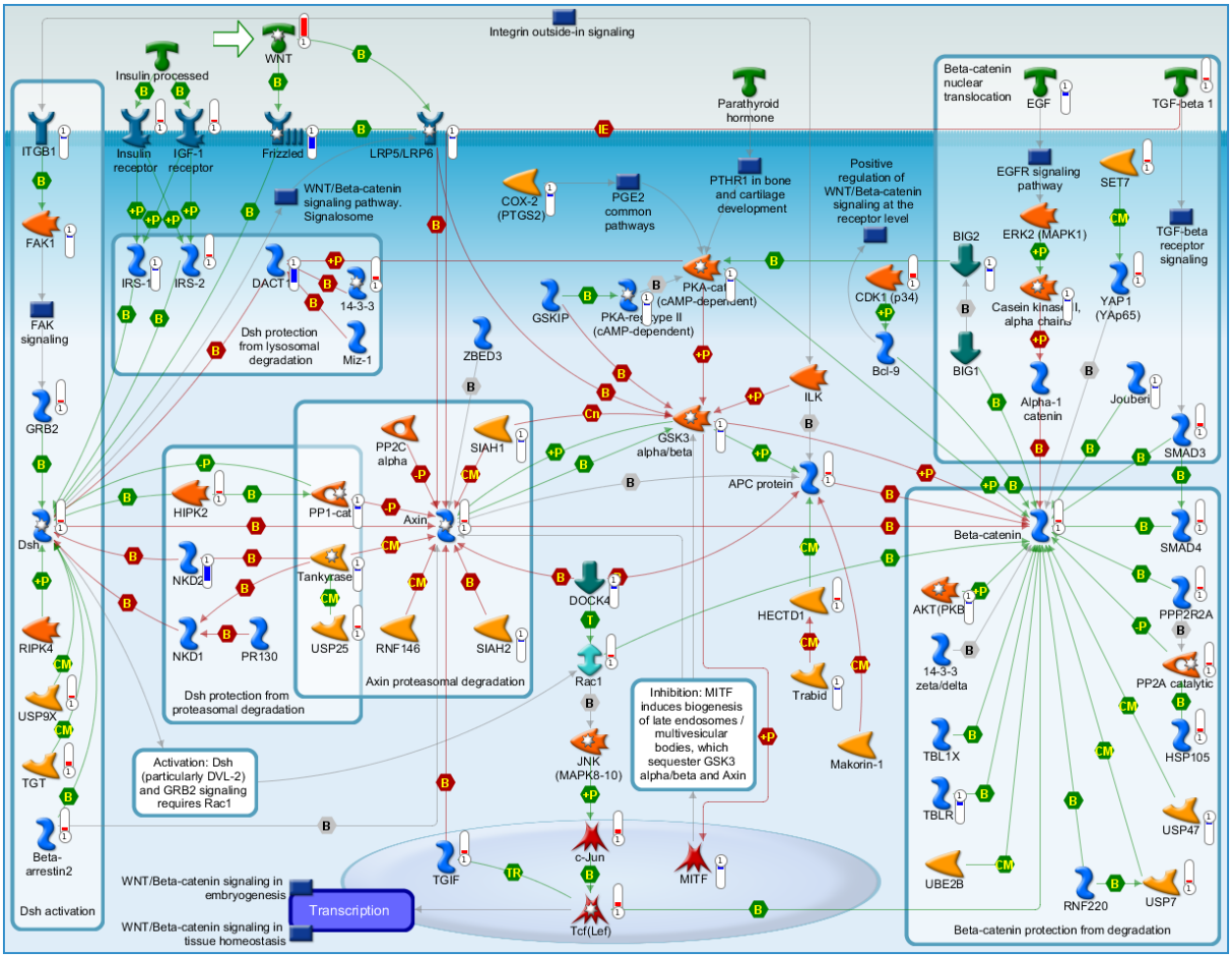


Figure 14. Developmental positive regulation of WNT/Beta-catenin signaling in the cell cytoplasm. The blue thermometer is genes that were downregulated in CRISPRmutated_1, while the red thermometer is representing upregulated genes. Red lines are representing genes inactivating other genes, green lines are representing genes activation and grey lines are representing unspecific relationship between genes. Analysis performed by MetaCore software.

3.4.8 Functional Enrichment analysis comparing CRISPRcontrol x CRISPRmutated_2

The functional enrichment analysis was performed comparing the CRISPRcontrol x CRISPRmutated_2. In the GO terms analysis, process related with cellular process, metabolic process, system development and regulation of cell communication were found (Figure 15). According with the networks that these DE genes are participating it is possible to see three with relation in the muscle development and hypertrophy as Signal transduction NOTCH signaling, Cytoskeleton_actin filaments and muscle contraction. Figure 16 shows the signal transduction NOTCH signaling and important genes participating in this pathway were DE when compared CRISPRcontrol x CRISPRmutated_2. The figure is showing the same map that mentioned before (in the CRISPRmutated_1) but here, DE genes found in the CRISPRmutated_2 group and their expression levels are represented. In the cytoskeleton actin

filaments, there are genes as *Myosin I, Actin, Troponin I, Tropomyosin I, Actin cytoskeletal*, while in the muscle contraction important genes related with this trait were found: *Myosin II, Alpha-actinin, MyHC,* also *Actin, MyLA* and *Actin muscle.*

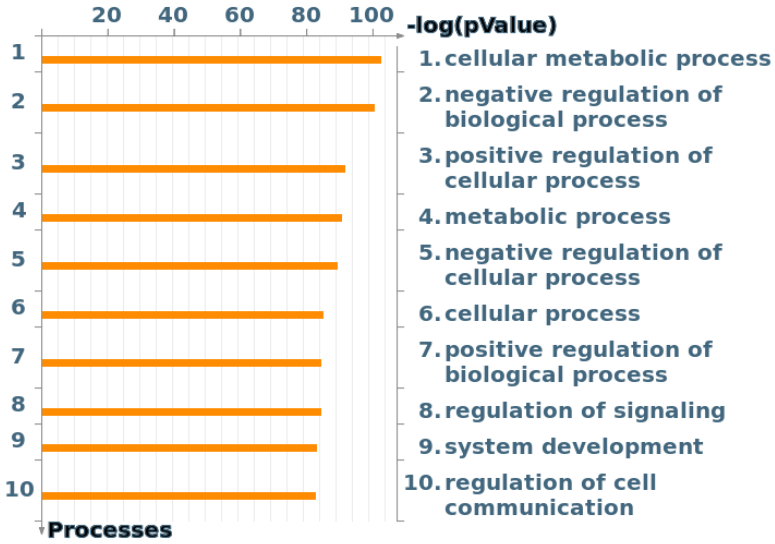


Figure 15: Gene Ontology terms showing the top 10 significantly ($p<0.05$) processes that the DE genes for the CRISPRmutated_2 clone are related. Image created by MetaCore software.

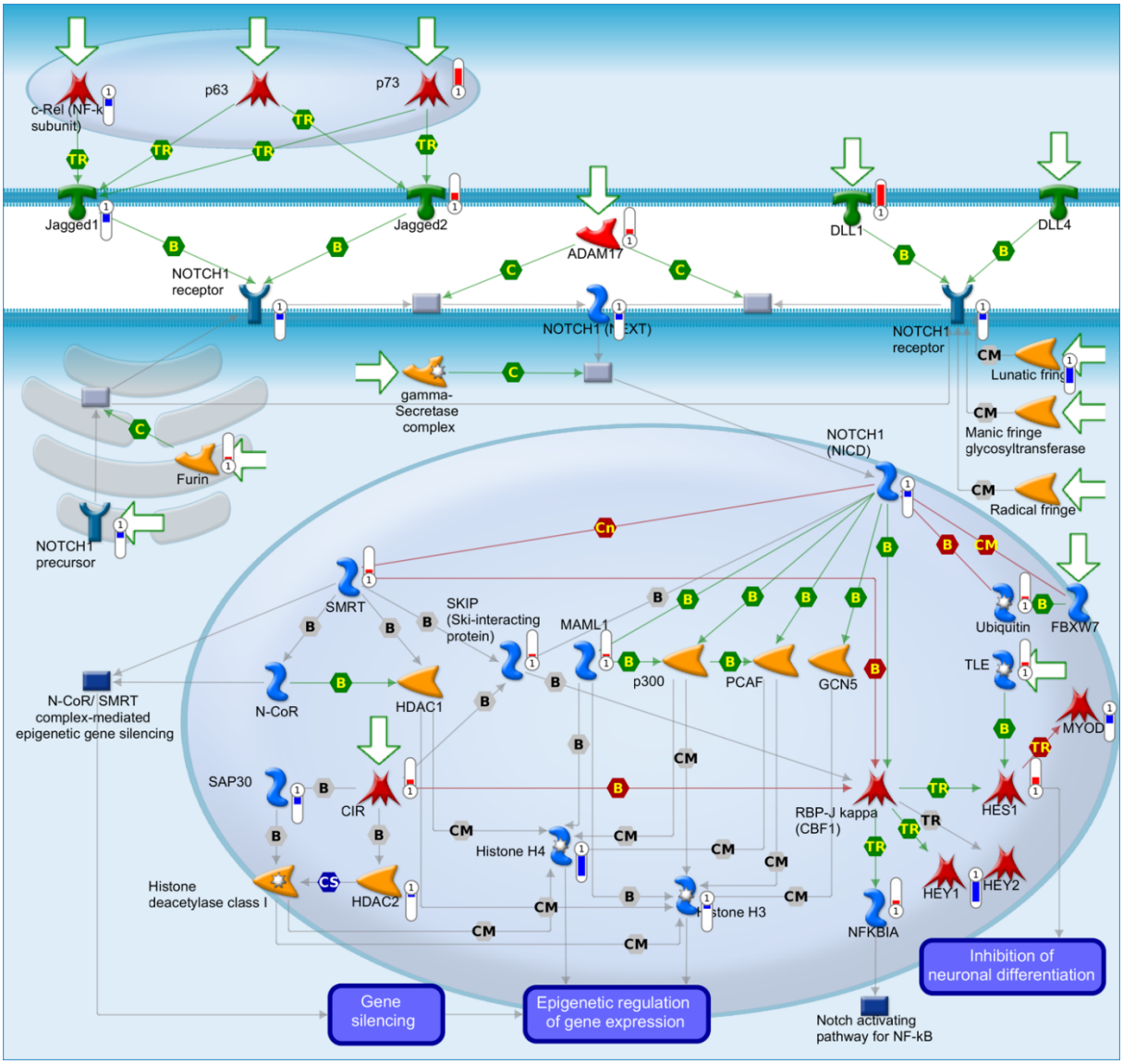


Figure 16: developmental NOTCH signaling pathway showing the differentially expressed genes for the CRISPRmutated_2. The blue thermometer is genes that were downregulated in CRISPRmutated_2, while the red thermometer is representing upregulated genes. Red lines are representing genes inactivating other genes, green lines are representing genes activation and grey lines are representing unspecific relationship between genes. Analysis performed by MetaCore software.

3.4.9 Functional Enrichment analysis comparing CRISPRcontrol x CRISPRmutated_3

Functional enrichment analysis for the CRISPRmutated_3 shows GO terms related with cellular metabolic process, cellular component organization, organic substance metabolic process, negative regulation of biological process (Figure17). The networks analysis revealed maps also related with Signal transduction NOTCH signaling and Developmental positive regulation of WNT/Beta-catenin signaling in the cell cytoplasm (Figure 18 and 19, respectively), and cytoskeleton actin filaments and regulation of cytoskeleton rearrangement. Figure 18 shows the signal transduction NOTCH signaling and it is possible to see the *SAP30* gene represented downregulated as *MYOD*, other

gene involved in the muscle regulation. The Developmental positive regulation of WNT/Beta-catenin signaling in the cell cytoplasm represented in the figure 19 shows genes as *WNT, Axin, DAB2, p53, E2F1*, and *Dsh* participating in the map.

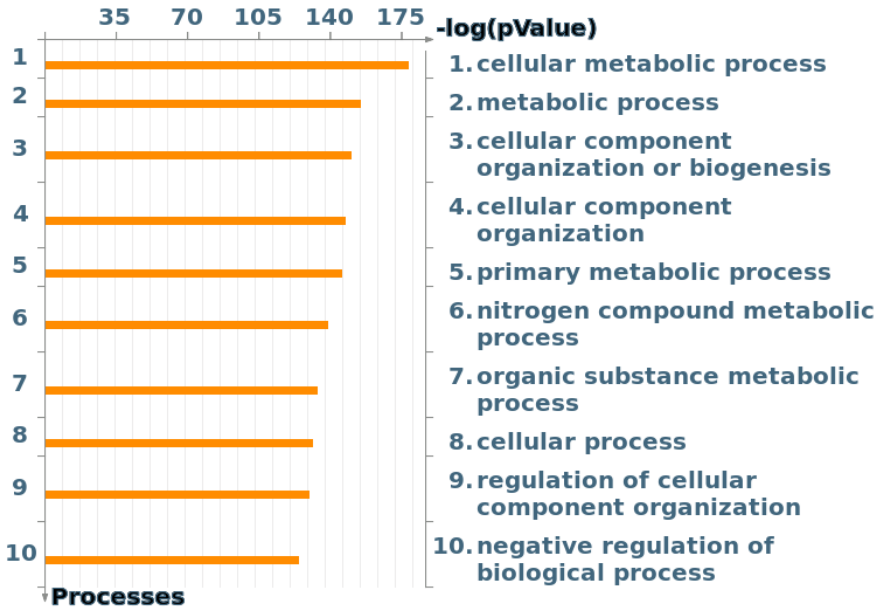


Figure 17: Figure 15: Gene Ontology terms showing the top 10 significantly ($p<0.05$) processes that the DE genes for the CRISPRmutated_3 clone are related. Image created by MetaCore software.

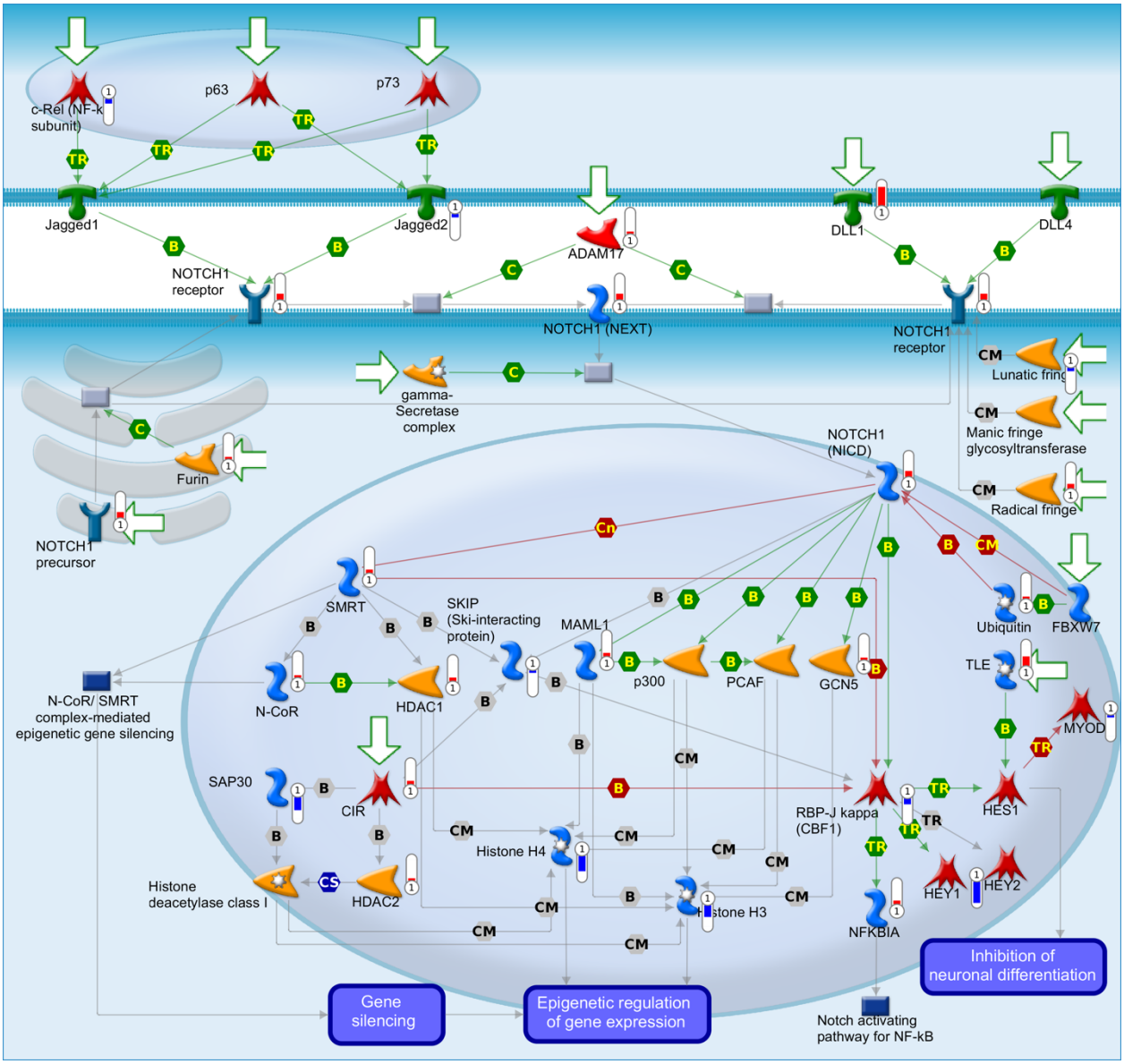


Figure 18. Developmental NOTCH signaling pathway showing the differentially expressed genes for the CRISPRmutated_3. The blue thermometer is genes that were downregulated in CRISPRmutated_3, while the red thermometer is representing upregulated genes. Red lines are representing genes inactivating other genes, green lines are representing genes activation and grey lines are representing unspecific relationship between genes. Analysis performed by MetaCore software.

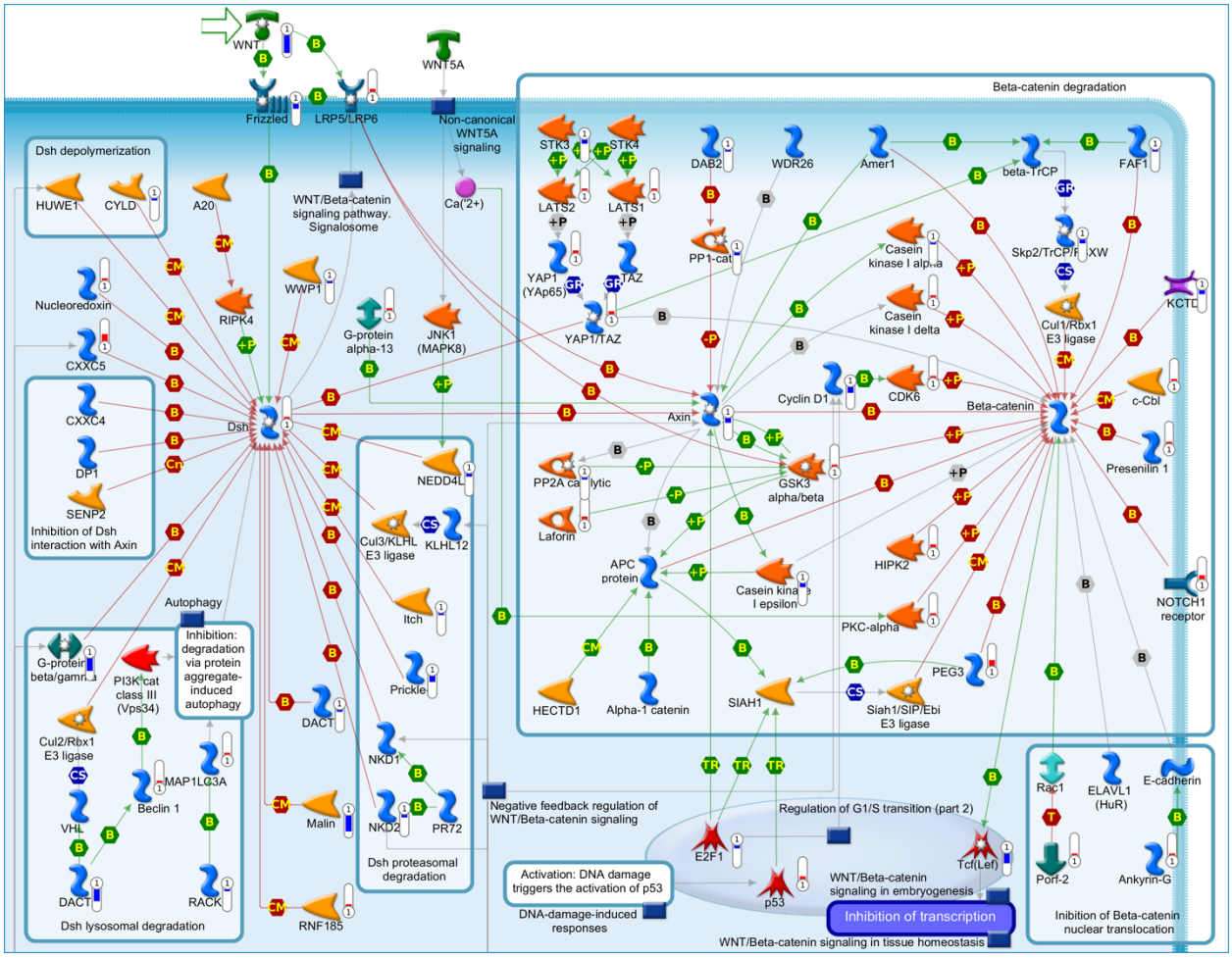


Figure 19. Developmental positive regulation of WNT/Beta-catenin signaling in the cell cytoplasm showing the differentially expressed genes for the CRISPRmutated_3. The blue thermometer is genes that were downregulated in CRISPRmutated_3, while the red thermometer is representing upregulated genes. Red lines are representing genes inactivating other genes, green lines are representing genes activation and grey lines are representing unspecific relationship between genes. Analysis performed by MetaCore software.

3.5 Discussion

Even with different mutations in the treated clones, we could observe hypertrophy in all *SAP30* gene deletions. This result corroborates with a previous study (Petry et al., 2021) where the cells treated with siRNA for *SAP30* had an increase in the myotube cell area. The differentially expressed analysis for the mutated cells revealed a difference in the expression level of genes involved in muscle development.

The participation of the *SAP30* gene in muscle hypertrophy was already reported (Petry et al., 2021). Here, *SAP30* was downregulated in all the treated groups (CRISPRmutated_1, CRISPRmutated_2 and CRISPRmutated_3). This gene is an important Sin3 multiprotein complex member and play a central role in gene silencing by the deacetylation of histones (Cowley et al., 2005). The SAP30 is a 30kDa protein with a poorly conserved N-terminal region and a highly conserved central region with a zinc finger domain (Xie et al., 2011). The CRISPRmutated_1 was the only one where the mutations caused by CRISPR/Cas9 system could change the SAP30 protein structure.

The functional enrichment analysis revealed that all the three mutated clones had DE genes related with developmental NOTCH signaling pathway, an important metabolic pathway where *SAP30* is inserted, but little to know about the gene function in this pathway. The NOTCH via is an important signing for the early phases in the muscle regeneration, promoting the cell proliferation and the activation of satellite cells (Conboy et al., 2003). In mature muscle cells, NOTCH is important in the *Pax7* and *MyoD* activation, promoting the myogenic lineage and differentiation (Zammit et al., 2006). In our study, *Pax7* was not DE, however, the *MyoD* gene was downregulated in all the CRISPRmutated cells, given evidence that the *SAP30* mutations are interfering in the hypertrophy phenotype by its interference in the NOTCH pathways.

The forced inactivation of the *Notch1* gene in satellite cells showed an increase in the cell proliferation and a reduction in the myogenic differentiation (Conboy et al., 2003). This finding may be corroborating with our results where, probably, the NOTCH pathway is activated, when *SAP30* mutated, inducing the myogenic differentiation and hypertrophy. Other evidence corroborating with these thoughts is that we did not have any significant difference in the fusion index analysis, providing that the mutations are not affecting the proliferation stages, only differentiation.

Another important pathway that the DE genes are involved is the regulation of WNT/Beta-catenin, significant for two treatments: CRISPRmutated_1 and CRISPRmutated_3. The WNT signaling has been characterized as important in the muscle formation (Cossu and Borello, 1999) and essential for the muscle differentiation (Anakwe et al., 2003). This pathway play an important role in the myogenic postnatal, acting directly in the muscle stem cells and controlling the myogenic progression and muscle regeneration (Brack et al., 2008). The enrichment analysis is also corroborating with the increase in myotube area observed because the participation of the activation in the WNT/B-catenin pathway in the muscle cell hypertrophy was already reported by Armstrong et al. (2005), where there was a significant increase in the B-catenin activity when in this phenotype.

CRISPRmutated_2 has pathways related with NOTCH signaling, as mentioned before, and for cytoskeleton_actin filaments and muscle contraction. The *SAP30* mutations induced a difference in the expression levels of several genes involved in these two processes. As examples, *Troponin T, Tropomyosin, Alpha-actin, Myosin II and ACTA1* being upregulated, and *Desmin, Ryanodine receptor 2, ACTA2, ACTC and Smooth muscle myosin* being downregulated. All these genes are participating in the muscle contraction pathway (Appendix A).

Actin cytoskeleton filaments are essential in the cell regeneration dynamic (Lee and Dominguez, 2010), being important also for the regulation and cell migration in the smooth muscle process (Tang and Gerlach, 2017). Muscle contraction is an important event to the animal movements. There are several kinds of muscle contraction, and all of them are regulated by essential signal transductor called PCG1α that activates some transcription factors as FOXO and Mef2 (Kuo and Ehrlich, 2015). Between the DE genes in the CRISPRmutated_2 there are some members of the FOXO family, as *Foxo3*, *Foxo4* and *Foxo6* that were upregulated, while the *Foxo1* was downregulated. In the Mef2 family, the *Mef2d* was downregulated. Together, these evidence may be important to clear the *SAP30* gene function, given us an idea of its participation in the skeletal muscle development, since the embryogenic stages, by NOTCH and WNT pathways and the importance of the relationship between these maps in the muscle hypertrophy.

3.6 Conclusion

Morphometric analysis revealed the important contribution of *SAP30* gene, regulating the hypertrophy phenotype. Moreover, it was clear to see its importance controlling the expression level of several genes, as myogenic markers, transcription factors and muscle contraction proteins. The differences observed in the genes and,

consequently, in the cell phenotype are affecting important pathways related with muscle growth, regeneration, and regulation.

3.7 References

Anakwe, K., Robson, L., Hadley, J., Buxton, P., Church, V., Allen, S., et al. (2003). Wnt signalling regulates myogenic differentiation in the developing avian wing. *Development* 130, 3503–3514. doi:10.1242/dev.00538.

Armstrong, D. D., Esser, K. A., and Wnt, K. A. E. (2005). Wnt/-catenin signaling activates growth-control genes during overload-induced skeletal muscle hypertrophy. *Am J Physiol Cell Physiol* 289, 853–859. doi:10.1152/ajpcell.00093.2005.--Catenin.

Brack, A. S., Conboy, I. M., Conboy, M. J., Shen, J., and Rando, T. A. (2008). A Temporal Switch from Notch to Wnt Signaling in Muscle Stem Cells Is Necessary for Normal Adult Myogenesis. *Cell Stem Cell* 2, 50–59. doi:10.1016/j.stem.2007.10.006.

Conboy, I. H., Conboy, M. J., Smythe, G. M., and Rando, T. A. (2003). Notch-Mediated Restoration of Regenerative Potential to Aged Muscle. *Science (80-.).* 302, 1575–1577. doi:10.1126/science.1087573.

Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* 16, 10881–10890. doi:10.1093/nar/16.22.10881.

Cossu, G., and Borello, U. (1999). Wnt signaling and the activation of myogenesis in mammals. *EMBO J.* 18, 6867–6872. doi:10.1093/emboj/18.24.6867.

Costa, A. do C., Gonçalves Lio Copola, A., Carvalho Souza, C., Meireles Nogueira, J., Aparecida Borges Silva, G., and Cristina Jorge, E. (2021). RGMa can induce skeletal muscle cell hyperplasia via association with neogenin signalling pathway. doi:10.1007/s11626-021-00555-9.

Cowley, S. M., Iritani, B. M., Mendrysa, S. M., Xu, T., Cheng, P. F., Yada, J., et al. (2005). The mSin3A Chromatin-Modifying Complex Is Essential for Embryogenesis and T-Cell Development. *Mol. Cell. Biol.* 25, 6990–7004. doi:10.1128/mcb.25.16.6990-7004.2005.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi:10.1093/bioinformatics/bts635.

Doudna, J. A., and Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science (80-.).* 346. doi:10.1126/science.1258096.

Douglas, M., and Pinheiro, J. C. (1998). NONLINEAR MIXED-EFFECTS MODELS LINEAR AND NONLINEAR MIXED-EFFECTS MODELS. *Conf. Appl. Stat. Agric.* doi:10.4148/2475-7772.1273.

Doynova, M. D., Markworth, J. F., Cameron-Smith, D., Vickers, M. H., and O'Sullivan, J. M. (2017). Linkages between changes in the 3D organization of the genome and transcription during myotube differentiation in vitro. *Skelet. Muscle* 7, 1–14. doi:10.1186/s13395-017-0122-1.

Hothorn, T., Bretz, F., and Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical J.* 50, 346–363. doi:10.1002/bimj.200810425.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakatura, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isoenzyme conversion in Escherichia coli, and identification of the gene product. *J. Bacteriol.* 169, 5429–5433. doi:10.1128/jb.169.12.5429-5433.1987.

Koslová, A., Trefil, P., Mucksová, J., Reinišová, M., Plachý, J., Kalina, J., et al. (2020). Precise CRISPR/Cas9 editing

of the NHE1 gene renders chickens resistant to the J subgroup of avian leukosis virus. *Proc. Natl. Acad. Sci. U. S. A.* 117, 2108–2112. doi:10.1073/pnas.1913827117.

Kuo, I. Y., and Ehrlich, B. E. Signaling in Muscle Contraction. doi:10.1101/cshperspect.a006023.

Lee, S. H., and Dominguez, R. (2010). Regulation of Actin Cytoskeleton Dynamics in Cells. *Mol. Cells* 29, 311–325. doi:10.1007/s10059-010-0053-8.

Liao, Y., Smyth, G. K., and Shi, W. (2014). Sequence analysis featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. 30, 923–930. doi:10.1093/bioinformatics/btt656.

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15. doi:10.1186/s13059-014-0550-8.

Magnusson, C., Svensson, A., Christerson, U., and Tågerud, S. (2005). Denervation-induced alterations in gene expression in mouse skeletal muscle. *Eur. J. Neurosci.* 21, 577–580. doi:10.1111/j.1460-9568.2005.03855.x.

Marchesi, J. A. P., Buzanskas, M. E., Cantão, M. E., Ibelli, A. M. G., Peixoto, J. O., Joaquim, L. B., et al. (2018). Relationship of runs of homozygosity with adaptive and production traits in a paternal broiler line. Animal 12, 1126–1134. doi:10.1017/S1751731117002671.

Martins, A.F., José Xavier Neto, Ana Azambuja, Maria Lorena Sereno, Antonio Figueira, Paulo Henrique Campos-Junior, et al. (2014). Repulsive Guidance Molecules a, b and c Are Skeletal Muscle Proteins, and Repulsive Guidance Molecule a Promotes Cellular Hypertrophy and Is Necessary for Myotube Fusion. *RGMb RGMc Cells Tissues Organs* 200, 326–338. doi:10.1159/000433491.

MetaCore MetaCore Login | Clarivate. Available at: https://portal.genego.com/ [Accessed May 11, 2021].

Mosmann, T. (1983). Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays. *J. Immunol. Methods* 65, 55–63. doi:10.1016/0022-1759(83)90303-4.

National Center for Biotechnology Information (NCBI)[Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; [1988] – [cited 2017 Apr 06]. Available from: https://www.ncbi.nlm.nih.gov/

Petry, B.P.; Moreira, G.C.M.; Copola, A.L.; Souza, M.M.; Veiga, F.C.; Jorge, E.C.; Peixoto, J.O.; Ledur, M.C.; Koltes, J.E.; Coutinho, L.L. SAP30 gene is a probable regulator of muscle hypertrophy in chickens. Frontiers in Genetics (2021).

R Core Team (2019). R: The R Project for Statistical Computing. Available at: https://www.r-project.org/ [Accessed May 11, 2021].

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* 8, 2281–2308. doi:10.1038/nprot.2013.143.

Sievers F, Wilm A, Dineen DG, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins D . Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega . Molecular Systems Biology 7 Article number: 539 doi:10.1038/msb.2011.75

Sorelle, E. D., Dai, J., Bonglack, E. N., Heckenberg, E. M., Zhou, J. Y., Giamberardino, S. N., et al. (2021). Single-cell rna-seq reveals transcriptomic heterogeneity mediated by host-pathogen dynamics in lymphoblastoid cell lines. *Elife* 10, 1–57. doi:10.7554/eLife.62586.

Tang, D. D., and Gerlach, B. D. The roles and regulation of the actin cytoskeleton, intermediate filaments and microtubules in smooth muscle cell migration. doi:10.1186/s12931-017-0544-7.

Trevisoli, P.A., Moreira, G.C.M., C, Cesar, A.S.M., Petrini, J., Ledur, M.C., Mourão, G.B., Garrick, D. and Coutinho, L.L. (2021). A Missense Mutation in the MYBPH Gene Is Associated With Abdominal Fat Traits in Meat-Type Chickens. Front. Genet. 12:698163. doi: 10.3389/fgene.2021.698163.

Xie, T., He, Y., Korkeamaki, H., Zhang, Y., Imhoff, R., Lohi, O., et al. (2011). Structure of the 30-kDa Sin3-associated protein (SAP30) in complex with the mammalian Sin3A corepressor and its role in nucleic acid binding. *J. Biol. Chem.* 286, 27814–27824. doi:10.1074/jbc.M111.252494.

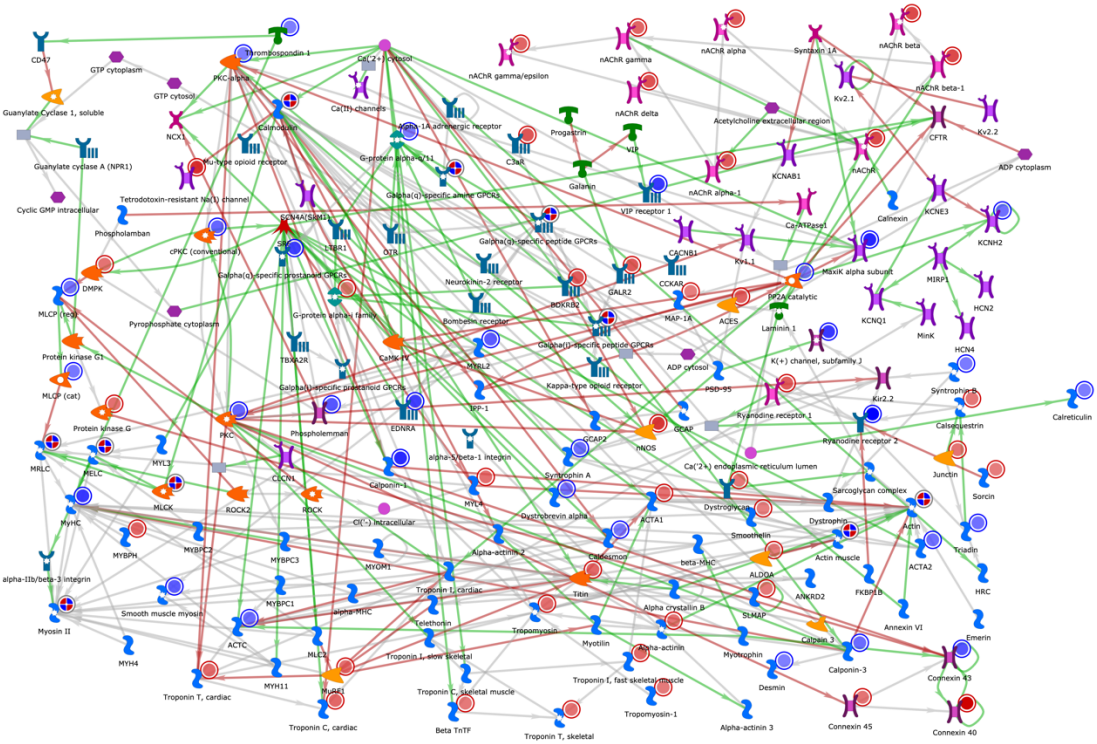
Zammit, P. S., Partridge, T. A., and Yablonka-Reuveni, Z. (2006). The skeletal muscle satellite cell: The stem cell that came in from the cold. *J. Histochem. Cytochem.* 54, 1177–1191. doi:10.1369/jhc.6R6995.2006.

Zhang, Y., and Iratni, R. (1997). Histone Deacetylases and SAP18, a Novel Polypeptide, Are Components of a Human Sin3 Complex.

Zhang, Y., Sun, Z. W., Iratni, R., Erdjument-Bromage, H., Tempst, P., Hampsey, M., et al. (1998). SAP30, a novel protein conserved between human and yeast, is a component of a histone deacetylase complex. *Mol. Cell* 1, 1021–1031. doi:10.1016/S1097-2765(00)80102-1.

Zhang, Y., Wang, Y., Zuo, Q., Li, D., Zhang, W., Wang, F., et al. (2017). CRISPR/Cas9 mediated chicken Stra8 gene knockout and inhibition of male germ cell differentiation. *PLoS One* 12, e0172207. doi:10.1371/journal.pone.0172207.

APPENDIX



Appendix A. Muscle contraction pathway. Genes marked with red circle were upregulated in the CRISPRmutated_2, while genes with the blue circle were downregulated. Figure created by Metacore software.