Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”

Identificação e caracterização de regiões de eucromatina associadas à regulação da expressão gênica e à gordura intramuscular em bovinos da raça Nelore

Natalia Silva Morosini

Dissertação apresentada para obtenção do título de
Mestra em Ciências. Área de concentração: Ciência
Animal e Pastagens

Piracicaba
2018

Natalia Silva Morosini
Farmacêutica-Bioquímica

Identificação e caracterização de regiões de eucromatina associadas à regulação da expressão gênica e à gordura intramuscular em bovinos da raça Nelore

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

Dissertação apresentada para obtenção do título de
Mestra em Ciências. Área de concentração: Ciência
Animal e Pastagens

Piracicaba
2018

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

Morosini, Natália Silva

Identificação e caracterização de regiões de eucromatina associadas à regulação da expressão gênica e à gordura intramuscular em bovinos da raça Nelore / Natalia Silva Morosini. - - Piracicaba, 2018.

55 p.

Dissertação (Mestrado) - USP / Escola Superior de Agricultura "Luiz de Queiroz".

1. Regulação gênica 2. *Bos indicus* 3. Gordura intramuscular 4. ATAC-seq
I. Título

DEDICATÓRIA

*À minha família, meu maior
alicerce, minha maior riqueza.
Dedico.*

AGRADECIMENTOS

À Deus, por ter me permitido finalizar mais uma etapa importante de minha caminhada. Agradeço pela força concedida para que eu pudesse vencer todos os obstáculos, pela luz enviada à cada dúvida e incerteza. Agradeço, também, pela fé renovada que me manteve de cabeça erguida diante toda dificuldade. Obrigada pela coragem para sempre seguir em frente. Toda a minha gratidão por Ele nunca falhar comigo e ser meu maior e mais fiel companheiro.

À Escola Superior de Agricultura “Luiz de Queiroz” pelo exemplo de profissionalismo, por agregar oportunidades e conhecimentos em minha história acadêmica e, principalmente, por ser Gloriosa em sua essência, representação plena de lealdade e compromisso aos ideais científicos educacionais.

Ao Prof. Dr. Luiz Lehmann Coutinho pelo aprendizado e críticas construtivas. Por ser meu guia acadêmico, por me dar voz e referência.

À Prof^a. Márcia Regina von Zeska Kress por ter me inserido no mundo da expressão gênica, por ter direcionado minhas decisões profissionais sempre com muita ética e cautela.

À equipe técnica Fátima, Sílvia, Pilar Mariani e Marcela Paduan. Obrigada pelo ombro amigo, pela torcida, pelas palavras de otimismo e por todo o incentivo. Agradeço por todas as tentativas, por fazerem das minhas dúvidas suas preocupações e por terem me oferecido todo o conhecimento teórico-prático que tenho hoje.

Aos meus pais, Carlos Alberto Morosini e Olívia Maria B.S. Morosini, pela vida. Obrigada pela estrutura familiar que me permitiu ser, cada dia mais, uma pessoa convicta de meus ideais. Obrigada por serem sempre leais a nós quatro e à nossa união. Por terem estado sempre ao meu lado me ensinando, amorosamente, a enfrentar a vida com maturidade. Obrigada pela confiança e por me propiciarem momentos tão ricos em felicidade e amor. Agradeço, também, por nunca terem medido esforços em me oferecer colo, palavras de consolo. Ouvir vocês falarem “vai dar tudo certo”, é o primeiro passo para que tudo seja concretizado com sucesso. Obrigada por serem exemplo de resiliência. Sem vocês eu nada seria.

À minha irmã, Júlia Morosini, pelo exemplo a ser seguido. Obrigada por ser rumo para a minha vida acadêmica e profissional. Obrigada pelo carinho e pelo amor, pelos abraços apertados, pelas longas conversas e por compartilhar comigo sonhos, planos, medos e esperanças.

Ao meu namorado, Lucas Delfini. Obrigada pelos sorrisos, pela paz transmitida e por todo o seu equilíbrio.

Agradeço aos amigos do laboratório de Biotecnologia Animal pela companhia e pelos ensinamentos. Em especial, gostaria de agradecer a Dr^a Aline Silva Mello Cesar pelo profissionalismo, por ter corrido contra o tempo junto à mim e por ter tornado possível o aprimoramento desse trabalho. Agradeço também a Dr^a Mirele Daiana Poleti por todo o suporte nas inúmeras tentativas de otimização de protocolo. Sem vocês, a finalização desse estudo não teria sido possível.

Ao açougue Bardela e toda sua equipe. Obrigada por terem sido tão receptivos, por me propiciarem a base material do meu projeto e por fazer do sucesso da minha pesquisa, o seu objetivo.

À EMBRAPA – pecuária sudeste por todo o apoio técnico e por todas as críticas e sugestões.

Agradeço à CAPES e à FAPESP (processo nº: 2016/13773-6), pelo auxílio financeiro e confiança em minha pesquisa.

Por fim, agradeço a Luiz de Queiroz (*in memorian*) pelo privilégio concedido, a partir de seu sonho e grande idealização, de ter me tornado filha da terra Esalqueana.

EPÍGRAFE

*“Oh Escola nascida no monte!
Joia rara de fino labor!
Esse nome que trazes na fronte
é o nome do teu sonhador,
do que teve feliz privilégio,
qual Anchieta de ampla visão,
de prever, na montanha, um colégio
que crescesse por toda nação!
Desta Escola, por ele sonhada
e dos jovens que a ela vêm ter,
eis que surge a legião denodada
de uma gente que aspira vencer.
Cavaleiros que odeiam a guerra,
bem armados de sãos ideais,
converteram o humo da terra
na pujança dos seus cafezais.
Oh Escola! Oh flor da montanha!
Oh insigne "Luiz de Queiroz"!
Tua história é uma força tamanha,
que nos faz avançar mais veloz.
Tua vida, o passado escreveu!
Tua glória, o futuro dirá!
Teu presente assinala o apogeu
do grandioso amanhã que virá.
Ao cantarmos as nossas conquistas,
numa vida de intenso labor,
outra coisa não temos em vista,
que pagar-te um tributo de amor.
Eia, pois, esalqueanos, sem guerra!
Co'a bandeira da Escola na mão,
ensinai que plantar nesta terra
é lutar pela grande Nação ”*

Salvador de Toledo Piza Jr

SUMÁRIO

RESUMO.....	9
ABSTRACT.....	10
1. INTRODUÇÃO	11
2. REVISÃO DE LITERATURA.....	13
2.1. Maquinaria transcricional.....	13
2.2. Remodelagem da cromatina	13
2.3. Elementos regulatórios	14
2.4. Arquitetura genética	15
2.5. Objetivos	16
2.5.1. <i>Objetivos específicos</i>	16
2.6. Hipóteses.....	17
3. MATERIAIS E MÉTODOS.....	19
3.1. Materiais	19
3.2. Animais e amostras	19
3.3. Extração e Biblioteca ATAC-DNA	19
3.4. Sequenciamento ATAC-DNA.....	19
3.5. Controle de qualidade dos dados do sequenciamento	19
3.6. Alinhamento contra o genoma de referência	20
3.7. Identificação de regiões de eucromatina	20
3.8. Caracterização de regiões de eucromatina	20
3.8.1. <i>Sobreposição entre regiões de eucromatina, Transcriptional Start Sites (TSS), quantitative trait loci (QTL), expression quantitative trait loci (eQTL), genes expressos, genes não expressos e genes diferencialmente expressos (GDE)</i>	20
3.8.2. <i>Análises de overlap e geração de gráfico</i>	21
3.8.3. <i>Anotação dos picos de eucromatina</i>	21
4. RESULTADOS	23
4.1. Obtenção de região de cromatina aberta	23
4.1.1. <i>Extração de eucromatina</i>	23
4.1.2. <i>Determinação da proporção de núcleos e enzimas</i>	26
4.2. Periodicidade nucleossômica	27
4.3. Enriquecimento de TSSs ativos ordenados por expressão	28
4.4. Identificação de regiões de eucromatina	29
4.4.1. <i>Anotação dos picos de eucromatina</i>	30
4.5. Caracterização de regiões de eucromatina	31
4.5.1. <i>Sobreposições entre regiões de eucromatina e TSS anotados, eQTL associados à gordura intramuscular (GIM) e genes expressos em músculo esquelético</i>	31
4.5.2. <i>Sobreposição entre regiões de eucromatina e QTL associados à gordura intramuscular (GIM)</i>	32
4.5.3. <i>Sobreposição entre regiões de eucromatina e genes diferencialmente expressos (GDE)</i>	33
4.5.4. <i>Visualização da correspondência entre picos de eucromatina e genes diferencialmente expressos (GDE)</i>	34
5. DISCUSSÃO	39
5.1. Protocolo ATAC-Seq	39
5.2. Periodicidade nucleossômica	39
5.3. Enriquecimento de TSSs ativos ordenados por expressão	40
5.4. Identificação de regiões de eucromatina	40
5.4.1. <i>Anotação dos picos de eucromatina</i>	40
5.5. Caracterização de regiões de eucromatina	40

5.5.1. Sobreposições entre regiões de eucromatina e TSS, eQTL e genes expressos	40
5.5.2. Sobreposição entre regiões de eucromatina e genes diferencialmente expressos.....	41
6. CONCLUSÃO	43
REFERÊNCIAS.....	45
ANEXOS	51

RESUMO

Identificação e caracterização de regiões de eucromatina associadas a regulação da expressão gênica e a gordura intramuscular em bovinos da raça Nelore

Em eucariotos, o DNA é organizado juntamente com histonas em um complexo nucleoproteico conhecido como cromatina, cuja unidade fundamental corresponde aos nucleossomos. A cromatina apresenta-se de duas maneiras: eucromatina, região estruturalmente menos condensada e, portanto, mais facilmente transcrita, e heterocromatina, região muito condensada e transcricionalmente silenciosa. Na forma de eucromatina, o acesso dos fatores de transcrição a regiões de DNA livres de nucleossomos é facilitado, enquanto que na forma de heterocromatina os fatores de transcrição não conseguem acessar o DNA para ativar ou reprimir a expressão gênica inferindo, assim, que o grau de compactação da cromatina interfere na regulação da expressão gênica e que o nucleossomo atua como silenciador gênico. Neste contexto, os objetivos foram identificar, mapear e caracterizar regiões em eucromatina na musculatura esquelética de bovinos da raça Nelore. As análises foram realizadas em relação ao músculo *Longissimus dorsi* pela técnica *Assay for Transposase-Accessible Chromatin* (ATAC-Seq), capaz de isolar regiões livres de nucleossomos a partir do mecanismo enzimático de transposição. A fim de otimizar o protocolo dessa metodologia para tecido muscular, foram testadas concentrações de 50 mil, 75 mil e 100 mil núcleos tratados com transposase. Destes, foram encontrados 6.811, 11.121 e 11.473 picos de eucromatina, respectivamente, e 6.212 regiões de cromatina aberta foram coincidentes nas três amostras. A associação entre regiões eucromáticas, expressão gênica e gordura intramuscular foi confirmada a partir da análise de sobreposição com *transcriptional start sites* (TSS), genes expressos em músculo esquelético, genes diferencialmente expressos (GDE) para gordura intramuscular e regiões de *expression quantitative trait loci* (eQTL) de tecido muscular reforçando, assim, o potencial regulatório das regiões de eucromatina.

Palavras-chave: Regulação gênica; *Bos indicus*; Gordura intramuscular; ATAC-seq

ABSTRACT

Identification and characterization of euchromatic regions associated with gene expression and intramuscular fat in Nelore cattle

In eukaryotes, DNA is organized along with histones in nucleoproteins complexes known as chromatin, which has nucleosomes as their fundamental unit. Chromatin exists in two forms: euchromatin, corresponding to a lightly condensed structure and an easily transcribed region, and heterochromatin, a highly condensed and transcriptionally silent region. In euchromatin form, transcription factors have free access to nucleosome-depleted DNA regions, while in heterochromatin the transcription factors can not access the DNA for activate or repress genic expression, which suggests that the chromatin compaction degree interferes with regulation of gene expression and that nucleosomes act as gene silencer. In this context, the aims of the present project were to identify, map and characterize euchromatin regions in the skeletal musculature of Nelore cattle. Analyzes were performed considering the muscle *Longissimus dorsi* using *Assay for Transposase-Accessible Chromatin* technique (ATAC-Seq), which isolates nucleosome-depleted regions through transposition enzymatic mechanism. Different transposase-treated nuclei concentrations were tested: 50 thousand, 75 thousand and 100 thousand. From these, 6.811, 11.121, and 11.473 euchromatin peaks were found, respectively, and 6.212 open chromatin regions were coincident among them. The association between euchromatic regions, gene expression and intramuscular fat was confirmed from the overlap analysis with transcriptional start sites (TSS), genes expressed in skeletal muscle, differentially expressed genes (GDE) for intramuscular fat and regions of expression quantitative trait loci (eQTL) of muscle tissue, reinforcing the regulatory potential of the euchromatin regions.

Keywords: Genic regulation; *Bos indicus*; Intramuscular fat; ATAC-seq

1. INTRODUÇÃO

O Brasil possui aproximadamente 210 milhões de cabeças de gado distribuídos em 167 milhões de hectares resultando, em 2015, em uma produção nacional de carne bovina de aproximadamente 39 milhões de equivalente-carcaça e em uma movimentação monetária de mais de 480 bilhões de reais - cerca de 3% da receita de exportação brasileira (ABIEC, 2016). Os animais de corte brasileiros são, majoritariamente, da raça Nelore (origem *Bos indicus* - Zebu) considerada resistente ao calor e à parasitas sendo, portanto, mais bem adaptada às condições ambientais brasileiras (LUCHIARI FILHO, 2000).

As ferramentas de biotecnologia animal, tal como *next generation sequencing* (NGS), vêm sendo associadas a técnicas bioquímicas com o objetivo de investigar a relação entre a estrutura da cromatina e os fenômenos que regem a expressão gênica de características de interesse (DE JAGER et al., 2013; TIZIOTO et al., 2015; BERGET et al., 2010).

O genoma eucariótico envolto estruturalmente no interior celular por proteínas histonas, recebe o nome de cromatina cuja unidade fundamental corresponde aos nucleossomos (BARKSI et al., 2007; KOUZARIDES, 2007 e BERNENSTEIN et al., 2005) compostos por aproximadamente 146 pares de bases (SRIVASTAVA et al., 2016; GREWAL; JIA, 2007). A cromatina pode ser dividida em dois grupos extremos: i) forma ativa (induzível) conhecida como eucromatina e ii) forma inativa (silenciada) chamada de heterocromatina sendo que a primeira modifica seu grau de compactação ao longo dos ciclos celulares enquanto a segunda se mantém condensada (GRIGORYEV; WOODCOCK, 2012; MORALES et al., 2001). O fácil acesso de fatores de transcrição (FT) a regiões de eucromatina, menos condensadas, viabiliza a transcrição enquanto o silenciamento gênico é associado a heterocromatina devido a presença de nucleossomos e de domínios estruturais essenciais à manutenção de cromossomos intactos, nessa estrutura (BASSETT et al., 2009; LE et al., 2004 e BINTU et al., 2012).

Atualmente, diferentes metodologias foram descritas a fim de se identificar regiões de eucromatina e componentes regulatórios. A técnica *DNase I High Sensivity* consiste em digerir concentração adequada de núcleos celulares de modo a expor à DNase regiões genéticas hipersensíveis consideradas como cromatina aberta (SONG; CRAWFORD, 2010). A metodologia *Formaldehyde-assisted Isolation of Regulatory Elements* (FAIRE-seq) consiste em

isolar nucleosome-depleted regions (NDRs) por meio da utilização de formaldeído (PAUL et al., 2011). Chromatin immunoprecipitation (ChIP-Seq), por sua vez, caracteriza-se por identificar regiões de ligação entre proteínas associadas à DNA a partir da imunoprecipitação de cromatina (PEPKE et al., 2009). Por fim, *Assay for Transposase-Accessible Chromatin* (ATAC-seq) utiliza transposase bacteriana, Tn5, como forma de inserir pequeno fragmento de DNA, transposon, em região de eucromatina devido à incapacidade enzimática de acessar molécula de DNA vinculada a nucleossomos (DAVIE, 2015).

O objetivo deste trabalho foi utilizar a metodologia ATAC-Seq para identificar e caracterizar as regiões de eucromatina presentes no genoma do músculo *Longissimus dorsi* de bovinos da raça Nelore a fim de determinar possíveis regiões regulatórias associadas a expressão gênica e à gordura intramuscular (GIM).

2. REVISÃO DE LITERATURA

2.1. Maquinaria transcricional

A transcrição é o processo inicial da expressão de um gene. Essa etapa é realizada pela enzima RNA polimerase II que, por meio da anulação da barreira nucleossomal, permite o acesso de complexos proteicos ao DNA, sintetizando uma molécula de RNA (PETESCH; LIS, 2012; DAHAN; CHODER, 2013). Os FT e as proteínas reguladoras de DNA são imprescindíveis na identificação dos genes a serem transcritos pela *Pol2* (HEYN et al., 2015). A maquinaria transcricional utiliza elementos *cis-acting* (promotores e elementos regulatórios distais: *enhancers*, *silencers*, *insulators* ou *locus control regions* (LCR)) codificados pela rede reguladora de genes (GNR) (MASTON et al., 2006) que representam locais específicos à fatores de transcrição *trans-acting* (PETER; DAVIDSON, 2015). Os FT podem ser divididos em: *fatores de transcrição gerais* (GTF), *ativadores e co-ativadores*. Os primeiros se ligam aos promotores formando o complexo pré-iniciação da transcrição (PIC) - responsável pelo direcionamento da *Pol2* ao *transcription start site* (TSS) enquanto os segundos situam-se nos promotores aumentando a formação de PIC e modulando a estrutura da cromatina. O terceiro e último grupo, por fim, tende a modificar a habilidade dos ativadores de regular positiva ou negativamente a transcrição (MASTON et al., 2006).

2.2. Remodelagem da cromatina

Proteínas co-ativadoras ativam elementos regulatórios ao remodelar nucleossomos (CALO; WYSOCKA, 2013) o que sugere a associação entre as histonas modificadas e as atividades regulatórias (GROSS; GARRARD, 1988 e CHAI et al., 2013, ROEDER, 2005 e WEAKE; WORKMAN, 2010). Essa informação baseia-se no fato de que as histonas apresentam cauda amino-terminal susceptível a alterações epigenéticas pós-traducionais (acetilação, metilação e fosforilação). Estas alterações modulam, dinâmica e estruturalmente, a arquitetura da cromatina, interferindo na acessibilidade dos elementos regulatórios ativos (BERNENSTEIN et al., 2002; CHEUNG et al., 2000 e LUGER et al., 1997).

A distância entre nucleossomos corresponde de 10 a 50 pares de bases, sendo esse espaçamento regulado a partir do reposicionamento de nucleossomos em regiões de controle *cis-acting* de acordo com a exigência transcricional do *locus*. Para isso,

remodeladores responsáveis pela acessibilidade da maquinaria transcricional ao DNA, reorganizam o genoma ao remover ou 'deslizar' nucleossomos que possam interferir no processo de transcrição ou, ainda, ao expor elementos transitórios na superfície destes. Eles se ancoram aos nucleossomos enquanto o domínio translocase, ATPase, se liga ao DNA envolto por esses na posição fixa do octâmero conduzindo, assim, a translocação de DNA. Esse mecanismo ocorre a partir da ação de domínios DBD e Tr responsáveis por direcionar DNA para dentro do nucleossomo (criando looping de 100 pb de DNA) e por bombear esse material genético em direção à díade nucleossômica, respectivamente. O looping formado apresenta conformação desfavorável à ligação nucleossômica criando uma região 'livre' de nucleossomos (NFR) localizadas, geralmente, nos TSS e compostas por sítios de ligação para fatores de transcrição (CLAPIER; CAIRNS, 2009).

2.3. Elementos regulatórios

Os *Enhancers* compõem uma importante classe de elementos regulatórios que possuem entre 200 e 1000 pares de base específicos à FT e agrupam-se próximos à *transcriptional start site* (TSS) e apresentam funcionalidade quando situados dentro íntrons (LEVINE, 2010, CALO; WYSOCKA, 2013; ISTRAIL; DAVIDSON, 2005, WUNDERLE et al., 1998; SPITZ et al., 2001; GOLDHAMER et al., 1992, HEITZMAN et al., 2007 e WELBOREN et al., 2009; SCHWARTZ; MESHORER, 2009; LEVINE, 2010). Por possuírem ampla habilidade de comunicação genética (centenas de kilobases), permitem interação entre DNA e fatores de transcrição aumentando o nível de expressão dos genes associados (BULGER; GROUDINE, 2011; ONG; CORCES, 2011; SPITZ, 2016; JIN et al., 2013 e SANYAL et al., 2012).

Para controlar a expressão gênica, os *enhancers* são sujeitos a regulação dinâmica via modificações epigenéticas (ENGEL et al., 2016). Estudos sugerem que a maioria dos *enhancers* atuam em *cis* apesar de não precisarem residir na mesma molécula de DNA de seu promotor alvo uma vez que apresentam capacidade em rastrear e localizar promotores receptivos (DONG; WENG, 2013 e MUELLER-STORM et al., 1989) a partir do engajamento de genes distantes, pulando ou ignorando promotores próximos e distintos. (SANYAL et al., 2012). Além disso, são susceptíveis a rearranjos nucleotídicos que interferem na afinidade dos fatores de transcrição e, conseqüentemente, na expressão de genes vizinhos indicando que *enhancers* são seletivos em suas interações regulatórias (LETTICE et al., 2011). *Enhancers* distais, porém, podem ser envolvidos por nucleossomos requerendo acetilação de histonas e a

ação de remodeladores – perda nucleossomal perto dos TSS e movimentação nucleossômica por regiões promotoras e codante – para serem expostos e ativados (SCHWARTZ; MESHORER, 2009).

A função dos *enhancers* pode ser explicada por dois modelos: o Binário e o Progressivo. O primeiro propõe que os *enhancers* aumentam a ativação transcricional de um determinado *locus* dentro de uma população celular enquanto o segundo sugere que os *enhancers* aumentam o número de moléculas de RNA transcritas a partir de um determinado gene embora não interfira no número de células iniciadoras da transcrição (GARCIA-GONZALEZ et al., 2016). Complexos macromoleculares formados por proteínas específicas, FT e cofatores proteicos são diretamente envolvidos na regulação transcricional (ENGEL et al., 2016). Essas proteínas favorecem a regulação gênica por atuarem como pontes moleculares que ligam genes promotores à elementos regulatórios distais a partir de um modelo *looping*. Segundo SANYAL et al., 2012; DEKKER, 2008; CUBENAS-POTTS; CORCES, 2015; PTASHNE, 1986). Já a ligação entre enhancers e FT pode ser discutida por dois modelos diferentes: modelo Enhanceosome e modelo Billboard. No primeiro modelo acredita-se que a sequência de DNA recrute fatores de transcrição formando um complexo nucleoproteico ativador da transcrição. Nesse cenário a ação do *enhancer* emerge a partir de uma rede de interações e essa ação só se rompe quando uma proteína é perdida (MERIKA et al., 1998 e PANNE, 2008). O modelo Billboard, por sua vez, supõe que as ligações dos fatores de transcrição são independentes e que esses fatores não atuam como unidades únicas (ARNOSTI; KULKARNI, 2005).

2.4. Arquitetura genética

Outro fator que contribui para a organização do material genético baseia-se na estrutura arquitetônica do DNA que pode ser mantida por diferentes unidades de éxons e íntrons (SCHWARTZ; MESHORER, 2009). Podemos citar, por exemplo, a distribuição da quantidade e comprimento desses dois grupos uma vez que há indícios de que células de ciclagem rápida são compostas por genes com poucos íntrons, facilitando a expressão gênica eficiente, enquanto genes com longos íntrons demonstram expressão tardia. Somado a esse fato, estudos indicam que a maior densidade de éxons em um segmento de DNA responde à uma taxa de alongamento mais lenta visto que a Pol2 fisicamente se posiciona sobre éxons,

onde nucleossomos possam estar presentes, pausando a atividade transcrricional em íntrons contendo genes aumentando, assim, o tempo total necessário para a expressão gênica (HEYN et al., 2015; SCHWARTZ; MESHORER, 2009; CARRILLO et al., 2011; VELOSO et al., 2014; JONKERS et al., 2014).

Locais de splicing (porção 3') geralmente contêm sequências gênicas desfavoráveis aos nucleossomos e, portanto, tendem a deslocá-los em direção aos éxons que, concordantemente, demonstram níveis aumentados de ocupação nucleossômica. Justificadamente, os fragmentos de DNA envoltos por nucleossomos apresentam valor semelhante ao comprimento médio dos éxons, 146 pares de base (SCHWARTZ; MESHORER, 2009).

O posicionamento nucleossômico no DNA pode afetar o reconhecimento de éxons ao nível de RNA uma vez que atuam como redutores de velocidade diminuindo a atividade da *Pol2*. No entanto, a redução na taxa de transcrição parece aumentar a inclusão de éxons justapostos pós-splicing. Os nucleossomos utilizam histonas com caudas modificadas para marcar éxons a fim de aumentar o reconhecimento destes.

2.5. Objetivos

Identificar, mapear e caracterizar regiões em eucromatina na musculatura esquelética de bovinos da raça Nelore.

2.5.1. Objetivos específicos

- 1) Identificar regiões de eucromatina usando a metodologia ATAC-seq;
- 2) Validar as regiões encontradas verificando o tamanho e a periodicidade dos fragmentos sequenciados, bem como a sobreposição com regiões de TSSs;
- 3) Identificar regiões de eucromatina em regiões de genes expressos, em regiões de QTLs e em regiões de eQTLs;
- 4) Identificar regiões de eucromatina próximas a genes associados a deposição de gordura intramuscular.

2.6. Hipóteses

A metodologia ATAC-seq permite identificar regiões genômicas associadas a regulação da expressão de gênica no tecido muscular de bovinos.

3. MATERIAIS E MÉTODOS

3.1. Materiais

Vide anexo A.

3.2. Animais e amostras

Foi utilizado tecido muscular de bovinos Nelore, com idade de 24 meses, provenientes de abatedouro comercial. Uma amostra de tecido muscular foi obtida a partir do músculo *Longissimus dorsi* localizado entre a 12ª e 13ª vértebra de carcaças refrigeradas por 30 horas após abate. O tecido muscular foi adequadamente transportado em caixas de isopor preenchidas com gelo a fim de se prevenir a deterioração ou congelamento da amostra.

3.3. Extração e Biblioteca ATAC-DNA

Para ser aplicado em amostras de tecido muscular, o protocolo ATAC-seq descrito por BUENROSTRO et al., 2013 foi otimizado seguindo as etapas retratadas na tese de BREWER, 2015 e no trabalho de GOUBIL, 2015. Embora este último estudo, realizado no INRA, ainda não tenha sido publicado, o artigo nos foi enviado pelo Professor James M. Reecy membro do “*Department of Animal Science of Iowa State University*” com o qual nosso laboratório possui colaboração (anexos). O protocolo ATAC-seq otimizado foi dividido em 7 etapas (vide Anexo B).

3.4. Sequenciamento ATAC-DNA

Foram gerados aproximadamente 40 milhões de *reads* por biblioteca sequenciadas em HiSeq 2500. O tipo de sequenciamento utilizado foi paired-end 2 x 100.

3.5. Controle de qualidade dos dados do sequenciamento

A ferramenta Trimmomatic (v:0.36) foi utilizada para retirar adaptadores Nextera das extremidades dos *reads* (BOLGER et al., 2014) e posteriormente a qualidade dessa remoção foi avaliada pela ferramenta FastQC (MARTIN, 2011; LEGGETT et al., 2013).

3.6. Alinhamento contra o genoma de referência

O mapeamento de *reads* contra o genoma de referência “*Bos taurus*” (UMD, 3.1) foi realizado por meio do *software* Bowtie2 (v:2.2.9). As *reads* mapeadas em regiões do DNA mitocondrial e classificadas como duplicatas de PCR foram removidas por meio dos *softwares* Picard (v:2.6.0) e Samtools (v:1.3.1) (EBBERT, 2016).

3.7. Identificação de regiões de eucromatina

A identificação das regiões de eucromatina foi realizada pelo programa *Model-based Analysis of Chip-Seq* (MACS2 v. 2.1.1) que estima a distribuição dos picos potencialmente eucromáticos (FENG et al., 2011). O controle de qualidade tanto dos dados de mapeamento quanto dos gerados pelo MACS foi realizado pela plataforma Qualimap (v. 2.2.1) (GARCÍA-ALCALDE, 2012). Para a comparação entre os resultados de ATAC-seq para os diferentes ensaios foram utilizadas as ferramentas do programa Atacv v. 0.9.4) (UNIVERSITY OF MICHIGAN).

3.8. Caracterização de regiões de eucromatina

3.8.1. Sobreposição entre regiões de eucromatina, Transcriptional Start Sites (TSS), quantitative trait loci (QTL), expression quantitative trait loci (eQTL), genes expressos, genes não expressos e genes diferencialmente expressos (GDE)

O arquivo contendo as informações de anotação (cromossomo e posição) dos TSS para o genoma *Bos taurus* UMD_3.1, foi gerado a partir do programa BIOMART ENSEMBL (v. 91). A partir destas informações foram extendidos 1500 bases *upstream* dos TSS anotados para que sequências promotoras fossem abrangidas (HERMAN-IZYCKA, 2017). Os eQTL associados a funções biológicas de deposição de gordura de bovinos, publicados identificados por CESAR et al., 2018 (in review), foram sobrepostos com os picos obtidos para eucromatina para analisar a correspondência entre essas regiões. Foi realizada a interseção entre os genes expressos em músculo *Longissimus dorsi* publicados por CESAR et al., 2016 e as regiões de cromatina aberta uma vez que preferencialmente genes expressos devem situar-se em regiões transcritas e, portanto, em eucromatina. Foi realizado, também, o *overlap* entre os genes diferencialmente expressos (GDE) para gordura intramuscular (GIM) - publicados por

CESAR et al., 2015 e para ácido oleico (C18:1), ácido linoléico conjugado cis9 trans11 (CLA) e ácido palmítico (C16:0) - publicados por CESAR et al., 2016. Todas esses *overlaps* foram realizados a partir da ferramenta *bedtools intersect* (v:2.26.0).

3.8.2. Análises de *overlap* e geração de gráfico

A contagem do número de sobreposições (*overlap*) entre os picos de eucromatina e outras regiões de interesse foram realizadas pelo pacote RegioneR, pertencente à ferramenta R/Bioconductor, específico para se analisar regiões genômicas (GEL et al., 2016). O conjunto de ferramentas deepTools foi utilizado para gerenciar e explorar os dados sequenciados permitindo, assim, a análise gráfica dos resultados (RAMÍREZ et al., 2014).

3.8.3. Anotação dos picos de eucromatina

A análise da localização genética das regiões de eucromatina foi realizada a partir da ferramenta VEP-ENSEMBL.

4. RESULTADOS

4.1. Obtenção de região de cromatina aberta

4.1.1. Extração de eucromatina

A obtenção de um bom resultado de ATAC-seq depende de uma boa preparação de núcleos. Uma pequena quantidade de núcleos limita a quantidade de DNA disponível para a construção da biblioteca. O aumento do número de núcleos pode ser conseguido com métodos agressivos de dissociação do tecido muscular, no entanto isto pode resultar em núcleos rompidos, o que prejudica a qualidade da biblioteca porque desfavorece a atividade enzimática da Tn5.

Para otimizar esta etapa, iniciamos com 10 mg de tecido fresco que foi macerado em cadinho juntamente com nitrogênio líquido. Utilizamos o equipamento douncer em duas etapas subsequentes cada qual com 50 movimentos 'stroke'. Com o auxílio de microscopia invertida, foi possível visualizar núcleos em pouquíssima quantidade (Figura 1).

Figura 1. Microscopia invertida de 10 mg de tecido muscular fresco obtida a partir do Microscópio LEICA DMIL, 200X. Presença de núcleos fragmentados e em pouquíssima quantidade.

A partir desse resultado tentamos aumentar para 650 mg a quantidade inicial de tecido fresco e esse foi submetido às mesmas condições anteriores. A partir da microscopia, observamos a presença de núcleos aglomerados e totalmente fragmentados (Figura 2).

Figura 2. Microscopia invertida de 650 mg de tecido fresco obtida a partir do Microscópio LEICA DMIL, 200X. Presença de núcleos aglomerados e totalmente fragmentados.

Novamente, a quantidade de tecido inicial foi aumentada para 1g mas dessa vez a maceração em cadinho juntamente com nitrogênio líquido foi substituída pela utilização de bisturi e tesoura cirúrgica a fim de tentar reduzir o rompimento e a fragmentação nuclear. Mais uma vez, com o auxílio de microscopia invertida foi possível observar presença de núcleos em grumos (aglomerados) e totalmente fragmentados (Figura 3).

Figura 3. Microscopia invertida de 1g de tecido muscular fresco obtida a partir do Microscópio LEICA DMIL, 200X. Presença de núcleos fragmentados, aglomerados B: Microscopia de fluorescência obtida a partir do Microscópio Axiophot, 100X – núcleos corados com DAPI. Visualização de núcleo fragmentado.

De acordo com as observações anteriores, aumentamos para 2g a quantidade de tecido fresco e este foi fragmentado com bisturi e tesoura cirúrgica, em buffer A. O meio obtido foi homogeneizado em Turrax com velocidades 3, 2 e 1 durante 15 segundos cada uma subsequentemente. Posteriormente realizou-se filtração seguido de ação mecânica com douncer de vidro - 50 movimentos 'stroke'. Na etapa de formação de gradiente de separação, utilizou-se Percoll a 27% e a programação da centrífuga foi ajustada de acordo com a dimensão do rotor (95 cm de raio) para 20817g a fim de se obter 25000g de rotação (sugerida pelo protocolo). Na microscopia foi possível contabilizar 2.880.000 núcleos intactos em 1mL de suspensão (Figura 4).

Figura 4. A: Microscopia invertida de 2g de tecido fresco obtida a partir do Microscópio LEICA DMIL, 200X. Núcleos íntegros, conservados e passíveis de serem tratados com Tn5. B: Microscopia de fluorescência obtida a partir do Microscópio Axiophot, 100X – núcleos corados com DAPI. Visualização de núcleo íntegro e conservado.

Foram realizados, também, testes para comparar as eficiências da utilização de 2g de tecido fresco e de tecido muscular congelado há 4 anos (amostras da EMBRAPA). Após serem mantidas as etapas de fragmentação e isolamento de núcleos, a microscopia invertida apresentou-se de maneira límpida (núcleos sem presença de miofibrilas) e com núcleos intactos (Figura 5) para tecido fresco e quantidade considerável de núcleos estourados para tecido congelado (Figura 6).

Figura 5. Microscopia invertida de 2g de tecido fresco obtida a partir do Microscópio LEICA DMIL, 200X. Presença de núcleos íntegros, conservados e em quantidade adequada para o tratamento com transposase. B e C: Microscopia de fluorescência obtida a partir do Microscópio Axiophot, 100X – núcleos corados com DAPI. Visualização de núcleo íntegro e conservado.

Figura 6. A: Núcleos estourados obtidos para 2g de tecido congelado e visualizados em Microscópio LEICA DMIL, 200X (à esquerda) e em Microscópio Axiophot, 100X – núcleos corados com DAPI (à direita).

4.1.2. Determinação da proporção de núcleos e enzimas

Com o intuito de testar a quantidade nuclear ideal para a adequada atividade enzimática, alíquotas correspondentes a quatro diferentes concentrações de núcleos: 1X = 50.000 núcleos; 1,5X = 75.000 núcleos; 2X = 100.000 núcleos e 4X = 200.000 núcleos, foram obtidas e tratadas com 2,5 μ L de Tn5. A biblioteca foi posteriormente amplificada com um

total de 17 ciclos de *polymerase chain reaction* (PCR). A concentração final de DNA foi mensurada em QUBIT HS apresentando valor de 275,12 ng; 287,28 ng; 532 e 327 ng. Posteriormente foram testadas 2g de três amostras de tecido muscular proveniente de bovinos Nelore recém-abatidos na Universidade de São Paulo - campus Pirassununga, pertencentes à linha de pesquisa do laboratório Avaliação Animal e Qualidade de carne responsável Prof. Dr. Saulo da Luz e Silva (Figura 7).

Figura 7. Microscopia invertida de 2g de tecido recém-abatido (processado 4 horas após abate) obtida a partir do Microscópio LEICA DMIL, 200X. Presença de núcleos aglomerados e estourados, em maior quantidade.

Para analisar a repetibilidade dos resultados anteriormente considerados satisfatórios, foi realizado um último experimento com 2 gs de tecido muscular fresco. As mesmas etapas foram seguidas e foram contabilizados 840.000 núcleos intactos em câmara de Neubauer. Duas alíquotas (I e II) de 100.000 núcleos foram obtidas e tratadas com Tn5. As bibliotecas foram posteriormente construídas e os resultados da qPCR sugeriram a adição de 12 ciclos de PCR normal (sendo esse valor padronizado para ATAC-seq de tecido muscular). A utilização do QUBIT HS mensurou 810 ng e 509 ng de DNA.

4.2. Periodicidade nucleossômica

Para verificar se a eficácia da enzima Tn5 foi satisfatória ao estudo de ATAC-seq, os fragmentos obtidos pelo sequenciamento das bibliotecas foram analisados por meio do sinal nucleossômico resultante do isolamento periódico das diferentes regiões de eucromatina. As amostras 1X, 1,5X e 2X apresentaram maior frequência periódica para fragmentos de

tamanho inferior à 200 pares de base e menor frequência para fragmentos de tamanho correspondente à 1, 2 ou 3 nucleossomos (150 à 500 pares de base) (Figura 8-A,B,C).

Figura 8. Análise gráfica da periodicidade nucleossômica. A: amostra 1X; B: Amostra 1,5X e C: Amostra 2X.

4.3. Enriquecimento de TSSs ativos ordenados por expressão

Buscamos reforçar a idéia de que regiões de eucromatina isoladas pela ATAC-seq regulam a expressão gênica a partir de um teste Heatmap para enriquecimento de TSSs ativos ordenados por expressão. Na figura 9, observamos que as regiões de eucromatina das amostras 1X, 1,5X e 2X enriquecem, majoritariamente, em regiões de TSSs ativos e em regiões *upstream* aos genes expressos em músculo *Longissimus dorsi*.

Figura 9. Heatmap das amostras 1X, 1,5X e 2X demonstrando maior enriquecimento de TSSs ativos e de regiões *upstream* ao gene expresso.

4.4. Identificação de regiões de eucromatina

Um desafio do ensaio ATAC-Seq é diferenciar o número de leituras que caracteriza uma região de cromatina aberta versus leituras randômicas que são obtidas no sequenciamento. Para isso, o programa MACS2 (v:2.1.1) foi usado para identificar a distribuição das regiões de eucromatina no músculo esquelético de Nelore. As amostras analisadas por esse programa foram as de 50.000 núcleos (1X), 75.000 núcleos (1,5X) e 100.000 núcleos (2X) (vide resultados – 4.1.2) resultando em um total de 33.991.425 leituras

para a amostra 1X, 40.448.209 reads para a amostra 1,5X e 39.257.295 para a amostra 2X. A chamada de picos de eucromatina indicou que o total de reads acima citado para 1X, 1,5X e 2X núcleos, identificou 6.811 picos, 11.121 picos e 11.473 picos, respectivamente. A quantidade de picos de eucromatina correspondentes nas três amostras, foi estimada a partir da sobreposição das informações da localização física de cada pico presente em cada amostra. Essa análise indicou 6.309 regiões de eucromatina em comum para 1X e 1,5X; 6.302 picos similares para 1X e 2X; 9.024 regiões de cromatina aberta equivalentes para 1,5X e 2X e 6.212 picos correlatos nas três amostras.

4.4.1. Anotação dos picos de eucromatina

A anotação dos picos de eucromatina correspondentes nas três amostras foi realizada por meio da ferramenta VEP-ENSEMBL para determinar, de acordo com o genoma de referência (*Bos taurus* UMD_3.1), a localização genética onde cada região de eucromatina se situa. A partir da figura 10 é possível inferir que 30% dos picos associados à cromatina aberta correspondem à regiões intergênicas enquanto apenas 5% deles situam-se em posição downstream de genes próximos.

Figura 10. Ordenação da localização genética anotada para o conjunto dos picos de eucromatina. A maioria dos picos associados à cromatina aberta correspondem a regiões intergênicas.

4.5. Caracterização de regiões de eucromatina

4.5.1. Sobreposições entre regiões de eucromatina e TSS anotados, eQTL associados à gordura intramuscular (GIM) e genes expressos em músculo esquelético

As regiões de eucromatina identificadas para *Longissimus dorsi* foram sobrepostas com regiões de cis e trans-eQTL (Cesar et al., 2018, *in review*), de genes expressos no músculo esquelético (Cesar et al., 2016) e de TSSs anotados pelo BIOMART ENSEMBL (v.91), para analisarmos se as regiões enriquecidas para ATAC-Seq se encontram em regiões associadas à regulação da expressão gênica. Esse teste foi realizado por meio do pacote estatístico Regioner do R a partir do qual foram consideradas regiões sobrepostas aquelas que apresentaram valores de significância ($p < 0.05$). Para que regiões promotoras também fossem analisadas foram adicionadas 1.500 bases *upstream* às regiões. Além disso, os TSSs foram separados conforme o nível de expressão dos genes do músculo esquelético (baixa, média e alta expressão). Em nossos resultados foi observada uma sobreposição significativa entre picos de eucromatina e todos os itens analisados exceto cis-eQTL, distantes não mais do que 1Mb do gene expresso. A maior região de *overlap* (2,3 Mb) foi observada entre as regiões de eucromatina e as regiões pertencentes à trans-eQTL (distante mais do que 1Mb do gene associado) seguida da sobreposição entre regiões de cromatina aberta e TSSs associados à genes de média e alta expressão (Figura 11). Diferente do que foi observado para genes expressos, a sobreposição entre regiões de cromatina aberta e TSS de genes não expressos em músculo esquelético revelou uma menor proporção de picos de eucromatina do que seria esperado ao acaso (Figura 11). Este resultado reforça a validade da metodologia empregada, visto que não é esperada a presença de cromatina aberta em genes não expressos.

Figura 11. Tamanho do overlap (em Mb) observado (em vermelho) e aleatório (em azul) entre os picos de eucromatina e as regiões descritas de cima para baixo como: trans = eQTL distante há mais de 1Mb do gene associado; tss_ens = TSSs anotados para *Bos taurus* (UMD 3.1); cis = eQTL distantes até 1Mb do gene associado; me_exp = TSS de genes de média expressão em Longissimus dorsi; hi_exp = TSS de genes de alta expressão em Longissimus dorsi; ne = genes não expressos em músculo; lo_exp = TSS de genes de baixa expressão em Longissimus dorsi.

4.5.2. Sobreposição entre regiões de eucromatina e QTL associados à gordura intramuscular (GIM)

As regiões enriquecidas para ATAC-Seq foram sobrepostas com as regiões de QTL associadas à GIM por meio do pacote estatístico RegioneR. Esse teste teve como objetivo verificar se as regiões de eucromatina coincidem com regiões de QTL descritas para GIM. A partir do overlap apresentado na Figura 12, é possível observar que não houve sobreposição significativa entre regiões de cromatina aberta e regiões de QTL. O tamanho do *overlap*, em Mb, randômico (aleatório) superou o *overlap* observado demonstrando que regiões de eucromatina não foram enriquecidas para os QTL associados à GIM.

Figura 12. Tamanho do overlap (em Mb) observado (em vermelho) e aleatório (em azul) entre as regiões de eucromatina e regiões de QTL associadas à gordura intramuscular.

4.5.3. Sobreposição entre regiões de eucromatina e genes diferencialmente expressos (GDE)

Tendo como objetivo identificar regiões de eucromatina associadas à GIM, foi realizada a sobreposição entre regiões enriquecidas para ATAC-Seq e GDE para GIM (CESAR et al., 2015); ácido oleico (C18:1), ácido linoleico conjugado cis9trans11 (CLA) e ácido palmítico (C16:0) (CESAR et al., 2016). Foi observada uma sobreposição significativa entre picos de eucromatina e GDE para todos os fenótipos analisados. A maior região de sobreposição foi entre picos de cromatina aberta e GDE para ácido oleico resultando em um enriquecimento de aproximadamente 0,07 Mb. A segunda maior região sobreposta foi entre picos de eucromatina e CLA representando um *overlap* de 0,06 Mb. A associação entre picos de eucromatina e GDE para gordura intramuscular, apesar de significativo, apresentou valor de sobreposição de 0,005 Mb indicando que regiões de cromatina aberta foram pouco enriquecidas para GIM (Figura 13).

Figura 13. Tamanho do overlap (em Mb) observado (em vermelho) e aleatório (em azul) entre os picos de eucromatina e as regiões descritas de cima para baixo como: C18:1 = ácido oleico; CLA = ácido linoleico conjugado cis9trans11; C16:0 = ácido palmítico e IMF = gordura intramuscular (GIM).

4.5.4. Visualização da correspondência entre picos de eucromatina e genes diferencialmente expressos (GDE)

O resultado de sobreposição significativa entre regiões de eucromatina, eQTL e regiões de GDE em músculo esquelético (Figuras 11 e 13) foi analisado em IGV (v:2.3.81) para que pudéssemos visualizar como as regiões de cromatina aberta se correspondem aos GDE. Essa observação foi realizada comparativamente entre as três amostras (1X, 1,5X e 2X) para que fosse possível identificar, também, a repetibilidade do alinhamento das *reads* nas regiões de eucromatina enriquecida para as três amostras.

Quando analisamos o posicionamento físico de duas regiões cis-eQTL identificadas por CESAR et al., 2018 (*in review*), observamos que, no cromossomo 13, a região de eucromatina (Peak_1349) enriquecida para ATAC-seq nas três amostras, coincide com região *upstream* do gene ZNF335 que foi associado à GIM (Figura 14). No cromossomo 3, também encontramos uma região de cromatina aberta (Peak_4010) upstream do gene ETV3 que sobrepõem com uma região de trans-eQTL identificada por CESAR et al (2018, *in review*) (Figura 15).

Figura 14. Visualização da sobreposição entre picos de eucromatina e região cis-eQTL associada à GIM no cromossomo 13.

Figura 15. Visualização da sobreposição entre picos de eucromatina e região trans-eQTL associada à GIM no cromossomo 3.

No cromossomo 7, encontramos um pico de eucromatina (Peak_5697), localizado upstream do gene o gene FAM151B que foi descrito como GDE associado com GIM (CESAR et al., 2015). (Figura 16).

Figura 16. Visualização da sobreposição, no cromossomo 7, entre picos de eucromatina e gene FAM151B diferencialmente expresso para GIM.

No cromossomo 3 encontramos outra região de eucromatina (Peak_4090), localizada upstream do gene DDX20 que foi diferencialmente expresso e associado à concentração de ácido linoleico cis9trans11 (CLA) (Figura 17).

Figura 17. Visualização da sobreposição, no cromossomo 3, entre picos de eucromatina e gene diferencialmente expresso para ácido graxo CLA.

Já o GDE SCD situado no cromossomo 26 e associado à ácido oleico (C18:1), teve seu posicionamento genômico abrangido por duas regiões de eucromatina aberta (Peak_3524 e Peak_3525) indicando que uma se localiza na porção 5' enquanto a outra na porção 3' desse gene (Figura 18).

Figura 18. Visualização da sobreposição, no cromossomo 26, entre picos de eucromatina e gene diferencialmente expresso para ácido graxo C18:1

Os GDE para ácido graxo C16:0 situados no cromossomo 15 também foram visualizados em IGV de modo a determinar se alguma região de eucromatina se sobrepõe nessa localização genômica. O Peak_1551 teve overlap em região intergênica entre os genes CRYAB e HSPB2 e *upstream* ao gene CRYAB (Figura 19).

Figura 19. Visualização da sobreposição, no cromossomo 15, entre picos de euromatina e genes diferencialmente expressos para ácido graxo C16:0.

Para mostrarmos como se manifesta a distribuição das *reads* em regiões de cromatina aberta e como estas correspondem à genes anotados para *Bos taurus*, selecionamos a região do cromossomo 13 na qual foram encontrados picos de euromatina com maior número de leituras alinhadas. Na figura 20 podemos analisar a sobreposição visual entre estes picos (em vermelho) eem qual região gênica as regiões de cromatina aberta se associam.

Figura 20. Distribuição das *reads* em regiões de cromatina aberta associadas, no cromossomo 13, à genes anotados para *Bos taurus*.

5. DISCUSSÃO

5.1. Protocolo ATAC-Seq

Os resultados mostraram que a quantidade de tecido inicial não compromete a integridade da estrutura dos núcleos, mas interfere na quantidade de núcleos observados em microscopia invertida. Além disso, os experimentos confirmaram a importância da utilização de força mecânica uma vez que a utilização de douncer de vidro possibilitou fragmentação adequada de tecido muscular e, consequentemente, maior quantidade de núcleos obtidos.

O protocolo otimizado mostrou ser mais eficiente quando executado em alíquotas de 100.000 núcleos uma vez que essa quantidade corresponde à concentração de DNA adequada para a atividade enzimática e para a construção de bibliotecas.

5.2. Periodicidade nucleossômica

O sinal nucleossômico corresponde à capacidade enzimática de isolar periodicamente diferentes perfis de região de eucromatina classificados pelo tamanho do fragmento observado. O tamanho desses fragmentos direciona informações a respeito do posicionamento de nucleossomos uma vez que a periodicidade de aproximadamente 200 pares de base indica a fração real de enovelamento da cromatina inferindo o quanto que uma região é abrangida por nucleossomos íntegros (BUENROSTRO, 2013). O perfil nucleossomal das regiões eucromáticas exposto pela ATAC-Seq permite prever a acessibilidade da cromatina tornando possível a distinção entre regiões protegidas por 1, 2 ou mais nucleossomos. Em nossas análises obtivemos maior fração de *reads* com comprimento inferior à 100 pares de base (Figura 8) mas uma ampla distribuição, em menor frequência, de fragmentos de até 250 pares de base sugerindo que a maior parte do DNA foi submetido à transposição entre 2 nucleossomos visto que a distância entre estes consiste entre 10 e 50 pares de base (CLAPIER; CAIRNS, 2009). Esse resultado se assemelha aos descritos nos trabalhos de BUENROSTRO et al., 2015, BUENROSTRO et al., 2013 nos quais bibliotecas de células linfoblásticas humanas enriquecidas para ATAC-Seq, apresentaram maior frequência de *reads* com comprimento subnucleossômico (menor que 150 pares de base) e ampla distribuição periódica de fragmentos de tamanho de 200 pares de base, respectivamente. O estudo de DAVIE et al., 2015 também demonstrou periodicidade nucleossômica parecida com

a encontrada em nosso trabalho visto que ao estudarem desenvolvimento tumoral em *Drosophila*, obtiveram maior fração de reads de comprimento inferior à 147 pares de base e uma distribuição mais homogênea de fragmentos de até 294 pares de base.

5.3. Enriquecimento de TSSs ativos ordenados por expressão

A análise de enriquecimento de TSSs ativos ordenados por expressão, propôs verificar se as regiões de cromatina aberta se relacionam com as regiões promotoras e responsáveis pelo início da transcrição e validar a confiabilidade de as regiões de eucromatina enriquecidas para ATAC-Seq corresponderem à regiões associadas à expressão gênica. Nossos resultados mostraram que maior distribuição de *reads* em regiões promotoras e de TSSs ativos que, ao apresentarem maior acessibilidade transcricional, revelam a associação entre região de cromatina aberta e expressão gênica.

5.4. Identificação de regiões de eucromatina

5.4.1. Anotação dos picos de eucromatina

As informações a respeito da posição física dos picos de eucromatina despertaram nosso interesse em identificar a localização genética das regiões de cromatina aberta. O entendimento das características dos locais onde regiões de eucromatina são encontradas favorece a associação dessas regiões com elementos regulatórios anotados para o genoma de referência. A partir dos nossos resultados, cerca de 30% das regiões de eucromatina encontram-se em região intergênica enquanto 20% e 11% dos picos correspondem à regiões de íntrons e regiões *upstream*, respectivamente. Estudos indicam que 90% do genoma eucariótico é composto por regiões intergênicas (HEYEN et al., 2015), que a grande maioria dos enhancers distribuem-se em íntrons específicos (WUNDERLE et al., 1998; SPITZ et al., 2001) e que regiões de cromatina aberta localizam-se em regiões *upstream* ao início da transcrição (ALLEMAND et al., 2008; DE LA MATA et al., 2003; KORNBLIHTT, 2007) justificando, assim, a anotação dos picos enriquecidos para ATAC-Seq no músculo esquelético.

5.5. Caracterização de regiões de eucromatina

5.5.1. Sobreposições entre regiões de eucromatina e TSS, eQTL e genes expressos

O teste de *overlap* entre as regiões que compreendem eucromatina com regiões de eQTL (cis e trans), TSSs, genes expressos e genes não expressos reforçou a confiabilidade dos picos encontrados. Os resultados indicam que as regiões associadas ao processo transcricional sobrepõem regiões reafirmando a relação existente entre remodelagem de cromatina e expressão gênica. O fato de regiões de cromatina aberta coincidirem significativamente com trans-eQTL, sustentam a ideia de que a maquinaria transcricional é também regulada por regiões e elementos regulatórios distais (SANYAL et al., 2012; DEKKER, 2008; CUBENAS-POTTS; CORCES, 2015; PTASHNE, 1986). O estudo da sobreposição de TSS de genes não expressos no tecido muscular com regiões de eucromatina confirma identificação de regiões de eucromatina, visto que observamos uma menor proporção de picos de cromatina aberta próximas a TSS de genes não expresso no tecido muscular.

5.5.2. Sobreposição entre regiões de eucromatina e genes diferencialmente expressos

Uma maneira de explicar se as regiões de eucromatina encontradas se associam à gordura intramuscular de *Longissimus dorsi*, é analisando se essas regiões coincidem com regiões pertencentes à GDE para GIM e para ácidos graxos atuantes nos processos fisiológicos do músculo esquelético. Estudos indicam que além de os ácidos graxos de cadeia longa participarem da regulação transcricional muscular, os GDE estão mais associados ao ácido oleico (C18:1) e ao ácido linoleico conjugado cis9 trans11 (CLA) do que ao ácido palmítico (C16:0) (CESAR et al., 2016). CESAR et al., 2016 elucidaram que C18:1 e CLA contribuem significativamente para a expressão de genes associados à importantes processos biológicos. No ano anterior, CESAR et al., 2015, constataram que animais com maior GIM apresentam maior nível de expressão de genes relacionados à remodelagem da cromatina. Em nosso estudo, as regiões de eucromatina coincidiram de maneira significativa com todas as regiões de TSSs ativos pertencentes aos GDE acima citados. Além disso, os significativos *overlaps* observados entre picos de eucromatina e C16:0 e CLA, confirmam a atuação das primeiras regiões na maioria dos processos fisiológicos associados à regulação e à arquitetura genética. A correspondência entre as regiões enriquecidas para ATAC-Seq e as regiões de GDE foram confirmadas a partir da visualização gráfica em IGV. Os GDE associados com GIM, demonstraram enriquecimento para regiões de eucromatina em regiões *upstream* do TSS.

Esse resultado indicou que as regiões de cromatina aberta identificadas nesse estudo coincidem com regiões promotoras responsáveis pelo início da transcrição e com regiões de *enhancers* situados na porção 5' ou 3' de genes associados à característica de gordura intramuscular em *Longissimus dorsi*.

6. CONCLUSÃO

A técnica ATAC-Seq demonstrou ser eficiente em identificar regiões de eucromatina, em músculo *Longissimus dorsi*, associadas à expressão gênica e à GIM uma vez que verificamos o enriquecimento de regiões de início de transcrição (TSS), de genes expressos em músculo esquelético, de genes diferencialmente expressos para GIM e de regiões expressas associadas a características quantitativas, às regiões de cromatina aberta descritas nesse trabalho.

REFERÊNCIAS

- ALLEMAND, E.; BATSCHE, E.; MUCHARDT, C. Splicing, transcription, and chromatin: a ménage à trois. **Curr Opin Genet Dev**, v. 18, n. 2, p. 145-51, Apr 2008.
- ARNOSTI, D. N.; KULKARNI, M. M. Transcriptional enhancers: Intelligent enhanceosomes or flexible billboards? **J Cell Biochem**, v. 94, n. 5, p. 890-8, Apr 2005.
- Associação Brasileira das Indústrias Exportadoras de Carne – ABIEC. **Perfil da Pecuária no Brasil**. Relatório anual 2016.
- BARSKI, A. et al. High-resolution profiling of histone methylations in the human genome. **Cell**, v. 129, n. 4, p. 823-37, May 2007
- BASSETT, A.; COOPER, S.; WU, C.; TRAVERS, A. The folding and unfolding of eukaryotic chromatin. **Current Opinion in Genetics & Development**, v. 19, p. 159– 165, 2009
- BERGET, I. et al. Caprine CSN1S1 haplotype effect on gene expression and milk composition measured by Fourier transform infrared spectroscopy. **J Dairy Sci**, v. 93, n. 9, p. 4340-50, Sep 2010
- BERNSTEIN, B. E. et al. Methylation of histone H3 Lys 4 in coding regions of active genes. **Proc Natl Acad Sci U S A**, v. 99, n. 13, p. 8695-700, Jun 2002
- BERNSTEIN, B.E. et al. Genomic maps and comparative analysis of histone modifications in human and mouse. **Cell**, v. 120, n. 2, p. 169-81, Jan 2005.
- BINTU, L. et al. Nucleosomal elements that control the topography of the barrier to transcription. **Cell**, v. 151, n. 4, p. 738-49, Nov 2012.
- BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114-20, Aug 2014.
- BREWER, O. Isolation of Myogenic Nuclei from Whole Muscle Tissue. Tese, 2015; p.32
- BUENROSTRO, J. D. et al. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. **Curr Protoc Mol Biol**, v. 109, p. 21.29.1-9, Jan 2015.
- BUENROSTRO, J. D. et al. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. **Nat Methods**, v. 10, n. 12, p. 1213-8, Dec 2013
- BULGER, M.; GROUDINE, M. Functional and mechanistic diversity of distal transcription enhancers. **Cell**, v. 144, n. 3, p. 327-39, Feb 2011.

- CALO, E.; WYSOCKA, J. Modification of enhancer chromatin: what, how, and why? **Mol Cell**, v. 49, n. 5, p. 825-37, Mar 2013.
- CARRILLO OESTERREICH, F.; BIEBERSTEIN, N.; NEUGEBAUER, K. M. Pause locally, splice globally. **Trends Cell Biol**, v. 21, n. 6, p. 328-35, Jun 2011.
- CESAR, A. S. et al. Differences in the skeletal muscle transcriptome profile associated with extreme values of fatty acids content. **BMC Genomics**, v. 17, n. 1, p. 961, 11 2016.
- CESAR, A. S. et al. Putative regulatory factors associated with intramuscular fat content. **PLoS One**, v. 10, n. 6, p. e0128350, 2015.
- CHAI, X. et al. Regulation of the boundaries of accessible chromatin. **PLoS Genet**, v. 9, n. 9, p. e1003778, 2013.
- CHEUNG, P.; ALLIS, C. D.; SASSONE-CORSI, P. Signaling to chromatin through histone modifications. **Cell**, v. 103, n. 2, p. 263-71, Oct 2000.
- CLAPIER, C. R.; CAIRNS, B. R. The biology of chromatin remodeling complexes. **Annu Rev Biochem**, v. 78, p. 273-304, 2009.
- CUBEÑAS-POTTS, C.; CORCES, V. G. Architectural proteins, transcription, and the three-dimensional organization of the genome. **FEBS Lett**, v. 589, n. 20 Pt A, p. 2923-30, Oct 2015.
- DAHAN, N.; CHODER, M. The eukaryotic transcriptional machinery regulates mRNA translation and decay in the cytoplasm. **Biochim Biophys Acta**, v. 1829, n. 1, p. 169-73, Jan 2013.
- DAVIE, K. et al. Discovery of transcription factors and regulatory regions driving in vivo tumor development by ATAC-seq and FAIRE-seq open chromatin profiling. **PLoS Genet**, v. 11, n. 2, p. e1004994, Feb 2015.
- DE JAGER, N. et al. Gene expression phenotypes for lipid metabolism and intramuscular fat in skeletal muscle of cattle. **J Anim Sci**, v. 91, n. 3, p. 1112-28, Mar 2013.
- DE LA MATA, M. et al. A slow RNA polymerase II affects alternative splicing in vivo. **Mol Cell**, v. 12, n. 2, p. 525-32, Aug 2003.
- DEKKER, J. Gene regulation in the third dimension. **Science**, v. 319, n. 5871, p. 1793-4, Mar 2008.
- DONG, X.; WENG, Z. The correlation between histone modifications and gene expression. **Epigenomics**, v. 5, n. 2, p. 113-6, Apr 2013.

- EBBERT, M. T. et al. Evaluating the necessity of PCR duplicate removal from next-generation sequencing data and a comparison of approaches. **BMC Bioinformatics**, v. 17 Suppl 7, p. 239, Jul 2016.
- ENGEL, K. L. et al. Decoding transcriptional enhancers: Evolving from annotation to functional interpretation. **Semin Cell Dev Biol**, v. 57, p. 40-50, Sep 2016.
- FENG, J.; LIU, T.; ZHANG, Y. Using MACS to identify peaks from ChIP-Seq data. **Curr Protoc Bioinformatics**, v. Chapter 2, p. Unit 2.14, Jun 2011
- GARCÍA-ALCALDE, F. et al. Qualimap: evaluating next-generation sequencing alignment data. **Bioinformatics**, v. 28, n. 20, p. 2678-9, Oct 2012.
- GARCÍA-GONZÁLEZ, E. et al. Chromatin remodeling effects on enhancer activity. **Cell Mol Life Sci**, v. 73, n. 15, p. 2897-910, Aug 2016.
- GEL, B. et al. regioneR: an R/Bioconductor package for the association analysis of genomic regions based on permutation tests. **Bioinformatics**, v. 32, n. 2, p. 289-91, Jan 2016.
- GOLDHAMER, D. J. et al. Regulatory elements that control the lineage-specific expression of myoD. **Science**, v. 256, n. 5056, p. 538-42, Apr 1992.
- GOUBIL, A, 2015: **ATAC-SEQ Protocol**. Disponível em: <ftp://ftp.faang.ebi.ac.uk/ftp/protocols/assays/INRA_SOP_ATAC-seq_AG_v1_20160805.pdf>. Acesso em 07.jun.de 2017.
- GREWAL, S. I.; JIA, S. Heterochromatin revisited. **Nat Rev Genet**, v. 8, n. 1, p. 35-46, Jan 2007.
- GRIGORYEV, S. A.; WOODCOCK, C. L. Chromatin organization - the 30 nm fiber. **Exp Cell Res**, v. 318, n. 12, p. 1448-55, Jul 2012.
- GROSS, D. S.; GARRARD, W. T. Nuclease hypersensitive sites in chromatin. **Annu Rev Biochem**, v. 57, p. 159-97, 1988.
- HEINTZMAN, N. D. et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. **Nat Genet**, v. 39, n. 3, p. 311-8, Mar 2007.
- HERMAN-IZYCKA, J.; WLASNOWOLSKI, M.; WILCZYNSKI, B. Taking promoters out of enhancers in sequence based predictions of tissue-specific mammalian enhancers. **BMC Med Genomics**, v. 10, n. Suppl 1, p. 34, May 2017.
- HEYN, P. et al. Introns and gene expression: cellular constraints, transcriptional regulation, and evolutionary consequences. **Bioessays**, v. 37, n. 2, p. 148-54, Feb 2015.

- ISTRAIL, S.; DAVIDSON, E. H. Logic functions of the genomic cis-regulatory code. **Proc Natl Acad Sci U S A**, v. 102, n. 14, p. 4954-9, Apr 2005.
- JIN, F. et al. A high-resolution map of the three-dimensional chromatin interactome in human cells. **Nature**, v. 503, n. 7475, p. 290-4, Nov 2013.
- JONKERS, I.; KWAK, H.; LIS, J. T. Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. **Elife**, v. 3, p. e02407, Apr 2014
- KORNBLIHTT, A. R. Coupling transcription and alternative splicing. **Adv Exp Med Biol**, v. 623, p. 175-89, 2007.
- KOUZARIDES, T. Chromatin modifications and their function. **Cell**, v. 128, n. 4, p. 693-705, Feb 2007.
- LE, H. D. et al. A high proportion of genes involved in position effect variegation also affect chromosome inheritance. **Chromosoma**, v. 112, n. 6, p. 269-76, Mar 2004.
- LEGGETT, R. M. et al. Sequencing quality assessment tools to enable data-driven informatics for high throughput genomics. **Front Genet**, v. 4, p. 288, 2013
- LETTICE, L. A. et al. Enhancer-adoption as a mechanism of human developmental disease. **Hum Mutat**, v. 32, n. 12, p. 1492-9, Dec 2011.
- LEVINE, M. Transcriptional enhancers in animal development and evolution. **Curr Biol**, v. 20, n. 17, p. R754-63, Sep 2010.
- LUCHIARI FILHO, A. **Pecuária da carne bovina**. 1a. ed. Nova Odessa: A. Luchiari Filho, 2000
- LUGER, K. et al. Crystal structure of the nucleosome core particle at 2.8 Å resolution. **Nature**, v. 389, n. 6648, p. 251-60, Sep 1997.
- MARTIN, M: Cutadapt removes adapter sequences from high-throughput sequencing *reads*. EMBnetjournal, North America 2011, 17.
- MASTON, G. A.; EVANS, S. K.; GREEN, M. R. Transcriptional regulatory elements in the human genome. **Annu Rev Genomics Hum Genet**, v. 7, p. 29-59, 2006.
- MERIKA, M. et al. Recruitment of CBP/p300 by the IFN beta enhanceosome is required for synergistic activation of transcription. **Mol Cell**, v. 1, n. 2, p. 277-87, Jan 1998.
- MORALES, V. et al. Chromatin structure and dynamics: functional implications. **Biochimie**, v. 83, n. 11-12, p. 1029-39, 2001 Nov-Dec 2001.
- MÜELLER-STORM, H. P.; SOGO, J. M.; SCHAFFNER, W. An enhancer stimulates transcription in trans when attached to the promoter via a protein bridge. **Cell**, v. 58, n. 4, p. 767-77, Aug 1989.

- ONG, C. T.; CORCES, V. G. Enhancer function: new insights into the regulation of tissue-specific gene expression. **Nat Rev Genet**, v. 12, n. 4, p. 283-93, Apr 2011.
- PANNE, D. The enhanceosome. **Curr Opin Struct Biol**, v. 18, n. 2, p. 236-42, Apr 2008.
- PAUL, D. S. et al. Maps of open chromatin guide the functional follow-up of genome-wide association signals: application to hematological traits. **PLoS Genet**, v. 7, n. 6, p. e1002139, Jun 2011.
- PEPKE, S.; WOLD, B.; MORTAZAVI, A. Computation for ChIP-seq and RNA-seq studies. **Nat Methods**, v. 6, n. 11 Suppl, p. S22-32, Nov 2009.
- PETER, I.S; DAVIDSON, E.H. Genomic Control Process: Development and Evolution. Estados Unidos: Academic Press, 2013; p.257-321.
- BROOKS, J. W. Postmortem Changes in Animal Carcasses and Estimation of the Postmortem Interval. **Vet Pathol**, v. 53, n. 5, p. 929-40, Sep 2016.
- PETESCH, S. J.; LIS, J. T. Overcoming the nucleosome barrier during transcript elongation. **Trends Genet**, v. 28, n. 6, p. 285-94, Jun 2012.
- PTASHNE, M. Gene regulation by proteins acting nearby and at a distance. **Nature**, v. 322, n. 6081, p. 697-701, 1986 Aug 21-27 1986.
- RAMÍREZ, F. et al. deepTools: a flexible platform for exploring deep-sequencing data. **Nucleic Acids Res**, v. 42, n. Web Server issue, p. W187-91, Jul 2014.
- ROEDER, R. G. Transcriptional regulation and the role of diverse coactivators in animal cells. **FEBS Lett**, v. 579, n. 4, p. 909-15, Feb 2005
- SANYAL, A. et al. The long-range interaction landscape of gene promoters. **Nature**, v. 489, n. 7414, p. 109-13, Sep 2012.
- SCHWARTZ, S.; MESHORER, E.; AST, G. Chromatin organization marks exon-intron structure. **Nat Struct Mol Biol**, v. 16, n. 9, p. 990-5, Sep 2009.
- SONG, L.; CRAWFORD, G. E. DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. **Cold Spring Harb Protoc**, v. 2010, n. 2, p. pdb.prot5384, Feb 2010.
- SPITZ, F. et al. Large scale transgenic and cluster deletion analysis of the HoxD complex separate an ancestral regulatory module from evolutionary innovations. **Genes Dev**, v. 15, n. 17, p. 2209-14, Sep 2001

- SPITZ, F. Gene regulation at a distance: From remote enhancers to 3D regulatory ensembles. **Semin Cell Dev Biol**, v. 57, p. 57-67, Sep 2016.
- SRIVASTAVA, R.; AHN, S. H. The Epigenetic Pathways to Ribosomal DNA Silencing. **Microbiol Mol Biol Rev**, v. 80, n. 3, p. 545-63, Sep 2016.
- TIZIOTO, P. C. et al. Global liver gene expression differences in Nelore steers with divergent residual feed intake phenotypes. **BMC Genomics**, v. 16, p. 242, 2015.
- UNIVERSITY OF MICHIGAN. The Parker Lab. Disponível em <<https://github.com/ParkerLab/ataqv>>. Acesso em: 03 ago. 2017.
- VELOSO, A. et al. Rate of elongation by RNA polymerase II is associated with specific gene features and epigenetic modifications. **Genome Res**, v. 24, n. 6, p. 896-905, Jun 2014.
- WEAKE, V. M.; WORKMAN, J. L. Inducible gene expression: diverse regulatory mechanisms. **Nat Rev Genet**, v. 11, n. 6, p. 426-37, Jun 2010.
- WELBOREN, W. J. et al. ChIP-Seq of ERalpha and RNA polymerase II defines genes differentially responding to ligands. **EMBO J**, v. 28, n. 10, p. 1418-28, May 2009.
- WUNDERLE, V. M. et al. Deletion of long-range regulatory elements upstream of SOX9 causes campomelic dysplasia. **Proc Natl Acad Sci U S A**, v. 95, n. 18, p. 10649-54, Sep 1998.

ANEXOS

Anexo A: Materiais ATAC-Seq

- 1X PBS, pH 7.4 (Thermo Fisher Scientific 10010-023)
- Acetato de Magnésio (Sigma M2545)
- AMPure XP beads (Agilent)
- Bioanalyzer 2100 (Agilent)
- Bisturi
- Bovine Serum Albumin (BSA) (Sigma A7906)
- Buffer A
- Buffer B
- Câmara de Neubauer
- Celltrics 100 μ M (Sysmex 04-0042-2328)
- Celltrics 30 μ M (Sysmex 04-0042-2316)
- Centrífuga (Eppendorf 5427R)
- Cloreto de Cálcio (Merck 102382)
- Cloreto de Magnésio (Merck 105833)
- Cloreto de Sódio (Fisher Scientific BP358-212)
- Complete EDTA-free protease inhibitor tablets (Roche 05056489001)
- DAPI (4',6-Diamidine-2'-phenylindole dihydrochloride) (Roche – 10236276001)
- DL-Dithiothreitol (DTT) (Sigma D0632)
- Douncer tissue grinder (Sigma D9063)
- EDTA (Fisher Scientific 120-500)
- Etanol 70% (diluição a partir de etanol absoluto Merck 107017)
- Falcon 15 mL
- Homogeneizador (T10 Basic Ultra Turrax)
- IGEPAL CA630 (Sigma I8896)
- Light Cycler 480 II (Roche)
- Light Cycler 480 SYBR Green I Master (Roche 04707516901)
- Lysis Buffer
- Microscópio de Fluorescência Axiophot - Zeiss

- Microscópio LEICA DMIL
- Nextera DNA Library Preparation Kit (Illumina FC-121-1030)
- Nextera Index Kit (Illumina FC-121-1011)
- Percoll (Sigma P1644)
- Placa de Petri
- Placa magnética (Ambion)
- QIAquick PCR purification Kit (QIAGEN 28104)
- Qubit HS (High Sensitivity) e BR (Broad Range) (Invitrogen Q32866 1107007702)
- Sucrose (Sigma S0389)
- Termociclador (Biorad)
- Tesoura cirúrgica
- Tris-HCl pH 7,4 (Invitrogen 15504-020)
- Triton-X-100 (Sigma X100)
- Trypan Blue Stain 0.4% (Thermo-Fischer 15250061)
- Vortex (Fisherbrand)

Reagentes :

LYSIS BUFFER: 10 mM Tris-HCl, pH 7,4; 10 mM NaCl; 3 mM MgCl₂; 0,1% IGEPAL CA-630

BUFFER A: 0,25 M Sucrose; 10 mg/mL BSA; 5 mM MgCl₂; Inibidor protease (1 tablet em 50 mL)

BUFFER B: 0,32M sucrose; 3mM CaCl; 2mM Acetato de magnésio; 0,1mM EDTA; 1mM DTT; 10mg/ML BSA; 10 mM Tris-HCl; Inibidor de protease (1 tablet em 50 mL)

Anexo B: Protocolo ATAC-Seq otimizado

I) Homogeneização do tecido muscular: 2g de tecido muscular (24 horas após abate) foram picados em placa de petri, em 6 mL de meio Buffer A, com bisturi e tesoura cirúrgica. Logo em seguida, esse tecido foi homogeneizado com Turrax em velocidade 3, 2 e 1, durante 15 segundos, respectivamente, e filtrado com filtro de 100 μ M. O filtrado foi, então, acrescido de 0,5% de Triton X-100, submetido à 50 'strokes' em douncer de vidro e filtrado novamente, dessa vez com filtro 30 μ M.

II) Separação de núcleos celulares: A solução obtida após filtragem foi dividida em alíquotas de 1ml que foram centrifugadas a 3000g por 10 minutos. Os pellets formados foram ressuspensos em 1ml de Buffer B, cada, onde adicionou-se Percoll a 27% para criar gradiente de separação. Essas amostras foram centrifugadas a 20817g por 15 minutos e cada pellet formado foi novamente ressuspenso em 1 ml de Buffer B. Posteriormente a esta etapa, centrifugou-se os tubos a 450g por 5 minutos e novamente os pellets formados foram ressuspensos em 1ml de Buffer B. Juntou-se as alíquotas em falcon de 15 ml e realizou-se contagem em Microscópio eletrônico com câmara de Neubauer.

III) Preparação da amostra (para realizar ATAC-seq): Após contagem a partir de microscopia invertida, retirou-se alíquotas correspondentes à 100.000 núcleos intactos que foram centrifugadas a 500g por 5 minutos a 4°C. O pellet de cada alíquota foi ressuspenso e, 100 μ L de Lysis Buffer e centrifugado a 500g por 10 minutos a 4°C.

IV) Reação de transposição: Preparou-se o mix da reação de transposição com os reagentes do Nextera DNA Kit:

- 2x TD Buffer ----- 25 μ L
- Tn5 Transposase ----- 2,5 μ L
- Nuclease Free H2O ----- 22,5 μ L
- Total----- 50 μ L

O volume total (50 µL) foi distribuído para cada pellet formado na etapa anterior, incubado por 30 minutos a 37°C e purificado com QIAquick PCR purification Kit.

V) Amplificação por PCR: A amplificação dos fragmentos de DNA transposto foi realizada seguindo o seguinte parâmetro:

- 10 µL DNA (produto da purificação)
- 5 µL i5
- 5 µL i7
- 25 µL Master Mix
- 5 µL PPC
- Total = 50µL

Nas seguintes condições:

- 1 ciclo: 5 minutos	72°C
30 segundos	98°C
- 5 ciclos: 10 segundos	98°C
30 segundos	63°C
1 minuto	72°C

VI) PCR quantitativa: Uma vez que o número apropriado de ciclos de PCR (N) é determinado utilizando Qpcr, calculous-se o número adicional de ciclos necessários mensurando o número de ciclos correspondentes a 1/3 da intensidade máxima de fluorescência. Para isso, a seguinte reação foi preparada:

- 5 µl de DNA amplificado por PCR
- 3,9 µL PPC
- 0,5 µL i5
- 0,5 µL i7
- 0,1 µL SYBR Green I
- 5 µL Master Mix

- Total = 15 μ L

e submetida as seguintes condições:

- 1 ciclo: 30 segundos	98°C
- 20 ciclos: 10 segundos	98°C
30 segundos	63°C
1 minuto	72°C

Os 45 μ L remanescentes da reação de PCR foram submetidos novamente a amplificação da seguinte maneira:

- 1 ciclo: 30 segundos	98°C
10 segundos	98°C
- N ciclos: 10 segundos	98°C
30 segundos	63°C
1 minuto	72°C

VII) Purificação final e controle de qualidade: O produto da PCR foi então submetido a purificação com QIAquick PCR purification Kit, eluído em 20 μ L de Buffer EB (presente no kit) e purificado novamente com AMPure XP beads. A concentração do volume final de eucromatina, 38 μ L, foi mensurado em QUBIT HS e a qualidade do material genético foi analisada em bioanalyzer.